Title
Controlling Quantum Systems for Quantum Information Processing

Author
Young, Kevin Christopher

Publication Date
2010

Peer reviewed|Thesis/dissertation

Controlling Quantum Systems for Quantum Information Processing

by

Kevin Christopher Young

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Physics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor K. Birgitta Whaley, Co-Chair

Professor Irfan Siddiqi, Co-Chair

Professor Ashvin Vishwanath

Professor Jeffrey Bokor

Fall 2010

Controlling Quantum Systems for Quantum Information Processing

Kevin Christopher Young

Abstract

Controlling Quantum Systems for Quantum Information Processing

by

Kevin Christopher Young

Doctor of Philosophy in Physics

University of California, Berkeley

Professor K. Birgitta Whaley, Co-Chair

Professor Irfan Siddiqi, Co-Chair

For several decades it has been appreciated that quantum computers hold incredible promise to perform calculations intractable to classical computation. However, this promise has be slow to realize. Dozens of quantum systems are currently being investigated for use in quantum information processing - none of which have yet demonstrated algorithms involving more than a handful of qubits and it remains unclear which, if any, of these systems will ultimately compose a scalable, robust quantum information processing architecture. In this thesis we employ analytical, optimal and algebraic control techniques to evaluate various quantum systems for their potential use in quantum information processing. In doing so, we have additionally identified several novel characterization procedures capable of probing both the coherent and incoherent dynamics of quantum systems. The first part of this thesis discusses work motivated by attempts to utilize donor qubits in silicon as quantum bits. We first propose a measurement of the state of a single donor electron spin using two-dimensional electron gas of a field-effect transistor and electrically detected magnetic resonance. We analyze the potential sensitivity of this measurement and show that it is a quantum nondemolition measurement of an electron-encoded state.

We then present the first of two novel qubit characterization procedures. We consider the problem of rapidly characterizing a large number of similarly prepared qubits using techniques from optimal experiment design. All qubits are assumed to evolve according to the same physical processes, though the Hamiltonian parameters may vary from device to device - an inevitability in solid state qubits. We use the Cramér-Rao bound on the variance of a point estimator to construct the optimal series of experiments to estimate these free parameters, and present a complete analysis of the optimal experimental configuration. Though applied to dipole- and exchange-coupled qubits, this technique is widely applicable to other systems.

The second part of the thesis discusses the role that control can play in measuring and mitigating noise in qubit systems. Our first result describes a method for quickly simulating the effects of arbitrary markovian noise on qubit systems through the use of a numerically optimized, multi-state Markovian fluctuator. This ability to rapidly simulate the noisy qubit

evolution allows us to compute control sequences capable of maximally decoupling the qubit from the noise source.

We then introduce the second characterization procedure of the these, showing that a single measurable and controllable qubit may act as a spectrometer of dephasing noise. We show that the formalism of dynamical decoupling can be used to estimate the short-time correlation function of the noise source, while long time correlations may be estimated by a very simple series of free evolution experiments. This technique is applicable to the wide range of physical implementations which suffer from dephasing noise.

The final part of this thesis demonstrates that trapped neutral atoms may be utilized for the robust simulation of complex systems exhibiting a topological phase. We present a method to simulate the toric code Hamiltonian stroboscopically, and demonstrate that our technique preserves the ground state degeneracy . Furthermore, we introduce a dissipative mechanism allowing for thermalization of the system to a finite temperature or direct cooling to the ground state manifold.

For my wife, Michelle Alison Young,
who thought she was done with physics forever.

Contents

List of Figures

List of Tables

Acknowledgments

The past five years of my life have been spent in the company of some truly amazing people, and I am honored to recognize their contributions to this thesis here. I have been privileged to have many wonderful physics teachers over the years, and I thank Patricia Armillei, John Hakkila, Jeffrey Wragg, and Robert Littlejohn for inspiring lectures and sage advice. I, and hundreds of graduate students before me, owe much to Anne Takizawa, Donna Sakima and Claudia Trujillo; the Berkeley Physics Department would certainly grind to a halt were it not for their heroic efforts. I will always be indebted to my advisor, K. Birgitta Whaley, for introducing me to a fascinating field and always having an interesting problem for me to work on. My collaborators in and beyond the Whaley Group have included Chris Herdman, Jon Aytac, Vito Scarola, Robert Kosut, Brendan Abolins, Travis Beals, Jan Ivar Korsbakken, and Dylan Gorman, and I thank each of them for years of advice and inspiration. Mohan Sarovar deserves special acknowledgement both for his seemingly endless patience as well as his astounding breadth of knowledge. Furthermore, the work in this thesis has benefited immeasurably through close contact with experimentalists, Thomas Schenkel, Jeff Bokor, Cheuk Chi Lo, Dave Weiss, and my co-advisor, Irfan Siddiqi.

On a more personal level, I have been enriched and sustained by the lifelong friendships I have made during my time in graduate school. I thank Erwan Lescolan, Ansley Scott and Mike Webber, for many great meals and entertaining conversation. I will be forever thankful for Jennie Guzman, Josh Meyers, Lauren Tompkins, Kyle Barbary and especially to Jessica Kirkpatrick, who brought us all together so often. I'm very much indebted to Yolanda Hagar, a friend and colleague, for her tireless explanations of statistics. I thank Shannon McCurdy, Leslie Elwell, and Sarah Placella for opportunities to vent once in a while. I am incredibly grateful to Dan Slichter and Jesse Noffsinger, and I will sincerely miss our weekly lunches. And my very dear friends, Shannon Lambert-Ryan and Daniel Gordon, have been like family to me for as long as I can remember and I cannot thank them enough for their many years of loving support.

I must also thank my family: my father, Scott Young; my mother, Sandy Erickson; and my sister, Courtney Young, for a lifetime of encouragement. But most of all, I owe everything I have accomplished to the unremitting support of my amazing wife, Michelle Young.

Part I

Introduction

Chapter 1

General introduction

1.1 Overview

That quantum systems can be especially useful for performing calculations was first pointed out in 1982 [1]. However, it wasn't until Shor's factoring algorithm [2] appeared in 1994 that the field of quantum information began to attract mainstream attention. In the years that followed, nearly every existing quantum technology has been investigated for its potential use in quantum information processing (QIP), and several new technologies have been developed specifically for this purpose. The field is currently engaged in a furious search for a viable experimental system in which nontrivial quantum algorithms may be implemented; among those considered are:

- Liquid-state nuclear magnetic resonance (NMR) [3]
- Trapped ions [4]
- Trapped neutral atoms in an optical lattice [5]
- Cavity quantum electrodynamics (QED) with atoms [6]
- Linear optics [7]
- Solid state (quantum dots [8], donor spin qubits [9], NV centers in diamond [10])
- Josephson junctions (charge, flux, phase) [11]
- Electrons on the surface of liquid helium [12]

However, each of these approaches has encountered challenges. Liquid-state NMR quantum computing, though providing several spectacular early results [3], is inherently not scalable. Solid state qubits, while easily controlled, are not easily measured and the route to inter-qubit coupling is unclear. Noticing the successes and failures of the various implementations, and to aid in the design of future QIP devices, DiVincenzo described a set of criteria [13] that must be satisfied by a physical system for it to be useful for QIP. These criteria require that the system:

1. Be a scalable physical system with well-defined qubits.
2. Be initializable to a simple fiducial state such as $|000\ldots\rangle$.
3. Have long coherence times.
4. Have a universal set of quantum gates.
5. Permit high quantum efficiency, qubit-specific measurements.
6. Allow the intraconversion of stationary and flying qubits.
7. Allow the faithful transmission of flying qubits between specified stationary qubits.

For a given a physical system, it is rarely clear at first which, if any, of these criteria are satisfiable. Furthermore, the failure of a system to satisfy the any of above requirements may not be due not to an intrinsic inadequacy of the approach - more likely is that the field has simply not found a way to make the system behave appropriately. or example, in liquid-state NMR quantum computing, it is impossible to initialize the to a fiducial state at finite temperature. However, a clever technique involving the construction of so-called pseudopure initial states allowed for the successful implementation of the most intricate quantum algorithm ever attempted [3], Shor's algorithm on . By considering novel quantum computing architectures, and by exploiting analytical, optimal and algebraic control techniques, the work presented in this thesis suggests several new approaches to the construction of a functional, scalable quantum information processing devices.

Part II of this thesis presents work motivated by a long-term collaboration with Dr. Thomas Schenkel and Professor Jeffrey Bokor studying the feasibility of isolated donor atoms in silicon to act as qubits. Donors in silicon are an attractive route to scalable quantum information processing because of the potential to bootstrap the massive infrastructure developed by the semiconductor industry. In Chapter 3 we propose a technique for measuring the state of a single donor electron spin using a field-effect transistor induced two-dimensional electron gas and electrically detected magnetic resonance techniques. The scheme is facilitated by hyperfine coupling to the donor nucleus. We analyze the potential sensitivity and outline experimental requirements. Our measurement provides a single-shot, projective, and quantum non-demolition measurement of an electron-encoded qubit state.

