Engineering Effective Hamiltonians for Magnetic Resonance

by

Holger Haas

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Physics (Quantum Information)

Waterloo, Ontario, Canada, 2019

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:	Dr. Carlos A. Meriles Professor, Deptartment of Physics, The City College of New York
Supervisor:	Dr. David G. Cory Professor, Department of Chemistry, University of Waterloo
Internal Member:	Dr. Raymond Laflamme Professor, Department of Physics, University of Waterloo
Internal-External Member:	Dr. Jonathan Baugh Associate Professor, Department of Chemistry, University of Waterloo
Other Member(s):	Dr. Joseph Emerson Associate Professor, Department of Applied Mathematics, University of Waterloo

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Effective Hamiltonian engineering is a powerful technique that utilises time-dependent perturbation theory to suppress or enhance certain effects that arise from otherwise weak Hamiltonian terms, when an experimenter attempts to control the evolution of a quantum system. With this thesis we conduct theoretical, numerical and experimental explorations of two somewhat different kinds of effective Hamiltonian engineering in magnetic resonance, aimed at elongating spin coherence times and increasing spin polarization.

First, we provide a general framework for computing arbitrary time-dependent perturbation theory terms relevant for effective Hamiltonian engineering, as well as their gradients with respect to control variations, that immediately enables the use of gradient methods for optimizing these terms. We demonstrate the method in various numerical control engineering examples, including decoupling, recoupling, and robustness to control errors and stochastic errors. We then utilize the method in high-resolution nanoscale magnetic resonance experiments that use high-sensitivity silicon nanowire based force detection. With our numerically engineered control sequences, that are optimized to suppress spin decoherence under dipolar and chemical shift Hamiltonians, we achieve an increase by a factor of 500 in the proton spin coherence time for low-temperature, $(50\text{-nm})^3$ polystyrene sample. We harness the enhanced spin coherence times for magnetic resonance imaging experiments on the nanoscale, and achieve one-dimensional imaging resolution below 2 nm on the same sample.

Secondly, we study low-temperature, high magnetic field nuclear hyperpolarization of phosphorus donor nuclei in isotopically enriched 28-silicon crystal irradiated with non-resonant above band gap laser light. We demonstrate hyperpolarization with a time constant of 51.5 s, and introduce a phonon-mediated effective Hamiltonian model detailing the process. We verify our model by designing and conducting a series of magnetic resonance experiments.

Acknowledgements

A PhD is, in so many ways, an incredibly selfish pursuit, and yet this thesis would not have materialized without the selfless support, generous guidance, patient collaboration and kind help from more people than I will be able to list. First and foremost, I would like to thank my supervisor Professor David Cory who has taught me a huge amount of science and an equally great deal about science over these years. There are not many labs in the world, where I could have been exposed to such a diversity of ideas, technologies, facets of physics and science in general, where I could have learned and developed the skills that I was fortunate enough to acquire under David's guidance.

I would also like to express enormous gratitude to the three other closest collaborators of mine on the numerical Hamiltonian engineering project: Dr. Daniel Puzzuoli, Professor Raffi Budakian and William Rose, all of whom are absolutely spectacular scientists. It has been a true privilege to be able to work with you and learn from you. Finally, I need to thank my fellow lab monkeys on the silicon project: Thomas Alexander and Rahul Deshpande. It was such a great time, lads, I could not have asked for a better team to spend all these days and nights leaning over the cryostat with!

There are so many other Corylab members I still need to thank. I am truly grateful to Pat and George for teaching me everything I know about low temperature physics, cryogenic and RF equipment. Moreover, and much more importantly, thank you both for your incredible friendship over these years. Also, Ivar – an absolute wizard of a man when it comes to scientific instrumentation – and Alex – a machinist with a tremendous can-do attitude – thank you both for the time and energy you invested into the apparatus of the silicon project. Carly, thanks so much for always having my back and for all your kind assistance along the way. Joachim, thank you for all these excellent hours shared in our office. I am thankful of all the other Corylab members, past and present, that I had the pleasure to work with, but I would like to specifically mention Ian, Chris Grenade, Sarah, Troy, Mohamad, Natsumi, Dusan, Maryam, Don and Peter for teaching, helping and inspiring me throughout these years. Also, I am especially thankful to have met Chris Wood and Saba, who have become great friends.

I am incredibly appreciative of the educational and research experience at the Institute for Quantum Computing and the University of Waterloo in general, the multifaceted approach to quantum science and technology I have developed here is such a valuable asset. Furthermore, over the years I have spent in Waterloo I have made a number of amazing friends for life outside of the institute. Alex, Madalina, Lori, Mansi, Bill, Jessica, Shane and Anisha, thank you so much for the memorable moments and for your most generous friendships.

Katelyn, I am eternally grateful that I met you in Waterloo, and that you were there for me when I most needed it – at the time when I had entered the most horrific depths of depression. Above all, thank you for simply believing me, without a trace of judgement, that I had totally lost my nut – because I had lost it good and proper – and thank you for your truly unwavering, unbelievably kind and considerate support over the time that it took me to regain my mind. This PhD ended up being a journey to hell and back for me. And God, it is nice to be back! I do not write this to imply that there was someone or something that deliberately designed and erected, or perhaps excavated, nine circles of personal inferno for me. Absolutely not. But it is a sad, uncomfortable truth that there is a huge number of graduate students, not just me, that experience depression in some form or another. The blessing and curse of mine was that I managed to stay seemingly functional and somewhat productive throughout these times, and yet, a genuine, all-consuming feeling that I have nothing, absolutely nothing, to look forward to in life is the single most painful experience I have ever had. Any physical pain that I have ever felt totally pales in comparison. I would, quite genuinely, rather have my arm chopped off than go through anything like this again. I had some difficult negotiations over the course of my PhD, but none were anywhere near as difficult than having to negotiate with myself over the worth of living. Katelyn, you were there when I was contemplating, in total agony, whether it was time to get myself locked up in a self-harm-prevention cell, while searching for hostels around the world I could disappear to. Your smile and disposition calmed me down that night and set me on the path of recovery, you are one of the dearest, most amazing friends I have.

In the same vein, I also need to thank my mum. I have no idea what it feels like to have your thirty-year-old son cry in utmost despair on Skype for an hour, pausing only to pronounce, 'I can't do this any more', every once in a while, knowing that there is very little you can do to help him. And heaven knows I really hope I will never find that out. Thank you for standing strong and calm, thank for your soothing words and caring smile throughout this PhD.

Liina, I have missed you so, so much over these years here. But I know that you are this one friend of mine, who would drop anything and everything, at whatever time, and take the first flight to wherever I happen to be, if only I asked for it, or even if I did not, but you deemed it necessary. I would do the same exact thing for you, too! Thank you so much for always being there.

Finally, I would like to thank my sister, my godmother and my dad. I could not have done this without your ever-so-steady support!

To Harri

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Ever since the introduction of nuclear magnetic resonance by Rabi in 1938 [94], magnetic resonance has proven to be a truly transformative tool in fields as diverse as physics, materials science, chemistry, biology, medicine, engineering and quantum computing. Magnetic resonance utilises the fact that intrinsic spin angular momentum of electrons and of nuclei, that possess non-zero spin, couples to magnetic fields, which gives rise to energy differences between quantized spin angular momentum states. These spins can then be driven between the quantized levels with external electro-magnetic radiation – typically in the MHz to GHz frequency range – and the quantized state of the spins can be detected by having the magnetic moment of the spins couple to some external physical system. The fact that the spins are generally localized to atomic or nuclear length-scales makes them incredibly localized magnetic field sensors, that can be operated in a completely non-destructive way due to the low photon energies involved in magnetic resonance experiments. Moreover, in the case of nuclear spins, the coupling to magnetic fields depends on the chemical species and the particular isotope of the nucleus, therefore, such magnetic field sensors provide chemically and isotopically specific information about the sample, which forms the foundation for a variety of magnetic resonance applications including spectroscopy, sensing and imaging.

Two crucial challenges in further expanding the range of magnetic resonance applications are:

- improving spin coherence times in a particular physical system and for a particular experimental configuration, and
- increasing spin polarization in a fast manner.

The former usually translates to enhanced spectral and imaging resolution, improved sensing ability or lower error rates in the case of quantum computing, while the latter can result tremendous gains in device and experimental efficiencies. These challenges are especially critical when magnetic resonance techniques are shrunk down to nanometre scale for sensing, spectroscopy and imaging at molecular length-scales, at which point conventional spin detection and coherence preservation methods become impractical. Increasing spin polarization becomes particularly important also for applications such as magnetometry, which benefit from diluting the particles with non-zero spin to enhance the precision of sensing, because using dilute spin samples reduces the spin interactions that would hamper the sensing ability. With this thesis we address both of these challenges through effective Hamiltonian engineering. Effective Hamiltonian engineering drives a spin system to undertake a desired evolution while suspending the effects of unwanted couplings and system Hamiltonians, and it can be utilised for elongating effective spin coherence times, or driving the spin into states of increased polarization.

1.1 Thesis Outline and Main Results

Here, we give the outline for the rest of this thesis.

- With the rest of Chapter 1 we give a brief treatment of spin dynamics in magnetic fields. This entire thesis is concerned with time evolution of spin systems, hence, we introduce the Schrödinger and Lindblad equations that govern the time evolution of spin systems, treated as quantum systems in finite dimensional Hilbert spaces. In this chapter, we will only focus on spin-1/2 dynamics; we will introduce more complicated spin Hamiltonians when they become necessary throughout the rest of the thesis. Importantly, we give a brief sketch of the rotating wave approximation that is used when computing the dynamics of spins in static fields that are considerably stronger than the amplitudes of time-dependent magnetic fields employed for driving the spin system.

- Chapter 2 will introduce the concept of effective Hamiltonians by considering two analytical paradigms for Hamiltonian engineering: average Hamiltonian theory (AHT) [41] and stochastic Liouville theory [58]. We will argue that effective Hamiltonian engineering translates to setting certain perturbative expressions to desired values. Specifically, we say that if a quantum system is controlled over a period $0 \leq t \leq T$, and the unitary evolution over this time period, as is generated by the experimenter's best characterization of the system's internal and control Hamiltonians,

is $U(0 \leq t \leq T)$, then successful engineering of the desired effective evolution boils down to ensuring that $U(T)$ and certain variations of it take a desired form. These variations can generally be expressed as time-dependent perturbation theory expressions of the following form:

$$\int_0^T dt_1 \int_0^{t_1} dt_2 ... \int_0^{t_{n-1}} dt_n \; f(t_1, t_2, ..., t_n) \; U^{-1}(t_1) A(t_1) U(t_1) \; ... \; U^{-1}(t_n) A(t_n) U(t_n), \tag{1.1}$$

where $f(t_1, t_2, ..., t_n)$ is a scalar function and $\{A(t_i)\}$ is a set of, possibly time-dependent, operators.

With Chapter 2 we review that AHT prescribes a systematic approach for setting perturbation theory integrals of the kind in Equation (1.1) with $f = 1$ to some desired values. Here, we remark that AHT proved an indispensable tool for the development of a vast number of magnetic resonance control sequences, e.g., dipolar sequences [121, 71, 95, 79, 109, 17], composite pulses for control and internal Hamiltonian variations [61], imaging sequences [18] and many more. We will also show in Chapter 2 that perturbation theory terms with $f \neq 1$ in Equation (1.1) appear when solving for the ensemble averaged evolution of a quantum system under stochastic operators, as is the case for stochastic Liouville theory [58]. $f(t_1, t_2, ..., t_n)$ will then be composed of correlation functions that characterize the stochastic operators. Analytic control design seeking to minimize nested integrals of that kind was performed in [13].

- Chapter 3 is devoted to developing a framework for efficient evaluation of expressions appearing in Equation (1.1), which is very easy to incorporate into numerical control design algorithms.

  Achieving the most efficient and accurate control of any quantum system, or an ensemble of quantum systems, requires tailoring of control sequences for the particular experimental setup and physical system at hand. When it comes to flexible tailored control design, numerical control optimization has a number of advantages over analytical control design: (i) it can easily deal with simultaneous control of an ensemble [7, 63], (ii) even though it is limited to smaller Hilbert space dimensions, it is not specific to any Hilbert space dimension [48], (iii) it can accommodate any experimental constraints present for the specific hardware configuration, e.g., amplitude and bandwidth constraints for the control waveform [82, 105, 6], (iv) it can account for deterministic control distortions due to control hardware [47, 26], and (v) it stands a better chance of yielding control sequences that are closer to being time optimal than analytic solutions, given (iii) and (iv).

Given the benefits of both – perturbative tools and numerical control design – there has been increasing interest in numerical engineering of control sequences that implement effective Hamiltonians. A filter function formalism for mitigating the effect of stochastic noise in quantum control was introduced by Green *et al* [34, 35, 89], and has been combined with gradient free numerical optimization, which lead to experimental advancements [104]. Although there are individual, problem specific, numerical approaches that have previously been taken [88, 33], a complete framework for numerical control optimization that would yield a desired value for $U(T)$, simultaneously with values for an arbitrary set of perturbation terms, has so far been lacking.

With Chapter 3 we provide a general method for the numerical evaluation of $U(T)$ simultaneously with the evaluation, or arbitrarily close approximation, of any number of nested integrals in Equation (1.1). Furthermore, the method also enables straightforward computation of gradients of these integrals, which is crucial for efficiently searching large control landscapes when numerically optimizing quantum control sequences. We demonstrate that such integrals can be evaluated by solving particular block matrix differential equations, that we call Van Loan equations, which have the same form as the Schrödinger equations, that determine $U(t)$. We demonstrate the efficiency and applicability of our framework by conducting four control searches that involve various nested integral terms and one control search that deals with stochastic noise process. The latter also implements various experimentally realistic constraints on the pulse waveform, including amplitude and pulse bandwidth limitations, and a demand that the ends of the pulse waveform go smoothly to zero.

Chapter 3 is largely based on the results in [39], a research project conducted by the author in collaboration with Daniel Puzzuoli.

- Chapter 4 comprises an experimental implementation and verification of the Van Loan numerical control engineering framework developed in Chapter 3. We used our tools for nanoscale magnetic resonance imaging experiments which posed a very challenging control setting – we were dealing with an ensemble of strongly dipolar coupled proton spins that experienced a vast Rabi frequency variation from 0.6 MHz to 1.3 MHz, while the phase coherence time T_2 of the coupled spins was 11 μs. Our numerical tools helped us find control sequences that yielded a $\pi/2$ unitary rotation that was insensitive to first order perturbations due to dipolar and chemical shift Hamiltonians for the entire spin ensemble simultaneously. Even though the rotation took 7.5 μs to implement, it enabled an increase of spin T_2 by a factor of $\sim$500. We also demonstrate numerical simulations which indicate strongly that such elongation of

the phase coherence time would not have been possible without our effective Hamiltonian engineering tools. We made use of this enhanced coherence time by performing one-dimensional magnetic resonance imaging experiments on the same proton spin sample with a resolution of ∼2 nm. Crucially, we performed the experiments on virtually the same setup that was limited to ∼10 nm spatial resolution in previous imaging experiments due to phase coherence time resolution limitations [85].

Chapter 4 is largely based on the published results in [96], which the author conducted in collaboration with William Rose and Raffi Budakian.

- Chapter 5 introduces low temperature properties of phosphorus defects in silicon at low doping concentrations, with a particular focus on properties relevant for magnetic resonance experiments. We will specifically discuss phosphorus nuclear spins (^{31}P) embedded at dilute concentrations in high quality isotopically purified silicon-28 (^{28}Si) single crystals that have been demonstrated to have extremely long coherence times both at cryogenic and room temperatures [98]. These nuclei in that material are the subject of our interest in Chapter 6. Consequently, the nuclei have been proposed as a candidate for quantum information storage in a scalable, fault-tolerant quantum computer architecture [56, 106, 113]. The same long coherence time, coupled with various efficient read out schemes, also make the ^{31}P nuclei in silicon a viable candidate for extremely high accuracy, high sensitivity magnetometry. A crucial challenge for realizing the latter is to establish a fast and simple means for hyperpolarizing the nuclear spins in order to enhance the measurement efficiency and sensitivity. It is known experimentally that the ^{31}P nuclear spins can be hyperpolarized by illuminating the silicon crystal with above band gap light at high magnetic fields and cryogenic temperatures [77, 38]. With Chapter 5 we will summarize the aforementioned excellent coherence properties of the defect nuclei, we will also outline two proposed candidate mechanisms explaining the low temperature optical hyperpolarization schemes.

- Chapter 6 presents our studies of nuclear hyperpolarization of ^{31}P donor nuclei in an isotopically enriched 28-silicon crystal under the illumination of above band gap 1047 nm laser light at a high magnetic field of 6.7 T at low temperature of 1.3 K. We demonstrate spin polarization growth with a time constant of 51.5 s and introduce a microscopic effective Hamiltonian, phonon mediated model detailing the process. In order to verify our model, we demonstrate experimentally that the nuclear polarization build up is independent of the local light intensity variations inside the crystal and uncorrelated with donor ionization events under the above band gap illumination. We attribute the photoionization to Auger recombination that follows free

exciton capture at the donor defect sites, hence, implying that the hyperpolarization mechanism is unrelated to the exciton capture. To study such events we introduce a magnetic resonance technique that directly characterizes the donor defect ionization process. This previously unexploited way of probing the photoionization enables us to measure both the rate and the average duration of the defect site ionization under our illumination, temperature and field conditions.

The studies in Chapter 6 were conducted by the author in collaboration with Thomas Alexander and Rahul Deshpande.

- Chapter 7 provides concluding remarks.

1.2 Brief Introduction to Magnetic Resonance

In this thesis, we deal with Hamiltonian engineering in the context of magnetic resonance, hence, all problems will be stated in terms of spin Hamiltonians. Nevertheless, it should be understood that in most cases, as far as the mathematical treatment goes, the the word 'spin' is interchangeable with 'finite level quantum system'. Here, we will give a brief mathematical treatment of magnetic resonance concepts that govern the dynamics of a single isolated spin-1/2 nucleus and a single isolated electron, that experiences negligible spin-orbit coupling, when placed in magnetic fields [62]. Nuclear spins that have spin numbers greater than 1/2 also couple to electric field gradients through their non-zero electric quadrupolar moment, and electrons in solid state matter with non-zero spin-orbit coupling can exhibit effective couplings to electric as well as strain fields. For the sake of brevity, we will ignore these aspects of spin dynamics for now along with multi-spin Hamiltonians, which will be introduced as they become necessary in the rest of this thesis. Furthermore, this section does not touch on the experimental aspects of magnetic resonance, such explanations will also be reserved for the upcoming chapters.

The spin degree of freedom for a single isolated spin-1/2 particle forms a two-dimensional complex Eucledian vector space, usually referred to as the Hilbert space. All through this thesis, we work in the Schrödinger picture, i.e., we say that the spin state of a system at any time t is fully characterized by state vector $|\psi(t)\rangle$. $|\psi(t)\rangle$ can be decomposed as a linear superposition of two ortho-normal basis vectors, which we will denote as $|\uparrow\rangle$ and $|\downarrow\rangle$ throughout this thesis: $|\psi(t)\rangle = c_\uparrow(t)\,|\uparrow\rangle + c_\downarrow(t)\,|\downarrow\rangle$, where $c_\uparrow(t)$ and $c_\downarrow(t)$ are complex numbers for which it holds that $|c_\uparrow(t)|^2 + |c_\downarrow(t)|^2 = 1$. We often work with an explicit vector representation of $|\psi(t)\rangle$ in that basis, in which case we identify $|\psi(t)\rangle$ with a complex unit

vector

$$\psi(t) = \begin{pmatrix} \langle\uparrow|\psi(t)\rangle \\ \langle\downarrow|\psi(t)\rangle \end{pmatrix} = \begin{pmatrix} c_\uparrow(t) \\ c_\downarrow(t) \end{pmatrix}. \tag{1.2}$$

When a spin-1/2 particle is placed in magnetic field $\vec{B}(t)$, the spin state $|\psi(t)\rangle$ evolves according to the Schrödinger equation:

$$i\hbar\frac{d}{dt}|\psi(t)\rangle = H(t)|\psi(t)\rangle, \tag{1.3}$$

where the Hamiltonian $H(t)$ is the Zeeman Hamiltonian:

$$H(t) = -\vec{\mu}\cdot\vec{B}(t), \tag{1.4}$$

$\vec{\mu}$ being the magnetic moment proportional to spin angular momentum of the particle: $\vec{\mu} = \frac{\gamma\hbar}{2}(\sigma_x, \sigma_y, \sigma_z)$. γ is called the gyromagnetic ratio and is specific to a free electron or a particular nuclear isotope, while σ_i, $i \in \{x, y, z\}$ denotes a Pauli operator. From this point on, throughout this thesis, we take $\hbar = 1$, which implies that we express our Hamiltonians in units of frequency rather than energy. Furthermore, we use a convention that designates $|\uparrow\rangle$ and $|\downarrow\rangle$ as the $+1$ and -1 eigenvalue eigenstates of the σ_z operator, and we generally use the matrix representation of $\{\sigma_i\}$ for the $\{|\uparrow\rangle, |\downarrow\rangle\}$ basis interchangeably with the operators themselves, i.e., we assert that

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},\ \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix},\ \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.5}$$

Since the evolution of $|\psi(t)\rangle$ in Equation (1.3) is generated by a Hermitian operator, the state vector $|\psi(t=0)\rangle$ is transformed to $|\psi(t>0)\rangle$ by a unitary operator $U(t)$, such that $|\psi(t)\rangle = U(t)|\psi(0)\rangle$, and $U(t)U^\dagger(t) = \mathbb{1}$, where $\mathbb{1}$ is the identity operator. $U(t)$ itself is a solution to the Schrödinger equation

$$\frac{d}{dt}U(t) = -iH(t)U(t), \tag{1.6}$$

with $U(0) = \mathbb{1}$. If $H(t_1)$ commutes with itself for all times $t_1 \in [0, t]$, then

$$U(t) = \exp\left[-i\int_0^t dt_1\ H(t_1)\right], \tag{1.7}$$

otherwise, we can express $U(t)$ either in terms of Dyson series [22]:

$$U(t) = \mathbb{1} - i\int_0^t dt_1\ H(t_1) + (-i)^2\int_0^t dt_1\int_0^{t_1} dt_2\ H(t_1)H(t_2) + ..., \tag{1.8}$$

or in some cases through Magnus series [70, 5]:

$$U(t) = \exp\left(-i\int_0^t dt_1\ H(t_1) + \frac{(-i)^2}{2}\int_0^t dt_1\int_0^{t_1} dt_2[H(t_1), H(t_2)] + ...\right), \tag{1.9}$$

where $[A, B] = AB - BA$ denotes a matrix commutator.

1.2.1 Rotating Wave Approximation

With this subsection we tie the mathematical nomenclature introduced so far in this section to the physical phenomenon of magnetic resonance. In the case of most magnetic resonance experiments, the spins are placed in a strong static external magnetic field [62], the strength of which is denoted by B_0 and its direction is conventionally taken to be along the z-axis, i.e., $\vec{B} = (0, 0, B_0)$. The resulting *drift* Hamiltonian is then $H_{\text{drift}} = \gamma B_0\sigma_z/2 = \omega_0\sigma_z/2$, where the quantity ω_0 is called the Larmor frequency of the spins. It easy to see that the energy difference between the ground and the excited state of H_{drift} – $|\downarrow\rangle$ and $|\uparrow\rangle$, respectively – is $\hbar\omega_0$.

We will now show that electro-magnetic radiation, specifically alternating magnetic field resonant with ω_0, drives the spin state between $|\downarrow\rangle$ and $|\uparrow\rangle$. It is this exact resonant phenomenon that gives magnetic resonance its name. We begin by considering a different, time-dependent, component to the $\vec{B}(t)$ vector, that points orthogonal to the z-axis. Here, with no loss of generality, we take it to point along the x-axis:

$$\vec{B}_1(t) = B_1\left(\cos(\omega_t t - \phi), 0, 0\right), \tag{1.10}$$

where B_1 is referred to as the *Rabi field* and ϕ is some phase constant, while ω_t is the *transmitter frequency* of the AC magnetic field, named so, because it is typically generated by radio-frequency (RF) or microwave ($\mu\omega$) synthesizers. The corresponding spin Hamiltonian becomes

$$\begin{aligned} H(t) &= \omega_1\cos(\omega_t t - \phi)\sigma_x + \omega_0\frac{\sigma_z}{2} \\ &= e^{i\omega_t\frac{\sigma_z}{2}t}\left[\omega_1\left(\frac{\sigma_x}{2}\cos(\phi) + \frac{\sigma_y}{2}\sin(\phi)\right) + \omega_0\frac{\sigma_z}{2} + \omega_1 e^{-i(\omega_t t-\phi/2)\sigma_z}\frac{\sigma_x}{2}e^{i(\omega_t t-\phi/2)\sigma_z}\right]e^{-i\omega_t\frac{\sigma_z}{2}t}, \end{aligned} \tag{1.11}$$

with $\omega_1 = \gamma B_1/2$ being referred to as the *Rabi frequency*. For solving Equation (1.3) for $H(t)$ given by Equation (1.11), we first make an ansatz that $|\psi(t)\rangle = e^{i\omega_t \frac{\sigma_z}{2} t} |\widetilde{\psi}(t)\rangle$, where $|\widetilde{\psi}(t)\rangle$ is called the state vector in the *frame* of $\omega_t \frac{\sigma_z}{2}$ Hamiltonian. Inserting this ansatz into Equation (1.3), and collecting some terms, yields a differential equation for $|\widetilde{\psi}(t)\rangle$:

$$\begin{aligned} i\hbar \frac{d}{dt} |\widetilde{\psi}(t)\rangle &= \left[\omega_1 \left(\frac{\sigma_x}{2}\cos(\phi) + \frac{\sigma_y}{2}\sin(\phi)\right) + \omega_0 \frac{\sigma_z}{2} + \omega_1 e^{-i(\omega_t t - \phi/2)\sigma_z} \frac{\sigma_x}{2} e^{i(\omega_t t - \phi/2)\sigma_z}\right] |\widetilde{\psi}(t)\rangle \\ &\approx \left[\omega_1 \left(\frac{\sigma_x}{2}\cos(\phi) + \frac{\sigma_y}{2}\sin(\phi)\right) + \Delta\omega \frac{\sigma_z}{2}\right] |\widetilde{\psi}(t)\rangle \,, \end{aligned} \tag{1.12}$$

with $\Delta\omega = \omega_0 - \omega_t$; the omission of the time-dependent term on the second line is called the *rotating wave approximation* (RWA). RWA is a very good approximation as long as $\omega_1 \ll \omega_0$ and $|\Delta\omega| \ll \omega_0$, which is a regime where most magnetic resonance experiments are performed. It is now clear that

$$|\psi(t)\rangle = e^{i\omega_t \frac{\sigma_z}{2} t} e^{-i\left[\omega_1\left(\frac{\sigma_x}{2}\cos(\phi) + \frac{\sigma_y}{2}\sin(\phi)\right) + \Delta\omega\frac{\sigma_z}{2}\right]t} |\psi(0)\rangle \tag{1.13}$$

or

$$|\widetilde{\psi}(t)\rangle = U(t)\,|\widetilde{\psi}(0)\rangle = e^{-i\left[\omega_1\left(\frac{\sigma_x}{2}\cos(\phi) + \frac{\sigma_y}{2}\sin(\phi)\right) + \Delta\omega\frac{\sigma_z}{2}\right]t} |\widetilde{\psi}(0)\rangle \,. \tag{1.14}$$

In the rest of this thesis, unless stated otherwise, we work implicitly in the frame of $\omega_0 \frac{\sigma_z}{2}$ of the spin Hamiltonian, i.e., we only concern ourselves with the evolution of $|\widetilde{\psi}(t)\rangle$ for resonant $\Delta\omega = 0$ drive fields. This makes physical sense because the detection of $|\psi(t)\rangle$ in magnetic resonance experiments happens generally either at the transmitter frequency ω_t, as is the case for inductive detection, or as an effective measurement of σ_z expectation value which coincides for $|\widetilde{\psi}(t)\rangle$ and $|\psi(t)\rangle$. Consequently, we regard the Rabi fields defined by Equation (1.10) to always be modulated by transmitter or *carrier* frequency $\omega_t = \omega_0$, the resulting rotating frame Hamiltonian being $\omega_1 \left(\frac{\sigma_x}{2}\cos(\phi) + \frac{\sigma_y}{2}\sin(\phi)\right)$. It is easy to see that such Rabi field drives the state $|\widetilde{\psi}(t)\rangle$ between $|\downarrow\rangle$ and $|\uparrow\rangle$ with Rabi frequency ω_1. The transitions $|\downarrow\rangle \leftrightarrow |\uparrow\rangle$ are called *Rabi oscillations*.

1.2.2 Open Quantum System Dynamics

For most real spin systems, the state vector $|\widetilde{\psi}(t)\rangle$ description given above does not fully encapsulate the measurement results in magnetic resonance experiments. In general a density matrix description is necessary to account for ensemble effects or to describe the dynamics of some particular spin degree of freedom within a larger quantum system. With

this subsection, we give a brief sketch of the density matrix description in the case of open quantum system setting. The density matrix $\rho(t)$ is a trace-one, positive semi-definite operator, meaning that it can always be written as a sum of orthogonal rank-one projectors, the real eigenvalues of which are always non-negative and add to one. In that sense $\rho(t)$, describing a spin-1/2 system, can be regarded as a probabilistic mixture of two orthogonal states $\{|a(t)\rangle, |b(t)\rangle\}$:

$$\rho(t) = p(t) |a(t)\rangle \langle a(t)| + [1 - p(t)] |b(t)\rangle \langle b(t)|, \tag{1.15}$$

where $0 \leq p(t) \leq 1$. $\rho(t)$ is a mathematical object that packages expectation values for any measurement performed only on the spin degree of freedom of the particle. For $p(t) = \{0, 1\}$, the density matrix description reduces back to *pure state* description given above, such that $\rho(t) = |\widetilde{\psi}(t)\rangle \langle \widetilde{\psi}(t)|$.

Under certain assumptions about the underlying quantum system dynamics [10, 64], which ensure that the evolution of $\rho(t)$ follows a quantum dynamical semigroup, the time-dependence of $\rho(t)$ is governed by Lindblad equation:

$$\frac{d}{dt}\rho(t) = -i[H(t), \rho(t)] + \sum_j a_j(t) \left[A_j(t)\rho(t)A_j^\dagger(t) - \frac{1}{2}\{A_j^\dagger(t)A_j(t), \rho(t)\} \right], \tag{1.16}$$

where $\{A, B\} = AB + BA$ denotes a matrix anti-commutator, $\{a_j(t)\}$ is a set of non-negative real constants and $\{A_j(t)\}$ is a set of operators. If $A_j(t) = 0$ for all j, Equation (1.16) reduces to Liouville equation and the evolution of $\rho(t)$ is unitary. Equation (1.16) enforces that for any initial trace-one, positive semi-definite $\rho(0)$ the final density matrix $\rho(t)$ is also a trace-one, positive semi-definite operator.

1.2.3 Bloch Equations

In this subsection, we deduce a particular set of phenomenological equations describing the evolution of a spin-1/2 system through $\rho(t)$, called the Bloch equations. Let $H(t) = \omega_1 \left(\frac{\sigma_x}{2}\cos(\phi) + \frac{\sigma_y}{2}\sin(\phi)\right) + \Delta\omega\frac{\sigma_z}{2}$ and $a_1(t) = s/T_1$, $A_1(t) = \sigma_+$, $a_2(t) = (1-s)/T_1$, $A_2(t) = \sigma_-$, $a_3(t) = 1/T_2 - 1/(2T_1)$, $A_3(t) = \sigma_z/\sqrt{2}$ and $0 < T_2 \leq 2T_1$, $0 \leq s \leq 1$ in the Lindblad equation defined by Equation (1.16), then

$$\begin{aligned}\frac{d}{dt}\rho(t) = &-i\left[\omega_1\left(\frac{\sigma_x}{2}\cos(\phi) + \frac{\sigma_y}{2}\sin(\phi)\right) + \Delta\omega\frac{\sigma_z}{2}, \rho(t)\right] + \left(\frac{1}{2T_2} - \frac{1}{4T_1}\right)(\sigma_z\rho(t)\sigma_z - \rho(t)) \\ &+ \frac{s}{T_1}\left(\sigma_+\rho(t)\sigma_- - \frac{1}{2}\{\sigma_-\sigma_+, \rho(t)\}\right) + \frac{1-s}{T_1}\left(\sigma_-\rho(t)\sigma_+ - \frac{1}{2}\{\sigma_+\sigma_-, \rho(t)\}\right). \end{aligned} \tag{1.17}$$

σ_+ and $\sigma_- = \sigma_+^\dagger$ are the spin raising and lowering operators, respectively,

Equation (1.17) is a first order operator differential equation for $\rho(t)$, which can be expressed as a matrix differential equation, just as the Schrödinger equation in Equation (1.3), when a particular operator basis is chosen for $\rho(t)$. Here, we choose $\rho(t) = \frac{1}{2}\left(r_{\mathrm{I}}(t)\mathbb{1} + r_x(t)\sigma_x + r_y(t)\sigma_y + r_z(t)\sigma_z\right)$, where $\vec{r}(t) = (r_x(t), r_y(t), r_z(t))$ is a vector in $\mathbb{R}^3$, and is called the *Bloch vector*, which has to satisfy $|\vec{r}(t)| \leq 1$, as well as $r_{\mathrm{I}}(t) = 1$ for $\rho(t)$ to be a trace-one semi-definite operator. Fixing an operator basis for $\rho(t)$ and expressing the density matrix as a four-dimensional vector will be referred to as *vectorization* throughout this thesis, and we will often denote it as $|\rho(t)\rangle$, importantly, in many cases we do not explicitly write out the basis but merely rely on its existence. For deriving Bloch equations, we identify $|\rho(t)\rangle = (r_x(t), r_y(t), r_z(t), r_{\mathrm{I}}(t))$, where the elements of $|\rho(t)\rangle$ can be found as $r_i(t) = \mathrm{Tr}\left[\sigma_i^\dagger \rho(t)\right]$, for $i \in \{x, y, z\}$, while $r_{\mathrm{I}}(t) = \mathrm{Tr}\left[\mathbb{1}^\dagger \rho(t)\right] = 1$. Using this identification, we can rewrite Equation (1.17) as

$$\frac{d}{dt}|\rho(t)\rangle = \begin{pmatrix} -\frac{1}{T_2} & -\Delta\omega & \omega_1 \sin(\phi) & 0 \\ \Delta\omega & -\frac{1}{T_2} & -\omega_1 \cos(\phi) & 0 \\ -\omega_1 \sin(\phi) & \omega_1 \cos(\phi) & -\frac{1}{T_1} & \frac{2s-1}{T_1} \\ 0 & 0 & 0 & 0 \end{pmatrix} |\rho(t)\rangle\,. \tag{1.18}$$

The 4×4 matrix above is referred to as the Bloch equations for a spin-1/2 system. Solving the equations for $\omega_1 = \Delta\omega = 0$ and any $\rho(0) = \frac{1}{2}\left(\mathbb{1} + r_x(0)\sigma_x + r_y(0)\sigma_y + r_z(0)\sigma_z\right)$ yields

$$\begin{aligned} &\rho(t) = \\ &\frac{1}{2}\left(\mathbb{1} + r_x(0)e^{-\frac{t}{T_2}}\sigma_x + r_y(0)e^{-\frac{t}{T_2}}\sigma_y + \left[r_z(0)e^{-\frac{t}{T_1}} + (2s-1)\left(1 - e^{-\frac{t}{T_1}}\right)\right]\sigma_z\right), \end{aligned} \tag{1.19}$$

which elucidates the physical interpretation of the quantities s, T_1 and T_2, and the corresponding operators, that were inserted into the Lindblad equation above in a seemingly ad hoc manner. T_2 is called the dephasing, or the transverse relaxation time, and it determines the time scale over which the coherence, i.e., x and y components of $\rho(t)$, decay to zero. Such dephasing is a universal phenomenon in the case of magnetic resonance experiments, and arises due to ensemble effects and/or due to noise, furthermore, dephasing also appears when the spin degree of freedom couples to external systems. The fact that x and y operators are treated on an equal footing is also universal and arises because the external magnetic field B_0 merely determines a symmetry axis for the spin system. T_1 is called the longitudinal relaxation time, and it characterizes the time scale it takes for the spin system to return to thermal equilibrium state $\rho_{\mathrm{thermal}} = \frac{1}{2}\left[\mathbb{1} + (2s-1)\sigma_z\right]$, where $0 \leq s \leq 1$ is often determined by Boltzmann distribution such that $s = 1/2\left(1 - \tanh\left[\Theta\omega_0\hbar/(2k_B)\right]\right)$.

Θ is inverse temperature and k_B the Boltzmann constant. Finally, we remark that this thesis is largely concerned with elongating T_2 values and engineering effective T_1 processes with s-values greater than $1/2\,(1 - \tanh\left[\Theta\omega_0\hbar/(2k_B)\right])$ in order to enhance the sensitivity and efficiency of various magnetic resonance experiments.

Chapter 2

Coherent Averaging and Effective Hamiltonians

With this short chapter we give a brief review of the earliest two effective Hamiltonian analytical treatments: average Hamiltonian theory (AHT) introduced by Haeberlen and Waugh [41] in 1968 and stochastic Liouville theory introduced by Kubo [58] in 1963. We do not intend our treatment here to be comprehensive, since the use of AHT and stochastic Lioville methods is incredibly broad. Our main aim is to motivate the next two chapters, and specifically, to demonstrate that both of the analytical treatments work with nested integrals of the following form:

$$\int_0^T dt_1 \int_0^{t_1} dt_2 ... \int_0^{t_{n-1}} dt_n \; f(t_1, t_2, ..., t_n) \; U^{-1}(t_1)A(t_1)U(t_1) \; ... \; U^{-1}(t_n)A(t_n)U(t_n), \quad (2.1)$$

where f is a scalar function. In the case of AHT $f = 1$, whereas for stochastic Lioville theory f is generally a non-trivial function of $\{t_1, t_2, ..., t_n\}$. We also give an explicit AHT treatment of the magic echo sequence [74, 42] that will be used extensively in Chapter 4.

This chapter contains no significant original contributions by the author.

2.1 Average Hamiltonian Theory

AHT was introduced in the context of spin Hamiltonians in magnetic resonance experiments. Haeberlen and Waugh [41] considered a spin system driven by some time-dependent

Hamiltonian $H_1(t)$, that evolves under a total Hamiltonian

$$H(t) = H_1(t) + H_{\text{int}}, \tag{2.2}$$

where H_{int} is some time-independent internal Hamiltonian, the evolution under which is unwanted. In the next chapter, it will be shown by Equation (3.12) that the evolution under $H(t)$ can be partially separated into evolution under $H_1(t)$ and another term. This is done through the concept of toggling frame. Here, we just assert that we can write

$$U(t) = U_1(t)U_{\text{int}}(t), \tag{2.3}$$

where $U(t)$, $U_1(t)$ and $U_{\text{int}}(t)$ are solutions to the following respective Schrödinger equations: $\dot{U}(t) = -iH(t)U(t)$, $\dot{U}_1(t) = -iH_1(t)U_1(t)$ and $\dot{U}_{\text{int}}(t) = -i\widetilde{H}_{\text{int}}(t)U_{\text{int}}(t)$, with $U(0) = U_1(0) = U_{\text{int}}(0) = \mathbb{1}$, and $\widetilde{H}_{\text{int}}(t) = U_1^\dagger(t)H_{\text{int}}U_1(t)$ being called the toggling frame Hamiltonian.

Because $H_1(t)$ is assumed to be totally under the experimenter's control, hence, resulting in an ability to generate any arbitrary $U_1(t)$, Haeberlen and Waugh [41] analysed the effect of $U_{\text{int}}(t)$ on the spin system evolution. They assumed that $H_1(t)$ is *periodic* and *cyclic* over time t_c, implying that

$$H_1(t + Nt_c) = H_1(t) \text{ and} \tag{2.4}$$
$$U_1(Nt_c) = \mathbb{1}, \tag{2.5}$$

for any $N = 1, 2, \ldots$. These two conditions ensure that the toggling frame Hamiltonian is also periodic, i.e.,

$$\widetilde{H}_{\text{int}}(t + Nt_c) = \widetilde{H}_{\text{int}}(t), \tag{2.6}$$

with the same period as $H_1(t)$. The cycle time t_c is usually considered to be a period sufficiently shorter than the characteristic time scale for evolving under H_{int}, and the above implies $U_{\text{int}}(Nt_c) = [U_{\text{int}}(t_c)]^N$.

Haeberlen's and Waugh's [41] treatment considers spin system state at integer intervals of t_c, which translates to considering the properties of $U(Nt_c) = U_1(Nt_c)U_{\text{int}}(Nt_c) = [U_{\text{int}}(t_c)]^N$. In that way, only a short cycle period evolution under the toggling frame Hamiltonian has to be considered, which can be done using Magnus expansion introduced in the previous chapter:

$$U(t_c) = \exp\left(-it_cF\right), \tag{2.7}$$

where F is a time-independent *average Hamiltonian* given as $F = \overline{H}_{\text{int}} + \overline{H}^{(1)}_{\text{int}} + \overline{H}^{(2)}_{\text{int}} + \dots$, such that

$$\overline{H}_{\text{int}} = \frac{1}{t_c}\int_0^{t_c} dt_1\ \widetilde{H}_{\text{int}}(t_1), \tag{2.8}$$

$$\overline{H}^{(1)}_{\text{int}} = \frac{-i}{2t_c}\int_0^{t_c} dt_1 \int_0^{t_1} dt_2 [\widetilde{H}_{\text{int}}(t_1), \widetilde{H}_{\text{int}}(t_2)], \tag{2.9}$$

$$\begin{aligned}\overline{H}^{(2)}_{\text{int}} = &\frac{1}{6t_c}\int_0^{t_c} dt_1 \int_0^{t_1} dt_2 \int_0^{t_2} dt_3 [\widetilde{H}_{\text{int}}(t_1), [\widetilde{H}_{\text{int}}(t_2), \widetilde{H}_{\text{int}}(t_3)]] \\ &+ \frac{1}{6t_c}\int_0^{t_c} dt_1 \int_0^{t_1} dt_2 \int_0^{t_2} dt_3 [\widetilde{H}_{\text{int}}(t_3), [\widetilde{H}_{\text{int}}(t_2), \widetilde{H}_{\text{int}}(t_1)]] \\ &\vdots\end{aligned} \tag{2.10}$$

Since F is time independent we see from Equation (2.7) that $U(Nt_c) = \exp(-iNt_cF)$, which is the unitary operator that entirely determines the spin system dynamics. Finally, AHT prescription for making the spin system evolve as if the H_{int} Hamiltonian term is not present, is to pick a $H_1(t)$ that satisfies the conditions in Equation (2.4) and (2.5), while setting Magnus terms up to some order to zero, i.e, $\overline{H} = \overline{H}^{(1)} = \dots = \overline{H}^{(n)} = 0$. For short enough t_c, one then has $U(Nt_c) \approx \mathbb{1}$. We also point out that all Magnus terms appearing in Equation (2.8) can be composed of nested integrals in Equation (2.1), with $f = 1$, because $\widetilde{H}_{\text{int}}(t) = U_1^\dagger(t) H_{\text{int}} U_1(t)$.

2.1.1 Magic Echo Sequence

In this subsection, we provide an AHT analysis of a particular, idealized NMR pulse sequence, called the magic echo [74, 42]. The pulse sequence is depicted in Figure 2.1, and consists of two $\pi/2$ rotations, that are assumed to be infinitesimally short, sandwiching a rotary echo implemented with Rabi frequency Ω. The same pulse sequence will form the basis for most of the experiments in Chapter 4. We will specifically focus on the AHT terms arising due to resonance offset/chemical shift Hamiltonian and spin-spin dipolar coupling Hamiltonian – the former could be written as $H_{\text{int}} = \Delta\omega \frac{\sigma_z}{2}$, $\Delta\omega$ denoting the resonance offset/chemical shift for a particular spin, while the latter is given as $H_{\text{int}} = \xi\left(3\frac{\sigma_z}{2}\otimes\frac{\sigma_z}{2} - \sum_{k\in\{x,y,z\}}\frac{\sigma_k}{2}\otimes\frac{\sigma_k}{2}\right)$, where ξ specifies the strength of the dipolar interaction between a pair of spins.

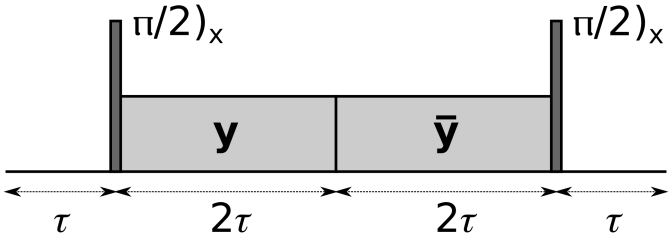

Figure 2.1: Magic echo sequence: the sequence consists of τ-long zero-amplitude periods at the beginning and at the end of the sequence, and two rotary echo periods with opposite phases, each of length 2τ that are sandwiched by $\pi/2$ rotations around the x-axis. The rotary echo periods are denoted by $\mathbf{y}$ and $\bar{\mathbf{y}}$, and stand for a continuous Rabi drive along the y and negative y-axis, respectively.

We now find the zeroth order average Hamiltonian $\overline{H}_{\text{int}} = \frac{1}{6\tau}\int_0^{6\tau} dt\ \widetilde{H}_{\text{int}}(t)$ for the two H_{int}'s over the 6τ-long duration of the magic echo sequence. The unitary $U_1(t)$ is generated by the Hamiltonian $H_1(t)$, which is determined by the pulse amplitudes in Figure 2.1. We can therefore write $U_1(t)$ as

$$U_1(t) = \begin{cases} \mathbb{1}, \text{ for } 0 \le t < \tau \\ \exp\left[-i\Omega\frac{\sigma_y}{2}(t-\tau)\right]\exp\left(-i\frac{\pi}{2}\frac{\sigma_x}{2}\right), \text{ for } \tau \le t < 3\tau \\ \exp\left[-i\Omega\frac{\sigma_y}{2}(5\tau-t)\right]\exp\left(-i\frac{\pi}{2}\frac{\sigma_x}{2}\right), \text{ for } 3\tau \le t < 5\tau \\ \exp\left(-i\pi\frac{\sigma_x}{2}\right), \text{ for } 5\tau \le t \le 6\tau. \end{cases} \tag{2.11}$$

We first evaluate $\widetilde{H}_{\text{int}}(t) = U_1^\dagger(t) H_{\text{int}} U_1(t)$ for $H_{\text{int}} = \Delta\omega\frac{\sigma_z}{2}$:

$$\widetilde{H}_{\text{int}}(t) = \begin{cases} \Delta\omega\frac{\sigma_z}{2}, \text{ for } 0 \le t < \tau \\ \Delta\omega\left(\cos\left[\Omega(t-\tau)\right]\frac{\sigma_y}{2} - \sin\left[\Omega(t-\tau)\right]\frac{\sigma_x}{2}\right), \text{ for } \tau \le t < 3\tau \\ \Delta\omega\left(\cos\left[\Omega(5\tau-t)\right]\frac{\sigma_y}{2} - \sin\left[\Omega(5\tau-t)\right]\frac{\sigma_x}{2}\right), \text{ for } 3\tau \le t < 5\tau \\ -\Delta\omega\frac{\sigma_z}{2}, \text{ for } 5\tau \le t \le 6\tau. \end{cases} \tag{2.12}$$

Correspondingly, the zeroth order average Hamiltonian for chemical shifts and resonance offsets under the magic echo sequence is

$$\overline{H}_{\text{int}} = \frac{1}{6\tau}\int_0^{6\tau} dt\ \widetilde{H}_{\text{int}}(t) = \frac{\Delta\omega}{3\tau\Omega}\left(\sin(2\Omega\tau)\frac{\sigma_y}{2} + \left[\cos(2\Omega\tau) - 1\right]\frac{\sigma_x}{2}\right). \tag{2.13}$$

It is clear that $\overline{H}_{\text{int}}$ becomes zero in the limit of $\Omega\tau \to \infty$.

For $H_{\text{int}} = \xi\left(3\frac{\sigma_z}{2}\otimes\frac{\sigma_z}{2} - \sum_{k\in\{x,y,z\}}\frac{\sigma_k}{2}\otimes\frac{\sigma_k}{2}\right)$, we follow an analogous procedure as above. This time, we have $\widetilde{H}_{\text{int}}(t) = \left[U_1^\dagger(t)\otimes U_1^\dagger(t)\right] H_{\text{int}}\left[U_1(t)\otimes U_1(t)\right]$ since H_{int} is a two-spin Hamiltonian. Furthermore, we notice that $\sum_{k\in\{x,y,z\}}\left(U_1^\dagger\otimes U_1^\dagger\right)(\sigma_k\otimes\sigma_k)\left(U_1\otimes U_1\right) = \sum_{k\in\{x,y,z\}}\sigma_k\otimes\sigma_k$ for any unitary operation U_1. The last identity arises because $\sum_{k\in\{x,y,z\}} = 2U_{\text{SWAP}} - \mathbb{1}\otimes\mathbb{1}$, where U_{SWAP} denotes the swap operator. Therefore, we only need to explicitly evaluate $\widetilde{H}_{zz}(t) = 3\xi\left[U_1^\dagger(t)\otimes U_1^\dagger(t)\right]\left(\frac{\sigma_z}{2}\otimes\frac{\sigma_z}{2}\right)\left[U_1(t)\otimes U_1(t)\right]$ to calculate $\widetilde{H}_{\text{int}}(t)$:

$$\widetilde{H}_{zz}(t) = \begin{cases} 3\xi\frac{\sigma_z}{2}\otimes\frac{\sigma_z}{2}, \text{ for } 0\le t<\tau \\ \frac{3\xi}{2}\left[\left(\frac{\sigma_x}{2}\otimes\frac{\sigma_x}{2}+\frac{\sigma_y}{2}\otimes\frac{\sigma_y}{2}\right)+\cos\left[2\Omega(t-\tau)\right]\left(\frac{\sigma_y}{2}\otimes\frac{\sigma_y}{2}-\frac{\sigma_x}{2}\otimes\frac{\sigma_x}{2}\right)\right] \\ -\frac{3\xi}{2}\sin\left[2\Omega(t-\tau)\right]\left(\frac{\sigma_x}{2}\otimes\frac{\sigma_y}{2}+\frac{\sigma_y}{2}\otimes\frac{\sigma_x}{2}\right), \text{ for } \tau\le t<3\tau \\ \frac{3\xi}{2}\left[\left(\frac{\sigma_x}{2}\otimes\frac{\sigma_x}{2}+\frac{\sigma_y}{2}\otimes\frac{\sigma_y}{2}\right)+\cos\left[2\Omega(5\tau-t)\right]\left(\frac{\sigma_y}{2}\otimes\frac{\sigma_y}{2}-\frac{\sigma_x}{2}\otimes\frac{\sigma_x}{2}\right)\right] \\ -\frac{3\xi}{2}\sin\left[2\Omega(5\tau-t)\right]\left(\frac{\sigma_x}{2}\otimes\frac{\sigma_y}{2}+\frac{\sigma_y}{2}\otimes\frac{\sigma_x}{2}\right), \text{ for } 3\tau\le t<5\tau \\ 3\xi\frac{\sigma_z}{2}\otimes\frac{\sigma_z}{2}, \text{ for } 5\tau\le t\le 6\tau \end{cases} \tag{2.14}$$

Consequently,

$$\begin{aligned}\overline{H}_{\text{int}} &= \frac{1}{6\tau}\int_0^{6\tau} dt\ \widetilde{H}_{\text{int}}(t) = \frac{1}{6\tau}\int_0^{6\tau} dt\ \widetilde{H}_{zz}(t) - \xi\sum_{k\in\{x,y,z\}}\frac{\sigma_k}{2}\otimes\frac{\sigma_k}{2} \\ &= \frac{\xi}{4\tau\Omega}\sin\left(4\Omega\tau\right)\left(\frac{\sigma_y}{2}\otimes\frac{\sigma_y}{2}-\frac{\sigma_x}{2}\otimes\frac{\sigma_x}{2}\right)+\frac{\xi}{4\tau\Omega}\left[\cos\left(4\Omega\tau\right)-1\right]\left(\frac{\sigma_x}{2}\otimes\frac{\sigma_y}{2}+\frac{\sigma_y}{2}\otimes\frac{\sigma_x}{2}\right)\end{aligned} \tag{2.15}$$

which also becomes zero in the limit of $\Omega\tau\to\infty$. That implies that for an ideal magic echo sequence – with its two infinitesimally short $\pi/2$ rotations as well as a rotary echo Rabi frequency that substantially exceeds both $\Delta\omega$ and ξ – will remove the effect of these Hamiltonians in AHT zeroth order. In Chapter 4, we will study the effects of introducing finite length $\pi/2$ rotations in the magic echo sequence. Finally, we remark that while checking the AHT properties for a specific pulse sequence is usually not too involved, it is generally not obvious which pulse sequence would result the desired values for $\overline{H}_{\text{int}}, \overline{H}^{(1)}_{\text{int}}, \overline{H}^{(2)}_{\text{int}}, \ldots$. Much of the existing NMR pulse sequences have been derived based on physical and mathematical insights, nevertheless, such derivations become very difficult for non-standard control settings, or when experimental pulse distortions are taken into account, too. With the next chapter we provide a numerical control engineering method that enables one to search for an $H_1(t)$ that generates a $U_1(t)$, that yields desired values for various AHT integrals.

2.2 Stochastic Liouville Theory

With this section we give an overview of another perturbative paradigm called stochastic Liouville theory, that can be used for effective Hamiltonian engineering in the presence Hamiltonian terms that can be modelled as stochastic operators described through their moment functions. Stochastic Liouville Theory is appropriate for evaluating the effect of stochastic Hamiltonians, that are substantially smaller (in norm) than deterministic system Hamiltonians, for the evolution of a quantum system over short times.

We start with a Hamiltonian $H(t)$ governing the evolution of our quantum system determined by $\rho(t)$:

$$H(t) = H_{\text{det}}(t) + \sum_i \varepsilon_i(t) A_i, \tag{2.16}$$

where $H_{\text{det}}(t)$ is some, potentially time-dependent, deterministic Hamiltonian, $\{A_i\}$ is a set of constant Hermitian operators and $\{\varepsilon_i(t)\}$ is a set of stochastic functions characterised by their moments $\{\langle\varepsilon_i(t)\rangle\}$, $\{\langle\varepsilon_i(t)\varepsilon_j(t)\rangle\}$, ... We note that $\{A_i\}$ are not assumed to form an operator basis nor even be orthogonal to each other. $\rho(t)$ evolves according to Liouville equation $\dot{\rho}(t) = -i\,[H(t), \rho(t)]$, where $[,]$ stands the matrix commutator. Here, we denote a vectorized density matrix as $|\rho(t)\rangle$, that should be understood as a particular element-wise mapping of $d \times d$ complex matrices to d^2 dimensional complex vectors, which we will not explicitly state. We say that $\mathcal{A}_i$ is the Liouville operator that acts on the density matrix $\rho(t)$, such that $\mathcal{A}_i\,|\rho(t)\rangle = |[A_i, \rho(t)]\rangle$. Therefore, we can write the Liouville equation for the evolution of $\rho(t)$ as

$$\frac{d}{dt}\,|\rho(t)\rangle = -i\left[\mathcal{H}_{\text{det}}(t) + \sum_i \varepsilon_i(t)\mathcal{A}_i\right]|\rho(t)\rangle\,, \tag{2.17}$$

where $\mathcal{H}_{\text{det}}(t)$ is also defined implicitly through the relationship $\mathcal{H}_{\text{det}}(t)\,|\rho(t)\rangle = |[H_{\text{det}}(t), \rho(t)]\rangle$.

Just like earlier in this chapter, we move into the toggling frame of the deterministic Liouville operator $\mathcal{H}_{\text{det}}(t)$. The remaining stochastic Lioville equation for a particular realisation of noise processes $\{\varepsilon_i(t)\}$ is given as

$$\frac{d}{dt}\,|\widetilde{\rho}(t), \{\varepsilon_i(t)\}\rangle = -i\sum_i \varepsilon_i(t)\widetilde{\mathcal{A}}_i(t)\,|\widetilde{\rho}(t), \{\varepsilon_i(t)\}\rangle\,, \tag{2.18}$$

where $\widetilde{\mathcal{A}}_i(t) = \mathcal{U}^\dagger(t)\mathcal{A}_i\mathcal{U}(t)$ and $\mathcal{U}(t)$ is a solution to $\dot{\mathcal{U}}(t) = -\mathcal{H}_{\text{det}}(t)\mathcal{U}(t)$, with $\mathcal{U}(0) = \mathbb{1}$. The stochatic Liouville equation above has a formal solution given by Dyson series or

Magnus expansion

$$|\widetilde{\rho}(t),\{\varepsilon_i(t)\}\rangle = \left[\mathbb{1} - i\int_0^t dt_1\ E(t_1) + (-i)^2\int_0^t dt_1\int_0^{t_1} dt_2\ E(t_1)E(t_2) + ...\right]|\widetilde{\rho}(0)\rangle\,, \tag{2.19}$$

where $E(t)=\sum_i \varepsilon_i(t)\widetilde{\mathcal{A}}_i(t)$ and we have assumed the initial density matrix $\widetilde{\rho}(0)=\rho(0)$ to be independent of the stochastic processes $\{\varepsilon_i(t)\}$.

We are generally interested in the ensemble averaged behaviour of $|\widetilde{\rho}(t),\{\varepsilon_i(t)\}\rangle$, therefore, we would like to calculate $\langle|\widetilde{\rho}(0),\{\varepsilon_i(t)\}\rangle\rangle = S(t)\,|\widetilde{\rho}(0)\rangle$, where

$$S(t) = \mathbb{1} - i\int_0^t dt_1\langle E(t_1)\rangle + (-i)^2\int_0^t dt_1\int_0^{t_1} dt_2\langle E(t_1)E(t_2)\rangle + ...\ . \tag{2.20}$$

We can now rewrite $S(t)$ in terms of cumulant averages introduced by Kubo [58], and defined implicitly through the equation

$$S(t) = \exp\left[\sum_{m=1}^{\infty}(-i)^m\int_0^t dt_1\int_0^{t_1} dt_2\ ...\int_0^{t_{m-1}} dt_m\langle E(t_1)E(t_2)...E(t_m)\rangle_c\right], \tag{2.21}$$

whereas the explicit first few cumulant averages can be extracted from the following equations

$$\langle E(t_1)\rangle = \langle E(t_1)\rangle_c, \tag{2.22}$$

$$\langle E(t_1)E(t_2)\rangle = \langle E(t_1)E(t_2)\rangle_c + \frac{1}{2}\{\langle E(t_1)\rangle_c,\langle E(t_2)\rangle_c\}, \tag{2.23}$$

$$\begin{aligned}\langle E(t_1)E(t_2)E(t_3)\rangle &= \langle E(t_1)E(t_2)E(t_3)\rangle_c \\ &+ \frac{1}{2}\{\langle E(t_1)\rangle_c,\langle E(t_2)E(t_3)\rangle_c\} + \frac{1}{2}\{\langle E(t_2)\rangle_c,\langle E(t_1)E(t_3)\rangle_c\} \\ &+ \frac{1}{2}\{\langle E(t_3)\rangle_c,\langle E(t_1)E(t_2)\rangle_c\} + \frac{1}{6}\langle E(t_1)\rangle_c\{\langle E(t_2)\rangle_c,\langle E(t_3)\rangle_c\} \\ &+ \frac{1}{6}\langle E(t_2)\rangle_c\{\langle E(t_1)\rangle_c,\langle E(t_3)\rangle_c\} + \frac{1}{6}\langle E(t_3)\rangle_c\{\langle E(t_1)\rangle_c,\langle E(t_2)\rangle_c\},\end{aligned} \tag{2.24}$$

with $\{,\}$ being the anti-commutator of matrices.

From now on, we assume that $\{\epsilon_i(t)\}$ is a set of independent Gaussian random processes. That implies that all correlation functions $\langle\epsilon_{i_1}(t_1)\epsilon_{i_2}(t_2)...\epsilon_{i_n}(t_n)\rangle$ separate as long as $i_j\neq i_k$; it also implies that any correlation function $\langle\epsilon_i(t_1)\epsilon_i(t_2)...\epsilon_i(t_n)\rangle$ can always be written in terms of the first and second order correlation functions $\langle\epsilon_i(t_1)\rangle$ and $\langle\epsilon_i(t_1)\epsilon_i(t_2)\rangle$,

respectively. Furthermore, we also assume that the stochastic processes have zero mean, i.e., $\langle \epsilon_i(t_1) \rangle = 0$, for all i, and hence, all odd order correlation functions are zero, which implies that all odd order cumulant averages are zero, too. The motivation for using cumulants, instead of evaluating Dyson series in terms of moments of the stochastic process, arises from the fact that cumulant averages enter the exponential in Equation (2.21) and, hence, reduce the risk of secular terms as is argued below.

The cumulant averages are constructed in such a way that due to statistical independence the nth order cumulant disappears if for a set of times $\{t_1, t_2, ..., t_n\}$ any of the gaps $|t_1 - t_2|$, $|t_2 - t_3|$, ..., $|t_{n-1} - t_n|$ is larger than the correlation time τ_c of the stochastic process. This fact was proved by Kubo in [58]. Since cumulants at every order vanish once the gap between the set of time points exceeds τ_c, an nth order cumulant

$$\int_0^t dt_1 \int_0^{t_1} dt_2 \ ... \int_0^{t_{n-1}} dt_n \langle E(t_1)E(t_2)...E(t_n) \rangle_c, \tag{2.25}$$

is effectively an integral over $n-1$ dimensional sphere with radius τ_c and therefore grows linearly in time t provided $t \gg \tau_c$. This fact enables one to put bounds on errors induced by truncating the cumulant series at any even order. Here, we look at truncating the cumulant expansion at second order. Given the discussion above, it is clear that for $t \gg \tau_c$ the norm of $\int_0^t dt_1 \int_0^{t_1} dt_2 \ ... \int_0^{t_{n-1}} dt_n \langle E(t_1)E(t_2)...E(t_n) \rangle_c$ scales as $\tau_c^{n-1} e^n t$, where e is the norm of $E(t)$. Comparing the norms of second and forth order cumulants yields $\frac{\tau_c^3 e^4 t}{\tau_c e^2 t} = \tau_c^2 e^2$, meaning that if $\tau_c e \ll 1$ and $t \gg \tau_c$ we can safely truncate the cumulant expansion at second order.

Lastly, we point out that nth order cumulants are functions of statistical averages up to nth order, while the latter can be written explicitly as

$$\begin{aligned} &\int_0^t dt_1 \int_0^{t_1} dt_2 \ ... \int_0^{t_{n-1}} dt_n \langle E(t_1)E(t_2)...E(t_n) \rangle = \\ &\sum_{i_1, i_2, ..., i_n} \int_0^t dt_1 \int_0^{t_1} dt_2 \ ... \int_0^{t_{n-1}} dt_n \langle \varepsilon_{i_1}(t_1)\varepsilon_{i_2}(t_2)...\varepsilon_{i_n}(t_n) \rangle \widetilde{\mathcal{A}}_{i_1}(t_1)\widetilde{\mathcal{A}}_{i_2}(t_2)...\widetilde{\mathcal{A}}_{i_n}(t_n), \end{aligned} \tag{2.26}$$

having used the definition of $E(t)$ given above, where $\langle \varepsilon_{i_1}(t_1)\varepsilon_{i_2}(t_2)...\varepsilon_{i_n}(t_n) \rangle$ is a scalar function determined by the correlation functions of the stochastic process. This makes it clear that cumulants are functions of nested integrals of Equation (2.1) kind, with $f \neq 1$ unlike the terms appearing in AHT.

2.3 Conclusions

With this chapter we have reviewed two analytical treatments of effective Hamiltonian Engineering. Importantly, we have shown that both of these schemes work with perturbation expressions which take a form of Equation (2.1). Nevertheless, it is also clear that, while AHT and stochastic Liouville theory provide a list of nested integral terms that have to be set to zero or to some specific value, they do not necessarily prescribe a way for doing it. This motivates the development of control engineering tools that would enable using numerical optimization for setting any number of expressions in Equation (2.1) to certain desired values. In the next chapter, we will do just that.

Chapter 3

Numerical Engineering of Effective Hamiltonians

With the past two chapters we have argued that in order to successfully engineer some desired effective evolution for a quantum system, we need to ensure that the system evolution operator $U(T)$ and certain time dependent perturbation theory expressions of the following form:

$$\int_0^T dt_1 \int_0^{t_1} dt_2 ... \int_0^{t_{n-1}} dt_n \; f(t_1, t_2, ..., t_n) \; [U'(t_1)]^{-1} A(t_1) U'(t_1) \; ... \; [U'(t_n)]^{-1} A(t_n) U'(t_n), \tag{3.1}$$

take some specific values. Here, $f(t_1, t_2, ..., t_n)$ is a scalar function and $\{A(t_i)\}$ is a set of, potentially time-dependent, operators. $U'(t)$ is either equal to $U(t)$ or is a function of it. In Chapter 1, we also highlighted the numerous benefits of numerical control engineering tools. Specifically, their ability to handle arbitrary system Hamiltonians of arbitrary Hilbert space dimension, find control sequences that adhere to experimental constraints and easily extend to ensembles of quantum systems that are controlled with the same control sequence.

Although there exist problem specific numerical control engineering approaches that have been used [88, 33] to optimize the value of $U(T)$ along with some perturbation terms in the literature, a complete framework for such optimization has been lacking. In this chapter, we develop a general framework which enables efficient and easy to implement numerical evaluation of $U(T)$, simultaneously with the evaluation or arbitrarily close approximation of any number of nested integrals in Equation (3.1), for piecewise constant control sequences. We do this by setting up and solving a first order matrix differential equation, which we call the Van Loan equation. The Van Loan equation incorporates the

same control amplitudes that appear in the Schrondinger equation for $U(t)$ and is constructed by replacing the Hamiltonian matrices in the Schrödinger equation with certain larger block matrices. We accomplish this by generalizing the work of Van Loan [116], Carbonell *et al* [14] and, more recently, Goodwin and Kuprov[31], who noticed that various nested operator integrals can be evaluated by exponentiating certain block matrices.

Our generalizations of the existing work are two-fold: we first introduce a way of calculating the perturbation terms in Equation (3.1) of time dependent operators under time dependent control amplitudes, we also develop tools for approximating nested integrals involving arbitrary scalar functions $f(t_1, t_2, ..., t_n)$. The Van Loan matrix differential equation formulation has two immediate benefits: for piecewise constant control amplitudes the integrals can be evaluated accurately and efficiently by matrix exponentiation. Furthermore, such formulation immediately enables the use of numerical control search methods that use gradient information. Most of this chapter is devoted to demonstrating how to exploit the differential equation formulation for the purpose of numerical control searches that involve various perturbation expressions. The numerical tools developed in this chapter will be deployed in the next chapter for engineering effective nuclear spin Hamiltonians for nanoscale magnetic resonance imaging experiments that were published in [96].

In this chapter, we first introduce our notation and terminology in Section 3.1 and give some computational background for matrix differential equations in Section 3.2. In Section 3.2, we also present a solution for a general time dependent upper triangular block matrix differential equation and highlight how it can be used for calculating these nested integrals in Equation (3.1), for which $f = 1$. After that, in Section 3.3, we introduce the GRAPE algorithm [57] for numerical control searches, that was used for finding all examples within this thesis. Subsequently, we exemplify the construction of Van Loan block matrix differential equations and numerical control optimizations that include perturbation integrals with four simple, yet non-trivial, examples in Section 3.4. For some of these simple examples the searches converge to well-known analytical control sequences. We then outline our approach for tackling general experimental control problems in Section 3.5, followed by a description of our matrix methods for implementing transfer functions in Section 3.6. In Section 3.7, we give a prescription for how to approximate perturbation terms with $f \neq 1$ in Equation (3.1). Finally, in Section 3.8, we provide an experimentally realistic example of numerical control engineering that deals with stochastic 1/f noise using control waveforms that are subject to bandwidth limitations.

The methods to be introduced here were developed by the author in collaboration with Daniel Puzzuoli, contributions that were not the author's have been clearly indicated in the footnotes.

3.1 Notation and Terminology

In this section, we describe the notation that we will use throughout this thesis for denoting various matrices, block matrices and nested integral expressions. We also introduce various definitions, including the definitions for the matrix norms and fidelity functions that we will use within this and the next chapter, typically as metrics characterizing the performance of our control sequences.

We use $\mathrm{M}_{i,j}(S)$ to denote a set of matrices of dimension $i \times j$ that have their entries drawn from some set S. Since most of the matrices we deal with are square matrices, we will use $\mathrm{M}_i(S)$ as a shorthand for $\mathrm{M}_{i,i}(S)$. We can write each element $A \in \mathrm{M}_{n,m}(S)$ in its component form as $A = (S_{i,j})$, where $1 \leq i \leq n$ and $1 \leq j \leq m$. For example, the set S could be the set of complex numbers, in which case $\mathrm{M}_{i,j}(\mathbb{C})$ stands for the set of $i \times j$ complex matrices. We introduce another shorthand, denoting $\mathrm{M}_{i,j}(\mathbb{C})$ simply as $\mathrm{M}_{i,j}$. Nevertheless, the set S is not restricted to numbers and given our notation $\mathrm{M}_k(\mathrm{M}_n(\mathbb{C})) = \mathrm{M}_k(\mathrm{M}_n)$ would denote the set of $k \times k$ block matrices composed from $n \times n$ complex matrices, hence, an element $A \in \mathrm{M}_k(\mathrm{M}_n)$ is also an element of $\mathrm{M}_{k \times n}$. For any set S, we denote n-dimensional vectors with elements drawn from that set as S^n, although, we generally only consider either real or complex vectors, $\mathbb{R}^n$ or $\mathbb{C}^n$, respectively. Finally, we take $\mathbb{1}_n$ to stand for the n-dimensional identity operator, i.e., $\mathbb{1}_n = \mathrm{diag}\underbrace{(1, 1, \ldots, 1)}_{n \text{ elements}}$.

For denoting functions that map an element from some space $\mathcal{A}$ to an element of another space $\mathcal{B}$, we frequently use a notation $\mathcal{A} \to \mathcal{B}$. Specifically, since most of the functions used throughout this thesis are matrix valued functions of a single real valued (time) parameter t, where $0 \leq t \leq T$, we denote such functions as $A(t)$, $A : [0, T] \to \mathrm{M}_n$. We will often use the trace function $\mathrm{Tr} : \mathrm{M}_n \to \mathbb{C}$, where $\mathrm{Tr}[A] = \sum_i^n A_{i,i}$ for any $A \in \mathrm{M}_n$. We also make frequent use of the transpose function $^{\mathrm{T}} : \mathrm{M}_n \to \mathrm{M}_n$ and the conjugate transpose function $^{\dagger} : \mathrm{M}_n \to \mathrm{M}_n$, which combines the transpose and entry-wise complex conjugation. An inverse of matrix $A \in \mathrm{M}_n$ will be denoted as A^{-1}. Finally, we say that the inner product between a pair of matrices $A, B \in \mathrm{M}_n$ is given as $\mathrm{Tr}[A^\dagger B]$, and the matrices A and B are called *orthogonal* whenever their inner product is zero.

Since much of this and the next chapter deals with nested integrals of various time dependent matrix valued functions $A_1, \ldots, A_n : [0, T] \to \mathrm{M}_m$ that have been transformed by some evolution operator $U(t)$, $U : [0, T] \to \mathrm{M}_m$, we denote

$$\mathcal{D}_U(A_1, \ldots, A_n) = \tag{3.2}$$
$$U(T) \int_0^T dt_1 \cdots \int_0^{t_{n-1}} dt_n U^{-1}(t_1) A_1(t_1) U(t_1) \ldots U^{-1}(t_n) A_n(t_n) U(t_n),$$

where $U^{-1}(t)A_i(t)U(t) \in \mathrm{M}_m$, for all i, represents a particular matrix valued function that has been transformed by U. $\mathcal{D}$ could be read as *Dyson term*, for which we give some context in the next section.

Here, we define the only matrix norm and the matrix fidelity function to be used in this thesis. We take $\|A\|$ to stand for the Hilbert-Schmidt norm [120] for a matrix $A \in \mathrm{M}_n$; the norm is defined as

$$\|A\| = \sqrt{\mathrm{Tr}\,(A^\dagger A)}, \tag{3.3}$$

and $\|A\| = \sqrt{\sum_{i,j}^{n} |A_{i,j}|^2}$. Note that $\|A\| \geq 0$, for all $A \in \mathrm{M}_n$, and $\|A\| = 0$ if and only if $A = 0$; it is this property that we will generally exploit throughout this and the next chapter. Furthermore, if $U(T)$ is a unitary matrix, i.e $UU^\dagger = \mathbb{1}_m$, then $\|\mathcal{D}_U(A_1, \ldots, A_n)\| = \|U^{-1}(T)\mathcal{D}_U(A_1, \ldots, A_n)\|$, the latter being the quantity generally appearing in various time dependent perturbation theory expressions. We also define a fidelity function $\mathcal{F}(U, V)$ for a pair of matrices $U, V \in \mathrm{M}_n$

$$\mathcal{F}(U, V) = \sqrt{\frac{\mathrm{Tr}\,(U^\dagger V)\,\mathrm{Tr}\,(V^\dagger U)}{\mathrm{Tr}\,(U^\dagger U)\,\mathrm{Tr}\,(V^\dagger V)}}. \tag{3.4}$$

The fidelity function is useful for characterizing the closeness of two unitary matrices since all eigenvalues of unitary matrices have a modulus of one, i.e., they lie on the complex unit circle. It can be seen that for two unitary matrices $U, V \in \mathrm{M}_n$, it always holds that $0 \leq \mathcal{F}(U, V) \leq 1$, whereas $\mathcal{F}(U, V) = 1$ if and only if $U = V$, and $\mathcal{F}(U, V) = 0$ if the matrices are orthogonal.

3.2 Matrix Differential Equations and Computational Methods

We start this section by introducing our treatment of matrix differential equations and the toggling frame. We then show that the nested integral expressions appearing in AHT, that we discussed in the previous chapter, can be thought of as various directional derivatives of the evolution operator. We finish the section by introducing our block matrix methods for calculating a certain class of such nested integral, while keeping our most general treatment of various perturbation integrals for Section 3.7.

Given a matrix valued function $G(t)$, $G : [0, T] \to \mathrm{M}_n$, we consider initial value problems of the following form:

$$\dot{U}(t) = G(t)U(t), \tag{3.5}$$

where $U : [0,T] \to \mathrm{M}_n$, $U(0) = \mathbb{1}_n$, and where $\dot{U}$ denotes the time derivative of U. We call $G(t)$ the *generator* and $U(t)$ the *propagator* of the initial value problem. For quantum control problems we can often, but not always, relate $G(t)$ to the time dependent Hamiltonian $H(t)$ for the quantum system by setting $G(t) = -iH(t)$. In such cases $U(t)$ would be a unitary matrix for all $0 \leq t \leq T$. Equation (3.5) has a unique solution [43], which we write using the *time-ordered exponential* notation:

$$U(t) = \mathcal{T} \exp\left(\int_0^t dt_1 \; G(t_1)\right). \tag{3.6}$$

Here, the time-ordered exponential should be understood as the solution to the matrix differential equation above. By definition, the time-ordered exponential has the property $\frac{d}{dt}\mathcal{T} \exp\left(\int_0^t dt_1 \; G(t_1)\right) = G(t)\mathcal{T} \exp\left(\int_0^t dt_1 \; G(t_1)\right)$; it is this property that we will use extensively throughout this chapter, importantly, not just in the case of Hamiltonian generators.

It is possible to express the solution of Equation (3.5) as a series of nested integrals called the *Dyson series* [22]

$$U(t) = \mathbb{1}_n + \int_0^t dt_1 \; G(t_1) + \int_0^t dt_1 \int_0^{t_1} dt_2 \; G(t_1)G(t_2) + \dots \; . \tag{3.7}$$

Under certain conditions [70, 5] we can also express the propagator through the *Magnus expansion* giving $U(t) = \exp\left[\Omega(t)\right]$ for

$$\Omega(t) = \int_0^t dt_1 G(t_1) + \int_0^t dt_1 \int_0^{t_1} dt_2 [G(t_1), G(t_2)] + \dots \; . \tag{3.8}$$

3.2.1 The Toggling Frame and Variations of the Generator

Given two matrix valued functions $G(t)$ and $G_{\mathrm{v}}(t)$, such that $G, G_{\mathrm{v}} : [0,T] \to M_n$, we consider the evolution of a system under the combined generator $G(t) + G_{\mathrm{v}}(t)$, where we regard $G_{\mathrm{v}}(t)$ as a *variation* of the generator $G(t)$. Even though our treatment here is abstract and independent of any particular system, we remark that in the case of quantum computing we typically think of $G_{\mathrm{v}}(t)$ as the unwanted or the uncharacterized parts of the system Hamiltonian, while in the case of sensing and spectroscopy $G_{\mathrm{v}}(t)$ might be regarded as some unknown part of the Hamiltonian that we intend to characterize.