With Chapter 4 we consider the problem of quantum multi-parameter estimation with experimental constraints and formulate the solution in terms of a convex optimization. Specifically, we outline an efficient method to identify the optimal strategy for estimating multiple unknown parameters of a quantum process and apply this method to a realistic example. The example is two electron spin qubits coupled through the dipole and exchange interactions with unknown coupling parameters – explicitly, the position vector relating the two qubits and the magnitude of the exchange interaction are unknown. This coupling Hamiltonian generates a unitary evolution which, when combined with arbitrary single-qubit operations, produces a universal set of quantum gates. However, the unknown parameters must be known precisely to generate high-fidelity gates. We use the Cramér-Rao bound on the variance of a point estimator to construct the optimal series of experiments to estimate these free parameters, and present a complete analysis of the optimal experimental configuration. Our method of transforming the constrained optimal parameter estimation problem into a convex optimization is powerful and widely applicable to other systems.

Part III describes the various roles that control may play in studying and mitigating the effects of noise in quantum systems. In Chapter 5 we address the experimentally relevant problem of robust mitigation of dephasing noise acting on a qubit. We first present an extension of a method for representing $1/\omega^{\alpha}$ noise developed by Kuopanportti et al. [14] to the efficient representation of arbitrary Markovian noise. We then add qubit control pulses to enable the design of numerically optimized, two-dimensional control functions with bounded amplitude, that are capable of decoupling the qubit from the dephasing effects of a broad

variety of Markovian noise spectral densities during arbitrary one qubit quantum operations. We illustrate the method with development of numerically optimized control pulse sequences that minimize decoherence due to a combination of $1/\omega$ and constant offset noise sources. Comparison with the performance of standard dynamical decoupling protocols shows that the numerically optimized pulse sequences are considerably more robust with respect to the noise offset, rendering them attractive for application to situations where homogeneous dephasing noise sources are accompanied by some extent of heterogeneous dephasing. Application to the mitigation of dephasing noise on spin qubits in silicon indicates that high fidelity single qubit gates are possible with current pulse generation technology.

In Chapter 6 we investigate the extent to which a single qubit can act as a spectrometer of classical dephasing noise. By analyzing the limitations of single-qubit initialization, control and measurement, we find a natural separation of the problem of measuring the noise correlation function into two distinct timescales. By bootstrapping dynamical decoupling sequences, we show that a qubit can be used to estimate the short-time correlation function of the noise source, while long time correlations may be estimated by a very simple series of free evolution experiments. Our technique is applicable to all qubit implementations which suffer from dephasing noise.

Finally, Part IV shows that trapped neutral atoms offer a powerful route to the robust simulation of complex quantum systems. We present here a stroboscopic scheme for realization of a Hamiltonian with n-body interactions on a set of neutral atoms trapped in an addressable optical lattice, using only 1- and 2-body physical operations together with a dissipative mechanism that allows thermalization to finite temperature or cooling to the ground state. We demonstrate this scheme with application to the toric code Hamiltonian, ground states of which can be used to robustly store quantum information when coupled to a low temperature reservoir.

1.2 Additional information for the dissertation committee

The chapters that comprise this thesis each represent a separate research project. Several of these projects were undertaken in collaboration with various coauthors, and we discuss the division of labor here.

Chapter 3 is based on material from "Quantum nondemolition measurements of single donor spins in semiconductors," written in collaboration with Mohan Sarovar, Thomas Schenkel and K. Birgitta Whaley, published in Phys. Rev. B 78, 245302 (2008). The scattering calculations were performed by M. Sarovar, while the present author worked out the details of the electron-nuclear state transfer gate. The remaining mathematical work was carried out jointly by M. Sarovar and the present author.

Chapter 4 presents material from "Optimal quantum multiparameter estimation and application to dipole- and exchange-coupled qubits." This work was performed with Mohan Sarovar, Robert Kosut and K. Birgitta Whaley, and was published as Phys. Rev. A 79,

062301 (2009). This work was inspired by R. Kosut's suggestion that classical experiment design protocols could be implemented in a quantum setting. Nearly all mathematical work and all simulations were carried out by the present author.

Chapter 5 is based on "Fighting dephasing noise with optimal control," written with Dylan Gorman and K. Birgitta Whaley, available on the preprint arXiv, as arXiv:1005.5418 and submitted for publication to Phys. Rev. A. The technique of modeling of arbitrary Markovian spectra was developed by the present author, while many of the optimal control calculations were done by Dylan Gorman as part of his work with UC LEADS under the supervision of the author.

Chapter 6 is based on as-yet-unpublished work that was performed entirely by the present author under the supervision of K. Birgitta Whaley. This work was motivated by a desire to understand dephasing dynamics of donor qubits in silicon due to fluctuations at the silicon-silicon dioxide interface.

Chapter 7 is based on "Stroboscopic Generation of Topological Protection," work performed in collaboration with Chris Herdman, Vito Scarola, Mohan Sarovar and K. Birgitta Whaley and was published in Phys. Rev. Lett. 104, 230501 (2010). Application of the Magnus expansion to simulating many-body interactions was developed by the present author and C. Herdman performed the numerical experiments demonstrating that topological degeneracy was preserved by the stroboscopic implementation. C. Herdman and the present author independently developed different approaches to the thermalization procedure, though C. Herdman's proved to be more general. The present author's approach is presented in Appendix. 7.7.4.

Chapter 2

The model spin qubit

In this chapter we introduce the spin-$\frac{1}{2}$ system as a model qubit. We begin by reviewing the Bloch sphere representation of one qubit states and magnetic resonance. We close the chapter with a brief discussion of decoherence and show that control is capable of making decoherence effects less severe.

2.1 Single qubit states: the Bloch sphere

The density matrix of the spin-$\frac{1}{2}$ system can be decomposed into a weighted sum of Pauli spin matrices as

$$\rho = \frac{1}{2}\left(\mathbb{1} + \langle\vec{\sigma}\rangle \cdot \vec{\sigma}\right)$$

Here, and throughout this thesis, the notation, $\mathbb{1}$, will be used to represent the identity matrix and, $\vec{\sigma}$, the vector of Pauli matrices. The above expression defines the *coherence vector*, $\langle\vec{\sigma}\rangle$ and provides us with a one-to-one mapping of single qubit states to the unit 3-sphere. This mapping is called the Bloch sphere, and is illustrated in Fig. 2.1. Points on

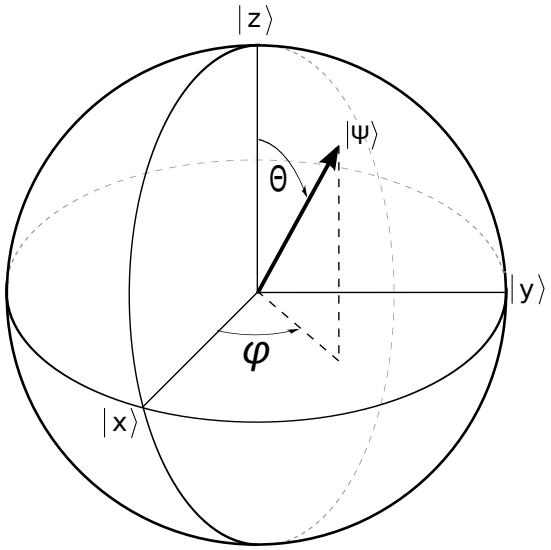


Figure 2.1: The Bloch sphere.

the surface of the Bloch sphere represent pure quantum states,

$$|\hat{n}\rangle = \sin(\theta/2)\,|0\rangle + \cos(\theta/2)e^{-i\phi}\,|1\rangle\,. \tag{2.1}$$

Here ϕ and θ correspond to the spherical coordinates of the point on the surface. The state Eq. 2.1 is the $+1$ eigenstate of the operator $\hat{n}\cdot\vec{\sigma}$. Points on the interior of the Bloch sphere are mixed states.

2.2 Spin resonance

2.2.1 Larmor precession

In the presence of a magnetic field, the qubit will evolve under the Hamiltonian

$$H = -\vec{\mu} \cdot \vec{B} = \frac{1}{2}\gamma \vec{\sigma} \cdot \vec{B}$$

where γ is the gyromagnetic ratio.

We shall now consider the time evolution of a state $|\psi(t)\rangle$, with the initial condition $|\psi(0)\rangle = |\hat{x}\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$, when placed in a constant magnetic field, $\vec{B} = B_0\hat{z}$. The time evolution is given by solving the Schrödinger equation:

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

Here, and for the remainder of this thesis we set $\hbar = 1$. In terms of the above Hamiltonian, the state evolves as,

$$\begin{aligned}
|\psi(t)\rangle &= e^{-i\gamma B_0 \sigma_z/2}\left(\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\right) \\
&= \frac{1}{\sqrt{2}}\left(e^{-i\gamma B_0 t/2}|0\rangle + e^{i\gamma B_0 t/2}|1\rangle\right) \\
&= \frac{1}{\sqrt{2}}\left(|0\rangle + e^{i\gamma B_0 t}|1\rangle\right)
\end{aligned}$$

In the last equality we noted that we exploited our freedom to multiply the wavefunction by an arbitrary phase, in this case, $\exp(i\gamma Bt)$. Comparison with Eq. (2.1) shows that this state is precessing about the applied magnetic field at the Larmor frequency, $\omega_L = \gamma B_0$. This is illustrated in Fig. 2.2. In fact, any state will evolve a relative phase, $\phi(t) = \gamma B_0 t$, and will precess at the same frequency.

2.2.2 Resonance

We now consider the problem of computing the qubit evolution in the presence of an *oscillating* field, $B_1\cos(\omega t)\,\hat{x}$, in addition to the applied static field. The Hamiltonian thus becomes:

$$H(t) = \frac{1}{2}\gamma\left(B_0\sigma_z + B_1\cos(\omega t)\,\sigma_x\right)$$

Note that, in general, $[H(t), H(t')] \neq 0$, so we cannot write the time evolution operator simply as, $U(t) = \exp\left(-i\int_0^t H(t')\,dt'\right)$. Instead, we would have to use the time ordering operator, $\underleftarrow{\mathcal{T}}$, but this does not help us evaluate the evolution operator in closed form.

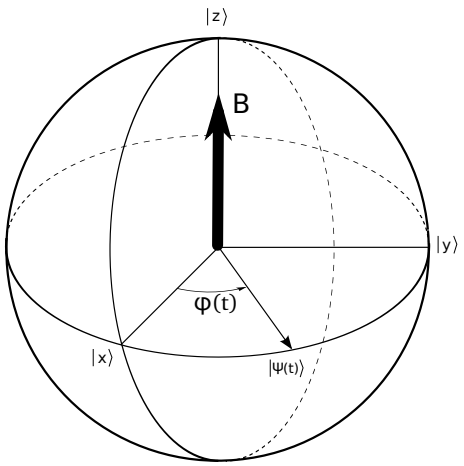


Figure 2.2: Bloch sphere representation of spin procession in constant field. The state, $|\psi(t)\rangle$ is precessing about the applied field at frequency, $\omega_L = \gamma B_0$.

Instead we work in an interaction picture, in reference frame rotating at the Larmor frequency. We begin by defining a new state variable in this rotating frame,

$$|\psi'(t)\rangle = e^{i\omega_L \sigma_z t/2} |\psi(t)\rangle$$

The time evolution of this state is

$$\begin{aligned}
i\frac{d}{dt}|\psi'(t)\rangle &= (i)i\omega_L\sigma_z/2\,|\psi'(t)\rangle + e^{i\omega_L\sigma_z t/2} i\frac{d}{dt}|\psi(t)\rangle \\
&= -\frac{\omega_L\sigma_z}{2}|\psi'(t)\rangle + e^{i\omega_L\sigma_z t/2}H(t)\,|\psi(t)\rangle \\
&= -\frac{\omega_L\sigma_z}{2}|\psi'(t)\rangle + e^{i\omega_L\sigma_z t/2}H(t)e^{-i\omega_L\sigma_z t/2}e^{i\omega_L\sigma_z t/2}\,|\psi(t)\rangle \\
&= \left(-\frac{\omega_L\sigma_z}{2} + e^{i\omega_L t/2}H(t)e^{-i\omega_L\sigma_z t/2}\right)|\psi'(t)\rangle \\
&= H'(t)\,|\psi'(t)\rangle
\end{aligned}$$

Which defines the Hamiltonian in the rotating frame:

$$H'(t) = \left(-\frac{\omega_L}{2}\sigma_z + e^{i\omega_L\sigma_z t/2}H(t)e^{-i\omega_L\sigma_z t/2}\right)$$

Focusing on this last term:

$$\begin{aligned}
e^{i\omega_L\sigma_z t/2}H(t)e^{-i\omega_L\sigma_z t/2} &= \frac{\gamma}{2}e^{i\omega_L\sigma_z t/2}\left(B_0\sigma_z + B_1\cos(\omega t)\,\sigma_x\right)e^{-i\omega_L\sigma_z t/2} \\
&= \frac{\gamma}{2}\left(B_0\sigma_z + e^{i\omega_L\sigma_z t/2}B_1\cos(\omega t)\sigma_x e^{-i\omega_L\sigma_z t/2}\right)
\end{aligned}$$

We now rewrite the linearly polarized field as the sum of two counter rotating circular waves:

$$\begin{aligned}\cos(\omega t)\sigma_x &= \frac{1}{2}\left((\cos(\omega t)\sigma_x + \sin(\omega t)\sigma_y) + (\cos(\omega t)\sigma_x - \sin(\omega t)\sigma_y)\right) \\ &= \frac{1}{2}\left(\left(e^{-i\omega t}\sigma_+ + e^{i\omega t}\sigma_-\right) + \left(e^{i\omega t}\sigma_+ + e^{-i\omega t}\sigma_-\right)\right)\end{aligned}$$

Here we have used the raising and lowering Pauli operators, $\sigma_\pm = (\sigma_x \pm i\sigma_y)/2$. We calculate the transformations, $\exp(i\alpha\sigma_z)\sigma_\pm \exp(-i\alpha\sigma_z)$:

$$e^{i\alpha\sigma_z}\sigma_+ e^{-i\alpha\sigma_z} = \begin{pmatrix} e^{i\alpha} & 0 \\ 0 & e^{-i\alpha} \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} e^{-i\alpha} & 0 \\ 0 & e^{i\alpha} \end{pmatrix} = \begin{pmatrix} 0 & e^{2i\alpha} \\ 0 & 0 \end{pmatrix} = e^{2i\alpha}\sigma_+$$

$$e^{i\alpha\sigma_z}\sigma_+ e^{-i\alpha\sigma_z} = \begin{pmatrix} e^{i\alpha} & 0 \\ 0 & e^{-i\alpha} \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} e^{-i\alpha} & 0 \\ 0 & e^{i\alpha} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ e^{-2i\alpha} & 0 \end{pmatrix} = e^{-2i\alpha}\sigma_-$$

This gives,

$$\begin{aligned}H'(t) &= -\frac{\omega_L}{2}\sigma_z + \frac{\gamma}{2}\left(B_0\sigma_z + B_1 e^{i\omega_L\sigma_z t/2}\cos(\omega t)\,\sigma_x e^{-i\omega_L\sigma_z t/2}\right) \\ &= -\frac{\omega_L}{2}\sigma_z + \frac{\omega_L}{2}\sigma_z \\ &\quad + \frac{\gamma}{4}\left(B_1 e^{i\omega_L\sigma_z t/2}\left(\left(e^{-i\omega t}\sigma_+ + e^{i\omega t}\sigma_-\right) + \left(e^{i\omega t}\sigma_+ + e^{-i\omega t}\sigma_-\right)\right)e^{-i\omega_L\sigma_z t/2}\right) \\ &= \frac{\gamma}{4}B_1\left(\left(e^{i(\omega_L-\omega)t}\sigma_+ + e^{i(\omega-\omega_L)t}\sigma_-\right) + \left(e^{i(\omega+\omega_L)t}\sigma_+ + e^{-i(\omega+\omega_L)t}\sigma_-\right)\right)\end{aligned}$$

If the frequency of the applied field is equal to the Larmor frequency, ω_L, a condition known as *resonance*, then this simplifies:

$$H'(t) = \frac{\gamma}{4}B_1\left((\sigma_+ + \sigma_-) + \left(e^{2i\omega_L t}\sigma_+ + e^{-2i\omega_L t}\sigma_-\right)\right)$$

However, the second term oscillates so fast that its effect on the qubit dynamics are negligible, so we drop it. This is known as the *rotating wave approximation*. The Hamiltonian under this approximation is then,

$$H'(t) = \frac{\gamma}{4}B_1(\sigma_+ + \sigma_-) = \frac{\gamma}{4}B_1\sigma_x$$

This Hamiltonian will cause the Bloch sphere representation of the state to rotate about the $\hat{x}$-axis at the Rabi frequency, $\omega_R = \gamma B_1/2$. The path on the Bloch sphere of a state experiencing a *pi* rotation is illustrated in Fig. 2.3.

2.3 Decoherence

2.3.1 The Lindblad master equation

While unitary evolution corresponds to ridged rotations of the Bloch sphere, physics permits much more general operations on qubit states. Many of these transformations may be

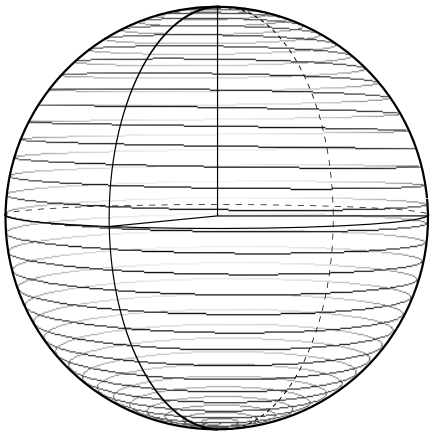


Figure 2.3: Lab frame Bloch sphere representation of Rabi oscillation.

described phenomenologically through the use of quantum master equations [15],

$$\dot{\rho}(t) = \mathcal{L}(t)\rho(t).$$

Here we have introduced the Liouville superoperator, $\mathcal{L}$. Its is a linear operator on the elements of the density matrix. For instance, one may describe the usual Hamiltonian evolution in terms of superoperators as,

$$\dot{\rho}(t) = -i[H(t), \rho(t)] = \mathcal{L}_H(t)\rho(t).$$

The Liouville-Hamiltonian operator is thus defined as $\mathcal{L}_H \cdot = -i[H, \cdot]$. However, more general, nonunitary evolution is also possible, and this behavior may be generally modeled by Lindblad form of the master equation,

$$\dot{\rho}(t) = -i[H(t), \rho(t)] + \sum_j \left(2L_j\rho(t)L_j^\dagger - \left\{L_j^\dagger L_j, \rho(t)\right\}\right) = L_H(t)\rho(t) + \sum_j \mathcal{D}(L_j)\rho(t)$$

Here we have used the anticommutator $\{A, B\} = AB + BA$, L_j is a Lindblad operator and $\mathcal{D}(L)$ is the dissipator superoperator. Lindblad master equations are especially useful for describing the effective dynamics of a reduced density matrix.