We now introduce the concept of a *toggling frame*, which enables us to partially separate the evolution under $G(t)$ and $G_{\mathrm{v}}(t)$, hence, illuminating the effect of the variation generator.

We say that $U(t)$ is the propagator under $G(t)$ alone:

$$U(t) = \mathcal{T} \exp\left(\int_0^t dt_1 G(t_1)\right), \tag{3.9}$$

whereas the propagator under $G(t) + G_{\mathrm{v}}(t)$ is

$$U_{\mathrm{total}}(t) = \mathcal{T} \exp\left(\int_0^t dt_1 [G(t_1) + G_{\mathrm{v}}(t_1)]\right). \tag{3.10}$$

In addition, we define the *toggling frame propagator*

$$U_{\mathrm{tog}}(t) = \mathcal{T} \exp\left(\int_0^t dt_1 \widetilde{G}_{\mathrm{v}}(t_1)\right), \tag{3.11}$$

where $\widetilde{G}_{\mathrm{v}}(t) = U^{-1}(t) G_{\mathrm{v}}(t) U(t)$. We now show that

$$U_{\mathrm{total}}(t) = U(t) U_{\mathrm{tog}}(t) \tag{3.12}$$

by first noting that at time $t = 0$ the equality holds because $U_{\mathrm{total}}(0) = U(0) U_{\mathrm{tog}}(0) = \mathbb{1}_n$. If the time derivatives of $U_{\mathrm{total}}(t)$ and $U(t) U_{\mathrm{tog}}(t)$ are also equal for all times, the equality in Equation (3.12) is true. Differentiating the right-hand side of the equation and employing the property of the time-ordered exponential that we highlighted above yields

$$\begin{aligned} \frac{d}{dt} U(t) U_{\mathrm{tog}}(t) &= \dot{U}(t) U_{\mathrm{tog}}(t) + U(t) \dot{U}_{\mathrm{tog}}(t) = G(t) U(t) U_{\mathrm{tog}}(t) + U(t) \widetilde{G}_{\mathrm{v}}(t) U_{\mathrm{tog}}(t) \\ &= [G(t) + G_{\mathrm{v}}(t)] U(t) U_{\mathrm{tog}}(t), \end{aligned} \tag{3.13}$$

demonstrating that the time derivative of $U(t) U_{\mathrm{tog}}(t)$ is equal to that of $U_{\mathrm{total}}(t)$.

It can now be seen that, with the decomposition $U_{\mathrm{total}}(t) = U(t) U_{\mathrm{tog}}(t)$, we have encapsulated the effect of the variation on $U_{\mathrm{total}}(t)$ into $U_{\mathrm{tog}}(t)$ only. This enables us to conveniently analyse various derivatives with respect to the variation. As an example, we consider the directional derivative of $U(t)$ arising from a variation in $G(t)$ in the direction $G_{\mathrm{v}}(t)$:

$$\frac{d}{d\epsilon}\Big|_{\epsilon=0} \mathcal{T} \exp\left(\int_0^t dt_1 [G(t_1) + \epsilon G_{\mathrm{v}}(t_1)]\right). \tag{3.14}$$

Expanding $U_{\mathrm{tog}}(t)$ into the Dyson series we arrive at a power series for $U_{\mathrm{total}}(t)$ in ϵ

$$U_{\mathrm{total}}(t) = U(t) + \epsilon U(t) \int_0^t dt_1 \widetilde{G}_{\mathrm{v}}(t_1) + \epsilon^2 U(t) \int_0^t dt_1 \int_0^{t_1} dt_2 \widetilde{G}_{\mathrm{v}}(t_1) \widetilde{G}_{\mathrm{v}}(t_2) + \ldots, \tag{3.15}$$

which enables reading off the directional derivative in Equation (3.14) simply as the multiplier of the ϵ term

$$\frac{d}{d\epsilon}\Big|_{\epsilon=0}\mathcal{T}\exp\left(\int_0^t dt_1[G(t_1)+\epsilon G_{\rm v}(t_1)]\right)=U(t)\int_0^t dt_1\widetilde{G}_{\rm v}(t_1)$$
$$=U(t)\int_0^t dt_1 U^{-1}(t_1)G_{\rm v}(t_1)U(t_1). \tag{3.16}$$

Of course, the second directional derivative then reads as

$$\frac{d^2}{d\epsilon^2}\Big|_{\epsilon=0}\mathcal{T}\exp\left(\int_0^t dt_1[G(t_1)+\epsilon G_{\rm v}(t_1)]\right)=2U(t)\int_0^t dt_1\int_0^{t_1} dt_2\widetilde{G}_{\rm v}(t_1)\widetilde{G}_{\rm v}(t_2)$$
$$=2U(t)\int_0^t dt_1\int_0^{t_1} dt_2 U^{-1}(t_1)G_{\rm v}(t_1)U(t_1)U^{-1}(t_2)G_{\rm v}(t_2)U(t_2). \tag{3.17}$$

We can also apply the same exact analysis for multiple variations, for that we consider

$$\mathcal{T}\exp\left(\int_0^t dt_1[G(t_1)+\epsilon_1 G_{\rm v1}(t_1)+\epsilon_2 G_{\rm v2}(t_1)]\right)=U(t)+U(t)\int_0^t dt_1(\epsilon_1\widetilde{G}_{\rm v1}(t_1)+\epsilon_2\widetilde{G}_{\rm v2}(t_1))$$
$$+U(t)\int_0^t dt_1\int_0^{t_1} dt_2(\epsilon_1\widetilde{G}_{\rm v1}(t_1)+\epsilon_2\widetilde{G}_{\rm v2}(t_1))(\epsilon_1\widetilde{G}_{\rm v1}(t_2)+\epsilon_2\widetilde{G}_{\rm v2}(t_2))+\ldots, \tag{3.18}$$

which enables us to identify

$$\frac{d}{d\epsilon_1}\Big|_{\epsilon_1=0}\mathcal{T}\exp\left(\int_0^t dt_1[G(t_1)+\epsilon_1 G_{\rm v1}(t_1)+\epsilon_2 G_{\rm v2}(t_1)]\right)=U(t)\int_0^t dt_1\,\widetilde{G}_{\rm v1}(t_1) \tag{3.19}$$

and

$$\frac{d}{d\epsilon_2}\Big|_{\epsilon_2=0}\mathcal{T}\exp\left(\int_0^t dt_1[G(t_1)+\epsilon_1 G_{\rm v1}(t_1)+\epsilon_2 G_{\rm v2}(t_1)]\right)=U(t)\int_0^t dt_1\,\widetilde{G}_{\rm v2}(t_1), \tag{3.20}$$

as well as

$$\frac{d}{d\epsilon_1}\Big|_{\epsilon_1=0}\frac{d}{d\epsilon_2}\Big|_{\epsilon_2=0}\mathcal{T}\exp\left(\int_0^t dt_1[G(t_1)+\epsilon_1 G_{\rm v1}(t_1)+\epsilon_2 G_{\rm v2}(t_1)]\right) \tag{3.21}$$
$$=U(t)\int_0^t dt_1\int_0^{t_1} dt_2\,\widetilde{G}_{\rm v1}(t_1)\widetilde{G}_{\rm v2}(t_2)+U(t)\int_0^t dt_1\int_0^{t_1} dt_2\,\widetilde{G}_{\rm v2}(t_1)\widetilde{G}_{\rm v1}(t_2).$$

As we said at the beginning of this chapter, engineering of quantum control typically aims to implement a particular propagator $U(T)$ on the system or an ensemble of

systems, which would correspond as closely as possible to the experimenters best characterization of the system generator $G(t)$. Nevertheless, because of additional system generators, couplings to the environment or variations in $G(t)$ – which can all be thought of as various $\{G_{\mathrm{v}}(t)\}$ – the final system propagator will be perturbed from the value $U(T) = \mathcal{T}\exp\left(\int_0^t dt_1 G(t_1)\right)$. In the context of quantum computing, the experimenter typically wants to make the control robust against such perturbations, which translates to ensuring that various directional derivatives with respect to the variations present are set to zero or near zero. In the case of sensing and spectroscopy the usual goal is to have the perturbation from a certain $G_{\mathrm{v}}(t)$ generate a very well characterized effect on the system propagator, while potentially ensuring that the perturbations from other $\{G_{\mathrm{w}}(t)\}$ have no observable effect. Both of these demands can be fulfilled by ensuring that the directional derivatives that arise from the variations present for the system take some specific values.

Furthermore, in the context of control engineering the directional derivatives can also be seen as a way of calculating the change in $U(T)$ corresponding to a variation in the system generator, this information could be exploited when searching for a $G(t)$, $G : [0, T] \to \mathrm{M}_n$, that would yield the desired propagator $U(T)$. In particular, we will show in the next section that for piecewise constant control sequences one is typically interested in directional derivatives of $U(T)$ for constant $G(t)$ and $G_{\mathrm{v}}(t)$. This can be done by setting $t = 1$, $G(t) = A$ and $G_{\mathrm{v}}(t) = B$ in Equation (3.16), which gives

$$\frac{d}{d\epsilon}|_{\epsilon=0} e^{A+\epsilon B} = e^{A}\int_0^1 d\alpha\; e^{-\alpha A} B e^{\alpha A}. \tag{3.22}$$

3.2.2 Computing Directional Derivatives of Propagators Using Block Matrices[1]

We now introduce our linear algebra tool set for numerically evaluating integral expressions of the following form: $U(t)\int_0^t dt_1 U^{-1}(t_1) G_{\mathrm{v}}(t_1) U(t_1)$. First, we note that naively one might expect that the evaluation of such integrals requires integral approximation methods, which can become computationally very expensive, since it would necessitate many evaluations of $U(t)$ for various $t \in [0, T]$. In order to familiarize the reader with our block matrix methods we first consider the case of constant generators, i.e., $G(t) = A$ and $G_{\mathrm{v}}(t) = B$, $A, B \in \mathrm{M}_n$, such that the integral above becomes $e^{At}\int_0^t dt_1 e^{-At_1} B e^{At_1}$. A convenient method for calculating that integral was pointed out by Van Loan [116] and it stems from

[1]The theorem and its proof appearing in this subsection was shown to the author by Daniel Puzzuoli.

the following block matrix equation:

$$\exp\left[\begin{pmatrix} A & B \\ 0 & A \end{pmatrix} t\right] = \begin{pmatrix} e^{At} & e^{At}\int_0^t dt_1 e^{-At_1} B e^{At_1} \\ 0 & e^{At} \end{pmatrix}. \tag{3.23}$$

From this identity, it is evident that the integral $e^{At}\int_0^t dt_1 e^{-At_1} B e^{At_1}$ can be evaluated by performing a single matrix exponential operation for a matrix in $\mathrm{M}_2(\mathrm{M}_n)$. This observation is incredibly useful for accurate numerical evaluation of such integrals. The equality above can be easily verified by noticing that both sides of the equation are equal to $\mathbb{1}_{2\times n}$ for $t = 0$ and the time derivatives on the left and right-hand side of the equation are equal. In the same vein, we can evaluate the integral $e^{At}\int_0^t dt_1 \int_0^{dt_1} dt_2 e^{-At_1} B e^{At_1} e^{-At_2} C e^{At_2}$ by noticing that the following is true:

$$\exp\left[\begin{pmatrix} A & B & 0 \\ 0 & A & C \\ 0 & 0 & A \end{pmatrix} t\right] = \tag{3.24}$$
$$\begin{pmatrix} e^{At} & e^{At}\int_0^t dt_1 e^{-At_1} B e^{At_1} & e^{At}\int_0^t dt_1 \int_0^{dt_1} dt_2 e^{-At_1} B e^{At_1} e^{-At_2} C e^{At_2} \\ 0 & e^{At} & e^{At}\int_0^t dt_1 e^{-At_1} C e^{At_1} \\ 0 & 0 & e^{At} \end{pmatrix}.$$

This equality can be verified the same way as the one above. Again, the evaluation of the nested integral expression is converted to a problem of exponentiating a single matrix in $\mathrm{M}_3(\mathrm{M}_n)$.

The block matrix expressions above were first pointed out by Van Loan [116], who deduced the expressions for exponentials of arbitrary upper triangular block matrices in $\mathrm{M}_4(\mathrm{M}_n)$. Such expressions were then generalized by Carbonell *et al* [14] for exponentials of upper triangular block matrices in $\mathrm{M}_m(\mathrm{M}_n)$, where m is arbitrary. These block matrix methods have already found use in quantum theory and control applications [83, 31], where there exist a demand for efficient and accurate numerical evaluation schemes of various nested integrals. Specifically, these block matrix methods have been employed for calculating first and second order derivatives of the unitary propagator $U(T)$ with respect to piecewise-constant control amplitude variations in gradient ascent pulse engineering [32].

Such block matrix methods are especially attractive for numerical evaluation of various nested integrals since they are very easy to implement and they exploit numerical exponentiation algorithms, which are heavily optimized [60]. Furthermore, it will be demonstrated in the upcoming sections that the block matrix differential equation representation works extremely well for numerical optimization algorithms that exploit gradient information.

This lays the foundation for numerical control engineering techniques that enable finding $U(T)$, which has its directional derivatives with respect to certain variations set to desired values.

With this work, we first extend the existing methods by introducing a time dependence for the operators generating the propagator $U(t)$ and the operators appearing in the nested integral expressions. As the first example, we point out that Equation (3.23) holds also in the time-dependent case:

$$\mathcal{T}\exp\left[\begin{pmatrix} A(t) & B(t) \\ 0 & A(t) \end{pmatrix} t\right] = \begin{pmatrix} U(t) & U(t)\int_0^t dt_1 U^{-1}(t_1)B(t_1)U(t_1) \\ 0 & U(t) \end{pmatrix}, \tag{3.25}$$

where $U(t) = \mathcal{T}\exp\left(\int_0^t dt_1 A(t_1)\right)$. The equation above is easily verified by first noticing that at $t = 0$ both sides of the equation are equal to $\mathbb{1}_{2n}$ and then confirming that the time derivatives for the block matrices on either side of the equation are equal.

We now give a full time-ordered generalization for the theorems of Van Loan [116] and Carbonell *et al* [14] as an expression for the time-ordered exponential of an arbitrary time-dependent upper triangular block matrix through a set of recursive relations. The theorem is proved in Appendix A, and will be employed in our first numerical control searches that include AHT terms in Section 3.4, as well as in the next chapter.

Theorem 1. Let $B_{i,j} : [0,T] \to M_n$ for $1 \leq i \leq j \leq m$. And let $C_{i,j} : [0,T] \to M_n$ be defined implicitly by the equation

$$\begin{pmatrix} C_{1,1}(t) & C_{1,2}(t) & \dots & C_{1,m}(t) \\ 0 & C_{2,2}(t) & \dots & C_{2,m}(t) \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & C_{m,m}(t) \end{pmatrix} = \tag{3.26}$$
$$\mathcal{T}\exp\left[\int_0^t dt_1 \begin{pmatrix} B_{1,1}(t_1) & B_{1,2}(t_1) & \dots & B_{1,m}(t_1) \\ 0 & B_{2,2}(t_1) & \dots & B_{2,m}(t_1) \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & B_{m,m}(t_1) \end{pmatrix}\right].$$

For all $1 \leq k \leq m$, $t \in [0,T]$, it holds that

$$C_{k,k}(t) = U_k(t) = \mathcal{T}\exp\left(\int_0^t dt_1 B_{i,i}(t_1)\right), \tag{3.27}$$

while all $C_{k,k+j}(t)$, for $1 < j \leq m-k$, can be found recursively as

$$C_{k,k+j}(t) = \sum_{i=1}^{j} U_k(t) \int_0^t dt_1 U_k^{-1}(t_1) B_{k,k+i}(t_1) C_{k+i,k+j}(t_1). \tag{3.28}$$

In order to evaluate all $\{C_{k,k+j}(t)\}$, one has to apply the recursion relation above sequentially for $j = \{0, 1, \ldots, m-k\}$.

3.3 GRAPE Algorithm

In this section, we describe the Gradient Ascent Pulse Engineering (GRAPE) algorithm that was introduced by Khaneja *et al* [57]. [57] points out a way for numerically solving bilinear control theory problems for piecewise constant control amplitudes. Even though Khaneja *et al* only considered problems of unitary engineering and state to state transfer, the same approach is equally applicable for any bilinear control theory problem including every example presented in this thesis.

We say that a control problem is *bilinear* if the control sequence $a(t)$ used for controlling the system over a time interval $[0, T]$, such that $a : [0, T] \to \mathbb{R}^k$, determines the system evolution through a first order matrix differential equation

$$\dot{U}(t) = \left[G_0 + \sum_{i=1}^{k} a_i(t) G_i \right] U(t), \tag{3.29}$$

where $G_i \in \mathrm{M}_n$, for all $i \in \{0, 1, \ldots, k\}$, and $U(0) = \mathbb{1}_n$. It is clear that the Schrödinger equation is a special instance of bilinear control with $\{G_i\}$ being anti-Hermitian matrices, i.e., $G_i = -G_i^\dagger$, for all i. We remark here that Van Loan equations, that will be introduced in the upcoming sections, also have the bilinear form of Equation (3.29), yet their generators $\{G_i\}$ are not anti-Hermitian.

Throughout this thesis, we will only deal with piecewise constant control amplitudes $a(t)$, for which we split the interval $[0, T]$ into N subintervals with respective durations ΔT_j such that $\Delta T_j \geq 0$, for all $j \in \{1, 2, \ldots, N\}$, and $\sum_{j=1}^{N} \Delta T_j = T$. We say that the control sequence $a(t)$ takes a constant value over each of the subintervals, i.e.,

$$a_i \left(\sum_{s=1}^{j-1} \Delta T_s \leq t < \sum_{s=1}^{j-1} \Delta T_s + \Delta T_j \right) = \alpha_{i,j}, \tag{3.30}$$

where $\alpha \in \mathrm{M}_{k,N}(\mathbb{R})$ is a real valued matrix that contains all piecewise control elements $\{\alpha_{i,j}\}$. In such case

$$U(T) = \prod_{j=1}^{N} \exp\left[\left(G_0 + \sum_{i=1}^{k} \alpha_{i,j} G_i\right) \Delta T_j\right], \tag{3.31}$$

where the product symbol denotes a sequential matrix multiplication.

A bilinear control theory problem for piecewise constant control amplitudes is stated as: find an $a(t)$, $a : [0, T] \to \mathbb{R}^k$, or equivalently an α, that adheres to certain problem specific constraints and yields a $U(T)$ with some desired properties. The constraints on $a(t)$ can vary from problem to problem, a common constraint sets a maximum amplitude for $a(t)$. Given the desired properties for $U(T)$, we can always write down a target function $\Phi : \mathrm{M}_n \to [0, 1]$ that is an analytic function, and takes the value 1 if and only if its argument has the properties that we want from $U(T)$. Having defined such a target function, the problem of finding an α that yields $\Phi[U(T)] = \Phi(\alpha) = 1$ becomes a multivariable optimization problem. Because Φ is an analytic function of $U(T)$, knowing $U(T)$ and its partial derivatives $\{\frac{\partial}{\partial \alpha_{i,j}} U(T)\}$, $i \in \{1, \ldots, k\}$ and $j \in \{1, \ldots, N\}$, is enough for determining both $\Phi[U(T)]$ and $\{\frac{\partial}{\partial \alpha_{i,j}} \Phi[U(T)]\}$. The key insight by Khaneja *et al* [57] was to point out that the computational cost for the simultaneous evaluation of $U(T)$ and $\{\frac{\partial}{\partial \alpha_{i,j}} U(T)\}$ is not too much more expensive than that for the evaluation of $U(T)$ only. We will demonstrate this fact explicitly below. First, however, we note that such relative computational cost of $U(T)$ and $\{\frac{\partial}{\partial \alpha_{i,j}} U(T)\}$ evaluation makes it advantageous to use numerical multivariable optimization algorithms that make use of the gradient information, e.g., conjugate gradient algorithm [86], rather than just $\Phi(\alpha)$.

Here, we proceed by evaluating a partial derivative

$$\frac{\partial}{\partial \alpha_{r,s}} U(T) = \tag{3.32}$$
$$\left(\prod_{j=s+1}^{N} \exp\left[\left(G_0 + \sum_{i=1}^{k} \alpha_{i,j} G_i\right) \Delta T_j\right]\right) \Upsilon_{rs} \prod_{j=1}^{s-1} \exp\left[\left(G_0 + \sum_{i=1}^{k} \alpha_{i,j} G_i\right) \Delta T_j\right],$$

with

$$\Upsilon_{rs} = \frac{\partial}{\partial \alpha_{r,s}} \exp\left[\left(G_0 + \sum_{i=1}^{k} \alpha_{i,s} G_i\right) \Delta T_s\right] \tag{3.33}$$
$$= \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \exp\left[\left(G_0 + \sum_{i=1}^{k} \alpha_{i,s} G_i\right) \Delta T_s + \epsilon G_r \Delta T_s\right] = \exp\left[\left(G_0 + \sum_{i=1}^{k} \alpha_{i,s} G_i\right) \Delta T_s\right]$$
$$\cdot \int_0^{\Delta T_s} dt \ \exp\left[-\left(G_0 + \sum_{i=1}^{k} \alpha_{i,s} G_i\right) t\right] G_r \exp\left[\left(G_0 + \sum_{i=1}^{k} \alpha_{i,s} G_i\right) t\right],$$

where we have introduced the last integral using the identity in Equation (3.22). The combination of Equation (3.32) and (3.33) makes it apparent that the evaluation of $\{\frac{\partial}{\partial \alpha_{i,j}} U(T)\}$ involves evaluating a set of matrix products

$$\left\{\prod_{j=1}^{s} \exp\left[\left(G_0 + \sum_{i=1}^{k} \alpha_{i,j} G_i\right) \Delta T_j\right]\right\}, \tag{3.34}$$

as well as

$$\left\{\prod_{j=s}^{N} \exp\left[\left(G_0 + \sum_{i=1}^{k} \alpha_{i,j} G_i\right) \Delta T_j\right]\right\}, \tag{3.35}$$

for all $s \in \{1, \ldots, N\}$, and all kN integral terms given by the last line of Equation (3.33). If we ignore the cost of calculating the integral terms, the evaluation of $U(T)$ costs N matrix exponentials and N matrix multiplications, while the simultaneous evaluation of $U(T)$ and $\{\frac{\partial}{\partial \alpha_{i,j}} U(T)\}$ costs N matrix exponentials and $2(k+1)N$ matrix multiplications – an extra computational cost which is almost always outweighed by the faster convergence of gradient based optimization algorithms.

The integral term in Equation (3.33) can be evaluated either through block matrix techniques introduced in Section 3.2 with Equation (3.23) or by noticing that

$$\int_0^{\Delta T_s} dt \ \exp\left[-\left(G_0 + \sum_{i=1}^{k} \alpha_{i,s} G_i\right) t\right] G_r \exp\left[\left(G_0 + \sum_{i=1}^{k} \alpha_{i,s} G_i\right) t\right] \tag{3.36}$$
$$= G_r \Delta T_s + \frac{1}{2}\left[G_r \Delta T_s, \left(G_0 + \sum_{i=1}^{k} \alpha_{i,s} G_i\right) \Delta T_s\right]$$
$$+ \frac{1}{6}\left[\left[G_r \Delta T_s, \left(G_0 + \sum_{i=1}^{k} \alpha_{i,s} G_i\right) \Delta T_s\right], \left(G_0 + \sum_{i=1}^{k} \alpha_{i,s} G_i\right) \Delta T_s\right] + \ldots,$$

and approximating the integral with a finite sum of commutators. During the optimization of the examples presented in this thesis, we always approximated the integral with a finite sum of 15 commutators, since the speed of the optimization was not our main concern. We point out that a very comprehensive treatment of bilinear control optimization performance is given in [30], this work also investigates the use of block matrix methods for evaluating the integral expression in Equation (3.36).

Khaneja *et al* [57] suggested that unitary engineering, i.e., the search of an α that would yield $U(T) = U_{\text{target}}$, where $U_{\text{target}} \in \mathrm{M}_n$ is some desired final unitary operation, could be carried out by maximizing the fidelity function $\mathcal{F}\left[U(T), U_{\text{target}}\right]$ defined in Equation (3.4). The explicit form for such a target function then becomes

$$\Phi(\alpha) = \frac{1}{n}\sqrt{\mathrm{Tr}\left[U(T)U_{\text{target}}^{\dagger}\right]\mathrm{Tr}\left[U(T)^{\dagger}U_{\text{target}}\right]}, \tag{3.37}$$

where $U(T)$ is given by Equation (3.31). The partial derivatives of $\{\frac{\partial}{\partial\alpha_{i,j}}\Phi(\alpha)\}$ can be evaluated simply using the chain rule, yielding

$$\frac{\partial}{\partial\alpha_{i,j}}\Phi(\alpha) = \frac{1}{n}\frac{\mathrm{Re}\left(\mathrm{Tr}\left[\frac{\partial}{\partial\alpha_{i,j}}U(T)U_{\text{target}}^{\dagger}\right]\mathrm{Tr}\left[U(T)^{\dagger}U_{\text{target}}\right]\right)}{\sqrt{\mathrm{Tr}\left[U(T)U_{\text{target}}^{\dagger}\right]\mathrm{Tr}\left[U(T)^{\dagger}U_{\text{target}}\right]}}, \tag{3.38}$$

where Re denotes the real part of a complex number and the partial derivatives of $U(T)$ are given by Equation (3.32). As we said in Section 3.1, the fidelity function satisfies $0 \leq \mathcal{F}\left[U(T), U_{\text{target}}\right] \leq 1$ for any unitary $U(T)$, and $\mathcal{F}\left[U(T), U_{\text{target}}\right] = 1$ if and only if $U(T) = U_{\text{target}}$. Khaneja *et al* [57] described a basic implementation of a gradient ascent algorithm that could be used for maximizing $\Phi(\alpha)$ starting from some initial $\alpha^{(0)} \in \mathrm{M}_{k,N}(\mathbb{R})$. The control finding procedure then entails generating a set of $\{\alpha^{(0)}\}$ and running a gradient ascent algorithm for each of these $\alpha^{(0)}$s until the algorithm yields a $\Phi(\alpha)$ value sufficiently close to one.

In this work, we do not just use the fidelity function as the target function for our optimizations, because our generators do not yield a $U(T)$ that is a unitary matrix. Moreover, we are frequently interested in having only certain parts of $U(T)$ satisfy certain properties. Therefore, the target functions we will employ are more elaborate than the one in Equation (3.37). Nevertheless, the general scheme of our control searches is much the same as the one proposed by Khaneja *et al* [57]. We always define target functions that are analytic functions of $U(T)$, i.e., $\Phi(\alpha) = \Phi[U(T)]$, $\Phi : \mathrm{M}_n \to [0, 1]$. The partial derivatives of $\Phi(\alpha)$ with respect to $\{\alpha_{i,j}\}$ are then evaluated in terms of $\{\frac{\partial}{\partial\alpha_{i,j}}U(T)\}$ using the chain

rule. Finally, we use an off-the-shelf gradient ascent optimizer to maximize $\Phi(\alpha)$ starting from a set of initial control sequences $\{\alpha^{(0)}\}$ just as it was described above. Unlike Khaneja *et al* [57], we did not attempt to implement our own gradient ascent optimizer since Mathematica's FindMaximum conjugate gradient ascent implementation proved to work extremely well.

3.4 Simple Examples

The aim of this section is to familiarize the reader with the basic idea of our block matrix methods with as little notational overhead and complexity as possible. We do this by demonstrating the use of block matrix computational methods for control sequences that fulfill certain AHT conditions with four simple examples that have an increasing level of complexity. We reserve the complete description of our methods and the treatment of Van Loan differential equations in their most general form for Section 3.5 and 3.7. All of the examples involve typical spin Hamiltonians. Nevertheless, we point out that there is nothing spin physics specific about the method nor the solutions, hence, a 'spin' here could also be understood as a 'two level quantum system'.

First, we introduce another shorthand that makes the definitions of our target functions in this section more concise. For any nested operator integral $\mathcal{D}_U(A_1, \ldots, A_n)$, we denote its maximum Hilbert-Schmidt norm maximized over all permissible control sequences $a(t)$, $t \in [0, T]$, as $\max_{a(t)} \|\mathcal{D}_U(A_1, \ldots, A_n)\|$. Because all examples in this section involve AHT terms with $n = 1$, we also point out that for any time independent matrix A_1 it holds that $\max_{a(t)} \|\mathcal{D}_U(A_1)\| = T\|A_1\|$.

We proceed by setting up three rather standard decoupling problems with known analytical solutions and in some cases arrive at pulses which resemble control sequences that have been known for some time. Our aim is not to reiterate these solutions, rather it is to demonstrate that the length of the control sequences found using block matrix numerical tools does not significantly exceed that of the analytically derived ones. This demonstration provides an encouraging starting point for employing the same tools to tackle far harder control problems, for which the search of analytical solutions is intractable. With the fourth example, we provide an illustration for a problem that demands simultaneous minimization of some AHT terms, while preserving or maximizing other AHT terms. Such control problems are very common in many sensing and spectroscopy applications. Furthermore, here, we do not search for control sequences that are intended to be experimentally applicable – we apply no transfer functions, control field distributions, pulse waveform bandwidth constraints etc. Only maximum amplitude constraints are used which generally yield pulses of

rather jagged form, however, such constraints are typically the only constraints considered when deriving analytical solutions. In Section 3.8 we give a much more experimentally realistic control search example and the next chapter is devoted on control engineering for nanoscale magnetic resonance experiments.

For our numerical pulse searches, we always use a target function Φ that is a linear combination of different matrix norms for various AHT terms and the fidelity of $U(T)$ with the target unitary U_{target}. A typical target function takes the following form:

$$\Phi = p_0 \left(\mathcal{F}[U(T), U_{\text{target}}]\right)^2 + \sum_{i>0} p_i \left(1 - \frac{\|\mathcal{D}_U(A_i)\|^2}{\left[\max_{a(t)} \|\mathcal{D}_U(A_i)\|\right]^2}\right), \tag{3.39}$$

where $0 \leq p_i \leq 1$, for all i, such that $\sum_i p_i = 1$. $\{p_i\}$ denote various weights of constituent optimization targets. It easy to see that $0 \leq \Phi(\alpha) \leq 1$, for all α, and equal to one if and only if $U(T) = U_{\text{target}}$, as well as $\mathcal{D}_U(A_i) = 0$, for all i. Contrary to the target function defined earlier in Equation (3.37), here, we work with the square of the fidelity metric and the matrix norms because this somewhat simplifies the evaluation of $\{\frac{\partial}{\partial \alpha_{i,j}}\Phi\}$, while yielding the same behaviour for the optimization. The linear form of the target function in Equation (3.39) is not strictly necessary, however, it does greatly simplify the calculation of partial derivatives of Φ with respect to the control amplitudes. We generally pick $\{p_i\}$ that assigns relatively equal weights to all optimization targets, although, we typically pick a p_0 value that is slightly lower than the weights of the AHT terms, because generating the desired final unitary is a somewhat easier optimization task than the rest. For the examples demonstrated in this section, we always define target functions, for which it is known from analytic solutions that the value $\Phi = 1$ is attainable.

All examples in this section concern either one or two spins. In the case of two spins, the control is taken to be global, i.e., the Rabi fields for either spin are identical. Furthermore, we say that our control sequence $a(t)$ only couples to the x and y angular momentum operators such that $a : [0, T] \to \mathbb{R}^2$. For notational convenience, we assign $a_x(t) = a_1(t)$ and $a_y(t) = a_2(t)$. For all searches in this section, we use amplitude constraints: $-\frac{1}{\sqrt{2}} \leq a_i(t) \leq \frac{1}{\sqrt{2}}$ for $i = \{x, y\}$, to ensure $|a(t)| \leq 1$, for all $t \in [0, T]$. The latter amplitude constraint is common for analytical pulse derivations. In the case of single spin problems, we define a generator

$$G_1(t) = -i\frac{a_x(t)}{2}\sigma_x - i\frac{a_y(t)}{2}\sigma_y, \tag{3.40}$$

which generates a unitary $U_1(t) = \mathcal{T} \exp\left[\int_0^t dt_1 G_1(t_1)\right]$. Correspondingly, in the case of

two spins identical spins, we define a generator

$$G_2(t) = -i\frac{a_x(t)}{2}\left(\sigma_x \otimes \mathbb{1}_2 + \mathbb{1}_2 \otimes \sigma_x\right) - i\frac{a_y(t)}{2}\left(\sigma_y \otimes \mathbb{1}_2 + \mathbb{1}_2 \otimes \sigma_y\right), \tag{3.41}$$

which generates the following separable unitary propagator $U_2(t) = \mathcal{T}\exp\left[\int_0^t dt_1 G_2(t_1)\right] = U_1(t) \otimes U_1(t)$. Since our numerical pulse searches assume piecewise constant control amplitudes, we specify the control amplitudes by a matrix $\alpha \in \mathrm{M}_{2,N}(\mathbb{R})$, where the elements of α are given by Equation (3.30).

All control searches were undertaken using the modified GRAPE algorithm for evaluating partial derivatives with respect to piecewise constant control amplitudes, the details of which were given in Section 3.3. Our general procedure for finding a working control sequence was to first pick a pulse length T and thereafter a number of time steps N. We kept all subintervals of $[0, T]$ of equal length $\Delta T = T/N$ and always picked an N for which $\Delta T < \tau_{\text{Rabi}}/30$, where τ_{Rabi} is the length of a Rabi cycle. Given some T and N, we generated $\sim$40 seed waveforms $\alpha^{(0)}$, the pulse amplitudes $\{\alpha^{(0)}_{i,j}\}$ of which were drawn from independent uniform distributions. We used Mathematica's FindMaximum function for multivariate conjugate gradients optimization on these seeds with the maximum number of target function evaluations set to 1000. If none of the $\sim$40 searches yielded $\Phi < 0.9999$ we increased T and N and repeated the same procedure. For the seeds that reached $\Phi < 0.9999$ in under 1000 Φ evaluations, we let the optimization run until FindMaximum was terminated by machine precision.

When presenting the control sequences found in the figures below, we rotate the $G(t)$ basis in the following way: $\sigma_x \to \frac{1}{\sqrt{2}}(\sigma_x + \sigma_y)$ and $\sigma_y \to \frac{1}{\sqrt{2}}(\sigma_x - \sigma_y)$. This is done since the basis change tends to highlight the similarities between the pulses found and known analytic solutions. Furthermore, because for all examples in this section, we are either not concerned about the final unitary $U_1(T)$ generated or we demand that $U_1(T) = \mathbb{1}_2$, such basis change does not affect the desired characteristics of the control sequence.

3.4.1 Dipolar Decoupling

The simplest numerical AHT example that we consider is the problem of dipolar decoupling. For such a problem, we imagine two spins coupled via the dipolar Hamiltonian $D = 3\sigma_z \otimes \sigma_z - \sum_{i\in\{x,y,z\}} \sigma_i \otimes \sigma_i$. The control task is to engineer a sequence that enables the spins to evolve effectively uncoupled. Zeroth order AHT solution to the problem dictates setting the first derivative of $U_2(T)$ in the direction of D to zero, i.e., we want

$U_2^{-1}(T)\mathcal{D}_{U_2}(D) = \int_0^T dt_1\ U_2^{-1}(t_1)DU_2(t_1) = 0$. The dipolar decoupling problem is a simple yet non-trivial problem – it is easy to show that the desired $U_2(t)$, $t \in [0,T]$, that yields $U_2^{-1}(T)\mathcal{D}_{U_2}(D) = 0$ cannot be generated with either $a_x(t) = 0$ for all $t \in [0,T]$ or $a_y(t) = 0$ for all $t \in [0,T]$.

Here, we are not concerned about the final unitary $U_1(T)$ generated on either of the spins. Accordingly, our target function for the optimization is

$$\Phi = 1 - \frac{\|U_2^{-1}(T)\mathcal{D}_{U_2}(D)\|^2}{\left[\max_{a(t)} \|U_2^{-1}(T)\mathcal{D}_{U_2}(D)\|\right]^2}, \tag{3.42}$$

where the denominator is a normalization factor ensuring that $0 \leq \Phi \leq 1$. We now set up a block matrix differential equation for $V(t) \in \mathrm{M}_2(\mathrm{M}_4)$, that will be used for evaluating Φ. It easy to show either by differentiation or by employing Theorem 1 that

$$V(t) = \mathcal{T}\exp\left[\int_0^t dt_1 \begin{pmatrix} G_2(t_1) & D \\ 0 & G_2(t_1) \end{pmatrix}\right] = \begin{pmatrix} U_2(t) & \mathcal{D}_{U_2}(D) \\ 0 & U_2(t) \end{pmatrix}. \tag{3.43}$$

Consequently, our target could also be given as

$$\Phi = 1 - \frac{\mathrm{Tr}\left[V_{1,2}^\dagger(T)V_{1,2}(T)\right]}{24T^2}, \tag{3.44}$$

where we have used the fact that $\left[\max_{a(t)} \|\mathcal{D}_{U_2}(D)\|\right]^2 = 24T^2$ and that $\|U_2^{-1}(T)\mathcal{D}_{U_2}(D)\|^2 = \|\mathcal{D}_{U_2}(D)\|^2$. Following the procedure outlined in Section 3.3, we can evaluate $V(T)$ as a function of α along with the partial derivatives $\{\frac{\partial}{\partial\alpha_{i,j}}V(T)\}$ with respect to the piecewise constant control amplitudes $\{\alpha_{i,j}\}$.

Given $V(T)$ and $\{\frac{\partial}{\partial\alpha_{i,j}}V(T)\}$, we can write

$$\frac{\partial}{\partial\alpha_{i,j}}\Phi = -\frac{2}{24T^2}\mathrm{Re}\left(\mathrm{Tr}\left[V_{1,2}^\dagger(T)\left(\frac{\partial}{\partial\alpha_{i,j}}V(T)\right)_{1,2}\right]\right), \tag{3.45}$$

for all i and j, where we have used the fact that $\frac{\partial}{\partial\alpha_{i,j}}$ and the projection operations commute. We employ these partial derivatives in the optimization protocol described above. We find a pulse with a total length of $T = 6.2$ consisting of $N = 100$ subintervals shown in Figure 3.1 which yields the following norm for $U_2^{-1}(T)\mathcal{D}_{U_2}(D)$:

$\frac{\|U_2^{-1}(T)\mathcal{D}_{U_2}(D)\|}{\max_{a(t)} \|U_2^{-1}(T)\mathcal{D}_{U_2}(D)\|}$	3.09×10^{-7}

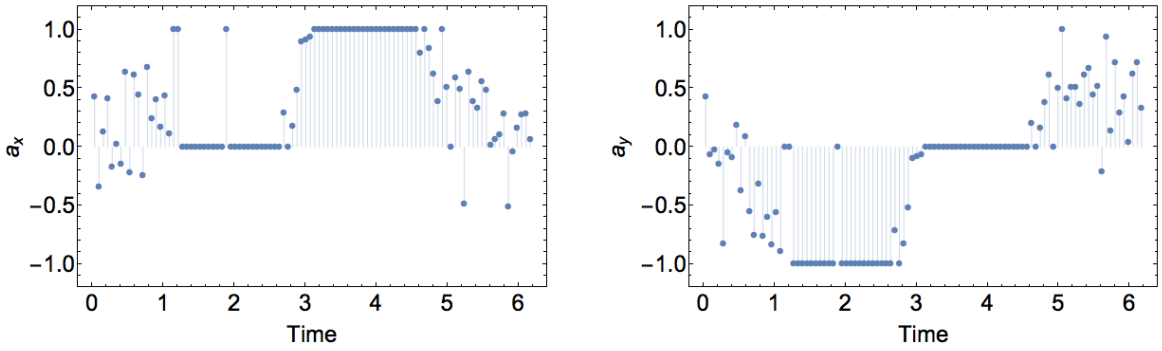

Figure 3.1: Numerically found dipolar decoupling control sequence: $a_x(t)$ on the left and $a_y(t)$ on the right.

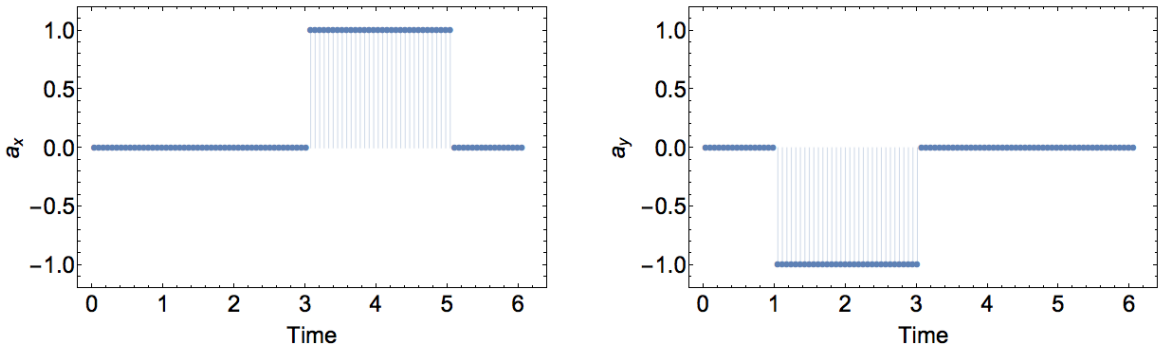

Figure 3.2: Dipolar decoupling pulse sequence introduced by Mehring: $a_x(t)$ on the left and $a_y(t)$ on the right.

We point out the rough similarity to the dipolar sequence proposed by Mehring [79] consisting of two square 116°14′ pulses with orthogonal phases depicted in Figure 3.2. Our numerically found sequence is only 1.02 times longer than the known analytical solution. We also remark that further optimization of the target would likely be senseless for any practical case, since at that point the second derivative of $U_2(T)$ in the direction of D is expected to become more important than $U_2^{-1}(T)\mathcal{D}_{U_2}(D)$.

3.4.2 Universal Decoupling

With the next two subsections, we turn to single spin pulses. As the first example, we consider a universal decoupling sequence which would decouple all non-identity operators $\sigma_x, \sigma_z, \sigma_y$ acting on a single spin at the zeroth order AHT level. This translates to setting $U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_x) = U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_y) = U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_z) = 0$. A sequence known to satisfy these conditions, called the XY8 sequence [36], is shown in Figure 3.5. For that sequence $U(T) = \mathbb{1}_2$. Hence, in addition to the AHT integral terms we also require the final unitary operation on the spin to be an identity. To achieve this, we use the following target:

$$\Phi = \frac{4}{5}\left(1 - \frac{1}{3}\sum_{i\in\{x,y,z\}} \frac{\|U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_i)\|^2}{\left[\max_{a(t)} \|U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_i)\|\right]^2}\right) + \frac{1}{5}\left(\mathcal{F}\left[\mathbb{1}_2, U_1(T)\right]\right)^2. \tag{3.46}$$

Again, we can define a block matrix propagator $V(t) \in \mathrm{M}_4(\mathrm{M}_2)$ that enables us to calculate the integral terms as well as the final unitary. It is easy to see that

$$\begin{aligned} V(t) &= \mathcal{T}\exp\left[\int_0^t dt_1 \begin{pmatrix} G_1(t_1) & \sigma_x & 0 & 0 \\ 0 & G_1(t_1) & \sigma_y & 0 \\ 0 & 0 & G_1(t_1) & \sigma_z \\ 0 & 0 & 0 & G_1(t_1) \end{pmatrix}\right] \\ &= \begin{pmatrix} U_1(t) & \mathcal{D}_{U_1}(\sigma_x) & \mathcal{D}_{U_1}(\sigma_x,\sigma_y) & \mathcal{D}_{U_1}(\sigma_x,\sigma_y,\sigma_z) \\ 0 & U_1(t) & \mathcal{D}_{U_1}(\sigma_y) & \mathcal{D}_{U_1}(\sigma_y,\sigma_z) \\ 0 & 0 & U_1(t) & \mathcal{D}_{U_1}(\sigma_z) \\ 0 & 0 & 0 & U_1(t) \end{pmatrix} \end{aligned} \tag{3.47}$$

is true and suitable. Our target Φ can therefore be given as

$$\Phi = \frac{4}{5}\left(1 - \frac{1}{3}\sum_{i=1}^{3}\frac{\mathrm{Tr}\left[V_{i,i+1}^{\dagger}(T)V_{i,i+1}(T)\right]}{2T^2}\right) + \frac{1}{5}\frac{\mathrm{Tr}\left[V_{1,1}^{\dagger}(T)\right]\mathrm{Tr}\left[V_{1,1}(T)\right]}{4}, \tag{3.48}$$

where we have used the fact that $\max_{a(t)} \|U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_i)\| = \sqrt{2}T$, for all $i \in \{x, y, z\}$, and that $\mathrm{Tr}\left[V_{1,1}^{\dagger}(T)V_{1,1}(T)\right]\mathrm{Tr}\left[\mathbb{1}_2\right] = 4$. Using the chain rule, we can again evaluate the partial derivatives of Φ with respect to the piecewise constant control amplitudes $\{\alpha_{i,j}\}$:

$$\begin{aligned}\frac{\partial}{\partial\alpha_{i,j}}\Phi = -\frac{4}{15T^2}\sum_{i=1}^{3}\mathrm{Re}\left(\mathrm{Tr}\left[V_{i,i+1}^{\dagger}(T)\left(\frac{\partial}{\partial\alpha_{i,j}}V(T)\right)_{i,i+1}\right]\right) \\ + \frac{1}{10}\mathrm{Re}\left(\mathrm{Tr}\left[V_{1,1}^{\dagger}(T)\right]\mathrm{Tr}\left[\left(\frac{\partial}{\partial\alpha_{i,j}}V(T)\right)_{1,1}\right]\right),\end{aligned} \tag{3.49}$$

for all i and j. Equipped with $\Phi(\alpha)$ and $\{\frac{\partial}{\partial\alpha_{i,j}}\Phi(\alpha)\}$ we use the procedure described above to search for a piecewise constant control sequence $a(t)$ that yields the maximum $\Phi(\alpha)$ value. We find a sequence with a total length of $T = 19.6$ consisting of $N = 200$ subintervals which has the following characteristics:

$\frac{\|U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_x)\|}{\max_{a(t)}\|U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_x)\|}$	8.22×10^{-7}
$\frac{\|U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_y)\|}{\max_{a(t)}\|U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_y)\|}$	4.72×10^{-7}
$\frac{\|U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_z)\|}{\max_{a(t)}\|U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_z)\|}$	7.03×10^{-6}
$1 - \mathcal{F}\left[\mathbb{1}_2, U_1(T)\right]$	1.66×10^{-12}

The control sequence is shown in Figure 3.3. It is interesting to note that the sequence found is 0.78 times the length of the XY8 sequence.

3.4.3 Universal Decoupling with Control Variations

While the XY8 sequence [36] satisfies the conditions stated in the previous subsection, it is also robust against variations in the controls, i.e., it satisfies $U_1^{-1}(T)\mathcal{D}_{U_1}\left[a_x(t)\sigma_x\right] = 0$

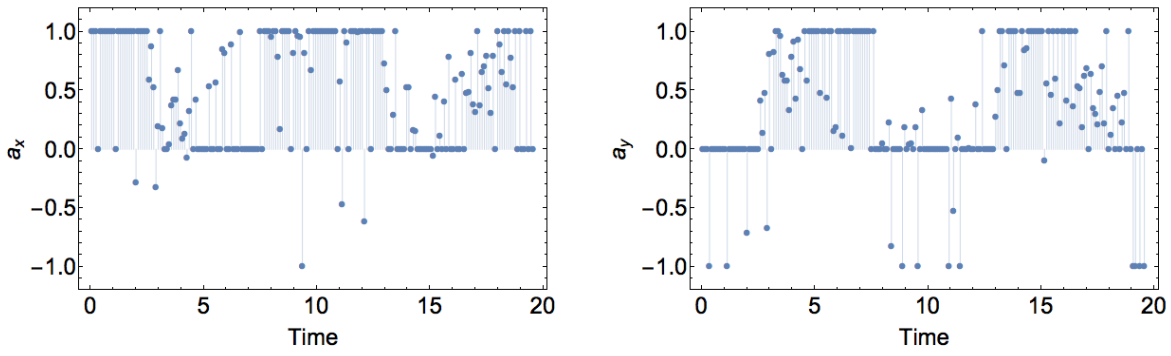

Figure 3.3: Numerically found universal decoupling sequence implementing an $\mathbb{1}_2$ gate: $a_x(t)$ on the left and $a_y(t)$ on the right.

and $U_1^{-1}(T)\mathcal{D}_{U_1}\left[a_y(t)\sigma_y\right] = 0$. This robustness is the property that ensures its excellent performance under realistic experimental conditions. To demonstrate the ability of our numerical control searches in incorporating averaging of time dependent operators, we set up a search that would simultaneously set $U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_x) = U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_y) = U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_z) = U_1^{-1}(T)\mathcal{D}_{U_1}\left[a_x(t)\sigma_x\right] = U_1^{-1}(T)\mathcal{D}_{U_1}\left[a_y(t)\sigma_y\right] = 0$, while implementing an identity operation. We search for a pulse with the following target:

$$\Phi = \frac{2}{5}\left(1 - \frac{1}{3}\sum_{i\in\{x,y,z\}} \frac{\|U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_i)\|^2}{\left[\max_{a(t)}\|U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_i)\|\right]^2}\right) + \frac{2}{5}\left(1 - \frac{1}{2}\sum_{i\in\{x,y\}} \frac{\|U_1^{-1}(T)\mathcal{D}_{U_1}\left[a_i(t)\sigma_i\right]\|^2}{\left[\max_{a(t)}\|U_1^{-1}(T)\mathcal{D}_{U_1}\left[a_i(t)\sigma_i\right]\|\right]^2}\right) + \frac{1}{5}\left(\mathcal{F}\left[\mathbb{1}_2, U_1(T)\right]\right)^2. \tag{3.50}$$

A suitable block matrix propagator $V(t) \in \mathrm{M}_6(\mathrm{M}_2)$ for such a target function is

$$V(t) = \tag{3.51}$$
$$\mathcal{T}\exp\left[\int_0^t dt_1 \begin{pmatrix} G_1(t_1) & \sigma_x & 0 & 0 & 0 & 0 \\ 0 & G_1(t_1) & \sigma_y & 0 & 0 & 0 \\ 0 & 0 & G_1(t_1) & \sigma_z & 0 & 0 \\ 0 & 0 & 0 & G_1(t_1) & a_x(t_1)\sigma_x & 0 \\ 0 & 0 & 0 & 0 & G_1(t_1) & a_y(t_1)\sigma_y \\ 0 & 0 & 0 & 0 & 0 & G_1(t_1) \end{pmatrix}\right] =$$
$$\begin{pmatrix} U_1(t) & \mathcal{D}_{U_1}(\sigma_x) & \cdot & \cdot & \cdot & \cdot \\ 0 & U_1(t) & \mathcal{D}_{U_1}(\sigma_y) & \cdot & \cdot & \cdot \\ 0 & 0 & U_1(t) & \mathcal{D}_{U_1}(\sigma_z) & \cdot & \cdot \\ 0 & 0 & 0 & U_1(t) & \mathcal{D}_{U_1}\left[a_x(t)\sigma_x\right] & \cdot \\ 0 & 0 & 0 & 0 & U_1(t) & \mathcal{D}_{U_1}\left[a_y(t)\sigma_y\right] \\ 0 & 0 & 0 & 0 & 0 & U_1(t) \end{pmatrix}.$$

Note that, we have not specified the $V(t)$ elements that are not relevant for our control problem. Given the $V(t)$ above, Φ can be determined as

$$\Phi = \frac{2}{5}\left(1 - \frac{1}{3}\sum_{i=1}^{3}\frac{\mathrm{Tr}\left[V_{i,i+1}^{\dagger}(T)V_{i,i+1}(T)\right]}{2T^2}\right) + \frac{2}{5}\left(1 - \frac{1}{2}\sum_{i=4}^{5}\frac{\mathrm{Tr}\left[V_{i,i+1}^{\dagger}(T)V_{i,i+1}(T)\right]}{T^2}\right)$$
$$+ \frac{1}{5}\frac{\mathrm{Tr}\left[V_{1,1}^{\dagger}(T)\right]\mathrm{Tr}\left[V_{1,1}(T)\right]}{4}, \tag{3.52}$$

where we have used the fact that $\max_{a(t)} \|U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_i)\| = \sqrt{2}T$, for all $i \in \{x, y, z\}$, and that $\mathrm{Tr}\left[V_{1,1}^{\dagger}(T)V_{1,1}(T)\right]\mathrm{Tr}\left[\mathbb{1}_2\right] = 4$. In the above, we also used that

$$\max_{a(t)} \|U_1^{-1}(T)\mathcal{D}_{U_1}(a_i(t)\sigma_i)\| = T, \tag{3.53}$$

for $i \in \{x, y\}$, which corresponds to having $a_i(t)$ take its maximum value $a_i(t) = 1/\sqrt{2}$ for all $t \in [0, T]$. Again, we can evaluate the partial derivatives of $\Phi(\alpha)$ with respect to the

piecewise constant control amplitudes:

$$\begin{aligned}\frac{\partial}{\partial\alpha_{i,j}}\Phi = &-\frac{2}{15T^2}\sum_{i=1}^{3}\mathrm{Re}\left(\mathrm{Tr}\left[V_{i,i+1}^{\dagger}(T)\left(\frac{\partial}{\partial\alpha_{i,j}}V(T)\right)_{i,i+1}\right]\right)\\ &-\frac{2}{5T^2}\sum_{i=4}^{5}\mathrm{Re}\left(\mathrm{Tr}\left[V_{i,i+1}^{\dagger}(T)\left(\frac{\partial}{\partial\alpha_{i,j}}V(T)\right)_{i,i+1}\right]\right)\\ &+\frac{1}{10}\mathrm{Re}\left(\mathrm{Tr}\left[V_{1,1}^{\dagger}(T)\right]\mathrm{Tr}\left[\left(\frac{\partial}{\partial\alpha_{i,j}}V(T)\right)_{1,1}\right]\right).\end{aligned} \tag{3.54}$$

Using the above, we find a pulse with a total length of $T = 30$ consisting of $N = 200$ subintervals with its pulse metrics given below:

$\frac{\|U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_x)\|}{\max_{a(t)}\|U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_x)\|}$	2.16×10^{-6}
$\frac{\|U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_y)\|}{\max_{a(t)}\|U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_y)\|}$	2.35×10^{-6}
$\frac{\|U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_z)\|}{\max_{a(t)}\|U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_z)\|}$	1.62×10^{-7}
$\frac{\|U_1^{-1}(T)\mathcal{D}_{U_1}[a_x(t)\sigma_x]\|}{\max_{a(t)}\|U_1^{-1}(T)\mathcal{D}_{U_1}[a_x(t)\sigma_x]\|}$	6.24×10^{-6}
$\frac{\|U_1^{-1}(T)\mathcal{D}_{U_1}[a_y(t)\sigma_y]\|}{\max_{a(t)}\|U_1^{-1}(T)\mathcal{D}_{U_1}[a_y(t)\sigma_y]\|}$	6.24×10^{-6}
$1-\mathcal{F}\left[\mathbb{1}_2, U_1(T)\right]$	2.83×10^{-14}

We present the pulse waveform in Figure 3.4. We note that, while the pulse found is 1.19 times the length of the XY8 sequence [36], it does display definite similarities to the latter shown in Figure 3.5.

3.4.4 Exchange Interaction Recoupling

With the following example, we wish to highlight that the block matrix method does not only enable the removal of unwanted Hamiltonian terms; it is equally easy to set up optimization targets which retain or reshape parts of the Hamiltonian, while possibly

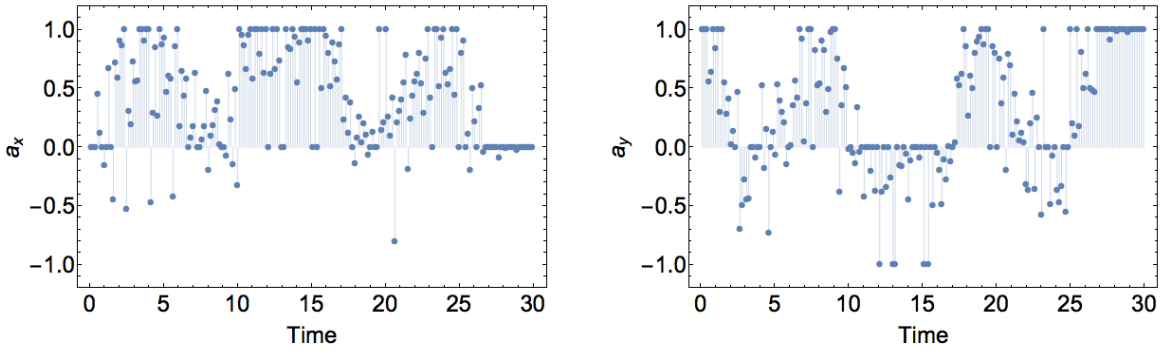

Figure 3.4: Numerically found universal decoupling pulse robust to control variations implementing an $\mathbb{1}_2$ gate: $a_x(t)$ on the left and $a_y(t)$ on the right. It should be seen that this sequence bares significant resemblance to the XY8 sequence shown in Figure 3.5.

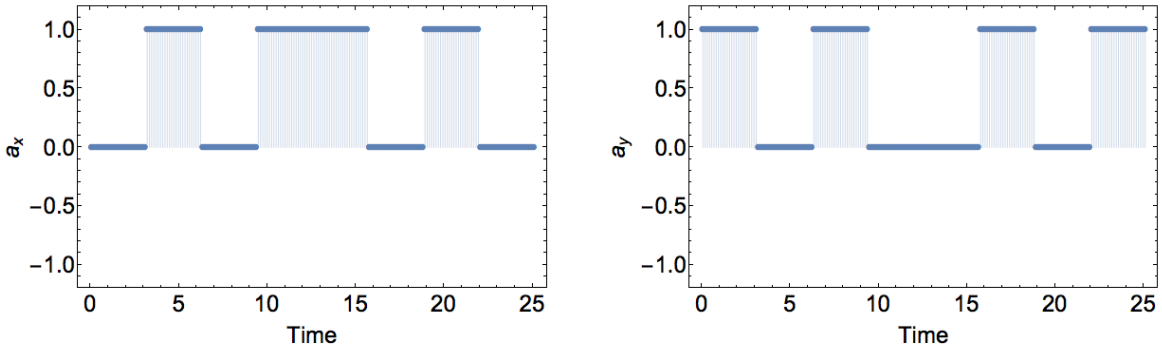

Figure 3.5: XY8 sequence: $a_x(t)$ on the left and $a_y(t)$ on the right.

removing others. A problem, which arises in many situations involving ensembles of spins, is removing pairwise dipolar interactions between all members of the ensemble as well as inhomogeneities in their energy level splittings, while retaining exchange interaction with some other system or systems the spins are interacting with.

Such a situation would be described by the following Hamiltonian:

$$H_{\text{exchange}} = \sum_i \Delta\omega_i \sigma_z^{(i)} + \sum_{\langle i,j\rangle} \xi_{i,j} D^{(i,j)} + \sum_i g_i \left(\sigma_+^{(i)} \otimes q^{(i)} + \sigma_-^{(i)} \otimes \left[q^{(i)}\right]^\dagger \right), \tag{3.55}$$

where $\sigma_\pm = (\sigma_x \pm i\sigma_y)/2$ and $\langle i,j\rangle$ denotes a sum over all spin pairs. The first sum in Equation (3.55) specifies these spin dependent energy level splitting variations $\{\Delta\omega_i\}$, the second sum gives all dipolar interaction strengths $\{\xi_{i,j}\}$ that correspond to the dipolar Hamiltonian $D^{(i,j)}$ for a pair of spins. The third sum contains the aforementioned exchange interactions $\sigma_+^{(i)} \otimes q^{(i)} + \sigma_-^{(i)} \otimes \left[q^{(i)}\right]^\dagger$ that can often be substantially weaker than the other two terms, yet this is frequently the term in the Hamiltonian that leads to desirable spin dynamics. The exchange interactions could, for example, arise from spin cavity interactions in which case $q^{(i)} = q^{(j)} = q$, for all i and j, would stand for the annihilation operator acting on the cavity. Alternatively, the exchange interaction could also arise from interactions with a different ensemble of spins, e.g., the main ensemble in Equation (3.55) could be an ensemble of electron spins, which is used for polarization enhancement of some nuclear spin ensemble through an effective exchange interaction. In such case $q^{(i)} = \sigma_-^{\prime(i)}$, where $\sigma_-^{\prime(i)}$ is a lowering operator acting on the ith nuclear spin.

Our block matrix tools enable us to search for control sequences that would effectively remove the spin-spin dipolar couplings and variations in level splittings (σ_z offsets), while retaining the form of the exchange interaction and performing an identity operation. To achieve this, zeroth order AHT prescribes setting $U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_z) = U_2^{-1}(T)\mathcal{D}_{U_2}(D) = 0$, $U_1(T) = \mathbb{1}_2$ and $U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_+) = c\mathcal{D}_{U_1}(\sigma_+)$, where $c \in \mathbb{R}$ is a constant. In order to set up a target function Φ that can reach its maximum value, we do not set up the optimization with any specific value for c. Instead, we merely enforce that the integral $I = U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_+)$ is proportional to $\mathcal{D}_{U_1}(\sigma_+)$. This is achieved by demanding that I is orthogonal to σ_z and σ_-, i.e., $\text{Tr}\left(\sigma_z^\dagger I\right) = \text{Tr}\left(\sigma_-^\dagger I\right) = 0$. We do not demand that I be orthogonal to $\mathbb{1}_2$ since $\text{Tr}(I) = \int_0^T Tr[\sigma_+] = 0$. Correspondingly, the optimization target

for this problem is

$$\Phi = \frac{2}{5}\left(1 - \frac{\|U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_z)\|^2}{2\left[\max_{a(t)}\|U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_z)\|\right]^2} - \frac{\|U_2^{-1}(T)\mathcal{D}_{U_2}(D)\|^2}{2\left[\max_{a(t)}\|U_2^{-1}(T)\mathcal{D}_{U_2}(D)\|\right]^2}\right)$$
$$+ \frac{2}{5}\left(1 - \frac{\left|\mathrm{Tr}\left[\sigma_z^\dagger U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_+)\right]\right|^2}{2\left[\max_{a(t)}\|U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_+)\|\right]^2} - \frac{\left|\mathrm{Tr}\left[\sigma_-^\dagger U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_+)\right]\right|^2}{2\left[\max_{a(t)}\|U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_+)\|\right]^2}\right)$$
$$+ \frac{1}{5}\left(\mathcal{F}\left[\mathbb{1}_2, U_1(T)\right]\right)^2. \tag{3.56}$$

In order to simplify the calculation of the target function and its partial derivatives below we, will replace the second line of Equation (3.56) with

$$\frac{2}{5}\left(1 - \frac{\left|\mathrm{Tr}\left[\sigma_z^\dagger \mathcal{D}_{U_1}(\sigma_+)\right]\right|^2}{2\left[\max_{a(t)}\|U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_+)\|\right]^2} - \frac{\left|\mathrm{Tr}\left[\sigma_-^\dagger \mathcal{D}_{U_1}(\sigma_+)\right]\right|^2}{2\left[\max_{a(t)}\|U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_+)\|\right]^2}\right), \tag{3.57}$$

which is allowed because we are looking for solutions for which $U_1(T)$ and its inverse will be very close to $\mathbb{1}_2$. If our target unitary was not an identity operation this replacement would not be appropriate. Nevertheless, at the expense of slightly more involved derivative terms, this scenario could also be handled in much the same way.

This time, we define a block matrix propagator $V(t) \in \mathrm{M}_{14}$ which decomposes into a direct sum of $\mathrm{M}_3(\mathrm{M}_2)$ and $\mathrm{M}_2(\mathrm{M}_4)$ and helps us evaluate Φ:

$$V(t) = \mathcal{T}\exp\left[\int_0^t dt_1 \begin{pmatrix} G_1(t_1) & \sigma_z & 0 & 0 & 0 \\ 0 & G_1(t_1) & \sigma_+ & 0 & 0 \\ 0 & 0 & G_1(t_1) & 0 & 0 \\ 0 & 0 & 0 & G_2(t_1) & D \\ 0 & 0 & 0 & 0 & G_2(t_1) \end{pmatrix}\right] \tag{3.58}$$
$$= \begin{pmatrix} U_1(t) & \mathcal{D}_{U_1}(\sigma_z) & \mathcal{D}_{U_1}(\sigma_z,\sigma_+) & 0 & 0 \\ 0 & U_1(t) & \mathcal{D}_{U_1}(\sigma_+) & 0 & 0 \\ 0 & 0 & U_1(t) & 0 & 0 \\ 0 & 0 & 0 & U_2(t) & \mathcal{D}_{U_2}(D) \\ 0 & 0 & 0 & 0 & U_2(t) \end{pmatrix}.$$

In the following, when writing our target function Φ in terms of $V(T)$, we will slightly abuse our notation for specifying the components of a block matrix. We take $V_{i,j}(T)$ to mean the

ith row and jth column of $V(T)$ as it is specified above. However, note that not all blocks of $V(T)$ are of the same size, e.g, $V_{4,5}(T) = \mathcal{D}_{U_2}(D) \in \mathrm{M}_4$ while $V_{2,3}(T) = \mathcal{D}_{U_1}(\sigma_+) \in \mathrm{M}_2$. It can now be seen that

$$\Phi = \frac{2}{5}\left(1 - \frac{1}{2}\frac{\mathrm{Tr}\left[V_{1,2}^\dagger(T)V_{1,2}(T)\right]}{2T^2} - \frac{1}{2}\frac{\mathrm{Tr}\left[V_{4,5}^\dagger(T)V_{4,5}(T)\right]}{24T^2}\right) + \frac{1}{5}\frac{\mathrm{Tr}\left[V_{1,1}^\dagger(T)\right]\mathrm{Tr}\left[V_{1,1}(T)\right]}{4}$$
$$+ \frac{2}{5}\left(1 - \frac{1}{2}\frac{\mathrm{Tr}\left[\sigma_z^\dagger V_{2,3}(T)\right]\mathrm{Tr}\left[\sigma_z V_{2,3}^\dagger(T)\right]}{T^2} - \frac{1}{2}\frac{\mathrm{Tr}\left[\sigma_-^\dagger V_{2,3}(T)\right]\mathrm{Tr}\left[\sigma_- V_{2,3}^\dagger(T)\right]}{T^2}\right), \quad (3.59)$$

where we have used the fact that $\left[\max_{a(t)} \|\mathcal{D}_{U_2}(D)\|\right]^2 = 24T^2$, $\left[\max_{a(t)} \|\mathcal{D}_{U_1}(\sigma_z)\|\right]^2 = 2T^2$, $\left[\max_{a(t)} \|\mathcal{D}_{U_1}(\sigma_+)\|\right]^2 = T^2$ and $\mathrm{Tr}\left[V_{1,1}^\dagger(T)V_{1,1}(T)\right]\mathrm{Tr}\left[\mathbb{1}_2\right] = 4$. The partial derivatives of Φ with respect to $\{\alpha_{i,j}\}$ evaluate to

$$\begin{aligned}
\frac{\partial}{\partial\alpha_{i,j}}\Phi = &-\frac{1}{5T^2}\mathrm{Re}\left(\mathrm{Tr}\left[V_{1,2}^\dagger(T)\left(\frac{\partial}{\partial\alpha_{i,j}}V(T)\right)_{1,2}\right]\right) \\
&- \frac{1}{60T^2}\mathrm{Re}\left(\mathrm{Tr}\left[V_{4,5}^\dagger(T)\left(\frac{\partial}{\partial\alpha_{i,j}}V(T)\right)_{4,5}\right]\right) \\
&+ \frac{1}{10}\mathrm{Re}\left(\mathrm{Tr}\left[V_{1,1}^\dagger(T)\right]\mathrm{Tr}\left[\left(\frac{\partial}{\partial\alpha_{i,j}}V(T)\right)_{1,1}\right]\right) \\
&- \frac{2}{5T^2}\mathrm{Re}\left(\mathrm{Tr}\left[\sigma_z^\dagger\left(\frac{\partial}{\partial\alpha_{i,j}}V(T)\right)_{2,3}\right]\mathrm{Tr}\left[\sigma_z V_{2,3}^\dagger(T)\right]\right) \\
&- \frac{2}{5T^2}\mathrm{Re}\left(\mathrm{Tr}\left[\sigma_-^\dagger\left(\frac{\partial}{\partial\alpha_{i,j}}V(T)\right)_{2,3}\right]\mathrm{Tr}\left[\sigma_- V_{2,3}^\dagger(T)\right]\right). \quad (3.60)
\end{aligned}$$

Using the partial derivatives in our optimization scheme, we arrive at a control sequence with a total length of $T = 24$ consisting of $N = 200$ subintervals with the following characteristics:

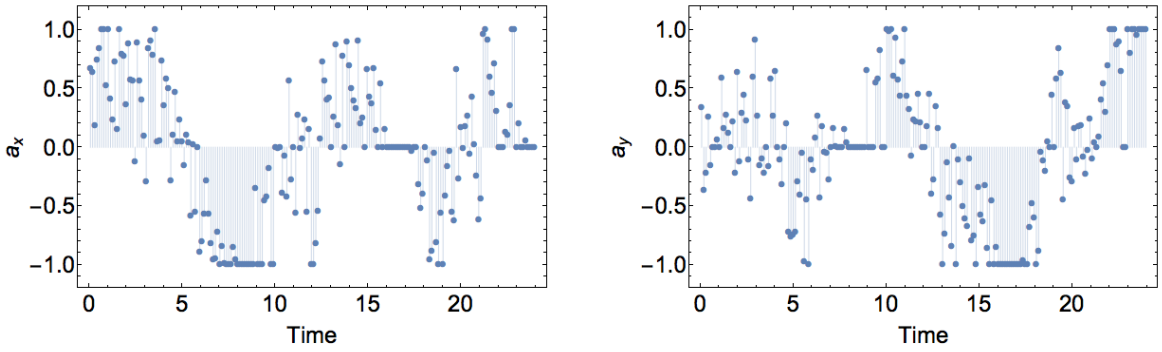

Figure 3.6: Exchange recoupling pulse sequence: $a_x(t)$ on the left and $a_y(t)$ on the right.

$\frac{\|U_2^{-1}(T)\mathcal{D}_{U_2}(D)\|}{\max_{a(t)}\|U_2^{-1}(T)\mathcal{D}_{U_2}(D)\|}$	5.41×10^{-6}
$\frac{\|U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_z)\|}{\max_{a(t)}\|U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_z)\|}$	3.11×10^{-7}
$\frac{\|U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_+)\|}{\max_{a(t)}\|U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_+)\|}$	0.476
$\frac{\left\|\mathrm{Tr}\left[\sigma_-^\dagger \mathcal{D}_{U_1}(\sigma_+)\right]\right\|}{\max_{a(t)}\|U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_+)\|}$	1.49×10^{-7}
$\frac{\left\|\mathrm{Tr}\left[\sigma_z^\dagger \mathcal{D}_{U_1}(\sigma_+)\right]\right\|}{\max_{a(t)}\|U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_+)\|}$	1.66×10^{-8}
$1 - \mathcal{F}\left[\mathbb{1}_2, U_1(T)\right]$	$< 10^{-16}$

The sequence is presented in Figure 3.6. It can be seen from the table above that the pulse does virtually remove the dipolar and σ_z Hamiltonians while rescaling the exchange coupling $U_1^{-1}(T)\mathcal{D}_{U_1}(\sigma_+)$ by a factor of 0.47 with extremely small ($< 2 \times 10^{-7}$) unwanted orthogonal components.

3.5 Setup of the Full Control Problem

In this section, we build on the block matrix methods introduced so far and give a full description of the quantum control problems that are addressed by the tools developed in this work. We hope that the examples presented so far will provide a sufficient guide

for understanding the rather abstract descriptions in this section. We also add that the problem description outlined in this section maps almost one to one to our implementation of the control optimizer in Mathematica. In fact, a lot of our treatment and notation has been chosen specifically to simplify the implementation process, while retaining full generality.

With Figure 3.7, we illustrate the general control setting addressed. We say that we have a finite set of quantum systems labelled by $\gamma \in \Gamma$. In principle, this is true for any control setting, although in many practical cases one approximates macroscopic ensembles of quantum systems as being parametrized by some set of continuous variables – Rabi field strengths, resonance offsets, etc – in such cases, we think of Γ as a representative sample of the real ensemble. Note that γ is a single compound label, that could encapsulate any number of constituent system parametrizations, which we use in order to keep our notation as clear and concise as possible. Furthermore, the ensemble Γ could in some cases denote the same quantum system under different conditions for distinct experimental realizations, i.e., it might stand for an ensemble in time rather than a spatial ensemble of physical systems. For all examples in the previous section, we had $|\Gamma| = 1$ to simplify the initial discussions. In the next chapter, we will provide an example, for which $|\Gamma| \neq 1$. That example was used for nanoscale magnetic resonance experiments.

The ensemble Γ of quantum systems is controlled by the *control sequence* $a(t)$ – a real vector valued function specified over an interval $[0, T]$ and delivered by some control signal source. The source should be thought of as a physical device which outputs $a(t)$, $a : [0, T] \to \mathbb{R}^k$; usually, we think of it as an arbitrary waveform generator (AWG). In the context of this thesis, we regard $a(t)$ as the waveform generated by our numerical pulse search routines. Each quantum system labelled by γ has an associated *transfer function* $\Xi^{(\gamma)}$, $\Xi^{(\gamma)} : \phi^k \to \phi^l$, where we use ϕ^i to denote the space of real vector valued functions, i.e., $\phi^i : [0, T] \to \mathbb{R}^i$. $\Xi^{(\gamma)}$ is an analytic deterministic map which transforms the control sequence $a(t)$ to system specific *control amplitudes* $b^{(\gamma)}(t) = \Xi^{(\gamma)}[a(t)]$, $b^{(\gamma)} : [0, T] \to \mathbb{R}^l$. The components of $b^{(\gamma)}(t)$ are the real valued functions that appear in the matrix differential equation determining the evolution of system $\gamma \in \Gamma$. In the next section, we demonstrate how to construct $\Xi^{(\gamma)}$ for piece-wise constant control sequences and control amplitudes in the case of linear transfer functions.

All quantum control problems boil down to engineering quantum state trajectories with certain desired properties. Mathematically this corresponds to generating a *system propagator* $U^{(\gamma)}(t)$, $U^{(\gamma)} : [0, T] \to \mathrm{M}_n$, which satisfies some set of conditions. We want

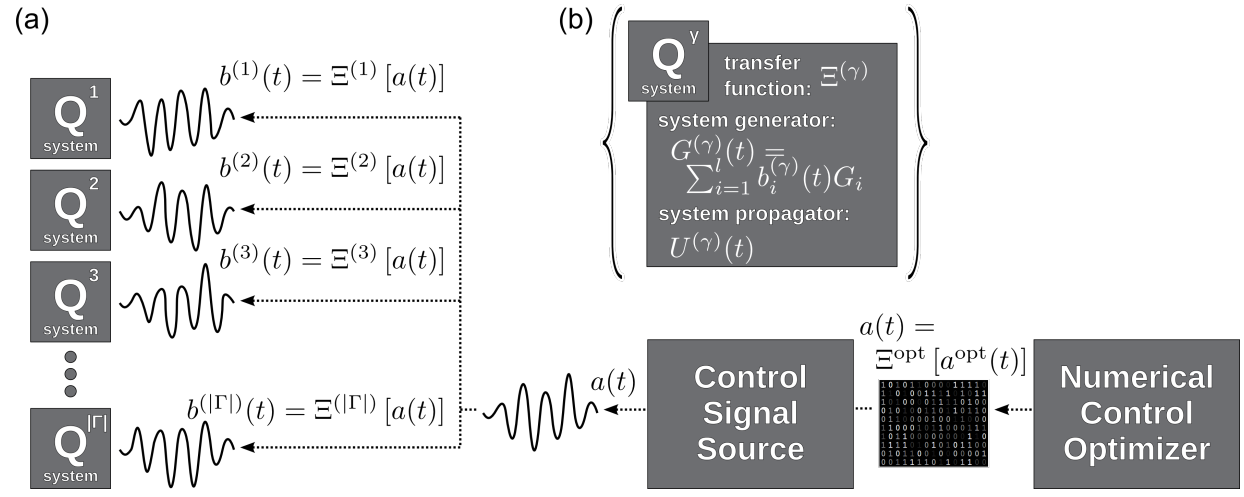

Figure 3.7: (a) Illustration of the control setting considered in this section. We say that we an ensemble Γ of quantum systems, the unique characteristics of each quantum system $\gamma \in \Gamma$ are captured by the transfer function $\Xi^{(\gamma)}$ associated with it. We carry out our numerical control finding searches on the optimization control sequence $a^{\mathrm{opt}}(t)$ that is converted into an experimentally implementable sequence $a(t)$ through the use of the optimization transfer function Ξ^{opt}. Ξ^{opt} is used for imposing the experimentally necessary constraints on $a(t)$, while $a^{\mathrm{opt}}(t)$ need not adhere to such restrictions. When performing experiments, the sequence $a(t)$ is outputted by some control signal source, typically thought of as an AWG. $a(t)$ is transformed by the set of transfer functions $\{\Xi^{(\gamma)}\}$ to a set of control amplitudes $\{b^{(\gamma)}(t) = \Xi^{(\gamma)}[a(t)]\}$ which dictate the evolution of each quantum system. (b) Each quantum system $\gamma \in \Gamma$ is identified by its unique transfer function $\Xi^{(\gamma)}$, whereas the evolution of it is determined by the system propagator $U^{(\gamma)}(t)$ generated by the system generator $G^{(\gamma)}(t) = \sum_{i=1}^{l} b_i^{(\gamma)}(t) G_i$, such that $U^{(\gamma)}(t) = \mathcal{T} \exp\left(\int_0^t dt_1 G^{(\gamma)}(t_1)\right)$.

to emphasize that the properties desired from $\{U^{(\gamma)}(t)\}$ need not be merely its value at time T, these could also be various integral expressions of $U^{(\gamma)}(t)$ over $0 \leq t \leq T$, which describe the trajectory of $U^{(\gamma)}(t)$. Here, a quantum system should be understood simply as a finite level system or one that can be treated as such; the time dependent state of the quantum system is determined by $U^{(\gamma)}(t)$. The time dependent value of $U^{(\gamma)}(t)$ itself is determined by a first order linear matrix differential equation which we refer to as the *system differential equation*:

$$\dot{U}^{(\gamma)}(t) = G^{(\gamma)}(t)U^{(\gamma)}(t), \tag{3.61}$$

where $G_\gamma : [0, T] \to \mathrm{M}_n$ is the *system generator*, while $U^{(\gamma)}(0) = \mathbb{1}_n$. The value of $G^{(\gamma)}(t)$ at each instant is determined by the control amplitudes

$$G^{(\gamma)}(t) = \sum_{i=1}^{l} b_i^{(\gamma)}(t)G_i, \tag{3.62}$$

where $G_i \in \mathrm{M}_n$ is a constant matrix for all i. This implies that the problem is a bilinear control theory problem. The system differential equation should be understood as the Schrödinger equation or some generalization of it, e.g., Liouville equation for vectorized density matrices. For the cases where Equation (3.61) stands for the Schrödinger equation, we have $G_\gamma(t) = -iH(t)$, where $H(t)$ is the system Hamiltonian, and accordingly $U^{(\gamma)}(t)$ is a unitary operator. However, this need not always be the case.

It is important to realize that we are using transfer functions in a somewhat unconventional way. Usually, one associates transfer functions with deterministic distortions of $a(t)$ specific to the experimental setup (electronics in use), whereas we use $\Xi^{(\gamma)}$ to describe these distortions as well as control amplitude variations that enter $G^{(\gamma)}(t)$ due to other factors, e.g., Rabi dispersion. Moreover, in some cases, we use transfer functions to introduce additional auxiliary amplitudes into $b^{(\gamma)}(t)$ which may be totally independent of $a(t)$. The simplest of such auxiliary amplitudes would be the amplitude for the drift term in $G^{(\gamma)}(t)$. By drift we mean time independent terms in the system generator, which are, at first sight, absent from Equation (3.62). Such formalism enables us to describe all inherent system variations (e.g., Rabi distribution, resonance offsets) through $\{\Xi^{(\gamma)}\}$, which lends for concise notation and, perhaps more importantly, reflects the structure of our implementation, while being totally general.

We also note that, here, we have assumed all $G^{(\gamma)}(t)$ to have identical generators for all $\gamma \in \Gamma$. Even though this may not be the case for all quantum control problems, one can always use our problem description by employing a direct sum of different sets of $\{G_i\}$ – similar to how we handled the example in Section 3.4.4. Such approach is computationally

not the most efficient, but it does substantially simplify implementing the algorithm while retaining total generality.

As was said in Section 3.2, Equation (3.61) has a formal solution

$$U^{(\gamma)}(t) = \mathcal{T} \exp\left(\int_0^t dt_1 G^{(\gamma)}(t_1)\right). \tag{3.63}$$

We are typically interested in finding an $a(t)$, $0 \leq t \leq T$, that yields some wanted final unitary operations $\{U^{(\gamma)}(T)\}$ as well as some desired values for nested integral expressions of the following form:

$$\int_0^T dt_1 \cdots \int_0^{t_{n-1}} dt_n \; f(t_1, \ldots, t_n) \; \left[U^{(\gamma)}(t_1)\right]^{-1} A(t_1) U^{(\gamma)}(t_1) \ldots \left[U^{(\gamma)}(t_n)\right]^{-1} A(t_n) U^{(\gamma)}(t_n). \tag{3.64}$$

Such expressions can be thought of as perturbations of $U(T)$, as it was presented in Section 3.2, and they appear in various perturbative treatments of quantum control, e.g., AHT and stochastic Liouville theory reviewed in the previous chapter. So far in this chapter, we have only treated control scenarios, for which $f = 1$, however, with Section 3.7 we extend our formalism to include arbitrary time dependent functions.

The main and the most significant result of this chapter is demonstrating that control problems of this kind can still be written as bilinear control theory problems that involve the same control amplitudes $\{b^{(\gamma)}(t)\}$ that appear in the system differential equation. In order to find $a(t)$ that yields the desired $\{U^{(\gamma)}(T)\}$, and the desired values for any set of integral expressions for $\{U^{(\gamma)}(t)\}$, $U^{(\gamma)} : [0, T] \to \mathrm{M}_n$, we can always construct a block matrix differential equation – called the *Van Loan differential equation* – which comprises the system generators $\{G_i\}$ and the objects that appear in the integral expressions for $\{U^{(\gamma)}(t)\}$. The Van Loan differential equation is expressed as

$$\dot{V}^{(\gamma)}(t) = L^{(\gamma)}(t) V^{(\gamma)}(t), \tag{3.65}$$

where $V^{(\gamma)}(t)$ is the *Van Loan propagator* and $L^{(\gamma)}(t)$ is the *Van Loan generator*

$$L^{(\gamma)}(t) = L_0 + \sum_{i=1}^{l} b_i^{(\gamma)}(t) L_i, \tag{3.66}$$

for $L_i \in \mathrm{M}_m(\mathrm{M}_n)$. In the previous section, we constructed four Van Loan differential equations, the propagators and generators of which can be seen in Equations (3.43), (3.47), (3.51) and (3.58). In all of these equations, the $L(t)$ term appears in the time ordered

exponential on the left-hand side of the equation, and $V(t)$ appears on the right; it is also clear that all of the integral expressions of interest appeared as various blocks of $V(T)$. The benefit of such block matrix methods is two-fold: it enables an accurate and efficient way for evaluating the integral expressions for piece-wise constant $b^{(\gamma)}(t)$ and, as we have demonstrated, is readily deployable within control finding routines that take advantage of the linear differential equation structure of the problem. We should note that the Van Loan generators, and hence, the propagator in Equation (3.58), were not members of $\mathrm{M}_m(\mathrm{M}_4)$ because we simplified the Van Loan differential equation using the fact that the two-spin control was global. This aspect is rather universal – while the Van Loan generators can always be written as $L_i \in \mathrm{M}_m(\mathrm{M}_n)$, where n is the dimension of the system differential equation, the dimension of $\{L_i\}$ can often be reduced from $m \times n$ by exploiting certain symmetries in the problem. Such reduction typically decreases the computational cost for the control search.

Having constructed the Van Loan differential equation that enables the evaluation of all terms of interest, we can always define a target function $\Phi^{(\gamma)}$ for each system in the ensemble. $\Phi^{(\gamma)}$ being a function of the final Van Loan propagator $V^{(\gamma)}(T)$ for system γ, i.e., $\Phi^{(\gamma)} : \mathrm{M}_m(\mathrm{M}_n) \to [0, 1]$, where $\Phi^{(\gamma)} = 1$ corresponds to having the desired properties from the system propagator $U^{(\gamma)}(T)$ and from any number of nested integral terms of interest. The construction of $\Phi^{(\gamma)}$ is done much the same way as it was described at the beginning of Section 3.4. Our examples for target functions expressed as functions of the final Van Loan propagator $V(T)$ are given in Equations (3.44), (3.48), (3.52) and (3.59). Finally, we combine $\{\Phi^{(\gamma)}\}$ into a target function Φ for the whole ensemble Γ: $\Phi = \sum_{\gamma\in\Gamma} p^{(\gamma)}\Phi^{(\gamma)}$, where $\{p^{(\gamma)}\}$ are the relative weights assigned to each member of Γ. We have assumed that $0 \leq p^{(\gamma)} \leq 1$, for all $\gamma \in \Gamma$, and that $\sum_{\gamma\in\Gamma} p^{(\gamma)} = 1$. Of course, the linear form of Φ is not necessary, but it does tremendously simplify the implementation. It is clear that Φ is a functional of $a(t)$ and its derivatives with respect to the control sequence are given as

$$\frac{\partial}{\partial a(t)}\Phi = \sum_{\gamma\in\Gamma} p^{(\gamma)} \frac{\partial}{\partial a(t)}\Phi^{(\gamma)} \left[V^{(\gamma)}(T)\left(\Xi^{(\gamma)}\left[a(t)\right]\right)\right]. \tag{3.67}$$

For all examples in Section 3.4 we regarded the transfer function Ξ as the identity functional, hence, we had $V^{(\gamma)}(T)\left(\Xi^{(\gamma)}\left[a(t)\right]\right) = V^{(\gamma)}(T)\left[a(t)\right]$, which simplified the evaluation of $\frac{\partial}{\partial a(t)}\Phi$, because the derivatives extracted from the Van Loan differential equation were $\frac{\partial}{\partial a(t)}V^{(\gamma)}(T)$. Given a non-identity transfer function Ξ, which is likely the case for any experimental setting, we have to use the chain rule to evaluate $\frac{\partial}{\partial a(t)}V^{(\gamma)}(T)\left(\Xi^{(\gamma)}\left[a(t)\right]\right)$. Since in this work we deal exclusively with piecewise constant control settings, we now define the matrices $\{\beta^{(\gamma)}\}$ that specify the piecewise constant amplitudes of $\{b^{(\gamma)}(t)\}$. Just

like in Section 3.3, we say that we split the interval $[0, T]$ into M subintervals with respective durations δT_j such that $\delta T_j \geq 0$, for all $j \in \{1, 2, \ldots, M\}$, and $\sum_{j=1}^{M} \delta T_j = T$. We note that M does not necessarily have to match N – the number of subintervals used for specifying $a(t)$. We can now specify $\beta^{(\gamma)} \in \mathrm{M}_{l,M}(\mathbb{R})$, the components of which correspond to the piecewise constant values of $b^{(\gamma)}(t)$:

$$b_i^{(\gamma)}\left(\sum_{s=1}^{j-1} \delta T_s \leq t < \sum_{s=1}^{j-1} \delta T_s + \delta T_j\right) = \beta_{i,j}^{(\gamma)}. \tag{3.68}$$

Because we treat the control sequence and the control amplitudes as piecewise constant functions, we will also, from now on, regard the transfer functions $\{\Xi^{(\gamma)}\}$ as matrix valued functions, i.e., $\Xi^{(\gamma)} : \mathrm{M}_{k,N}(\mathbb{R}) \to \mathrm{M}_{l,M}(\mathbb{R})$, such that $\beta^{(\gamma)} = \Xi^{(\gamma)}(\alpha)$. We can now write $\Phi(\alpha) = \sum_{\gamma \in \Gamma} p^{(\gamma)} \Phi^{(\gamma)}\left(V^{(\gamma)}(T)\left[\Xi^{(\gamma)}(\alpha)\right]\right)$. Solving the Van Loan differential equations gives us the values of $\frac{\partial}{\partial \beta_{i,j}^{(\gamma)}} V^{(\gamma)}(T)$, just as it was shown in Section 3.3. Hence, we can evaluate $\frac{\partial}{\partial \alpha_{i,j}} V^{(\gamma)}(T) = \sum_{s=1}^{l} \sum_{t=1}^{M} \frac{\partial}{\partial \alpha_{i,j}} \left(\Xi^{(\gamma)}(\alpha)\right)_{s,t} \frac{\partial}{\partial \beta_{s,t}^{(\gamma)}} V^{(\gamma)}(T)$. As we assume $\Xi^{(\gamma)}$ to be an analytic matrix valued function, we can always evaluate the elements of its Jacobian $\{\frac{\partial}{\partial \alpha_{i,j}} \left(\Xi^{(\gamma)}(\alpha)\right)_{s,t}\}$ in order to implement the maximization of Φ. For all examples considered in this thesis, $\{\Xi^{(\gamma)}\}$ were taken to be linear, meaning that their Jacobians are trivial.

We now tie the formalism introduced in this section together, and give an outline of our procedure for treating the problems described. The problem of engineering control for an ensemble of quantum systems Γ, where each system is specified by its unique transfer function $\Xi^{(\gamma)}$, $\gamma \in \Gamma$, breaks into the following steps.

- First, we need to identify the properties desired from the system propagators $\{U^{(\gamma)}(t)\}$, $U^{(\gamma)} : [0, T] \to \mathrm{M}_n$, and construct suitable Van Loan generators $\{L^{(\gamma)}(t)\}$ for evaluating the terms of interest.

- Secondly, we need to construct a set of $\Phi^{(\gamma)}$, which are analytic functions of $V^{(\gamma)}(T)$ and yield the value one if and only if all the terms of interest in $V^{(\gamma)}(T)$ take their desired values. In practice, this step usually takes some tweaking whenever multiple optimization targets are being considered. Determining the relative weights of the various optimization targets – such that no single one of them dominates the optimization – always takes some trial and error.

- After that, we need to pick the relative weights for the each of the ensemble members: $\{p^{(\gamma)}\}$. Typically, when all members of the ensemble are equally important, it is sufficient to pick $p^{(\gamma)} = 1/|\Gamma|$, for all $\gamma \in \Gamma$.

- Equipped with the above, we need to evaluate $\{\frac{\partial}{\partial\alpha_{i,j}}\Phi\}$, and proceed with the control search as it was described in Section 3.4. We first generate a set of seed control sequences $\{\alpha^{(0)}\}$ and then run a conjugate-gradient maximization algorithm for each seed until a sequence that yields a Φ value sufficiently close to one is found.

Finally, in most practical cases the experimentalist needs $a(t)$ to adhere to some constraints, e.g., pulse waveform bandwidth and amplitude constraints or periods for which $a(t) = 0$. Such constraints are often convenient to implement with the use of an *optimization transfer function* $\Xi^{\rm opt}$, $\Xi^{\rm opt} : \mathrm{M}_{k,N^{\rm opt}}(\mathbb{R}) \to \mathrm{M}_{k,N}(\mathbb{R})$, where $N^{\rm opt}$ is the number of time steps for the piecewise constant *optimization control sequence* $a^{\rm opt}(t)$. $\Xi^{\rm opt}$ is constructed in a way which ensures that all control sequences belonging to its output space adhere to the constraints – that regardless of the input $a^{\rm opt}(t)$. Having constructed $\Xi^{\rm opt}$, the numerical pulse searches are then carried out on $a^{\rm opt}(t)$, which does not need to adhere to the constraints. We represent the piecewise-constant $a^{\rm opt}(t)$ by a matrix $\alpha^{\rm opt} \in \mathrm{M}_{k,N^{\rm opt}}(\mathbb{R})$, just as we did for $a(t)$ and $\{b^{(\gamma)}(t)\}$ above. For finding a suitable control sequence using gradient based algorithms, one then needs to evaluate

$$\frac{\partial}{\partial\alpha^{\rm opt}_{i,j}}\Phi = \sum_{\gamma\in\Gamma} p^{(\gamma)} \frac{\partial}{\partial\alpha^{\rm opt}_{i,j}}\Phi^{(\gamma)} \left[V^{(\gamma)}(T)\left[\Xi^{(\gamma)}\left(\Xi^{\rm opt}\left[\alpha^{\rm opt}\right]\right)\right]\right], \tag{3.69}$$

which means evaluating the elements of the Jacobian for $\{\Xi^{(\gamma)}\}$, as well as for $\Xi^{\rm opt}$. After finding an $\alpha^{\rm opt}$ that yields a high enough Φ value, the control sequence, that is to be implemented experimentally, is calculated simply as $\alpha = \Xi^{\rm opt}(\alpha^{\rm opt})$. In the next section, we demonstrate explicitly how to construct $\Xi^{\rm opt}$ that introduces zero pulse amplitudes to the beginning and the end of the control sequence and how to construct $\Xi^{\rm opt}$ that limits the bandwidth of $a(t)$ waveform in frequency (Fourier) domain.

3.6 Transfer Functions

In this section, we describe the matrix methods that we used when performing control searches for the example in Section 3.8, as well as for the nanoscale magnetic resonance experiments that will be discussed in the next chapter. In both cases, the control sequences $a(t)$, $a : [0, T] \to \mathbb{R}^k$, are piecewise constant over intervals of equal length ΔT such that $T = N\Delta T$. Furthermore, in both cases, the control vectors $a(t)$ are of dimension two, i.e., $k = 2$. Consequently, the real valued control vectors could be mapped onto a complex scalar function $a'(t) : [0, T] \to \mathbb{C}$, with $a'(t) = a_1(t) - ia_2(t) = a_x(t) - ia_y(t)$. Given the

complex vector representation, and the fact that all transfer functions $\{\Xi^{(\gamma)}\}$ in this work are linear functions that treat $a_x(t)$ and $a_y(t)$ symmetrically, since $a_x(t)$ and $a_y(t)$ stand for the in-phase and quadrature components of the control waveform, we could represent each $\Xi^{(\gamma)}$ as an $N \times N$ matrix with complex entries such that

$$\beta_{1,j}^{(\gamma)} = \mathrm{Re}\left[\Xi^{(\gamma)}(\alpha_1 - i\alpha_2)\right]_j, \tag{3.70}$$

and

$$\beta_{2,j}^{(\gamma)} = -\mathrm{Im}\left[\Xi^{(\gamma)}(\alpha_1 - i\alpha_2)\right]_j, \tag{3.71}$$

where $\{\beta_{i,j}^{(\gamma)}\}$ are the piecewise constant control amplitudes as defined in Section 3.5, and α_i denotes the ith row of the matrix $\alpha \in \mathrm{M}_{2,N}(\mathbb{R})$ that specifies the control sequence $a(t)$.

Moreover, for all examples in this thesis $\{\Xi^{(\gamma)}\}$ are diagonal in the Fourier domain, and fully specified by two real valued scalar functions – the amplitude transfer function $\lambda^{(\gamma)}(\nu)$ and the phase transfer function $\phi^{(\gamma)}(\nu)$. In order to construct each $\Xi^{(\gamma)}$, we first calculated the discrete Fourier transform matrix, which is a unitary transformation $U^{\mathrm{Fourier}}(N) \in \mathrm{M}_N$, with its elements given as

$$\left[U^{\mathrm{Fourier}}(N)\right]_{s,t} = \frac{1}{\sqrt{N}} \exp\left[\frac{2\pi i(s-1)(t-1)}{N}\right], \tag{3.72}$$

for $1 \leq s, t \leq N$. We then construct a diagonal matrix $\Lambda^{(\gamma)} \in \mathrm{M}_N$, the diagonal elements of which are given as $\Lambda_{j,j}^{(\gamma)} = \lambda^{(\gamma)}(\nu_j)\ \exp\left[i\phi^{(\gamma)}(\nu_j)\right]$, with

$$\nu_j = \frac{1}{N\Delta T}\left(2\left[(j-1) \bmod \frac{N}{2}\right] - (j-1)\right), \tag{3.73}$$

for $1 \leq j \leq N$. Here, $x \bmod y$ denotes the remainder from diving x by y. Finally, we evaluate

$$\Xi^{(\gamma)}(N, \Delta T, \lambda^{(\gamma)}, \phi^{(\gamma)}) = \left[U^{\mathrm{Fourier}}(N)\right]^{-1} \Lambda^{(\gamma)}(N, \Delta T, \lambda^{(\gamma)}, \phi^{(\gamma)}) U^{\mathrm{Fourier}}(N), \tag{3.74}$$

for each $\gamma \in \Gamma$. The control amplitudes $\{\beta^{(\gamma)}\}$ are then evaluated via Equation (3.70) and (3.71). The elements of the Jacobian $\{\frac{\partial}{\partial \alpha_{i,t}} \beta_{j,s}^{(\gamma)}\}$, that are also necessary for the control

searches, evaluate to

$$\begin{aligned}
\frac{\partial}{\partial \alpha_{1,t}} \beta_{1,s}^{(\gamma)} &= \mathrm{Re}\left(\Xi^{(\gamma)}\right)_{s,t} \\
\frac{\partial}{\partial \alpha_{2,t}} \beta_{1,s}^{(\gamma)} &= \mathrm{Im}\left(\Xi^{(\gamma)}\right)_{s,t} \\
\frac{\partial}{\partial \alpha_{1,t}} \beta_{2,s}^{(\gamma)} &= -\mathrm{Im}\left(\Xi^{(\gamma)}\right)_{s,t} \\
\frac{\partial}{\partial \alpha_{2,t}} \beta_{2,s}^{(\gamma)} &= \mathrm{Re}\left(\Xi^{(\gamma)}\right)_{s,t},
\end{aligned} \tag{3.75}$$

for all $1 \leq s, t \leq N$.

3.6.1 Optimization Transfer Functions

In Section 3.5, we argued that to implement certain constraints on the control sequence $a(t)$, one can use an optimization transfer function Ξ^{opt} such that $\alpha = \Xi^{\mathrm{opt}}[\alpha^{\mathrm{opt}}]$, where α and α^{opt} specify the piecewise constant control sequences $a(t)$ and $a^{\mathrm{opt}}(t)$, respectively. Ξ^{opt} is constructed in such a way that its output functions always adhere to the necessary constraints. In this subsection, we demonstrate explicitly how to construct Ξ^{opt} that introduces periods of zero pulse amplitudes at the beginning and at the end of the control sequence $a(t)$, and limits the bandwidth of $a(t)$ in the Fourier domain. This optimization transfer function was used for control searches in Section 3.8 and in the next chapter.

First, we construct an optimization transfer function that ensures that the control sequence $a(t)$ has equal periods of zero amplitude at the beginning and at the end of the sequence. To introduce such periods, we define an optimization transfer function Ξ^0. Just like above, we map both $\alpha \in \mathrm{M}_{2,N}(\mathbb{R})$ and $\alpha^{\mathrm{opt}} \in \mathrm{M}_{2,N^{\mathrm{opt}}}(\mathbb{R})$ onto complex vectors $\alpha' = \alpha_1 - i\alpha_2$ and $\alpha^{\mathrm{opt}\prime} = \alpha_1^{\mathrm{opt}} - i\alpha_2^{\mathrm{opt}}$, respectively. $N^{\mathrm{opt}} = N - 2N_0$ is the number of piecewise constant elements of $a^{\mathrm{opt}}(t)$, while N_0 is the number of zero amplitude elements that are introduced at the beginning and at the end of $a(t)$. The action of $\Xi^0 \in \mathrm{M}_{N,N^{\mathrm{opt}}}$ is then defined implicitly as

$$\alpha' = \Xi^0 \alpha^{\mathrm{opt}\prime} = \begin{pmatrix} 0^{N_0} \\ \alpha^{\mathrm{opt}\prime} \\ 0^{N_0} \end{pmatrix}, \tag{3.76}$$

for any $\alpha^{\mathrm{opt}\prime}$; 0^{N_0} denotes a zero vector of length N_0. It is easy to see that a Ξ^0, which has the property above, can be constructed from three blocks:

$$\Xi^0(N, N_0) = \begin{pmatrix} 0 \\ \mathbb{1}_{N-2N_0} \\ 0 \end{pmatrix}, \tag{3.77}$$

where $0 \in M_{N_0,N-2N_0}$ is a rectangular matrix with all its entries being zero. The elements of α are then given by

$$\alpha_{1,j} = \mathrm{Re}\left[\Xi^0(N, N_0)(\alpha_1^{\mathrm{opt}} - i\alpha_2^{\mathrm{opt}})\right]_j \tag{3.78}$$

and

$$\alpha_{2,j} = -\mathrm{Im}\left[\Xi^0(N, N_0)(\alpha_1^{\mathrm{opt}} - i\alpha_2^{\mathrm{opt}})\right]_j . \tag{3.79}$$

Of course, the construction of $\Xi^0(N, N_0)$ generalizes easily for introducing an arbitrary number of zero amplitude periods of an arbitrary duration into $a(t)$.

Now, to limit the bandwidth of $a(t)$ frequency components, we construct an optimization transfer function Ξ^{bp} that acts as a low-pass filter. Here, a low-pass filter should be understood simply as some amplitude transfer function $\lambda^{\mathrm{bp}} : \mathbb{R} \to \mathbb{R}$ in Equation (3.74), that takes non-zero values only over some limited range $\Delta\nu$ centred around $\nu = 0$. In this thesis, we used a particular functional form:

$$\lambda^{\mathrm{bp}}(\nu, \Delta\nu) = \frac{1}{4}\left(1 + \tanh\left[\frac{20}{\Delta\nu}\left(\nu + \frac{\Delta\nu}{2}\right)\right]\right)\left(1 - \tanh\left[\frac{20}{\Delta\nu}\left(\nu - \frac{\Delta\nu}{2}\right)\right]\right), \tag{3.80}$$

that has smooth frequency cut-offs at $\pm\Delta\nu/2$ in order to prevent introducing long lasting ripples to the pulse waveform $a(t)$. Accordingly, we implement Ξ^{bp} as

$$\Xi^{\mathrm{bp}}(N, \Delta T, \Delta\nu) = \Xi(N, \Delta T, \lambda^{\mathrm{bp}}, 0), \tag{3.81}$$

where $\Xi(N, \Delta T, \lambda^{\mathrm{bp}}, 0)$ is given by Equation (3.74).

The optimization transfer function Ξ^{opt}, that we used in Section 3.8 and in the next chapter, combined the action of Ξ^0 and Ξ^{bp} and is calculated as

$$\Xi^{\mathrm{opt}}(N, N_0, \Delta T, \Delta\nu) = \Xi^{\mathrm{bp}}(N, \Delta T, \Delta\nu)\Xi^0(N, N_0). \tag{3.82}$$

The elements of α are consequently determined as

$$\begin{aligned} \alpha_{1j} &= \mathrm{Re}\left[\Xi^{\mathrm{opt}}(N, N_0, \Delta T, \Delta\nu)(\alpha_1^{\mathrm{opt}} - i\alpha_2^{\mathrm{opt}})\right]_j \\ \alpha_{2j} &= -\mathrm{Im}\left[\Xi^{\mathrm{opt}}(N, N_0, \Delta T, \Delta\nu)(\alpha_1^{\mathrm{opt}} - i\alpha_2^{\mathrm{opt}})\right]_j , \end{aligned} \tag{3.83}$$

for all $1 \leq j \leq N$, whereas the elements of the Jacobian $\{\frac{\partial}{\partial \alpha_{i,t}^{\mathrm{opt}}} \alpha_{j,s}\}$ evaluate to

$$\begin{aligned}
\frac{\partial}{\partial \alpha_{1,t}^{\mathrm{opt}}} \alpha_{1,s} &= \mathrm{Re}\left[\Xi^{\mathrm{opt}}(N, N_0, \Delta T, \Delta\nu)\right]_{s,t} \\
\frac{\partial}{\partial \alpha_{2,t}^{\mathrm{opt}}} \alpha_{1,s} &= \mathrm{Im}\left[\Xi^{\mathrm{opt}}(N, N_0, \Delta T, \Delta\nu)\right]_{s,t} \\
\frac{\partial}{\partial \alpha_{1,t}^{\mathrm{opt}}} \alpha_{2,s} &= -\mathrm{Im}\left[\Xi^{\mathrm{opt}}(N, N_0, \Delta T, \Delta\nu)\right]_{s,t} \\
\frac{\partial}{\partial \alpha_{2,t}^{\mathrm{opt}}} \alpha_{2,s} &= \mathrm{Re}\left[\Xi^{\mathrm{opt}}(N, N_0, \Delta T, \Delta\nu)\right]_{s,t},
\end{aligned} \tag{3.84}$$

for all $1 \leq s \leq N$ and $1 \leq t \leq N^{\mathrm{opt}}$.

3.7 Generalizations

So far in this chapter, we have demonstrated how to construct Van Loan differential equations that would enable numerical evaluation of nested integrals of the following kind:

$$\begin{aligned}
&\mathcal{D}_U(A_1, \ldots, A_n) = \\
&\quad U(T) \int_0^T dt_1 \cdots \int_0^{t_{n-1}} dt_n \; U^{-1}(t_1) A_1(t_1) U(t_1) \ldots U^{-1}(t_n) A_n(t_n) U(t_n),
\end{aligned} \tag{3.85}$$

where $U(t) = \mathcal{T} \exp\left(\int_0^t dt_1 G(t_1)\right)$ and $A_1, \ldots, A_n : [0, T] \to \mathrm{M}_n$. Nevertheless, as it was shown in the previous chapter, there exist control settings which require the evaluation of the following integral expressions:

$$U(T) \int_0^T dt_1 \cdots \int_0^{t_{n-1}} dt_n \; f(t_1, \ldots, t_n) U^{-1}(t_1) A_1(t_1) U(t_1) \ldots U^{-1}(t_n) A_n(t_n) U(t_n), \tag{3.86}$$

where f is a scalar valued function. This section will be devoted on constructing block matrix methods to address integrals of that kind, which appear frequently when treating stochastic noise operators. In that case, the scalar function f is typically either the noise correlation function or some function of that correlation function, e.g., the four-point correlator. In the next section, we put the generalized numerical tools developed here into use for an example that treats qubit control in the presence of stochastic 1/f noise.

It is clear that integrals of Equation (3.85) kind involve matrix valued integrands that have a particularly simple product structure, i.e., they can all be written as $\widetilde{A}_1(t_1)\ldots\widetilde{A}_n(t_n)$, where $\widetilde{A}_i(t_i) = U^{-1}(t_i)A_1(t_i)U(t_i)$. On the other hand, it is certain that an arbitrary scalar function $f(t_1,\ldots,t_n)$ can not be written in such product form. This makes the extension of our methods, to integrals such as Equation (3.86), not obvious. We approach this problem by first assuming that we can approximate any $f(t_1,\ldots,t_n)$ as a linear combination of functions that can be written in a product form:

$$f(t_1,\ldots,t_n) \approx \sum_i c_i\ g_{i1}(t_1)\ldots g_{in}(t_n), \tag{3.87}$$

with $\{c_i\}$ being constant complex coefficients and $\{g_{i,j}(t)\}$, $g_{i,j} : [0,T] \to \mathbb{C}$, being scalar valued functions of a single argument. Given this approximation, we can now write Equation (3.86) as

$$\begin{aligned} U(T)\int_0^T dt_1\cdots\int_0^{t_{n-1}} dt_n\ f(t_1,\ldots,t_n)U^{-1}(t_1)A_1(t_1)U(t_1)\ldots U^{-1}(t_n)A_n(t_n)U(t_n) \approx \\ \sum_i c_i\int_0^T dt_1\cdots\int_0^{t_{n-1}} dt_n U^{-1}(t_1)\left[g_{i1}(t_1)A_1(t_1)\right]U(t_1)\ldots U^{-1}(t_n)\left[g_{in}(t_n)A_n(t_n)\right]U(t_n). \end{aligned} \tag{3.88}$$

The equation above immediately suggest one way for evaluating the integrals in Equation (3.86): given the set of $\{A_i\}$ and the decomposition of $f(t_1,\ldots,t_n)$, we can define a new set of operators $\{B_{i,j}(t)\}$, where $B_{i,j}(t) = g_{i,j}(t)A_j(t)$, and evaluate each integral in the sum of Equation (3.88) independently. Nevertheless, such approach may not always be the most efficient one in the context of piecewise constant control amplitudes, especially if $\{A_i\}$ happen to be time independent operators; in that case the number of time steps needed to accurately approximate $\{g_{i,j}\}$ can become prohibitively large.

In this section, we consider two specific kinds of $\{g_{i,j}\}$ that can be integrated into Van Loan differential equations in a very natural way – without having to resort to the $B_{i,j}(t) = g_{i,j}(t)A_j(t)$ mapping described above. First, we describe block matrix integral evaluation techniques for the special case of $g_{i,j}(t) = \exp(d_{i,j}t)$, where $d_{i,j}$ is a complex constant. After that, we describe our techniques for the case of $g_{i,j}(t) = t^{s_{i,j}}$, where $s_{i,j} \in \mathbb{N}$ denotes the power of t. We note that the assumption that we can write $f(t_1,\ldots,t_n)$ as a linear combination of $\exp(d_{i,1}t_1)\ldots\exp(d_{i,n}t_n)$ and $t_1^{s_{i,1}}\ldots t_n^{s_{i,n}}$ is not overly restrictive because functions of this form, as a vector space, are dense in the set of all continuous functions. Of course, the last statement ignores the fact that if the number of non-zero c_i needed to approximate $f(t_1,\ldots,t_n)$ becomes too large, our numerical integral evaluation

methods might become infeasible. However, we expect that in most practical cases – with reasonably well behaved $f(t_1, \ldots, t_n)$ and short time intervals $[0, T]$ – the number of non-zero c_i needed will not be too large. In the next section, we provide an example which supports this assumption.

3.7.1 Products of Exponentials[2]

With this subsection we introduce the Van Loan generator $L(t)$, $L : [0, T] \to \mathrm{M}_{n+1}(\mathrm{M}_m)$, that can be used for evaluating the following nested integral expressions

$$\mathcal{D}_U\left(e^{d_1 t}A_1, \ldots, e^{d_n t}A_n\right) = \tag{3.89}$$
$$U(T)\int_0^T dt_1 \cdots \int_0^{t_{n-1}} dt_n\; U^{-1}(t_1)\left[e^{d_1 t_1}A_1(t_1)\right]U(t_1)\ldots U^{-1}(t_n)\left[e^{d_n t_n}A_n(t_n)\right]U(t_n),$$

where $U(t) = \mathcal{T}\exp\left(\int_0^t dt_1 G(t_1)\right)$ is the system propagator generated by $G(t)$, $G : [0, T] \to \mathrm{M}_m$, and $A_1, \ldots, A_n : [0, T] \to \mathrm{M}_m$. The block matrix expression for $L(t)$ is

$$L(t) = \tag{3.90}$$
$$\begin{pmatrix} G(t) & A_1(t) & 0 & \ldots & 0 & 0 \\ 0 & G(t)+d_1\mathbb{1}_m & A_2(t) & \ldots & 0 & 0 \\ 0 & 0 & G(t)+(d_1+d_2)\mathbb{1}_m & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & G(t)+\sum_{i=1}^{n-1} d_i\mathbb{1}_m & A_n(t) \\ 0 & 0 & 0 & \ldots & 0 & G(t)+\sum_{i=1}^{n} d_i\mathbb{1}_m \end{pmatrix},$$

and the corresponding Van Loan propagator can be written as

$$V(t) = \tag{3.91}$$
$$\begin{pmatrix} U(t) & \mathcal{D}_U\left(e^{d_1 t}A_1\right) & \mathcal{D}_U\left(e^{d_1 t}A_1, e^{d_2 t}A_2\right) & \ldots & \mathcal{D}_U\left(e^{d_1 t}A_1, \ldots, e^{d_n t}A_n\right) \\ 0 & e^{d_1 t}\,U(t) & e^{d_1 t}\mathcal{D}_U(e^{d_2 t}A_2) & \ldots & e^{d_1 t}\mathcal{D}_U\left(e^{d_2 t}A_2, \ldots, e^{d_n t}A_n\right) \\ 0 & 0 & e^{(d_1+d_2)t}\,U(t) & \ldots & e^{(d_1+d_2)t}\mathcal{D}_U\left(e^{d_3 t}A_3, \ldots, e^{d_n t}A_n\right) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & e^{\sum_{i=1}^{n-1} d_i t}\mathcal{D}_U\left(e^{d_n t}A_n\right) \\ 0 & 0 & 0 & \ldots & e^{\sum_{i=1}^{n} d_i t}\,U(t) \end{pmatrix},$$

[2]The matrix identity in this subsection was shown to the author by Daniel Puzzuoli.

such that $\mathcal{D}_U\left(e^{d_1 t}A_1, \ldots, e^{d_n t}A_2\right) = V_{1,n+1}(T)$. Equation (3.91) can be checked by employing Theorem 1 given in Section 3.2. First, we notice that its diagonal elements $[V(t)]_{k,k}$, for $1 \leq k \leq n+1$, can be evaluated as

$$V_{k,k}(t) = \mathcal{T} \exp\left(\int_0^t dt_1 \left[G(t_1) + \sum_{i=1}^{k-1} d_i \mathbb{1}_m\right]\right) = e^{\sum_{i=1}^{k-1} d_i t} U(t), \tag{3.92}$$

because $\mathbb{1}_m$ commutes with $G(t)$ at all times. Next, we use the recursive expression in Theorem 1 noticing that for all $1 \leq k \leq n$: $L_{k,k+1}(t) = A_k(t)$; and $L_{k,k+i}(t) = 0$, for all $i > 1$, hence,

$$V_{k,k+j}(t) = e^{\sum_{i=1}^{k-1} d_i t} U(t) \int_0^t dt_1 e^{-\sum_{i=1}^{k-1} d_i t_1} U^{-1}(t_1) A_k(t_1) V_{k+1,k+j}(t_1). \tag{3.93}$$

The equation above enables us to write

$$\begin{aligned} &V_{k,k+j}(t) = e^{\sum_{i=1}^{k-1} d_i t}\ U(t) \\ &\cdot \int_0^t dt_1 \cdots \int_0^{t_{j-1}} dt_j\ e^{d_k t_1}\ U^{-1}(t_1) A_k(t_1) U(t_1) \ldots e^{d_{k+j-1} t_j}\ U^{-1}(t_1) A_{k+j-1}(t_2) U(t_2), \end{aligned} \tag{3.94}$$

which verifies Equation (3.91). In the next section, we make use of $L(t)$ in Equation (3.90) for $n = 2$.

3.7.2 Products of Monomials

With this subsection we give a procedure for evaluating nested integral expressions of the following kind:

$$\begin{aligned} &\mathcal{D}_U\left(t^{s_1}A_1, \ldots, t^{s_n}A_n\right) = \\ &\quad U(T) \int_0^T dt_1 \cdots \int_0^{t_{n-1}} dt_n\ U^{-1}(t_1)\left[t_1^{s_1} A_1(t_1)\right] U(t_1) \ldots U^{-1}(t_n)\left[t_n^{s_n} A_n(t_n)\right] U(t_n), \end{aligned} \tag{3.95}$$

where $s_1, \ldots, s_n \in \mathbb{N}$, while $U(t) = \mathcal{T} \exp\left(\int_0^t dt_1 G(t_1)\right)$ is the system propagator generated by $G(t)$, $G : [0,T] \to \mathrm{M}_m$, and $A_1, \ldots, A_n : [0,T] \to \mathrm{M}_m$. Unlike the exponential case considered above, the monomial case requires the evaluation of $\prod_{i=1}^{n-1}(s_i + 1)$ Van Loan propagators. Here, we will first describe how to construct the necessary Van Loan propagators, and then outline a procedure for evaluating $\mathcal{D}_U\left(t^{s_1}A_1, \ldots, t^{s_n}A_n\right)$.

Let $z_1, \ldots, z_n \in \mathbb{N}$, for which we denote $Z_n = \sum_i^n z_i$. We also say that

$$\widetilde{A}_i = U^{-1}(t) A_i(t) U(t). \tag{3.96}$$

The main result in this subsection relies on a simple observation that

$$\begin{aligned}\int_0^T dt_1 \cdots \int_0^{t_{n-1+Z_n}} dt_{n+Z_n}\ \widetilde{A}_1(t_1)\widetilde{A}_2(t_{2+z_1})\widetilde{A}_3(t_{3+z_1+z_2})\ldots\widetilde{A}_n(t_{n+Z_{n-1}}) = \\ \int_0^T dt_1 \cdots \int_0^{t_{n-1}} dt_n\ \frac{(t-t_1)^{z_1}}{z_1!}\widetilde{A}_1(t_1)\frac{(t_1-t_2)^{z_2}}{z_2!}\widetilde{A}_2(t_2)\frac{(t_2-t_3)^{z_3}}{z_3!}\widetilde{A}_3(t_3)\ldots\frac{t_n^{z_n}}{z_n!}\widetilde{A}_n(t_n),\end{aligned} \tag{3.97}$$

where we have used the identities $\int_0^t dt_1 \cdots \int_0^{t_{z-1}} dt_z = \frac{t^z}{z!}$ and $\int_0^t dt_1 \cdots \int_0^{t_z} dt_{1+z} F(t_{1+z}) = \int_0^t dt_1 \frac{(t-t_1)^z}{z!} F(t_1)$, for any function $F : [0,T] \to \mathrm{M}_m$; the latter identity can be obtained via sequential integration by parts.

To construct a Van Loan generator $L(t)$ for evaluating the expression in Equation (3.97), we first identify that

$$\begin{aligned}U(T)\int_0^T dt_1 \cdots \int_0^{t_{n-1+Z_n}} dt_{n+Z_n}\ \widetilde{A}_1(t_1)\widetilde{A}_2(t_{2+z_1})\widetilde{A}_3(t_{3+z_1+z_2})\ldots\widetilde{A}_n(t_{n+Z_{n-1}}) = \\ \mathcal{D}_U\left(A_1, \underbrace{\mathbb{1}_m, ..., \mathbb{1}_m}_{z_1 \text{ copies}},\ A_2, \underbrace{\mathbb{1}_m, ..., \mathbb{1}_m}_{z_2 \text{ copies}}, \ldots, A_n, \underbrace{\mathbb{1}_m, ..., \mathbb{1}_m}_{z_n \text{ copies}}\right),\end{aligned} \tag{3.98}$$

since $U^{-1}(t)\mathbb{1}_m U(t) = \mathbb{1}_m$, for any $0 \leq t \leq T$. The $L(t) \in \mathrm{M}_{n+Z_n+1}(\mathrm{M}_m)$ necessary for evaluating the expression above has all of its diagonal elements $L_{k,k}(t)$, $1 \leq k \leq n+Z_n+1$, equal to $G(t)$, and the first off diagonal given as

$$\left(A_1, \underbrace{\mathbb{1}_m, ..., \mathbb{1}_m}_{z_1 \text{ copies}},\ A_2, \underbrace{\mathbb{1}_m, ..., \mathbb{1}_m}_{z_2 \text{ copies}}, \ldots, A_n, \underbrace{\mathbb{1}_m, ..., \mathbb{1}_m}_{z_n \text{ copies}}\right), \tag{3.99}$$

so that

$$\begin{aligned}U(T)\int_0^T dt_1 \cdots \int_0^{t_{n-1}} dt_n\ \frac{(t-t_1)^{z_1}}{z_1!}\widetilde{A}_1(t_1)\frac{(t_1-t_2)^{z_2}}{z_2!}\widetilde{A}_2(t_2)\frac{(t_2-t_3)^{z_3}}{z_3!}\widetilde{A}_3(t_3)\ldots\frac{t_n^{z_n}}{z_n!}\widetilde{A}_n(t_n) \\ = V_{1,n+Z_n+1}(T) = \left[\mathcal{T}\exp\left(\int_0^t dt_1 L(t_1)\right)\right]_{1,n+Z_n+1}.\end{aligned} \tag{3.100}$$

The last fact can easily be checked by employing Theorem 1 in Section 3.2 in the same way as it was done in the previous subsection.

Finally, we point out that the term of interest $\mathcal{D}_U\left(t^{s_1}A_1, \ldots, t^{s_n}A_n\right)$ appears as one of the terms in Equation (3.97), after one expands the binomials and sets $z_i = s_i$, for all i. Nevertheless, to isolate $\mathcal{D}_U\left(t^{s_1}A_1, \ldots, t^{s_n}A_n\right)$, one has to evaluate Equation (3.97) for all $0 \leq z_1 \leq s_1, \ldots, 0 \leq z_{n-1} \leq s_{n-1}$, which yields a system of $\prod_{i=1}^{n-1}(s_i+1)$ linear equations. Therefore, we need to generate $\prod_{i=1}^{n-1}(s_i+1)$ Van Loan propagators, as described above, in order to evaluate $\mathcal{D}_U\left(t^{s_1}A_1, \ldots, t^{s_n}A_n\right)$. This set of Van Loan generators $\{L_i(t)\}$, $1 \leq i \leq \prod_{i=1}^{n-1}(s_i+1)$, can be combined into single a direct sum generator $L(t) = L_i(t) \oplus \cdots \oplus L_{\prod_{i=1}^{n-1}(s_i+1)}(t)$ the same way as we did in Section 3.4.4. At first sight, this might seem hugely inefficient, however, it should be noted that solving such a system of linear equations yields the values for all $\mathcal{D}_U\left(t^{z_1}A_1, \ldots, t^{z_n}A_n\right)$, $0 \leq z_1 \leq s_1, \ldots, 0 \leq z_{n-1} \leq s_{n-1}$. And because the goal of evaluating such Dyson terms is generally to approximate some $f(t_1, \ldots, t_n)$ as a linear combination $\sum_i c_i\, t_1^{s_{i1}} \ldots t_n^{s_{in}}$, then the fact that our protocol could provide all monomials up to $t_1^{s_{i1}} \ldots t_n^{s_{in}}$ is not an unreasonable overhead.

3.8 1/f Noise Decoupling

In this section, we turn to an experimentally realistic control example. We will demand that the pulse ends go smoothly to zero, and that the spectral width of the pulse waveform be limited. The example we consider demonstrates our ability to engineer control sequences that are designed to decouple stochastic noise characterized by its spectral density function. For that, we employ the tools for that were developed in the previous section for evaluating generalized nested integrals of Equation (3.86) kind. We pick 1/f noise due to its ubiquity in solid state devices [12]. We proceed by evaluating the first non-zero term in the perturbative cumulant expansion for the Liouville equation, which was described in the previous chapter. We then demonstrate how such a toggling frame term can be approximated, and consequently minimized, using Van Loan differential equations.

In this section, unlike for the examples in Section 3.4, $G(t)$ will denote the Liouville generator rather than the system Hamiltonian. Accordingly, we start with a single spin Liouville generator which includes a stochastic noise term $G_{\mathrm{n}}(t) = \varepsilon(t)G_z$ and dictates the evolution of the system

$$G(t) + G_{\mathrm{n}}(t) = a_x(t)G_x + a_y(t)G_y + \varepsilon(t)G_z, \tag{3.101}$$

where $G_i = -i\left(\frac{\sigma_i}{2} \otimes \mathbb{1} - \mathbb{1} \otimes \frac{\sigma_i^{\mathrm{T}}}{2}\right)$, for $i \in \{x, y, z\}$, and $\varepsilon(t)$ is a stationary, zero mean, Gaussian stochastic function capturing the fluctuations in the spin level spacing. This

implies that $\langle \varepsilon(t) \rangle = 0$, where the angle brackets denote an ensemble average over noise realizations. Here, we take the power spectral density of $\varepsilon(t)$ to be given by $P(\nu) = \frac{2}{\pi\nu}\left[\arctan\left(\frac{\nu}{\Lambda_1}\right) - \arctan\left(\frac{\nu}{\Lambda_2}\right)\right]$, where Λ_1 and Λ_2 are smooth low and high frequency cutoffs for $P(\nu)$, respectively. It is easy to show that $\lim_{\Lambda_1 \to 0, \Lambda_2 \to \infty} P(\nu) = \frac{1}{|\nu|}$. For this example, we use $\Lambda_1 = 2\pi$ Hz and $\Lambda_2 = 2\pi \cdot 10^{10}$ Hz. According to the Wiener-Khinchin theorem:

$$\langle \delta\varepsilon(t_1)\varepsilon(t_2) \rangle = \int_{-\infty}^{\infty} d\nu \; P(\nu) e^{i\nu|t_1 - t_2|} = -2\left[\mathrm{Ei}\left(-\Lambda_1 |t_1 - t_2|\right) - \mathrm{Ei}\left(-\Lambda_2 |t_1 - t_2|\right)\right], \tag{3.102}$$

where $\mathrm{Ei}(z) = -\int_{-z}^{\infty} dt \; e^{-t}/t$ stands for the exponential integral function.

Using stochastic Liouville theory, we can treat the noise perturbatively and show that the toggling frame propagator $U_{\mathrm{tog}}(T)$ is given as

$$\begin{aligned} U_{\mathrm{tog}}(T) &= \left\langle \mathcal{T} \exp\left[\int_0^T dt_1 \; U^{-1}(t_1) G_{\mathrm{n}}(t_1) U(t_1)\right]\right\rangle \\ &= \mathbb{1}_4 + \int_0^T dt_1 \int_0^{t_1} dt_2 \; \langle \varepsilon(t_1)\varepsilon(t_2) \rangle \, U^{-1}(t_1) G_z U(t_1) U^{-1}(t_2) G_z U(t_2) + ..., \end{aligned} \tag{3.103}$$

where $U(t) = \mathcal{T} \exp\left(\int_0^t dt_1 \; G(t_1)\right)$ and we have made use of the fact that

$$\int_0^t dt_1 \; \langle \varepsilon(t_1) \rangle \, U^{-1}(t_1) G_z U(t_1) = 0. \tag{3.104}$$

We remark that $U_{\mathrm{tog}}(T)$ is not a unitary matrix, in fact, it is precisely the operator that encapsulates the non-unitary decoherence effects induced by $G_{\mathrm{n}}(t)$. In order to reduce such decoherence at its lowest perturbative order, we need to minimize the nested double integral term in Equation (3.103). Because the generator $G(t)$ is a Liouville generator, it also holds that $U(t) = \mathcal{T} \exp\left(\int_0^t dt_1 \; G(t_1)\right) = U_1(t) \otimes \overline{U_1(t)}$, where the bar denotes an entry-wise conjugation of the matrix and $U_1(t) = \mathcal{T} \exp\left(-i \int_0^t dt_1 \left[a_x(t_1)\frac{\sigma_x}{2} + a_y(t_1)\frac{\sigma_y}{2}\right]\right)$.

To employ the method developed in the previous section, we first note that a linear combination of exponential functions provides a good approximation for $\langle \varepsilon(t_1)\varepsilon(t_2) \rangle$ over a region of integration $0 \leq t_2 \leq t_1 \leq T$, with $T = 400$ ns. Using a least squares fit to the correlation function over that region, we find an approximation that combines seven

exponential functions

$$\langle \epsilon(t_1)\varepsilon(t_2)\rangle = -2\left[\mathrm{Ei}\left(-\Lambda_1\left|t_1-t_2\right|\right)-\mathrm{Ei}\left(-\Lambda_2\left|t_1-t_2\right|\right)\right] \approx \sum_{i=1}^{7} c_i\ e^{d_i(t_1-t_2)}, \tag{3.105}$$

where $c_1 = 7.49448$, $d_1 = -1.11796\cdot 10^8$ Hz, $c_2 = 0.947027$, $d_2 = -3.37122\cdot 10^7$ Hz, $c_3 = -0.490555$, $d_3 = -4.69721\cdot 10^6$ Hz, $c_4 = -0.163987$, $d_4 = -3.77087\cdot 10^6$ Hz, $c_5 = 29.83$, $d_5 = -577865$ Hz, $c_6 = -0.102058$, $d_6 = 122339$ Hz, $c_7 = 0.00035238$ and $d_7 = 2.05605\cdot 10^7$ Hz. Given the approximation, we can now write

$$\begin{aligned} I_{1/f}(t) &= \int_0^t dt_1 \int_0^{t_1} dt_2\ \langle \varepsilon(t_1)\varepsilon(t_2)\rangle\, U^{-1}(t_1)G_zU(t_1)U^{-1}(t_2)G_zU(t_2) \\ &\approx \sum_{i=1}^{7} c_i \int_0^t dt_1 \int_0^{t_1} dt_2\ e^{d_i(t_1-t_2)}U^{-1}(t_1)G_zU(t_1)U^{-1}(t_2)G_zU(t_2) \\ &= \sum_{i=1}^{7} c_i U^{-1}(T)\mathcal{D}_U\left(e^{d_it}G_z, e^{-d_it}G_z\right). \end{aligned} \tag{3.106}$$

It is important to realize that in the case of actual experimental scenarios either the noise correlation function or its power spectral density would be characterized before embarking on control engineering. In such cases, $\langle \varepsilon(t_1)\varepsilon(t_2)\rangle$ is extremely unlikely to fit to some simple and specific analytic function. Therefore, our procedure, for fitting a set of functions to a set of data – in this case an analytic function – in order to approximate the noise correlation, closely matches a real control engineering protocol.

We search for a pulse implementing a Y gate, i.e., we wish to set $U(T) = e^{-i\pi\sigma_y/2}\otimes \overline{e^{-i\pi\sigma_y/2}}$. Consequently, we use the following target function:

$$\Phi = \frac{4}{5}\left(1-\frac{\|U(T)I_{1/f}(t)\|^2}{\left[\max_{a(t)}\|U(T)I_{1/f}(t)\|\right]^2}\right) + \frac{1}{5}\mathcal{F}\left[e^{-i\pi\sigma_y/2}\otimes\overline{e^{-i\pi\sigma_y/2}}, U(T)\right]. \tag{3.107}$$

Combining Equation (3.106) with Equation (3.90) in the previous section, we set up a Van

Loan differential equation for $V(t) \in \mathrm{M}_{21}(\mathrm{M}_4)$:

$$V(t) = \tag{3.108}$$

$$\mathcal{T}\exp\left[\int_0^t dt_1 \begin{pmatrix} G(t_1) & G_z & 0 & \dots & 0 & 0 & 0 \\ 0 & G(t_1)+d_1\mathbb{1}_4 & G_z & \dots & 0 & 0 & 0 \\ 0 & 0 & G(t_1) & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & G(t_1) & G_z & 0 \\ 0 & 0 & 0 & \dots & 0 & G(t_1)+d_7\mathbb{1}_4 & G_z \\ 0 & 0 & 0 & \dots & 0 & 0 & G(t_1) \end{pmatrix}\right] =$$

$$\begin{pmatrix} U(t) & \mathcal{D}_U\left(e^{d_1 t}G_z\right) & \mathcal{D}_U\left(e^{d_1 t}G_z, e^{-d_1 t}G_z\right) & \dots & 0 & 0 \\ 0 & e^{d_1 t}U(t) & e^{d_1 t}\mathcal{D}_U\left(e^{-d_1 t}G_z\right) & \dots & 0 & 0 \\ 0 & 0 & U(t) & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \mathcal{D}_U\left(e^{d_7 t}G_z\right) & \mathcal{D}_U\left(e^{d_7 t}G_z, e^{-d_7 t}G_z\right) \\ 0 & 0 & 0 & \dots & e^{d_7 t}U(t) & e^{d_7 t}\mathcal{D}_U\left(e^{-d_7 t}G_z\right) \\ 0 & 0 & 0 & \dots & 0 & U(t) \end{pmatrix}.$$

We can now approximate the target as a function of $V(T)$:

$$\Phi = \frac{4}{5}\left(1 - \frac{\mathrm{Tr}\left(\left[\sum_{i=1}^7 c_i V_{3(i-1)+1,3i}(T)\right]^\dagger \sum_{i=1}^7 c_i V_{3(i-1)+1,3i}(T)\right)}{2\left(\int_0^T dt_1 \int_0^{t_1} dt_2 \ \langle\varepsilon(t_1)\varepsilon(t_2)\rangle\right)^2}\right) + \frac{1}{5}\frac{\mathrm{Tr}\left[(e^{-i\pi\sigma_y/2})^\dagger \otimes (e^{-i\pi\sigma_y/2})^{\mathrm{T}} V_{1,1}(T)\right]}{4}, \tag{3.109}$$

where we have used the fact that $\sqrt{\mathrm{Tr}\left[V_{1,1}^\dagger(T)V_{1,1}(T)\right]\mathrm{Tr}\left[\mathbb{1}_4\right]} = 4$ and that $Tr[G_z^\dagger G_z] = 2$, while $\int_0^T dt_1 \int_0^{t_1} dt_2 \ \langle\varepsilon(t_1)\varepsilon(t_2)\rangle$ is evaluated numerically for any particular T. The partial derivatives of Φ with respect to the control amplitudes $\{\beta_{i,j}\}$ are given as

$$\frac{\partial}{\partial\beta_{i,j}}\Phi = -\frac{4}{5}\frac{\mathrm{Re}\left[\mathrm{Tr}\left(\left[\sum_{i=1}^7 c_i V_{3(i-1)+1,3i}(T)\right]^\dagger \sum_{i=1}^7 c_i \left(\frac{\partial}{\partial\beta_{i,j}}V(T)\right)_{3(i-1)+1,3i}\right)\right]}{\left(\int_0^T dt_1 \int_0^{t_1} dt_2 \ \langle\varepsilon(t_1)\varepsilon(t_2)\rangle\right)^2} + \frac{1}{10}\mathrm{Re}\left(\mathrm{Tr}\left[(e^{-i\pi\sigma_y/2})^\dagger \otimes (e^{-i\pi\sigma_y/2})^{\mathrm{T}}\left(\frac{\partial}{\partial\beta_{i,j}}V(T)\right)_{1,1}\right]\right). \tag{3.110}$$

Here, we attempt to closely mimic a control search procedure that would be undertaken when searching for an experimentally implementable sequence. Hence, we impose three distinct constraint on the pulse waveform: maximum amplitude constraint, bandwidth limitations for the pulse waveform frequency components and zero amplitude periods at the beginning and at the end of the sequence. The last two constraints are implemented by introducing an optimization transfer function Ξ^{opt}, as was described in Section 3.5, and the explicit construction of Ξ^{opt} was given in Section 3.6. Here, we do not consider any ensemble effects, i.e., $|\Gamma| = 1$, and we take the only experimental transfer function to act as the identity, such that $\beta = \beta^{(1)} = \Xi^{(1)}(\alpha) = \alpha$. Accordingly, the numerical control searches are conducted for α^{opt}, with $\alpha = \Xi^{\text{opt}}(\alpha^{\text{opt}})$.

For the searches, we limit the Rabi frequency $|a^{\text{opt}}(t)|/(2\pi)$ to be less than or equal to 200 MHz by enforcing that

$$-\frac{1}{\sqrt{2}}2\pi \cdot 200 \cdot 10^6 \text{ Hz} \leq a_i^{\text{opt}}(t) \leq \frac{1}{\sqrt{2}}2\pi \cdot 200 \cdot 10^6 \text{ Hz} \tag{3.111}$$

for $i = \{x, y\}$. We take the pulse length T to be 10 Rabi cycles or 50 ns, which is divided into $N = 300$ intervals of equal length $\Delta T = 1.67 \cdot 10^{-10}$ s, whereas the zero amplitude periods at the beginning and at the end of the pulse $a(t)$ have a length of 8.33 ns, corresponding to $N_0 = 50$. Therefore, the numerical control search is conducted on 33.33 ns-long $a^{\text{opt}}(t)$ that is divided into $N^{\text{opt}} = N - 2N_0 = 200$ equal steps. Using Ξ^{opt}, we constrain all spectral components of the pulse $a(t)$ to lie within a $\Delta\nu = 400$ MHz bandwidth around the carrier frequency. The optimization transfer function $\Xi^{\text{opt}}(N, N_0, \Delta T, \Delta\nu)$, that is employed to enforce the constraints, is defined by Equation (3.82) in Section 3.6.

Because our control optimization was carried out on the optimization control sequence $a^{\text{opt}}(t)$, we needed to evaluate $\{\frac{\partial}{\partial\alpha_{i,j}^{\text{opt}}}\Phi\}$, for all $i = \{1, 2\}$ and $j = \{1, \ldots, N^{\text{opt}}\}$, in order perform our conjugate gradient minimization procedure. With Equation (3.110) we have evaluated $\{\frac{\partial}{\partial\beta_{i,j}}\Phi\}$; we have also identified that $\frac{\partial}{\partial\alpha_{i,j}}\Phi = \frac{\partial}{\partial\beta_{i,j}}\Phi$, for all $i = \{1, 2\}$ and $j = \{1, \ldots, N\}$. Finally, we link $\{\frac{\partial}{\partial\alpha_{i,j}^{\text{opt}}}\Phi\}$ with $\{\frac{\partial}{\partial\beta_{i,j}}\Phi\}$ through Equation (3.84) in Section 3.6:

$$\frac{\partial}{\partial\alpha_{1,j}^{\text{opt}}}\Phi = \tag{3.112}$$
$$\sum_{t=1}^{N}\left(\text{Re}\left[\Xi^{\text{opt}}(N, N_0, \Delta T, \Delta\nu)\right]_{t,j}\frac{\partial}{\partial\beta_{1,t}}\Phi - \text{Im}\left[\Xi^{\text{opt}}(N, N_0, \Delta T, \Delta\nu)\right]_{t,j}\frac{\partial}{\partial\beta_{2,t}}\Phi\right)$$

and

$$\frac{\partial}{\partial \alpha_{2,j}^{\mathrm{opt}}}\Phi = \tag{3.113}$$
$$\sum_{t=1}^{N}\left(\mathrm{Im}\left[\Xi^{\mathrm{opt}}(N,N_0,\Delta T,\Delta\nu)\right]_{t,j}\frac{\partial}{\partial\beta_{1,t}}\Phi + \mathrm{Re}\left[\Xi^{\mathrm{opt}}(N,N_0,\Delta T,\Delta\nu)\right]_{t,j}\frac{\partial}{\partial\beta_{2,t}}\Phi\right).$$

Given the target function and the partial derivatives above, we search for a control sequence just like for the examples in Section 3.4. The resulting waveform $a(t)$ is shown in Figure 3.8 and the pulse characteristics are given as:

$\left\|I_{1/f}(T)\right\| \Big/ \sqrt{2\left(\int_0^T dt_1 \int_0^{t_1} dt_2 \ \langle\varepsilon(t_1)\varepsilon(t_2)\rangle\right)}$	0.0127
$1-\mathcal{F}\left(\sigma_y, U(T)\right)$	1.25×10^{-7}

We point out that, while 8.33 ns-long zero amplitude periods were introduced to the beginning and the end of $a(t)$ with the use of Ξ^{opt}, the waveform in Figure 3.8 does not immediately display such character. That is because the low-pass filter, also incorporated into Ξ^{opt}, creates ripples at the beginning and at the end of $a(t)$. Nevertheless, by adjusting $N_0\Delta T$, and the functional shape of the low-pass filter, one can always ensure that the waveform dies off to any desired degree at its ends.

Finally, in the case of a stochastic operator $G_{\mathrm{n}}(t)$, one cannot expect to be able to set the integral $I_{1/f}(t)$ in Equation (3.106) equal to zero, since the high frequency components of the noise always retain their decoherence inducing effect. This last fact can easily be verified by considering the noise correlation function for uncorrelated white noise: $\langle\varepsilon(t_1)\varepsilon(t_2)\rangle = \delta(t_1-t_2)$, where $\delta(t_1-t_2)$ stands for the Dirac delta function. In that case, it is obvious that the perturbative terms in $I_{1/f}(t)$ are completely independent of the control sequence. Nevertheless, for a reasonably low amplitude noise, our sequence in Figure 3.8 would extend the spin coherence time by a factor of $1/0.0127 \approx 80$.

3.9 Conclusions

With this chapter, we have developed an extremely versatile Van Loan block matrix differential equation framework for efficient and accurate numerical evaluation of any nested

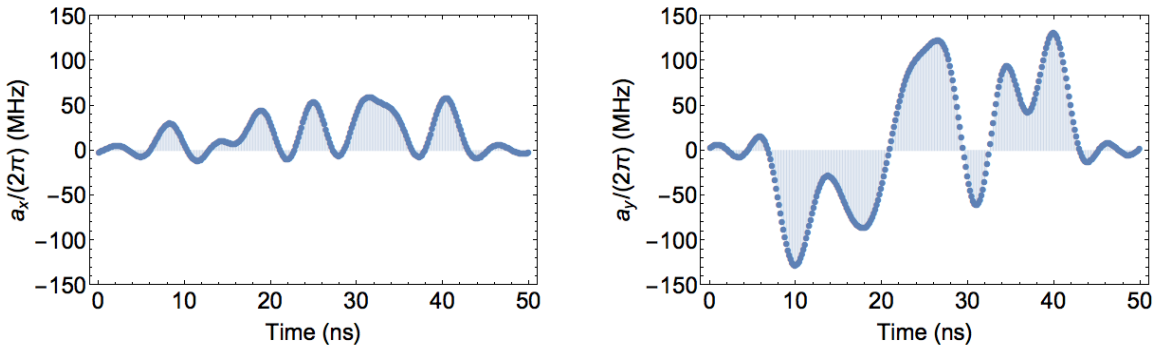

Figure 3.8: Control sequence robust to 1/f noise that implements a Y gate: $a_x(t)$ on the left and $a_y(t)$ on the right. The control waveform was optimized to take zero values at its beginning and its end, furthermore, the limited range of frequency components within the waveform ensures its smoothness.

matrix integral expression of the following kind:

$$U(T)\int_0^T dt_1 \cdots \int_0^{t_{n-1}} dt_n \; f(t_1,\ldots,t_n)U^{-1}(t_1)A_1(t_1)U(t_1)\ldots U^{-1}(t_n)A_n(t_n)U(t_n), \tag{3.114}$$

where the system propagator $U(t)$, $0 \le t \le T$, is generated by piecewise constant control sequences $a(t)$. We have shown that Van Loan differential equation formalism is particularly suitable and easy to incorporate into gradient based control finding schemes that seek to set $U(T)$ and various integrals of the above kind to certain desired values. We have demonstrated the latter property of our formalism by conducting four simple, yet non-trivial, control searches in Section 3.4. For many of these examples our numerical searches converged to control sequences that bore definite resemblances to known analytic solutions. This is a very reassuring starting point for deploying our block matrix methods in experimental control situations, where the constraints on the control sequences, experimental distortions and ensemble effects do not allow for analytic control engineering. In order to mimic such a control engineering situation as well as to demonstrate our ability to handle noise resulting from stochastic processes, we generated the example in Section 3.8 that has a bandwidth limited waveform $a(t)$ with zero amplitude ends.

The fact that our formalism provides a natural way for evaluating partial derivatives of the integral expressions with respect to the control amplitudes is critical for ensuring fast

convergence of our control search implementation. The control sequences presented in this chapter typically converged to the solutions presented within a few thousand target function and partial derivative evaluations. Gradient information becomes especially crucial at the stage of the optimization when the target function Φ – and hence the values for the integral expressions and the final propagator $U(T)$ – approach their desired values, which in the case of Φ is one of its local maxima. Consequently, we expect our methods to provide faster convergence of Φ to its local maximum than any method for control engineering that does not utilize gradient information of Φ. Nevertheless, in the case of applications that require high fidelity operations, optimal or near optimal Φ values are absolutely imperative.

The examples presented in this chapter should not be understood as demonstrative of the full capacity of our framework. In order to keep our treatment illustrative and concise, we did not present control searches that included Dyson terms of orders higher than the first. However, such extensions follow from our examples in a very obvious way and can easily be included to any control search protocol. We have tried searching for single spin sequences that include higher order Dyson terms and have found them to also converge to desired solutions. In this chapter, we did not consider control searches for an ensemble of quantum systems or control searches with non-trivial transfer functions. A control problem of this kind will be dealt with in the next chapter, where we demonstrate an experimental implementation of our control sequences that were designed for an ensemble of proton spins in force detected nanoscale nuclear magnetic resonance experiments. In the case of these experiments, we were also dealing with non-trivial transfer functions for the control electronics.

The methods developed here could, in principle, be used in any setting that requires robust coherent control of quantum systems, whether it be quantum computation, sensing or spectroscopy. The main requirement for successful implementation is an accurate model of the system generators and a precise knowledge of the control amplitudes seen by the quantum system, i.e., sufficiently good characterization of the experimental transfer function. The methods work best if the experimenter has an ability to generate continuously varying arbitrary waveforms $a(t)$, although it is also possible to implement the framework in the case of more restrictive control electronics, e.g., an ability to only generate pulses with their phases drawn from a discrete set $\{0, \pi/2, \pi, 3\pi/2\}$. Quantum control problems that have not yet been addressed with tools of this kind include, but are not limited to: minimizing the effect of cross-talk in the case of simultaneous multiple system control, reducing the effect of counter-rotating terms in the cases where the Rabi strength approaches the spin level spacing, i.e., Bloch-Siegert type effects, and adding robustness against variations in the experimental transfer functions. Another potential avenue for using this framework would be the inclusion of non-Hamiltonian Lindblad terms for cross-polarization problems.

In the next section, we discuss some further possible refinements and extensions of the methods described here.

3.10 Outlook

In this section, we first briefly describe some technical aspects related to our framework that this work has not addressed. Finally, we give an example of a control problem, that involves integral expressions of the system propagator $U(t)$, which are not derived from perturbation expansions. Nevertheless, we can still employ the Van Loan differential equation formalism to evaluate such integrals and numerically search for control sequences that would set them to some desired values. We expect this property to generalize to a variety of integral expressions involving $U(t)$.

Although, in this chapter, we did not concern ourselves with the speed and efficiency of the optimization, we remark that the computations involved in our optimizations are frequently highly parallelizable. We could parallelize the Van Loan propagator evaluations at an ensemble member level, or break up Van Loan generators that are of direct sum form, so to evaluate them in parallel. Furthermore, the upper triangular block matrix structure of the Van Loan generators often allows for substantial reductions in the computational cost of their multiplication and exponentiation. Another path for future refinements, which was not explored within this work, is the use of block matrix methods for evaluating partial derivatives of the integral expressions with respect to the control amplitudes. The convergence of the numerical searches to the optimal target function values might also be substantially sped up by the inclusion of second order partial derivatives (Hessians), which can also be evaluated using the same block matrix techniques.

3.10.1 Adiabatic Control

With this subsection we wish to highlight that the numerical methods developed in this chapter are not only useful for numerical control engineering involving perturbation expressions. In fact, the method is potentially applicable for control searches that require the system propagator $U(t)$, $U : [0, T] \to \mathrm{M}_n$, to have any arbitrary property or a set of properties that can be expressed as a set of integral expressions. To exemplify this, we consider the design of adiabatic passages. Here, we will not carry out explicit control searches, rather we state the conditions necessary for performing an adiabatic state to state transfer and demonstrate how they can be rewritten in a way that enables the use of Van Loan differential equations for finding control sequences to implement it.

Adiabatic control sequences have proven to be an indispensable tool in many applications of quantum state to state transfers which require extreme robustness against some system variation [29]. Nevertheless, adiabatic pulse waveforms are almost always postulated analytically and parametrized by a small number of parameters, which are then adjusted for the problem at hand [29, 117]. Such control design could have a few drawbacks. It may not always be obvious how to engineer adiabatic control sequences in unconventional control settings: either in the case of non-standard system generators or in cases where the control waveform has some limitations due to control electronics. Furthermore, analytic waveforms may not easily extend to systems of large Hilbert space dimensions, and they may not always yield the most time-efficient control sequences. Lastly, numerical control engineering has an ability to deal with various ensembles of quantum systems, which is often difficult to handle analytically.

The basic working principle of adiabatic control is that the system Hamiltonian, and accordingly the system generator $G(t) = -iH(t)$, $G : [0, T] \to \mathrm{M}_n$, is modulated in such a way that the quantum state of the system $|\psi(t)\rangle$ is taken from its initial state $|\psi(0)\rangle$ to the final state $|\psi(T)\rangle$ while remaining an approximate eigenstate of $G(t)$ for all $0 \leq t \leq T$. In this subsection, we will consider the case of a spin-1/2 system, just like in Sections 3.4 and 3.8, however, it should be understood that generalizations to larger systems follow easily. The spin will be controlled through a sequence $a : [0, T] \to \mathbb{R}^2$, where again we identify $a_x(t) = a_1(t)$ and $a_y(t) = a_2(t)$, that couple to the x and y angular momentum operators:

$$G(t) = -i\frac{a_x(t)}{2}\sigma_x - i\frac{a_y(t)}{2}\sigma_y. \tag{3.115}$$

The system generator generates a unitary propagator $U(t) = \mathcal{T}\exp\left[\int_0^t dt_1 G(t_1)\right]$. Successful adiabatic passage relies on two aspects: (a) that the time-dependent state of the of the quantum system $|\psi(t)\rangle = U(t)\,|\psi(0)\rangle$ is very close to being an eigenstate of $G(t)$, for all $0 \leq t \leq T$, and (b) that the norm of $G(t)$ is as high as possible throughout $0 \leq t \leq T$. The latter condition ensures that the energy cost for leaving the instantaneous eigenstate $|\psi(t)\rangle$ would remain as high as possible.

We first sketch how could one enforce condition (a) in numerical control searches that utilize Van Loan differential equations. Here, we will not concern ourselves with the particular $|\psi(0)\rangle$ and without loss of generality assume it to be the $i\sqrt{\mathrm{Tr}\,[G^\dagger(0)G(0)]}/\sqrt{2}$ eigenvalue eigenstate of $G(0)$. Requirement (a) can then be expressed as $G(t)\,|\psi(t)\rangle = G(t)U(t)\,|\psi(0)\rangle = i\lambda(t)U(t)\,|\psi(0)\rangle$, where $\lambda(t) = \sqrt{\mathrm{Tr}\,[G^\dagger(t)G(t)]}/\sqrt{2}$. It is clear that this

condition can be rewritten as

$$\left[i\sqrt{\frac{\mathrm{Tr}\left[G^{\dagger}(t)G(t)\right]}{2}}\mathbb{1}_2 - U^{-1}(t)G(t)U(t)\right]|\psi(0)\rangle = 0. \tag{3.116}$$

Multiplying both sides of Equation (3.116) with $\langle\psi(0)|$ reveals that we seek to find an $a(t)$ that would minimize the difference

$$i\sqrt{\frac{\mathrm{Tr}\left[G^{\dagger}(t)G(t)\right]}{2}} - \langle\psi(0)|\, U^{-1}(t)G(t)U(t)\,|\psi(0)\rangle\,, \tag{3.117}$$

for all times $0 \leq t \leq T$. This can be done by minimizing the expression

$$\begin{aligned}\int_0^T dt_1 \frac{\mathrm{Tr}\left[G^{\dagger}(t_1)G(t_1)\right]}{2} \\ + \int_0^T dt_1\, \langle\psi(0)|\, U^{-1}(t_1)G(t_1)U(t_1)\,|\psi(0)\rangle\, \langle\psi(0)|\, U^{-1}(t_1)G(t_1)U(t_1)\,|\psi(0)\rangle\,,\end{aligned} \tag{3.118}$$

where we have used the fact that $G(t)$ is anti-Hermitian, i.e., $G(t) = -G^{\dagger}(t)$.

At first sight, Equation (3.118) does not resemble the integral expressions we have considered so far in this chapter. Specifically, the appearance of the square terms in the integrand seems problematic. Nevertheless, we will now demonstrate a way for rewriting Equation (3.118), such that it can be treated within our framework. We first note that

$$\mathrm{Tr}\left[G^{\dagger}(t)G(t)\right] = -\mathrm{Tr}\left[G(t)G(t)\right] = -\mathrm{Tr}\left[G(t)\otimes G(t)\; U_{\mathrm{SWAP}}\right], \tag{3.119}$$

where U_{SWAP} is the swap operator:

$$U_{\mathrm{SWAP}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{3.120}$$

and we have used the fact that $G(t)$ is an anti-Hermitian operator. Secondly, we point out that

$$\begin{aligned}\langle\psi(0)|U^{-1}(t_1)G(t_1)U(t_1)\,|\psi(0)\rangle\, \langle\psi(0)|\, U^{-1}(t_1)G(t_1)U(t_1)\,|\psi(0)\rangle = \\ \langle\psi(0)|^{\otimes 2}\left[U^{-1}(t_1)G(t_1)U(t_1)\right]^{\otimes 2}|\psi(0)\rangle^{\otimes 2}\,.\end{aligned} \tag{3.121}$$

Combining the above, we can now rewrite Equation (3.118) as

$$\langle\psi(0)|^{\otimes 2}\int_0^T dt_1\left[U^{-1}(t_1)\otimes U^{-1}(t_1)\right]\left[G(t_1)\otimes G(t_1)\right]\left[U(t_1)\otimes U(t_1)\right]|\psi(0)\rangle^{\otimes 2} - \frac{1}{2}\mathrm{Tr}\left[\int_0^T dt_1\; G(t_1)\otimes G(t_1)\; U_{\mathrm{SWAP}}\right], \tag{3.122}$$

which bears closer resemblance to the integral expressions of $U(t)$ that we have been considered so far. Finally, we identify that $U(t)\otimes U(t)$ is generated by $G(t)\otimes\mathbb{1}_2+\mathbb{1}_2\otimes G(t)$, which enables us to immediately write down the Van Loan generator $L(t)\in \mathrm{M}_4(\mathrm{M}_4)$ for the problem:

$$L(t)=\begin{pmatrix} G(t)\otimes\mathbb{1}_2+\mathbb{1}_2\otimes G(t) & G(t)\otimes G(t) & 0 & 0\\ 0 & G(t)\otimes\mathbb{1}_2+\mathbb{1}_2\otimes G(t) & 0 & 0\\ 0 & 0 & 0 & G(t)\otimes G(t)\\ 0 & 0 & 0 & 0\end{pmatrix}. \tag{3.123}$$

Equation (3.118) can now be expressed as

$$\langle\psi(0)|^{\otimes 2}\left[U^{-1}(T)\otimes U^{-1}(T)\right]V_{1,2}(T)\,|\psi(0)\rangle^{\otimes 2} - \frac{1}{2}\mathrm{Tr}\left[V_{3,4}(T)\; U_{\mathrm{SWAP}}\right], \tag{3.124}$$

where $V(T)=\mathcal{T}\exp\left[\int_0^T dt_1 L(t_1)\right]$.