2.3.2 Dephasing and the Hahn Echo

Dephasing

We consider here a qubit interacting with a large (many degrees of freedom) environment with initial state, ρ_B. We suppose that after a time, t, the system has evolved under the unitary operator,

$$U = |0\rangle\langle 0| \otimes U_0 + |1\rangle\langle 1| \otimes U_1$$

The initial state of the system is assumed separable,

$$\rho(0) = \rho_S^0 \otimes \rho_B^0$$

where

$$\rho_S^0 = a_{00} \left|0\right\rangle\!\left\langle 0\right| + a_{01} \left|0\right\rangle\!\left\langle 1\right| + a_{10} \left|1\right\rangle\!\left\langle 0\right| + a_{11} \left|1\right\rangle\!\left\langle 1\right|$$

and

$$\rho_B^0 = \sum_i \left|\psi_i\right\rangle\!\left\langle\psi_i\right| .$$

At time $t = 0$, the reduced density matrix of the qubit is the trace of the full density matrix over the bath degrees of freedom.

$$\rho_S(0) = \mathrm{Tr}_B \left(\rho_S^0 \otimes \rho_B^0\right) = \rho_S^0$$

At later times, however, the state evolves as

$$\begin{aligned}
\rho(t) &= U(\rho_S^0 \otimes \rho_B^0)U^\dagger \\
&= U((a_{00} \left|0\right\rangle\!\left\langle 0\right| + a_{01} \left|0\right\rangle\!\left\langle 1\right| + a_{10} \left|1\right\rangle\!\left\langle 0\right| + a_{11} \left|1\right\rangle\!\left\langle 1\right|) \otimes \rho_B^0)U^\dagger \\
&= \Big(a_{00} \left|0\right\rangle\!\left\langle 0\right| \otimes U_0\rho_B^0U_0^\dagger + a_{01} \left|0\right\rangle\!\left\langle 1\right| \otimes U_0\rho_B^0U_1^\dagger \\
&\qquad + a_{10} \left|1\right\rangle\!\left\langle 0\right| \otimes U_1\rho_B^0U_0^\dagger + a_{11} \left|1\right\rangle\!\left\langle 1\right| \otimes U_1\rho_B^0U_1^\dagger\Big)
\end{aligned}$$

Tracing out the bath now gives,

$$\begin{aligned}
\rho_S(t) &= \mathrm{Tr}_B \left(\rho(t)\right) \\
&= a_{00} \left|0\right\rangle\!\left\langle 0\right| + a_{11} \left|1\right\rangle\!\left\langle 1\right| + a_{10} \left|1\right\rangle\!\left\langle 0\right| \mathrm{Tr} \left(U_1\rho_B^0U_0^\dagger\right) + a_{01} \left|0\right\rangle\!\left\langle 1\right| \mathrm{Tr} \left(U_0\rho_B^0U_1^\dagger\right)
\end{aligned}$$

For systems with large bath degrees of freedom, this trace is likely to vanish,

$$\begin{aligned}
\mathrm{Tr} \left(U_0\rho_B^0U_1^\dagger\right) &= \mathrm{Tr} \left(U_1\rho_B^0U_0^\dagger\right)^* \\
&= \sum_k \sum_i \left\langle k\right| U_0 \left|\psi_i\right\rangle \left\langle\psi_i\right| U_1^\dagger \left|k\right\rangle \\
&= \sum_i \left\langle\psi_i\right| U_1^\dagger U_0 \left|\psi_i\right\rangle \\
&\simeq 0
\end{aligned}$$

This implies that the final reduced density matrix of the qubit is

$$\rho_S(t) \simeq a_{00} \left|0\right\rangle\!\left\langle 0\right| + a_{11} \left|1\right\rangle\!\left\langle 1\right| .$$

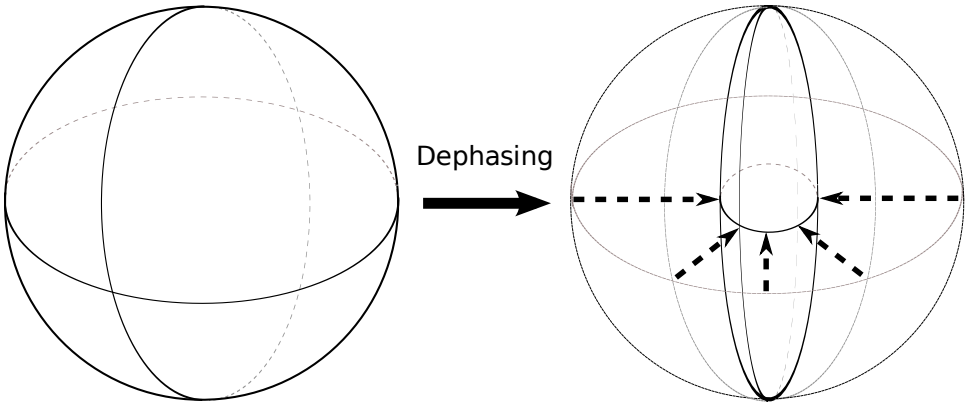


Figure 2.4: Illustration of the effect of dephasing on the Boch sphere. The qubit loses quantum coherences, and the Bloch sphere becomes more prolate.

We see that the off diagonal entries of the density matrix have decayed away through the interaction with the bath. This process is known generally as dephasing, specific example of decoherence. The effect on an aribitrary state is illustrated in Fig. 2.4.

Dephasing can be described quite easily in the Lindblad master equation picture. Let $L = \sqrt{\gamma}\sigma_z/2$. The master equation for a qubit subject to dephasing is then,

$$\dot{\rho}(t) = \frac{\gamma}{4}\left(2\sigma_z\rho(t)\sigma_z - \sigma_z^2\rho - \rho\sigma_z^2\right) = \frac{\gamma}{2}\left(\sigma_z\rho(t)\sigma_z - \rho\right)$$

The density matrix at any time may be written,

$$\rho(t) = \begin{pmatrix} a(t) & c(t) \\ c(t)^* & b(t) \end{pmatrix}.$$

Under the above master equation, this evolves as,

$$\dot{\rho}(t) = \begin{pmatrix} 0 & -\gamma c(t) \\ -\gamma c(t)^* & 0 \end{pmatrix}$$

The coherences, as measured by $\langle\sigma_+\rangle$ evolve as

$$\langle\dot{\sigma}_+(t)\rangle = \mathrm{Tr}\,\sigma_+\dot{\rho}(t) = -\gamma c(t) = -\gamma\left\langle\sigma_+(t)\right\rangle.$$

Leading to $\langle\sigma_+(t)\rangle = \langle\sigma_+(0)\rangle\, e^{-\gamma t}$, an exponential decay in the coherence with rate γ.

Hahn Echo

To illustrate the possible benefits of control for mitigating decoherence, we shall now briefly review the Hahn echo [16] as applied to an ensemble of uncoupled spin-$\frac{1}{2}$ systems precessing in

an spatially inhomogeneous magnetic field. Each qubit, which we'll label with the subscript i, evolves under its own Hamiltonian,

$$H(\omega) = \frac{1}{2}\omega\sigma_z.$$

The inhomogineity of the magnetic field results in a precession rate, ω_i, which will be different for each qubit. We assume that each qubit begins in the same initial state,

$$\rho(0) = \frac{1}{2}\left(\mathbb{1} + \sigma_x\right).$$

and that the precession rates are distributed as $P(\omega) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\omega^2/2\sigma^2}$. Any given qubit will then evolve in time as

$$\begin{aligned}\rho(t;\omega) &= \frac{1}{2}\left(\mathbb{1} + e^{-i\omega\, t\,\sigma_z/2}\sigma_x e^{i\omega\, t\,\sigma_z/2}\right)\\ &= \frac{1}{2}\left(\mathbb{1} + \cos(\omega t)\sigma_x + \sin(\omega t)\sigma_y\right)\end{aligned}$$

The full density matrix of the system is then the average over all possible precession frequencies,

$$\begin{aligned}\rho(t) &= \int_{-\infty}^{\infty}\rho(t;\omega)P(\omega)d\omega\\ &= \frac{1}{2}\left(\mathbb{1} + \int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\omega^2/2\sigma^2}d\omega\left(\cos(\omega t)\sigma_x + \sin(\omega t)\sigma_y\right)\right)\\ &= \frac{1}{2}\left(\mathbb{1} + e^{-\sigma^2 t^2/2}\sigma_x\right)\end{aligned}$$

We see that the total coherence of the system decays exponentially in t^2.

What if instead of the free evolution we had applied a π pulse in the middle of the free evolution at time $T/2$. At times longer than $T/2$, the unitary evolution experienced by the qubit is,

$$U(t > T/2;\omega) = e^{-i\omega\,(t-T/2)\,\sigma_z/2}\sigma_x e^{-i\omega\, T\,\sigma_z/4} = \cos(\omega(t-T)/2)\sigma_x + \sin(\omega(t-T)/2)\sigma_y$$

Therefore, the state of the system at times greater than $T/2$ is given by,

$$\begin{aligned}\rho(t > T/2;\omega) &= \frac{1}{2}\left(\mathbb{1} + U(t;\omega)\sigma_x U^\dagger(t;\omega)\right)\\ &= \frac{1}{2}\left(\mathbb{1} + \cos(\omega(t-T))\sigma_x + \sin(\omega(t-T))\sigma_y\right)\end{aligned}$$

Averaging over the environment gives,

$$\rho(t > T/2) = \frac{1}{2}\left(\mathbb{1} + e^{-\sigma^2(t-T)^2/2}\sigma_x\right)$$

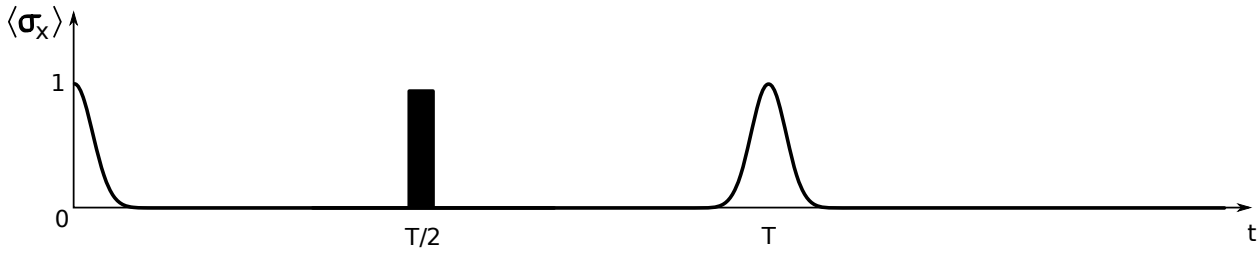


Figure 2.5: Trace of $\langle \sigma_x(t) \rangle$ with Hahn echo. Bold black line represents the π-pulse.

Which, at time $t = T$, is

$$\rho(t) = \frac{1}{2}\left(\mathbb{1} + \sigma_x\right)$$

Thus by applying a π-pulse, we have recovered the coherence of the system. This is the famous Hahn echo [16]. More complicated pulse sequences, like those described in Part III, are capable of recovering the coherence of a system subjected to much more damaging noise.

2.3.3 Other decoherence processes

Depolarizing

We saw in the previous section how noise along the z-axis can cause the qubit to dephase. Noise along the x- or y-axes instead cause probabilistic bit-flip errors. The Lindblad operators that model such a situation are $L_1 = \sqrt{\gamma}\sigma_x/2$ and $L_2 = \sqrt{\gamma}\sigma_y/2$. The result on the Bloch sphere is shown in Fig. 2.6.

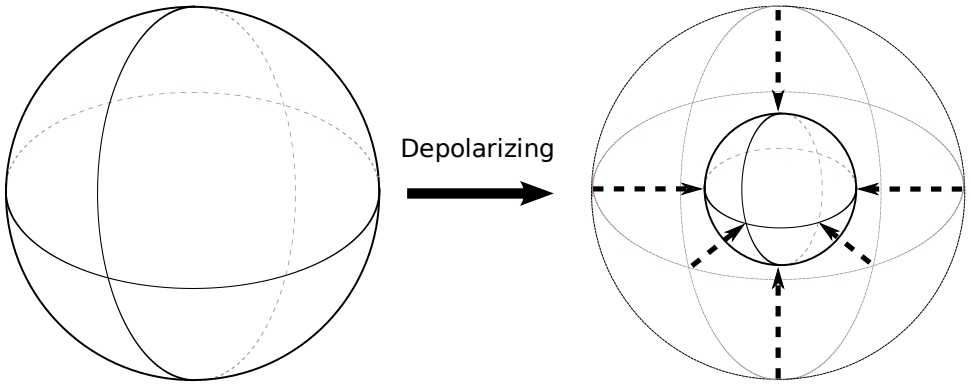


Figure 2.6: Illustration of the effect of depolarizing on the Bloch sphere. This may be thought of as a probabilistic bit flip error, and causes the Bloch sphere to shrink.

Amplitude damping

Amplitude damping is caused by relaxation processes, such as spontaneous emission or coupling to a low-temperature reservoir, which result in the upper state decaying to the lower state. This process may be modeled by the relaxation Lindblad operator, $L_R = \sqrt{(}\gamma)\sigma_-$, where γ is the classical relaxation rate. The effect of this operation on the Bloch sphere is shown in Fig. 2.7.

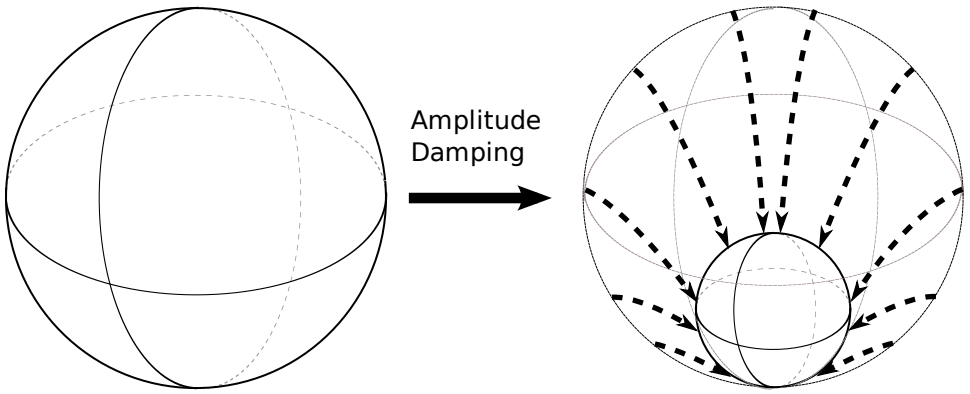


Figure 2.7: Illustration of the effect of amplitude damping on the Bloch sphere. Similar to spontaneous emission, the upper state probabilistically decays to the lower state, causing the Bloch sphere to shrink and move towards the lower state. This is an example of a nonunital channel.

Part II

Donor qubits in silicon

Chapter 3

Quantum nondemolition measurement of donor spins in silicon

3.1 Introduction

Semiconductor implementations of quantum computation have become a vibrant subject of study in the past decade because of the promise quantum computers (QCs) hold for radically altering our understanding of efficient computation, and the appeal of bootstrapping the wealth of engineering experience that the semiconductor industry has accumulated. A promising avenue for implementing quantum computing in silicon was proposed by Kane [17], suggesting the use of phosphorous nuclei to encode quantum information. However, while the long coherence times of the nuclei are advantageous for information storage tasks, their weak magnetic moment also results in long gate operation times. In contrast, donor *electrons* in Si couple strongly to microwave radiation and permit the fast execution of gates; and while electron spin decoherence times are shorter than their nuclear counterparts, the tradeoff of decreased robustness to noise for faster operation times could be appropriate to implementing a fault-tolerant QC. This has led several authors to suggest the use of electron spin qubits as a variant on the original Kane proposal (e.g. [18, 19, 20]), and we focus on such a modified Kane architecture here.

An integral part of any quantum computation architecture is the capacity for high-fidelity qubit readout. While small ensembles of donor spins have been detected [21, 22] and single spin measurements have been demonstrated (e.g. [23, 24]), detection of spin states of single donor electrons and nuclei in silicon has remained elusive. In this paper we analyze spin dependent scattering between conduction electrons and neutral donors [25, 26] as a spin-to-charge-transport conversion technique, and show that quantum non-demolition (QND) measurements of single electron spin-encoded qubit states are realistically achievable when mediated via nuclear spin states. Such a measurement will also be of value to the developing field of spintronics [27] where the electrical detection of spin states is valuable. Our readout takes advantage of two features: i) the ability to perform electron spin resonance spectroscopy using a two-dimensional electron gas (2DEG), and ii) the hyperfine shift induced on dopant electron Zeeman energies by the dopant nuclear spin state.

In the next section we describe the experimental apparatus and the techniques of 2DEG mediated spin resonance spectroscopy. In section 3.3 we present our proposal for spin state measurement in detail, and then in section 3.4 we analyze the sensitivity of the measurement scheme and establish the key factors that determine signal-to-noise. Then section 3.5 concludes with a discussion.