For the adiabatic passage to proceed as desired, we also need to ensure that the instantaneous eigenvalue of the system generator $\lambda(t)$ remains as close as possible to its maximum value $\lambda_{\max}$ for all times. Therefore, we also need to minimize the integral $\int_0^T dt_1\;|\lambda_{\max}|^2+\frac{1}{2}\mathrm{Tr}\left[\int_0^T dt_1\; G(t_1)\otimes G(t_1)\; U_{\mathrm{SWAP}}\right]$, which translates to

$$T|\lambda_{\max}|^2+\frac{1}{2}\mathrm{Tr}\left[V_{3,4}(T)\; U_{\mathrm{SWAP}}\right]. \tag{3.125}$$

The expressions that appear in Equation (3.124) and (3.125) can now be minimized, along with the fidelity function for $U(T)$, which determines the final state $|\psi(T)\rangle$, by defining a suitable target function $\Phi\left[V(T)\right]$ and evaluating its derivatives as we did for all examples in this chapter. We expect that various generalizations of the procedure outlined in this subsection for rewriting integral expressions involving $U(t)$, such that they could be handled by Van Loan differential equations, to be extremely versatile and not at all limited to the example presented here.

Chapter 4

High-resolution Nanoscale Magnetic Resonance

With the previous chapter, we developed a framework for numerical control engineering that includes Dyson terms. Nevertheless, we only demonstrated the ability to generate control sequences that generate the desired system propagators without any numerical or experimental tests evaluating the performance of such sequences. In this chapter, we provide an experimental implementation of such numerically optimized sequences, as well as thorough numerical evaluation of the numerically found controls. We employ our control engineering tools in force detected magnetic resonance experiments working with nanometer scale proton spin ensembles and experimentally demonstrate a 500-fold increase in the effective spin coherence time. We make use of this coherence time by performing one-dimensional magnetic resonance imaging (MRI) experiments with a resolution of $\sim$2 nm.

The experiments presented here demonstrate a neat feature of the numerically engineered control sequences that include Dyson terms. Namely, the operations $U(T)$ that result from these optimizations approximate zero-length control pulses, since the unwanted Hamiltonian terms are suspended over the pulse duration $0 \leq t \leq T$. For zero-length pulses, it is easy to show that all directional derivatives of $U(T)$ with respect to time-independent Hamiltonian terms become zero. In the case of the current experiment, these variations of $U(T)$ are mainly with respect to the dipolar Hamiltonian $D = \left(\sum_{k\in\{x,y,z\}} \sigma_k \otimes \sigma_k - 3\ \sigma_z \otimes \sigma_z\right)$ and σ_z Hamiltonian, which represents either resonance offsets or chemical shifts. This feature of the numerically engineered control sequences implies that the pulses can be used in any existing NMR sequences by substituting the existing unitary rotations with our numerically optimized pulses. The major

benefit of that approach is, that it enables the use of the enormous wealth of existing NMR sequences, e.g., multidimensional spectroscopy and imaging sequences, in much more challenging control settings, such as the one described in this chapter.

In this chapter, we will briefly describe force detected magnetic resonance or magnetic resonance force microscopy (MRFM) in Section 4.1 as well as the properties of the sample and the control electronics used, the latter includes experimental characterization of the transfer functions. Since the author was not directly involved with conducting the experiments, here, we keep the descriptions of the experimental setup and protocols rather brief, we will mostly focus on the aspects relevant to the control engineering. With Section 4.2, we provide full details of the numerical control engineering that was undertaken to perform the MRFM experiments presented in Section 4.3. With Section 4.4 we provide numerical studies of our control sequences, in order to argure for their necessity in these experiments. Furthermore, in Appendix C we show some illustrative simulations of the control action on the full spin ensemble.

All experiments in this chapter were conducted by William Rose and Raffi Budakian, the author conducted all of the control engineering and simulations. Data analysis was conducted by the author, William Rose and Raffi Budakian.

4.1 Force Detected Magnetic Resonance

Force detected magnetic resonance experiments can essentially be thought of as localized Stern-Gerlach experiments. A single spin, or a spin ensemble, of interest are fixed onto a mechanical resonator that has some resonance frequency f and acts as a transducer; applying a magnetic field gradient at the location of the spin, or the spin ensemble, generates a force on the mechanical resonator the spins are fixed to. If that force is periodic and resonant with f, an increase in the oscillator amplitude ensues. For the atto-Newton level forces involved in typical MRFM experiments, the mechanical oscillator generally behaves as a classical damped harmonic oscillator that starting from standstill, hence, the application of a resonant force for some period of time t results in an oscillator amplitude that is proportional to both the net magnetic moment attached to the resonator and the magnetic field gradient strength; the phase of the oscillator reflects the phase of the force, i.e., the direction of the net magnetic moment. Given a means for detecting the oscillator amplitude and phase, a projection of the total angular momentum for the spin sample can be inferred.

There are several ways of generating the necessary oscillatory force on the mechanical oscillators in MRFM experiments. One widely employed way is to apply a magnetic field

gradient by bringing a magnetic (nano)particle near the spin sample, and flip the spins at intervals equalling half the period of the oscillator [97, 20]. Nevertheless, for high-frequency mechanical oscillators – advantageous for high-sensitivity detection – this protocol tends to become impractical since the Rabi field strengths required to yield sufficiently fast spin flips can become prohibitively large [23], that especially in the presence of large ω_0 inhomogeneity, which results from the presence of the magnetic field gradient source.

For the experiments presented in this chapter, the mechanical oscillator was a single crystal silicon nanowire grown via vapour-liquid-solid method. The roughly elliptical nanowire tip dimensions were 60×80 nm^2 and its length was approximately 20 μm. The elliptical cross-section was achieved by asymmetrically depositing, and thereafter etching away, amorphous silicon from the nanowires. The lowest frequency flexural modes, employed for spin detection, had frequencies $f = \{315 \text{ kHz}, 369 \text{ kHz}\}$, with respective quality factors of $Q = \{8000, 8300\}$ at the experimental operating temperature of 4.2 K, and the oscillator spring constant for the $f = 315$ kHz mode was $\sim 1.0 \times 10^{-4}$ N/m. The $f = 315$ kHz mode of the nanowire was used for all MRFM measurements here; the position of the mechanical oscillator was detected with a 2 μm wavelength laser interferometer, the beam of which was polarized along the length of the nanowire, and focused on a roughly 2 μm-diameter spot in the middle of the nanowire.

To circumvent the aforementioned problem of having to flip the sample spins at intervals of $1/(2f) = 1.6$ μs in the presence of a static magnetic field gradient, the resonant time dependent force on the mechanical resonator was induced by a periodic modulation of the field gradients. This was achieved by driving electrical currents of extremely high current densities past the spin ensemble through a nano-fabricated metal device that had a narrow – $\sim 150 \times 80$ nm^2 cross-section, 100 nm-long – constriction on the current's path. The device was fabricated lithographically by liftoff of 5 nm/80 nm-thick Ti/Ag film off of a MgO substrate. A scanning electron micrograph of the device can be seen in Figure 4.1(b). Because the device generates high current densities, and hence, high magnetic field gradients, by focusing electrical currents, we refer to the device as a *current-focusing field gradient source* or CFFGS through the rest of this chapter.

An illustration of the full experimental setup used can be seen in Figure 4.1(a); it includes the CFFGS and the silion nanowire transducer, the axis of which was positioned perpendicular to the CFFGS surface approximately 50 nm above the centre of the constriction. The tip of the nanowire was coated with a thin layer of polystyrene, which contains the proton spins that constituted the spin ensemble used for the MRFM experiments. We will discuss the properties of polystyrene relevant for the experiments and simulations in Section 4.1.3. A scanning electron micrograph of a sample silicon nanowire end, that has been coated with polystyrene, is given as an inset of Figure 4.1(a). The MRFM setup –

kept at ultra-high vacuum environment, and placed in the bore of a superconducting magnet, the $B_0 = 1.13$ T field of which coincides with the nanowire axis – was cooled down to 4.2 K. During spin detection, AC currents of 70 mA peak amplitude at the resonance frequency of the nanowire oscillator were driven through the CFFGS, which generated field gradients of $dB_z/d_x = 1.0 \times 10^6$ T/m at a height of 55 nm above the surface of the device. The resulting resonant force on the nanowire oscillator generated Ångstrom scale motion, which was detected by the laser interferometer. In the next subsection, we discuss the exact spin detection protocol employed during the experiments.

In addition to producing high magnetic field gradients for the spin detection, CFFGS was also used for delivering RF control sequences that were resonant with the Larmor frequency ω_0 of the spin ensemble. This experimental configuration could, at first sight, appear problematic, because the design consideration for the CFFGS to generate high magnetic field gradients across the spin ensemble seems at odds with the demand of generating accurate control operations for the whole ensemble, that is because the resulting Rabi field ω_1 dispersion can end up being very large. In Figure 4.1(c), we show a contour plot of the Rabi frequencies $\omega_1/(2\pi)$ as a function of distance from the CFFGS surface that were produced by COMSOL finite element simulations for 50 mA peak amplitude drive at $\omega_0/(2\pi)$ frequency. The spin control problem is further complicated by the high proton spin density in the solid-state polystyrene sample, which results rather short – dipolar interaction limited – spin coherenece times on the order of 11 μs. Hence, not only need the control sequences $a(t)$ yield the same unitary operation $U(T)$ over the spin ensemble, they also need to suspend the spin evolution under the dipolar Hamiltonian and the chemical shift Hamiltonian during the pulse period $0 \leq t \leq T$, in order to successfully result high fidelity operations.

We solved all of these seeming issues by employing the numerical control engineering tools developed in the previous chapter. In fact, we used the approximate one-to-one mapping between $\omega_1/(2\pi)$ – averaged over the cross-section of the polystyrene sample – and the height z from the CFFGS to our advantage, for performing one-dimensional MRI experiments. Detailed COMSOL simulations of the ω_1 strengths and dB_z/dx values, that were used for analysing the data from these experiments, are given in Appendix B.

4.1.1 Spin Detection

For the experiments presented here, we worked with spin noise, which results from the statistical imbalance of spins that are in $|\downarrow\rangle$ state versus $|\uparrow\rangle$ state and scales with $\sqrt{N_{\text{spins}}}$,

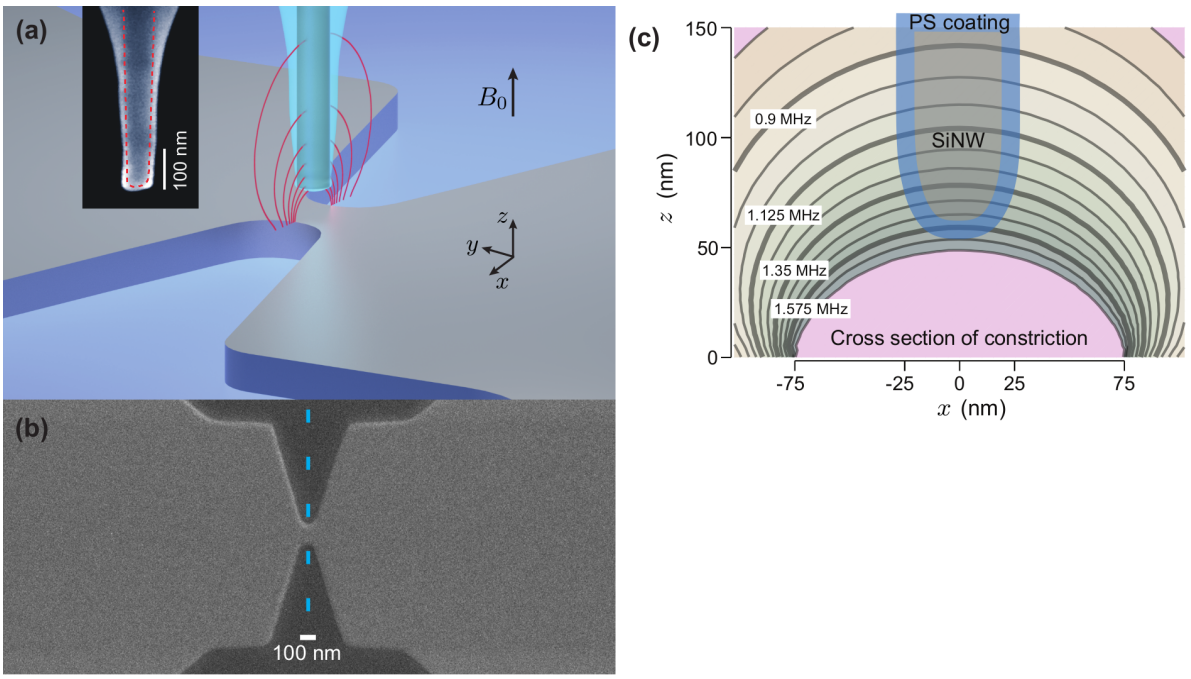

Figure 4.1: (a) A schematic illustrating the silicon nanowire coated in polystyrene and current focusing field gradient source (CFFGS) configuration during the force deteced magnetic resonance experiments. The nanowire axis is aligned with the external magnetic field marked as B_0. The red loops near the constriction of the CFFGS denote contours of constant Rabi frequency $\omega_1/(2\pi)$ generated by currents flowing in the CFFGS. (Inset) Representative scanning electron micrograph (SEM) of a polystyrene-coated silicon nanowire. The dashed red line marks the perimeter of the silicon nanowire itself. (b) SEM of the CFFGS used in experiments. (c) A plot of constant Rabi frequency $\omega_1/(2\pi)$ contours generated by the CFFGS at 50-mA peak RF current in a plane along the dashed line in sub-figure (b). The axis in the figure correspond to those in sub-figure (a). Besides the Rabi contours, we also display an illustrative schematic of the polystyrene-coated silicon nanowire positioned around 50 nm above the CFFGS top surface, which coincides with the horizontal axis in the figure.

where N_{spins} is the total number of spins in the ensemble. For small enough spin ensembles, statistical polarization can greatly exceed thermal polarization. Furthermore, using statistical polarization enables one to perform experiments continuously without waiting for the spins to thermalize, a feature which can substantially reduce experiment times for low temperature nuclear spin systems that can often have very long relaxation times T_1's. The main disadvantage of spin noise detection is the fact that the resulting signal cannot be coherently averaged. This is because for spin noise measurements, the initial spin polarization has a random magnitude and direction, and hence, one has to always measure and compare overall spin magnetization before and after a period of spin control operations.

As we mentioned already, the overall magnetization of the spin ensemble is measured by driving AC currents, resonant with the frequency f of the nanowire oscillator, through the CFFGS and detecting the amplitude of the oscillator. Nevertheless, when driving the CFFGS with AC currents at frequency f, the resonant AC voltage that accompanies the AC currents can cause electrostatic excitation of the mechanical oscillator by coupling to stray charges present on the nanowire. To prevent such electrostatic excitation of the nanowire a measurement scheme called *modulated alternating gradients generated with currents* (MAGGIC) was used. The MAGGIC readout protocol is designed such that the AC current, and hence, the AC voltage modulation has minimal frequency components at f, which is achieved by inverting the phase of the gradient modulation periodically at every 0.2 ms. To ensure that the force generated on the nanowire is still resonant with f, the sample spins are also inverted simultaneously with the gradient modulation phase inversions. We display the exact timings of the measurement scheme in Figure 4.2. Given the extremely broad range of Rabi strengths ω_1 within the sample demonstrated by Figure 4.1(c), *adiabatic full passages* (AFPs) were used to invert the spins of interest, we discuss the action of the AFPs in the next subsection.

Spin noise encoding means that the z-projection of the spin angular momentum has to be measured both before and after performing the unitary operations. In Figure 4.2, we indicate that the experiment splits into three blocks: two measurement periods of duration 130 ms before and after an encoding period of duration τ_e. We now work out the explicit details of the MAGGIC protocol. We use a convention that the average force on the mechanical oscillator generated by a single spin at location $\vec{r}$ in state $|\uparrow\rangle$, and sinusoidally modulated field gradient that has a peak value of $g(\vec{r}) = \frac{dB_z(\vec{r})}{dx}$, is given by $F = \mu g(\vec{r})/\sqrt{2}$, where μ is the magnetic moment of the spin. We model the instantaneous spin z-angular momentum projection under MAGGIC readout protocol for a proton at location $\vec{r}$ as a stationary zero-mean Markovian process, described by the random telegraph function $h(\vec{r}, t)$. Hence, $h(\vec{r}, t)$ is a stochastic function, that any instant takes values of ± 1. For some initial value $h(\vec{r}, 0) = \pm 1$, the probability of Q spin flips over an interval $[0, t]$ is given

by the Poisson process, such that $\text{prob}(Q) = (t/\tau_m)^Q \exp(-t/\tau_m)/Q!$. The last property determines the correlation function $\langle h(\vec{r},t')\, h(\vec{r},t)\rangle = \exp(-|t-t'|/\tau_m)$. We note that the correlation time τ_m encapsulates the effects of both the thermal T_1 relaxation of the nuclear spins and the spin flips occurring under the application of AFPs.

Given the protocol in Figure 4.2(a), we can express the integrated force correlation function for a single spin at location $\vec{r}$ inside the sample:

$$\begin{aligned} c[U(\vec{r})] = & \\ & \Theta[U(\vec{r})]\,\chi[\omega_1(\vec{r})]\, e^{-\tau_e/T_c}\, \frac{1}{\tau_0^2}\int_0^{\tau_0} dt_1 \int_0^{\tau_0} dt_2 \left\langle \frac{\mu g(\vec{r})}{\sqrt{2}} h(\vec{r},t_1)\, \frac{\mu g(\vec{r})}{\sqrt{2}} h(\vec{r},\tau_0+t_2)\right\rangle = \\ & \frac{1}{2}\mu^2 g^2(\vec{r})\,\Theta[U(\vec{r})]\,\chi[\omega_1(\vec{r})]\, e^{-\tau_e/T_c}\left(\frac{\tau_m}{\tau_0}\right)^2 \left(1-e^{-\tau_0/\tau_m}\right)^2, \end{aligned} \tag{4.1}$$

where

$$\Theta[U(\vec{r})] = \langle\uparrow| U^\dagger(\vec{r})\,\sigma_z U(\vec{r})\,|\uparrow\rangle, \tag{4.2}$$

and $U(\vec{r})$ is the unitary operation applied on the nuclear spin over the encoding period τ_e. T_c is the effective spin dephasing time during the encoding period. The high-pass filter $\chi[\omega_1(\vec{r})]$ in Equation 4.1 arises because of the AFPs present in the detection scheme, specifically, due to the fact that AFPs have a high yield for flipping spins only for sufficiently high Rabi strengths $\omega_1(\vec{r})$, since that is necessary for fulfilling the adiabatic condition. We derive the properties of χ in the next section, here, we note that by choosing a specific amplitude and modulation rate for the AFPs, we can effectively set $\chi[\omega_1(\vec{r})] = 1$ only for the spins of interest, and $\chi[\omega_1(\vec{r})] = 0$ for the rest of the spins. Finally, as we mentioned already, assuming that the nanowire motion is described by a classical driven harmonic oscillator that is stationary at the beginning of the measurement period, the integrated resonant force and the amplitude of the nanowire have one-to-one correspondence, hence, the interferometric measurement tracks the spatial integral of the quantity in Equation 4.1:

$$C = \frac{1}{2}\mu^2 e^{-\tau_e/T_c}\left(\frac{\tau_m}{\tau_0}\right)^2 \left(1-e^{-\tau_0/\tau_m}\right)^2 \int d\vec{r}\,\rho(\vec{r})\, g^2(\vec{r})\,\Theta[U(\vec{r})]\,\chi[\omega_1(\vec{r})], \tag{4.3}$$

where $\rho(\vec{r})$ is the spatial spin density inside the sample. Accordingly, C will be the quantity we quote for all experimental results in Section 4.3. It is clear that in the limit $\tau_0/\tau_m \ll 1$ and $\tau_e/\tau_m \ll 1$ the measurement C is essentially a weighted projection of spin z-angular momentum.

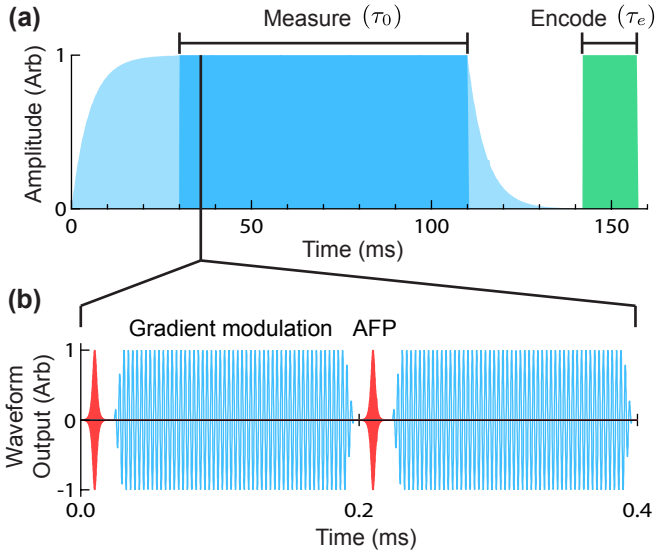

Figure 4.2: (a) Timings of the MAGGIC detection protocol. The measurement period, indicated in blue, consists of gradient modulation at the frequency f of the nanowire, which is turned on and off with an exponential rise and fall times of $Q/(\pi f) = 8$ ms to minimize spurious electrostatic excitation of the silicon nanowire. Only once the gradient modulation has reached its peak value, the AFPs are also turned on for a duration of τ_0, yielding a resonant net spin magnetization dependent force on the silicon nanowire oscillator. The measurement window is followed by an encoding period of duration τ_e, during which RF pulses resonant with the spin Larmor frequency are applied. The encoding period is always followed by a second measurement period not shown here. (b) Pulse diagram for the MAGGIC waveform during the measurement period which includes periods of gradient modulation interspersed with AFPs.

4.1.2 High-Pass Filter

In the previous subsection, we introduced the high-pass filter $\chi\left(\omega_1(\vec{r})\right)$ that arises because the adiabaticity condition for the AFPs is fulfilled only by spins experiencing sufficiently high Rabi strengths. Here, we will outline how to calculate the characteristics of $\chi\left(\omega_1(\vec{r})\right)$, used for the imaging experiments presented in Section 4.3.

The AFPs work by turning on the effective magnetic field along the z-direction, and thereafter slowly inverting the direction of the magnetic field. The desired spin flips ensue as long as the effective magnetic field is strong enough, and modulated slowly enough, that the spins will remain in an eigenstate of the effective field Hamiltonian throughout the AFP pulse duration. An effective field along the z-direction is implemented by starting and finishing the pulses far off resonance. Specifically, we used hyperbolic secant and hyperbolic tangent modulation of the x and z-components of the effective field, respectively [29]. The rotating frame spin Hamiltonian during the AFPs is given by

$$H_{\mathrm{AFP}}(\omega_1, \lambda, t) = \omega_1 \left(a_1(t)\frac{\sigma_x}{2} + a_2(t)\frac{\sigma_y}{2} \right) + 2\pi a_3(\lambda, t)\frac{\sigma_z}{2}, \tag{4.4}$$

where ω_1 is the Rabi strength and λ is an adjustable parameter that determines the resonance offset at the beginning and at the end of the AFP. The pulse envelope for $0 \leq t \leq T$ is given by

$$\{a_1(t), a_2(t), a_3(t,\lambda)\} = \{\mathrm{sech}\left[\mathrm{arcsech}(\kappa)\left(\frac{2t}{T}-1\right)\right], 0, \lambda \tanh\left[\mathrm{arcsech}(\kappa)\left(\frac{2t}{T}-1\right)\right]\}, \tag{4.5}$$

with $\kappa = 0.001$ and $T = 20$ μs. For our imaging experiments we used AFPs with three different λ values: $\lambda = \{\lambda_1 = 10^6, \lambda_2 = 2\times10^6, \lambda_3 = 3\times10^6\}$. To evaluate the performance of the AFPs we calculated the Rabi strength ω_1 and the λ-dependent unitary at the end of the AFP pulse by discretizing $a_1(t)$ and $a_3(t,\lambda)$ with a time step $\Delta T = T/N$ for $N = 60,000$, and calculated a product of unitary propagators:

$$U_{\mathrm{AFP}}\left(\omega_1, \lambda\right) = \prod_{i=1}^{N} \exp\left[-i\left(\omega_1 a_1^{(N)}\frac{\sigma_x}{2} + 2\pi a_3^{(N)}(\lambda)\frac{\sigma_z}{2} \right)\Delta T \right], \tag{4.6}$$

where $a_1^{(i)} = a_1(i\Delta t)$ and $a_3^{(i)}(\lambda) = a_3(i\Delta t, \lambda)$, for $i \in \{1, ..., N-1, N\}$. Furthermore, since we assume full dephasing between consecutive AFPs, we only needed to evaluate $\upsilon\left(\omega_1, \lambda\right)$ – the remnant z-axis polarization after a single AFP for a spin with Rabi strength ω_1. $\upsilon\left(\omega_1, \lambda\right) = 1 - 2\left|\langle\downarrow| U_{\mathrm{AFP}}\left(\omega_1, \lambda\right)|\uparrow\rangle\right|^2$, where $\left|\langle\downarrow| U_{\mathrm{AFP}}\left(\omega_1, \lambda\right)|\uparrow\rangle\right|^2$ is the probability for an AFP to yield a flip.

Amongst the work resented later in this chapter, we conducted two separate nanoscale MRI experiments: one with 75-kHz resolution and one with 25-kHz resolution, these experiments used different χ functions that will be given below. For these one-dimensional imaging experiments we approximate $\chi\left(\omega_1(\vec{r})\right) \approx \chi\left(\omega_1(z)\right)$. The MAGGIC protocol for the 75-kHz resolution imaging experiments involved performing 400 AFPs, each separated by 200 μs, with $\lambda = \lambda_1$ during the two MAGGIC readout periods before and after the encoding period. The corresponding high-pass filter function for the 75-kHz resolution experiment is then given as

$$\begin{aligned}\chi_{75}\left(\omega_1\right) &= \frac{1}{401}\sum_{n=0}^{400}\left|\upsilon\left(\omega_1,\lambda_1\right)\right|^n \frac{1}{401}\left|\upsilon\left(\omega_1,\lambda_1\right)\right|^{400}\sum_{n=0}^{400}\left|\upsilon\left(\omega_1,\lambda_1\right)\right|^n \\ &= \left|\upsilon\left(\omega_1,\lambda_1\right)\right|^{400}\left(\frac{1}{401}\frac{1-\left|\upsilon\left(\omega_1,\lambda_1\right)\right|^{401}}{1-\left|\upsilon\left(\omega_1,\lambda_1\right)\right|}\right)^2. \end{aligned} \tag{4.7}$$

The 25-kHz resolution imaging experiments used 400 AFPs, each separated by 200 μs, with $\lambda = \lambda_2$ during the detection periods. Also, to further reduce the signal from low ω_1 spins, an additional 100 AFPs with $\lambda = \lambda_3$ were applied at the beginning of the encoding period. Hence, the $\chi(\omega_1)$ function for the 25-kHz resolution experiments is given as

$$\begin{aligned}\chi_{25}\left(\omega_1\right) &= \frac{1}{401}\sum_{n=0}^{400}\left|\upsilon\left(\omega_1,\lambda_2\right)\right|^n \frac{1}{401}\left|\upsilon\left(\omega_1,\lambda_3\right)\right|^{100}\left|\upsilon\left(\omega_1,\lambda_2\right)\right|^{400}\sum_{n=0}^{400}\left|\upsilon\left(\omega_1,\lambda_2\right)\right|^n \\ &= \left|\upsilon\left(\omega_1,\lambda_3\right)\right|^{100}\left|\upsilon\left(\omega_1,\lambda_2\right)\right|^{400}\left(\frac{1}{401}\frac{1-\left|\upsilon\left(\omega_1,\lambda_2\right)\right|^{401}}{1-\left|\upsilon\left(\omega_1,\lambda_2\right)\right|}\right)^2. \end{aligned} \tag{4.8}$$

4.1.3 Sample

The sample used for the experiments was a thin, ∼1 to ∼100 nm thick, polystyrene coating applied to the last 1 to 2 μm portion of the silicon nanowire. The particular coating procedure resulted an amorphous crystal structure. Polystyrene has a high density of proton spins, which is advantageous for spin detection, yet it also results a rather short – dipolar interaction limited – nuclear spin dephasing time of $T_2 = 11$ μs. Here, we give the relative coordinates for the proton spins in a single polystyrene monomer – styrene. Styrene molecule consists of eight carbon and eight hydrogen atoms, we denote the coordinates of the latter in a single styrene molecule by $\vec{r}_i$ for $1 \leq i \leq 8$. We retrieved $\{\vec{r}_i\}$ from Wolfram Mathematica Chemical Data database [124], and we give the relative proton positions

$\vec{r}_{ij} = \vec{r}_i - \vec{r}_j$ for each pair $\{i, j\}$, with $1 \leq i \leq 7$ and $i+1 \leq j \leq 8$, as a column vector in units of Ångstrom in the table below.

	1	2	3	4	5	6	7
2	$\begin{pmatrix} 0.0 \\ 2.4 \\ -3.5 \end{pmatrix}$ -0.35	$\begin{pmatrix} -2.2 \\ -0.92 \\ -0.64 \end{pmatrix}$ 5.43	$\begin{pmatrix} -2.2 \\ 1.6 \\ -4.1 \end{pmatrix}$ -0.49	$\begin{pmatrix} -3.3 \\ -0.12 \\ -2.7 \end{pmatrix}$ -0.06	$\begin{pmatrix} 1.9 \\ 2.7 \\ -1.7 \end{pmatrix}$ 1.57	$\begin{pmatrix} 2.0 \\ 0.08 \\ -0.25 \end{pmatrix}$ 5.04	$\begin{pmatrix} 3.3 \\ 1.4 \\ -0.3 \end{pmatrix}$ 2.11
3		$\begin{pmatrix} -2.2 \\ -3.3 \\ 2.9 \end{pmatrix}$ 0.49	$\begin{pmatrix} -2.2 \\ -0.87 \\ -0.60 \end{pmatrix}$ 5.34	$\begin{pmatrix} -3.4 \\ -2.5 \\ 0.86 \end{pmatrix}$ 1.42	$\begin{pmatrix} 1.9 \\ 0.32 \\ 1.8 \end{pmatrix}$ -0.12	$\begin{pmatrix} 1.9 \\ -2.3 \\ 3.3 \end{pmatrix}$ -0.66	$\begin{pmatrix} 3.3 \\ -1.0 \\ 3.2 \end{pmatrix}$ -0.40
4			$\begin{pmatrix} -0.02 \\ 2.5 \\ -3.5 \end{pmatrix}$ -0.35	$\begin{pmatrix} -1.1 \\ 0.8 \\ -2.0 \end{pmatrix}$ -3.93	$\begin{pmatrix} 4.1 \\ 3.7 \\ -1.1 \end{pmatrix}$ 0.67	$\begin{pmatrix} 4.2 \\ 1.0 \\ 0.39 \end{pmatrix}$ 0.89	$\begin{pmatrix} 5.5 \\ 2.3 \\ 0.33 \end{pmatrix}$ 0.45
5				$\begin{pmatrix} -1.1 \\ -1.7 \\ 1.5 \end{pmatrix}$ 3.82	$\begin{pmatrix} 4.1 \\ 1.2 \\ 2.4 \end{pmatrix}$ 0.35	$\begin{pmatrix} 4.2 \\ -1.5 \\ 3.9 \end{pmatrix}$ -0.21	$\begin{pmatrix} 5.5 \\ -0.17 \\ 3.8 \end{pmatrix}$ -0.02
6					$\begin{pmatrix} 5.2 \\ 2.9 \\ 0.96 \end{pmatrix}$ 0.46	$\begin{pmatrix} 5.3 \\ 0.2 \\ 2.4 \end{pmatrix}$ 0.09	$\begin{pmatrix} 6.6 \\ 1.5 \\ 2.4 \end{pmatrix}$ 0.14
7						$\begin{pmatrix} 0.09 \\ -2.7 \\ 1.5 \end{pmatrix}$ 0.23	$\begin{pmatrix} 1.4 \\ -1.4 \\ 1.4 \end{pmatrix}$ -4.28
8							$\begin{pmatrix} 1.3 \\ 1.3 \\ -0.06 \end{pmatrix}$ 18.87

In Section 4.4, we make use of $\{\vec{r}_{ij}\}$ when simulating the evolution of eight dipolar coupled proton spins in a styrene molecule under our numerically engineered control sequences. $\{\vec{r}_{ij}\}$ determine the strength of the dipolar interaction ξ_{dipolar} between a pair of proton spins:

$$\xi_{\text{dipolar}}\left(\vec{r}_{ij}, \hat{\zeta}\right) = \frac{\mu_0 \gamma_p^2 \hbar}{8\pi^2 \left|\vec{r_{ij}}\right|^3} \frac{1}{2} \left(1 - 3\left[\frac{\vec{r}_{ij} \cdot \hat{\zeta}}{|\vec{r}_{ij}|}\right]^2\right), \tag{4.9}$$

where $\mu_0 = 4\pi \times 10^{-7}$ kg $\cdot$ m/ $\left(\text{s}^2 \cdot \text{A}^2\right)$ is the permeability of free space, $\gamma_p = 2.675 \times 10^8$ rad/ $(\text{s} \cdot \text{T})$ is the proton gyromagnetic ratio and $\hat{\zeta}$ denotes a unit vector pointing in the direction of the external field B_0, evaluated in the same coordinate system as $\{\vec{r}_{ij}\}$. The dipolar Hamiltonian for a pair of spins is therefore expressed as $H_{\text{dipolar-pair}} = 2\pi\hbar \ \xi_{\text{dipolar}} \frac{1}{4} \left(\sum_{k \in \{x,y,z\}} \sigma_k \otimes \sigma_k - 3 \ \sigma_z \otimes \sigma_z\right)$. In the table above, we have also quoted the values for $\left\{\xi_{\text{coupling}}\left(\vec{r}_{ij}, \hat{\zeta}\right)\right\}$ in units of kHz for a particular $\hat{\zeta} = (-0.71, -0.70, 0.03)$,

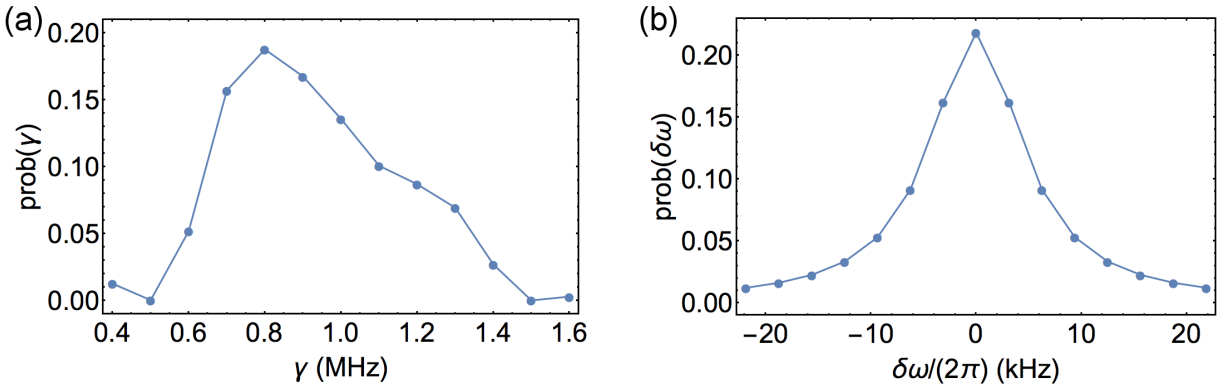

Figure 4.3: (a) Experimentally determined distribution of Rabi field strengths parametrized by γ such that $\omega_1 = 2\pi\gamma$. (b) 10.6 kHz wide Lorentzian distribution corresponds to the $\omega_0/(2\pi)$ inhomogeneity measured. Because we use the points displayed, and their respective probabilities, in our spin signal simulations in Section 4.4, the probability distributions are normalized such that $\sum_{\gamma\in\Gamma} \text{prob}\,(\gamma) = 1$ and $\sum_{\delta\omega\in\Delta} \text{prob}\,(\delta\omega) = 1$.

which appear outside of the $\vec{r}_{ij}$ column matrices. The ξ_{coupling} values that reach tens of kHz explain the relatively short dephasing times for the proton spins in our experiments.

For the initial experiments, we determined the probability distribution of polystyrene Rabi strengths $\omega_1 = 2\pi\gamma$ and chemical shifts/resonance offsets in the sample volume of interest to be given by the values in Figure 4.3.

4.1.4 Transfer Function[1]

In order to ensure that we have a precise knowledge of control amplitudes $\{b^{(\gamma)}(t)\}$ seen by the ensemble of proton spins, as a function of control sequence $a(t)$, a set of separate experiments were conducted to determine the transfer function of the electronics. These experiments involved bringing a different electrically grounded, gold-coated mechanical resonator near the CFFGS, which coupled electrostatically to the currents passing through the CFFGS. The electrostatically induced force $F(t)$ on the mechanical resonator scales as the square of the current $I(t)$ that is driven through the CFFGS. We will demonstrated

[1]The transfer function measurement protocol was developed by Raffi Budakian and William Rose, who also conducted its measurement.

below that the transfer function for the electronics can be determined by driving the CFFGS with two pure current tones separated in frequency by the resonance frequency f_{Au} of the mechanical oscillator, i.e.,

$$I(t) = I_0 \left(\cos\left[2\pi(\nu + f_{\mathrm{Au}}/2)t\right] + \cos\left[2\pi(\nu - f_{\mathrm{Au}}/2)t\right]\right), \tag{4.10}$$

and measuring the amplitude response of the oscillator.

The time-dependent force on the gold-coated nanowire that results from the current specified by Equation (4.10) is

$$\begin{aligned} F(t) \propto I(t)^2 = I_0^2 \cos\left(2\pi f_{\mathrm{Au}} t\right) \\ + \frac{I_0^2}{2}\left(2 + \cos\left[2\pi(2\nu + f_{\mathrm{Au}})t\right] + \cos\left[2\pi(2\nu - f_{\mathrm{Au}})t\right] + 2\cos\left[4\pi\nu t\right]\right), \end{aligned} \tag{4.11}$$

the term on the first line of the equation being resonant with f_{Au}, while all terms on the second line of the equation are vastly off resonance with the mechanical oscillator, since $f_{\mathrm{Au}} \approx 300$ kHz, while the ν of interest for the magnetic resonance experiments is in the range of 40 MHz to 70 MHz. It was first experimentally verified that for such ν range, the amplitude response of the gold-coated resonator was indeed quadratic in I_0, implying that the control electronics behaved as a linear system and could be modelled the way we described in Section 3.6 of the previous chapter. The quadratic behaviour of the oscillator amplitude response also confirmed that the amplitude transfer function $\lambda(\nu)$ for the electronics did not vary significantly over frequency intervals of width f_{Au}, therefore, $\lambda(\nu)$ could be directly inferred from the resulting amplitude of the mechanical oscillator. For determining the phase transfer function $\phi(\nu)$, quadrature detection of the mechanical oscillator was necessary, as the phase of the force on the oscillator depends on the first derivative of $\phi(\nu)$. A set of measurements for CFFGS currents described by Equation (4.10) at different ν values yielded $\lambda(\nu)$ and $\phi(\nu)$ values displayed in Figure 4.4(a) and (b), respectively.

For our spin control, the carrier frequency of the control waveform $a(t)$ was made to coincide with $\omega_0/(2\pi) = \gamma_{^1\mathrm{H}} B_0/(2\pi)$, where ω_0 is the proton Larmor frequency, $\gamma_{^1\mathrm{H}}$ denotes the proton gyromagnetic ratio and B_0 the external magnetic field applied. The control waveform $a(t)$ is distorted the least by $\{\Xi^{(\gamma)}\}$, if the value of $\lambda(\nu)$ is the same for all of its spectral components and the value of $\phi(\nu)$ only varies linearly with ν. The latter corresponds merely to a time delay between the output of $a(t)$ waveform by the AWG and the arrival of $\{b^{(\gamma)}(t)\}$ at the site of the spins. For our experiments, we were free to pick the value of ω_0 by changing B_0, and we could also use optimization transfer function techniques, described in Section 3.6.1 of the previous chapter, to limit the frequency range

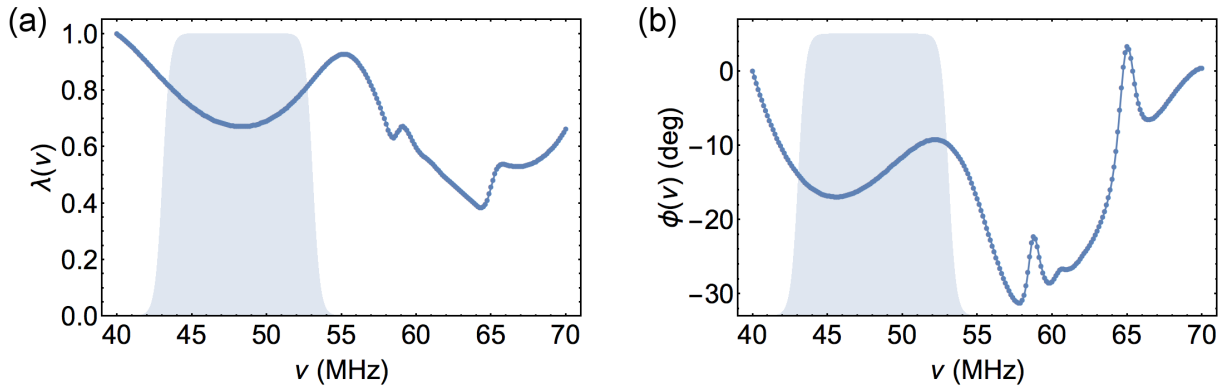

Figure 4.4: (a) Amplitude transfer function λ as a function of frequency ν for the control electronics, determined through electrostatic force detection scheme. The shaded area marks the bandpass filter described by Equation 3.80 for restricting the control sequence frequency components within $\Delta\nu = 10$ MHz of the carrier frequency of $\omega_0/(2\pi) = 48$ MHz. (b) Phase transfer function ϕ as a function of frequency ν for the control electronics that was also determined through the electrostatic force detection scheme. The experimentally determined amplitude and phase transfer functions have been normalized to have values one and zero at $\nu = 40$ MHz, respectively.

of the pulse waveform. This freedom gave us an ability to reduce the control waveform distortions. Given the experimental measurements of $\lambda(\nu)$ and $\phi(\nu)$ displayed in Figure 4.4, we decided to choose a B_0 value that corresponded to $\omega_0/(2\pi) = 48$ MHz, and limit the frequency components of $a(t)$ within a $\Delta\nu = 10$ MHz bandwidth around that carrier frequency. The functional form of such bandpass filter, or a low-pass around the carrier frequency ω_0, is illustrated by the shaded areas in Figure 4.4. Finally, the deviations of $\lambda(\nu)$ and $\phi(\nu)$ from a constant function and a linear function, respectively, were accounted for by including the particular λ and ϕ values into the definitions of $\{\Xi^{(\gamma)}\}$, given by Equation 3.74, when conducting the control searches.

4.2 Control Engineering

In this section, we describe numerical control optimization for three separate control sequences. The experiments that will be introduced in the next section were done at two different nanowire positions with respect to the CFFGS. This yielded two slightly different distributions for proton spin Rabi frequencies, therefore, we engineered two separate control sequences for these experiments. In addition, to explore the necessity of Dyson terms for our control engineering, we carried out comparisons of control performance in numerical simulations that contrasted one of the pulses that implemented Dyson term minimization with a third sequence that was not engineered with such targets. We will present these simulations in Section 4.4. Here, we first define a general optimization target for the three aforementioned sequences. We will follow this with the exact details of each optimization, along with the pulse waveforms found and various metrics characterizing the control performance.

We will define our control problem exactly according to the ensemble control setup abstraction that was laid out in Section 3.5 of the previous chapter. We say that we have an ensemble of proton spins labelled by $\gamma \in \Gamma$. Here, we consider this ensemble to be a representative ensemble of spins in the sample volume of interest. Each γ has an associated unique transfer function $\Xi^{(\gamma)}$ that determines the control amplitudes $b^{(\gamma)}(t)$ seen by γ, as a function of the control sequence $a(t)$. We constrain the maximum amplitude $|a(t)|$ of the control sequence to be equal to one. Our transfer functions $\{\Xi^{(\gamma)}\}$ reflect the effects of both the amplitude and phase transfer functions shown in Figure 4.4 and the Rabi field distributions displayed in Figure 4.3.

In addition to the transfer functions $\{\Xi^{(\gamma)}\}$ defined by Equation 4.12, we also use an optimization transfer function Ξ^{opt} in order to limit the range of frequency components in $a(t)$ to within a bandwidth of $\Delta\nu$, as well as to enforce that the pulse starts and ends

with zero amplitude. Each of the three control sequences $a(t)$ has piecewise constant pulse amplitudes for N equal periods of duration ΔT, so that the total pulse length is $T = N\Delta T$. The bandpass filter that is incorporated into the optimization transfer function is illustrated by the shaded area in Figure 4.4(a). $\Xi^{\rm opt}(N, N_0, \Delta T, \Delta\nu)$ is defined by Equation 3.82 in Section 3.6. The control sequence is then determined by $\alpha^{\rm opt} \in {\rm M}_{2,N^{\rm opt}}(\mathbb{R})$, where $N^{\rm opt} = N - 2N_0$, and N_0 determines the number of zero amplitude intervals of duration ΔT at the beginning and at the end of the sequence.

It is natural to label $\{\Xi^{(\gamma)}\}$ according to the maximum Rabi frequency they yield on the nuclear spins, i.e., according to the $|b^{(\gamma)}(t)|/(2\pi)$ value corresponding to $a(t) = 1$ for all $0 \leq t \leq T$. The transfer functions could then be written explicitly as

$$\Xi^{(\gamma)}(N, \Delta T) = 2\pi\gamma\ \Xi(N, \Delta T, \lambda, \phi), \tag{4.12}$$

where the function $\Xi(N, \Delta t, \lambda, \phi)$ is given by Equation (3.74) in Section 3.6. $\lambda(\nu)$ and $\phi(\nu)$ in Equation (4.12) are evaluated by interpolating the experimental transfer function data in Figure 4.4. The particular set of γ values, along with N and ΔT, depend on the specific control setting and will be quoted individually for the three control searches below. Because our control electronics was the same for all experiments, we take λ and ϕ in Equation (4.12) to be the same for each search.

For the following problems, the system generators $\{G^{(\gamma)}(t)\}$ are given by

$$G^{(\gamma)}(t) = -i\frac{b_1^{(\gamma)}(t)}{2}\sigma_x - i\frac{b_2^{(\gamma)}(t)}{2}\sigma_y, \tag{4.13}$$

while the system propagators evaluate to $U^{(\gamma)}(t) = \mathcal{T}\exp\left[\int_0^t dt_1 G^{(\gamma)}(t_1)\right]$, for each $\gamma \in \Gamma$. Our control amplitudes $b^{(\gamma)}(t)$ are taken to be resonant with the proton spin Larmor frequency ω_0. This implies that we take there to be no drift part, i.e., no z angular momentum operator, for the system generators. The control amplitudes $\{b^{(\gamma)}(t)\}$ that are specified by a matrix $\beta^{(\gamma)} \in {\rm M}_{2,N}(\mathbb{R})$, for each $\gamma \in \Gamma$, equate to

$$\beta_{1,j}^{(\gamma)} = {\rm Re}\left[\Xi^{(\gamma)}(N, \Delta t)\ \Xi^{\rm opt}(N, N_0, \Delta T, \Delta\nu)(\alpha_1^{\rm opt} - i\alpha_2^{\rm opt})\right]_j \tag{4.14}$$

and

$$\beta_{2,j}^{(\gamma)} = -{\rm Im}\left[\Xi^{(\gamma)}(N, \Delta t)\ \Xi^{\rm opt}(N, N_0, \Delta T, \Delta\nu)(\alpha_1^{\rm opt} - i\alpha_2^{\rm opt})\right]_j, \tag{4.15}$$

where $\alpha^{\rm opt} \in {\rm M}_{2,N^{\rm opt}}(\mathbb{R})$ is the matrix specifying the optimization waveform $a^{\rm opt}(t)$. In order to enforce $|a(t)| \leq 1$, for all $0 \leq t \leq T$, we use the following constraint: $-1/\sqrt{2} \leq$

$\alpha_i^{\rm opt} \leq 1/\sqrt{2}$ for $i \in \{1,2\}$. The control sequence $a(t)$, that is implemented experimentally, is specified by another matrix $\alpha \in \mathrm{M}_{2,N}(\mathbb{R})$ the elements of which are calculated as

$$\alpha_{1,j} = \mathrm{Re}\left[\Xi^{\rm opt}(N, N_0, \Delta T, \Delta\nu)(\alpha_1^{\rm opt} - i\alpha_2^{\rm opt})\right]_j \tag{4.16}$$

and

$$\alpha_{2,j} = -\mathrm{Im}\left[\Xi^{\rm opt}(N, N_0, \Delta T, \Delta\nu)(\alpha_1^{\rm opt} - i\alpha_2^{\rm opt})\right]_j, \tag{4.17}$$

after a suitable $\alpha^{\rm opt}$ is found.

Our most general objective is to find a control sequence that would yield $U^{(\gamma)}(T) = U_{\rm target} = \exp\left(-i\frac{\pi}{2}\frac{\sigma_x}{2}\right)$, as well as

$$\left[U^{(\gamma)}(T) \otimes U^{(\gamma)}(T)\right]^{-1} \mathcal{D}_{U^{(\gamma)}\otimes U^{(\gamma)}}(D) = \left[U^{(\gamma)}(T)\right]^{-1} \mathcal{D}_{U^{(\gamma)}}(\sigma_z) = 0, \tag{4.18}$$

for each $\gamma \in \Gamma$. We assign equal weights $p^{(\gamma)} = \frac{1}{|\Gamma|}$ to each member of the ensemble and define the following combined target function:

$$\begin{aligned}\Phi = 1 &- \frac{1}{|\Gamma|}\sum_{\gamma\in\Gamma} p_U \sqrt{1 - \left(\mathcal{F}\left[U_{\rm target}, U^{(\gamma)}(T)\right]\right)^2} \\ &- \frac{1}{|\Gamma|}\sum_{\gamma\in\Gamma}\left(p_D \frac{\|\mathcal{D}_{U^{(\gamma)}\otimes U^{(\gamma)}}(D)\|}{\max_{a(t)}\|\mathcal{D}_{U^{(\gamma)}\otimes U^{(\gamma)}}(D)\|} + p_{\sigma_z}\frac{\|\mathcal{D}_{U^{(\gamma)}}(\sigma_z)\|}{\max_{a(t)}\|\mathcal{D}_{U^{(\gamma)}}(\sigma_z)\|}\right),\end{aligned} \tag{4.19}$$

where p_U, p_D and p_{σ_z} are the relative weights for the three-part target function, such that $0 \leq p_U, p_D, p_{\sigma_z} \leq 1$ and $p_U + p_D + p_{\sigma_z} = 1$. The weights depend on the particular search, and will be stated for each search separately.

We construct a set of Van Loan generators $L^{(\gamma)} \in \mathrm{M}_{12}$ that decompose into a direct sum of $\mathrm{M}_2(\mathrm{M}_2)$ and $\mathrm{M}_2(\mathrm{M}_4)$:

$$L^{(\gamma)}(t) = \tag{4.20}$$
$$\begin{pmatrix} G^{(\gamma)}(t) & \sigma_z & 0 & 0 \\ 0 & G^{(\gamma)}(t) & 0 & 0 \\ 0 & 0 & G^{(\gamma)}(t)\otimes\mathbb{1} + \mathbb{1}\otimes G^{(\gamma)}(t) & D \\ 0 & 0 & 0 & G^{(\gamma)}(t)\otimes\mathbb{1} + \mathbb{1}\otimes G^{(\gamma)}(t) \end{pmatrix},$$

the corresponding Van Loan propagators of which are given as

$$V^{(\gamma)}(t) = \mathcal{T}\exp\left[\int_0^t dt_1 L(t_1)\right] \tag{4.21}$$
$$= \begin{pmatrix} U^{(\gamma)}(t) & \mathcal{D}_{U^{(\gamma)}}(\sigma_z) & 0 & 0 \\ 0 & U^{(\gamma)}(t) & 0 & 0 \\ 0 & 0 & U^{(\gamma)}(t)\otimes U^{(\gamma)}(t) & \mathcal{D}_{U^{(\gamma)}\otimes U^{(\gamma)}}(D) \\ 0 & 0 & 0 & U^{(\gamma)}(t)\otimes U^{(\gamma)}(t) \end{pmatrix}.$$

Here, we will slightly abuse our sub-matrix index notation, just as it was done in Section 3.4.4 of the previous chapter, and write the target function in Equation 4.19 as a function of $\{V^{(\gamma)}(T)\}$:

$$\Phi = 1 - \frac{1}{|\Gamma|}\sum_{\gamma\in\Gamma}\frac{p_U}{2}\sqrt{4 - \mathrm{Tr}\left[V_{1,1}^{(\gamma)}(T)\; e^{i\frac{\pi}{2}\frac{\sigma_x}{2}}\right]\mathrm{Tr}\left[e^{-i\frac{\pi}{2}\frac{\sigma_x}{2}}\left(V_{1,1}^{(\gamma)}(T)\right)^\dagger\right]} \tag{4.22}$$
$$- \frac{1}{|\Gamma|}\sum_{\gamma\in\Gamma}\left(\frac{p_D}{2\sqrt{6}T}\sqrt{\mathrm{Tr}\left[V_{3,4}^{(\gamma)}(T)\left(V_{3,4}^{(\gamma)}(T)\right)^\dagger\right]} + \frac{p_{\sigma_z}}{\sqrt{2}T}\sqrt{\mathrm{Tr}\left[V_{1,2}^{(\gamma)}(T)\left(V_{1,2}^{(\gamma)\dagger}(T)\right)^\dagger\right]}\right),$$

where we have used the fact that $\max_{a(t)}\|\mathcal{D}_{U^{(\gamma)}\otimes U^{(\gamma)}}(D)\| = 2\sqrt{6}T$, $\max_{a(t)}\|\mathcal{D}_{U^{(\gamma)}}(\sigma_z)\| = \sqrt{2}T$, and that $\mathrm{Tr}\left[V_{1,1}^{(\gamma)}(T)\left(V_{1,1}^{(\gamma)}(T)\right)^\dagger\right] = \mathrm{Tr}[\mathbb{1}_2] = 2$.

We now evaluate the partial derivatives of Equation (4.22) with respect to the elements

of $\beta^{(\gamma)}$, for each $\gamma \in \Gamma$:

$$\frac{\partial}{\partial \beta_{i,j}^{(\gamma)}}\Phi = \frac{p_U}{2|\Gamma|}\frac{\mathrm{Re}\left(\mathrm{Tr}\left[\left(\frac{\partial}{\partial \beta_{i,j}^{(\gamma)}}V^{(\gamma)}(T)\right)_{1,1} e^{i\frac{\pi}{2}\frac{\sigma_x}{2}}\right]\mathrm{Tr}\left[e^{-i\frac{\pi}{2}\frac{\sigma_x}{2}}\left(V_{1,1}^{(\gamma)}(T)\right)^\dagger\right]\right)}{\sqrt{4-\mathrm{Tr}\left[V_{1,1}^{(\gamma)}(T)\, e^{i\frac{\pi}{2}\frac{\sigma_x}{2}}\right]\mathrm{Tr}\left[e^{-i\frac{\pi}{2}\frac{\sigma_x}{2}}\left(V_{1,1}^{(\gamma)}(T)\right)^\dagger\right]}}$$
$$- \frac{p_D}{2\sqrt{6}T|\Gamma|}\frac{\mathrm{Re}\left(\mathrm{Tr}\left[\left(\frac{\partial}{\partial \beta_{i,j}^{(\gamma)}}V^{(\gamma)}(T)\right)_{3,4}\left(V_{3,4}^{(\gamma)}(T)\right)^\dagger\right]\right)}{\sqrt{\mathrm{Tr}\left[V_{3,4}^{(\gamma)}(T)\left(V_{3,4}^{(\gamma)}(T)\right)^\dagger\right]}}$$
$$- \frac{p_{\sigma_z}}{\sqrt{2}T|\Gamma|}\frac{\mathrm{Re}\left(\mathrm{Tr}\left[\left(\frac{\partial}{\partial \beta_{i,j}^{(\gamma)}}V^{(\gamma)}(T)\right)_{1,2}\left(V_{1,2}^{(\gamma)}(T)\right)^\dagger\right]\right)}{\sqrt{\mathrm{Tr}\left[V_{1,2}^{(\gamma)}(T)\left(V_{1,2}^{(\gamma)}(T)\right)^\dagger\right]}}, \tag{4.23}$$

where we have used that fact that for any $\rho \neq \gamma$: $\frac{\partial}{\partial \beta_{i,j}^{(\rho)}}V^{(\gamma)}(T) = 0$, for all i and j. In order to carry out the gradient ascent searches to find α^{opt} that yields $\Phi \approx 1$, we evaluate partial derivatives of Φ with respect to the elements of α^{opt}:

$$\frac{\partial}{\partial \alpha_{1,j}^{\mathrm{opt}}}\Phi = \sum_{\gamma\in\Gamma}\sum_{t=1}^{N}\mathrm{Re}\left[\Xi^{(\gamma)}(N,\Delta T)\,\Xi^{\mathrm{opt}}(N,N_0,\Delta T,\Delta\nu)\right]_{t,j}\frac{\partial}{\partial \beta_{1,t}^{(\gamma)}}\Phi \tag{4.24}$$
$$- \sum_{\gamma\in\Gamma}\sum_{t=1}^{N}\mathrm{Im}\left[\Xi^{(\gamma)}(N,\Delta t)\,\Xi^{\mathrm{opt}}(N,N_0,\Delta T,\Delta\nu)\right]_{t,j}\frac{\partial}{\partial \beta_{2,t}^{(\gamma)}}\Phi$$

and

$$\frac{\partial}{\partial \alpha_{2,j}^{\mathrm{opt}}}\Phi = \sum_{\gamma\in\Gamma}\sum_{t=1}^{N}\mathrm{Im}\left[\Xi^{(\gamma)}(N,\Delta t)\,\Xi^{\mathrm{opt}}(N,N_0,\Delta T,\Delta\nu)\right]_{t,j}\frac{\partial}{\partial \beta_{1,t}^{(\gamma)}}\Phi \tag{4.25}$$
$$+ \sum_{\gamma\in\Gamma}\sum_{t=1}^{N}\mathrm{Re}\left[\Xi^{(\gamma)}(N,\Delta t)\,\Xi^{\mathrm{opt}}(N,N_0,\Delta T,\Delta\nu)\right]_{t,j}\frac{\partial}{\partial \beta_{2,t}^{(\gamma)}}\Phi.$$

Equipped with $\Phi\left(\{\alpha_{i,j}^{\mathrm{opt}}\}\right)$, that is given by the combination of Equations (4.22), (4.14) and (4.15), and the partial derivatives of Φ with respect to $\{\alpha_{i,j}^{\mathrm{opt}}\}$, that are given by the

combination of Equations (4.23), (4.24) and (4.25), we can now proceed with the specific pulse searches.

Our control searches are conducted in the way it was described in Section 3.4 of the previous chapter – we keep on increasing the pulse length T until sufficiently rapid convergence towards the maximum of Φ is observed, while keeping the number of time steps N high enough to ensure that ΔT is substantially shorter than the Rabi cycle. All searches are started with $\{\alpha_{i,j}^{\text{opt}}\}$ drawn from independent uniform pseudorandom distributions over a range $\left[-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right]$. The three individual quantities that appear in Equation (4.22) are the unitary metric $\Psi_U^{(\gamma)}$, dipolar metric $\Psi_D^{(\gamma)}$, and σ_z metric $\Psi_{\sigma_z}^{(\gamma)}$, which are defined as

$$\Psi_U^{(\gamma)} = \sqrt{1 - \left(\mathcal{F}\left[\exp\left(-i\frac{\pi}{2}\frac{\sigma_x}{2}\right), U^{(\gamma)}(T)\right]\right)^2}, \tag{4.26}$$

$$\Psi_D^{(\gamma)} = \frac{\|\mathcal{D}_{U^{(\gamma)}\otimes U^{(\gamma)}}(D)\|}{\max_{a(t)}\|\mathcal{D}_{U^{(\gamma)}\otimes U^{(\gamma)}}(D)\|}, \tag{4.27}$$

$$\Psi_{\sigma_z}^{(\gamma)} = \frac{\|\mathcal{D}_{U^{(\gamma)}}(\sigma_z)\|}{\max_{a(t)}\|\mathcal{D}_{U^{(\gamma)}}(\sigma_z)\|}, \tag{4.28}$$

respectively. p_U, p_D and p_{σ_z} – the relative weights of $\Psi_U^{(\gamma)}$, $\Psi_D^{(\gamma)}$ and $\Psi_{\sigma_z}^{(\gamma)}$ in Equation (4.22) – are picked so as to give approximately equal minimization rates for $\Psi_U^{(\gamma)}$, $\Psi_D^{(\gamma)}$ and $\Psi_{\sigma_z}^{(\gamma)}$ during the control optimization procedure. In the next three subsections, we give the specific values for the parameters N, N_0, $\Delta\nu$, ΔT, p_U, p_D and p_{σ_z} that were used for engineering the three control sequences used for the experiments and simulations in the upcoming sections. For each control search, we also state the representative set of Rabi strengths Γ, and present a figure displaying the pulse waveform and the values of $\Psi_U^{(\gamma)}$, $\Psi_D^{(\gamma)}$ and $\Psi_{\sigma_z}^{(\gamma)}$ as functions of γ.

4.2.1 Pulse #1

The first pulse for the nanoscale magnetic resonance measurements was designed to act on a spin ensemble with proton spin Rabi strengths $\omega_1/(2\pi)$ ranging from 0.6 MHz to 1.2 MHz. Therefore, we picked a representative set Γ, the elements of which are given as $\gamma \in \{0.6\text{ MHz}, 0.65\text{ MHz}, 0.69\text{ MHz}, 0.74\text{ MHz}, 0.78\text{ MHz}, 0.83\text{ MHz}, 0.88\text{ MHz}, 0.92\text{ MHz}, 0.97\text{ MHz}, 1.02\text{ MHz}, 1.06\text{ MHz}, 1.11\text{ MHz}, 1.15\text{ MHz}, 1.2\text{ MHz}\}$. The control sequence was designed to suspend the evolution under both the dipolar Hamiltonian and the chemical shift Hamiltonian, hence, we used the following parameters specifying the control problem:

N	522	$\Delta\nu$	10 MHz	p_U	5/9
N_0	12	ΔT	0.0248 μs	p_D	3/9
N^{opt}	498	T	12.95 μs	p_{σ_z}	1/9

The resulting control sequence and its figures of merit are shown in Figure 4.5. We remark that the sequence length of ∼13 μs corresponds to only 7.8 Rabi cycles for the spins experiencing $\omega_1/(2\pi) = 0.6$ MHz. The desired action of the sequence is clear from Figure 4.5(d,e,f) that shows the values for $\Psi_U^{(\gamma)}$, $\Psi_D^{(\gamma)}$ and $\Psi_{\sigma_z}^{(\gamma)}$, respectively, for a range of γ values.

4.2.2 Pulse #2

The second control pulse was designed for a relocated ensemble of polystyrene proton spins, with their Rabi strengths $\omega_1/(2\pi)$ ranging from 0.9 MHz to 1.75 MHz. Our representative set for these strengths was $\gamma \in \{0.9$ MHz, 0.965 MHz, 1.03 MHz, 1.095 MHz, 1.16 MHz, 1.225 MHz, 1.29 MHz, 1.355 MHz, 1.42 MHz, 1.485 MHz, 1.55 MHz, 1.615 MHz, 1.68 MHz, 1.745 MHz$\}$. Again, the sequence was designed to suspend the evolution under the dipolar Hamiltonian and the chemical shift Hamiltonian. We used the following parameters that specified the control problem:

N	360	$\Delta\nu$	10 MHz	p_U	5/9
N_0	30	ΔT	0.0208 μs	p_D	3/9
N^{opt}	300	T	7.49 μs	p_{σ_z}	1/9

The resulting control sequence and its figures of merit are shown in Figure 4.6. The sequence length of ∼7.5 μs corresponds to 6.75 Rabi cycles for the spins experiencing $\omega_1/(2\pi) = 0.9$ MHz. Figure 4.6(d,e,f) gives the values for $\Psi_U^{(\gamma)}$, $\Psi_D^{(\gamma)}$ and $\Psi_{\sigma_z}^{(\gamma)}$, respectively, for a range of γ values.

4.2.3 Pulse #3

This control sequence was generated only to argue for the necessity of including Dyson terms to our numerical control engineering. We demonstrated this by searching for a control sequence with the same Γ as pulse #1, yet we set $p_D = p_{\sigma_z} = 0$, such that the search target was exclusively determined by $\Psi_U^{(\gamma)}$. Accordingly, the set of Rabi strength parameters for the pulse search was $\gamma \in \{0.6$ MHz, 0.65 MHz, 0.69 MHz, 0.74 MHz,

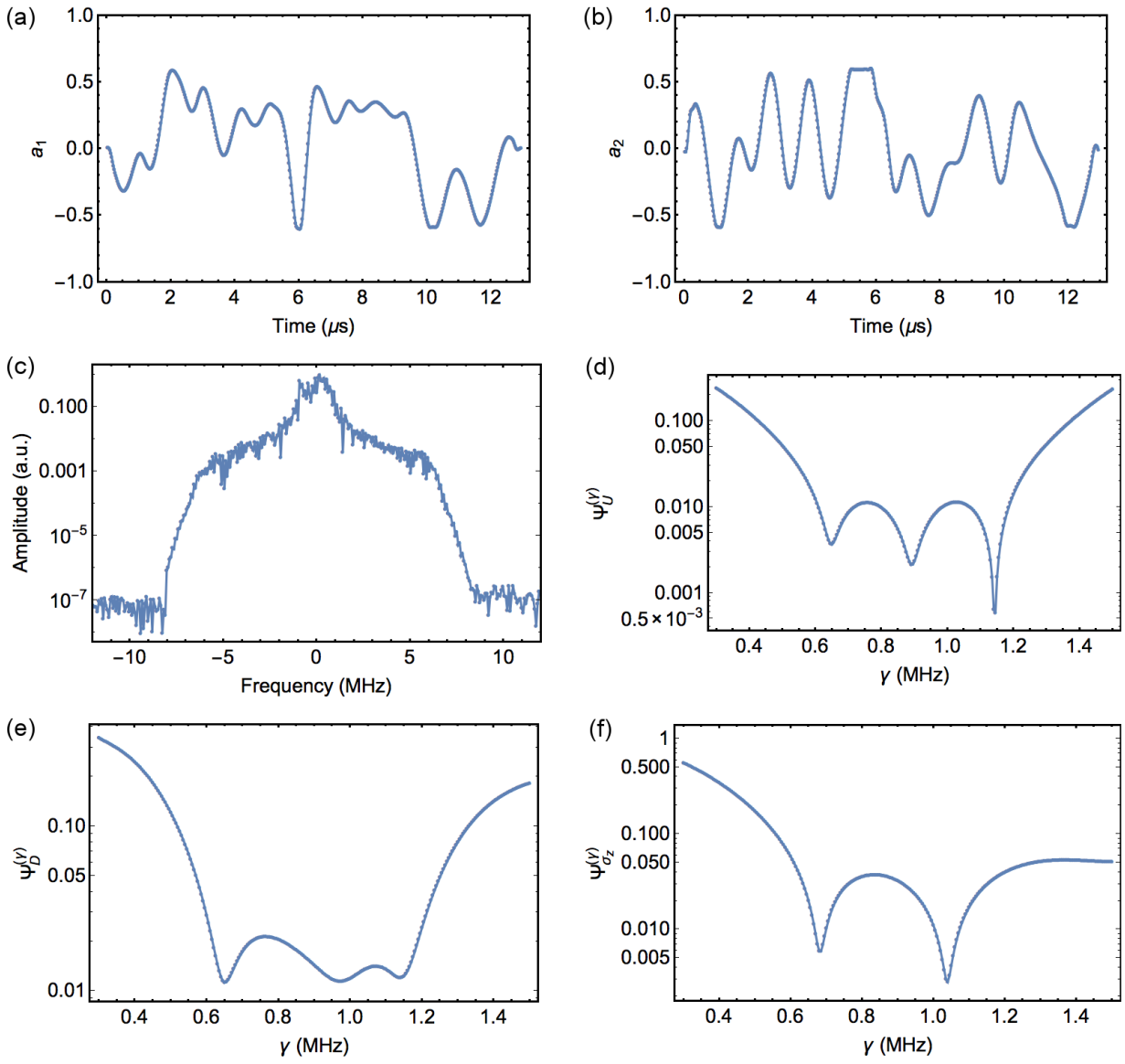

Figure 4.5: (a) $a_1(t)$ for pulse #1. (b) $a_2(t)$ for pulse #1. (c) Absolute value of the Fourier-transformed pulse centred at the carrier frequency. The limited spectral range of the sequence waveform is clearly visible. (d) Unitary metric $\Psi_U^{(\gamma)}$ defined in Equation (4.26) as a function of the Rabi strength parameter γ. It can be seen that the pulse targets the range 0.6 MHz to 1.2 MHz. (e) Dipolar metric $\Psi_D^{(\gamma)}$ defined in Equation (4.27) as a function of γ. (f) σ_z metric $\Psi_{\sigma_z}^{(\gamma)}$ defined in Equation (4.28) as a function of γ.

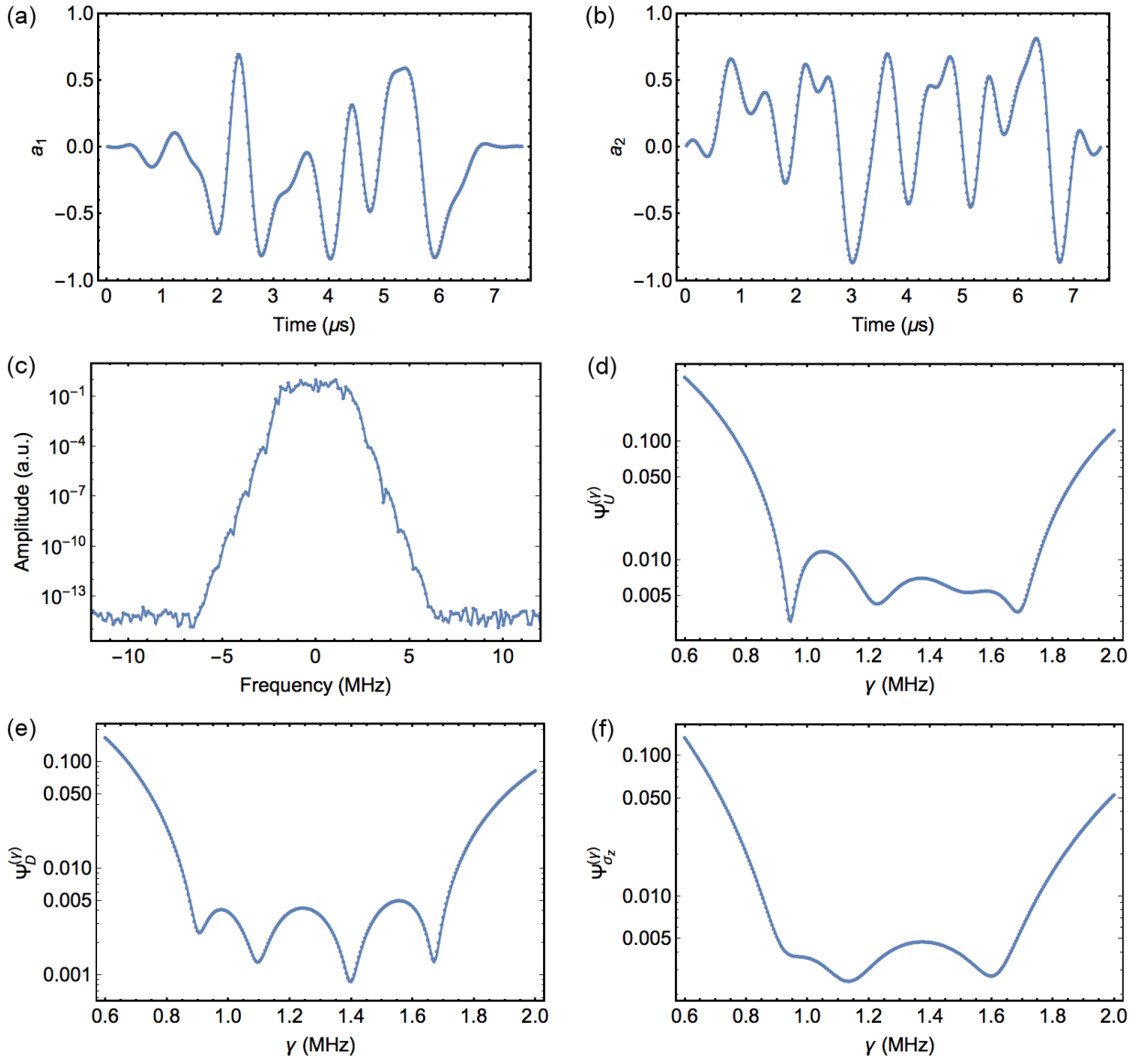

Figure 4.6: (a) $a_1(t)$ for pulse #2. (b) $a_2(t)$ for pulse #2. (c) Absolute value of the Fourier-transformed pulse centered at the carrier frequency. (d) Unitary metric $\Psi_U^{(\gamma)}$ defined in Equation (4.26) as a function of γ; the target range $2\pi \times 0.9$ MHz to $2\pi \times 1.75$ MHz is clearly seen. (e) Dipolar metric $\Psi_D^{(\gamma)}$ defined in Equation (4.27) as a function of γ. (f) σ_z metric $\Psi_{\sigma_z}^{(\gamma)}$ defined in Equation (4.28) as a function of γ.

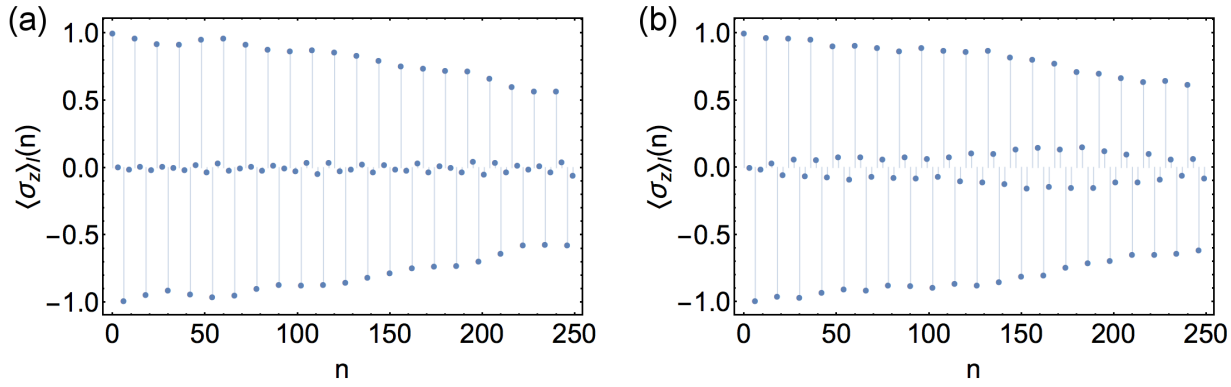

Figure 4.7: (a) Single spin simulation results of $\langle\sigma_z\rangle(n)$, defined in Equation (4.29) for pulse #1, and (b) simulations of $\langle\sigma_z\rangle(n)$ for pulse #3 as a function of n. n is the number of pulses applied.

0.78 MHz, 0.83 MHz, 0.88 MHz, 0.92 MHz, 0.97 MHz, 1.02 MHz, 1.06 MHz, 1.11 MHz, 1.15 MHz, 1.2 MHz}, the rest of the search parameters being:

N	292	$\Delta\nu$	10 MHz	p_U	1
N_0	26	ΔT	0.0248 μs	p_D	0
N^{opt}	240	T	7.24 μs	p_{σ_z}	0

Because the sequence in this subsection was engineered for comparing it with pulse #1, we terminated the control optimization once the performance of the pulse, without any dipolar effects nor resonance offsets, mimicked the performance of pulse #1. To determine that, we evaluate the expectation value $\langle\sigma_z\rangle(n)$, that has been averaged over all $\gamma\in\Gamma$, for $0\leq n\leq 250$ back-to-back applications of either pulse on a single spin. We used this kind of termination condition, rather than matching the $\Psi_U^{(\gamma)}$ curves for the pulses, to ensure that the eight spin simulations – that will be presented in Section 4.4 in order to compare the performance of pulse #1 and pulse #3 – will only highlight errors resulting from evolution under the dipolar and σ_z Hamiltonians, rather than incoherence from the Rabi dispersion. The ensemble averaged σ_z expectation value is

$$\langle\sigma_z\rangle(n)=\sum_{\gamma\in\Gamma}\text{prob}(\gamma)\,\langle\uparrow|\left(\left[U^{(\gamma)}(T)\right]^{\dagger}\right)^{n}\sigma_z\left[U^{(\gamma)}(T)\right]^{n}|\uparrow\rangle\,, \tag{4.29}$$

where

$$U^{(\gamma)}(T) = \prod_{j=1}^{N} \exp\left[2\pi\gamma\left(-i\alpha_{1,j}\frac{\sigma_x}{2} - i\alpha_{2,j}\frac{\sigma_y}{2}\right)\Delta T\right], \tag{4.30}$$

and the values of $\gamma \in \Gamma$, and their relative weights $\{\text{prob}(\gamma)\}$ in Equation (4.29), are taken from an experimentally determined probability distribution given in Figure 4.3(a). We evaluated Equation (4.29) for pulse #1 and pulse #3 and show the results in Figure 4.7.

Pulse #3 and its figures of merit are shown in Figure 4.8. The sequence length of $\sim$7.24 μs is a little longer than half of that for pulse #1. Figure 4.8(d,e,f) gives the values for $\Psi_U^{(\gamma)}$, $\Psi_D^{(\gamma)}$ and $\Psi_{\sigma_z}^{(\gamma)}$, respectively, for pulse #3 in blue and for pulse #1 in red. It can be seen that while the unitary metric $\Psi_U^{(\gamma)}$ for pulse #3 is actually somewhat better than the one for pulse #1, its dipolar and σ_z metrics, $\Psi_D^{(\gamma)}$ and $\Psi_{\sigma_z}^{(\gamma)}$, are substantially worse.

4.3 Experimental Results[2]

In this section, we introduce, and present the data, for all of the MRFM experiments that were undertaken using our numerically engineered effective Hamiltonian control sequences introduced in the previous section. We will demonstrate three sets of increasingly more elaborate experiments, that successively build on the results of each other. First, we briefly present initial experimental tests for pulse #1; we then demonstrate the use of these pulses within an existing NMR sequence – the magic echo. Finally, we compose the magic echo sequences into more elaborate sequences – called symmetric magic echos – which were employed for our nanoscale MRI experiments.

Our numerically engineered control sequences were first tested by applying a cascade of n pulse #1's – introduced in Section 4.2.1 – on the spin ensemble and measuring the spin correlation function C as a function of the number of pulses n applied. The experimentally determined relative probabilities for the Rabi strength parameter γ, and the chemical shifts/resonance offsets $\delta\omega$ for the proton spins in this experiment, are given in Figure 4.3(a) and (b), respectively. The pulses were applied in increments of five pulses for $n = \{0, 5, 10, ..., 250\}$, and the experimental results can be seen in Figure 4.9 where we display the normalized spin correlation quantity $C(n)/C(n = 0)$. In the case of perfect $\pi/2$ operations for the entire spin ensemble, the spin correlation would behave sinusoidally, i.e.,

[2]All experimental measurements presented in this section were carried out by William Rose. Data analysis and interpretation was carried out by the author, William Rose and Raffi Budakian.

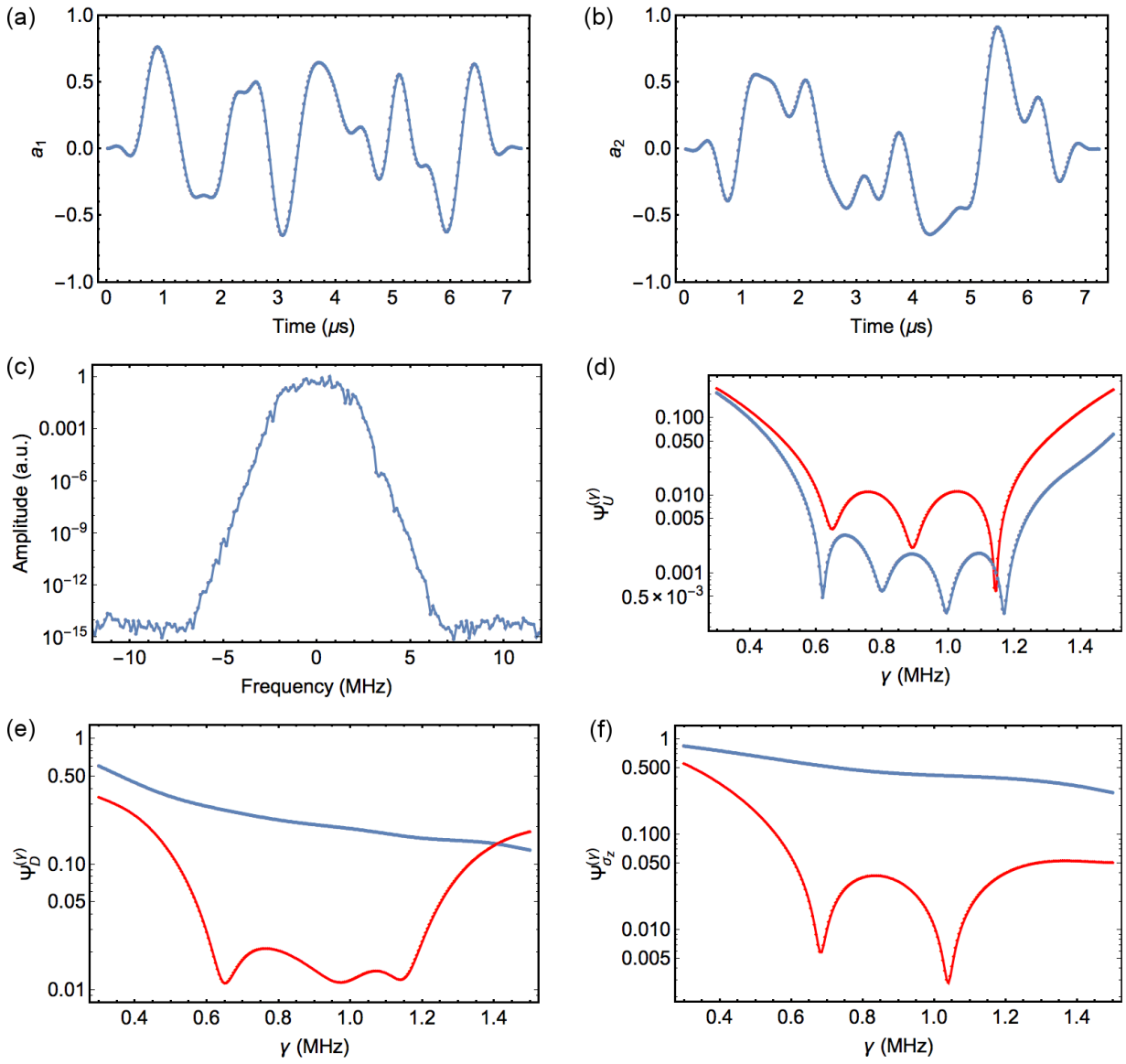

Figure 4.8: (a) $a_1(t)$ for pulse #3. (b) $a_2(t)$ for pulse #3. (c) Absolute value of the Fourier transformed pulse centred at the carrier frequency. (d) Unitary metric $\Psi_U^{(\gamma)}$ defined in Equation (4.26) for pulse #3 (blue) and for pulse #1 (red). (e) Dipolar metric $\Psi_D^{(\gamma)}$ defined in Equation (4.27) for pulse #3 (blue) and for pulse #1 (red). (f) σ_z metric $\Psi_{\sigma_z}^{(\gamma)}$ defined in Equation (4.28) for pulse #3 (blue) and for pulse #1 (red).

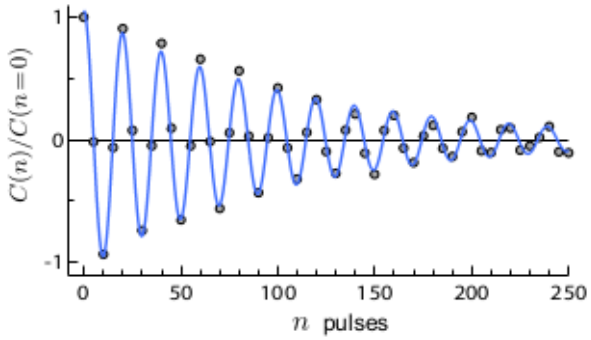

Figure 4.9: Normalized spin correlation $C(n)/C(n=0)$ as a function of n, the number of pulse #1's applied on the spin ensemble. C is defined by Equation (4.3). The data has been fitted to the function $\cos(n\pi/10)\exp(-cn)$, with $c = 0.01$ yielding the best least squares fit.

$C(n)/C(n=0) = \cos(n\pi/2)\exp(-cn)$, with $c = 0$. Any loss of coherence during the pulse, or imperfect implementation of the desired $\pi/2$ rotation, would yield a non-zero c-value. In Figure 4.9, we fit a function $\cos(n\pi/10)\exp(-cn)$, rather than $\cos(n\pi/2)\exp(-cn)$, to the experimental data since the former provides a better guide to the eye over the range of n-values used while enabling us to extract the value of c equally well.

From the data in Figure 4.9, we extract $c = 0.01$, which means that for spins starting from the z-axis, the pulse operations happen with ~99% accuracy, all while the ~13 μs duration of pulse #1 exceeds the dipolar limited coherence time of 11 μs. This experiment shows that the numerical control engineering can yield high-fidelity unitary operations for an ensemble of spins that experience vastly different Rabi strengths and are controlled by control electronics that has a non-trivial transfer function. In Section 4.4, we present numerical simulations that demonstrate that control sequences, which did not include Dyson term optimization, would not have yielded performance fidelities as high as the experimental results in Figure 4.9 demonstrate.

Figure 4.9 also indicates that our control engineering methods provide us with unitary operations that suspend the spin evolution under the dipolar and chemical shift Hamiltonians, and hence, they act as approximations for control pulses of infinitesimal duration. That feature enables us to substitute the unitary rotations into any existing magnetic resonance sequence. With the next two subsections we demonstrate an experimental usage of

our control waveforms within the magic echo sequence that was introduced in Section 2.1.1 of Chapter 2, and that is particularly useful for magnetic resonance imaging experiments.

4.3.1 Magic Echo Experiments

In our implementation of the magic echo sequence, we replaced the usual square pulses implementing the $\pi/2$ spin rotations, that were shown in Figure 2.1, with pulse #1. Pulse #1 implements a $\pi/2$ rotation around x-axis; the in-phase $a_1(t)$ and quadrature $a_2(t)$ amplitudes of the pulse are shown in Figure 4.5(a) and (b), respectively. For the rest of this section we adopt a notational shorthand which specifies Ω_x as an implementation of pulse #1. Furthermore, we use Ω_y to denote an implementation of $\pi/2$ rotation around y-axis that is achieved by swapping $a_1(t)$ and $a_2(t)$ of pulse #1, while $\Omega_{\bar{x}}$ implements an inverse rotation of pulse #1 by multiplying $a_1(t)$ and $a_2(t)$ with -1, and $\Omega_{\bar{y}}$ implements an analogous inverse rotation around y-axis. For the rotary echo, we use a notation that designates $\mathbf{x}$, $\bar{\mathbf{x}}$, $\mathbf{y}$ and $\bar{\mathbf{y}}$ as waveforms for which $(a_1, a_2) = (1, 0)$, $(a_1, a_2) = (-1, 0)$, $(a_1, a_2) = (0, 1)$ and $(a_1, a_2) = (0, -1)$, respectively.

We illustrate our implementation of the magic echo sequence with Figure 4.10. The blue blocks designate the numerically engineered $\pi/2$ pulses, while the pink blocks designate the rotary echo part of the sequence. The parameter τ, that determines the rotary echo $\mathbf{x} : \bar{\mathbf{x}}$ length, was chosen to be $\tau = \{15\ \mu\text{s}, 25\ \mu\text{s}, 35\ \mu\text{s}\}$, depending on the experiment. For all experiments, we picked τ to be greater than the dipolar spin dephasing time of 11 μs in order to avoid spin locking effects [74, 75]. Since the 4τ-long rotary echo is symmetric about its centre, the ω_1 inhomogeneity throughout the sample does not have an effect on the final unitary operation of the full magic echo sequence. In the next subsection, we show how introducing an asymmetry into the rotary echo can be employed for MRI experiments.

Our first experiment in this subsection verified that the numerically engineered pulses can indeed be used within the magic echo sequence. Our experimental protocol first brought the spins to the xy plane by applying a Ω_x pulse, this was followed by $\Omega_{\bar{x}}$ pulse either within the initial or the final zero amplitude interval of the magic echo sequence, the resulting data and the pulse sequence used are depicted in Figure 4.10. The $\Omega_{\bar{x}}$ pulse was applied to read out the spin z-projection and the time axis in the figure indicates at what point the $\Omega_{\bar{x}}$ pulse was applied. The length of the whole experimental sequence as we as the individual Ω pulses was chosen to be an integer number of Larmor cycles to ensure an accurate characterization of the spin coherence.

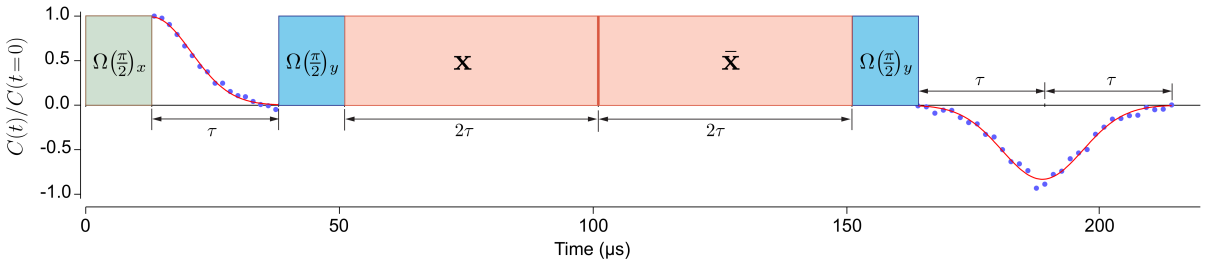

Figure 4.10: Magic echo sequence: the green (Ω_x) and the blue (Ω_y) blocks denote the application of pulse #1 and pink blocks ($\mathbf{x}$ and $\bar{\mathbf{x}}$) denote a rotary echo. The blue dots represent the normalized spin correlation $C(t)/C_0$ measurements after applying a $\Omega_{\bar{x}}$ pulse, which is not shown, at various times during the magic echo sequence. C is defined by Equation (4.3). The data has been fitted with Gaussian functions that have a decay time of $T_2 = 11$ μs.

In Figure 4.10, we have fitted a Gaussian decay envelope $C(t)/C(0) = \exp\left[-(t/T_2)^2\right]$, with $T_2 = 11$ μs, to the initial decay in the first zero amplitude window of the magic echo and the revival of the spin signal at the end of the sequence. For the second fit, we reference t to the end of the magic echo sequence, which happens $6\tau+2T = 176$ μs after the beginning of the sequence, where $\tau = 25$ μs is the period used for the particular magic echo experiment and $T = 13$ μs is the length of pulse #1. $C_0 = C(0)$ denotes the spin correlation value at the start of the sequence. The inverted nature of the revived spin signal arises because the two Ω_y pulses within the magic echo sequence yield a cumulative π-rotation. The near full recovery of the spin signal at the end of the sequence indicates that the magic echo sequence did indeed work in an expected manner, and that the numerically engineered pulses serve as viable and useful substitutes for rectangular pulses that are conventionally used within NMR sequences.

Having confirmed the performance of our magic echo sequences, we now turn to our initial experiments that tested symmetric magic echo (SME) [8] sequences, which form the platform for our imaging experiments in the next subsection. SME sequences are constructed from consecutively applied magic echo sequences, where the phases for the Ω pulses, and for the rotary echos, are alternated in a systematic way. SME sequences have been demonstrated to have excellent line narrowing properties with conventional NMR experiments [8]. Furthemore, the long zero-amplitude pulse periods, and rotary echos, that the SME sequences inherit from their magic echo constituents make them particularly suitable for Fourier-transform imaging experiments. In Figure 4.11, we show the pulse

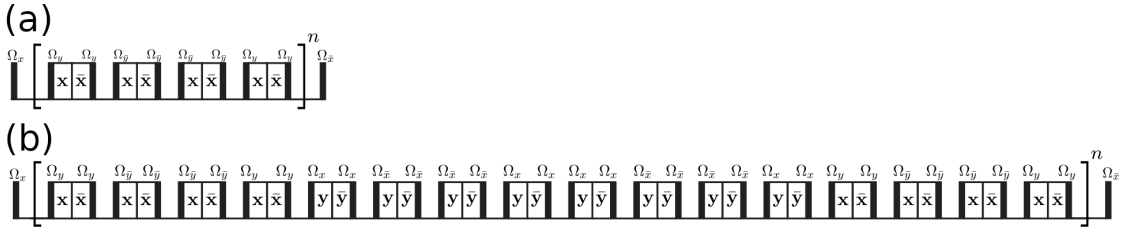

Figure 4.11: Protocols for (a) the SME4 and (b) SME16 experiments. The initial $\pi/2$ rotation creates a coherent state, after which, n blocks of SME sequences are applied with their respective $\pi/2$ rotations and rotary echo phases indicated in the figure. The experiments end with an inverse $\pi/2$ rotation that converts the coherences to populations, which are then measured through the MAGGIC protocol.

diagrams for two specific SME sequences inside the square brackets – SME4 and SME16 that comprise 4 and 16 individual magic echo blocks, respectively. Such combinations of magic echo sequences have several benefits. The symmetrization turns sequences that set $\mathcal{D}_{U^{(\gamma)}\otimes U^{(\gamma)}}(D) = \mathcal{D}_{U^{(\gamma)}}(\sigma_z) = 0$, over a single application of magic echo sequence, to ones that set $\mathcal{D}_{U^{(\gamma)}\otimes U^{(\gamma)}}(D) = \mathcal{D}_{U^{(\gamma)}\otimes U^{(\gamma)}}(D,D) = \mathcal{D}_{U^{(\gamma)}}(\sigma_z) = \mathcal{D}_{U^{(\gamma)}}(\sigma_z,\sigma_z) = 0$ over each block of two back-to-back magic echo sequences. Hence, the symmetrization generally yields longer coherence times. Furthermore, the phasing of the constituent Ω pulses within the SME sequences is chosen so that small under and over-rotations in the $\pi/2$ implementations would cancel each other out over the full sequence. Such phasing ensures that the final unitary generated for different γ values, even for imperfect Ω pulses, comes close to an identity operation.

We now present the data from the various SME experiments. Our experimental protocols for testing the performance of SME4 and SME16 sequences are given in Figure 4.11(a) and (b), respectively. Some of the resulting spin correlation measurements C, defined by Equation (4.3), after applying n back-to-back SME blocks, with particular τ values that determine the duration of rotary echos, are shown in Figure 4.12. For each τ value, the decay of C was exponential in n, and equivalently in the total length τ_e of the sequence, as is indicated in the figure. We fitted en exponential decay function $\exp(-\tau_e/T_c)$ to the data that resulted from a set of SME4 and SME16 experiments with different τ parameters, and give the best fit T_c values below:

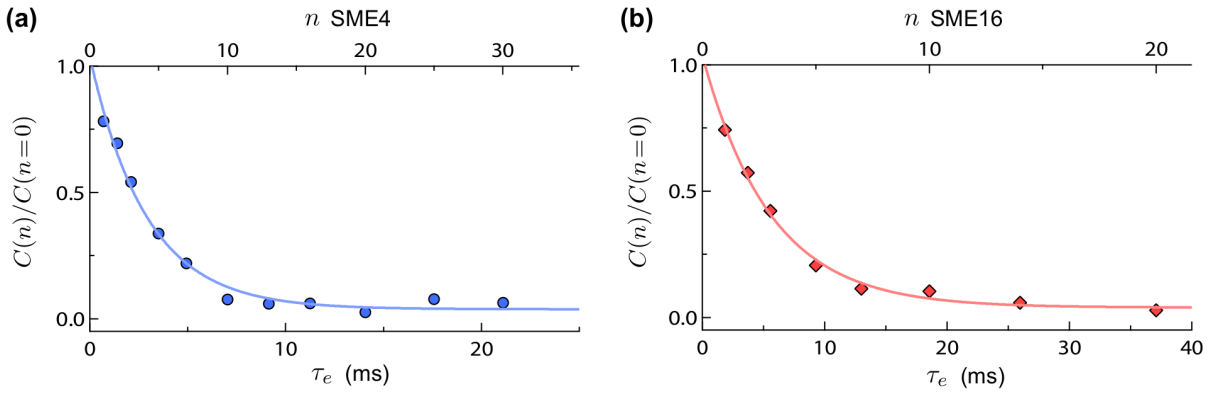

Figure 4.12: Normalized spin correlations $C(n)/C(n=0)$ for n applications of: (a) SME4 with $\tau = 25$ μs and (b) SME16 with $\tau = 15$ μs. The total duration of n blocks of SME sequences is denoted by τ_e, and displayed along the horizontal axis below the data. The data has been fitted to a decaying exponential function $C(n)/C(n=0) = \exp(-\tau_e/T_c)$, with T_c denoting the effective coherence time.

τ:	15 μs	25 μs	35 μs
T_c (SME4):	—	2.9 ms	—
T_c (SME16):	5.6 ms	4.4 ms	3.2 ms

It can be seen that the longest effective coherence time $T_c = 5.6$ ms is achieved by SME16 sequence with the shortest rotary echo period, $\tau = 15$ μs, and it constitutes an increase by a factor of 509 from the $T_2 = 11$ μs, that was measured without applying the numerical effective Hamiltonian engineering techniques, and is shown in Figure 4.10.

4.3.2 Nanoscale Magnetic Resonance Imaging Experiments

In this subsection, we build on the methods and experiments presented so far in this section to develop a protocol, and carry out nanoscale magnetic resonance imaging (MRI) experiments on our polystyrene sample. Fourier transform MRI relies on encoding the spatial coordinate of a spin to the phase value of the spin state by the use of magnetic fields, the strength of which varies across the spin ensemble, and provides a mapping between the spatial coordinate and the amount of phase accumulated.

We proceed by demonstrating how a spatially-varying phase encoding across the sample can be used for calculating the spatially-varying spin density $\rho(\vec{r})$ from ensemble averaged measurements of the total spin magnetization. The spatially varying magnetic fields used in MRI experiments could be either static $B_0 = \omega_0/\gamma_{^1\mathrm{H}}$ or Rabi field, $B_1 = 2\omega_1/\gamma_{^1\mathrm{H}}$. In some cases linear field variations are desirable, however, any monotonic spatially-varying field profile suffices, as it provides a one-to-one map between the magnetic field strength and a spatial coordinate.

For the experiments in this subsection, we used spatially varying Rabi fields, or ω_1's, that could be approximated to vary only along the height, or z-coordinate, of the silicon nanowire shown in Figure 4.1. Let $\rho(z)$ be the one-dimensional spatial density of spins in our ensemble. Ensemble averaged magnetization expectation value for spins starting from $|\uparrow\rangle$ state, and experiencing a location dependent Rabi strength $\omega_1(z)$ for some encoding period of duration τ_e, can then be written as

$$\begin{aligned}
\langle J_z\rangle(\tau_e) &= \int dz\ \rho(z) e^{-\tau_e/T_c} \langle\uparrow| e^{i\frac{\omega_1(z)\sigma_y\tau_e}{2}} \sigma_z e^{-i\frac{\omega_1(z)\sigma_y\tau_e}{2}} |\uparrow\rangle \\
&= \int dz\ \rho(z) e^{-\tau_e/T_c} \cos\left[\omega_1(z)\tau_e\right] \\
&= e^{-\tau_e/T_c} \int_0^\infty d\omega_1\ \rho(\omega_1) \cos\left(\omega_1\tau_e\right). \qquad (4.31)
\end{aligned}$$

For deducing the last row in Equation 4.31, we have used the fact that ω_1 is an invertible function of z, so that we can write $z = z(\omega_1)$ and $dz = d\omega_1 \frac{d}{d\omega 1} z$. T_c denotes the effective dephasing time. For the experiments presented in this chapter the functions $\omega_1(z)$ and $z(\omega_1)$ were first found through finite element simulations, as is indicated in Figure 4.1(c) and expanded in Appendix B.

The last row of Equation (4.31) makes it transparent that the spin density $\rho[\omega_1(z)]$ can be extracted from τ_e dependent magnetization measurements through a cosine Fourier transform:

$$\begin{aligned}
&\frac{2}{\pi}\int_0^\infty d\tau_e\ \langle J_z\rangle(\tau_e)\cos\left(\omega_1\tau_e\right) = \\
&\qquad \frac{1}{\pi}\left[\int_0^\infty d\nu \frac{\rho(\nu)T_c}{1+(\omega_1+\nu)^2T_c^2} + \int_0^\infty d\nu \frac{\rho(\nu)T_c}{1+(\nu-\omega_1)^2T_c^2}\right], \qquad (4.32)
\end{aligned}$$

which in the limit of $\mathrm{T}_c \to \infty$ converges to $\rho(\omega_1)$. For the experiments in this subsection we always chose the maximum τ_e value sufficiently smaller than the effective dephasing time T_c, such that the Fourier cosine transformation of the ensemble averaged spin signal

$\langle J_z \rangle(\tau_e)$ corresponded approximately to the one-dimensional spin density $\rho[\omega_1(z)]$ within the sample.

Equation (4.32) immediately suggests an experimental protocol for measuring the one-dimensional spin density of the sample. Assuming that the spin location dependent read-out field gradient $g(\vec{r})$ is approximately a function of z only, and hence a function of ω_1 only, the spin correlation function defined by Equation (4.3), measured through the MAGGIC protocol, can be rewritten in the limit of $\frac{\tau_m}{\tau_0} \ll 1$ as

$$C(\tau_e) = \frac{1}{2}\mu^2 e^{-\tau_e/T_c} \int d\omega_1 \; \rho(\omega_1)\, g^2(\omega_1) \cos(\omega_1 \tau_e)\, \chi(\omega_1). \tag{4.33}$$

Consequently, measuring $C(\tau_e)$ for a number of τ_e values $\{0, \tau_e^{\max}/N_{\text{encode}}, ..., \tau_e^{\max}\}$, such that $\tau_e^{\max} \ll T_c$, where T_c is the proton spin dephasing-time under continuous Rabi drive, and applying a discrete Fourier cosine transformation to the array

$$\{C(0), C(\tau_e^{\max}/N_{\text{encode}}), ..., C(\tau_e^{\max})\} \tag{4.34}$$

will yield

$$\tilde{C}(\omega_1) = \frac{1}{\pi}\mu^2 \rho(\omega_1) g^2(\omega_1)\, \chi(\omega_1), \tag{4.35}$$

for $\omega_1 = \{0, \frac{1}{2\tau_e^{\max}}, ..., N_{\text{encode}} \frac{1}{2\tau_e^{\max}}\}$. Because $g(\omega_1)$ is a known function from finite element simulations, that are shown in Appendix B, and the high-pass filter $\chi(\omega_1)$ is determined by the AFP properties, which are chosen such that $\chi(\omega_1) \approx 1$ for an ω_1 range of interest, we can extract $\rho(\omega_1)$, and correspondingly $\rho(z)$, for that ω_1 range of interest. This observation forms the basis for all of the experiments that will be presented here.

It is clear that the imaging spatial resolution, that results from measuring $C(\tau_e)$, is dependent on the maximum encoding time $\tau_e^{\max}$, which itself depends on the dephasing time T_c, under the particular experimental protocol, and the spin detection sensitivity of the experiment since the $C(\tau_e^{\max})$ measurement has to be carried out with sufficient signal-to-noise ratio. For the initial spin density measurements, an experimental procedure that followed exactly the steps described above was used, with the results given in Figure 4.3(a). It can be seen from that figure that the ω_1 resolution under such protocol was limited to $\sim$150 kHz, which corresponds to z-axis spatial resolution of $\sim$10 nm near the tip of the nanowire and a resolution of $\sim$20 nm at 50 nm above the tip.

The imaging resolution for the experiments presented in Figure 4.3(a) was determined by T_c, and coincides with the imaging resolution in a previous experiment by Nichol *et*

al [85], that was conducted on essentially the same apparatus. In order to increase the imaging resolution, an increase of the effective depahasing time was necessary, which in turn required effective removal of dipolar couplings and chemical shifts for the sample spins. As we mentioned already, SME sequences have been shown to achieve effective removal of unwanted Hamiltonian terms in the case of conventional NMR experiments, and are therefore ideal candidates for Fourier-transform imaging experiments. In addition, with the previous subsection we successfully demonstrated the use of numerically engineered control sequences within the SME sequences, while maintaining their spin coherence time increasing properties. Consequently, we employed particular modified SME sequences for imaging experiments of significantly higher spatial resolution, that will now be discussed for the rest of this section.

By introducing an asymmetry to the rotary echo periods, such that the period of $\mathbf{x}$ is $2\tau + \Delta T$ long and the period of $\bar{\mathbf{x}}$ is $2\tau - \Delta T$, as is shown in Figure 4.13(a), we can exploit the long coherence times demonstrated in the previous subsections to improve on the imaging resolution that makes use of Rabi field gradients. The resulting phase encoding under such sequence is exactly equivalent to the phase encoding scheme described above. The Ω_y and Ω_x $\pi/2$ rotations appearing before the $\mathbf{x}$ drive effectively convert it to a $\mathbf{y}$ drive, just as above, and by keeping the total SME4 sequence length fixed a varying ΔT has the following correspondence to the encoding period above: $\tau_e = 8\Delta T$. The maximum possible encoding period τ_e within a single SME4 sequence with a fixed τ value is 16τ.

For the experiments that will be presented here, and that make use of the asymmetric SME4 sequence described above, the nanowire was slightly repositioned to be exactly aligned with the CFFGS centre, which increased the maximum $\omega_1/(2\pi)$ experienced by the sample spins from 1.2 MHz to 1.6 MHz, hence, pulse #2 was used instead of pulse #1 as before. Pulse #2 is 7.5 μs long and targets an $\omega_1/(2\pi)$ range of 0.9 MHz to 1.75 MHz. Pulse #2 was incorporated into a SME4 sequence with $\tau = 15$ μs, yielding a total sequence length of 420 μs, and a maximum Rabi field encoding time of 240 μs. Three separate imaging experiments were conducted using the modified SME4 sequence: two experiments measured $C(\tau_e)$ for $N_{\text{encode}} = 21$ different τ_e values $\{0\ \mu\text{s}, \frac{6.7}{21}\ \mu\text{s}, ..., 6.7\ \mu\text{s}\}$, and one measured $C(\tau_e)$ for $N_{\text{encode}} = 31$ different τ_e values $\{0\ \mu\text{s}, \frac{20}{31}\ \mu\text{s}, ..., 20\ \mu\text{s}\}$, the data for these measurements is given in Figure 4.13(b) denoted by the red diamonds, the blue circles and the green triangles, respectively. The frequency resolution for the experiments using $\tau_e^{\max} = 6.7$ μs was $\Delta\omega_1/(2\pi) = 1/(2 \times 6.7\ \mu\text{s}) = 75$ kHz, and for the experiments using $\tau_e^{\max} = 20$ μs it was $\Delta\omega_1/(2\pi) = 1/(2 \times 20\ \mu\text{s}) = 25$ kHz.

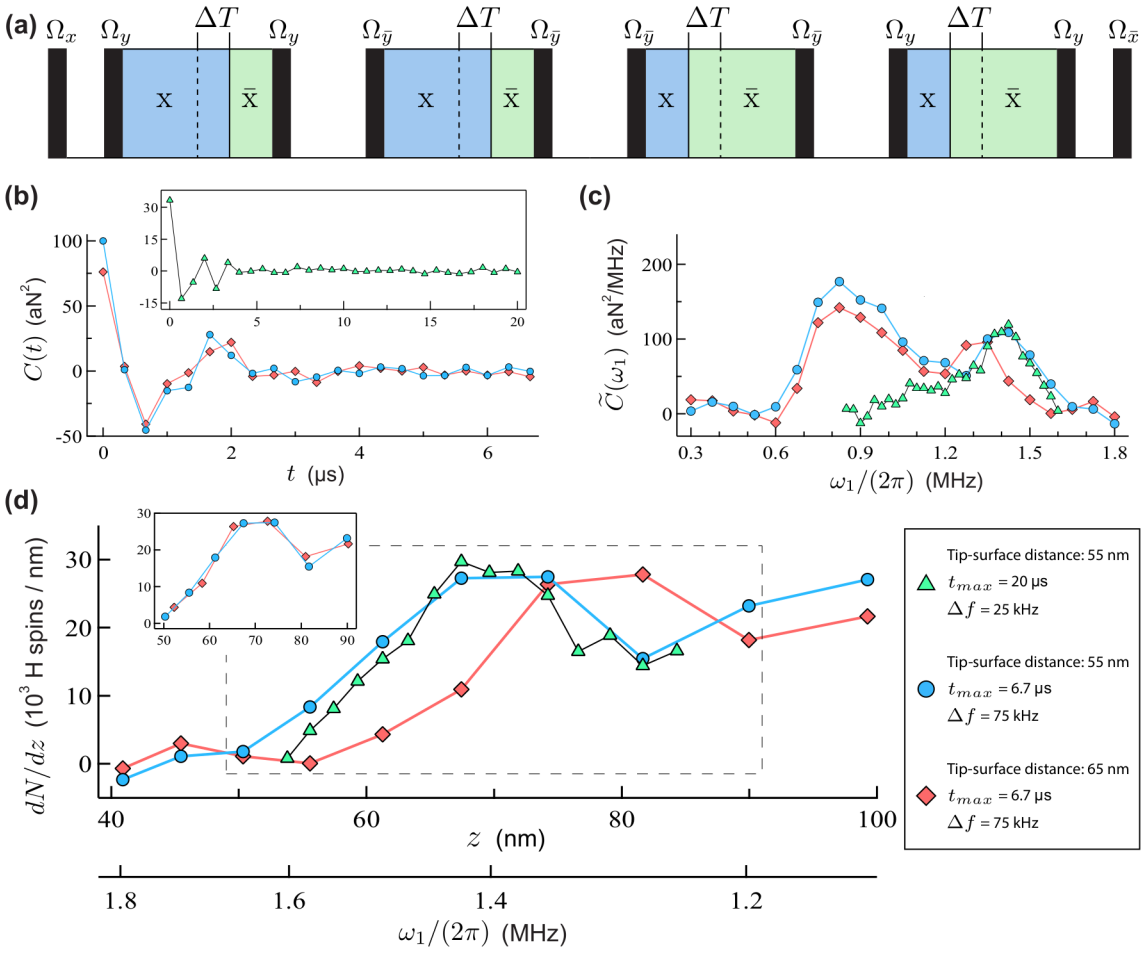

Figure 4.13: (a) Pulse diagram for the modified asymmetric SME4 sequences used for one-dimensional imaging experiments in this figure. ΔT determines the asymmetry offset for the rotary echos. (b) Spin correlation function $C(t)$ measured with a single modified SME4 sequence for various $t = 8\Delta T$ values. The red diamonds and the blue circles correspond to 75 kHz resolution data at nanowire tip-CFFGS surface distance of 55 nm and 65 nm, respectively. (Inset) The green triangles correspond to 25 kHz resolution data for nanowire tip-CFFGS distance of 55 nm. (c) Fourier transformed spin correlation functions $C(t)$. The absence of force signal from low ω_1-value spins arises because of the high-pass filter discussed in Section 4.1.2. (d) One-dimensional proton spin density in the polystyrene sample as a function of z measured from the top surface of the CFFGS. The nanowire tip-CFFGS distances are the same as in (b). (Inset) 75-kHz resolution data, where the 65 nm trace, denoted by the red diamonds, has been shifted left by 9 nm.

The measurements shown in Figure 4.13 were carried out using time-proportional phase increment procedure, which meant phase shifting the last $\Omega_{\bar{x}}$ pulse in Figure 4.13(a) by $\alpha(\tau_e) = \Delta\omega\tau_e$ with respect to the initial Ω_x pulse. Such modification to the sequence replaces the $\cos(\omega_1\tau_e)$ term in Equation (4.33) by $\cos[(\omega_1 - \Delta\omega)\tau_e]$, and turns its corresponding cosine Fourier transform into $\frac{1}{\pi}\mu^2\rho(\omega_1+\Delta\omega)g^2(\omega_1+\Delta\omega)\chi(\omega_1+\Delta\omega)$. Therefore, by choosing an appropriate $\Delta\omega$ value, the phase shifting enabled imaging the spin density $\rho(\omega_1)$ for the spins experiencing the highest Rabi drives. For the 25 kHz and the 75 kHz resolution data we picked $\Delta\omega/(2\pi) = 850$ kHz and $\Delta\omega/(2\pi) = 300$ kHz, respectively, which yielded corresponding $\tilde{C}$ values for 0.85 MHz $\leq \omega_1/(2\pi) \leq$ 1.6 MHz and 0.3 MHz $\leq \omega_1/(2\pi) \leq$ 1.8 MHz.

The time domain imaging data in Figure 4.13(b) was measured in units of force squared, which were deduced from the terminal amplitude of the nanowire mechanical resonator knowing its stiffness k and it quality factor Q. Figure 4.13(c) displays the discrete cosine Fourier transform $\tilde{C}\omega_1$ of the time domain data. Because of the time-proportional phase increment procedure employed in collecting the data, $\tilde{C}(\omega_1)$ is shifted up in frequency by $\Delta\omega$. To prevent aliased spin signal from sample spins experiencing ω_1 values lower than $\Delta\omega$ we used a high-pass filters $\chi(\omega_1)$ introduced Section 4.1.2, which suppressed any spin correlations arising from sample spins experiencing low Rabi strengths. The filters are labelled as $\chi_{25}(\omega_1)$ and $\chi_{75}(\omega_1)$ for the 25 kHz and 75 kHz resolution experiments, and are given by Equation (4.8) and Equation (4.7), respectively. We plot $\chi_{75}(\omega_1)$ in Figure 4.14(a) along with the imaging data collected with the silicon nanowire tip located 55 nm above the constriction. $\chi_{25}(\omega_1)$ is plotted in Figure 4.14(b) along with the 25 kHz resolution data from Figure 4.13. In Figure 4.14(c), we also demonstrate that multiplying the 75 kHz resolution data by $\chi_{25}(\omega_1)/\chi_{75}(\omega_1)$ closely matches the measured 25 kHz resolution data.

Finally, in Figure 4.13(d), the $\tilde{C}(\omega_1)$ function is converted to one-dimensional spin density $\rho(z)$ by using the mapping between $\omega_1(z)$ and $g(z)$ derived from finite element simulations. We display both coordinates ω_1 and z on the horizontal axis of the figure. The vertical axis in Figure 4.13(d) corresponds to one-dimensional spin density dN/dz in the polystyrene sample. The z-coordinate in that figure denotes vertical distance from the CFFGS device. It can be seen that for the spins nearest to the nanowire tip, i.e., experiencing the highest ω_1 values, the 25 kHz resolution data corresponds to a one-dimensional imaging resolution with the slice thickness below 2 nm, which increases to just above 2 nm at 30 nm above the nanowire tip. Such result marks a five-fold gain from a previous nanoscale imaging experiment conducted on essentially the same experimental setup by Nichol *et al* [85] without the numerical control techniques developed and implemented within this thesis.

The imaging resolution for Nichol *et al* [85] was set by the effective dephasing time under

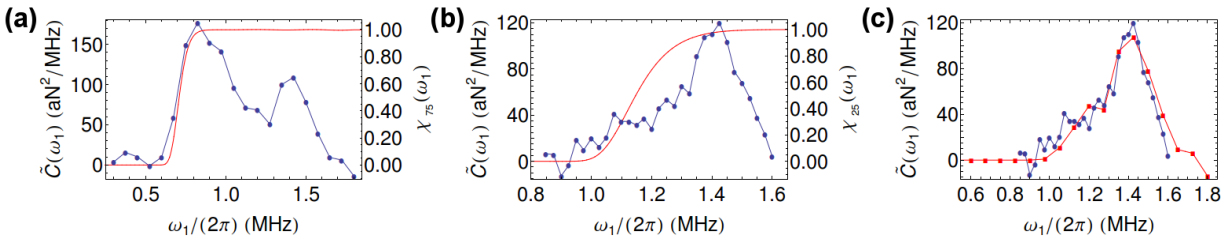

Figure 4.14: (a) Vertical axis on the left: $\widetilde{C}(\omega_1)$ for the 75 kHz resolution data from Figure 4.13 for nanowire tip-CFFGS surface separation of 55 nm (blue). Vertical axis on the right: $\chi_{75}(\omega_1)$ as defined by Equation (4.7) (red). (b) Vertical axis on the left: $\widetilde{C}(\omega_1)$ for the 25 kHz resolution data from Figure 4.13 for nanowire tip-CFFGS surface separation of 55 nm (blue). Vertical axis on the right: $\chi_{25}(\omega_1)$ as defined in Equation (4.8) (red). (c) $\widetilde{C}(\omega_1)$ for the 25 kHz resolution data in (b) (blue), $\widetilde{C}(\omega_1)\chi_{25}(\omega_1)/\chi_{75}(\omega_1)$ for the 75 kHz resolution data in (a) (red).

Rabi drive, whereas we were able to utilise the elongated dephasing time under SME4 protocol demonstrated by the previous subsection. As we said, the maximum available encoding time for the sequence we used for performing the experiments in Figure 4.13 was $\tau_e^{\max} = 240\ \mu$s, which would correspond to an imaging slice width of $\Delta\omega_1/(2\pi) = 1/(2 \times 240\ \mu\text{s}) = 2.1$ kHz, or $\sim$0.17 nm, near the tip of the nanowire. We did not choose to image the sample with such resolution since the averaging times for measuring $C(\tau_e)$ with sufficient signal-to-noise ratio would have become impractical, and also because we did not expect the amorphous polystyrene sample, used in the measurements, to have significant features over sub-nanometre length-scales. Nevertheless, such sub-nanometre imaging slice widths for our experiments were not prevented by the spin dephasing time as they were in [85]. In the case of our experiments, sub-nanometre imaging slices would have been attainable for sufficiently long averaging times, or in the presence of higher spin polarization that would have increased the spin signal strength for the experiment.

Our imaging experiment was certified by repeating the 75 kHz resolution experiment twice – once with the nanowire-CFFGS distance being 55 nm and once after moving the nanowire 10 nm vertically upwards using a calibrated piezoelectric actuator, the two experiments correspond to the blue circles and the red diamonds in Figure 4.13, respectively. It can be seen in Figure 4.13(d) that the spin density image of the translated sample has indeed shifted along the z-axis. A least square fitting for the two data sets puts an estimate for the shift at 9 ± 2 nm, we display that fit in the inset of Figure 4.13(d). This procedure verifies that our finite element field simulations, as well as the assumptions about the ω_1

and g profiles being functions of z-coordinate only, were valid. The coherent averaging performance of the modified SME4 sequences was also checked by applying two modified SME4 sequences back-to-back – one with a positive $\Delta T = 3.12$ μs, and one with an opposite ΔT which cancelled out the Rabi strength dependent rotations of the first one. The force correlation measured after the application of the two modified SME4 sequences was identical to the one measured after two unmodified $\Delta T = 0$ SME4 sequences. This implies that introducing the asymmetry into the rotary echos within the sequences did not degrade their dipolar Hamiltonian averaging properties, and hence, the sequences are perfectly suitable for imaging experiments the way they were implemented.

4.4 Simulations

With this section we provide complementary numerical studies of the control sequences that we have presented throughout this chapter. To demonstrate the necessity for including Dyson terms into our control engineering, we compare the performance of pulse #1 with that of pulse #3, which was optimized without the Dyson term targets. We will show that despite the fact that pulse #3 is roughly half the length of pulse #1, it performs substantially worse in numerical simulations with an eight spin network. For our simulations, we compute the Hamiltonian evolution for eight proton spins under dipolar couplings, as well as various Rabi strengths and resonance offset/chemical shift terms. Such simulations provide a good way for assessing the performance of the sequences, as the effects of experimental distortions and transients are totally removed. The ability to change the strength of various Hamiltonian terms also enables us to easily discern between different contributors to pulse performance. Here, such contributors might be Rabi field dispersion, resonance offesets/chemical shifts, dipolar evolution or any combination of the former. Disentangling the effects for such a variety of control performance degrading factors experimentally is a challenging task, although, the merits of numerical simulations also come with their own limitations. The main disadvantage of the numerical simulations is the severely restricted Hilbert space size, hence, we can only expect the coherence times found in our eight spin simulations to put an upper bound for the experimental results. In addition to the simulations in this section, in Appendix C, we also simulate the resulting Bloch sphere trajectories for select spins under pulse #1 and pulse #3 to visually illustrate the action of the control sequences.

The eight spin numerical simulations are performed using a nuclear spin dipolar network representing the atomic structure of the styrene molecule. The vectors $\{\vec{r}_{ij}\}$ connecting the eight protons in a single molecule are given in Section 4.1.3. Here, we ignore the ^{13}C

nuclei in the molecule. Because the sample used in our experiments was an amorphous sample, we average the simulation results over a number of possible molecular orientations. A single such orientation is denoted by a unit vector $\hat{\zeta}$, which is taken to point along the direction of the B_0 field in the coordinate system $\{\vec{r}_{ij}\}$ are evaluated in. By extending Equation (4.9) in Section 4.1.3, we can express the dipolar Hamiltonian for an eight spin network for a particular molecular orientation $\hat{\zeta}$:

$$\frac{H_{\text{dipolar}}(\hat{\zeta})}{\hbar} = \sum_{i=1}^{7}\sum_{j=i+1}^{8}\frac{\mu_0\gamma_p^2\hbar}{4\pi\left|\vec{r_{ij}}\right|^3}\frac{1}{2}\left(1-3\left[\frac{\vec{r}_{ij}.\hat{\zeta}}{|\vec{r}_{ij}|}\right]^2\right)\frac{1}{4}\left(\sum_{k\in\{x,y,z\}}\sigma_k^{(i)}\otimes\sigma_k^{(j)}-3\,\sigma_z^{(i)}\otimes\sigma_z^{(j)}\right). \tag{4.36}$$

Since our experimental results indicated a broad distribution of resonance offsets, we also include a corresponding term $\delta\omega$ into our simulations. The eight spin unitaries for a particular pulse, determined by $\alpha \in \mathrm{M}_{2,N}(\mathbb{R})$, are functions of γ, $\hat{\zeta}$ and $\delta\omega$:

$$U_{\text{VIII}}^{(\gamma,\delta\omega,\hat{\zeta})}(T) = \prod_{j=1}^{N}\exp\left(-i\left[2\pi\gamma\left(\alpha_{1,j}J_x+\alpha_{2,j}J_y\right)+\frac{H_{\text{dipolar}}(\hat{\zeta})}{\hbar}+\delta\omega J_z\right]\Delta T\right), \tag{4.37}$$

with $J_x = \sum_{i=1}^{8}\sigma_x^{(i)}/2$, $J_y = \sum_{i=1}^{8}\sigma_y^{(i)}/2$ and $J_z = \sum_{i=1}^{8}\sigma_z^{(i)}/2$. For our simulations, we calculate the eight spin z angular momentum expectation value $\langle 2J_z\rangle_{\text{VIII}}$ after n applications of the sequence as a weighted average over experimentally determined distributions of resonance offsets prob $(\delta\omega)$ and Rabi strengths prob (γ) given in Figure 4.3. In addition, we also include an equiprobable average of 50 random orientations $\{\hat{\zeta}\}$ that sample the surface of a unit sphere into the definition of

$$\langle 2J_z\rangle_{\text{VIII}}(n) = \sum_{\gamma\in\Gamma}\sum_{\delta\omega\in\Delta}\sum_{\hat{\zeta}\in\Xi}\frac{\text{prob}\,(\gamma)\;\;\text{prob}\,(\delta\omega)}{50}\langle\uparrow|^{\otimes 8}\left(\left[U_{\text{VIII}}^{(\gamma,\delta\omega,\hat{\zeta})}(T)\right]^{\dagger}\right)^{n}2J_z\left[U_{\text{VIII}}^{(\gamma,\delta\omega,\hat{\zeta})}(T)\right]^{n}|\uparrow\rangle^{\otimes 8}. \tag{4.38}$$

We compare the performance of pulse #1 and pulse #3 by evaluating $\langle 2J_z\rangle_{\text{VIII}}$, given by Equation (4.38), for either of the pulses using 250 back-to-back applications of the pulses. The results are given in Figure 4.15. It can be seen that pulse #3 performs about 7.1 times worse, despite being nearly two times shorter. Importantly, the performance of pulse #3 is sufficiently bad that it would not have enabled the experiments discussed in the previous section. Separate simulations, not presented here, reveal that the signal decay

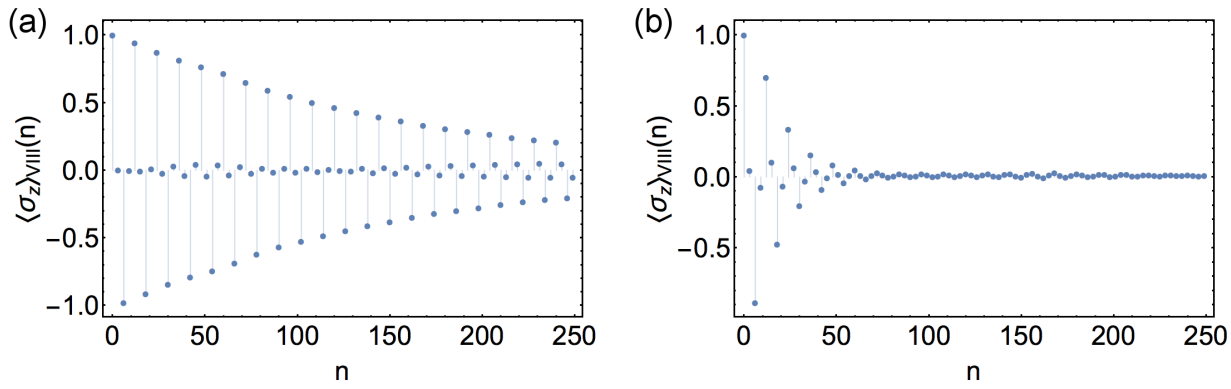

Figure 4.15: (a) Eight spin simulation results for $\langle 2J_z \rangle_{\mathrm{VIII}}(n)$, defined by Equation (4.38), using pulse #1 and (b) simulations for $\langle 2J_z \rangle_{\mathrm{VIII}}(n)$ using pulse #3. $\langle 2J_z \rangle_{\mathrm{VIII}}(n)$ are quoted as a function of n – the number of pulses applied.

for pulse #3 was mostly due to dipolar evolution – the removal of the resonance offsets for pulse #3 increases the effective e^{-1} time by a factor of 1.3; the removal of resonance offsets for pulse #1 increases the e^{-1} time by a factor of 1.5.

4.5 Conclusions and Outlook

With the work in this chapter we have provided an experimental demonstration for the applicability and usefulness of numerically engineered control sequences that are designed to minimize a set of Dyson terms. Importantly, we have managed to bring together the high-sensitivity spin detection, enabled by the high field gradients generated by CFFGS, and high-fidelity spin control, enabled by the pulse engineering methods developed in the previous chapter, to form a powerful nanoscale magnetic resonance spectroscopy and imaging platform. Furthermore, by incorporating our control sequences into existing SME sequences, we have demonstrated that we can adapt existing NMR pulse sequences to be used on our platform. With the adapted SME sequences we demonstrated a 500-fold nuclear spin coherence time enhancement for a densely packed proton sample, which is comparable to some of the best conventional NMR experiments [8]. Such adaptation of existing NMR sequences could enable a variety of experiments on nanoscale spin ensembles, such as atomic scale NMR diffraction [72] as well as multi-dimensional (correlation) spectroscopy. Both of the experiments, if successfully undertaken, could prove to be transformative.

In Section 4.3.2, we argued that the very high field gradients, and long coherence times, that we demonstrated, provided us with a technical ability to encode Ångstrom scale spatial information to the phase value of our nuclear spin ensemble. For a previous imaging experiment conducted on a similar experimental setup by Nichol *et al* [85] the resolution was limited to 10 nm, due to short coherence times. On the other hand, the one-dimensional slice thickness for the imaging experiments presented here was chosen to be $\sim$2 nm, because of spin detection sensitivity limitations, and also due to lack of sample features, and thereby a lack of non-zero Fourier components, at shorter length scales. The imaging resolution, for a given measurement period, could be greatly improved by utilizing the fact that electron spin polarization in our B_0 field and temperature is 12%. Hence, a perfect transfer of spin polarization from the electrons to the nuclei, using one of the many dynamical nuclear polarization schemes, would enhance our signal levels by a factor of nearly a 100. The nuclear spin polarization for our sample could be even further enhanced, through the aforementioned protocol, by operating at lower temperatures. Finally, the techniques presented in this chapter can be readily extended to three-dimensional imaging by fabricating additional modulation coils near the CFFGS for applying gradients in three directions.

Chapter 5

Phosphorus Donor Defects in Silicon

With this chapter we give a brief overview of phosphorus donor defects in silicon at low temperatures. We focus primarily on the magnetic resonance properties of the ^{31}P nuclear spin, and the spin of the electron that becomes bound to the defect site, at low temperatures. We will argue that the so-called semiconductor vacuum that encases the ^{31}P nuclear spins in high-purity isotopically enriched ^{28}Si crystals, makes the nuclear spins an excellent candidate as a platform for efficient, high-precision magnetometry. That is because the nuclear spins see very little local magnetic field variations that would dephase them. One crucial component for exploiting the nuclei for magnetometry applications is an ability to efficiently polarize them. We will describe some existing methods for achieving nuclear hyperpolarization, and focus on non-resonant optical pumping with above band gap light, which is known to yield hyperpolarization [77, 38]. Nevertheless, the microscopic working mechanism for such hyperpolarization scheme is yet to be experimentally confirmed. We summarize two explanations, that exist in the literature, and set the scene for our work in the next chapter, where we introduce a microscopic effective Hamiltonian model describing the non-resonant optical pumping hyperpolarization scheme, and present extensive experimental tests of our model.

This chapter contains significant original contributions by the author.

5.1 Neutral Phosphorus Defects

5.1.1 Wave Function and Orbital Levels

Phosphorus defects are substitutional defects that appear when a single silicon atom in the lattice is replaced by a phosphorus atom. Silicon crystal has a diamond cubic lattice structure, such that each silicon site has a full tetrahedral point group symmetry [127]. Tetrahedral point group, or T_d group is a discrete group with 24 elements, in Appendix E, we give an explicit irreducible matrix representation for the group. Phosphorus defect atom donates its extra valence electron to the silicon crystal, as well as deforms the atomic potential seen by the electrons. Such deformed atomic potential gives rise to a localized electronic state with biding energy, referenced to the bottom of the bulk crystal conduction band, of 45.3 meV at the site of the defect [2]. This shallow potential is not strong enough to localize electrons at room temperature since at 300 K 45.3 meV $< k_B T$, therefore, doping silicon with phosphorus atoms yields an excess of negatively charged carriers at room temperature. Nevertheless, at temperatures below the insulating temperature of $\sim$35 K, the excess electrons become localized at defect sites for silicon crystals doped with phosphorus at moderate concentrations of $< 10^{17}$ cm^{-3}. In Figure 5.1, we give an illustration of a defect electron localized to a phosphorus site in a silicon lattice. The phosphorus atom - localized electron system can be thought of as a hydrogen atom with a lattice point symmetry and an effective mass for the electron. In Figure 5.2, we show the orbital energies for the electron in that hydrogenic potential, the $1s$ ground state of the electron wave function has T_d point group symmetry and Bohr radius of approximately 2 nm [68, 2, 122]. Finally, the indirect band gap of silicon at low temperatures is 1.172 eV [15, 127], which corresponds to infra-red (IR) 1058 nm wavelength.

5.1.2 Magnetic Resonance Properties

Silicon has three stable isotopes ^{28}Si, ^{29}Si and ^{30}Si, the natural abundances of which are 92.23%, 4.67% and 3.1% [37], and nuclear spin numbers of 0, 1/2 and 0, respectively. For the work presented in this thesis, we used an isotopically enriched ^{28}Si crystal which forms a non-magnetic environment around the ^{31}P defect nucleus and the electron, it is

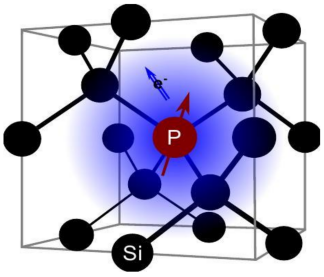

Figure 5.1: Illustration of a neutral phosphorus donor defect at low temperatures. ^{31}P atom substitutes a single silicon atom in the silicon crystal lattice, whereas a localized, represented by the blue cloud, surrounds the defect site.

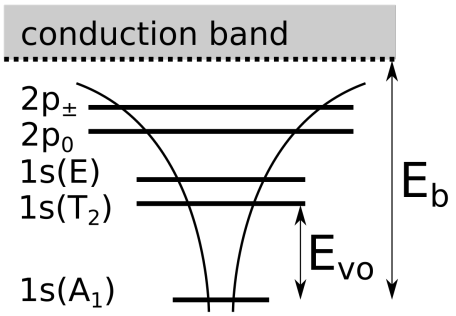

Figure 5.2: Orbital energy levels for a phosphorus donor defect electron in silicon. $E_{\mathrm{vo}} = 12.95$ eV denotes the energy splitting between the orbital ground state and the first excited states. The $1s$ ground state of the electron lies $E_{\mathrm{b}} = 45.3$ meV below the bottom of the conduction band.

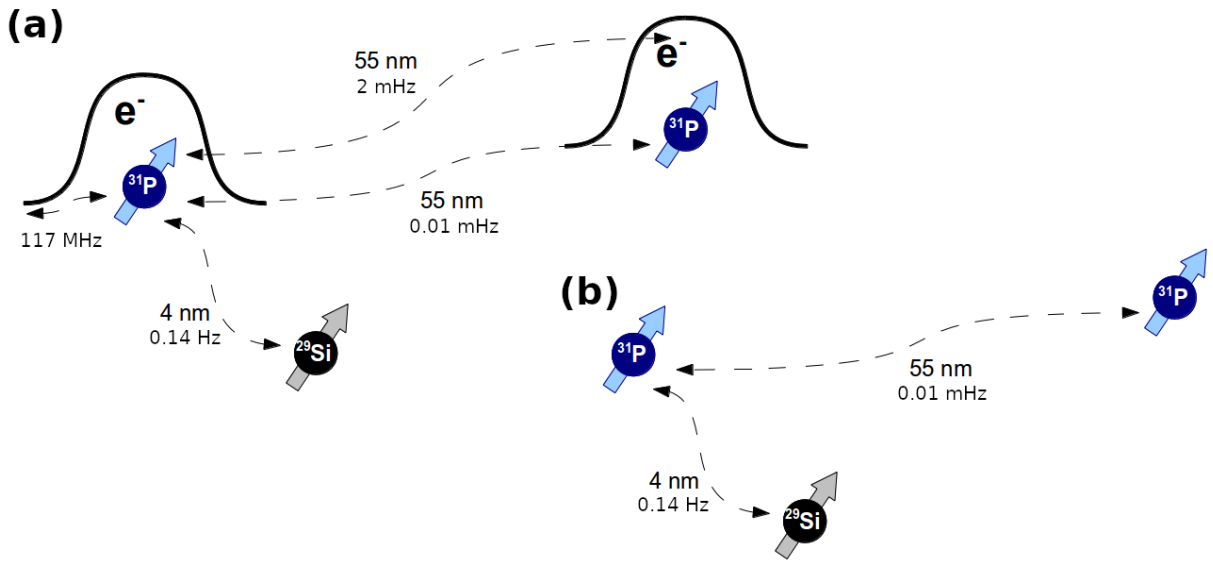

Figure 5.3: Approximate mean distances, and the corresponding spin interaction strengths, for nearest neighbour defect spins in silicon for 10^{15} cm^{-3} concentration in 46 ppm isotopically enriched ^{28}Si crystals: (a) for neutral donor defects, and (b) for ionized defect sites. The interaction strengths give an approximate idea for the ^{31}P nuclear spin coherence times arising from these interactions.

therefore sometimes referred to as a semiconductor vacuum [106, 98]. Because the mean nearest neighbour distance for two defects at low doping concentrations of 10^{15} cm^{-3} is approximately 55 nm, the electron-nuclear spin system forms a relatively isolated two-body system for magnetic resonance purposes. In Figure 5.3, we illustrate the mean nearest neighbour distances, and their respective mean interaction strengths which determine the spin decoherence rates. In the same figure, we show the mean nearest neighbour distance for ^{31}P and ^{29}Si atoms, which has been calculated for the isotopic concentration of ^{29}Si within our particular sample, containing 46 ppm ^{29}Si and ^{30}Si isotopes, with the latter being much less abundant.

The localization of the defect electron at the site of the phosphorus gives rise to a substantial hyperfine interaction between the defect electron and nuclear spins. It is shown

explicitly in Appendix E, that the T_d point symmetry of the wave function constrains the hyperfine interaction to be Fermi contact interaction:

$$H_{\mathrm{HF}} = \frac{\mu_0}{6\hbar}\gamma_e\gamma_{^{31}P}|\psi(0)|^2 \sum_{i\in\{x,y,z\}} \sigma_i^{(e)} \otimes \sigma_i^{(n)} = \frac{\mu_0}{6\hbar}\gamma_e\gamma_{^{31}\mathrm{P}}|\psi(0)|^2\sigma^{(e)}\cdot\sigma^{(n)}, \tag{5.1}$$

where $\psi(0)$ is the magnitude of the electron wave function at the site of the nucleus, $\mu_0 = 4\pi\cdot 10^{-7}$ T m A^{-1} is the vacuum permeability and $\gamma_e/(2\pi) = -28.024$ GHz T^{-1} and $\gamma_{^{31}\mathrm{P}} = 17.23$ MHz T^{-1} are the gyromagnetic ratios of the electron and nuclear spins, respectively. The upper indices for Pauli operators in Equation (5.1) denote whether the operators act on the electron, or the nuclear spin degrees of freedom. The work in this thesis was conducted at high magnetic fields of $B_0 = 6.69$ T, so that the full two-spin Hamiltonian also has strong Zeeman interaction terms:

$$H = \frac{\omega_e}{2}\sigma_z^{(e)} - \frac{\omega_{^{31}\mathrm{P}}}{2}\sigma_z^{(n)} + \frac{A_e}{4}\sigma^{(e)}\cdot\sigma^{(n)}. \tag{5.2}$$

$\omega_e/(2\pi) = \gamma_e B_0 = 187.5$ GHz and $\omega_{^{31}\mathrm{P}}/(2\pi) = \gamma_{^{31}\mathrm{P}}B_0 = 115.3$ MHz, are the electron and ^{31}P Larmor frequencies, respectively, and $A_e/(2\pi) = 117.52$ MHz is the hyperfine interaction constant that arises from Equation (5.1) [107]. Because ω_e is more than three orders of magnitude greater than A_e, the eigenstates of H in Equation 5.2 are almost exactly separable eigenstates of the Zeeman part of the Hamiltonian, i.e., $|{\downarrow_e}{\downarrow_n}\rangle$, $|{\downarrow_e}{\uparrow_n}\rangle$, $|{\uparrow_e}{\downarrow_n}\rangle$ and $|{\uparrow_e}{\uparrow_n}\rangle$. In Figure 5.4, we display the corresponding energy level diagram for the two-spin system and indicate the four dipole allowed transitions that can be driven with resonant AC magnetic fields – for both the electron and the nuclear spin we have two possible transitions at $\omega_e \pm A_e/2$ and $\omega_{^{31}\mathrm{P}} \pm A_e/2$, respectively.

At our temperatures of 1.3 K and magnetic fields $B_0 = 6.69$ T, the thermal spin polarization for free electrons is $\tanh[\Theta\omega_e\hbar/(2k_B)] = 0.998$ and for free ^{31}P nuclei it is $\tanh[\Theta\omega_{^{31}\mathrm{P}}\hbar/(2k_B)] = 0.0021$. This implies that the electron occupies its spin ground state $|{\downarrow_e}\rangle$ with near unity probability, and the nuclear spins can be driven and detected in the $\omega_{^{31}\mathrm{P}} + A_e/2 = 2\pi\cdot 174.1$ MHz manifold depicted in Figure 5.4. All of our experiments presented in the next chapter were performed on ^{31}P nuclei in that manifold.

It was show by Saeedi *et al* in [98] that for ionized phosphorus defect sites the ^{31}P nuclei in isotopically purified ^{28}Si crystals have a dephasing time of $T_2^{(n)} = 39$ min at room temperature and $T_2^{(n)} = 180$ min at temperatures of 2.2 K. These extremely long coherence

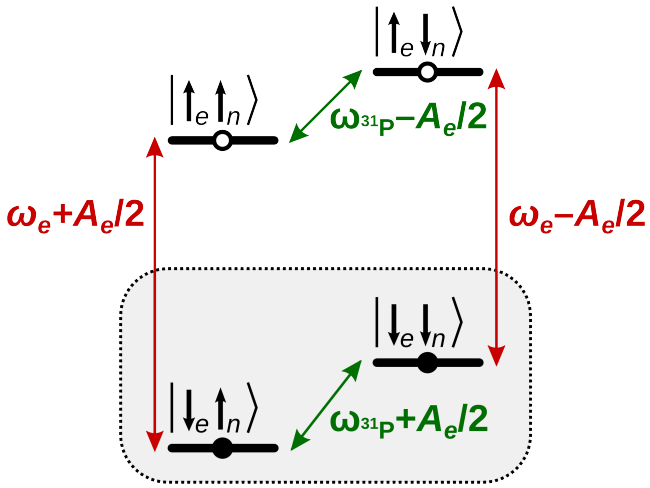

Figure 5.4: Energy levels for the neutral phosphorus defect spin degrees of freedom in high external magnetic fields. ω_e and $\omega_{^{31}\mathrm{P}}$ are the electron and ^{31}P nuclear spin Larmor frequencies, respectively. A_e is the hyperfine constant between the electron and the nuclear spins. The shaded area corresponds to electron spin $|\downarrow_e\rangle$ state manifold, which was used for all of our experiments.

times arise because of the extremely low levels of local magnetic field variations in such samples after the electrons have been removed. We illustrate that property of the ionized silicon crystals in Figure 5.3(b), where we indicate the mean nearest neighbour interactions for ionized defect nuclei. Saeedi *et al* [98] achieved the defect ionization trough resonant optical pumping of bound exciton transitions. With our work in [38] it was shown that at 1.7 K, and for $B_0 = 6.69$ T, the neutral defect dephasing time $T_2^{(n)}$ was 1.2 s, which is considerably shorter than the dephasing time for ionized nuclear spins in [98]. The $T_2^{(n)}$ in [38] was caused by the electron spin, that underwent flips between its ground and excited states as a result of its thermal relaxation $T_1^{(e)}$ process, and therefore, lead to an accumulation of random phase for the ^{31}P spin state. In Appendix D, we derive that the effective nuclear spin dephasing time $T_2^{(n)}$ as a function of electron thermal relaxation time $T_1^{(e)}$:

$$T_2^{(n)} = \frac{2T_1^{(e)}}{1-m}, \tag{5.3}$$

where m is the thermal electron spin polarization. In the case of our experiments, this implies that $T_2^{(n)} \approx 1000 T_1^{(e)}$. The nuclear dephasing times in the second range, that were observed for neutral donor defects, arise because the donor bound electron spins at high magnetic fields can have rather short $T_1^{(e)}$ values, which are often in the ms-range [80, 76].

Finally, by considering the properties of ^{31}P nuclei in isotopically purified ^{28}Si described in this section, one can imagine utilizing the nuclei as a platform for high-precision magnetometry applications, provided the one can successfully realize an efficient hyperpolarization scheme, to maximize detection efficiency, and a means for ionizing the defect sites. With this thesis we focus on the former ingredient necessary for magnetometry applications. We remark that the latter has been achieved by resonant optical ionization [98], as well as electric ionization for arsenic donors in silicon crystals [66].

5.2 Excitons in Low Temperature Silicon

In this short section, we briefly sketch the dynamics of the free exciton capture process in low temperature silicon irradiated with above band gap light, in order to then provide a comparison of the existing explanations for the non-resonant optical hyperpolarization mechanism with the next section.

The formation of bound excitons at donor defect sites in silicon irradiated with above band gap light has been known of for nearly 60 years through photoluminescence experiments [46]. Free electron-hole pairs generated by above band gap photons in silicon at low

temperatures recombine predominantly by creating free excitons – charge neutral quasi-particles travelling freely inside the crystal [127]. Since radiative lifetimes for such excitons reach milliseconds many free excitons end up forming energetically favoured bound exciton complexes at various impurity sites [101]. The simplest and most common of such complexes for phosphorus donor defects is a localized state of two electrons and a hole called D^0X [111]. The radiative recombination channel for bound excitons is greatly suppressed and they almost always recombine via the Auger mechanism [100], whereby the energy released by one of the electrons recombining with the hole is transferred to the second electron which becomes delocalized and leaves behind an ionized defect site D^+. The presence of above band gap light quickly photoneutralizes the ionized sites which capture one of the free electrons generated by the photons absorbed.

5.3 Optical Hyperpolarization of Phosphorus Donor Defects

This section provides a brief summary of the two candidate mechanisms for non-resonant optical hyperpolarization of ^{31}P defect nuclei in silicon, that have been suggested in the literature. The idea to exploit material specific properties of longitudinal electron relaxation, or $T_1^{(e)}$ process, and electron-nuclear cross relaxation, or $T_x^{(en)}$ process, for donor defects in silicon to achieve increased non-equilibrium nuclear polarization, without relying on the conventional Overhauser effect [87], dates back to the proposals of Bardeen, Slichter and Pines [91]. Since then several optical hyperpolarization schemes have been realized for the donor nuclei: optical pumping of the exciton transitions [126] and irradiation with non-resonant above-bandgap light [77, 38]. The former requires using multiple tunable narrow band lasers, while non-resonant optical schemes are particularly attractive for many experiments and devices owing to their implementation simplicity – only an above band gap light source, i.e., of wavelength 1057 nm and shorter [15], is needed.

A detailed microscopic description of the hyperpolarization mechanism underlying these non-resonant schemes, with conclusive experimental evidence supporting the description, has so far been lacking. Because non-resonant methods do not target any particular optical transition, they are notoriously hard to model and convincingly verify. There are currently two hypotheses explaining the ^{31}P spin polarization build up under above band gap light illumination at high magnetic fields and low temperatures: one relies on the idea of spin angular momentum conservation during the free exciton capture process [101], whereas the other makes a heat bath argument [77]. Definitive knowledge of the polarization mechanism would enable, amongst other things, to a priori predict the conditions necessary

for achieving fast and high nuclear spin polarization, hence, greatly assisting the design of experiments and devices harnessing such a dynamic nuclear polarization (DNP) scheme.

It has been proposed that the optical hyperpolarization of the ^{31}P nuclei results from net spin angular momentum conservation in the process of capturing free excitons when forming bound exctions. The electron spins of the D^0X ground state are antiparallel in a singlet state, due to the indistinguishability of the spatial parts of the electron wave functions [111], while the thermal electron spin polarization of 99.8% at our fields and temperatures is very close to unity. Hence, either the localized defect electron, or the free exciton electron, has to flip when creating a D^0X. Sekiguchi *et al* suggest that the formation of D^0X may only take place if the ^{31}P nuclear spin is in the $|\uparrow_n\rangle$ state, arguing that the capture must involve a nuclear spin flip to conserve net spin angular momentum [101]. The spin part of the two electron-nuclear wave function would in such case transform as $|\downarrow_e\downarrow_e\uparrow_n\rangle \rightarrow \frac{1}{\sqrt{2}}\left(|\downarrow_e\uparrow_e\rangle + |\uparrow_e\downarrow_e\rangle\right) \otimes |\downarrow_n\rangle$, and would eventually drive the ^{31}P nuclear spins into the $|\downarrow_n\rangle$ state.

Alternatively, it has been theorized that the above band gap optical illumination induced hyperpolarization may be driven via a phonon mediated Overhauser-like process [91, 77]. The photons absorbed in the crystal create free electron-hole pairs which scatter off the crystal lattice and generate non-thermal phonons that heat the phonon bath. As long as the effective heat baths yielding the electron $T_1^{(e)}$ and the electron-nuclear cross relaxation $T_x^{(en)}$ are distinct, any thermal imbalance between them will lead to a DNP process.

Most of the immediate predictions for the two proposed mechanisms are the same. They both predict the same terminal nuclear spin state of $|\downarrow_n\rangle$. The modelling presented in the next chapter also reveals that the nuclear polarization build up time for a phonon mediated process is inversely proportional to the light intensity – the same proportionality relationship as one would expect for a process involving the formation of bound excitons. Therefore, a simple nuclear magnetic resonance or an electron spin resonance experiment could not verify the correctness of one model over the other.

5.4 Conclusions

With this chapter we have introduced the physical properties of phosphorus donor defects at low temperatures along with a brief sketch of the exciton processes that take place when silicon is irradiated with above band gap light. We also outlined two possible explanations for the emergence of ^{31}P hyperpolarization under non-resonant optical pumping at low

temperatures and high magnetic fields. In the next chapter, we use previously unexplored magnetic resonance experimental design to directly observe the statistics of free exciton capture processes, and show that these processes can only very weakly correlate with the ^{31}P polarization build up.

Chapter 6

Non-resonant Optical Hyperpolarization

This chapter investigates ^{31}P nuclear spin polarization generated by optical pumping of silicon crystals with above band gap light at low temperatures and high magnetic fields. Specifically, we explore the two possible mechanisms that were outlined in the previous chapter – the phononic generation of hyperpolarization and the free exciton capture related dynamic nuclear polarization (DNP). Our study is carried out by comparing theoretical modelling and a series of experiments. We will construct a detailed model for the phononic mechanism and demonstrate its consistency with our experiments. The phononic model relies on the generation of an effective electron-nuclear Hamiltonian, however, unlike the effective Hamiltonians that have been considered so far in this thesis, the effective Hamiltonian of the hyperpolarization mechanism is generated indirectly by the light absorption in the silicon crystal, rather than by deliberate control sequence engineering.

Our argument in favour of the phonon mediated mechanism has two parts. First, we demonstrate experimentally that the ^{31}P polarization build up time constant is inversely proportional to the input laser intensity and, crucially, displays no detectable dependence on the substantial light intensity variations which extend over $\sim$mm length scales in our sample. Because sub-terahertz phonons are the only excitations with several-mm mean free paths in silicon at low temperatures, a phononic explanation for the DNP process is strongly supported. Secondly, we show experimentally that the photoionization events, which we assume to be mostly Auger ionization events, present in the illuminated crystal are largely uncorrelated with the nuclear spin polarization growth. To demonstrate the latter we design an experiment to study the ionization events in a previously unexplored manner. We exploit the fact that the reduced ^{31}P nuclear Hamiltonians for D^0, D^0X

and D^+ have distinct and far detuned Zeeman frequencies. Hence, by irradiating the silicon crystal with above band gap laser light simultaneously with a low amplitude radio frequency (RF) tone tuned on resonance with the ionized ^{31}P Zeeman splitting, we are able to deduce the average characteristics of the Auger ionization events from the modified build up dynamics induced by the saturation drive. Combining such novel experiment design with the theoretical tools of open quantum systems for analysing the results enables us to demonstrate that the Auger events, and hence the free exciton capture events, can only very weakly correlate with the nuclear polarization build up.