3.2 The physical setting

The use of electrical conductivity properties of semiconductors to investigate spin properties of (bulk-doped) impurities has a long history [28, 29], including studies of donor polarization using a 2DEG probe [25]. Figure 3.1 shows a cross section of a 2DEG spin readout device with a single implanted donor. Prior studies have used similar devices with bulk-doping [25, 30] or a large number of implanted donors (10^6) [26] in the 2DEG channel. The 2DEG

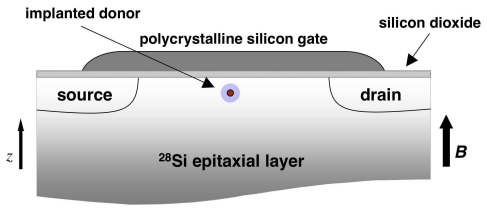


Figure 3.1: A cross section of the field-effect transistor (FET) used to create the 2DEG. In order to reduce qubit decoherence, it is beneficial to implant into isotopically purified silicon.

is operated in accumulation mode and thus conduction electrons scatter off the electron(s) bound to the shallow donor(s). The basic principle exploited in these studies is the role of the exchange interaction in electron-electron scattering. At a scattering event between a conduction electron and a loosely bound donor impurity electron, the Pauli principle demands that the combined wave function of the two electrons be antisymmetric with respect to coordinate exchange. This constraint, together with the fact that the combined spin state can be symmetric (triplet) or antisymmetric (singlet), imposes a correlation between the spatial and spin parts of the wave function and results in an effective spin dependence of the scattering matrix, leading to a spin dependent conductance. Application of a static magnetic field will partially polarize conduction and impurity electrons leading to excess triplet scattering. A microwave drive will alter these equilibrium polarizations when on resonance with impurity (or conduction) electron Zeeman energies and hence alter the ratio of singlet versus triplet scattering events, registering as a change in the 2DEG current. Thus, the spin dependent 2DEG current can be used as a detector of spin resonance and accordingly this technique is commonly known as *electrically detected magnetic resonance* (EDMR). Ghosh and Silsbee, and later Willems van Beveren *et. al.*, employed EDMR in bulk doped natural silicon to resolve resonance peaks corresponding to donor electron spins that are hyperfine split by donor P nuclei [25, 30]. Recently, Lo *et. al.* have used this technique to investigate spin dependent transport with micron-scale transistors on isotopically enriched ^{28}Si implanted with ^{121}Sb donors [26].

3.3 The proposal: EDMR based single spin measurement

From hereon we will consider the experimental setup described above in the particular situation where there is a single donor P nucleus present, the electron spin of which encodes the quantum information that we wish to measure. A crucial question in the context of quantum computing is whether the spin-dependent 2DEG current can be used to measure the state of an electron-spin qubit, as spin-dependent tunneling processes have been employed [23, 24]. The fundamental concern here is whether the spin exchange scattering interaction at the core of the spin-dependent 2DEG current allows for a quantum state measurement of

a single donor impurity electron spin.

A spin (1/2) state measurement couples the microscopic state of the spin, given in general by a (normalized) density matrix,

$$\rho_i = \begin{pmatrix} a & c \\ c^* & b \end{pmatrix} \tag{3.1}$$

(in the measurement basis, with $a + b = 1$), to a macroscopic meter variable I, the 2DEG current in our case. The meter variable can take one of two values, and at the conclusion of the measurement, a faithful measuring device would register each meter variable with the correct statistics, i.e., $I_\uparrow$ with probability a and $I_\downarrow$ with probability b. A QND measurement device will have the additional property that once a meter variable has been registered, the measured spin remains in the state corresponding to the value registered so that a second measurement gives the same result [31, 32].

One might expect that because the exchange interaction is destructive (in the sense that it will change the state of the target (donor electron) spin with some probability), it will only produce a faithful measurement if the time over which it acts is extremely short. We will now show that this is indeed the case and that direct measurement of the electron spin via the 2DEG current is consequently not possible within experimentally realizable times. This negative result will motivate our subsequent presentation in Section 3.3.2 of a more complex scheme for measurement that is both faithful and experimentally realizable.

3.3.1 Direct measurement of electron spin

To investigate the ability of the spin-dependent 2DEG current to measure the electronic spin state, we shall use a minimal model of the scattering process. Since we are primarily concerned with the spin state of the particles involved in the scattering, we examine the transformation that a single scattering event induces on the spinor components of the conduction and impurity electrons. We write this transformation as

$$\rho_{\rm out}(\mathbf{k}, \mathbf{k}') = \frac{\mathcal{T}_{\mathbf{k},\mathbf{k}'} \rho_{\rm in} \mathcal{T}^\dagger_{\mathbf{k},\mathbf{k}'}}{\mathrm{Tr}\left(\mathcal{T}_{\mathbf{k},\mathbf{k}'} \rho_{\rm in} \mathcal{T}^\dagger_{\mathbf{k},\mathbf{k}'}\right)}, \tag{3.2}$$

where $\rho_{\rm in/out}$ are the density operators for the spin state of the combined two-electron system, and: $\mathcal{T}_{\mathbf{k},\mathbf{k}'} = F_{\rm d}(\mathbf{k}, \mathbf{k}') + F_{\rm x}(\mathbf{k}, \mathbf{k}')\boldsymbol{\sigma}_c \cdot \boldsymbol{\sigma}_i$ [33]. Here $F_{\rm d}$ ($F_{\rm x}$) is the amplitude for un-exchanged (exchanged) conduction and impurity electron scattering [1]. Note that the spatial aspects of the problem only enter into the amplitudes. We will assume elastic scattering with the donor electron remaining bound, and no scattering of conduction electrons outside the 2DEG. Therefore the amplitudes can be parameterized by two parameters: $F_{\rm d/x} \equiv F_{\rm d/x}(\theta, k)$, the scattering angle within the 2DEG, θ; and the incoming momentum magnitude k (determined by the Fermi energy of the 2DEG electrons). These amplitudes are free parameters in our

[1] We assume throughout that the operating temperature is above the Kondo temperature of the device.

model and we explore a wide range of values for them in the simulations below. The direct and exchange scattering amplitudes are simply related to the more familiar singlet (f_s) and triplet (f_t) scattering amplitudes as:

$$F_d(\theta,k) = \frac{1}{4}\left(f_s + 3f_t\right)$$
$$F_x(\theta,k) = \frac{1}{4}\left(f_t - f_s\right) \tag{3.3}$$

Now, assume an initial state $\rho_{\rm in} = (p\,|\uparrow\rangle_c\langle\uparrow| + (1-p)\,|\downarrow\rangle_c\langle\downarrow|)\otimes\rho_i$, where the first term in the tensor product is the state of the conduction electron (the conduction band is assumed to be polarized to the degree $P_c^0 = 2p-1,\ 0\le p\le 1$), and the second term is the general state of the donor electron given above. After applying the scattering transformation and tracing out the conduction electron (because we have no access to its spin after the scattering event in this experimental scheme) we get a map that represents the transformation of the impurity electron state due to one scattering event:

$$\begin{aligned}\rho_i \to \rho_i'(\theta,k) = \frac{(1-p)}{\mathcal{N}}&\left[(F_{\rm d}+F_{\rm x}\sigma_z)\rho_i(F_{\rm d}^*+F_{\rm x}^*\sigma_z) + 4|F_x|^2\sigma_-\rho_i\sigma_+\right]\\ &+\frac{p}{\mathcal{N}}\left[(F_{\rm d}-F_{\rm x}\sigma_z)\rho_i(F_{\rm d}^*-F_{\rm x}^*\sigma_z) + 4|F_x|^2\sigma_+\rho_i\sigma_-\right]\end{aligned} \tag{3.4}$$

where $\mathcal{N}$ is a normalization constant to ensure $\mathrm{Tr}\,(\rho_i') = 1$, and we have not explicitly written the (θ,k) dependence of the scattering amplitudes for brevity.

In order for the measurement to be faithful, the diagonal elements of the impurity spin state (the population probabilities) must be preserved under the interaction – that is, the measurement interaction may induce dephasing (in the measurement basis), but no other decoherence. However, the terms proportional to $|F_{\rm x}|^2$ in Eq. (3.4) suggest that there will be population mixing. Using this equation, we can write the transformation of the diagonal elements (which are uncoupled from the off diagonal elements by the transformation Eq. (3.4)) as:

$$a \to a'(\theta,k) = \frac{(1-2P_c^0\Lambda\cos\chi + (2P_c^0-1)\Lambda^2)a + 2(1-P_c^0)\Lambda^2}{4P_c^0\Lambda(\Lambda-\cos\chi)a + 1 + 3\Lambda^2 + 2P_c^0\Lambda\cos\chi - 2P_c^0\Lambda^2}$$

where $\Lambda(\theta,k) \equiv |F_x(\theta,k)|/|F_d(\theta,k)|$, and $\chi(\theta,k) \equiv \arg F_x(\theta,k) - \arg F_d(\theta,k)$, and we have used the fact that $a+b=1$ to normalize the transformation.

We can iterate this recursion to simulate the effects of the repeated scattering events that contribute to the current. An appropriate quantification of measurement quality is the *measurement fidelity* [34]:

$$\mathcal{F}_n = 2|(\sqrt{a^{(n)}}\sqrt{a^{(0)}} + \sqrt{b^{(n)}}\sqrt{b^{(0)}})^2 - 0.5|, \tag{3.5}$$

where $a^{(n)}$ and $b^{(n)}$ are the diagonal elements of ρ_i after n scattering events. An ideal measurement has $\mathcal{F}_n = 1$, while $\mathcal{F}_n = 0$ indicates a measurement that yields no information

– i.e. no correlation between the original qubit state and the meter variables. Since the measurement should work for all initial states, we consider the worst-case measurement fidelity: $\mathcal{F}_n^w = \min_{a^{(0)},b^{(0)}} \mathcal{F}_n$.

In order to calculate this fidelity we need to determine how many scattering events will take place within the time required to do the measurement. Given the current state of the art, and factoring in improvements in 2DEG mobility [35] and conduction electron polarization, we estimate a shot-noise limited measurement time of $\tau_m \sim 10^{-3}$ s (this calculation is given below, in section 3.4). Within this time, there will be $\sim 10^9$ scattering events (see the Appendix for details on calculating the number of scattering events per second). Although we do not assume specific values of the scattering amplitudes, we find from iterating the above recursion for a broad range of values F_x/F_d that after $\sim 10^9$ scattering events $\mathcal{F}_n^w$ is $\ll 1$ for any non-zero value of the exchange amplitude $|F_x|$ and for any polarization, P_c^0. Figure 3.2 shows worst-case fidelity decay as a function of scattering amplitude parameters for various values of 2DEG polarization P_c^0. These simulations clearly show that the relaxation of a general electron spin state is rapid across virtually all reasonable parameter ranges. In fact, for realistic 2DEG polarization values $\mathcal{F}_n^w$ typically drops to near zero already after $\sim 10^3 - 10^4$ scattering events. Thus the measurement induced population mixing time is $T_{\rm mix} \sim 1 - 10$ns, which is drastically smaller than τ_m. These simulations thus show conclusively that whatever the precise values of the scattering amplitudes, under realistic experimental conditions the electron spin relaxation induced by the scattering interaction makes the 2DEG current an ineffective measurement of the electron spin state. This makes it impossible to faithfully map the electron spin state onto the meter variable, and hence impossible to perform a single electron spin state measurement using the 2DEG current directly.

For completeness we note that since we only have access to the total 2DEG current and no angle-resolving detectors, the actual impurity electron density matrix must also involve an average over θ and k over the 2DEG Fermi surface in Eq. (3.4). However, as we have shown that the direct measurement will not work for any value of θ and k, the averaged dynamics will only result in a worse performance analysis.

However, as we will now show, it is possible to make use of the nuclear spin degree of freedom in order to utilize EDMR for projective and QND measurement of single spin states. The key is that the state of the nuclear spin affects the Zeeman splitting of the electron spin (and thus its resonant frequency) via the mutual hyperfine coupling. Therefore our strategy is to transfer the qubit state from the electron to the nucleus and then to perform an EDMR readout.

3.3.2 Nuclear spin mediated electron spin state measurement

The low-energy, low-temperature Hamiltonian describing the electron and nuclear spins of a phosphorous dopant in a static magnetic field, $\mathbf{B} = B\hat{\mathbf{z}}$ is

$$H = \frac{1}{2}\left[g_e\mu_B B\sigma_z^e - g_n\mu_n B\sigma_z^n\right] + A\boldsymbol{\sigma}^e \cdot \boldsymbol{\sigma}^n \tag{3.6}$$

where μ_B and μ_n are the Bohr and nuclear magnetons, g_e (g_n) is the electron (nuclear) g-factor, and A characterizes the strength of the hyperfine interaction between the two spins [17] (we set $\hbar = 1$ throughout the paper). For moderate and large values of B, the σ_z terms dominate and we can make the *secular approximation*, to arrive at: $H \approx 1/2\left[g_e\mu_B B\sigma_z^e - g_n\mu_n B\sigma_z^n\right] + A\sigma_z^e\sigma_z^n$. The energy levels and eigenstates of this Hamiltonian are shown in Fig. 3.3. Note that we have ignored the coupling of both spins to uncontrolled degrees of freedom such as paramagnetic defects and phonons (coupling to lattice spins can be mitigated by the use of a ^{28}Si substrate). These environmental couplings will contribute to decoherence of the nuclear and electron spin states (e.g. [36]), and we will simply assume that this results in some effective relaxation and dephasing of the electron and nuclear spins.

We see that the resonance frequency (Zeeman energy) of the electron is a function of the nuclear spin state. Therefore, our strategy will be to transfer the qubit state from the electron to the nucleus and then use EDMR to measure the nuclear spin. This is in effect a spin-to-resonance-to-charge conversion measurement.

To perform the state transfer, we appeal to the qubit SWAP gate: $\mathrm{SWAP}[\rho_e\otimes\tau_n]\mathrm{SWAP}^\dagger = \tau_e \otimes \rho_n$. SWAP can be decomposed into the sequence of three controlled-not (CNOT) gates [15] $\mathrm{SWAP} = \mathrm{CNOT}_n\mathrm{CNOT}_e\mathrm{CNOT}_n$, where the subscript indicates which of the two qubits is acting as the control. However, the complete exchange of electron-nuclear states is unnecessary, since the spin state of the impurity electron is lost to the environment by the application of resonant pulses and elastic scattering with conduction electrons in the 2DEG. Therefore, the final operator in the sequence can be neglected since it only alters the state of the electron. This leads to the definition of the electron-to-nucleus transfer gate, $\mathrm{TRANS}_e = \mathrm{CNOT}_e\mathrm{CNOT}_n$. To apply these CNOT gates we use resonant pulses: CNOT_e interchanges the states $|\uparrow\rangle_e|\Uparrow\rangle_n$ and $|\uparrow\rangle_e|\Downarrow\rangle_n$ and so can be implemented by application of a resonant π-pulse at frequency ω_n (see Fig. 3.3), an RF transition; similarly, CNOT_n interchanges $|\uparrow\rangle_e|\Uparrow\rangle_n$ and $|\downarrow\rangle_e|\Uparrow\rangle_n$ and is implemented by a resonant π-pulse at ω_e, a microwave transition. Each of these transitions is dipole-allowed, ensuring that gate times are sufficiently fast. The ability to apply pulses faster than relevant decoherence times is required for successful implementation of the state transfer. We note that after this work was completed, a state swap scheme very similar to the state transfer scheme outlined above was successfully performed in experiments on bulk-doped Si:P samples [37].

Suppose the electron is in an initial (pure) state, $|\psi\rangle_e = \alpha|\uparrow\rangle_e + \beta|\downarrow\rangle_e$, while the nucleus is in a general mixed state,

$$\tau_n = \begin{pmatrix} u & w \\ w^* & v \end{pmatrix}. \tag{3.7}$$

After performing the state transfer on the combined state and tracing over the electron degrees of freedom (because it is lost to the environment), we are left with the reduced density matrix describing the nucleus,

$$\mathrm{Tr}_e\left(\mathrm{TRANS}_e\left[\rho_e\otimes\tau_n\right]\mathrm{TRANS}_e^\dagger\right) = \begin{pmatrix} |\alpha|^2 & \alpha\beta^*(w+w^*) \\ \alpha^*\beta(w+w^*) & |\beta|^2 \end{pmatrix}$$