We study an isotopically purified ^{28}Si single crystal sample using inductive readout of the phosphorus donor nuclear spins. The hyperpolarization occurs by illuminating the sample with 1047 nm laser light in a magnetic field of 6.7 T at 1.3 K temperature. In this chapter, we first derive the phononic polarization transfer mechanism in terms of time varying crystal strain fields generated by photocarriers scattering off of the crystal lattice in Section 6.1. In Section 6.2, we introduce our model which predicts the nuclear polarization dynamics under simultaneous laser illumination and saturation drive assuming that the optical DNP mechanism acting is phonon mediated. We also construct a model to describe polarization transfer resulting from free exciton capture events in Section 6.3, we do that in order to contrast the predictions of the two possible mechanisms. Finally, in Section 6.5, we introduce our measurements and data, and verify the self-consistency of our model. The data also demonstrates disagreement with exciton capture dependent mechanism. We further discuss our conclusions in Section 6.6 and lay out several opportunities and extensions of the work we present in Section 6.7.

Because some of the equipment assembly steps, data collection and simulations in this chapter were not conducted by the author alone, such collaborative contributions have been clearly indicated in the footnotes of the particular section or subsection.

6.1 Phononic Polarization Transfer

In this section, we introduce a model for phononic polarization transfer by first demonstrating that populating certain phonon modes with non-thermal phonons amounts to an effective enhancement of the cross-relaxation rate which is proportional to the light intensity. The non-thermal phonons in our experiments are generated by effective photon to phonon conversion events happening inside the silicon crystal irradiated with above gap light. In the second subsection, we derive explicit expressions for demonstrating that different effective temperatures for the electron thermal relaxation $T_1^{(e)}$ process and the

electron nuclear cross-relaxation $T_x^{(en)}$ process lead to an increased non-thermal nuclear polarization.

6.1.1 Light Induced Electron-Nuclear Cross Relaxation

Based on the experimental results in [123], we assume that the hyperfine interaction between the ^{31}P nuclear spin and the localised donor electron spin in strained silicon is almost entirely Fermi contact interaction. Accordingly, we Taylor expand the hyperfine constant A_e up to second order in strain tensor ϵ at the defect site. To simplify our notation we treat ϵ throughout this chapter as a 6-dimensional vector $\vec{\epsilon} = (\epsilon_{xx}, \epsilon_{yy}, \epsilon_{zz}, 2\epsilon_{yz}, 2\epsilon_{xz}, 2\epsilon_{xy})$, where the elements of the tensor have been evaluated in the crystallographic coordinate system. The Taylor expansion yields $A_e(\vec{\epsilon}) = A_e(0)\left[1 + \left(\vec{\nabla}A_e\right)\cdot\vec{\epsilon} + \vec{\epsilon}\cdot\mathbf{H}_{A_e}\cdot\vec{\epsilon} + ...\right]$, where $\vec{\nabla}A_e$ and $\mathbf{H}_{A_e}$ are a vector and a matrix proportional to the gradient vector and the Hessian matrix, respectively. Symmetry analysis introduced in [73] constrains $\vec{\nabla}A_e$ and $\mathbf{H}_{A_e}$ to a form that gives $\left(\vec{\nabla}A_e\right)\cdot\vec{\epsilon} = W(\epsilon_{xx} + \epsilon_{yy} + \epsilon_{zz})/3$ and $\vec{\epsilon}\cdot\mathbf{H}_{A_e}\cdot\vec{\epsilon} = X\left(\epsilon_{xx}^2 + \epsilon_{yy}^2 + \epsilon_{zz}^2\right) + Y\left(\epsilon_{yy}\epsilon_{zz} + \epsilon_{xx}\epsilon_{zz} + \epsilon_{xx}\epsilon_{yy}\right) + Z\left(\epsilon_{yz}^2 + \epsilon_{xz}^2 + \epsilon_{xy}^2\right)$, where W, X, Y and Z are constants.

In the following, we will ignore the linear term $\left(\vec{\nabla}A_e\right)\cdot\vec{\epsilon}$, since most experimental measurements [123, 54] and tight binding simulations [115] for hyperfine shifts under uniaxial stress have not reported shifts linear in strain, even though small linear shifts under hydrostatic strain have recently been suggested [73]. Furthermore, and more importantly for the purposes of this chapter, it was pointed out by Nakayama and Hasegawa [84] that if the phononic hyperfine shifts are predominantly linear in strain then the electron-nuclear cross relaxation would be a single-phonon process, and its time constant would scale as $T_x^{(en)} \propto \Theta/\omega_e^2$, where ω_e is the electron Larmor frequency and Θ is the inverse temperature. Such relationship was not supported by the measurements reported in [50], hence, we proceed in accordance with Nakayama and Hasegawa [84] by assuming that cross relaxation is a two-phonon process.

In treating the electron-phonon interactions below, we make the typical adiabatic approximation [127], and assume that the electron wave-function follows the atomic motion instantaneously. The first excited orbital state for neutral phosphorus donor defects lies 11.6 meV, or 2.8 THz, above the ground state [2]; for such an approximation to hold the contributing phonons have to be of substantially lower frequency.

The time-modulated strain field $\vec{\epsilon}$ at the site of the donor defect results from the effective photon to phonon conversion events taking place at random instances when

above band gap photons are absorbed by the silicon crystal and create electron-hole pairs. We model $\vec{\epsilon}$ as a sum of elastic waves with frequencies $\{\omega_i\}$ originating from a set of randomly distributed sources inside the crystal, each denoted by $\vec{r}_j$, such that $\vec{\epsilon} = \sum_i \sum_j^{N_i} \vec{\epsilon}_i\,(\vec{r}_j)\cos(\omega_i t + \xi_{ij})$, where we assume that N_i, the number of waves emitted at frequency ω_i, with amplitudes $\{\vec{\epsilon}_i\}$ independent of $\vec{r}_j$, is proportional to the light intensity I, whereas the phase ξ_{ij} is a uniformly distributed random variable. Summing over the random phases gives $\vec{\epsilon} = \sum_i \sqrt{N_i}\vec{\epsilon}_i \cos(\omega_i t + \xi_i)$. Furthermore, substituting $\sqrt{N_i}\vec{\epsilon}_i = \sqrt{I}\vec{\varepsilon}_i$ results $\vec{\epsilon} = \sqrt{I}\sum_i \vec{\varepsilon}_i \cos(\omega_i t + \xi_i)$, $\vec{\varepsilon}_i$ being an effective mode dependent photon to phonon conversion yield constant.

Introducing strain dependence into the hyperfine part of the electron-nuclear Hamiltonian gives

$$H = \frac{\omega_e}{2}\sigma_z^{(e)} - \frac{\omega_{^{31}\mathrm{P}}}{2}\sigma_z^{(n)} + \frac{A_e(0)}{4}\left(1 + \vec{\epsilon}\cdot\mathbf{H}_{A_e}\cdot\vec{\epsilon}\right)\sigma^{(e)}\cdot\sigma^{(n)}, \tag{6.1}$$

where ω_e and $\omega_{^{31}\mathrm{P}}$ are the electron and nuclear Larmor frequencies, respectively. In writing H we have omitted the strain dependent changes to the electron g-tensor. Such changes, while potentially much greater in magnitude than the hyperfine modulation [123], would either commute with $\omega_e\sigma_z^{(e)}/2$, and would therefore average out due to their oscillatory nature, or would lead to a $T_1^{(e)}$ process, which will be included in the analysis separately. Taking the secular part of H in Equation (6.1), with respect to $H_0 = \frac{\omega_e}{2}\sigma_z^{(e)} - \frac{\omega_{^{31}\mathrm{P}}}{2}\sigma_z^{(n)}$, yields

$$H_{\mathrm{sec}} \sim eq\frac{A_e(0)}{4}\sigma_z^{(e)}\sigma_z^{(n)} + \frac{IA_e(0)}{16}\left(\sum_{\substack{\omega_i+\omega_j=\\ \omega_e+\omega_{^{31}\mathrm{P}}}} + \sum_{\substack{\omega_i-\omega_j=\\ \omega_e+\omega_{^{31}\mathrm{P}}}}\right)\vec{\varepsilon}_i\cdot\mathbf{H}_{A_e}\cdot\vec{\varepsilon}_j\; H_{\text{flip-flop}}, \tag{6.2}$$

where $H_{\text{flip-flop}} = \sigma_x^{(e)}\sigma_x^{(n)} + \sigma_y^{(e)}\sigma_y^{(n)}$ is the flip-flop Hamiltonian. All terms in Equation (6.2) commute and the term on the second line is capable of transferring polarization from the electron to the nucleus. The summations in Equation (6.2) correspond to a two-phonon absorption process and Raman scattering, respectively, their particular energy conservation conditions being $\omega_i + \omega_j = \omega_e + \omega_{^{31}\mathrm{P}}$ and $\omega_i - \omega_j = \omega_e + \omega_{^{31}\mathrm{P}}$ [84, 1]. Consequently, the cross relaxation process in high magnetic fields involves a sum over the whole phonon spectrum, and one expects $1/T_x^{(en)} \propto I$. On the other hand, the phonon mode most responsible for electron $T_1^{(e)}$, via spin-orbit coupling modulation, is $\omega_i \approx \omega_e$ since the electron g-factor experiences shifts linear in crystal strain as demonstrated by [123].

It needs to be emphasized that the treatment above is not meant to imply that the cross relaxation process is a coherent Hamiltonian process. The strain fields inside the silicon

crystal result from random photon absorption processes which are distributed uniformly and isotropically within the crystal, when averaged over phonon mean free paths. Hence, there cannot be any phase coherence for the elastic waves present at a single defect site at different times, nor at different defect sites at the same time. We presented the calculation above to highlight two aspects of the phononic relaxation processes: the electron $T_1^{(e)}$ and the cross-relaxation $T_x^{(en)}$ arise from interactions with distinct phonon baths, and for strong enough I, for which the semi-classical treatment is appropriate, $1/T_x^{(en)}$ should be proportional to I.

6.1.2 Relaxation Rates

In the last subsection, we showed that the phonon baths responsible for the $T_1^{(e)}$ and $T_x^{(en)}$ are distinct, hence, their effective temperatures can also be different. In this subsection, we derive the steady state density matrix and the equilibration rates for an electron-nuclear spin system that is weakly coupled to two distinct thermal baths at inverse temperatures Θ_1 and Θ_x, which yield a $T_1^{(e)}$ process and a $T_x^{(en)}$ process, respectively. We will show that having $\Theta_1 \neq \Theta_x$ will generally lead to an increase in the nuclear spin polarization.

Let ρ be the electron-nuclear density matrix. The Lindblad equation [10, 64], written in the frame of $H_0 = \frac{\omega_e}{2}\sigma_z^{(e)} - \frac{\omega_{31_\mathrm{P}}}{2}\sigma_z^{(n)} + \frac{A_e(0)}{4}\sigma_z^{(e)} \cdot \sigma_z^{(n)}$, corresponding to such processes is then given by

$$\begin{aligned}\frac{\partial}{\partial t}\rho = L_{T_1^{(e)},T_x^{(en)}}[\rho] = &-\frac{q}{2T_1^{(e)}}\left(\sigma_-^{(e)}\sigma_+^{(e)}\rho + \rho\sigma_-^{(e)}\sigma_+^{(e)} - 2\sigma_+^{(e)}\rho\sigma_-^{(e)}\right) \\ &- \frac{1-q}{2T_1^{(e)}}\left(\sigma_+^{(e)}\sigma_-^{(e)}\rho + \rho\sigma_+^{(e)}\sigma_-^{(e)} - 2\sigma_-^{(e)}\rho\sigma_+^{(e)}\right) \\ &- \frac{p}{2T_x^{(en)}}\left(\sigma_-^{(en)}\sigma_+^{(en)}\rho + \rho\sigma_-^{(en)}\sigma_+^{(en)} - 2\sigma_+^{(en)}\rho\sigma_-^{(en)}\right) \\ &- \frac{1-p}{2T_x^{(en)}}\left(\sigma_+^{(en)}\sigma_-^{(en)}\rho + \rho\sigma_+^{(en)}\sigma_-^{(en)} - 2\sigma_-^{(en)}\rho\sigma_+^{(en)}\right),\end{aligned} \tag{6.3}$$

where $p = \left(1 + e^{\hbar\left(\omega_e + \omega_{31_\mathrm{P}}\right)\Theta_x/k_B}\right)^{-1}$ and $q = \left(1 + e^{\hbar\omega_e\Theta_1/k_B}\right)^{-1}$ are the thermal occupation probabilities for the excited state of a two-level system at inverse temperature Θ_x and Θ_1, respectively. The various operators in Equation (6.3) are defined as $\sigma_+^{(e)} = |\uparrow\rangle\langle\downarrow| \otimes \mathbb{1}$, $\sigma_-^{(e)} = |\downarrow\rangle\langle\uparrow| \otimes \mathbb{1}$, $\sigma_+^{(en)} = |\uparrow\downarrow\rangle\langle\downarrow\uparrow|$ and $\sigma_-^{(en)} = |\downarrow\uparrow\rangle\langle\uparrow\downarrow|$. It is easy to check that

$$\rho_{en} = \left[\frac{\mathbb{1}}{2} + (2q-1)\frac{\sigma_z}{2}\right] \otimes \left[\frac{\mathbb{1}}{2} + \frac{p-q}{p(2q-1)-q}\frac{\sigma_z}{2}\right] \tag{6.4}$$

yields $\frac{\partial}{\partial t}\rho_{en} = 0$, and is therefore the steady state electron-nuclear density matrix.

We define a partial basis set $\{B_1, ..., B_4\} = \{|\downarrow\rangle\langle\downarrow| \otimes |\downarrow\rangle\langle\downarrow|, |\uparrow\rangle\langle\uparrow| \otimes |\downarrow\rangle\langle\downarrow|, |\downarrow\rangle\langle\downarrow| \otimes |\uparrow\rangle\langle\uparrow|, |\uparrow\rangle\langle\uparrow| \otimes |\uparrow\rangle\langle\uparrow|\}$ for ρ in Equation (6.3), chiefly, because here we are only interested in the time scales pertaining to the diagonal elements of the electron-nuclear density matrix. Nevertheless, it also turns out that the evolution of such matrix elements is the only non-trivial part of the evolution under Equation (6.3). Accordingly, we write $\rho(t) = \sum_{i=1}^{4} \varrho_i(t) B_i$, where $\varrho_i(t)$ denotes the ith element of a four-dimensional vector $\vec{\varrho}(t)$. Because Equation (6.3) is a linear matrix differential equation, $\vec{\varrho}(t_0 + t) = \exp(St) \cdot \vec{\varrho}(t_0)$, where $S \in \mathrm{M}_4$ has its elements given as $S_{i,j} = \mathrm{Tr}\left(B_i^\dagger \cdot L_{T_1^{(e)}, T_x^{(en)}}[B_j]\right)$. Evaluating S explicitly gives

$$
\begin{aligned}
S &= -\begin{pmatrix}
\frac{q}{T_1^{(e)}} & -\frac{1-q}{T_1^{(e)}} & 0 & 0 \\
-\frac{q}{T_1^{(e)}} & \frac{1-q}{T_1^{(e)}} + \frac{1-p}{T_x^{(en)}} & -\frac{p}{T_x^{(en)}} & 0 \\
0 & -\frac{1-p}{T_x^{(en)}} & \frac{q}{T_1^{(e)}} + \frac{p}{T_x^{(en)}} & -\frac{1-q}{T_1^{(e)}} \\
0 & 0 & -\frac{q}{T_1^{(e)}} & \frac{1-q}{T_1^{(e)}}
\end{pmatrix} \\
&= -\frac{1}{T_1^{(e)}} Q_1 \cdot \left(\mathrm{P}_2 + \frac{T_1^{(e)}}{T_x^{(en)}} \Xi\right) \cdot Q_1^{-1} \qquad (6.5) \\
&= -\frac{1}{T_x^{(en)}} Q_x \cdot \left(\mathrm{P}_1 + \frac{T_x^{(en)}}{T_1^{(e)}} \left(\Xi + \mathrm{P}_2 - \mathrm{P}_1\right)\right) \cdot Q_x^{-1}, \qquad (6.6)
\end{aligned}
$$

where $\mathrm{P}_1 = \mathrm{diag}(1,0,0,0)$, $\mathrm{P}_1 = \mathrm{diag}(1,1,0,0)$ and

$$
\Xi = \begin{pmatrix}
1-r & 0 & r & 0 \\
0 & 0 & 0 & 0 \\
1-r & 0 & r & 0 \\
0 & 0 & 0 & 0
\end{pmatrix}, \qquad (6.7)
$$

with $r = p + q - 2pq$. It can be seen that $0 \leq r \leq 1/2$ for all $0 \leq p \leq 1/2$ and $0 \leq q \leq 1/2$. The matrices Q_1 and Q_x are given as

$$
Q_1 = \begin{pmatrix}
1-q & -\frac{pq}{1-p} & -(1-q) & \frac{p(1-q)^2}{r} \\
-(1-q) & \frac{pq}{1-p} & -q & \frac{pq(1-q)}{r} \\
q & q & 1-q & \frac{(1-p)q(1-q)}{r} \\
-q & -q & q & \frac{(1-p)q^2}{r}
\end{pmatrix}, \qquad (6.8)
$$

$$Q_x = \begin{pmatrix} 0 & -\frac{pq}{1-p} & \frac{1-q}{1-r} & \frac{p(1-q)^2}{r} \\ -1 & \frac{pq}{1-p} & \frac{q-r}{1-r} & \frac{pq(1-q)}{r} \\ 1 & q & -\frac{1-q-r}{1-r} & \frac{(1-p)q(1-q)}{r} \\ 0 & -q & -\frac{q}{1-r} & \frac{(1-p)q^2}{r} \end{pmatrix}, \tag{6.9}$$

while their inverses evaluate to

$$Q_1^{-1} = \begin{pmatrix} \frac{(1-p)q}{1-r} & -\frac{(1-p)(1-q)}{1-r} & \frac{pq}{1-r} & -\frac{p(1-q)}{1-r} \\ -\frac{(1-p)q}{1-r} & \frac{(1-p)(1-q)}{1-r} & \frac{(1-p)(1-q)}{1-r} & -\frac{(1-p)(1-q)^2}{q(1-r)} \\ -\frac{(1-p)q}{r} & -\frac{(1-p)q}{r} & \frac{p(1-q)}{r} & \frac{p(1-q)}{r} \\ 1 & 1 & 1 & 1 \end{pmatrix}, \tag{6.10}$$

$$Q_x^{-1} = \begin{pmatrix} 0 & -1+p & p & 0 \\ -\frac{(1-p)q}{1-r} & \frac{(1-p)(1-q)}{1-r} & \frac{(1-p)(1-q)}{1-r} & -\frac{(1-p)(1-q)^2}{q(1-r)} \\ \frac{(1-p)q}{r} & \frac{(1-p)(q-r)}{r} & -\frac{p(1-q-r)}{r} & -\frac{p(1-q)}{r} \\ 1 & 1 & 1 & 1 \end{pmatrix}. \tag{6.11}$$

Writing S as in Equation (6.5) makes it apparent that, for $T_1^{(e)} \ll T_x^{(en)}$, we can approximate

$$\exp(St) \approx Q_1 \cdot \mathrm{diag}\left(e^{-t/T_1^{(e)}}, e^{-t/T_1^{(e)}}, e^{-r\ t/T_x^{(en)}}, 1\right) \cdot Q_1^{-1}. \tag{6.12}$$

The exponential decaying the slowest in Equation (6.12) is $e^{-r\ t/T_x^{(en)}}$, hence, for a general initial electron-nuclear density matrix $\vec{\varrho}(0)$, the steady state density matrix $\vec{\varrho}(t \to \infty) = \left(\frac{p(1-q)^2}{r}, \frac{pq(1-q)}{r}, \frac{(1-p)q(1-q)}{r}, \frac{(1-p)q^2}{r}\right)$ is approached with time constant $T_x^{(en)}/r$.

From Equation (6.4) we see that when $\Theta_x = \Theta_1 = \Theta$ the thermal polarization of electrons and nuclei is $\tanh\left[\hbar\omega_e\Theta/(2k_B)\right]$ and $-\tanh\left[\hbar\omega_{^{31}\mathrm{P}}\Theta/(2k_B)\right]$, respectively. It is also clear that offsetting Θ_x and Θ_1 can lead to nuclear polarizations equal to that of the electrons. For this work we always have $T_1^{(e)} \ll T_x^{(en)}$, therefore, such terminal non-thermal is approached with time constant

$$T_{\mathrm{DNP}} = \frac{T_x^{(en)}}{p+q-2pq}. \tag{6.13}$$

Even though, it is difficult to conclude which exact phonon modes might contribute to cross relaxation in a particular experiment, since Equation (6.2) involves a summation over the

entire phonon spectrum, it is reasonable to assume that for high enough light intensities $\Theta_x \to 0$, such that $p \to 1/2$, which implies that $T_{\mathrm{DNP}} = 2T_x^{(en)}$ that is independent of Θ_1 and q. It has also been experimentally verified that in the presence of above band gap illumination at low temperatures the electron polarization $(1-2q)$ equilibrates very close to that of the sample temperature, since p and $T_1^{(e)}$ are mostly determined by factors other than phononic relaxation – photoionization and photoneutralization, accompanied by exchange interaction with conduction band electrons all act to thermalize localized donor electron spins with photocarrier electron spin temperature [77, 25]. Therefore, for high I one expects $1/T_{\mathrm{DNP}} \propto I$ and $p > q$ which yields nuclear spin population in the $|{\downarrow_n}\rangle$ state, aligned with the thermal electron polarization.

6.2 Photoionization and Photoneutralization Dynamics

In this section, we work out the ^{31}P magnetization dynamics under simultaneous irradiation with above band gap light and RF saturation drive near the Larmor frequency of the ionized phosphorus nucleus $\omega_{^{31}\mathrm{P}}$. When the silicon sample is irradiated with above band gap light the phosphorus donor defect sites are continuously ionized and neutralized through free exciton capture processes, Auger ionization and photoneutralization that were outlined in the previous chapter. The photoionization and neutralization cycle can be divided into three distinct steps as depicted in Figure 6.1. To explain the build up dynamics under RF saturation we introduce a three state rate equation model. When modelling the capture of free excitons and photoneutralization processes we make a few assumptions. First, we assume that all capture events happen instantaneously at a given rate without changing the phosphorus spin state and independently of earlier capture events, i.e., the capture is a Poisson process. We denote $\kappa_{D^0 \to D^0 X}$ as the rate of free exciton capture, $\kappa_{D^0 X \to D^+}$ as the rate of Auger recombination and $\kappa_{D^+ \to D^0}$ as the rate of photoionization, the assumptions above are valid as long as $\kappa_{D^0 \to D^0 X} \ll \kappa_{D^0 X \to D^+}\kappa_{D^+ \to D^0}/(\kappa_{D^0 X \to D^+} + \kappa_{D^+ \to D^0})$, and the electrostatically driven capture processes happen much faster than the characteristic nuclear spin evolution time scale $(\omega_{^{31}\mathrm{P}})^{-1}$. We ignore the $\sim 10^4$-fold weaker radiative recombination channel [100] and assume that every free exciton capture is followed by Auger recombination. We assume that the bound exciton hole spin state is thermalized [65], which at our temperature and magnetic field configuration implies that the hole occupies its spin ground state.

Each step in the photoionization and neutralization cycle yields a distinct effective Hamiltonian and relaxation process operators on the ^{31}P nuclear spin. With the following

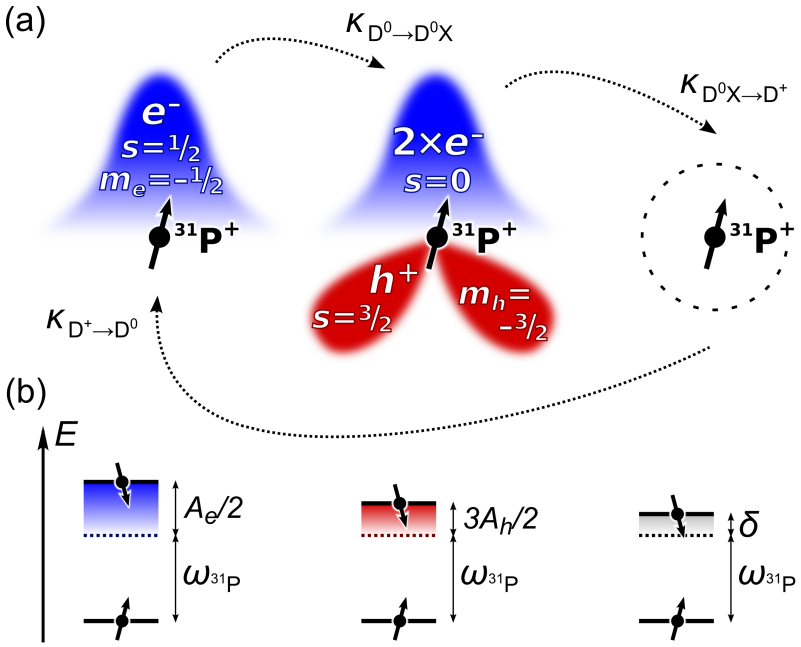

Figure 6.1: Photoionization and neutralization cycle taking place inside the silicon crystal illuminated with above band gap light. (a) Initially, the donor defect is in a neutral state D^0 with a single electron at the ^{31}P site. At a rate of $\kappa_{D^0 \to D^0 X}$ a free exciton is captured yielding a bound exciton D^0X that comprises a localized hole and two electrons in a spin singlet. The bound exciton recombines via Auger mechanism at a rate of $\kappa_{D^0X \to D^+}$ leaving behind an ionized defect site D^+. Photoneutralization returns the defect to its neutral state at a rate of $\kappa_{D^+ \to D^0}$. (b) The ^{31}P nuclear spin experiences a distinct energy level splitting for each step of the photoionization and neutralization cycle shown in (a). Given that the D^0 electron and the D^0X hole are thermalized to their spin ground states with probability close to unity, such splittings are given by a combination of bare Zeeman frequency $\omega_{^{31}\mathrm{P}}$ and either the electron hyperfine term A_e, the hole hyperfine term A_h or the paramagnetic shift δ.

we detail all of these reduced Hamiltonians and relaxation processes and derive the steady state magnetization $\mu_z^{D^0} = m_z^{D^0}(t \to \infty)$ for neutral donor spins under RF saturation drive of strength Ω in the limit of $\kappa_{D^0X \to D^+} \gg \kappa_{D^0 \to D^0X}$, provided that $\kappa_{D^0X \to D^+} \geq \kappa_{D^+ \to D^0}$, $\kappa_{D^0X \to D^+} \geq \Omega$ and $\kappa_{D^+ \to D^0} \geq \Omega$.

We say that the reduced nuclear spin density matrix ρ_D, $D \in \{D^0, D^0X, D^+\}$, during each stage of the free exciton capture and recombination process evolves under the Lindblad equation

$$\begin{aligned} \frac{\partial}{\partial t}\rho_D = L_D\left[\rho_D\right] = &-i\left[H_D, \rho_D\right] \\ &- \frac{1}{2T_1^D}\left(\sigma_-\sigma_+\rho_D + \rho_D\sigma_-\sigma_+ - 2\sigma_+\rho_D\sigma_-\right) \\ &- \frac{1}{2}\left(\frac{1}{T_2^D} - \frac{1}{2T_1^D}\right)\left(\rho_D - \sigma_z\rho_D\sigma_z\right), \end{aligned} \tag{6.14}$$

where the reduced nuclear Hamiltonian H_D, effective dephasing time T_2^D and effective longitudinal relaxation time T_1^D are specific to the particular stage of the process denoted by D. Here, we will write all Hamiltonians in the frame of the continuous saturation RF drive; the difference between the bare phosphorus frequency $\omega_{^{31}\mathrm{P}}$ and the drive frequency will be denoted by $\Delta\nu$. The raising and lowering operators are given as $\sigma_+ = |\uparrow\rangle\langle\downarrow|$ and $\sigma_- = |\downarrow\rangle\langle\uparrow|$. Because our experimental polarization measurements were relative measurements, blind to the absolute polarization of the ^{31}P spins, we have defined Equation (6.14) such that the terminal polarization is equal to one, i.e., the term on the second line of Equation (6.14) acts to make $\rho_D = |\uparrow\rangle\langle\uparrow|$ over a characteristic time scale T_1^D, while the term on the third line ensures that the off-diagonal terms of ρ_D decay with a time constant T_2^D.

We now introduce a (Bloch) basis $\{C_i\}$ for the reduced nuclear density matrix such that $\rho_D = \sum_i m_i^D C_i$, where $i \in \{x, y, z, I\}$ and m_i^D denotes the ith element of the magnetization vector $\vec{m}^D = \left(m_x^D, m_y^D, m_z^D, m_I^D\right)$. The basis elements are given as $C_i = \sigma_i/\sqrt{2}$, σ_I being a shorthand for two-dimensional identity matrix $\mathbb{1}_2$. The time evolution of ρ_D can then be found by exponentiating a four-dimensional generator matrix Λ_D, the elements of which are given as $(\Lambda_D)_{i,j} = \mathrm{Tr}\left(C_i^\dagger \cdot L_D[C_j]\right)$. In the subsequent subsections, we will evaluate the particular Λ_D for each of the three stages of the free exciton capture process.

Given the assumption that the photoneutralization, photoionization and bound exciton recombination are all instantaneous and do not affect the nuclear spin state, we proceed by adapting a method developed for modelling chemical exchange processes [78]. Accordingly, we combine the time evolution of all three magnetization vectors $\{\vec{m}^{D^+}, \vec{m}^{D^0X}, \vec{m}^{D^0}\}$

and describe this evolution by a set of coupled Bloch equations. We define a combined magnetization vector $\vec{m} \in \mathbb{R}^{3\times 4}$:

$$\vec{m} = \begin{pmatrix} \vec{m}^{D^+} \\ \vec{m}^{D^0 X} \\ \vec{m}^{D^0} \end{pmatrix}, \tag{6.15}$$

which evolves under coupled Bloch equations determined by $\Upsilon \in \mathrm{M}_3(\mathrm{M}_4)$:

$$\frac{\partial}{\partial t}\vec{m} = \Upsilon \cdot \vec{m} = \begin{pmatrix} \Lambda_{D^+} - \kappa_{D^+\to D^0}\mathrm{P}_3 & \kappa_{D^0 X\to D^+}\mathrm{P}_3 & 0 \\ 0 & \Lambda_{D^0 X} - \kappa_{D^0 X\to D^+}\mathrm{P}_3 & \kappa_{D^0\to D^0 X}\mathrm{P}_3 \\ \kappa_{D^+\to D^0}\mathrm{P}_3 & 0 & \Lambda_{D^0} - \kappa_{D^0\to D^0 X}\mathrm{P}_3 \end{pmatrix} \cdot \vec{m}, \tag{6.16}$$

where $\mathrm{P}_3 = \mathrm{diag}(1,1,1,0)$. The generator matrices $\{\Lambda_D\}$ will be explicitly evaluated in the next three subsections for the particular effective H_D, T_1^D and T_2^D in Equation (6.14). It should be noted that we set up the coupled equations in Equation (6.16) such that the nuclear spin polarization is measured relative to the full polarization of D^0 nuclei in the absence free exciton capture events, i.e., if $\kappa_{D^0 X\to D^+} = 0$ then $m_z^{D^0}(t\to\infty) = 1$ and $m_z^{D^0 X}(t) = m_z^{D^+}(t) = 0$ for all $t \geq 0$.

6.2.1 Neutral Donor Defect D^0

The reduced ^{31}P Hamiltonian H_{D^0} for a neutral donor defect, with the electron spin in its ground state, is given by

$$H_{D^0} = \Omega\frac{\sigma_x}{2} + \left(\frac{A_e(0)}{2} - \Delta\nu\right)\frac{\sigma_z}{2}, \tag{6.17}$$

where $A_e(0)/(2\pi) = 117.52$ MHz is the hyperfine constant at our temperature [107]. As we always operate in the the regime of $A_e(0) \gg \Delta\nu$ and $A_e(0) \gg \Omega$, Equation (6.17) is well approximated by $H_{D^0} \approx (A_e(0)/2 - \Delta\nu)\,\sigma_z/2$, where the RF drive term has been dropped after making the secular approximation. In addition to the Hamiltonian evolution, we also have the phononic DNP process acting to increase the nuclear spin polarization with a characteristic time scale $T_1^{D^0}$, as well as a dephasing $T_2^{D^0}$ process, hence, the Bloch

generator for the nuclear density matrix ρ_{D^0} becomes

$$\Lambda_{D^0} = \begin{pmatrix} -\frac{1}{T_2^{D^0}} & -\frac{A_e(0)}{2} + \Delta\nu & 0 & 0 \\ \frac{A_e(0)}{2} - \Delta\nu & -\frac{1}{T_2^{D^0}} & 0 & 0 \\ 0 & 0 & -\frac{1}{T_1^{D^0}} & \frac{1}{T_1^{D^0}} \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{6.18}$$

6.2.2 Bound Exciton D^0X

Because the two bound exciton electrons have identical spatial wave-functions, they have to be in a spin singlet state, which leads to a vanishing hyperfine interaction between the nucleus and the electrons [111]. The hyperfine interaction between the localized hole and the nucleus is in many experiments assumed to be insignificant, however, it could potentially be of influence to our results. It is hard to estimate its size, a reasonable upper bound would be $\sim 2\pi \times 20$ MHz from photoluminescence excitation experiments [125], while $2\pi \times 2$ MHz has been suggested in the literature [28]. Here, we denote the hole hyperfine strength by A_h. Assuming that the hole is thermalized to its spin ground state, the reduced ^{31}P Hamiltonian for a bound exciton becomes

$$H_{D^0X} = \Omega\frac{\sigma_x}{2} + \left(\frac{3A_h}{2} - \Delta\nu\right)\frac{\sigma_z}{2}. \tag{6.19}$$

Since the lifetime of bound excitons has been measured to be extremely short – $\kappa_{D^0X\to D^+} = 3.7$ MHz, corresponding to a lifetime of 272 ns [100] – we assume there to be no appreciable T_1 mechanism acting on the bound exciton nuclear state ρ_{D^0X}, yet the short T_1 time of the hole itself will introduce dephasing effects with time constant $T_2^{D^0X}$ on the ^{31}P. Accordingly, the Bloch generator for the nucleus becomes

$$\Lambda_{D^0X} = \begin{pmatrix} -\frac{1}{T_2^{D^0X}} & -\frac{3A_h}{2} + \Delta\nu & 0 & 0 \\ \frac{3A_h}{2} - \Delta\nu & -\frac{1}{T_2^{D^0X}} & -\Omega & 0 \\ 0 & \Omega & -\frac{1}{T_2^{D^0X}} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{6.20}$$

6.2.3 Ionized Donor Defect D^+

The ionized defect is modeled as a bare phosphorus spin with Zeeman splitting $\omega_{^{31}\mathrm{P}}$ that could be shifted away from $\omega_{^{31}\mathrm{P}}$ in the presence of a paramagnetic shift δ. Including this

term in our Hamiltonian for D^+ gives

$$H_{D^+} = \Omega\frac{\sigma_x}{2} + (\delta - \Delta\nu)\frac{\sigma_z}{2}. \tag{6.21}$$

Again, we ignore any T_1 effects due to the very short time the donors spend in this state. Furthermore, because the ionized defects see a very small local magnetic field we will also ignore any T_2 effects in this stage and write the evolution as purely unitary, such that

$$\Lambda_{D^+} = \begin{pmatrix} 0 & -\delta + \Delta\nu & 0 & 0 \\ \delta - \Delta\nu & 0 & -\Omega & 0 \\ 0 & \Omega & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{6.22}$$

The long T_2 assumption is supported by experimental results in [98].

6.2.4 Magnetization Dynamics

The only experimentally accessible quantity in $\vec{m}$ is $m_z^{D^0}$, hence, we are interested in its behaviour in time under different saturation conditions determined by Ω and $\Delta\nu$. Even though numerical simulations of evolution under Equation (6.16) are straight forward to carry out for any set of parameters, it is somewhat more illuminating to derive a few results analytically.

To approximate the steady state solution $\vec{m}(t \to \infty)$ we first notice that

$$\frac{\Upsilon}{\kappa_{D^0X\to D^+}} \begin{pmatrix} 0 \\ 0 \\ \frac{\kappa_{D^0\to D^0X}}{\kappa_{D^+\to D^0}} \\ 0 \\ 0 \\ 0 \\ \frac{\kappa_{D^0\to D^0X}}{\kappa_{D^0X\to D^+}} \\ 0 \\ 0 \\ 0 \\ 1 \\ 1/\mu_z^{D^0} \end{pmatrix} = \frac{\kappa_{D^0\to D^0X}}{\kappa_{D^0X\to D^+}} \begin{pmatrix} 0 \\ -\frac{\Omega}{\kappa_{D^+\to D^0}} \\ 0 \\ 0 \\ 0 \\ -\frac{\Omega}{\kappa_{D^0X\to D^+}} \\ 0 \\ 0 \\ 0 \\ 0 \\ o \\ 0 \end{pmatrix}, \tag{6.23}$$

where $o = \Omega^2 / \left[\Omega^2 + \zeta^2 + (\kappa_{D^+ \to D^0})^2\right]$, $\zeta = \delta - \Delta\nu$,

$$\mu_z^{D^0} = \left[1 + \frac{\Omega^2 \; T_1^{D^0} \kappa_{D^0 \to D^0 X}}{\Omega^2 + \zeta^2 + (\kappa_{D^+ \to D^0})^2}\right]^{-1} \tag{6.24}$$

and Υ is defined by Equation (6.16). It can be seen from Equation (6.23) that, if $\kappa_{D^0 X \to D^+} \geq \kappa_{D^+ \to D^0}$, $\kappa_{D^0 X \to D^+} \geq \Omega$, $\kappa_{D^+ \to D^0} \geq \Omega$ and $\kappa_{D^0 X \to D^+} \gg \kappa_{D^0 \to D^0 X}$, the vector on the left hand side of the equation is an approximate fixed point of Equation (6.16). For the relatively low Rabi strengths (≤ 4 kHz) used in this work we expect all such approximations to hold. Given that $m_I^{D^0}(t) = 1$, for all $t \geq 0$, we see that the steady state magnetization under Equation (6.16) for the stated parameter range is $m_z^{D^0}(t \to \infty) \approx \mu_z^{D^0}$. Furthermore, inspecting Equation (6.16) for $m_z^{D^0}(0) = 0$ under the same approximations reveals that $\partial/\partial t \; m_z^{D^0}(t)|_{t=0} = 1/T_1^{D^0}$, hence, the time dependent behaviour of $m_z^{D^0}$ for $m_z^{D^0}(0) = 0$ is very close to an exponential rise

$$m_z^{D^0}(t) \approx \mu_z^{D^0} \left[1 - \exp\left(-\frac{1}{\mu_z^{D^0}} \frac{t}{T_1}\right)\right]. \tag{6.25}$$

The experimentally controllable parameters in Equation (6.24) are the Rabi strength (Ω) of the RF saturation drive and its detuning ($\Delta\nu$) from the bare phosphorus Larmor frequency $\omega_{^{31}\mathrm{P}}$. It is apparent that $\mu_z^{D^0}$ is an inverted Lorentzian in ζ for constant Ω. Importantly, the initial growth of $m_z^{D^0}$ is always independent of the saturation drive as $\frac{d}{dt} m_z^{D^0}(t)|_{t=0} = 1/T_1^{D^0}$. This feature appears because in our model the polarization growth happens for neutral donor defects independently of the free exciton capture events. Conversely, it will be shown in the next section that when the polarization transfer to the nucleus happens during the free exciton capture, which is always followed by ionization and hence saturation, the initial slopes of $m_z^{D^0}(t)$ for different $\mu_z^{D^0}$ would vary.

Owing to the design of our experiment, we have appreciable variation of light intensity within our sample on length scales that exceed the mean free path of free electrons and excitons in silicon at cryogenic temperatures [25, 16]. Therefore, we have to account for local variations of $\kappa_{D^0 \to D^0 X}$ and $\kappa_{D^+ \to D^0}$ within our sample. Because we attribute the DNP to a phononic process, which would be non-local in nature [45], contrary to the photoionaization and photoneutralization processes, we take $T_1^{D^0}$ to be independent of local light intensity. Introducing a distribution for the light intensity $\eta(I)$, normalized such that $\int dI \; \eta(I) = 1$, we deduce an expression for overall nuclear magnetization for

neutral donor defects starting from a uniform $M_z^{D^0}(0)$ across the sample:

$$M_z^{D^0}(t) = \int dI\ \eta(I)\ \mu_z^{D^0}(I) \left[1 - \exp\left(-\frac{1}{\mu_z^{D^0}(I)} \frac{t}{T_1^{D^0}}\right)\right] + \int dI\ \eta(I)\ M_z^{D^0}(0) \exp\left(-\frac{1}{\mu_z^{D^0}(I)} \frac{t}{T_1^{D^0}}\right), \tag{6.26}$$

where

$$\mu_z^{D^0}(I) = \left[1 + \frac{\Omega^2\ T_1^{D^0}\ \kappa_{D^0 \to D^0 X}(I)}{\Omega^2 + \zeta^2 + \left[\kappa_{D^+ \to D^0}(I)\right]^2}\right]^{-1}, \tag{6.27}$$

$\kappa_{D^0 \to D^0 X}(I)$ being the local light intensity dependent free exciton capture rate and $\kappa_{D^+ \to D^0}(I)$ being the local light intensity dependent photoneutralization rate. All of our experiments were performed over time scales for which the free electron and free exciton densities in the sample have reached their equilibrium values, hence, we take $\kappa_{D^+ \to D^0}(I) = \alpha_{D^+ \to D^0} \sqrt{I}$ and $\kappa_{D^0 \to D^0 X} = \alpha_{D^0 \to D^0 X} I$, with $\alpha_{D^+ \to D^0}$ and $\alpha_{D^0 \to D^0 X}$ being proportionality constants. Such proportionality relations are derived under the same assumptions as in [118] – free electrons and holes are generated at the same rate, the predominant recombination channel for free electrons and holes is the formation of free excitons, free excitons recombine independently of one another and the number of bound excitons is proportional to the number of free excitons in a particular volume of the crystal.

6.3 Free Exciton Capture Dependent Hyperpolarization Model

In this section, we construct a model that describes the generation of nuclear hyperpolarization through free exciton capture events. We do this using the tools of open quantum systems theory [120, 10]; our reason for building the model is to enable us to contrast the phononic and the exciton capture induced nuclear hyperpolarization mechanisms in the light of our experimental data and to quantitatively investigate the correlation between the photoionization events and the polarization transfer to the ^{31}P nucleus. We keep the model as general as possible and only make a handful of assumptions that are listed below.

- Free exciton capture is a ^{31}P spin state dependent Poisson process – the capture events happen independently from each other and are on average separated by some fixed time τ_n, which represents the effective lifetime of the neutral donor state averaged over the two nuclear spin states.

- Polarization transfer occurs instantaneously and only during free exciton capture events, meaning that there is some ^{31}P spin state dependent probability for the nuclear spin to flip when an exciton is captured. Assuming this also implies that there are no other competing DNP mechanisms and that the thermal relaxation of ^{31}P is negligible over the duration of the experiment. The > 5 h long longitudinal spin relaxation time measured in our previous experiments [38] completely justifies the latter assumption.

- Every free exciton capture is followed by Auger ionization, and every D^0 ionization event is due to free exciton capture.

- Photoneutralization is a nuclear spin state independent Poisson process, such that the average time a donor spends in the ionized state is some fixed τ_i. Furthermore, ionization events resulting from free exciton capture are relatively infrequent as compared to the photoneutralization events, i.e., $\tau_i \ll \tau_n$.

- The Rabi strength (Ω) of the saturation drive is low enough, and τ_i is long enough, that the polarization loss during bound exciton lifetime is negligible as compared to the saturation effects during the ionized state.

Even though the assumptions above are idealized, they can be taken to be approximately true if the polarization transfer onto ^{31}P is indeed resulting from the free exciton capture events, and the number of donor ionization events other than Auger ionization does not substantially outnumber the Auger events.

Given our assumptions, we can say that the short time evolution of the nuclear spin density matrix ρ over a period Δt, $\tau_i \ll \Delta t \ll \tau_n$, happens under an effective map $\Phi(\Delta t)$, such that

$$\rho(t+\Delta t) = \Phi(\Delta t)[\rho(t)] = \left(1 - \frac{\Delta t}{\tau_n}\right)\rho(t) + \frac{\Delta t}{\tau_n}\Phi_{\mathrm{BE}}[\rho(t)], \tag{6.28}$$

where Φ_{BE} is a map acting on ρ upon the capture of a free exciton. Φ_{BE} encapsulates the effects of both the polarization transfer and the saturation in the ionized state. It can be seen that $\Phi(\Delta t)$ inherits all of its properties from Φ_{BE}. Specifically, they share the same steady-state density matrix ρ_{ss}, which satisfies $\Phi(\Delta t)(\rho_{\mathrm{ss}}) = \Phi_{\mathrm{BE}}(\rho_{\mathrm{ss}}) = \rho_{\mathrm{ss}}$, and the difference quotient $[\rho(t+\Delta t) - \rho(t)]/\Delta t = [\Phi_{\mathrm{BE}}[\rho(t)] - \rho(t)]/\tau_n$ is only a function of Φ_{BE} and τ_n.

We proceed by constructing the map Φ_{BE} from its constituents, starting with the effect of the saturation drive in the ionized state that we model as a pure polarization damping

process Φ_{sat}, with no coherences being left after the application of Φ_{sat}. Assuming that photoneutralization is a Poisson process with a characteristic time constant τ_i, the amount of polarization destroyed by saturation drive with amplitude Ω applied off resonance by ζ is quantified by

$$\int_0^\infty dt\ \frac{e^{-t/\tau_i}}{\tau_i} e^{-i\left(\Omega\frac{\sigma_x}{2}+\zeta\frac{\sigma_z}{2}\right)t}\sigma_z e^{i\left(\Omega\frac{\sigma_x}{2}+\zeta\frac{\sigma_z}{2}\right)t} = \left(1-p_{\text{sat}}\right)\sigma_z - \frac{p_{\text{sat}}}{\tau_i\ \Omega}\sigma_y + \frac{p_{\text{sat}}\zeta}{\Omega}\sigma_x, \tag{6.29}$$

where $p_{\text{sat}} = \tau_i^2\Omega^2/\left[1+\tau_i^2\left(\zeta^2+\Omega^2\right)\right]$. Accordingly, we can write Φ_{sat} for the nuclear spin density matrix ρ as

$$\begin{aligned}\Phi_{\text{sat}} = &\left(1-\frac{p_{\text{sat}}}{2}\right)|\uparrow\rangle\langle\uparrow|\rho|\uparrow\rangle\langle\uparrow| + \frac{p_{\text{sat}}}{2}|\downarrow\rangle\langle\uparrow|\rho|\uparrow\rangle\langle\downarrow| + \\ &\left(1-\frac{p_{\text{sat}}}{2}\right)|\downarrow\rangle\langle\downarrow|\rho|\downarrow\rangle\langle\downarrow| + \frac{p_{\text{sat}}}{2}|\uparrow\rangle\langle\downarrow|\rho|\downarrow\rangle\langle\uparrow|.\end{aligned} \tag{6.30}$$

Next, we say that over some short time Δt there are two possibly distinct nuclear spin state dependent probabilities – $p_{\text{BE}\uparrow}$ and $p_{\text{BE}\downarrow}$ – to catch a free exciton for states $|\uparrow\rangle$ and $|\downarrow\rangle$, respectively. Only if a free exciton is captured is the ^{31}P spin flipped with either of the nuclear state dependent probabilities $p_{\text{F}\uparrow}$ or $p_{\text{F}\downarrow}$, followed by polarization loss due to Φ_{sat}. Hence, we write Φ_{BE} in Equation (6.28) as

$$\begin{aligned}\Phi_{\text{BE}}[\rho] = &p_{\text{BE}\uparrow}\Phi_{\text{sat}}\left[p_{\text{F}\uparrow}|\downarrow\rangle\langle\uparrow|\rho|\uparrow\rangle\langle\downarrow| + \left(1-p_{\text{F}\uparrow}\right)|\uparrow\rangle\langle\uparrow|\rho|\uparrow\rangle\langle\uparrow|\right] + \\ &p_{\text{BE}\downarrow}\Phi_{\text{sat}}\left[p_{\text{F}\downarrow}|\uparrow\rangle\langle\downarrow|\rho|\downarrow\rangle\langle\uparrow| + \left(1-p_{\text{F}\downarrow}\right)|\downarrow\rangle\langle\downarrow|\rho|\downarrow\rangle\langle\downarrow|\right] + \\ &\left(1-p_{\text{BE}\uparrow}\right)|\uparrow\rangle\langle\uparrow|\rho|\uparrow\rangle\langle\uparrow| + \left(1-p_{\text{BE}\downarrow}\right)|\downarrow\rangle\langle\downarrow|\rho|\downarrow\rangle\langle\downarrow|.\end{aligned} \tag{6.31}$$

Substituting the definition of Φ_{sat} into the above, and evaluating $\Phi\left(\Delta t\right)[\rho]$ gives

$$\begin{aligned}\Phi\left(\Delta t\right)[\rho] = &|\uparrow\rangle\langle\uparrow|\rho|\uparrow\rangle\langle\uparrow| + |\downarrow\rangle\langle\downarrow|\rho|\downarrow\rangle\langle\downarrow| - \\ &\frac{\Delta t\ \chi_{\text{BE}\uparrow}}{2}\left[2p_{\text{F}\uparrow}\left(1-p_{\text{sat}}\right)+p_{\text{sat}}\right]|\uparrow\rangle\langle\uparrow|\rho|\uparrow\rangle\langle\uparrow| - \\ &\frac{\Delta t\ \chi_{\text{BE}\downarrow}}{2}\left[2p_{\text{F}\downarrow}\left(1-p_{\text{sat}}\right)+p_{\text{sat}}\right]|\downarrow\rangle\langle\downarrow|\rho|\downarrow\rangle\langle\downarrow| + \\ &\frac{\Delta t\ \chi_{\text{BE}\uparrow}}{2}\left[2p_{\text{F}\uparrow}\left(1-p_{\text{sat}}\right)+p_{\text{sat}}\right]|\downarrow\rangle\langle\uparrow|\rho|\uparrow\rangle\langle\downarrow| + \\ &\frac{\Delta t\ \chi_{\text{BE}\downarrow}}{2}\left[2p_{\text{F}\downarrow}\left(1-p_{\text{sat}}\right)+p_{\text{sat}}\right]|\uparrow\rangle\langle\downarrow|\rho|\downarrow\rangle\langle\uparrow|,\end{aligned} \tag{6.32}$$

where $\chi_{\text{BE}\uparrow} = p_{\text{BE}\uparrow}/\tau_n$ and $\chi_{\text{BE}\downarrow} = p_{\text{BE}\downarrow}/\tau_n$ are spin state dependent rates for exciton capture. It is rather easy to check, given $0 \leq \chi_{\text{BE}\uparrow}$, $0 \leq \chi_{\text{BE}\downarrow}$, $0 \leq p_{\text{F}\uparrow} \leq 1$, $0 \leq p_{\text{F}\downarrow} \leq 1$

and $\Delta t \leq \tau_n$, that Φ is indeed a completely positive trace preserving linear map [120] for all values of ζ, Ω and $\tau_i \ll \tau_n$. The steady state density matrix $\rho_{\mathrm{ss}}(\Omega, \zeta)$ of Φ is given by

$$\rho_{\mathrm{ss}}(\Omega,\zeta) = \frac{\mathbb{1}}{2} + \frac{2\left(\chi_{\mathrm{BE}\downarrow}p_{\mathrm{F}\downarrow} - \chi_{\mathrm{BE}\uparrow}p_{\mathrm{F}\uparrow}\right)\left(1+\tau_i^2\zeta^2\right) + \left(\chi_{\mathrm{BE}\downarrow} - \chi_{\mathrm{BE}\uparrow}\right)\tau_i^2\Omega^2}{2\left(\chi_{\mathrm{BE}\downarrow}p_{\mathrm{F}\downarrow} + \chi_{\mathrm{BE}\uparrow}p_{\mathrm{F}\uparrow}\right)\left(1+\tau_i^2\zeta^2\right) + \left(\chi_{\mathrm{BE}\downarrow} + \chi_{\mathrm{BE}\uparrow}\right)\tau_i^2\Omega^2}\frac{\sigma_z}{2}, \tag{6.33}$$

whereas the initial rate of polarization build up evaluates to

$$\begin{aligned}\frac{1}{\Delta t}\left(\mathrm{Tr}\left[\sigma_z\rho(\Delta t)\right] - \mathrm{Tr}\left[\sigma_z\rho(0)\right]\right) &= \frac{1}{\Delta t}\mathrm{Tr}\left[\sigma_z\Phi(\Delta t)\left[\mathbb{1}/2\right]\right] \\ &= \frac{1}{2}\frac{2\left(\chi_{\mathrm{BE}\downarrow}p_{\mathrm{F}\downarrow} - \chi_{\mathrm{BE}\uparrow}p_{\mathrm{F}\uparrow}\right)\left(1+\tau_i^2\zeta^2\right) + \left(\chi_{\mathrm{BE}\downarrow} - \chi_{\mathrm{BE}\uparrow}\right)\tau_i^2\Omega^2}{1+\tau_i^2\left(\zeta^2+\Omega^2\right)},\end{aligned} \tag{6.34}$$

for $\rho(0) = \mathbb{1}/2$. It is clear that, in the absence of very particular model conditions, the initial Ω and ζ dependent slopes are always different.

Combining the above, we see that the build up of magnetization starting from $m_z(0)$ under Ω strong Rabi drive applied ζ off resonance has the following behaviour in time:

$$\begin{aligned}m_z(t) \approx \mu_z\left[1 - \exp\left(-\frac{t}{2}\frac{2\left(\chi_{\mathrm{BE}\downarrow}p_{\mathrm{F}\downarrow} + \chi_{\mathrm{BE}\uparrow}p_{\mathrm{F}\uparrow}\right)\left(1+\tau_i^2\zeta^2\right) + \left(\chi_{\mathrm{BE}\downarrow} + \chi_{\mathrm{BE}\uparrow}\right)\tau_i^2\Omega^2}{1+\tau_i^2\left(\zeta^2+\Omega^2\right)}\right)\right] \\ + m_z(0)\exp\left(-\frac{t}{2}\frac{2\left(\chi_{\mathrm{BE}\downarrow}p_{\mathrm{F}\downarrow} + \chi_{\mathrm{BE}\uparrow}p_{\mathrm{F}\uparrow}\right)\left(1+\tau_i^2\zeta^2\right) + \left(\chi_{\mathrm{BE}\downarrow} + \chi_{\mathrm{BE}\uparrow}\right)\tau_i^2\Omega^2}{1+\tau_i^2\left(\zeta^2+\Omega^2\right)}\right),\end{aligned} \tag{6.35}$$

where

$$\mu_z = \frac{2\left(\chi_{\mathrm{BE}\downarrow}p_{\mathrm{F}\downarrow} - \chi_{\mathrm{BE}\uparrow}p_{\mathrm{F}\uparrow}\right)\left(1+\tau_i^2\zeta^2\right) + \left(\chi_{\mathrm{BE}\downarrow} - \chi_{\mathrm{BE}\uparrow}\right)\tau_i^2\Omega^2}{2\left(\chi_{\mathrm{BE}\downarrow}p_{\mathrm{F}\downarrow} + \chi_{\mathrm{BE}\uparrow}p_{\mathrm{F}\uparrow}\right)\left(1+\tau_i^2\zeta^2\right) + \left(\chi_{\mathrm{BE}\downarrow} + \chi_{\mathrm{BE}\uparrow}\right)\tau_i^2\Omega^2}. \tag{6.36}$$

It can be seen that the general form of Equation (6.36) and Equation (6.35) is quite different from the steady state magnetization, and build up behaviour, derived for the phononic model in Equations (6.24) and (6.25), respectively. Again, in the absence of particular values for $\chi_{\mathrm{BE}\downarrow}$, $\chi_{\mathrm{BE}\uparrow}$, $p_{\mathrm{F}\downarrow}$, $p_{\mathrm{F}\uparrow}$ and τ_i, the model predictions would differ. Specifically, for terminal polarization equalling one in the absence of saturation drive $\mu_z(\Omega = 0) = 1$ the equations are in agreement if and only if $\chi_{\mathrm{BE}\downarrow} = \kappa_{D^0\to D^0X} + 2/T_1^{D^0}$, $\chi_{\mathrm{BE}\uparrow} = \kappa_{D^0\to D^0X}$, $p_{\mathrm{F}\downarrow} = 1/\left(2 + \kappa_{D^0\to D^0X}T_1^{D^0}\right)$, $p_{\mathrm{F}\uparrow} = 0$ and $\tau_i = 1/\kappa_{D^+\to D^0}$. These particular parameters can be interpreted as describing a combined effect of spin state independent ionization events that also result no spin flips at a rate of $\kappa_{D^0\to D^0X}$ and exciton capture events at the rate of $2/T_1^{D^0}$ that only happen for nuclear spin state $|\downarrow\rangle$ with probability one.

Again, as in the previous section, we have to allow for variations in $\chi_{\mathrm{BE}\downarrow}$, $\chi_{\mathrm{BE}\downarrow}$ and τ_i due to local light intensity variations within our sample. Using the same arguments as

before we say that $\chi_{\mathrm{BE}\downarrow}(I) = \alpha_{\mathrm{BE}\downarrow} I$, $\chi_{\mathrm{BE}\uparrow}(I) = \alpha_{\mathrm{BE}\uparrow} I$ and $\tau_i(I) = 1/(\alpha_i \sqrt{I})$, where $\alpha_{\mathrm{BE}\downarrow}$, $\alpha_{\mathrm{BE}\uparrow}$ and α_i are constants, and the overall nuclear magnetization of the sample is given as

$$M_z(t) = \int dI\ \eta(I)\ \mu_z^{D^0}(I) \left[1 - \exp\left(-\frac{t}{T_{\mathrm{BE}}(\Omega, \zeta, I)}\right)\right] + \int dI\ \eta(I)\ M_z(0, I) \exp\left(-\frac{t}{T_{\mathrm{BE}}(\Omega, \zeta, I)}\right), \tag{6.37}$$

where

$$\mu_z(I) = \frac{2\left[\chi_{\mathrm{BE}\downarrow}(I) p_{\mathrm{F}\downarrow} - \chi_{\mathrm{BE}\uparrow}(I) p_{\mathrm{F}\uparrow}\right]\left[1 + \tau_i^2(I)\zeta^2\right] + \left[\chi_{\mathrm{BE}\downarrow}(I) - \chi_{\mathrm{BE}\uparrow}(I)\right]\tau_i^2(I)\Omega^2}{2\left[\chi_{\mathrm{BE}\downarrow}(I) p_{\mathrm{F}\downarrow} + \chi_{\mathrm{BE}\uparrow}(I) p_{\mathrm{F}\uparrow}\right]\left[1 + \tau_i^2(I)\zeta^2\right] + \left[\chi_{\mathrm{BE}\downarrow}(I) + \chi_{\mathrm{BE}\uparrow}(I)\right]\tau_i^2(I)\Omega^2} \tag{6.38}$$

and

$$T_{\mathrm{BE}}(\Omega, \zeta, I) = \frac{2\left[1 + \tau_i^2(I)\left(\zeta^2 + \Omega^2\right)\right]}{2\left[\chi_{\mathrm{BE}\downarrow}(I) p_{\mathrm{F}\downarrow} + \chi_{\mathrm{BE}\uparrow}(I) p_{\mathrm{F}\uparrow}\right]\left[1 + \tau_i^2(I)\zeta^2\right] + \left[\chi_{\mathrm{BE}\downarrow}(I) + \chi_{\mathrm{BE}\uparrow}(I)\right]\tau_i^2(I)\Omega^2} \tag{6.39}$$

and the initial polarization $M_z(0, I)$ is dependent on the local light intensity I. The key difference between Equation (6.37) and the overall magnetization evolution equations derived for the phononic model in Equation (6.26) is that for the phonic model the constant $T_1^{D^0}$ does not depend on the local light intensity. Hence, for any non-uniform light intensity distribution the two models cannot perfectly agree as the parameter matching shown above can only happen for the average local light intensity value $\langle I \rangle$.

6.4 Experimental Setup

In this section, we will introduce the low temperature NMR experimental setup that was used for validating the phononic hyperpolarization mechanism introduced in Section 6.1. The setup is illustrated by Figure 6.2, and can be divided into six main parts: a 6.7 T superconducting magnet, an NMR console, a liquid helium (LHe) dewar along with the pumping lines and pumps, cryogenic NMR probe insert, a room temperature optical setup and the single crystal silicon sample. The first three parts of the setup are commercial units and will be described only briefly, whereas the last three parts are not commercially available, so we will dedicate the next three subsections on discussing those. In the last subsection, we describe particular experimental configuration for all experiments to be introduced in this chapter along with our ray tracing simulations that account for the light intensity simulations inside the silicon crystal.

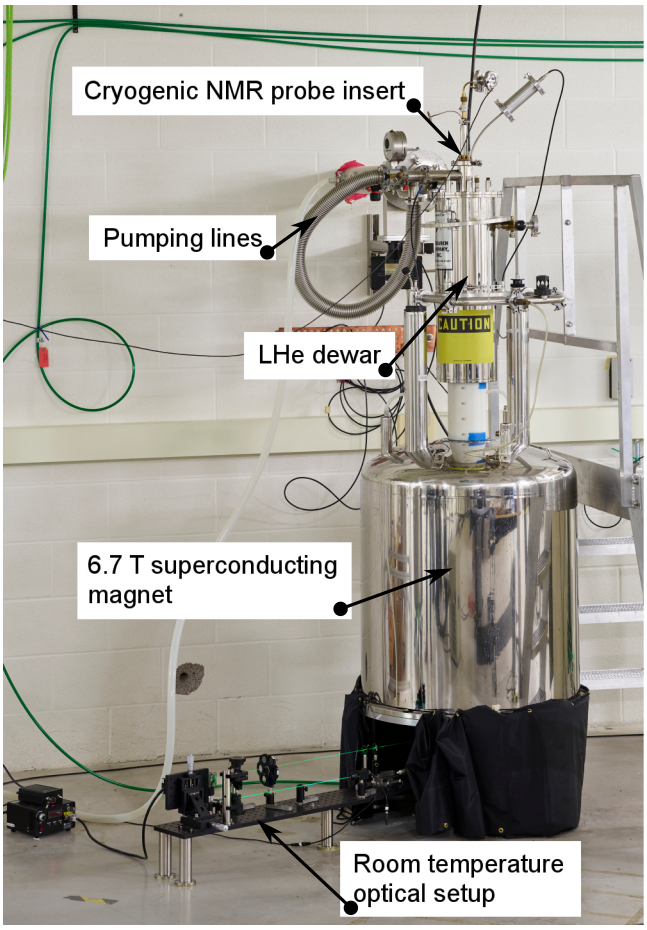

Figure 6.2: Experimental setup including the room temperature optical breadboard, Bruker superconducting magnet, Janis liquid helium dewar along with the pumping lines and the low temperature probe insert

The superconducting magnet is a 89 mm bore Bruker magnet, designed for NMR applications that has been energized to $B_0 = 6.7$ T. The NMR console is a Bruker Avance 300 spectrometer which combines an RF pulse generation unit and a 300 W Bruker BLAX300RS high power amplifier on the transmission side and a broadband preamplifier and a spectrometer unit on the receiver side. The cryostat is a wet Janis 8CNDT system with an 8 l bath space for LHe and an exterior liquid nitrogen shield. The cryostat is made mostly made of stainless steel and has an exterior tail of 3.375" diameter which is inserted into the superconducting magnet bore, and extends past the sweet spot of the magnet. The inner tail of the cryostat has a 1.88" diameter and fits the low temperature NMR probe insert, the inner tail is custom manufactured out of aluminium in order to reduce the presence of magnetic materials in the near vicinity of the sample. Both of the tails, as well as the middle radiation shield of the dewar, have sapphire windows mounted at the bottom to enable optical access to the sample from below. The bath space of the cryostat is built to withstand a vacuum of 0.1 Torr, which provided us with an ability to access temperature ranges between 1.3 K to 4.2 K by pumping on the helium space. The cold time of the cryostat was ∼28 h at atmospheric pressure, or 4.2 K, and ∼10 h at 1 Torr or 1.3 K. The pumping system was a RUVAC WS1001 roots pump backed by a TRIVAC B D 65 B rotary vane vacuum pump. The inside pressure of the dewar was monitored with a Granville-Phillips 275 Mini-Convectron pressure sensor, while the bath space temperature was monitored with a LakeShore 370 AC resistance bridge connected to a calibrated Cernox 1050-CU-1.4L thermometer designed for high magnetic field cryogenic applications.

6.4.1 Cryogenic NMR Probe Insert[1]

The cryogenic NMR probe insert is designed to fit inside the 1.88" inner diameter tail of the dewar. It consists of three main sections: the room temperature top of the probe insert that can be seen in Figure 6.2, mid-section of the insert, and the NMR probe head at the bottom. The entire probe insert is built around a 1/2" thin walled stainless steel tube, the inside of which is used for guiding the liquid helium fill line when cooling down the experiment. Stainless steel is used for construction due to its low heat conduction properties. At the top of the tube is a vacuum feed trough for the helium fill line which can be sealed with a brass plug after the fill is completed. The bottom 40 cm of the central tube is made from a brass tube of matching dimensions that is silver brazed to the stainless section, this is necessary because the weak magnetism of stainless steel could otherwise broaden the NMR spectra.

[1]The author designed and built the cryogenic probe insert with Rahul Deshpande.

The top of the probe insert accommodates: (i) four custom made vacuum feed-throughs for tuning rods of the probe head capacitors, (ii) two vacuum feed-throughs for 50 Ω impedance transmission lines for the probe head circuit and for a pick-up coil, (iii) a four-pin vacuum feed-trough for a thermometer. The top of the probe insert is clamped onto the cryostat neck with an o-ring vacuum seal between the two. The mid-section of the probe insert guides the two coaxial semi-rigid transmission lines along with the four tuning rods to the probe head through a serious of equally placed, ∼25 cm apart, copper baffles. The baffles are concentric with the central stainless steel tube, and are silver brazed onto it, the baffles are necessary to reduce the room temperature radiative heat load onto liquid helium and to break the convection path of helium vapour when the dewar is kept at atmospheric pressure. The walls of the central tube are perforated over the mid-section of the probe insert at equal lengths of ∼15 cm to prevent the formation of Taconis oscillations during or after liquid helium fills. The coaxial cable for the pick-up coil has a stainless steel outer conductor and a beryllium copper inner conductor to reduce the heat load on the liquid helium. The coaxial cable for the probe head has copper inner and outer conductors for all but the last ∼35 cm-long section closest to the probe top, that is done to minimize the signal attenuation from the probe head to the spectrometer pre-amplifier, whereas the last ∼35 cm section is again made from stainless steel/beryllium copper semi-rigid coaxial cable. The mid-section of the probe insert also houses a Cernox 1050-CU-1.4L thermometer that is mounted just above the probe head circuit to ensure accurate temperature readings throughout the whole cool-down period.

At the bottom of the probe insert lies the resonant NMR probe head circuit, which is used for delivering high power RF pulses to the ^{3}1P nuclear spins and also for the inductive signal detection of the same spins. This circuit incorporates an inductive coil that houses the sample and is adjusted to be located at the sweet spot of the superconducting magnet. In order to study the photoionization and photoneutralization processes happening under above gap light, we needed an NMR probe head that enabled us to drive, and possibly detect, both of the resonance frequencies $\omega_{^{31}\mathrm{P}}$ and ω_{D^0}. Given that we were dealing with the same nuclei, we did not, however, need an ability to address both of the resonances at the same time. This simplified the circuit design, and we opted for a single channel split resonance probe rather than a dual channel NMR probe. The lumped element circuit for the probe head is given in [40], and is illustrated by Figure 6.3. The only drawback of such a circuit is that the presence of two inductive coils doubles the effective mode volume of the resonator and therefore reduces the sensitivity of the probe if only one of the coils is used for spin detection, as it is in our case.

The circuit consist of two LC oscillators formed by C_{t1}, L_1 and C_{t2}, L_2, that have nearly equal resonance frequencies and are coupled through a coupling capacitor C_c. Such

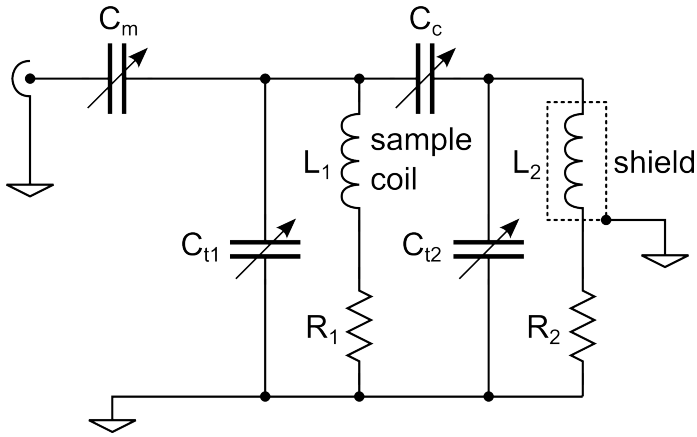

Figure 6.3: Circuit for the NMR probehead proposed in [40], and utilized for measurements here. The circuit consists of four variable capacitors C_{t1}, C_{t2}, C_m and C_c, two inductor coils L_1 and L_2 with respective resistances of R_1 and R_2. L_1 is dedicated as the sample or detection coil while L_2 is housed in a grounded copper shield in order to minimize mutual inductance between the two inductors.

a circuit has two nearby resonance frequencies, which can be made to coincide with $\omega_{^{31}P}$ and ω_{D^0} with the right choice for C_{t1}, L_1, C_{t2}, L_2 and C_c. For maximum signal transfer, the whole circuit is impedance matched to $Z_0 = 50\ \Omega$ with a matching capacitor C_m.

Both of the inductors are wound from rhodium flashed silver plated 0.7 mm gauge copper wire manufactured for NMR purposes. The coils have 7 turns and an identical inner diameter of 5.2 mm and pitch of 1.2 mm, which resulted $L_1 = L_2 = \sim$125 nH and $R_1 = R_2 = 0.1\ \Omega$. As it can be seen in Figure 6.3, only one of the coils is designated for housing the sample during experiments, that detection coil is held in place with a coil holder machined from PTFE. Mutual inductance between the two coils alters the desired circuit behaviour and has to be mitigated, hence, the coils are physically separated as far as possible within the probe head volume, and their axis are aligned to be perpendicular to each other. Nevertheless, this does not completely remove their mutual inductance, and therefore, a grounded copper shielding is added around L_2, as is indicated in Figure 6.3.

All four capacitors C_{t1}, C_{t2}, C_m, C_c are picked to be Voltronics NMCB10-5CKE trimmer capacitors with a $1 - 10$ pF capacitance range that are designed specifically for cryogenic NMR applications. The capacitors are manufactured out of strictly non-magnetic metals, which includes all platings, and made with low temperature compatible dielectric sapphire. All capacitors are connected axially to 1/8" G10 rods which are fed through vacuum fittings at the top of the probe insert. This allows for tuning and matching the probe circuit at

cryogenic temperatures, the upper and lower resonances of the circuit have tuning ranges of 165-190 MHz and 110-122 MHz, respectively, while maintaining good matching – $< 1\%$ reflections in power – to the external electronics.

Spin detection for this work was carried out only on the hyperfine shifted ^{31}P spins at the $\omega_{D^0}/(2\pi) \approx 174.1$ MHz frequency. Upon transmitting an RF pulse resonant with ω_{D^0} to the probe head circuit, the nuclear spins are nutated to the xy plane that is perpendicular to the direction of the external field B_0. The xy plane component of the spin magnetization in the sample will then precess around the z-axis at the Larmor frequency $\omega_{D^0}/(2\pi)$. The oscillating magnetic moment along the detection coil axis induces an AC electromotive force (emf) in the coil [51]. This emf resonantly drives the LC oscillator that is tuned to match the Larmor frequency, thereby generating a current that is guided to the Bruker Avance 300 preamplifier through the same coaxial transmission line that is used for delivering the RF excitation pulse. The preamplifier output is transmitted to the phase sensitive detection unit of the Bruker Avance 300 spectrometer which splits the signal from the probe and mixes the split signal down with reference and quadrature reference signals, respectively. The reference signal frequency corresponds to that of the RF excitation pulses, delivered by the spectrometer. Finally, the two mixed down signals are digitalized into two arrays of time dependent data. For all measurements in this work the digitalization time step was 1 μs and a total of 32,768 points for both the in-phase and quadrature components were recorded in each experiment.

The spin signal that is detected immediately after applying the RF excitation pulse is called the free induction decay (FID), since it always displays a decaying character with some characteristic time constant T_2^*. Frequently, this signal decay happens in an exponential or close to exponential fashion. It should be noted here that the spin signal detection always lags the RF excitation pulse end by an interval called the dead time, that happens because the resonant LC circuit dissipates the excitation energy over some finite time – immediately after the RF excitation one would only detect such probe ring-down rather than the spin signal. For the data presented in this chapter we used a dead time of 40 μs. The discrete Fourier transform of the FID is often referred to as the NMR spectrum, a representative NMR spectrum from the ^{31}P nuclei in the ^{28}Si sample is given in Figure 6.4. The 930 Hz linewidth seen in the figure is due inhomogeneous line broadening from silicon magnetic susceptibility and B_0 inhomogeneities, the ^{31}P nuclear spin $T_2^{D^0}$ in the same material has been measured to be 0.421 s at 1.7 K [38] with a Hahn echo experiment. All data points in the coming sections of this chapter correspond to various integrated amplitudes of the NMR spectra collected. The mean magnetization, or polarization, of nuclear spins is proportional to such integrals provided that the RF excitation pulses used are identical. Nevertheless, our inductive spin signal detection on its own does not enable

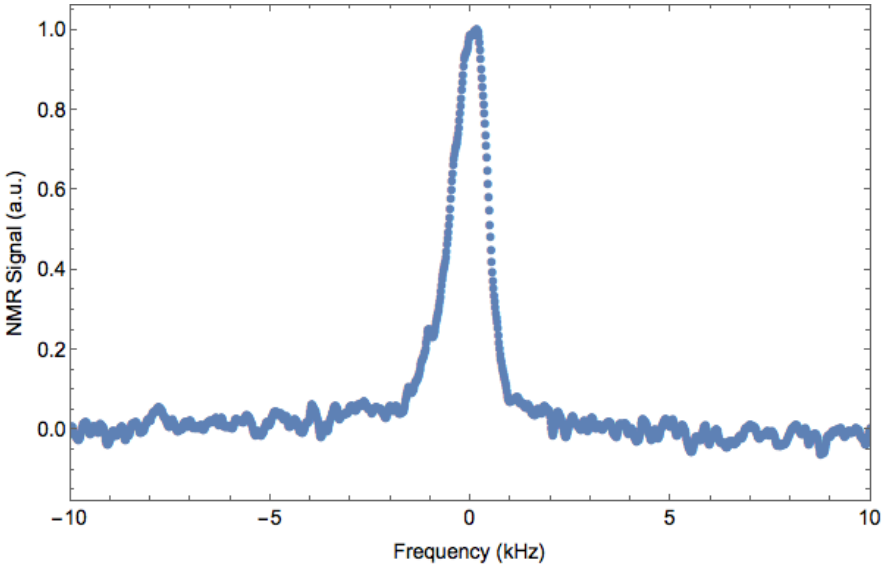

Figure 6.4: A sample NMR spectrum for the ^{31}P nuclei in the isotopically enriched silicon-28 crystal. Zero frequency in the figure corresponds to 174.10462 MHz, the FWHM of the spectrum is 930 Hz, the data was collected over 16 scans by applying a 2.5 μs excitation pulse after 400 s long illumination with above gap light.

quantifying the absolute magnitude nor the direction of the ^{31}P spin polarization.

In order to study the photoionization and neutralization processes, we needed to calibrate the Rabi field strengths at both ω_{D^0} and $\omega_{^{31}\mathrm{P}}$. At ω_{D^0} frequency, we needed the calibration to maximize the FID signal amplitude and thereby the signal to noise ratio (SNR) of our detection scheme. At $\omega_{^{31}\mathrm{P}}$ frequency, the calibration was necessary to test the model introduced in Section 6.3. For our probe circuit in Figure 6.3, both of the Rabi strengths depend a lot on the particular tuning configuration, and had to be measured before each experiment. ω_{D^0} can be calibrated simply by carrying out a Rabi experiment. For that, RF excitation pulses of different duration are applied at some certain fixed amplitude, and the resulting integrated signal amplitude is recoded. Plotting out the integrals as a function of pulse durations will reveal a decaying sinusoidal shape the inverse period of which corresponds to the Rabi frequency at ω_{D^0}. In practice, we usually searched for the first zero crossing of the Rabi curve, which happens at half the Rabi period.

Direct inductive detection of ^{31}P nuclei at $\omega_{^{31}\mathrm{P}}$ frequency was impossible since the number of ionized defect sites under illumination is very low, and the duration of the ionized state is around a few microseconds. Therefore, the Rabi field strength at that

frequency could not be calibrated like the one at ω_{D^0}. In order to calibrate the Rabi field strength at $\omega_{^{31}\mathrm{P}}$ frequency we used a pick up coil device that was weakly coupled to the NMR detection coil. This device is a single 3 mm loop of 1 mm gauge copper wire that is connected in series with a 50 Ω thin film resistor. The relatively low inductance of the wire loop ensures a relatively similar frequency response at both of these frequencies, as the complex impedance of the device is close to $Z_0 = 50\ \Omega$ real impedance at both frequencies. The thin film resistor, on the other hand, ensures that the resistance of the device, and hence its RF properties, are not much different between room temperature and cryogenic temperatures.

The pick up coil device enabled us to measure the relative voltage amplitudes generated in the detection coil at ω_{D^0} and $\omega_{^{31}\mathrm{P}}$ with a Tektronix DPO 4104 digital oscilloscope for any given RF excitation pulse amplitude and frequency. Thereby, we could convert the Rabi field strength at ω_{D^0} to that at $\omega_{^{31}\mathrm{P}}$. When doing so, we accounted for two factors that alter the conversion from plainly equating the pulse amplitudes for yielding equal Rabi strengths. First, the transmission properties from the pick up coil device to the digital oscilloscope were not exactly the same at both of these frequencies, this was corrected for by measuring the S_{11} transmission properties of the device with a vector network analyser. Secondly, and much more importantly, when converting the Rabi strength from ω_{D^0} to $\omega_{^{31}\mathrm{P}}$, we had to account for the difference in the effective gyromagnetic ratios for the ionized phosphorus nucleus and the hyperfine coupled ^{31}P. The ionized ^{31}P nuclei have a gyromagnetic ratio that is very close to that of the bare nuclei, i.e., $\gamma_{^{31}\mathrm{P}}/(2\pi) = 17.2515$ MHz/T [108]. Nevertheless, the hyperfine coupled ^{31}P nuclei experience an effective increased gyromagnetic ratio, which appears because the time-dependent magnetic field generated by the RF excitation pulse also adiabatically alters the spin state of the defect electron. This electronic spin state evolution translates to an extra effective Rabi field for the nuclear spin through an effective time-dependent modulation of the hyperfine coupling by the electron spin. Such enhanced gyromagnetic ratio is well known and exploited in electron nuclear double resonance experiments [93], the effective gyromagnetic ratio for ^{31}P at ω_{D^0} is then given by

$$\gamma_{D^0} = \left(1 + \frac{A_e(0)}{2\omega_{^{31}\mathrm{P}}}\right)\gamma_{^{31}\mathrm{P}} = 1.510\gamma_{^{31}\mathrm{P}} \tag{6.40}$$

for our values of $A_e(0) = 117.52$ MHz and $\omega_{^{31}\mathrm{P}} = 115.3$ MHz. This factor of 1.510 was accounted for when we calibrated the Rabi field strengths at $\omega_{^{31}\mathrm{P}}$ through the ones at ω_{D^0}.

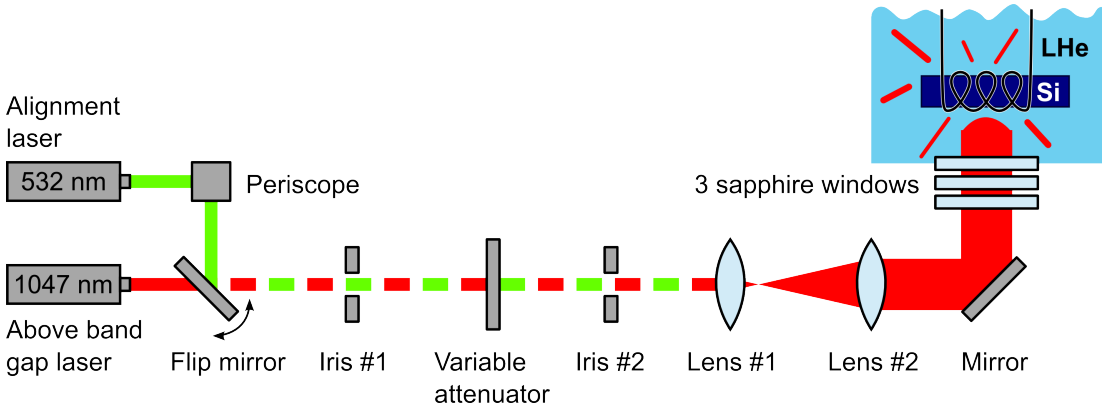

Figure 6.5: Room temperature optical assembly configuration used for this work comprising two lasers and a flip mirror to switch between them, a variable optical attenuator, two irises, two lenses and a mirror to guide the laser beams onto the silicon sample inside the helium bath space of the cryostat.

6.4.2 Optical Setup[2]

The room temperature optical setup can be seen at the bottom of the photograph in Figure 6.2, and the schematic configuration of it is shown in Figure 6.5. Since most of the parts for the optical setup have been purchased from Thorlabs, we specify their part numbers as (TL: #), where # stands for Thorlabs part number. The entire room temperature optical setup is mounted onto an aluminium optical breadboard (TL: MB648) that measures 6"×48"×1/2" and stands on five 5"-tall stainless steel posts (TL: P5) that have been epoxied onto the floor underneath the magnet, to ensure that the breadboard does not shift with respect to the magnet bore during, or in between, the experiments.

Here, we list all of the optical components shown in Figure 6.5, and describe their respective function.

- The 1047 nm above band gap laser is a LRS-1047 DPSS Laserglow laser with output power of 429.6 mW in a 2 mm ($1/e^2$ width) laser beam [59]. This laser is used for providing the optical excitation of the silicon sample, and is mounted onto a 4"×6"×1/2" aluminium breadboard (TL: MB4) in order to provide the laser with a necessary heat sink. The breadboard is, in turn, mounted onto a two-axis translation

[2]The author designed the optical setup described in this subsection. The assembly was done by the author and Natsumi Komatsu.

stage (TL: PT102 & PT1) which enables aligning the laser beam path precisely through the irises.

- The 532 nm green alignment laser is a LBS-532 Laserglow laser with output power of 5 mW in a 4 mm ($1/e^2$ width) laser beam. The beam width of the alignment laser can be changed with iris #1 in Figure 6.5 during the alignment procedure.

- The two irises (TL: ID36) are used for aligning the optical paths of the alignment laser and the 1047 nm laser. Once the optical paths of the two lasers are aligned, we can use the visible alignment laser for adjusting the laser spot position after the NMR probe head and the sample have been cooled down.

- The periscope (TL: RS99) that includes two silver mirrors (TL: PF10-03-P01) is used for setting the alignment laser beam height to be exactly the same as the 1047 nm beam height. One of the periscope mirrors can be positioned with three-axis adjustment screws.

- The flip mirror (TL: FM90, POLARIS-K1 & PF10-03-P01) is used for switching between the alignment laser and the 1047 nm laser, and can also be positioned with three-axis adjustment screws. The combination of the adjustable periscope mirror and the flip mirror makes guiding the alignment laser through the irises a rather simple task.

- The variable laser beam attenuator that is aligned perpendicular to the laser beam and comprises four neutral density filters – (TL: NE03B-B, NE06B-B, NE10B-B & NE13B-B) with respective optical densities of 0.3, 0.6, 1.0 and 1.3 – that are mounted within a filter wheel station (TL: FW1A).

- The two convex lenses (TL: LA1708-B & LA1131-B) with focal lengths of 20 cm and 5 cm are mounted perpendicular to the laser beam path, such that their focal points coincide. The lenses are used in the telescope configuration for magnifying the laser beam width four times. Lens #1 is mounted onto a linear translation stage (TL: PT1) which greatly simplifies the task of setting up the telescope. The collimation of the laser beams is carefully checked four meters away from the lenses because this is critical for our light distribution modelling that will be described in one of the following sections.

- The mirror (TL: PF10-03-P01) and its adjustable three-axis mount (TL: POLARIS-K1) are used to direct the horizontal laser beam through the three sapphire windows

at the bottom of the cryostat onto the sample. The mirror lies below the superconducting magnet on the axis of its bore. The three-axis adjustment screws are used for precise positioning of the laser spot before every experiment. It should be noted that the alignment laser spot on the sample is easily visible through the sapphire windows, both when the setup is at room temperature or when at cryogenic temperatures.

The alignment laser is not crucial, however, it is incredibly useful for two reasons. First, aligning the laser beam with the optical axes of the lenses is extremely tedious with an infrared laser, the spot of which is not visible. Secondly, since the interior tails of the liquid helium cryostat and the probe insert tend to move during the cool down process, sometimes as much as a few mm, it is extremely useful to be able to align the laser beam once the setup has reached its base temperature. Without a visible alignment laser this would not be possible. Before each experiment, we first check that both the 532 nm and the 1047 nm laser are perfectly aligned through the centres of the two irises. We then use the green laser to align the laser spot exactly on the sample and finally switch to the 1047 nm laser to perform the experiments.

6.4.3 Silicon-28 Sample[3]

The material is a float-zone grown, extremely pure, dislocation free, isotopically enriched ^{28}Si single crystal produced as part of the Avogadro project [3]. The ^{28}Si concentration in the material is 99.9954%, the residual 46 ppm is expected to be mostly ^{29}Si [4]. The crystal is doped with phosphorus at $N_{^{31}\mathrm{P}} = 1.5 \times 10^{15}$ cm^{-3}, which was introduced as phosphine gas during the final growth run. Such doping method should ensure that the phosphorus donor defects are distributed uniformly inside the material. The material also contains a residual boron acceptor concentration of 10^{14} cm^{-3} making it an n-type semiconductor. Assuming uniform distribution of defects, and ignoring the silicon lattice effects due to their much shorter length, the nearest neighbour pair distribution is given by

$$w(r) = 4\pi N_{^{31}\mathrm{P}} r^2 \ \exp\left(-\frac{4}{3}\pi r^3 N_{^{31}\mathrm{P}}\right), \tag{6.41}$$

corresponding to the most likely nearest neighbour distance of $(2\pi N_{^{31}\mathrm{P}})^{-1/3} = 47$ nm. Our sample was cut to its rectangular $2.12 \times 2.22 \times 8$ mm^3 size from the material roughly along the three crystallographic axis with a diamond saw. The cutting procedure creates surface damage, mechanical strain and electronic surface defects [24] – all of which would affect

[3]The sample was provided through a collaboration by Professor Micheal L. W. Thewalt.

our studies – therefore, the sample was etched in a 1 : 10 solution of HF/HNO_3 to remove such surface effects.

6.4.4 Experimental Configuration and Light Distribution[4]

All experiments were performed at 1.3 ± 0.05 K in a superfluid ^{4}He bath, this ensured that the sample remained well thermalized throughout the experiments, as well as the absence of bubbles that the laser beam would otherwise scatter off of. The latter is especially important to be able to accurately model the light deposition in the crystal. We measured the bare phosphorus Zeeman splitting to be $\omega_{^{31}\mathrm{P}}/(2\pi) = 115.31163$ MHz at room temperature with a liquid state trisodium phosphate water solution sample. This yields an $\omega_{D^0}/(2\pi) = (\omega_{^{31}\mathrm{P}} + A_e/2)/(2\pi) = 174.10462$ MHz resonance frequency for the hyperfine coupled phosphorus nuclear spin in the ground state manifold of the electron spin [107] as is seen in Figure 6.4. The resulting electron Larmor frequency of $\omega_e/(2\pi) = 188$ GHz implies that at 1.3 K the electron occupies its spin ground state with 99.9% probability. The same liquid state sample was also used for adjusting the cryostat and the dewar position inside the magnet bore, to precisely locate the sweet spot of the magnet by minimizing the spectral linewidth from that sample. Once the configuration yielding the narrowest and the most symmetric Lorentzian line shape was located, the dewar was fixed onto the magnet with double sided tape to prevent it from moving during and between the experiments.

The saturation drive at frequencies near $\omega_{^{31}\mathrm{P}}$ at high Rabi strengths $\Omega/(2\pi) = 4$ kHz was delivered directly by the Bruker Avance 300 spectrometer, and amplified by its 300 W amplifier. Conversely, low Rabi strength $\Omega/(2\pi) \leq 1$ kHz saturation drive was delivered by amplifying a clean RF tone generated by the spectrometer signal generation unit (SGU) with a Mini Circuits ZHL-32A low noise amplifier, the output of which was filtered by a K&L tunable bandpass filter and fed to the probe circuitry through the reverse port J4 of a Werlatone C5964 30dB dual directional coupler. The latter configuration provides a much less noisy saturation drive, which is necessary for experiments that have a longer duration over which the broadband noise outputted by the power amplifier had a measurable saturation effect on its own.

The sample was mounted inside a 4 mm ID polished sapphire tube in a strain free manner, this provided electrical isolation from the NMR coil wound tightly around the sapphire tube. An illustration of the sample mounting configuration, along with the laser beam can be seen in Figure 6.6(a). The laser beam was passed through the optical access

[4]The COMSOL light deposition simulations described in this subsection were carried out by Thomas Alexander, the data analysis was conducted by the author.

windows of the cryostat and hit the sample tube perpendicular to its axis. We measured the total laser power entering the ^{4}He bath space to be 218 mW. To adjust the laser power we used either absorptive neutral density filters or the power variation feature of the laser; in all cases, we characterized the output power and its drift with a power meter.

The reflective, silver plated, copper coil created a periodic grating of light intensity across the length of the sample. We simulated this three-dimensional pattern with COMSOL by tracing 200,000 rays in a 8 mm ($1/e^2$ width) Gaussian beam profile through the coil-sapphire tube-sample assembly and recording the laser power deposited in each roughly 100 μm $\times$ 100 μm $\times$ 100 μm-sized volume element of the Si crystal. The simulation tracked each ray until it excited the coil-tube-sample region of space. The mesh dimensions for the ray tracing simulations were picked to be four times greater than our upper bound for the diffusion length of free excitons and at least an order of magnitude shorter than the mean free path of 100 GHz phonons. Chen and Lyon [16] have measured the free exciton diffusion length for 5×10^{-15} cm^{-3} phosphorus doped silicon at 12 K to be 24 μm. We expect the free electron diffusion length to be shorter than that of the free excitons'. Electron mobility measurements at 1.25 K for 7×10^{-15} cm^{-3} estimate these lengths to be on the order of 10 μm [25]. On the other hand, Hao and Maris [45] quoted 100 GHz phonons to have a mean free path of 3.3 mm at 30 K in natural abundance silicon.

We show an image of the resulting surface light intensity pattern superimposed on the sample in Figure 6.6(a). This tracking included all reflections from any surface of the coil and any refraction due to the sapphire tube and the sample. For the ray tracing simulations we assumed the coil surface to be perfectly reflective; for silicon and sapphire we used their respective refractive indices $n_{\mathrm{Si}} = 3.41$ [27] and $n_{\mathrm{Sapphire}} = 1.75$ [21] for 1047 nm wavelength at cryogenic temperatures. For silicon we used a penetration depth $\lambda = 26.9$ cm quoted by Macfarlane *et al* [69] for 1047 nm light in intrinsic silicon at 4.2 K, which was used to calculate the laser power absorbed in each mesh element. We estimated the aggregate laser power deposited in the sample at full laser power to be 0.232 mW, which would correspond to uniform illumination with light intensity at 166 mW cm^{-2}. This would be a reasonable estimate even if the penetration depth is somewhat incorrect.

6.5 Data, Modelling, Results[5]

In this section, we first introduce the experiments performed along with the data collected, followed by the analysis and precise descriptions of our modelling and fits. All experiments

[5]The author collected the data presented in this section with Rahul Deshpande and Thomas Alexander. The modelling and analysis was conducted by the author.

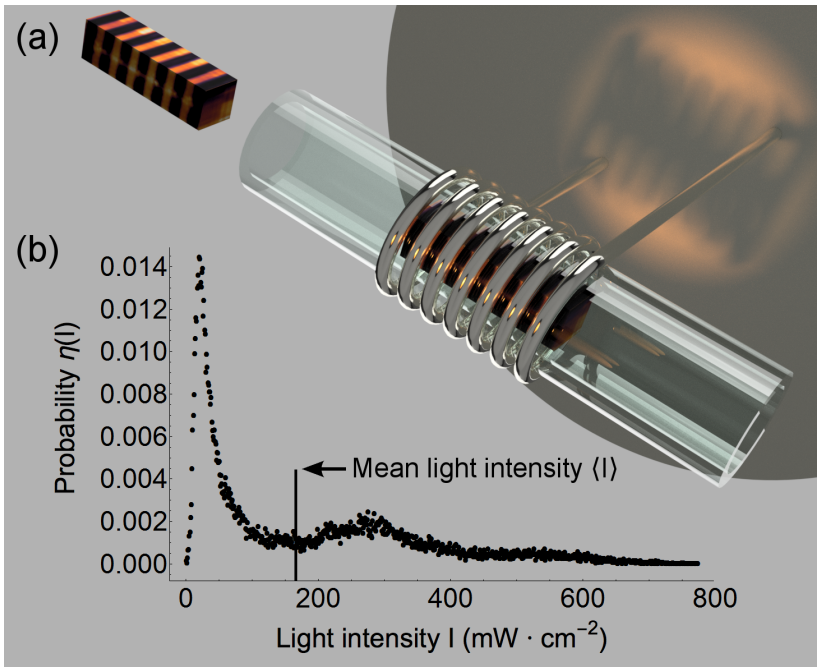

Figure 6.6: (a) The RF-coil, sapphire sample tube, silicon sample and laser beam configuration. The sample dimensions are $2 \times 2 \times 8$ mm^3. The coil blocks and reflects part of the light and the sample tube acts as a lens, this generates a highly non-uniform distribution of light across the sample. We have superimposed surface light intensity patterns from COMSOL ray tracing simulations on the surface of the sample, and we also include an identical copy of the sample outside of the tube to better illustrate the non-uniform nature of the light intensity. (b) The probability η of finding a $100 \times 100 \times 100$ μm^3 volume element inside the sample with a given light intensity I extracted from the ray tracing simulations illustrated above.

that we undertook followed the sequences given either in Figure 6.8(a) or Figure 6.9(a). In both cases, the experiment starts with a train of $\pi/2$ pulses resonant with D^0 nuclei to zero the ^{31}P spin polarization while the laser is off. This is followed by a period of laser illumination and simultaneous RF saturation drive near the resonance of the D^+ nuclei. Finally, the laser is switched off and a FID signal is collected after applying a readout $\pi/2$-pulse on the D^0 nuclei. All data points presented correspond to integrated FID spectra, the number of averages for each measurement was chosen to yield roughly the same SNR for all data points. In each case, we turned the laser off 500 ms before collecting the FID to ensure that transient effects of the photoexcitation, such as shortening of $T_2^{D^0}$, would not affect the measurement. Because thermal relaxation time scales of ^{31}P spins under no laser illumination exceed 5 hours [38], the signal decay over the delay period is negligible. Our first experiment followed the sequence in Figure 6.8(a). We measured the ^{31}P spin polarization build up in time at four different laser powers P_{laser} while applying no saturation pulse, i.e., $\Omega = 0$. The data collected is displayed in Figure 6.7 and exhibits exponential behaviour with the time constant $T_1^{D^0}(P_{\text{laser}})$ inversely proportional to P_{laser}, whereas the final polarization is independent of P_{laser}.

To minimize the experiment time, we performed the rest of the experiments at full laser power. Our second experiment again followed the sequence in Figure 6.8(a). This time we varied $\Delta\nu$ of the saturation drive at two distinct drive strengths $\Omega/(2\pi) = \{120\text{ Hz}, 240\text{ Hz}\}$, and measured the resulting spin polarization at the end of a 160 s long lasing period. The resulting data is given in Figure 6.8(b). Importantly, we observed a paramagnetic shift in the D^+ phosphorus resonance frequency, measured from $\omega_{^{31}\text{P}}$, which can be seen as a horizontal offset from the origin, denoted by δ, for the dips in Figure 6.8(b). We used this value for δ in both of the subsequent experiments.

We also measured build up data at three different saturation drive strengths $\Omega/(2\pi) = \{0\text{ Hz}, 120\text{ Hz}, 240\text{ Hz}\}$ applied on resonance $\Delta\nu = \delta$ with the ionized nuclei. This experiment employed the protocol in Figure 6.8(a) by varying the lasing duration, the resulting data is shown in Figure 6.8(c).

Finally, we designed an experiment, illustrated by Figure 6.9(a), to test our model at saturation drive strengths an order of magnitude stronger than the ones used for the experiments in Figure 6.8. For such high saturation drives, the terminal nuclear polarizations are very low and experiments of Figure 6.8(a) kind become impractical due to their low SNR. Consequently, we performed a measurement that involved first bringing the nuclei close to their terminal spin polarization by lasing the sample for 160 s, thereafter, we turned on a strong $\Omega/(2\pi) = 4$ kHz resonant $\Delta\nu = \delta$ saturation drive, which for half of the runs was preceded by a π pulse at the D^0 resonance to flip the spin orientation. The spin flip was performed in order to investigate whether the polarization decay starting from ^{31}P spin

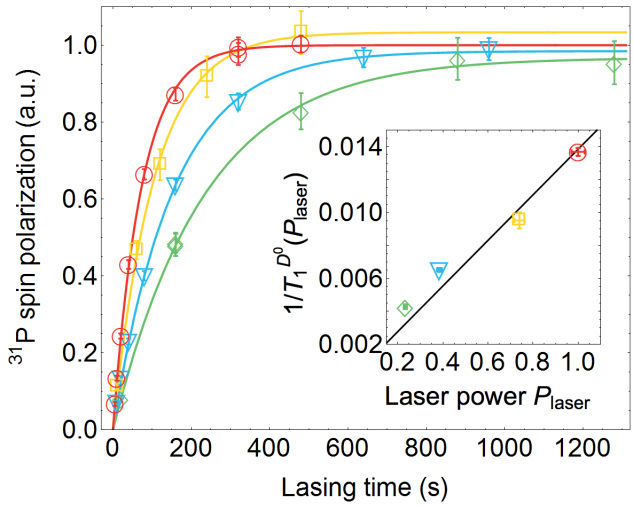

Figure 6.7: ^{31}P polarization build up data measured at four laser output powers as a function of lasing time. The relative laser power corresponding to a particular data set, denoted by a unique marker and color, is indicated on the horizontal axis of the inset. All build up curves are well described by exponentials that have been fitted to the data, and reach the same plateau value. We have extracted the build up time $T_1^{D^0}(P_{\text{laser}})$ for each curve, and show that the inverse of $T_1^{D^0}$ is linearly proportional to the laser power in the inset.

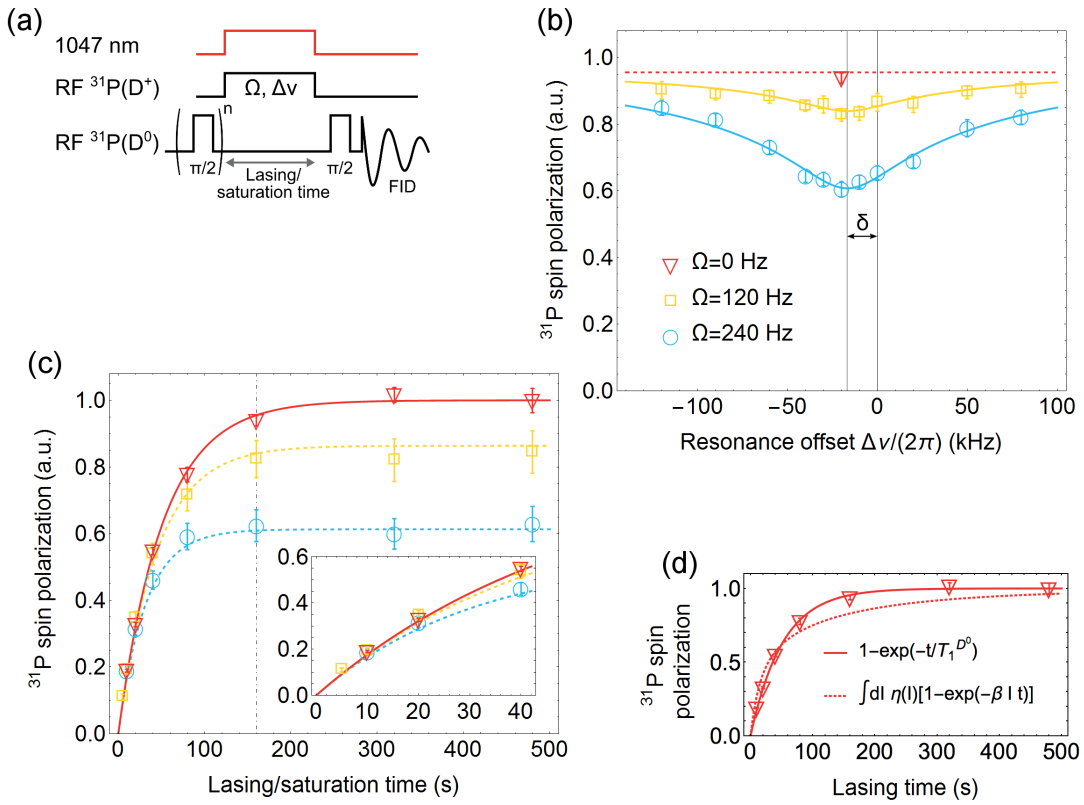

Figure 6.8: (a) Experimental sequence used for collecting the data in Figures 6.7, 6.8(b) and 6.8(c). (b) ^{31}P spin polarization at the end of a 160 s lasing/saturation period for $\Omega/(2\pi) = \{0 \text{ Hz}, 120 \text{ Hz}, 240 \text{ Hz}\}$ for various $\Delta\nu$ values. The horizontal axis origin is at the ^{31}P spin resonance $\omega_{^{31}\text{P}}/(2\pi) = 115.3116$ MHz, while the gray vertical line denotes our best estimate for the paramagnetic shift δ. The $\Omega/(2\pi) = \{120 \text{ Hz}, 240 \text{ Hz}\}$ data have been fitted to Equation 6.26, and the fits are presented in the figure. (c) ^{31}P spin polarization at the end of various lasing/saturation periods for $\Delta\nu = \delta$ and $\Omega/(2\pi) = \{0 \text{ Hz}, 120 \text{ Hz}, 240 \text{ Hz}\}$. The vertical dash dotted line marks the 160 s lasing/saturation period used for collecting the data in (b). The $\Omega/(2\pi) = 0$ Hz data has been fitted with an exponential, the time constant of which is $T_1^{D^0}$, while the dashed lines laid over the $\Omega/(2\pi) = \{120 \text{ Hz}, 240 \text{ Hz}\}$ data sets are our model predictions for the signal. The inset displays the ^{31}P spin polarization measurements at short lasing periods. (d) The same data as the $\Omega/(2\pi) = 0$ Hz measurements in (c) fitted with a single exponential (solid line) and local light intensity dependent integrated signal $\int dI\ \eta(I)\left(1 - e^{-\beta I t}\right)$ (dotted line), where β has been adjusted to yield the best least squares fit to the data.

Table 6.1: Parameter estimates extracted from the data presented in Figures 6.8(b) and 6.8(c)

Parameter	Estimate	Parameter	Mean value[1]
$T_1^{D^0}\langle I\rangle$	8540 ± 250 s $\frac{\text{mW}}{\text{cm}^2}$	$T_1^{D^0}$	51.5 ± 1.5 s
$\alpha_{D^0\to D^0X}$	7.2 ± 0.7 Hz $\frac{\text{cm}^2}{\text{mW}}$	$\kappa_{D^0\to D^0X}$	1.2 ± 0.1 kHz
$\alpha_{D^+\to D^0}$	36 ± 2 kHz $\frac{\text{cm}}{\sqrt{\text{mW}}}$	$\kappa_{D^+\to D^0}$	470 ± 30 kHz
$\delta/(2\pi)$	-17 ± 2 kHz		

[1]Calculated for mean light intensity of $\langle I\rangle = 166$ mW cm^{-2}.

state $|\uparrow_n\rangle$ is the same as it is for the spin state $|\downarrow_n\rangle$. We present the data collected in Figure 6.9(b).

For the analysis, we converted the laser power deposition data from COMSOL simulations into local light intensity I by assuming that the power deposited in each mesh element with area A and thickness Δz could be approximated by $IA\left(1-e^{-\Delta z/\lambda}\right) \approx IA\Delta z/\lambda$, where λ is the penetration depth. As we argued earlier, given that λ is two orders of magnitude longer than the sample thickness, our values for I should be good estimates even if the value we used for λ is not exact. We binned the light intensities for all mesh elements yielding the light intensity distribution $\eta(I)$ displayed in Figure 6.6(b).

To extract $\alpha_{D^0\to D^0X}$ and $\alpha_{D^+\to D^0}$ that determine the free exciton capture rate $\kappa_{D^0\to D^0X}(I)$ and the photoneutralization rate $\kappa_{D^+\to D^0}(I)$, we first found $T_1^{D^0}$ by fitting an exponential to the $\Omega = 0$ Hz dataset in Figure 6.8(c). Given $T_1^{D^0} = 51.5$ s and the simulated $\eta(I)$, we searched for δ, $\alpha_{D^0\to D^0X}$ and $\alpha_{D^+\to D^0}$ that would best fit Equation 6.26 to the data in Figure 6.8(b) as a function $\Delta\nu$, simultaneously for $\Omega/(2\pi) = 120$ Hz and $\Omega/(2\pi) = 240$ Hz, while having set $t = 160$ s and $M_z^{D^0} = 0$. We note that our ^{31}P FID spectra always had linewidths < 1 kHz, hence, we justifiably ignored any B_0 inhomogeneities in our model. The resulting values for δ, $\alpha_{D^0\to D^0X}$ and $\alpha_{D^+\to D^0}$ are quoted in Table 6.1, along with the light intensity dependent value for $T_1^{D^0}$, the latter is given as a function of average light intensity $\langle I\rangle = \int dI\ \eta(I)I$. In the same table, we also quote the rates $\kappa_{D^0\to D^0X}$ and $\kappa_{D^+\to D^0}$ corresponding to the average light intensity within our sample. It was stated in [101] that a free exciton capture should occur over a time period between 10 μs to 100 μs at their laser intensity of 400 mW and beam diameter of 3 mm. For such beam intensity our value for $\alpha_{D^0\to D^0X}$ predicts an average capture time of 25 μs, which lies perfectly within that range.

We used the parameters given in Table 6.1 along with $\eta(I)$ to simulate the magnetiza-

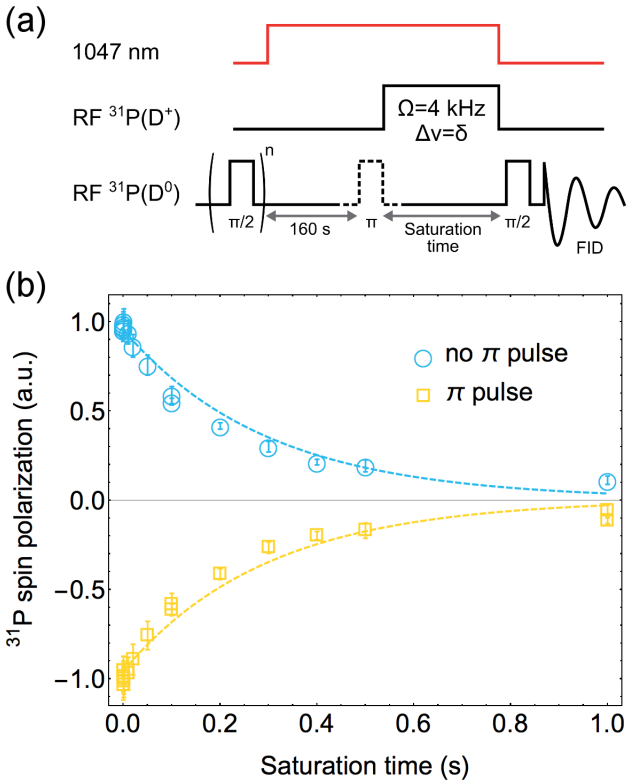

Figure 6.9: (a) Experimental sequence used for collecting the data in (b). The 160 s lasing period brought the spin polarization to 96% of its terminal value. For half of the measurements a π-pulse resonant with D^0 nuclear spins was included at the end of the lasing period to flip the spin orientation. The π pulse length was $\sim$13 μs, hence, its inclusion did not affect the effective experiment time. (b) ^{31}P spin polarization as a function of saturation time, the round/blue (square/yellow) markers correspond to measurements without (with) the application of the π pulse. The dashed lines correspond to our model prediction for the signal decay.

tion dynamics, as predicted by Equation (6.26), for build up curves that used $M_z^{D^0} = 0$, $\Delta\nu = \delta$ and $\Omega/(2\pi) = \{120 \text{ Hz}, 240 \text{ Hz}\}$. We present the model predictions with experimental data in Figure 6.8(c), both of the simulations follow the data very accurately providing a self-consistency check for our model. We also simulated the spin polarization decays in Figure 6.9(b) as predicted by Equation (6.26) for $\Omega/(2\pi) = 4$ kHz, $\Delta\nu = \delta$ and $M_z^{D^0} = \pm\left[1 - \exp\left(-160 \text{ s}/T_1^{D^0}\right)\right]$. Again, our predictions closely follow the data, the slight discrepancy present could result from finite $\kappa_{D^0X\rightarrow D^+}$ effects that may become important at higher saturation drives or small deviations between the simulated and experimental light intensity distributions within the sample.

In deriving Equation (6.26), we used an approximation $\kappa_{D^0X\rightarrow D^+} \gg \kappa_{D^0\rightarrow D^0X}$ and $\kappa_{D^0X\rightarrow D^+} \gg \kappa_{D^+\rightarrow D^0}$. From physical considerations we expect $\kappa_{D^+\rightarrow D^0} \gg \kappa_{D^0\rightarrow D^0X}$ since the former rate arises due to electrostatic attraction between an ionized nucleus and an electron, whereas the latter rate involves an effective attraction between two neutral objects – a neutral donor defect and a free exciton. This assumption was also reinforced by the absence of measurable signal difference between detecting an FID with or without the laser on. Furthermore, it was reasonable to assume $\kappa_{D^0X\rightarrow D^+} \gg \kappa_{D^+\rightarrow D^0}$ since the data presented in Figure 6.8(b) did not show any qualitative evidence of hole hyperfine couplings. The parameter estimates in Table 6.1 are self-consistent with the aforementioned approximations. We also verified that full simulations with the finite $\kappa_{D^0X\rightarrow D^+} = 3.7$ MHz [100], and a range of values for A_h and $T_2^{D^0X}$, yielded almost the same simulated curves as seen in Figures 6.8(b), 6.8(c) and 6.9(b). Consequently, we cannot rule out A_h values even within the frequency range displayed in Figure 6.8(b), that would require an experiment employing higher Rabi strengths or light intensities.

So far, in this section, we have presented our data and verified its consistency with our model. We will now highlight its features which support the argument for a phonon mediated hyperpolarization mechanism acting in our experiments. Our data shows two separate pieces of evidence for such a mechanism: the exponential nature of the build up curves in Figure 6.7 and the short time behaviour of the build up curves shown in the inset of Figure 6.8(c).

Our simulations of $\eta(I)$ in Figure 6.6(b) indicate that the local light intensity within our sample was highly inhomogeneous. Because the mean free path of free excitons in our crystal is likely of order 10 μm [16], we would expect a DNP process involving the formation of bound excitons to yield very different local polarization build up rates inside the sample. In such a case, the signal build up in Figure 6.7 should follow some non-trivial rather than the simple exponential trend observed. That is because nuclear spin diffusion cannot be assumed to average out the spatial variations in spin polarization over our experimental

time scales, given that the most likely nearest neighbour ^{31}P $\leftrightarrow^{31}$ P dipolar coupling at our donor concentration is $\sim$0.1 mHz. As discussed in Section 6.2, the number of bound excitons present in some portion of the sample is linearly proportional to the local light intensity I. Consequently, we expect the local light intensity dependent polarization build up to follow a trend $\int dI\ \eta(I)\left(1-e^{-\beta I t}\right)$, where β is some proportionality constant. We searched for the β that yielded the best least squares fit to the $\Omega/(2\pi) = 0$ Hz data in Figure 6.8(c), we display that fit along with the data in Figure 6.8(d). It can be seen that such integrated signal deviates considerably from the data. Meanwhile, sub-terahertz phonon mean free paths in silicon at 30 K have been measured to be several mm [45], almost equalling the dimensions of our sample. Phononic process would therefore be insensitive to the local light intensity within our sample, and result in a simple exponential build up curve as observed.

Furthermore, In Section 6.2 we showed that, if the polarization transfer happens for neutral donor defects D^0, the Rabi drive Ω dependent build up curves should all have equal Ω independent initial slopes, the experimental data shown in the inset of Figure 6.8(c) is consistent with that. Such behaviour is expected to appear whenever the DNP mechanism is largely uncorrelated with the exciton capture events since it implies that there exists a finite time, for which the number of defect sites that have experienced a capture is low enough, such that the polarization growth is fully determined by $T_1^{D^0}$.

In Section 6.3, we constructed a free exciton capture dependent hyperpolarization model under very general assumptions. It is clear from Equation (6.37) that $M_z(t)$ – the overall nuclear polarization of the sample – is fully specified by 5 independent parameters $\alpha_{\mathrm{BE}\downarrow}$, $\alpha_{\mathrm{BE}\uparrow}$, $p_{\mathrm{F}\downarrow}$, $p_{\mathrm{F}\uparrow}$, and α_i. We pointed out that for uniform light intensity within the sample the there exist a particular choice for these parameters which yields the same model predictions as the phonic model. For that parameter choice the process could be interpreted as a combined effect of spin state independent ionization events that result no spin flips at a rate of $\kappa_{D^0\rightarrow D^0X}$, and exciton capture events at the rate of $2/T_1^{D^0}$, that only happen for nuclear spin state $|\downarrow\rangle$ with probability one. Under that interpretation the quantity $2/(T_1^{D^0}\kappa_{D^0\rightarrow D^0X})$ gives us a ratio of spin flipping ionization events versus ionization events that do not alter the nuclear spin state. For the $T_1^{D^0}$, and the mean $\kappa_{D^0\rightarrow D^0X}$ values, given in Table 6.1 that ratio is 3.2×10^{-5}, which implies that the correlation between these events can only be very weak.

Nevertheless, for non-uniform light intensity distribution the two model predictions given by Equation (6.26) and Equation (6.37) cannot be made to agree due to the local light intensity independent polarization mechanism acting in the case of the phononic model. We searched for $\alpha_{\mathrm{BE}\downarrow}$, $\alpha_{\mathrm{BE}\uparrow}$, $p_{\mathrm{F}\downarrow}$, $p_{\mathrm{F}\uparrow}$, and α_i that would best fit the data presented in Figure 6.8(b) and 6.8(c) to Equation (6.37), as a function of t, Ω and ζ for $M_z^{D^0}(0) = 0$ and

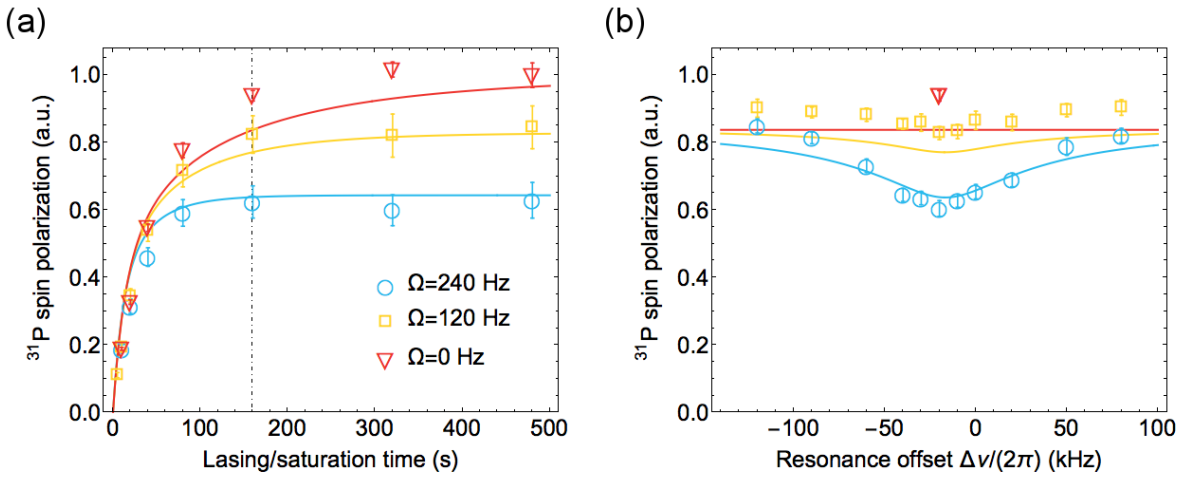

Figure 6.10: (a) Build up data from Figure 6.8(c) and our best fits to Equation (6.37) for $\Omega/(2\pi) = \{0 \text{ Hz}, 120 \text{ Hz}, 240 \text{ Hz}\}$, $\Delta\nu = \delta$ and $M_z^{D^0}(0) = 0$. (b) Data from Figure 6.8(b) and our best fits to Equation (6.37) for $t = 160$ s, $\Omega/(2\pi) = \{0 \text{ Hz}, 120 \text{ Hz}, 240 \text{ Hz}\}$ and $M_z^{D^0}(0) = 0$.

the δ value given in Table 6.1. We used this particular data set for our fits because the two experiments involved were guaranteed to start from uniform initial nuclear magnetization equalling zero; the parameters yielding the best fit to the data were $\alpha_{\text{BE}\downarrow} = 9 \pm 8$ Hz $\frac{\text{cm}^2}{\text{mW}}$, $\alpha_{\text{BE}\uparrow} = 6 \pm 5$ Hz $\frac{\text{cm}^2}{\text{mW}}$, $p_{\text{F}\downarrow} = 0.000027 \pm 0.000025$, $p_{\text{F}\uparrow} = 0 + 10^{-5}$, and $\alpha_i = 33 \pm 16$ kHz $\frac{\text{cm}}{\sqrt{\text{mW}}}$, our best fits along with the data are given in Figure 6.10(a) and 6.10(b). It is eminent that the fits to our data are rather poor, which is also reflected in the huge error bars for the parameters.

We also simulated the predicted magnetization dynamics for the experimental data presented in Figure 6.9(b), given the parameter values for $\alpha_{\text{BE}\downarrow}$, $\alpha_{\text{BE}\uparrow}$, $p_{\text{F}\downarrow}$, $p_{\text{F}\uparrow}$, and α_i quoted above. For this simulation the initial light intensity dependent polarization $M_z(0, I)$ in Equation (6.37) was taken to be the polarization generated by 160 s of irradiation under above gap light, i.e.,

$$M_z(0, I) = \left(\frac{2\alpha_{\text{BE}\downarrow} p_{\text{F}\downarrow}}{\alpha_{\text{BE}\downarrow} p_{\text{F}\downarrow} + \alpha_{\text{BE}\uparrow} p_{\text{F}\uparrow}} - 1 \right) \left(1 - \exp\left[-tI(\alpha_{\text{BE}\downarrow} p_{\text{F}\downarrow} - \alpha_{\text{BE}\uparrow} p_{\text{F}\uparrow})\right]\right), \tag{6.42}$$

where $t = 160$ s. We present our predictions for the polarization dynamics along with the experimental data in Figure 6.11. The starkest discrepancies between our experimental

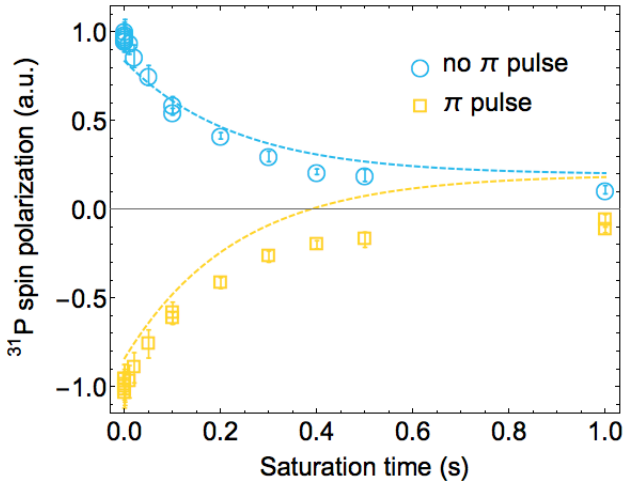

Figure 6.11: Data from Figure 6.9(b) and the predictions by Equation (6.37) for $\Omega/(2\pi) = 4$ kHz, $\Delta\nu = \delta$, given the parameters extracted from the fits in Figure 6.10.

data and the dynamics predicted by the free exciton capture dependent hyperpolarization model seen in Figures 6.10 and 6.11 arise because of the local light intensity I dependent nature of the that model. The $\Omega = 0$ build up curve, shown in Figure 6.10(a), reflects the same deviation from an exponential build up as was already shown in Figure 6.8(d), this is, of course, because Equation (6.42) matches the functional form of the I dependent build up dynamics used for produce the latter.

6.6 Conclusions and Discussion

We have demonstrated fast hyperpolarization of ^{31}P donor nuclei in silicon. Our polarization build up time constant of 51.5 s is a factor of three shorter than previously reported results under a very similar optical DNP scheme [77]. We have also shown that local light intensity dependent polarization mechanism is not consistent with our data and that our results would allow only for a very small $\sim 10^{-5}$ correlation between nuclear spin flip and photoionization events.

It has to be said that the ^{31}P spin probe, that we used for characterizing the free exciton capture events, detects only the rate of transitions into a donor state with vanishing hyperfine interaction, along with the average time spent in such state. Hence, we must rule out other mechanisms that would yield similar experimental signatures. There are

two distinct types of such processes: ionizations events of other origin and events that transform the electronic state of the donor to one with no hyperfine interaction, such as excitation into higher orbital states through phonon absorption or the formation of D^- states by trapping of free electrons. Ionization events other than D^0X Auger events could potentially arise from direct ionization of donor electrons with 1047 nm photons – a highly non-resonant process, as it requires electronic states 1139 meV above the conduction band minimum [127] – or impact ionization events that involve Auger electrons or high energy phonons, which could both have up to 1100 meV of energy.

In this paragraph, we will rule out processes that produce electronic states that have a vanishing hyperfine interaction with the ^{31}P nucleus. Excited orbital states of the defect electron have vanishing Fermi contact hyperfine interaction with the ^{31}P nucleus, and the electron could theoretically be driven into these states by the absorption of specific THz phonons. Nevertheless, the lifetime of $2p_0$ state in isotopically purified ^{28}Si at $\sim$5 K has been measured to be 235 ps [53], while the rest of the orbital states are expected to have shorter lifetimes [119, 114]. These lifetimes would be at least three orders of magnitude shorter than the $1/\kappa_{D^+\rightarrow D^0}$ measured. We can also ignore the trapping of free electrons at the phosphorus defect sites discussed and utilised in previously reported measurements by [112, 80], even though the two-electron D^- electronic state of the donor has near zero hyperfine interaction with the defect nucleus. The results in [80], carried out at very similar temperatures and magnetic fields as ours, indicate that in the presence of above band gap light the lifetimes of D^- states are up to two orders of magnitude longer than the $1/\kappa_{D^+\rightarrow D^0}$ in our experiments. Therefore, we are justified to conclude that the events that result vanishing hyperfine interaction for the ^{31}P nuclei are indeed ionization events.

It is reasonable to assume that for n-type silicon crystals, irradiated with above gap light at low temperatures, almost every free electron-hole pair generated will end up forming a free exciton [118] and almost every free exciton ends up bound to a defect site [16, 44], such bound exciton complexes do in turn almost always recombine through the Auger ionization process [100, 111]. Hence, we can safely assume that the direct photoionization events will be greatly outnumbered by the Auger ionization events. We are also justified to ignore the generation of bound multiexciton complexes, which do themselves lead to Auger ionization events, because such events are rather infrequent at our laser intensities and should not constitute more than 5% of the events even for the brightest spots in our sample [111]. Nevertheless, since the Auger electrons have to dissipate around 1100 meV of kinetic energy in the crystal it could theoretically lead to impact ionization events that accompany the Auger ionization events.

In order to conclusively attribute the photoionization observed to D^0X Auger events, we will now compare the photon absorption rate in our sample with our estimate for the

total rate of phosphorus donor defect photoionization events integrated across the sample. From the data we presented in the previous section we can estimate the rate of donor ionizations events integrated across the sample at full laser power as

$$N_{\text{ionized}} = \int dI\ \eta(I)\frac{\alpha_{D^0\to D^0X}I}{\alpha_{D^+\to D^0}\sqrt{I}}N_{^{31}\text{P}}V_{\text{sample}}\ \alpha_{D^0\to D^0X}I = 2.3\times 10^{14}\ \text{Hz}, \tag{6.43}$$

where $\eta(I)\alpha_{D^0\to D^0X}I/(\alpha_{D^+\to D^0}\sqrt{I})N_{^{31}\text{P}}V_{\text{sample}}$ stands for the number of ionized donors at a given light intensity I, while the sample volume $V_{\text{sample}} = 2.12\times 2.22\times 8\ \text{mm}^3$ and the donor defect concentration $N_{^{31}\text{P}} = 1.5\times 10^{15}\ \text{cm}^{-3}$ are the ones given in Section 6.4.3. Here, we have ignored the rate $\kappa_{D^X\to D^+}$, since it is much greater than $\kappa_{D^+\to D^0}$ for all values of I. At the same time, the photon absorption rate in our sample can be estimated from the total laser power deposited 0.232 mW, that was given in Section 6.4.4, and a single photon energy of 1184 meV for 1047 nm wavelength. The photon absorption rate evaluates to 0.232 mW/(1184 meV) $= 1.2\times 10^{15}$ Hz. The fact that these estimates are reasonably close, and most importantly, that the estimated photon absorption count does not outnumber the estimated ionization events makes a very strong case for assuming that the principal D^0 photoionization process observed converts a single photon absorbed into no more than a single ionization event. Finally, combining the arguments in this and the preceding paragraph imply that we can indeed attribute the photoionization events observed to D^0X Auger recombination events. In addition, the photoionization rate we have measured agrees with the expected free exciton capture times quoted in [101] when adjusted for the light intensity difference.

Of course, we cannot rule out the nuclear spin state dependent free exciton capture at some ($< 10^{-5}$) level, however, the fact that the free exciton capture proceeds relatively independently of the ^{31}P spin state is not necessarily a surprise. It was shown in [101] that ^{209}Bi donor bound exciton photoluminescence line does not disappear when nuclear hyperpolarization levels reach 79% at 6 T and 1.5 K, the same manuscript also suggests a pathway for the free exciton capture when the two electrons involved are both in their spin ground states. D^0X in its excited orbital state $1s\Gamma_{3,5}$ has a parallel electron spin configuration and this orbital state could serve as an intermediate step for the formation of a bound exciton in its electron spin singlet orbital ground state $1s\Gamma_1$ [101]. The spin angular momentum could then be lost through spin-orbit coupling.

We believe that this makes a strong case for the dominant DNP mechanism being the phononic mechanism, that was proposed in [77], and detailed in this chapter, for the non-resonant optical pumping protocol with 1047 nm light at high magnetic fields and low temperatures. The resulting precise understanding of the hyperpolarization mechanism enables optimization of devices and experiments which exploit the protocol for nuclear

spin signal enhancement, and could be especially useful for magnetometry applications given the extremely long coherence times of ^{31}P defect nuclei in isotopically enriched ^{28}Si crystals.

Finally, our experimental approach equips us with a novel technique for direct characterization of donor defect specific ionization events. We utilized this ability to measure the free exciton capture and photoneutralization rates. These rates play an important role for numerous experimental protocols that utilise above band gap light during measurement process [98, 67], as well as experiments that might suffer from exciton induced decoherence. We also measured the paramagnetic shift for ionized phosphorus donor defects which has not been previously reported. We will discuss other potential uses of our experimental technique in the next section.

6.7 Outlook

At first glance, it might seem that the phononic hyperpolarization mechanism and the free exciton capture mechanism would prescribe rather similar steps for device optimization. For both models the DNP time scales inversely with light intensity and they both require high electron spin polarization which, in the case of thermal polarization, implies having $|\hbar\gamma_e B_0\Theta| \gg 1$, where γ_e is the electron gyromagnetic ratio and Θ is inverse temperature. Nevertheless, there are some important differences between them. Most obviously the exciton capture mechanism requires high and reasonably uniform light intensities throughout the sample of interest, which the phononic mechanism does not. Secondly, and less obviously, the phononic mechanism is much more dependent on the particular values of Θ and B_0 than the exciton capture model, the latter being largely determined by their product. Specifically, the exciton capture mechanism does not rely on the short $T_1^{(e)}$ at high B_0 quite like the phononic mechanism. Lastly, the phononic mechanism could be substantially enhanced by increasing the strain induced modulation of the electron-nuclear hyperfine interaction, which could be achieved by engineering phononic resonator structures.

There exist an ever-increasing effort for improving the understanding and modelling accuracy of low temperature properties of isotopically enriched ultra-pure silicon and shallow donor defects within it, this is brought about by accelerating attempts to use this system as a platform for quantum computing [56, 106, 92, 113, 115, 11, 52]. Excitonic and phononic properties are certainly ones that matter for device engineering. It could prove to be very valuable to harness the tremendously powerful aspects magnetic resonance – its element selectivity and an ability to sense very weak interactions and rare events – for investigating

processes and interactions that are conventionally outside of the typical domain of NMR or ESR.

In this section, we discuss some neat opportunities that appear as natural extensions of the theoretical and experimental work presented in this chapter. First, our experimental technique for directly characterizing phosphorus donor ionization events in silicon at low temperatures could be used for measuring quantities such as the cross section for direct D^0 ionization with below gap light. Such quantities are often extracted from photoconductivity experiments [127], which are not defect specific, and also only characterize the resulting photocarrier densities rather than the photoionization and photoneutralization rates. One convenient aspect about below band gap light is that sufficiently below gap light is not expected to produce many ionization events other than the direct ionization, as it is unlikely to create electron-hole pairs and the accompanying Auger events. Although the expected ionization rates for below gap light can be very small, the very long $T_1^{D^0}$ at low temperatures provides a very long measurement window. For the average photoneutralization rate of 470 kHz, that we quote in Table 6.1, the average loss of nuclear spin polarization under $\Omega/(2\pi) = 4$ kHz strong resonant saturation drive from a single ionization event can be calculated from Equation 6.29 and evaluates to 0.003. 72 photoionization events would reduce the spin polarization to $(1-0.003)^{72} = 0.8$ of its initial value, if the 20% polarization reduction is detected over a period of 2 h – a period over which the thermal relaxation of ^{31}P spins does not reduce their polarization by more than 10% [38] – photoionization rates of 0.01 Hz are possible to detect. This rate is 10^5 times lower than the photoionization rates we have measured here. Furthermore, it is reasonable to expect that the photoneutralization rates under below gap irradiation are also substantially lower from the above gap values, this fact would serve to boost the sensitivity of the measurement described even further.

Our experimental protocol could also enable measuring other unreported quantities such as the bound exciton's hole hyperfine coupling to the donor nucleus, which would serve to improve fundamental understanding of silicon physics. An experiment for measuring that quantity will be outlined in the next subsection. Finally, it will be shown in Subsections 6.7.2, 6.7.3 and 6.7.4 that the effective Hamiltonian analysis carried out in Section 6.1 can be extended to analyse the interplay between external electric fields and strain fields in electron-nuclear Hamiltonians. Such interplay could potentially be harnessed to measure frequency specific properties of sub-terahertz phonons.

6.7.1 Hole hyperfine

Here, we outline how an experiment similar to the ones described in Section 6.5 can be used to measure the hyperfine coupling A_h between the ^{31}P nuclear spin and the bound exciton

hole. This quantity has never been reported as it is much too small to be able to quantify it in photoluminescence experiments that have been used to characterize the gyromagnetic ratios of the hole [55, 67]. Direct magnetic resonance detection of the hyperfine constant is also impossible due to the short 272 ns lifetime of the bound excitons. The knowledge of A_h would, nevertheless, be both of fundamental and practical interest as it could lead to a refined understanding of the hole wave function as well as help to quantify nuclear decoherence processes under exciton creation.

Because the lifetime of the bound excitons is extremely short, it is necessary to use high saturation drive strengths to measure the hole hyperfine coupling. For that we can use the same experimental protocol as in Figure 6.9(a) – first polarizing the nuclear spins with 160 s laser irradiation to near terminal 96% polarization and then applying the RF drive simultaneously with the laser for different frequency offsets $\Delta\nu$, for some fixed period t. Plotting the ^{31}P polarization at the end of the saturation period as a function of $\Delta\nu$ would produce dip plots similar to Figure 6.8(b) with an extra dip at the value of $-3A_h/2$.

We now explore the feasibility and the resolution of such an experiment by simulating its outcome for various values of A_h. To simulate the result of such an experiment we can use the full rate equations model given in Section 6.2 under the generator $\Upsilon(\kappa_{D^0\to D^0X}, \kappa_{D^+\to D^0}, \kappa_{D^+\to D^0}, \Delta\nu, A_h)$ defined by Equation (6.16). To fix the other parameters appearing in Υ we assume the saturation drive Rabi strength to be the same as that for the experiment in Figure 6.9, i.e, $\Omega/(2\pi) = 4$ kHz, we also use $A_e(0)/(2\pi) = 117.52$ MHz, $T_1^{D^0}$ and δ that are stated in Table 6.1 and $T_2^{D^0} = 5.4$ ms under 1047 nm laser light illumination reported in [38]. Lastly, we need to estimate $T_2^{D^0X}$ – the ^{31}P nuclear spin dephasing time during the bound exciton stage in Figure 6.1. In Appendix D we derive the electron $T_1^{(e)}$ induced effective $T_2^{D^0}$ for the nuclear spin, the same calculation is also applicable for hole T_1^h induced $T_2^{D^0X}$. Because it is known that the holes thermalize within their lifetime [65], i.e., $T_1^{(h)} \ll 272$ ns, and that the hole hyperfine interaction $A_h/(2\pi)$ cannot exceed $\sim$20 MHz we can assume that $T_1^{(h)} A_h \ll 1$. Therefore, the hole induced $T_2^{D^0X}$ can be ignored, and in our simulations we set $T_2^{D^0X} = 10 \times 272$ ns.

In order to calculate the resulting nuclear magnetization $M_z^{D^0}$ after a t-long lasing and saturation period for various A_h values, we have to account for the light intensity distribution in the sample. Accounting for that gives

$$M_z^{D^0}(\Delta\nu, A_h, t) = \left(\int \eta(I) \ \exp\left[\Upsilon(\alpha_{D^0\to D^0X} I, \kappa_{D^+\to D^0}, \alpha_{D^+\to D^0}\sqrt{I}, \Delta\nu, A_h) \ t \right] \cdot \vec{m}_0 \right)_{11}, \tag{6.44}$$

where $\vec{m}_0 = (0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0.96, 1)$ and the subscript denotes the 11th element of

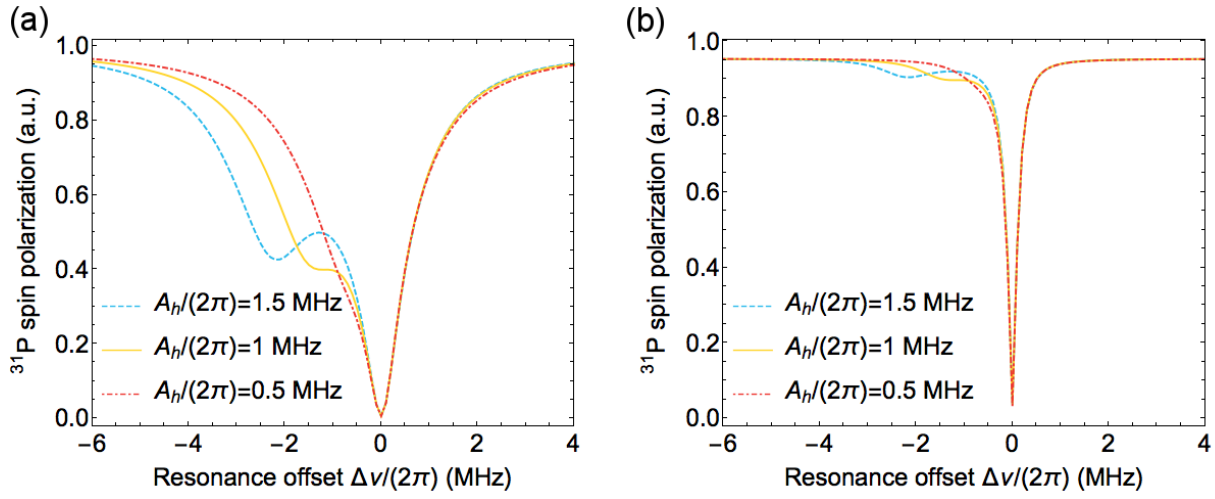

Figure 6.12: (a) Simulation of $M_z^{D^0}(\Delta\nu, A_h, t)$ in Equation 6.44 for $t = 100$ s and $A_h/(2\pi) = \{0.5 \text{ MHz}, 1 \text{ MHz}, 1.5 \text{ MHz}\}$. (b) Simulation of $M_z^{D^0}(\Delta\nu, A_h, t)$ for $t = 1$ s and $A_h/(2\pi) = \{0.5 \text{ MHz}, 1 \text{ MHz}, 1.5 \text{ MHz}\}$.

the integrated magnetization vector, $\kappa_{D^+\to D^0} = 1/(272 \text{ ns})$ and the values $\alpha_{D^0\to D^0 X}$ and $\alpha_{D^+\to D^0}$ are given in Table 6.1. We simulate the integrated magnetization $M_z^{D^0}(\Delta\nu, A_h, t)$ for two separate saturation periods $t = \{1 \text{ s}, 100 \text{ s}\}$ and three distinct hyperfine constants $A_h/(2\pi) = \{0.5 \text{ MHz}, 1 \text{ MHz}, 1.5 \text{ MHz}\}$ over a range of $-6 \text{ MHz} \leq \Delta\nu/(2\pi) \leq 4 \text{ MHz}$ and display our results in Figure 6.12.

It is evident from Figure 6.12 that hyperfine values of $A_h > 1$ MHz should be observable under our experimental protocol for a 100 s-long lasing and saturation period, even in the presence of our typical error bars of $\sim$5% for the terminal magnetization, whereas 1 s-long saturation period is too short to yield measurable features. Unfortunately, conducting such experiment was not feasible with our current apparatus due to the tuning range of our NMR probe head, and the fact that measuring $M_z^{D^0}$ for different $\Delta\nu$ values would have required retuning the probe and, hence, recalibrating the Ω value for each data point. Nevertheless, with an extra ability to sweep the external magnetic field B_0 between the saturation period and the signal read out step, while keeping the probe configuration the same, the experiment would have certainly become practical.

6.7.2 Strain and Electric Field

In the next two subsections, we point out two novel experimental opportunities stemming from the fact that Stark shifts due to external electric fields and crystal strain enter the electron-nuclear Hamiltonians in rather similar ways, since they both deform the electron wave function, and hence, affect both the hyperfine interaction strength and the electron g-tensor. Such interplay between crystal strain and electric fields has so far only been systematically studied by computational methods in [115], nevertheless, several experimental results [9, 90] have strongly hinted at the existence of an observable combined effect on the hyperfine constant. We highlight that the stated interplay could be incorporated into new device designs, or harnessed for measuring frequency specific phonon properties for sub-terahertz phonons in bulk silicon at currently unexplored temperatures.

Here, we do not assume any specific means for generating elastic waves that give rise to time varying crystal strain inside the device, as there is a multitude of ways that could be used for that, although care should be taken to eliminate other complementing, yet unwanted, effects of such phonon generation, e.g., free carrier and exciton creation with the use of above band gap light. We do, however, note that since we are dealing with effective electron-nuclear cross relaxation just like in Section 6.1, we do not necessarily assume such waves to have spatial nor temporal coherence. We treat the hyperfine interaction in the presence of crystal strain $\vec{\epsilon}$ and external electric field $\vec{E} = (E_x, E_y, E_z)$ in Appendix E by extending the symmetry analysis in [73]. Taylor expanding up to second order in $\vec{E}$ and $\vec{\epsilon}$ we find that

$$A_e\left(\vec{E},\vec{\epsilon}\right) = A_e(0)\left(1 + (\vec{\nabla}A_e)\cdot\vec{\epsilon} + \vec{\epsilon}\cdot\mathbf{H}_{A_e}\cdot\vec{\epsilon} + U\vec{E}\cdot\vec{\epsilon}^{\,(4,5,6)} + V\left|\vec{E}\right|^2 + ...\right), \tag{6.45}$$

where $\vec{\nabla}A_e$ and $\mathbf{H}_{A_e}$ are the same rescaled gradient vector and Hessian matrix as in Section 6.1, while $\vec{\epsilon}^{\,(4,5,6)} = \vec{\epsilon}^{\,(4,5,6)}(\vec{\epsilon}) = (2\epsilon_{yz}, 2\epsilon_{xz}, 2\epsilon_{xy})$ is a vector containing only the off-diagonal elements of the strain tensor evaluated in the crystallographic coordinate system. The value of $V = -2.7 \times 10^{-15}$ m^2/V^2 has been determined experimentally [102], while U, which represents the Stark shifts arising from the combined effect of strain and electric field, is unknown because the only current study dealing with the interplay of $\vec{\epsilon}$ and $\vec{E}$ [115] only simulated uniaxial stresses for which $\vec{\epsilon}^{\,(4,5,6)} = 0$.

Again, we make the adiabatic approximation discussed in Section 6.1, along with the assumption that the wavelength of phonons under consideration is sufficiently long, such that the strain dependent hyperfine modulation inherits all of its properties from static strains. Additionally, we once more ignore all strain and electric field dependent changes to the electron g-tensor, partly because such changes can be suppressed by working at

sufficiently low external fields, but more importantly, since here we are solely focused on the nuclear spin polarization, and the accompanying electron spin dynamics is rather irrelevant for us, this as long as the time dependent $\vec{\epsilon}$ and $\vec{E}$ do not appreciably saturate the electron spin transition ω_e.

We now concentrate on a single phosphorus defect in silicon in the presence of strain field $\vec{\epsilon} = \sum_i \vec{\epsilon}_i \cos(\omega_i t + \phi_i)$, where $\vec{\epsilon}_i$ denotes the amplitude of elastic waves with frequency and phase ω_i and ϕ_i, respectively. Furthermore, if that defect site is also experiencing an external AC electric field with frequency ω_{E} and amplitude $\vec{E}_0$, such that $\vec{E} = \vec{E}_0 \cos(\omega_{\mathrm{E}} t)$, then the secular part of

$$\begin{aligned} H = &\frac{\omega_e}{2}\sigma_z^{(e)} - \frac{\omega_{^{31}\mathrm{P}}}{2}\sigma_z^{(n)} + \frac{A_e(0)}{4}\left(1 + (\vec{\nabla}A_e)\cdot\vec{\epsilon} + \vec{\epsilon}\cdot\mathbf{H}_{A_e}\cdot\vec{\epsilon}\right)\sigma^{(e)}\cdot\sigma^{(n)} \\ &+ \frac{A_e(0)}{4}\left(U\vec{E}\cdot\vec{\epsilon}^{\,(4,5,6)} + V\left|\vec{E}\right|^2\right)\sigma^{(e)}\cdot\sigma^{(n)} \end{aligned} \tag{6.46}$$

with respect to $H_0 = \frac{\omega_e}{2}\sigma_z^{(e)} - \frac{\omega_{^{31}\mathrm{P}}}{2}\sigma_z^{(n)}$ becomes

$$\begin{aligned} H_{\mathrm{sec}} \sim eq\frac{A_e(0)}{4}\sigma_z^{(e)}\sigma_z^{(n)} &+ \frac{A_e(0)}{8}\sum_i \delta_{\omega_e+\omega_{^{31}\mathrm{P}}-\omega_i}(\vec{\nabla}A_e)\cdot\vec{\epsilon}_i\ H_{\text{flip-flop}} \\ &+ \frac{A_e(0)}{16}\left(\sum_{\substack{\omega_i+\omega_j=\\ \omega_e+\omega_{^{31}\mathrm{P}}}} + \sum_{\substack{\omega_i-\omega_j=\\ \omega_e+\omega_{^{31}\mathrm{P}}}}\right)\vec{\epsilon}_i\cdot\mathbf{H}_{A_e}\cdot\vec{\epsilon}_j\ H_{\text{flip-flop}} \\ &+ \frac{A_e(0)}{8}\sum_i \delta_{\omega_e+\omega_{^{31}\mathrm{P}}-\omega_{\mathrm{E}}-\omega_i} U\vec{E}_0\cdot\vec{\epsilon}_i^{\,(4,5,6)}\ H_{\text{flip-flop}} \\ &+ \frac{A_e(0)}{8}\sum_i \delta_{\omega_e+\omega_{^{31}\mathrm{P}}+\omega_{\mathrm{E}}-\omega_i} U\vec{E}_0\cdot\vec{\epsilon}_i^{\,(4,5,6)}\ H_{\text{flip-flop}} \\ &+ \frac{A_e(0)}{16}\delta_{\omega_e+\omega_{^{31}\mathrm{P}}-2\omega_{\mathrm{E}}} V\left|\vec{E}_0\right|^2 H_{\text{flip-flop}}, \end{aligned} \tag{6.47}$$

where δ_j is the Kronecker delta function, i.e., $\delta_0 = 1$ and $\delta_j = 0$ for all $j \neq 0$. In Equation (6.47), we have omitted all corrections to the $\sigma_z^{(e)}\sigma_z^{(n)}$ part of the Hamiltonian since they will always be much smaller than $A_e(0)/4$.

6.7.3 DNP via Electric Field Modulation

In this subsection, we assume all physical properties of the donor defect to be bulk properties, hence, we assume the envisioned devices to work with a sizeable ensemble of donor

defects in contrast to devices that utilize a single near-surface defect as in [92]. While single defect devices have several truly novel uses, ensemble devices are likely more fitting for some applications such as efficient high accuracy, high sensitivity magnetometry, where the incredibly long coherence times of ionized bulk ^{31}P nuclear spins in isotopically purified silicon [98] could be suitably employed.

All electric control of phosphorus donor nuclear spins in silicon was proposed not so long ago [19, 113], and has very recently been experimentally demonstrated for an ensemble of phosphorus defects [103]. This experiment relies on electron g-tensor modulation - by periodically tilting the axis of quantization for the electron spin with respect to that of the ^{31}P an effective Rabi field on the nucleus is realized, this is in essence equivalent to the hyperfine mediated Rabi field that was discussed in Section 6.4.1. While this technique does allow for universal control of the electron-nuclear system, it would require multiple pulse or optimal control methods over multiple transmitter frequencies to wholly transfer the electron polarization over onto the ^{31}P, such approach could be rather inefficient, especially in the presence of control field variations over the sample.

Nevertheless, one observes from the last line of Equation (6.47) that it is possible to directly drive the $|{\downarrow_e}{\uparrow_n}\rangle \leftrightarrow |{\uparrow_e}{\downarrow_n}\rangle$ transition with AC electric fields that modulate the hyperfine coupling at frequency $\omega_{\mathrm{E}}^{\mathrm{DNP}} = (\omega_e + \omega_{^{31}\mathrm{P}})/2$. Electric field induced perturbations to the hyperfine tensor, in contrast to the electron g-tensor perturbations, have the advantage of being B_0 independent, the maximum effective Rabi strength $\Omega_{\mathrm{E}}^{\mathrm{max}}$ for driving the $|{\downarrow_e}{\uparrow_n}\rangle \leftrightarrow |{\uparrow_e}{\downarrow_n}\rangle$ transition is limited only by the ionization voltage predicted to be $1.55 \cdot 10^6$ V/m [90]. Assuming that the shift in the hyperfine strength remains predominantly quadratic in the electric field magnitude up to the ionization point, we would estimate that $\Omega_{\mathrm{E}}^{\mathrm{max}} = 95$ kHz. Furthermore, if the objective is to merely transfer electron polarization onto the ^{31}P, electric fields only at a single frequency $\omega_{\mathrm{E}}^{\mathrm{DNP}}$ are required and field inhomogeneities become irrelevant because in the presence of $T_1^{(e)}$ they only affect the rate at which the electron and nuclear polarizations are equalized. We note that $\Omega_{\mathrm{E}}^{\mathrm{max}}$ is four orders of magnitude faster than the fastest reported hyperpolarization scheme for bulk ^{31}P nuclei [126], and is as fast as the incredibly rapid hyperpolarization rates reported in [49], the latter did, however, require introducing oxygen vacancy defects into the sample which come at the cost of reduced ^{31}P coherence time.

This method would be particularly useful for devices and experiments at relatively low (< 1 T) magnetic fields at which the frequency $\omega_{\mathrm{E}}^{\mathrm{DNP}}$ is not prohibitive and AC electric fields can be generated with relative ease. High electron polarization, necessary for practical DNP, could under such conditions be achieved either via cooling the system down to dilution fridge temperatures, where optical DNP schemes could be severely constrained due to input power limitations, or by hyperpolarizing the defect electrons at higher tem-

peratures, e.g., through resonant optical pumping [126] or spin injection [99].

6.7.4 Electric Field Assisted Phonon Detection

The fourth and the fifth line in Equation (6.47) have extremely interesting and potentially rather useful implications. We observe that by applying AC electric fields of sufficiently large amplitude $\vec{E}_0$, for which $\omega_{\mathrm{E}} \neq (\omega_e + \omega_{^{31}\mathrm{P}})/2$, one can eclipse all other terms in Equation (6.47), and make the polarization transfer rate from the electron to the nucleus dependent only on

$$H_{\text{el-phonon}}\left(\vec{E}_0, \vec{\epsilon}_k, \vec{\epsilon}_l\right) = \frac{A_e(0)}{8} P \vec{E}_0 \cdot \left(\vec{\epsilon}_k^{\,(4,5,6)} + \vec{\epsilon}_l^{\,(4,5,6)}\right), \tag{6.48}$$

where $\omega_k = \omega_e + \omega_{^{31}\mathrm{P}} - \omega_{\mathrm{E}}$ and $\omega_l = \omega_e + \omega_{^{31}\mathrm{P}} + \omega_{\mathrm{E}}$. This is significant because it provides a local frequency selective probe for elastic wave amplitudes inside silicon crystals, furthermore, rotating the direction of $\vec{E}_0$ yields information about the polarization of such waves. We note that, since $\vec{\epsilon}^{\,(4,5,6)}$ contains only the off-diagonal elements of the strain tensor evaluated in the crystallographic coordinate system, $\vec{\epsilon}^{\,(4,5,6)} = 0$ for longitudinal phonons travelling along any of the three crystal axes, nevertheless, for transverse phonons or for longitudinal phonons along other directions $\vec{\epsilon}^{\,(4,5,6)}$ is non-zero. For this combined effect to serve as a means for elastic wave amplitude detection one needs $H_{\text{el-phonon}}\left(\vec{E}_0, \vec{\epsilon}_k, \vec{\epsilon}_l\right) \gg 1/T_1^{D^0}$, where $T_1^{D^0}$ is the ^{31}P nuclear longitudinal relaxation time, encouragingly, these time scales have been measured to be ∼700 s at 4.2 K and $\gg 10^4$ s at 1.7 K for isotopically purified ^{28}Si samples [38].

Observing the ^{31}P polarization build up time constant inside a ^{28}Si sample, as a function of the direction, frequency and amplitude of the external AC electric field applied, gives one a time-resolved, frequency and polarization selective probe for low-frequency phonons at cryogenic temperatures. Conveniently, all of the listed parameters can be chosen and changed rather easily with an appropriate experimental apparatus. Such a probe could, for example, be used for characterizing the density of states or non-linear effects in phononic processes. Moreover, the local nature of donor defects, and the fact that inter-defect nuclear spin diffusion rates can be controlled by varying doping concentrations, could enable a new and unique way for measuring location specific phonon properties inside silicon crystals. For this, the AC electric field modified phononic DNP scheme could be paired with MRI so to spatially resolve the ^{31}P polarization time constants. The dephasing time limited spatial encoding resolution for MRI is given by $K = \pi / \left(\gamma_{^{31}\mathrm{P}}\, \partial B_0\, T_2^{D^0}\right)$ [72], where $\gamma_{^{31}\mathrm{P}}$ is the gyromagentic ratio of phosphorus, ∂B_0 is the static field gradient and $T_2^{D^0}$ is the

nuclear dephasing time. Importantly, $T_2^{D^0}$ is not necessarily the dephasing time during the DNP process, since the polarization transfer and MRI encoding periods can be separated. The measured $T_2^{D^0} = 1.75$ s at 5.5 K for $B_0 \approx 0.3$ T [81] would yield $K = 0.17$ μm even for rather modest field gradients of $\partial B_0 = 0.1$ T m^{-1}, which would certainly be sufficient resolution to enable the measurement of phonon attenuation lengths and possibly phonon scattering off surfaces or boundaries in heterostructures.

Chapter 7

Conclusion

The sensitivity and efficiency of experiments, as well as technologies, based on the phenomenon of magnetic resonance are intimately linked to the coherence time and polarization of the spins that are being detected. Therefore, increasing spin coherence time and polarization are paramount for advancing the scope and limitations of magnetic resonance applications. With this thesis we have made significant progress in addressing both of these aspects critical for magnetic resonance through effective Hamiltonian engineering.

We have developed tools for numerical effective Hamiltonian engineering, which can be used to engineer arbitrary spin dynamics with much fewer constraints than analytical Hamiltonian engineering would pose. This involved developing the Van Loan block matrix differential equation framework, that enables easy and numerically efficient evaluation of various perturbation expressions appearing in effective Hamiltonian engineering problems. Furthermore, we have demonstrated that, in the case of piecewise constant control amplitudes, the Van Loan differential equation framework provides a straight-forward way for calculating gradients of these perturbation expressions with respect to the control amplitudes, which in turn enabled us to use efficient numerical control optimization schemes for engineering effective Hamiltonians. Numerical Hamiltonian engineering is not specific to any specific Hilbert space dimension, and does not require any particular assumptions about the form of the control and drift Hamiltonians; it is also naturally suited for handling ensemble effects and deterministic control waveform distortions due to control hardware, that are both difficult to account for in analytic protocols. We expect our effective Hamiltonian engineering tools to be useful for a great range of settings including sensing, spectroscopy and quantum computing, and not necessarily limited to magnetic resonance applications, since the main requirements for employing the framework are a finite-dimensional first-order linear matrix differential equation description for the sys-

tem dynamics and sufficiently well characterized control amplitudes that can be efficiently modulated.

We put our effective Hamiltonian engineering tools into action in designing control sequences for force detected magnetic resonance experiments that worked with nuclear spin ensembles on the nanometre scale, and posed a very challenging control setting due to the short spin coherence time and great Rabi field dispersion across the spin ensemble. Our numerically engineered control sequences lead to 500 times increase in nuclear spin coherence time, a factor that directly translates to maximum achievable spatial resolution for MRI experiments. We demonstrated experimentally that control sequences, which are designed to coherently average unwanted Hamiltonian Dyson terms to near-zero values, approximate infinitesimally short spin rotations, as they suspend the unwanted Hamiltonian terms for the duration of the sequence. This fact unlocks the ability to use the immense wealth of various NMR sequences, developed for a wide range of applications, including imaging and multi-dimensional spectroscopy, also for unconventional and challenging control settings, just by swapping out the rotations within the sequences with numerically engineered controls. We expect this ability to be extremely useful in future experiments on the same force detected magnetic resonance setup in investigating nanoscale nuclear spin ensembles, as well as for other non-standard magnetic resonance experiments. Finally, the tools developed are not only useful for suspending the unwanted Hamiltonian evolution and thereby enhancing spin coherence times, it is conceivable that they could also aid with the second important challenge pertaining to magnetic resonance, that is spin polarization enhancement. By engineering effective Hamiltonians for multi-body systems polarization transfer schemes can potentially be implemented, which could lead to improved sensitivity in magnetic resonance experiments.

With the second half of this thesis we used a combination of theoretical modelling and experiments to demonstrate how effective spin Hamiltonians lead to nuclear spin hyperpolarization. The physical platform for our experiments were phosphorus donor nuclei in high-purity isotopically enriched silicon-28 crystals, which are known to have incredibly long coherence times in their ionized state, and are therefore an excellent candidate for sensing, and specifically for magnetometry applications. The non-resonant optical hyperpolarization scheme used for the work in this thesis is an attractive method for enhancing spin detection sensitivity by several orders of magnitude due to its particularly simple and relatively fast nature. Our detailed microscopic effective Hamiltonian model could enable optimizing experiments and devices making use of the defect nuclei.

References

[1] A. Abragam. *The Principles of Nuclear Magnetism*. Oxford University Press, London, 1961.

[2] R. L. Aggarwal and A. K. Ramdas. Optical determination of the symmetry of the ground states of group-V donors in silicon. *Physical Review*, 140(4A):A1246–A1253, 1965.

[3] B. Andreas, Y. Azuma, G. Bartl, P. Becker, H. Bettin, M. Borys, I. Busch, M. Gray, P. Fuchs, K. Fujii, H. Fujimoto, E. Kessler, M. Krumrey, U. Kuetgens, N. Kuramoto, G. Mana, P. Manson, E. Massa, S. Mizushima, A. Nicolaus, A. Picard, A. Pramann, O. Rienitz, D. Schiel, S. Valkiers, and A. Waseda. Determination of the Avogadro constant by counting the atoms in a ^{28}Si crystal. *Physical Review Letters*, 106(3):30801, 2011.

[4] P. Becker, H. J. Pohl, H. Riemann, and N. Abrosimov. Enrichment of silicon for a better kilogram. *Physica Status Solidi (A) Applications and Materials Science*, 207(1):49–66, 2010.

[5] S. Blanes, F. Casas, J. A. Oteo, and J. Ros. The Magnus expansion and some of its applications. *Physics Reports*, 470(5-6):151–238, 2009.

[6] T. W. Borneman and D. G. Cory. Bandwidth-limited control and ringdown suppression in high-Q resonators. *Journal of Magnetic Resonance*, 225:120–129, 2012.

[7] T. W. Borneman, M. D. Hürlimann, and D. G. Cory. Application of optimal control to CPMG refocusing pulse design. *Journal of Magnetic Resonance*, 207(2):220–233, 2010.

[8] G. S. Boutis, P. Cappellaro, H. Cho, C. Ramanathan, and D. G. Cory. Pulse error compensating symmetric magic-echo trains. *Journal of Magnetic Resonance*, 161(2):132–137, 2003.

[9] F. R. Bradbury, A. M. Tyryshkin, G. Sabouret, J. Bokor, T. Schenkel, and S. A. Lyon. Stark Tuning of Donor Electron Spins in Silicon. *Physical Review Letters*, 97(17):176404, 2006.

[10] H. P. Breuer and F. Petruccione. *The Theory of Open Quantum Systems*. OUP Oxford, 2007.

[11] M. A. Broome, S. K. Gorman, M. G. House, S. J. Hile, J. G. Keizer, D. Keith, C. D. Hill, T. F. Watson, W. J. Baker, L. C. L. Hollenberg, and M. Y. Simmons. Two-electron spin correlations in precision placed donors in silicon. *Nature Communications*, 9(1):980, 2018.

[12] J. Bylander, S. Gustavsson, F. Yan, F. Yoshihara, K. Harrabi, G. Fitch, D. G. Cory, Y. Nakamura, J.-S. Tsai, and W. D. Oliver. Dynamical decoupling and noise spectroscopy with a superconducting flux qubit. *Nature Physics*, 7(7):565–570, 2011.

[13] P. Cappellaro, J. S. Hodges, T. F. Havel, and D. G. Cory. Principles of control for decoherence-free subsystems. *Journal of Chemical Physics*, 125(4):44514, 2006.

[14] F. Carbonell, J. C. Jímenez, and L. M. Pedroso. Computing multiple integrals involving matrix exponentials. *Journal of Computational and Applied Mathematics*, 213(1):300–305, 2008.

[15] M. Cardona, T. A. Meyer, and M. L. W. Thewalt. Temperature Dependence of the Energy Gap of Semiconductors in the Low-Temperature Limit. *Physical Review Letters*, 92(19):196403, 2004.

[16] Y. H. Chen and S. A. Lyon. Photoluminescence and diffusivity of free excitons in doped silicon. *IEEE Journal of Quantum Electronics*, 25(5):1053–1055, 1989.

[17] D. G. Cory. A new multiple-pulse cycle for homonuclear dipolar decoupling. *Journal of Magnetic Resonance (1969)*, 94(3):526–534, 1991.

[18] D. G. Cory, J. B. Miller, R. Turner, and A. N. Garroway. Multiple-pulse methods of ^{1}H N.M.R. imaging of solids: second-averaging. *Molecular Physics*, 70(2):331–345, 1990.

[19] A. De, C. E. Pryor, and M. E. Flatté. Electric-field control of a hydrogenic donor's spin in a semiconductor. *Physical Review Letters*, 102(1):17603, 2009.

[20] C. L. Degen, Q. Lin, and B. H. Meier. Dipolar spin echoes in magnetic resonance force microscopy. *Physical Review B*, 74(10):104414, 2006.

[21] E. R. Dobrovinskaya, L. A. Lytvynov, and V. Pishchik. *Properties of Sapphire*. Springer US, 2009.

[22] F. J. Dyson. The radiation theories of Tomonaga, Schwinger, and Feynman. *Physical Review*, 75(3):486–502, 1949.

[23] K. W. Eberhardt, C. L. Degen, A. Hunkeler, and B. H. Meier. One- and two-dimensional NMR spectroscopy with a magnetic-resonance force microscope. *Angewandte Chemie - International Edition*, 47(46):8961–8963, 2008.

[24] G. Feher. Electron spin resonance experiments on donors in silicon. I. Electronic structure of donors by the electron nuclear double resonance technique. *Physical Review*, 114(5):1219–1244, 1959.

[25] G. Feher and E. A. Gere. Electron spin resonance experiments on donors in silicon. II. Electron spin relaxation effects. *Physical Review*, 114(5):1245–1256, 1959.

[26] G. Feng, F. H. Cho, H. Katiyar, J. Li, D. Lu, J. Baugh, and R. Laflamme. Gradient-based closed-loop quantum optimal control in a solid-state two-qubit system. *Phys. Rev. A*, 98:052341, 2018.

[27] B. J. Frey, D. B. Leviton, and T. J. Madison. Temperature-dependent refractive index of silicon and germanium. *Proceedings, Optomechanical Technologies for Astronomy*, 6273:62732J, 2006.

[28] K. M. C. Fu, T. D. Ladd, C. Santori, and Y. Yamamoto. Optical detection of the spin state of a single nucleus in silicon. *Physical Review B*, 69(12), 2004.

[29] M. Garwood and L. DelaBarre. The return of the frequency sweep: Designing adiabatic pulses for contemporary NMR. *Journal of Magnetic Resonance*, 153(2):155–177, 2001.

[30] D. L. Goodwin. *Advanced Optimal Control Methods for Spin Systems*. PhD thesis, University of Southampton, 2018.

[31] D. L. Goodwin and I. Kuprov. Auxiliary matrix formalism for interaction representation transformations, optimal control, and spin relaxation theories. *Journal of Chemical Physics*, 143(8):84113, 2015.

[32] D. L. Goodwin and I. Kuprov. Modified Newton-Raphson GRAPE methods for optimal control of spin systems. *Journal of Chemical Physics*, 144(20):204107, 2016.

[33] M. D. Grace, J. M. Dominy, W. M. Witzel, and M. S. Carroll. Optimized pulses for the control of uncertain qubits. *Physical Review A*, 85(5):52313, 2012.

[34] T. Green, H. Uys, and M. J. Biercuk. High-order noise filtering in nontrivial quantum logic gates. *Physical Review Letters*, 109(2):20501, 2012.

[35] T. J. Green, J. Sastrawan, H. Uys, and M. J. Biercuk. Arbitrary quantum control of qubits in the presence of universal noise. *New Journal of Physics*, 15(9):95004, 2013.

[36] T. Gullion, D. B. Baker, and M. S. Conradi. New, compensated Carr-Purcell sequences. *Journal of Magnetic Resonance (1969)*, 89(3):479–484, 1990.

[37] P. Gumann, H. Haas, S. Sheldon, L. Zhu, R. Deshpande, T. Alexander, M. L. W. Thewalt, D. G. Cory, and C. Ramanathan. Nmr study of optically hyperpolarized phosphorus donor nuclei in silicon. *Physical Review B*, 98:180405, 2018.

[38] P. Gumann, O. Patange, C. Ramanathan, H. Haas, O. Moussa, M. L.W. Thewalt, H. Riemann, N. V. Abrosimov, P. Becker, H. J. Pohl, K. M. Itoh, and D. G. Cory. Inductive measurement of optically hyperpolarized phosphorous donor nuclei in an isotopically enriched silicon-28 crystal. *Physical Review Letters*, 113(26):267604, 2014.

[39] H. Haas, D. Puzzuoli, F. Zhang, and D. G. Cory. Engineering Effective Hamiltonians. *arXiv*, 2019.

[40] J. Haase, N. J. Curro, and C. P. Slichter. Double Resonance Probes for Close Frequencies. *Journal of Magnetic Resonance*, 135(2):273–279, 1998.

[41] U. Haeberlen and J. S. Waugh. Coherent averaging effects in magnetic resonance. *Physical Review*, 175(2):453–467, 1968.

[42] S. Hafner, D. E. Demco, and R. Kimmich. Magic echoes and NMR imaging of solids. *Solid State Nuclear Magnetic Resonance*, 6(4):275–293, 1996.

[43] J. K. Hale. *Ordinary Differential Equations*. Dover Publications, 2009.

[44] R. B. Hammond and R. N. Silver. Temperature dependence of the exciton lifetime in high-purity silicon. *Applied Physics Letters*, 36(1):68–71, 1980.

[45] H. Y. Hao and H. J. Maris. Dispersion of the long-wavelength phonons in Ge, Si, GaAs, quartz, and sapphire. *Physical Review B*, 63(22):224301, 2001.

[46] J. R. Haynes. Experimental proof of the existence of a new electronic complex in silicon. *Physical Review Letters*, 4(7):361–363, 1960.

[47] I. N. Hincks, C. E. Granade, T. W. Borneman, and D. G. Cory. Controlling Quantum Devices with Nonlinear Hardware. *Physical Review Applied*, 4(2):24012, 2015.

[48] J. S. Hodges, J. C. Yang, C. Ramanathan, and D. G. Cory. Universal control of nuclear spins via anisotropic hyperfine interactions. *Physical Review A*, 78(1):10303, 2008.

[49] F. Hoehne, L. Dreher, D. P. Franke, M. Stutzmann, L. S. Vlasenko, K. M. Itoh, and M. S. Brandt. Submillisecond hyperpolarization of nuclear spins in silicon. *Physical Review Letters*, 114(11):117602, 2015.

[50] A. Honig and E. Stupp. Electron spin-lattice relaxation in phosphorus-doped silicon. *Physical Review*, 117(1):69–83, 1960.

[51] D. I. Hoult and R. E. Richards. The signal-to-noise ratio of the nuclear magnetic resonance experiment. *Journal of Magnetic Resonance (1969)*, 24(1):71–85, 1976.

[52] W. Huang, C. H. Yang, K. W. Chan, T. Tanttu, B. Hensen, R. C. C. Leon, M. A. Fogarty, J. C. C. Hwang, F. E. Hudson, K. M. Itoh, A. Morello, A. Laucht, and A. S. Dzurak. Fidelity benchmarks for two-qubit gates in silicon. *Nature*, 2019.

[53] H. W. Hübers, S. G. Pavlov, S. A. Lynch, Th. Greenland, K. L. Litvinenko, B. Murdin, B. Redlich, A. F. G. Van Der Meer, H. Riemann, N. V. Abrosimov, P. Becker, H. J. Pohl, R. Kh. Zhukavin, and V. N. Shastin. Isotope effect on the lifetime of the $2p_0$ state in phosphorus-doped silicon. *Physical Review B*, 88(3):35201, 2013.

[54] H. Huebl, A. R. Stegner, M. Stutzmann, M. S. Brandt, G. Vogg, F. Bensch, E. Rauls, and U. Gerstmann. Phosphorus donors in highly strained silicon. *Physical Review Letters*, 97(16):166402, 2006.

[55] A. S. Kaminskil, V. A. Karasyuk, and Ya. E. Pokrovskil. Luminescence of excitons bound to phosphorus atoms in silicon subjected to a magnetic field. *Soviet Journal of Experimental and Theoretical Physics*, 52(2):211–215, 1980.

[56] B. E. Kane. A silicon-based nuclear spin quantum computer. *Nature*, 393(6681):133–137, 1998.

[57] 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(2):296–305, 2005.

[58] R. Kubo. Stochastic Liouville equations. *Journal of Mathematical Physics*, 4(2):174–183, 1963.

[59] Laserglow. LRS-1047 DPSS Manual. www.laserglow.com/specsheets/RA43005FX.php, 2015.

[60] N. Leung, M. Abdelhafez, J. Koch, and D. Schuster. Speedup for quantum optimal control from automatic differentiation based on graphics processing units. *Physical Review A*, 95:042318, 2017.

[61] M. H. Levitt. Composite pulses. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 18(2):61–122, 1986.

[62] M. H. Levitt. *Spin dynamics: basics of nuclear magnetic resonance*. Wiley, 2001.

[63] J.-S. Li, J. Ruths, T.-Y. Yu, H. Arthanari, and G. Wagner. Optimal pulse design in quantum control: A unified computational method. *Proceedings of the National Academy of Sciences*, 108(5):1879–1884, 2011.

[64] G. Lindblad. On the generators of quantum dynamical semigroups. *Communications in Mathematical Physics*, 48(2):119–130, 1976.

[65] K. L. Litvinenko, J. Li, N. Stavrias, A. J. Meaney, P. C. M. Christianen, H. Engelkamp, K. P. Homewood, C. R. Pidgeon, and B. N. Murdin. The quadratic Zeeman effect used for state-radius determination in neutral donors and donor bound excitons in Si:P. *Semiconductor Science and Technology*, 31(4):45007, 2016.

[66] C. C. Lo, S. Simmons, R. Lo Nardo, C. D. Weis, A. M. Tyryshkin, J. Meijer, D. Rogalla, S. A. Lyon, J. Bokor, T. Schenkel, and J. J. L. Morton. Stark shift and field ionization of arsenic donors in 28si-silicon-on-insulator structures. *Applied Physics Letters*, 104(19):193502, 2014.

[67] C. C. Lo, M. Urdampilleta, P. Ross, M. F. Gonzalez-Zalba, J. Mansir, S. A. Lyon, M. L. W. Thewalt, and J. J. L. Morton. Hybrid optical-electrical detection of donor electron spins with bound excitons in silicon. *Nature Materials*, 14(5):490–494, 2015.

[68] J. M. Luttinger and W. Kohn. Motion of electrons and holes in perturbed periodic fields. *Physical Review*, 97:869–883, 1955.

[69] G. G. Macfarlane, T. P. McLean, J. E. Quarrington, and V. Roberts. Fine structure in the absorption-edge spectrum of Si. *Physical Review*, 111(5):1245–1254, 1958.

[70] W. Magnus. On the exponential solution of differential equations for a linear operator. *Communications on Pure and Applied Mathematics*, 7(4):649–673, 1954.

[71] P. Mansfield. Symmetrized pulse sequences in high resolution NMR in solids. *Journal of Physics C: Solid State Physics*, 4(11):1444–1452, 1971.

[72] P. Mansfield and P. K. Grannell. Nmr 'diffraction' in solids? *Journal of Physics C: Solid State Physics*, 6(22):L422–L426, 1973.

[73] J. Mansir, P. Conti, Z. Zeng, J. J. Pla, P. Bertet, M. W. Swift, C. G. Van de Walle, M. L. W. Thewalt, B. Sklenard, Y. M. Niquet, and J. J. L. Morton. Linear Hyperfine Tuning of Donor Spins in Silicon Using Hydrostatic Strain. *Physical Review Letters*, 120(16):167701, 2018.

[74] S. Matsui. Suppressing the zero-frequency artifact in magic-sandwich-echo proton images of solids. *Journal of Magnetic Resonance (1969)*, 98(3):618–621, 1992.

[75] S. Matsui and H. Miura. 1H13C cross-polarization using a modified magic echo sequence for 1H spin locking. *Chemical Physics Letters*, 242(1-2):163–168, 1995.

[76] D. R. McCamey, C. Boehme, G. W. Morley, and J. van Tol. Electrically detected spin echoes of donor nuclei in silicon. *Physical Review B*, 85:073201, 2012.

[77] D. R. McCamey, J. Van Tol, G. W. Morley, and C. Boehme. Fast nuclear spin hyperpolarization of phosphorus in silicon. *Physical Review Letters*, 102(2), 2009.

[78] H. M. McConnell. Reaction rates by nuclear magnetic resonance. *The Journal of Chemical Physics*, 28(3):430–431, 1958.

[79] M. Mehring. A four-pulse NMR experiment using maximum pulse width. *Review of Scientific Instruments*, 44(1):64–66, 1973.

[80] G. W. Morley, D. R. McCamey, H. A. Seipel, L. C. Brunel, J. Van Tol, and C. Boehme. Long-lived spin coherence in silicon with an electrical spin trap readout. *Physical Review Letters*, 101(20):207602, 2008.

[81] J. J. L. Morton, A. M. Tyryshkin, R. M. Brown, S. Shankar, B. W. Lovett, A. Ardavan, T. Schenkel, E. E. Haller, J. W. Ager, and S. A. Lyon. Solid-state quantum memory using the ^{31}P nuclear spin. *Nature*, 455(7216):1085–1088, 2008.

[82] F. Motzoi, J. M. Gambetta, S. T. Merkel, and F. K. Wilhelm. Optimal control methods for rapidly time-varying hamiltonians. *Physical Review A*, 84:022307, 2011.

[83] I. Najfeld and T. F. Havel. Derivatives of the matrix exponential and their computation. *Advances in Applied Mathematics*, 16(3):321–375, 1995.

[84] M. Nakayama and H. Hasegawa. Relaxation effects on donor spins in silicon and germanium. *Journal of the Physical Society of Japan*, 18(2):229–239, 1963.

[85] J. M. Nichol, T. R. Naibert, E. R. Hemesath, L. J. Lauhon, and R. Budakian. Nanoscale fourier-transform magnetic resonance imaging. *Physical Review X*, 3(3):31016, 2013.

[86] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer, second edition, 2006.

[87] A. W. Overhauser. Polarization of nuclei in metals. *Physical Review*, 92(2):411–415, 1953.

[88] S. Pasini, P. Karbach, C. Raas, and G. S. Uhrig. Optimized pulses for the perturbative decoupling of a spin and a decoherence bath. *Physical Review A*, 80(2):22328, 2009.

[89] G. A. Paz-Silva and L. Viola. General transfer-function approach to noise filtering in open-loop quantum control. *Physical Review Letters*, 113(25):250501, 2014.

[90] G. Pica, G. Wolfowicz, M. Urdampilleta, M. L. W. Thewalt, H. Riemann, N. V. Abrosimov, P. Becker, H. J. Pohl, J. J. L. Morton, R. N. Bhatt, S. A. Lyon, and B. W. Lovett. Hyperfine Stark effect of shallow donors in silicon. *Physical Review B*, 90(19):195204, 2014.

[91] D. Pines, J. Bardeen, and C. P. Slichter. Nuclear polarization and impurity-state spin relaxation processes in silicon. *Physical Review*, 106(3):489–498, 1957.

[92] J. J. Pla, K. Y. Tan, J. P. Dehollain, W. H. Lim, J. J. L. Morton, D. N. Jamieson, A. S. Dzurak, and A. Morello. A single-atom electron spin qubit in silicon. *Nature*, 489(7417):541–544, 2012.

[93] C. P. Poole. *Electron Spin Resonance: A Comprehensive Treatise on Experimental Techniques*. Courier Corporation, 1996.

[94] I. I. Rabi, J. R. Zacharias, S. Millman, and P. Kusch. A new method of measuring nuclear magnetic moment. *Physical Review*, 53:318–318, 1938.

[95] W. K. Rhim, D. D. Elleman, and R. W. Vaughan. Analysis of multiple pulse NMR in solids. *The Journal of Chemical Physics*, 59(7):3740–3749, 1973.

[96] W. Rose, H. Haas, A. Q. Chen, N. Jeon, L. J. Lauhon, D. G. Cory, and R. Budakian. High-Resolution Nanoscale Solid-State Nuclear Magnetic Resonance Spectroscopy. *Physical Review X*, 8(1):11030, 2018.

[97] D. Rugar, R. Budakian, H. J. Mamin, and B. W. Chui. Single spin detection by magnetic resonance force microscopy. *Nature*, 430(6997):329–332, 2004.

[98] K. Saeedi, S. Simmons, J. Z. Salvail, P. Dluhy, H. Riemann, N. V. Abrosimov, P. Becker, H. J. Pohl, J. J. L. Morton, and M. L. W. Thewalt. Room-temperature quantum bit storage exceeding 39 minutes using ionized donors in silicon-28. *Science*, 342(6160):830–833, 2013.

[99] T. Sasaki, T. Oikawa, Y. Suzuki, K. Tagami, M. Shiraishi, and T. Suzuki. Electrical Spin Injection into Silicon Using MgO Tunnel Barrier. *Applied Physics Express*, 2(5):053003, 2009.

[100] W. Schmid. Auger lifetimes for excitons bound to neutral donors and acceptors in Si. *Physica Status Solidi (B)*, 84(2):529–540, 1977.

[101] T. Sekiguchi, M. Steger, K. Saeedi, M. L. W. Thewalt, H. Riemann, N. V. Abrosimov, and N. Nötzel. Hyperfine structure and nuclear hyperpolarization observed in the bound exciton luminescence of Bi donors in natural Si. *Physical Review Letters*, 104(13):137402, 2010.

[102] A. J. Sigillito, A. M. Tyryshkin, and S. A. Lyon. Anisotropic stark effect and electric-field noise suppression for phosphorus donor qubits in silicon. *Physical Review Letters*, 114(21):217601, 2015.

[103] A. J. Sigillito, A. M. Tyryshkin, T. Schenkel, A. A. Houck, and S. A. Lyon. All-electric control of donor nuclear spin qubits in silicon. *Nature Nanotechnology*, 12(10):958–962, 2017.

[104] A. Soare, H. Ball, D. Hayes, J. Sastrawan, M. C. Jarratt, J. J. Mcloughlin, X. Zhen, T. J. Green, and M. J. Biercuk. Experimental noise filtering by quantum control. *Nature Physics*, 10(11):825–829, 2014.

[105] P. E. Spindler, Y. Zhang, B. Endeward, N. Gershernzon, T. E. Skinner, S. J. Glaser, and T. F. Prisner. Shaped optimal control pulses for increased excitation bandwidth in epr. *Journal of Magnetic Resonance*, 218:49 – 58, 2012.

[106] M. Steger, K. Saeedi, M. L. W. Thewalt, J. J. L. Morton, H. Riemann, N. V. Abrosimov, P. Becker, and H.-J. Pohl. Quantum information storage for over 180 s using donor spins in a 28si "semiconductor vacuum". *Science*, 336(6086):1280–1283, 2012.

[107] M. Steger, T. Sekiguchi, A. Yang, K. Saeedi, M. E. Hayden, M. L. W. Thewalt, K. M. Itoh, H. Riemann, N. V. Abrosimov, P. Becker, and H. J. Pohl. Optically-detected NMR of optically-hyperpolarized 31P neutral donors in 28Si. *Journal of Applied Physics*, 109(10):102411, 2011.

[108] N. J. Stone. Table of nuclear magnetic dipole and electric quadrupole moments. *Atomic Data and Nuclear Data Tables*, 90(1):75–176, 2005.

[109] K. Takegoshi and C. A. McDowell. A "magic echo" pulse sequence for the high-resolution NMR spectra of abundant spins in solids. *Chemical Physics Letters*, 116(2-3):100–104, 1985.

[110] A. J. Thakkar. *Molecular symmetry*. Morgan & Claypool Publishers, 2017.

[111] M. L. W. Thewalt. Details of the structure of bound excitons and bound multiexciton complexes in Si. *Canadian Journal of Physics*, 55(17):310–310, 1463-1480.

[112] D. D. Thornton and A. Honig. Shallow-donor negative ions and spin-polarized electron transport in silicon. *Physical Review Letters*, 30(19):909–912, 1973.

[113] G. Tosi, F. A. Mohiyaddin, V. Schmitt, S. Tenberg, R. Rahman, G. Klimeck, and A. Morello. Silicon quantum processor with robust long-distance qubit couplings. *Nature Communications*, 8(1):450, 2017.

[114] V. V. Tsyplenkov, K. A. Kovalevsky, and V. N. Shastin. Effect of uniaxial stress on intervalley phonon-assisted relaxation of excited shallow-donor states in silicon. *Semiconductors*, 43(11):1410, 2009.

[115] M. Usman, C. D. Hill, R. Rahman, G. Klimeck, M. Y. Simmons, S. Rogge, and L. C. L. Hollenberg. Strain and electric field control of hyperfine interactions for donor spin qubits in silicon. *Physical Review B*, 91(24):245209, 2015.

[116] C. F. Van Loan. Computing Integrals Involving the Matrix Exponential. *IEEE Transactions on Automatic Control*, 23(3):395–404, 1978.

[117] J. Vandermause and C. Ramanathan. Superadiabatic control of quantum operations. *Physical Review A*, 93:052329, 2016.

[118] A. S. Verhulst, I. G. Rau, Y. Yamamoto, and K. M. Itoh. Optical pumping of ^{29}Si nuclear spins in bulk silicon at high magnetic field and liquid helium temperature. *Physical Review B*, 71(23):235206, 2005.

[119] N. Q. Vinh, P. T. Greenland, K. Litvinenko, B. Redlich, A. F. G. van der Meer, S. A. Lynch, M. Warner, A. M. Stoneham, G. Aeppli, D. J. Paul, C. R. Pidgeon, and B. N. Murdin. Silicon as a model ion trap: Time domain measurements of donor Rydberg states. *Proceedings of the National Academy of Sciences*, 105(31):10649–10653, 2008.

[120] J. Watrous. *The Theory of Quantum Information*. Cambridge University Press, 2018.

[121] J. S. Waugh, L. M. Huber, and U. Haeberlen. Approach to high-resolution NMR in solids. *Physical Review Letters*, 20(5):180–182, 1968.

[122] C. J. Wellard and L. C. L. Hollenberg. Donor electron wave functions for phosphorus in silicon: Beyond effective-mass theory. *Physical Review B*, 72:085202, 2005.

[123] D. K. Wilson and G. Feher. Electron spin resonance experiments on donors in silicon. III. investigation of excited states by the application of uniaxial stress and their importance in relaxation processes. *Physical Review*, 124(4):1068–1083, 1961.

[124] Wolfram Research. Mathematica, Version 10.2.

[125] A. Yang, M. Steger, T. Sekiguchi, M. L W Thewalt, J. W. Ager, and E. E. Haller. Homogeneous linewidth of the 31P bound exciton transition in silicon. *Applied Physics Letters*, 95(12):122113, 2009.

[126] A. Yang, M. Steger, T. Sekiguchi, M. L. W. Thewalt, T. D. Ladd, K. M. Itoh, H. Riemann, N. V. Abrosimov, P. Becker, and H.-J. Pohl. Simultaneous Subsecond Hyperpolarization of the Nuclear and Electron Spins of Phosphorus in Silicon by Optical Pumping of Exciton Transitions. *Physical Review Letters*, 102(25):257401, 2009.

[127] P. Y. Yu and M. Cardona. *Fundamentals of Semiconductors*. Springer Berlin Heidelberg, 2010.

Appendix A

Proof of Theorem 1[1]

In this appendix, we give a proof for Theorem 1 given in Section 3.2 of the main text. The theorem states that, for $B_{i,j} : [0,T] \to M_n$ and for $1 \leq i \leq j \leq m$, we define $C_{i,j} : [0,T] \to M_n$ implicitly through the following expression:

$$\begin{pmatrix} C_{1,1}(t) & C_{1,2}(t) & \dots & C_{1,m}(t) \\ 0 & C_{2,2}(t) & \dots & C_{2,m}(t) \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & C_{m,m}(t) \end{pmatrix} = \tag{A.1}$$
$$\mathcal{T} \exp \left[\int_0^t dt_1 \begin{pmatrix} B_{1,1}(t_1) & B_{1,2}(t_1) & \dots & B_{1,m}(t_1) \\ 0 & B_{2,2}(t_1) & \dots & B_{2,m}(t_1) \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & B_{m,m}(t_1) \end{pmatrix} \right].$$

[1]The theorem and its proof appearing in this appendix was shown to the author by Daniel Puzzuoli.

Therefore, by differentiating both sides of Equation (A.1), and using the definition of the time-ordered exponential we arrive at

$$\begin{pmatrix} \dot{C}_{1,1}(t) & \dot{C}_{1,2}(t) & \dots & \dot{C}_{1,m}(t) \\ 0 & \dot{C}_{2,2}(t) & \dots & \dot{C}_{2,m}(t) \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & \dot{C}_{m,m}(t) \end{pmatrix} = \tag{A.2}$$
$$\begin{pmatrix} B_{1,1}(t) & B_{1,2}(t) & \dots & B_{1,m}(t) \\ 0 & B_{2,2}(t) & \dots & B_{2,m}(t) \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & B_{m,m}(t) \end{pmatrix} \begin{pmatrix} C_{1,1}(t) & C_{1,2}(t) & \dots & C_{1,m}(t) \\ 0 & C_{2,2}(t) & \dots & C_{2,m}(t) \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & C_{m,m}(t) \end{pmatrix}.$$

The above implies that $C_{k,k+j}$, for $1 \leq k \leq m$ and $0 \leq j \leq m-k$, satisfy the following differential equations:

$$\dot{C}_{k,k+j}(t) = \sum_{i=0}^{j} B_{k,k+i}(t) C_{k+i,k+j}(t), \tag{A.3}$$

with initial conditions

$$C_{k,k+j}(0) = \begin{cases} \mathbb{1}_n & \text{if } j = 0, \\ 0 & \text{else.} \end{cases} \tag{A.4}$$

From Equations (A.3) and (A.4) it is immediately clear that

$$C_{k,k}(t) = U_k(t) = \mathcal{T} \exp\left(\int_0^t dt_1 B_{i,i}(t_1) \right). \tag{A.5}$$

To show that the recursion relations

$$C_{k,k+j}(t) = \sum_{i=1}^{j} U_k(t) \int_0^t dt_1 U_k^{-1}(t_1) B_{k,k+i}(t_1) C_{k+i,k+j}(t_1) \tag{A.6}$$

hold for any $1 < j \leq m-k$, we first note that $C_{k,k+j}(0) = 0$, for any $1 < j \leq m-k$, hence, satisfying the initial conditions in Equation (A.4). In order to complete the proof, we differentiate Equation (A.6):

$$\begin{aligned} \dot{C}_{k,k+j}(t) &= \sum_{i=1}^{j} B_{k,k+i}(t) C_{k+i,k+j}(t) + B_{k,k}(t) \sum_{i=1}^{j} U_k(t) \int_0^t dt_1 U_k^{-1}(t_1) B_{k,k+i}(t_1) C_{k+i,k+j}(t_1) \\ &= \sum_{i=1}^{j} B_{k,k+i}(t) C_{k+i,k+j}(t) + B_{k,k}(t) C_{k,k+j}(t) = \sum_{i=0}^{j} B_{k,k+i}(t) C_{k+i,k+j}(t), \end{aligned} \tag{A.7}$$

which coincides with Equation (A.3) implying that solutions obtained through Equation (A.6) satisfy the system of differential equations.

Appendix B

Magnetic Field Values Above the CFFGS Device[1]

Magnetic fields and field gradients generated by the CFFGS, that were used for analysing the data in the main text of this thesis, were simulated using COMSOL Multiphysics finite element software. The simulations used CFFGS dimensions that were taken from the scanning electron micrograph image of the device shown in Figure 4.1(b). The B_1 and dB_1/dz values were calculated for 50 mA peak current, whereas the dB_z/dx values were calculated for 70 mA peak current through the device. We present the simulation results below in Figure B.1.

[1]The COMSOL simulations presented in this appendix were carried out by Raffi Budakian and William Rose.

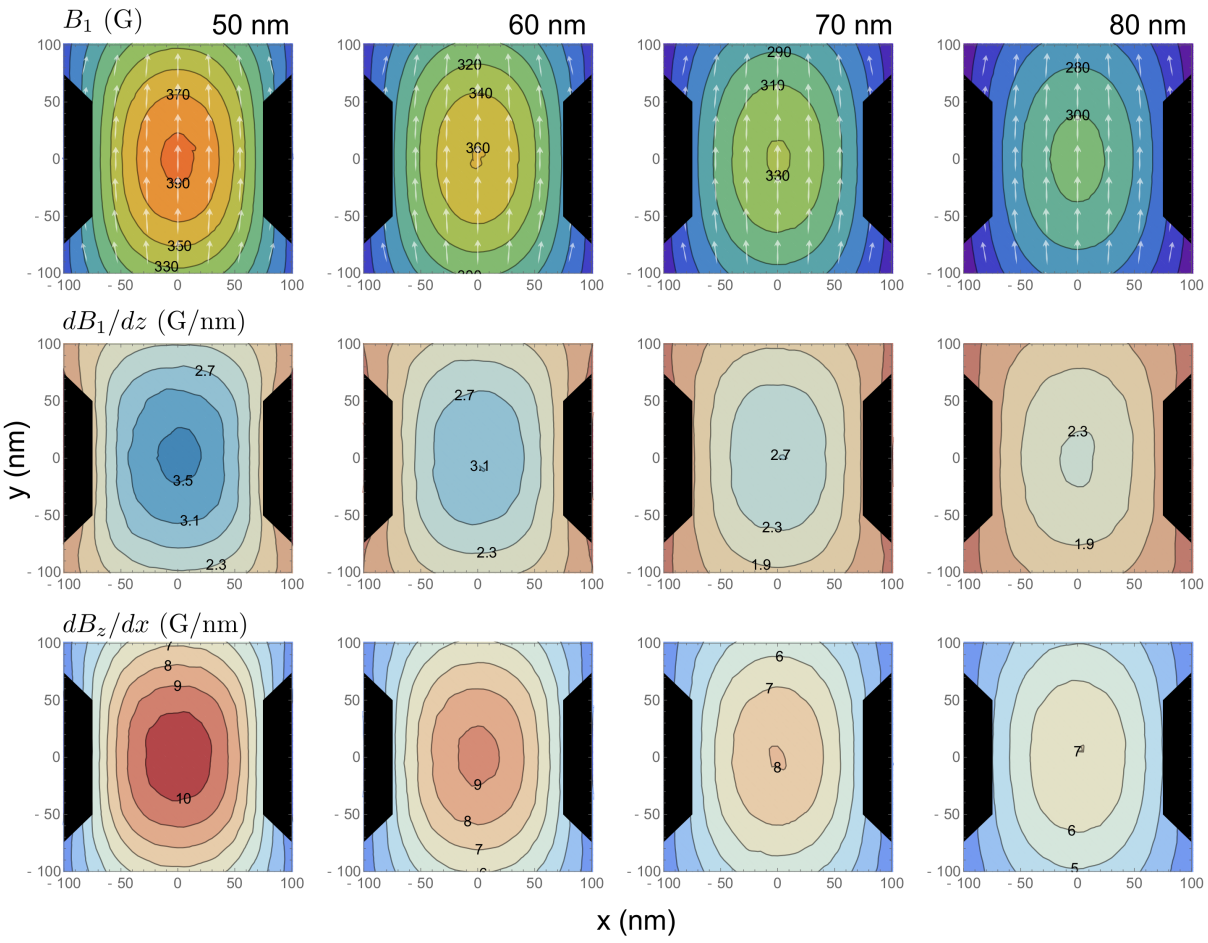

Figure B.1: Magnetic field and magnetic field gradients in the xy plane at constant heights z around the centre of the CFFGS constriction as indicated in Figure 4.1(a). The four heights picked for finite element simulations were $z \in \{50\text{ nm}, 60\text{ nm}, 70\text{ nm}, 80\text{ nm}\}$. In the top row, we display the Rabi fields $B_1(\vec{r}) = \sqrt{B_x^2(\vec{r}) + B_y^2(\vec{r})}/2$ at various locations in the horizontal plane of a fixed z-coordinate, as well as a vector plot that indicates the relative magnitudes of $B_x(\vec{r})$ and $B_y(\vec{r})$ at these locations. In the middle row, we give dB_1/dz values, and in the bottom row, we give dB_z/dx values at the four heights above the CFFGS top surface. The black regions in the contour plots denote the outside contours of the CFFGS near its centre.

Appendix C

Bloch Sphere Trajectories[1]

In order to visually illustrate how the numerically engineered control sequences achieve the same final unitary operation for a number of different Rabi strength parameters $\gamma \in \Gamma$ simultaneously, we simulate the Bloch sphere trajectories for spins starting in the $|\uparrow\rangle$ state for three different γ values: $\gamma = \{0.7 \text{ MHz}, 0.9 \text{ MHz}, 1.1 \text{ MHz}\}$ for pulse #1 and pulse #3, that were introduced in Section 4.2.1 and 4.2.3 of the main text. We present the trajectories by plotting Bloch vector projections onto x, y and z-axis as a function of time t over the pulse duration $0 \leq t \leq T$. The Bloch vector $\vec{m}^{(\gamma)}(t) = \left(m_x^{(\gamma)}(t), m_y^{(\gamma)}(t), m_z^{(\gamma)}(t)\right)$ has its elements given as

$$m_i^{(\gamma)}(t) = \frac{1}{2}\text{Tr}\left[\sigma_i \left[U^{(\gamma)}(t)\right]^\dagger \sigma_z U^{(\gamma)}(t)\right], \tag{C.1}$$

for $i \in \{x, y, z\}$. It is shown in Figures C.1 and C.2 that $\vec{m}^{(\gamma)}(t)$ traces out a unique trajectory for each γ value, nevertheless, the initial $\vec{m}^{(\gamma)}(t = 0) = (0, 0, 1)$ and the final $\vec{m}^{(\gamma)}(t = T) = (0, 1, 0)$ Bloch vectors coincide for all three γ values pictured. From the number of oscillations in these trajectories it is evident that the spins experiencing higher Rabi strengths end up completing a greater number of laps around the Bloch sphere. Although, we do not present plots for the Bloch vector trajectories for spins starting in the xy plane they would possess all the same characteristics.

By decomposing $\left[U^{(\gamma)}(T)\right]^\dagger \mathcal{D}_{U^{(\gamma)}}(\sigma_z) = \int_0^T dt \left[U^{(\gamma)}(t)\right]^\dagger \sigma_z U^{(\gamma)}(t)$ into a sum of traceless operator basis elements: $\{\sigma_i\}$, for $i \in \{x, y, z\}$, it is easy to show that $\Psi_{\sigma_z}^{(\gamma)}$ defined in

[1]The simulations presented in this appendix were conducted by the author and are published as part of Supplementary Materials for [96].

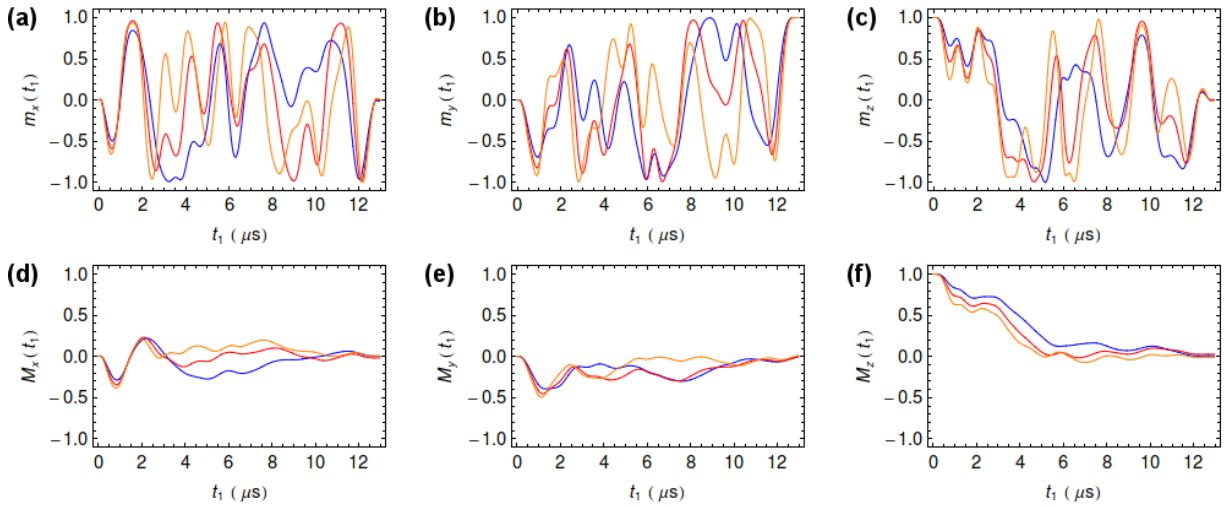

Figure C.1: Bloch sphere trajectories $\{m_i^{(\gamma)}(t)\}$, defined by Equation (C.1), and their integrals $\{M_i^{(\gamma)}(t)\}$ defined by Equation (C.3) for spins evolving under pulse #1 with three distinct Rabi strengths: $\gamma = 0.7$ MHz (blue), $\gamma = 0.9$ MHz (red), $\gamma = 1.1$ MHz (orange). (a) $m_x^{(\gamma)}(t)$, (b) $m_y^{(\gamma)}(t)$, (c) $m_z^{(\gamma)}(t)$, (d) $M_x^{(\gamma)}(t)$, (e) $M_y^{(\gamma)}(t)$, (f) $M_z^{(\gamma)}(t)$ for $0 \leq t \leq T$, with $T = 12.95$ μs.

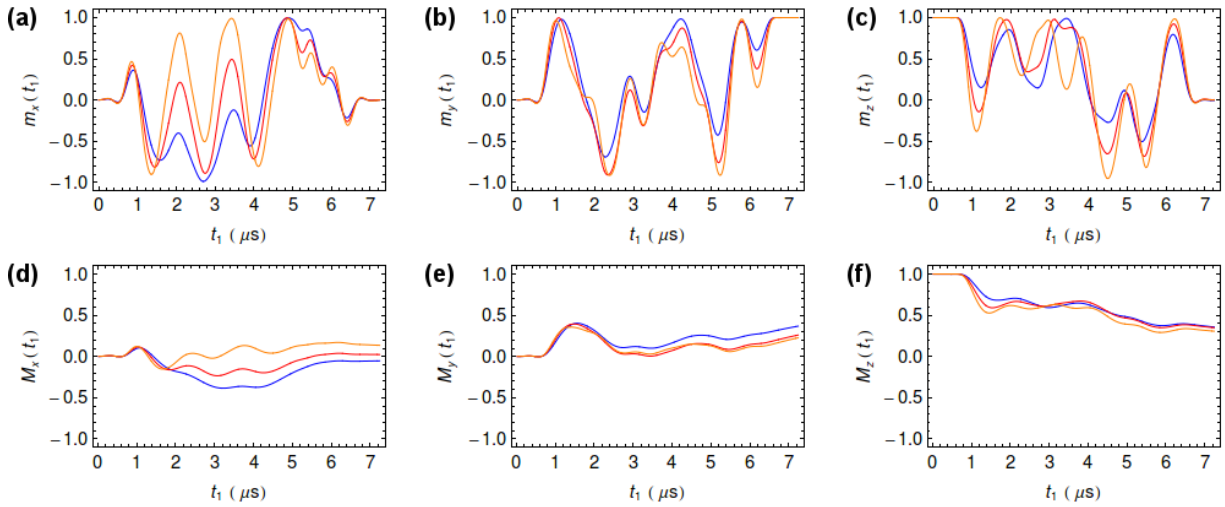

Figure C.2: Bloch sphere trajectories $\{m_i^{(\gamma)}(t)\}$, defined by Equation (C.1), and their integrals $\{M_i^{(\gamma)}(t)\}$ defined by Equation (C.3) for spins evolving under pulse #3 with three distinct Rabi strengths: $\gamma = 0.7$ MHz (blue), $\gamma = 0.9$ MHz (red), $\gamma = 1.1$ MHz (orange). (a) $m_x^{(\gamma)}(t)$, (b) $m_y^{(\gamma)}(t)$, (c) $m_z^{(\gamma)}(t)$, (d) $M_x^{(\gamma)}(t)$, (e) $M_y^{(\gamma)}(t)$, (f) $M_z^{(\gamma)}(t)$ for $0 \leq t \leq T$, with $T = 7.24$ μs.

Equation (4.28) can be written as

$$\Psi_{\sigma_z}^{(\gamma)} = \sqrt{\sum_i \left(\frac{1}{2T} \mathrm{Tr} \left[\sigma_i \left[U^{(\gamma)}(T) \right]^\dagger \mathcal{D}_{U^{(\gamma)}}(\sigma_z) \right] \right)^2} = \sqrt{\sum_i \left(\frac{1}{T} \int_0^T dt \; m_i^{(\gamma)}(t) \right)^2}. \quad \text{(C.2)}$$

Therefore, the σ_z averaging property of the pulses can conveniently be demonstrated by integrating the trajectories of the individual components of the Bloch vector $\vec{m}^{(\gamma)}$. We define

$$M_i^{(\gamma)}(t) = \frac{1}{t} \int_0^t dt_1 \; m_i^{(\gamma)}(t_1) = \frac{1}{2t} \int_0^t dt_1 \; \mathrm{Tr} \left[\sigma_i \left[U^{(\gamma)}(t_1) \right]^\dagger \sigma_z U^{(\gamma)}(t_1) \right] \quad \text{(C.3)}$$

as a normalized integral of $m_i^{(\gamma)}(t)$, and note that $\Psi_{\sigma_z}^{(\gamma)} = 0$ if and only if $M_i^{(\gamma)}(T) = 0$, for all $i \in \{x, y, z\}$. We evaluate these integrals for the three γ values mentioned, and present them in Figure C.1 and Figure C.2 for pulse #1 and pulse #3, respectively. It can be seen that the integrated trajectories in Figure C.1 are again unique for each Rabi strength, yet they all yield $M_i^{(\gamma)}(T) = 0$, for all $i \in \{x, y, z\}$. On the other hand, pulse #3 results non-zero $M_i^{(\gamma)}(T)$ values, as is shown in Figure C.2, since it was optimized without the σ_z Dyson term target. Lastly, we remark that the dipolar integral $\left[U^{(\gamma)}(T) \otimes U^{(\gamma)}(T) \right]^\dagger \mathcal{D}_{U^{(\gamma)} \otimes U^{(\gamma)}}(D)$ could have been decomposed and visualized entirely analogously to $\left[U^{(\gamma)}(T) \right]^\dagger \mathcal{D}_{U^{(\gamma)}}(\sigma_z)$. Nevertheless, the six operator basis elements required would have made the presentation rather cumbersome.

Appendix D

Nuclear spin T_2 due to electron spin undergoing T_1 relaxation[1]

The spin Hamiltonian that was given in Equation (6.1), and dictates the dynamics of our electron-nuclear spin system can, for high B_0-values and negligible strain fields, be approximated as

$$H = \frac{\omega_e}{2}\sigma_z \otimes \mathbb{1} - \frac{\omega_{^{31}\mathrm{P}}}{2}\mathbb{1} \otimes \sigma_z + \frac{A_e}{4}\sigma_z \otimes \sigma_z, \tag{D.1}$$

where $\omega_e = \gamma_e B_0$, $\omega_{^{31}\mathrm{P}} = \gamma_n B_0$. Here, we consider a situation where the nuclear spin thermal relaxation time is much longer than the duration of the experiments carried out on the spin system. In such case, the combined electron-nuclear density matrix ρ evolves under Lindblad equation introduced by Equation (1.16):

$$\frac{d}{dt}\rho(t) = -i[H, \rho(t)] + (L_e \otimes \mathbb{1})[\rho(t)], \tag{D.2}$$

$L_e[\rho_e]$ corresponding to the action of T_1 relaxation on the electron spin only:

$$\begin{aligned} L_e[\rho_e] = &-\frac{1}{2T_1}\ (1-s)\ \left[\sigma_+\sigma_-\rho_e + \rho_e\sigma_+\sigma_- - 2\sigma_-\rho_e\sigma_+\right] \\ &-\frac{1}{2T_1}\ s\ \left[\sigma_-\sigma_+\rho_e + \rho_e\sigma_-\sigma_+ - 2\sigma_+\rho_e\sigma_-\right]. \end{aligned} \tag{D.3}$$

The equilibrium electron spin density matrix of $L_e[\rho_e]$ is $\rho_{eq} = s\ \ket{\uparrow}\bra{\uparrow} + (1-s)\ \ket{\downarrow}\bra{\downarrow}$.

[1] The treatment presented in this appendix was conducted by the author and is published as part of Supplementary Materials for [38].

As we said in Chapter 5, we performed our experiments in the electron spin state $|\downarrow\rangle$ manifold due to the negligible s values at our magnetic fields and temperatures. Consequently, we go into an interaction frame of nuclear spin Hamiltonian $-\left(\frac{\omega_{31\mathrm{P}}}{2}+\frac{A_e}{4}\right)\ \mathbb{1}\otimes\sigma_z$ and arrive at an effective Hamiltonian

$$\begin{aligned}\widetilde{H} &= \frac{\omega_e}{2}\ \sigma_z\otimes\mathbb{1}+\frac{A_e}{4}\ \left(\mathbb{1}\otimes\sigma_z+\sigma_z\otimes\sigma_z\right)\\ &= \frac{\omega_e}{2}\ \sigma_z\otimes\mathbb{1}+\frac{A_e}{2}\ |\uparrow\rangle\langle\uparrow|\otimes\sigma_z.\end{aligned} \tag{D.4}$$

It can be seen that in this interaction frame we can interpret the nuclear spin Hamiltonian as a Hamiltonian conditional on the electron spin state: for electron spin state $|\downarrow\rangle$, the Hamiltonian is equal to zero, and for electron spin state $|\uparrow\rangle$ the Hamiltonian is $\frac{A_e}{2}\sigma_z$. Lindblad operators in Equation (D.3) commute with the interaction frame, hence, the electron-nuclear density matrix in the frame $\widetilde{\rho}$ evolves according to

$$\frac{d}{dt}\widetilde{\rho}(t) = -i[\widetilde{H},\widetilde{\rho}(t)]+\left(L_e\otimes\mathbb{1}\right)[\widetilde{\rho}(t)]\,. \tag{D.5}$$

Evolution under Equation (D.5) cannot create coherences between electron spin states $|\downarrow\rangle$ and $|\uparrow\rangle$. If we assume that the initial electron-nuclear density matrix has no electron coherences, we only need to solve dynamics for density matrices of the following form: $\widetilde{\rho}=p\,|\uparrow\rangle\langle\uparrow|\otimes\widetilde{\rho}_\uparrow+(1-p)\,|\downarrow\rangle\langle\downarrow|\otimes\widetilde{\rho}_\downarrow$, where $0\le p\le 1$, and $\widetilde{\rho}_\downarrow$ and $\widetilde{\rho}_\uparrow$ are conditional nuclear spin density matrices for electron spin states $|\downarrow\rangle$ and $|\uparrow\rangle$, respectively. We solve the dynamics by using a particular vectorization of such density matrices. First, we define a vector $\vec{r}(t)\in\mathbb{R}^{2\times 4}$, $\vec{r}(t)=(\vec{r}_\downarrow(t),\vec{r}_\uparrow(t))$, with $\vec{r}_{\downarrow(\uparrow)}=(r_{I_{\downarrow(\uparrow)}},r_{x_{\downarrow(\uparrow)}},r_{y_{\downarrow(\uparrow)}},r_{z_{\downarrow(\uparrow)}})$ such that $\widetilde{\rho}(t)=\frac{1}{\sqrt{2}}\,|\downarrow\rangle\langle\downarrow|\otimes\vec{r}_\downarrow(t)\cdot(\mathbb{1},\sigma_x,\sigma_y,\sigma_z)+\frac{1}{\sqrt{2}}\,|\uparrow\rangle\langle\uparrow|\otimes\vec{r}_\uparrow(t)\cdot(\mathbb{1},\sigma_x,\sigma_y,\sigma_z)$. We now rewrite Equation (D.5) in this basis:

$$\frac{d}{dt}\vec{r}(t)=\begin{pmatrix}
\frac{-s}{T_1} & 0 & 0 & 0 & \frac{1-s}{T_1} & 0 & 0 & 0\\
0 & \frac{-s}{T_1} & 0 & 0 & 0 & \frac{1-s}{T_1} & 0 & 0\\
0 & 0 & \frac{-s}{T_1} & 0 & 0 & 0 & \frac{1-s}{T_1} & 0\\
0 & 0 & 0 & \frac{-s}{T_1} & 0 & 0 & 0 & \frac{1-s}{T_1}\\
\frac{s}{T_1} & 0 & 0 & 0 & \frac{s-1}{T_1} & 0 & 0 & 0\\
0 & \frac{s}{T_1} & 0 & 0 & 0 & \frac{s-1}{T_1} & -A_e & 0\\
0 & 0 & \frac{s}{T_1} & 0 & 0 & A_e & \frac{s-1}{T_1} & 0\\
0 & 0 & 0 & \frac{s}{T_1} & 0 & 0 & 0 & \frac{s-1}{T_1}
\end{pmatrix}\cdot\vec{r}(t). \tag{D.6}$$

The equation above can be easily solved in either of the regimes: $T_1A_e\gg 1$ and $T_1A_e\ll 1$, in the former case, we can approximate $\vec{r}(t)\approx S_1(t)\cdot\vec{r}(0)$, whereas in the latter, we can

say $\vec{r}(t) \approx S_2(t) \cdot \vec{r}(0)$. The respective propagators – S_1 and S_2 – are given as

$$S_1(t) = \exp\left[\begin{pmatrix} \frac{-s}{T_1} & 0 & 0 & 0 & \frac{1-s}{T_1} & 0 & 0 & 0 \\ 0 & \frac{-s}{T_1} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{-s}{T_1} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{-s}{T_1} & 0 & 0 & 0 & \frac{1-s}{T_1} \\ \frac{s}{T_1} & 0 & 0 & 0 & \frac{s-1}{T_1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{s-1}{T_1} & -A_e & 0 \\ 0 & 0 & 0 & 0 & 0 & A_e & \frac{s-1}{T_1} & 0 \\ 0 & 0 & 0 & \frac{s}{T_1} & 0 & 0 & 0 & \frac{s-1}{T_1} \end{pmatrix} t\right] \tag{D.7}$$

and

$$S_2(t) = \exp\left[\begin{pmatrix} \frac{-s}{T_1} & 0 & 0 & 0 & \frac{1-s}{T_1} & 0 & 0 & 0 \\ 0 & \frac{-s}{T_1} & 0 & 0 & 0 & \frac{1-s}{T_1} & 0 & 0 \\ 0 & 0 & \frac{-s}{T_1} & 0 & 0 & 0 & \frac{1-s}{T_1} & 0 \\ 0 & 0 & 0 & \frac{-s}{T_1} & 0 & 0 & 0 & \frac{1-s}{T_1} \\ \frac{s}{T_1} & 0 & 0 & 0 & \frac{s-1}{T_1} & 0 & 0 & 0 \\ 0 & \frac{1-s}{T_1} & 0 & 0 & 0 & \frac{s-1}{T_1} & 0 & 0 \\ 0 & 0 & \frac{1-s}{T_1} & 0 & 0 & 0 & \frac{s-1}{T_1} & 0 \\ 0 & 0 & 0 & \frac{s}{T_1} & 0 & 0 & 0 & \frac{s-1}{T_1} \end{pmatrix} t\right]. \tag{D.8}$$

Hence, an initial electron-nuclear density matrix $\widetilde{\rho}_i(0) = |\downarrow\rangle\langle\downarrow| \otimes \left(\frac{1}{2}\mathbb{1} + r_x\sigma_x + r_y\sigma_y + r_z\sigma_z\right)$, $i \in \{1, 2\}$, evolves like

$$\begin{aligned} \widetilde{\rho}_1(t) = {} & |\downarrow\rangle\langle\downarrow| \otimes \left(e^{-\frac{st}{T_1}} r_x\sigma_x + e^{-\frac{st}{T_1}} r_y\sigma_y\right) \\ & + \left(\left[1 - s\left(1 - e^{-\frac{t}{T_1}}\right)\right]|\downarrow\rangle\langle\downarrow| + s\left(1 - e^{-\frac{t}{T_1}}\right)|\uparrow\rangle\langle\uparrow|\right) \otimes \left(\frac{1}{2}\mathbb{1} + r_z\sigma_z\right) \end{aligned} \tag{D.9}$$

for $T_1 A_e \gg 1$, whereas for $T_1 A_e \ll 1$ it evolves as

$$\begin{aligned} \widetilde{\rho}_2(t) = {} & \\ & \left(\left[1 - s\left(1 - e^{-\frac{t}{T_1}}\right)\right]|\downarrow\rangle\langle\downarrow| + s\left(1 - e^{-\frac{t}{T_1}}\right)|\uparrow\rangle\langle\uparrow|\right) \otimes \left(\frac{1}{2}\mathbb{1} + r_x\sigma_x + r_y\sigma_y + r_z\sigma_z\right). \end{aligned} \tag{D.10}$$

Tracing out the electron state for $\widetilde{\rho}_1(t)$ and $\widetilde{\rho}_2(t)$ above reveals that for $T_1 A_e \gg 1$, the electron spin T_1 process yields an effective nuclear spin T_2 process with time constant $T_2 = \frac{T_1}{s}$, and for $T_1 A_e \ll 1$ there is no effective decoherence of the nucleus due to the

electron spin. Finally, if we assume the equilibrium electron density matrix ρ_{eq} to be a thermal state, at inverse temperature Θ, then $\rho_{eq} = \frac{1}{\mathrm{Tr}\left[\exp\left(-\frac{\omega_e \hbar \Theta}{k_B}\frac{\sigma_z}{2}\right)\right]} \exp\left(-\frac{\omega_e \hbar \Theta}{k_B}\frac{\sigma_z}{2}\right)$ and $s = 1/\left(1 + e^{\frac{\omega_e \hbar \Theta}{k_B}}\right)$.

Appendix E

Symmetry Analysis

Since the ground state of the donor electron wave function, that was introduced in Section 5.1.1 of the main text, is non-degenerate, it has the full tetrahedral point symmetry of a cubic diamond lattice site in silicon crystals. A group that corresponds to the symmetry is called the tetrahedral point group, or T_d group, which is a finite group of order 24 [110]. In this appendix, we construct a number of T_d group representations, and explicitly demonstrate the group properties that are relevant for the arguments in Chapter 5 and Chapter 6. For a silicon lattice its with crystallographic axes given as $(1,0,0)$, $(0,1,0)$ and $(0,0,1)$, the elements $O_i \in \mathrm{M}_3(\mathbb{R})$, $i \in \{1,...,24\}$, of a three-dimensional representation for the T_d group can be divided into the following five conjugacy classes:

- identity,
- eight rotations by $2\pi/3$ around: $(1,1,1)$, $(-1,-1,-1)$, $(-1,1,1)$, $(1,-1,-1)$, $(1,-1,1)$, $(-1,1,-1)$, $(-1,-1,1)$ and $(1,1,-1)$,
- three rotations by π around: $(1,0,0)$, $(0,1,0)$ and $(0,0,1)$,
- six reflections in a mirror normal to $(1,1,0)$, $(1,-1,0)$, $(1,0,1)$, $(1,0,-1)$, $(0,1,1)$ and $(0,1,-1)$,
- six rotoreflections by $\pi/2$ around $(-1,0,0)$, $(1,0,0)$, $(0,1,0)$, $(0,-1,0)$, $(0,0,-1)$ and $(0,0,1)$.

The explicit matrix representations of the above are easily evaluated using Mathematica's functions: IdentityMatrix, RotationMatrix and ReflectionMatrix, and have the following

form

$$O^{(1)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},\ O^{(2)} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},\ O^{(3)} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \tag{E.1}$$

$$O^{(4)} = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & 1 \\ -1 & 0 & 0 \end{pmatrix},\ O^{(5)} = \begin{pmatrix} 0 & 0 & -1 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},\ O^{(6)} = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \end{pmatrix},$$

$$O^{(7)} = \begin{pmatrix} 0 & 0 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \end{pmatrix},\ O^{(8)} = \begin{pmatrix} 0 & 0 & -1 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \end{pmatrix},\ O^{(9)} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & -1 \\ -1 & 0 & 0 \end{pmatrix},$$

$$O^{(10)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix},\ O^{(11)} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix},\ O^{(12)} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$O^{(13)} = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},\ O^{(14)} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},\ O^{(15)} = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix},$$

$$O^{(16)} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix},\ O^{(17)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix},\ O^{(18)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$

$$O^{(19)} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix},\ O^{(20)} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix},\ O^{(21)} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ -1 & 0 & 0 \end{pmatrix}$$

$$O^{(22)} = \begin{pmatrix} 0 & 0 & -1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix},\ O^{(23)} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix},\ O^{(24)} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Let $\{o^{(i)}\}$ be a matrix representation of the T_d group, such that $\left(o^{(i)}\right)^\dagger \in \{o^{(k)}\}$, for all i, and $o^{(i)}o^{(j)} \in \{o^{(k)}\}$, for all i and j. Throughout this appendix, we are generally concerned with finding eigenvalue-one mutual eigenvectors for every element of a particular T_d group matrix representation, i.e., finding a vector $\vec{v}$ such that $o^{(i)} \cdot \vec{v} = \vec{v}$, for all i. These vectors represent particular invariants under the tetrahedral point symmetry. It easy to see that $P = \frac{1}{24}\sum_{i=1}^{24} o^{(i)}$ is a projector since $P^\dagger = P$ and $P^2 = \frac{1}{24^2}\sum_{i,j=1}^{24} o^{(i)}o^{(j)} = \frac{1}{24}\sum_{k=1}^{24} o^{(k)} = P$. We now show that a vector $\vec{v}$ satisfies $o^{(i)} \cdot \vec{v} = \vec{v}$, for all i, if and only if $P \cdot \vec{v} = \vec{v}$. It is trivial to show if $\vec{v}$ is a mutual eigenvector of all $o^{(i)}$ then $P \cdot \vec{v} = \frac{1}{24}\sum_{k=1}^{24} o^{(i)} \cdot \vec{v} = \vec{v}$. For the other direction, one needs to note that if $\vec{v} = P \cdot \vec{v}$ then $o^{(i)}\vec{v} = o^{(i)}P \cdot \vec{v} = P \cdot \vec{v} = \vec{v}$,

for all i. Hence, the problem of finding mutual eigenvalue-one eigenvectors for all $o^{(i)}$ reduces to finding the eigenvectors of P. As the simplest example, we could consider $P = \frac{1}{24}\sum_{i=1}^{24} O^{(i)} = 0$, where $O^{(i)}$ are the matrices given in Equation (E.1), therefore, the set $\{O^{(i)}\}$ has no mutual eigenvalue-one eigenvectors.

In Chapter 6, we used a particular vectorization of the strain tensor ϵ at the site of the phosphorus defect, evaluated in the crystallographic coordinate system. Because ϵ is a real symmetric tensor, we only need a 6-dimensional vector $\vec{\epsilon} = (\epsilon_{xx}, \epsilon_{yy}, \epsilon_{zz}, 2\epsilon_{yz}, 2\epsilon_{xz}, 2\epsilon_{xy})$ to represent ϵ. It is easy to check that the elements of $\vec{\epsilon}$ are given as $\vec{\epsilon}_i = \mathrm{Tr}[B_i^\dagger \epsilon]$, where

$$B_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; B_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; B_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{E.2}$$
$$B_4 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \; B_5 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \; B_6 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The strain tensor transforms like $\epsilon \to O^{(i)} \epsilon \left(O^{(i)}\right)^\dagger$ under the T_d group elements, equivalently, the vectorized strain tensor transforms like $\vec{\epsilon} \to \mathcal{O}^{(i)} \cdot \vec{\epsilon}$, where $\{\mathcal{O}^{(i)}\}$ is a six-dimensional faithful matrix representation of the T_d group. The matrix elements of $\mathcal{O}^{(k)} \in \mathrm{M}_6(\mathbb{R})$ are given as $\left(\mathcal{O}^{(k)}\right)_{i,j} = \mathrm{Tr}\left[C_i^\dagger O^{(k)} B_j \left(O^{(k)}\right)^\dagger\right]$, for $1 \leq i, j \leq 6$, where $C_i = B_i$, for $i \in \{1, 2, 3\}$, and $C_i = B_i/2$, for $i \in \{4, 5, 6\}$.

In Section 6.1 of Chapter 6, we said that the hyperfine constant for the ^{31}P donor nucleus and the electron localized to the nuclear site can be Taylor expanded in the strain tensor ϵ at the site of the donor, such that $A_e(\vec{\epsilon}) = A_e(\vec{\epsilon} = 0)\left[1 + \left(\vec{\nabla} A_e\right) \cdot \vec{\epsilon} + \vec{\epsilon} \cdot \mathbf{H}_{A_e} \cdot \vec{\epsilon} + ...\right]$. We now derive the result given in [73], where it was shown that $\vec{\nabla} A_e$ and $\mathbf{H}_{A_e}$ can only take certain symmetry-restricted forms. Because the defect site has tetrahedral point symmetry, the application of any strain $\vec{\epsilon}$ must perturb A_e the same way as the application of $\mathcal{O}^{(i)}\vec{\epsilon}$, for all i, i.e.,

$$\left(\vec{\nabla} A_e\right) \cdot \mathcal{O}^{(i)} \vec{\epsilon} = \left(\vec{\nabla} A_e\right) \cdot \mathcal{O}^{(j)} \vec{\epsilon}, \tag{E.3}$$

for any $\vec{\epsilon}$ and all $\{i, j\}$. Because $\left(\mathcal{O}^{(i)}\right)^\dagger \in \{\mathcal{O}^{(k)}\}$, for all i, and $\mathcal{O}^{(1)} = \mathbb{1}_6$, the equation above is equivalent to

$$P \cdot \left(\vec{\nabla} A_e\right) = \vec{\nabla} A_e, \tag{E.4}$$

i.e., $\vec{\nabla} A_e$ must be a linear combination of P eigenvectors, where $P = \frac{1}{24}\sum_{i=1}^{24} \mathcal{O}^{(i)}$. Given the construction for $\{\mathcal{O}^{(i)}\}$ above, it easy to check with Mathematica that P is a rank-one projector with its only eigenvalue-one eigenvector being $(1,1,1,0,0,0)$. This implies that the only symmetry-allowed first order perturbation to $A_e(\vec{\epsilon})$ in $\vec{\epsilon}$ is proportional to the trace of ϵ, i.e., $\left(\vec{\nabla} A_e\right) \cdot \vec{\epsilon} = W\left(\epsilon_{xx} + \epsilon_{yy} + \epsilon_{zz}\right)/3$, where W is a constant.

For second order perturbation of $A_e(\vec{\epsilon})$ it has to hold that

$$\vec{\epsilon} \cdot \mathcal{O}^{(i)} \mathbf{H}_{A_e} \mathcal{O}^{(i)} \cdot \vec{\epsilon} = \vec{\epsilon} \cdot \mathcal{O}^{(j)} \mathbf{H}_{A_e} \mathcal{O}^{(j)} \cdot \vec{\epsilon}, \tag{E.5}$$

for any $\vec{\epsilon}$ and all $\{i,j\}$. This, in turn, implies that the vectorized Hessian matrix $\mathbf{H}_{A_e}$ has to be a linear combination of P eigenvectors, where $P = \frac{1}{24}\sum_{i=1}^{24} \mathcal{O}^{(i)} \otimes \mathcal{O}^{(i)}$, and where $\{\mathcal{O}^{(i)} \otimes \mathcal{O}^{(i)}\}$ is a 36-dimesnional faithful representation of the T_d group. Again, it is easy to check with Mathematica that P is a rank-three projector, which means that there are three independent symmetry allowed terms such that $\vec{\epsilon} \cdot \mathbf{H}_{A_e} \cdot \vec{\epsilon} = X\left(\epsilon_{xx}^2 + \epsilon_{yy}^2 + \epsilon_{zz}^2\right) + Y\left(\epsilon_{yy}\epsilon_{zz} + \epsilon_{xx}\epsilon_{zz} + \epsilon_{xx}\epsilon_{yy}\right) + Z\left(\epsilon_{yz}^2 + \epsilon_{xz}^2 + \epsilon_{xy}^2\right)$, where X, Y and Z are constants.

In Section 6.7 of Chapter 6, we said that by extending the analysis in [73], we can look at the symmetry allowed terms for perturbations of A_e under both strain $\vec{\epsilon}$ and external electric field $\vec{E} = (E_x, E_y, E_z)$, evaluated in the crystallographic coordinate system. This time the symmetry sets a requirement that

$$A_e\left(O^{(i)} \cdot \vec{E}, \mathcal{O}^{(i)} \cdot \vec{\epsilon}\right) = A_e\left(O^{(j)} \cdot \vec{E}, \mathcal{O}^{(j)} \cdot \vec{\epsilon}\right), \tag{E.6}$$

for any $\vec{E}$ and $\vec{\epsilon}$ and all $\{i,j\}$. Consequently, we need to look at projectors P constructed of 9-dimensional faithful matrix representations of the T_d group that have a direct sum structure, and correspond to the simultaneous transformations of $\vec{E}$ and $\vec{\epsilon}$ seen above:

$$\{\Omega^{(i)} = \begin{pmatrix} O^{(i)} & 0 \\ 0 & \mathcal{O}^{(i)} \end{pmatrix}\}. \tag{E.7}$$

It is easy to see that

$$P = \frac{1}{24}\sum_{i=1}^{24} \Omega^{(i)} = \frac{1}{24}\begin{pmatrix} \sum_{i=1}^{24} O^{(i)} & 0 \\ 0 & \sum_{i=1}^{24} \mathcal{O}^{(i)} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{24}\sum_{i=1}^{24} \mathcal{O}^{(i)} \end{pmatrix}, \tag{E.8}$$

meaning that there are no symmetry allowed perturbations to $A_e\left(\vec{E}, \vec{\epsilon}\right)$ that are linear in $\vec{E}$, the only symmetry allowed linear term is the trace of the strain tensor as we showed above. For second order perturbations in $\vec{E}$ and $\vec{\epsilon}$, we need to look at $P = \frac{1}{24}\sum_{i=1}^{24} \Omega^{(i)} \otimes \Omega^{(i)}$

like before. It is easy to check with Mathematica that the eigenvalue-one eigenspace of such P introduces two extra second order symmetry allowed perturbations of $A_e\left(\vec{E},\vec{\epsilon}\right)$: one purely from the electric field, that is proportional to $\vec{E}\cdot\vec{E}$, and one from the interplay of the strain and the electric field that is proportional to $E_x\epsilon_{yz}+Ey\epsilon_{xz}+E_z\epsilon_{xy}$. Denoting $\vec{\epsilon}^{\,(4,5,6)}=(2\epsilon_{yz},2\epsilon_{xz},2\epsilon_{xy})$, we can now write

$$\begin{aligned}A_e\left(\vec{E},\vec{\epsilon}\right)=A_e\left(0,0\right)&\left[1+\frac{W}{3}\left(\epsilon_{xx}+\epsilon_{yy}+\epsilon_{zz}\right)+U\vec{E}\cdot\vec{\epsilon}^{\,(4,5,6)}+V\left|\vec{E}\right|^2\right]\\&+A_e\left(0,0\right)\left[X\left(\epsilon_{xx}^2+\epsilon_{yy}^2+\epsilon_{zz}^2\right)+Y\left(\epsilon_{yy}\epsilon_{zz}+\epsilon_{xx}\epsilon_{zz}+\epsilon_{xx}\epsilon_{yy}\right)\right]\\&+A_e\left(0,0\right)\left[Z\left(\epsilon_{yz}^2+\epsilon_{xz}^2+\epsilon_{xy}^2\right)+...\right],\end{aligned}\tag{E.9}$$

where U, V, W, X, Y and Z are constants.

Finally, the symmetry analysis presented in this appendix also fixes the Hamiltonian form of the hyperfine interaction between the ^{31}P nuclear spin and the localized defect electron spin in the presence of no strain or electric field. Most general hyperfine interaction Hamiltonian between two spin 1/2 particles can be written as

$$H_{\mathrm{HF}}=\frac{1}{4}\sum_{i,j\in\{x,y,z\}}A_{i,j}\sigma_i\otimes\sigma_j,\tag{E.10}$$

where the hyperfine tensor A is a real symmetric tensor. In the case of a phosphorus donor defect located in a perfect diamond cubic lattice, the hyperfine tensor A has to be invariant under all 24 tetrahedral group symmetry transformations, i.e., $o^{(i)}A\left(o^{(i)}\right)^\dagger=o^{(j)}A\left(o^{(j)}\right)^\dagger$, for all $\{i,j\}$, where $\{o^{(i)}\}$ is a matrix representation of the T_d group evaluated in the same coordinate system as the hyperfine tensor. As we showed before, the only real symmetric tensor that is invariant under all of these transformations is proportional to the identity operator such that

$$H_{\mathrm{HF}}=\frac{A_e}{4}\sum_{i\in\{x,y,z\}}\sigma_i\otimes\sigma_i,\tag{E.11}$$

where A_e is the hyperfine constant.