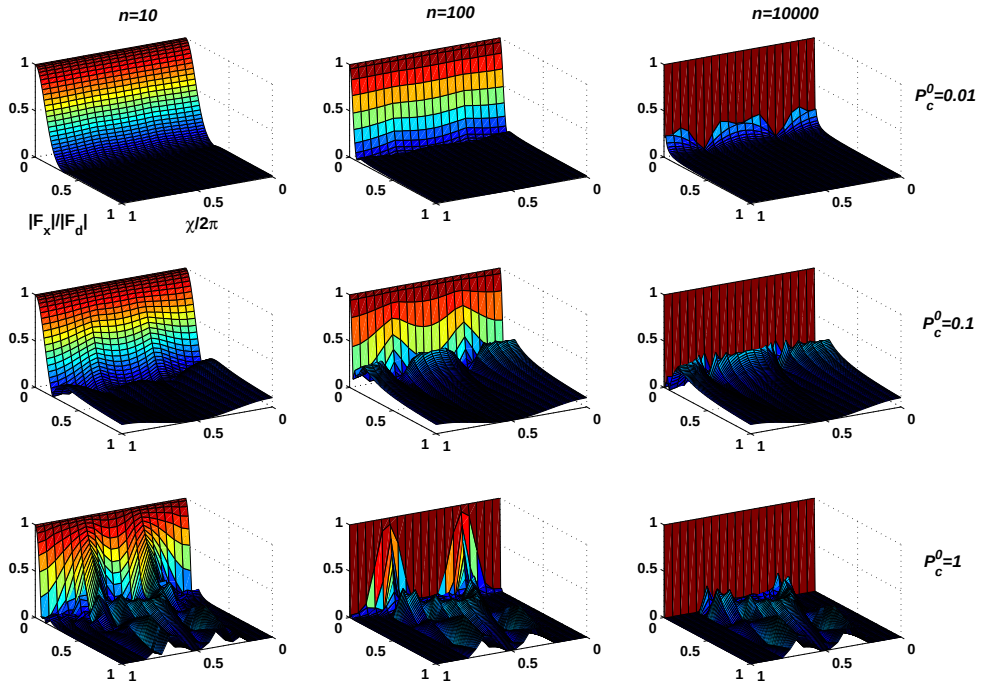


Figure 3.2: Evolution of worst-case measurement fidelity, $\mathcal{F}_n^w$, during 2DEG scattering dynamics as a function of the number of scattering events n, for a range of values of the ratio of direct and exchange scattering amplitudes and of 2DEG equilibrium polarization. The independent (base plane) axes on the plots parametrize the complex scattering amplitude ratio $F_x/F_d \equiv |F_x|/|F_d|e^{i\chi}$: one axis is the magnitude, $\Lambda \equiv |F_x|/|F_d|$ (shown for $0 < |F_x|/|F_d| < 1$; the plots are restricted to this range because the worst-case fidelity is negligibly small outside it), and the other is the phase, χ (shown for $0 < \chi/2\pi < 1$). The number of scattering events, n, varies across the columns, with values $n = 10, 10^2$ and 10^4 shown here. The 2DEG equilibrium polarization, P_c^0, varies across the rows, with values $P_c^0 = 0.01, 0.1$ and 1 shown here. For $P_c^0 \geq 0.1$, we see that there are fairly large regions in the F_x/F_d parameter space for which the worst-case measurement fidelity is non-zero: however, (i) $\mathcal{F}_n^w$ still decays rapidly with number of scattering events, and is rarely > 0.9 (the fidelities desirable for high-quality measurement), and (ii) $\mathcal{F}_n^w$ is highly sensitive to the precise value of F_x/F_d and P_c^0 in these regions.

Because of the hyperfine coupling (Fig. 3.3), electron resonance will occur at the lower frequency with probability $|\alpha|^2$ and at the higher frequency with probability $|\beta|^2$. The electrical detection of this shift from the free electron resonance frequency by EDMR constitutes a single-shot, projective measurement in the σ_z basis of the original electron state (and therefore, qubit state) with the correct statistics.

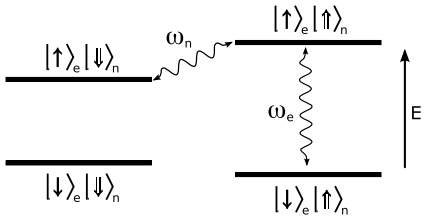


Figure 3.3: Four-level system of electron-nuclear spin degrees of freedom. The energy eigenstates in the secular approximation are the eigenstates of σ_z^e and σ_z^n. The transitions indicated by arrows are required for the state transfer described in the text.

In detail, the single qubit spin readout can proceed as follows. Following state transfer to the nuclear spin, one of the two hyperfine split electron spin resonance lines that corresponds to a given nuclear spin projection is addressed by dialing in the corresponding microwave frequency for resonant excitation of electron spin transitions. At the same time, the transistor is turned on and the channel current is monitored. Now, assume that the magnetic fields have been tuned to address the $|\uparrow\rangle_n$ nuclear state projection. Then with probability $|\alpha|^2$ the transistor current will differ from the off-resonant current value and with probability $|\beta|^2$ it will be just equal to the off-resonant channel current. In either case, monitoring the current at one hyperfine resonance for the τ_m measurement duration constitutes a readout of the nuclear spin. And due to the prior state transfer, it effectively measures the spin state of the original donor electron spin.

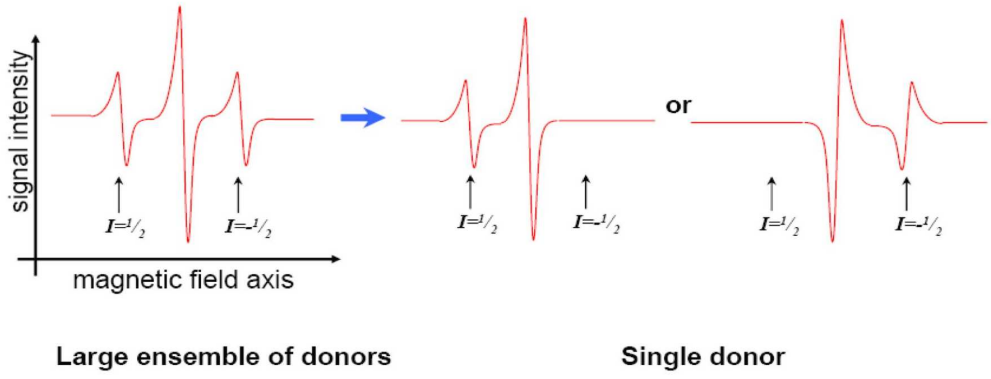


Figure 3.4: Illustration of single spin readout. In experiments with large ensembles of donor spin qubits, lines from all nuclear spin projections are present in EDMR measurements (left). In measurements with single donors (right), only single lines are present for measurement times shorter then the nuclear spin relaxation time. Monitoring the current at a given resonant field measures the spin state of the donor nucleus with the correct statistics.

3.4 Measurement sensitivity and measurement-induced decoherence

Two critical practical issues need to be addressed for realization of this protocol for measurement of the donor electron spin quantum state. These are: i) the sensitivity of the EDMR measurement needs to be sufficiently high to allow single donor spins to be detected, and ii) the measurement time τ_m must be small compared to the lifetime of the nuclear spin.

We first address the issue of the sensitivity of the differential EDMR current in the limit of single donor scattering. As detailed above, the spin state of the donor electron is continually changing due to the scattering interaction, and hence is time-dependent. However, ignoring the transient, we can approximate it with a time independent value given by the steady state solution of the recursion relation, Eq. (3.4). This approximation can be thought of as taking the equilibrium spin value, where the "spin temperature" of the impurity has equilibrated with that of conduction electrons via the scattering interaction. The explicit simulations of the scattering recursion Eq. (3.5) shown in Section 3.3.1 indicate that this equilibration happens within 10^4 scattering events for all possible values of scattering amplitudes. Thus the time scale for this equilibration is ~ 10ns, much faster than the observable times scales of the measurement, justifying our use of the steady-state solution of the recursion (see the Appendix for details on calculating the number of scattering events per second). Solving for the steady-state ($\rho_i^{(n)} = \rho_i^{(n-1)} \equiv \rho_i^{ss}$), gives us a time-independent, non-resonant single donor "polarization" equal to

$$\langle \sigma_z \rangle_i^{\rm ss} \equiv {\rm Tr}\left(\sigma_z \rho_i^{ss}\right) = \frac{\Lambda - \sqrt{(P_c^0)^2 \cos^2\chi + \Lambda^2(1-(P_c^0)^2)}}{P_c^0(\Lambda - \cos\chi)} \tag{3.8}$$

It should be noted that this expression for the steady state "polarization" is not valid when $P_c^0 = 0$ or $\Lambda = 0$, but neither of these limits is relevant to spin measurement. Now, we can follow the analysis of Ref. [25], using donor "polarization" $\langle \sigma_z \rangle_i^{\rm ss}$, to estimate the on-resonant (I) and off-resonant (I_0) current differential (normalized) as:

$$\frac{\Delta I}{I_0} \equiv \frac{I - I_0}{I_0} \approx -\alpha' \; s \; \langle \sigma_z \rangle_i^{\rm ss} \, P_c^0 \frac{1/\tau_n}{1/\tau_t}. \tag{3.9}$$

Here $\alpha' \equiv \langle \Sigma_s - \Sigma_t \rangle |_{z=z_i} / \langle \Sigma_s + 3\Sigma_t \rangle |_{z=z_i}$, Σ_s and Σ_t are singlet and triplet scattering cross sections, respectively, and $\langle \cdot \rangle |_{z=z_i}$ denotes an average over the scattering region with the donor location in z (see Fig. 3.1) held fixed [38]. $s = 1 - (1 - s_i)(1 - s_c)$, and s_i and s_c (both between 0 and 1) are saturation parameters which characterize how much of the microwave power is absorbed by the impurity and conduction electrons, respectively [25]. s_i is a function of the broadening at the *single* donor electron resonance frequency: if we work in a regime where this broadening is minimal (as required to perform the quantum state transfer described above), $s_i \approx 1$ and thus $s \approx 1$. The final term in Eq. (3.9) represents the ratio between impurity scattering ($1/\tau_n$) and total scattering ($1/\tau_t$) rates. We assume

$1/\tau_t = 1/\tau_0 + 1/\tau_n$, where $1/\tau_0$ is the scattering rate due to all other processes (such as surface roughness scattering and Coulomb scattering by charged defects).

To estimate the expected magnitude of this current differential, we begin by considering present state of the art 2DEG mediated EDMR experiments where this current differential is $\sim 10^{-7}$ (with T $\sim 5K, B \sim 0.3T$, a 2DEG channel area of $160 \times 20\mu\text{m}^2$, probe current 1μA, and a donor density of $2 \times 10^{11}\text{donors/cm}^2$) [26]. We assume that α' will be similar for the single donor device as in current experiments. Then in scaling down to a single donor, the first aspect to consider is the scattering rate ratio: $\varrho \equiv \frac{1/\tau_n}{1/\tau_t}$. To first order this ratio can be kept constant if we scale the 2DEG area concomitantly with the donor number. From the channel area and density of current experiments, we extrapolate that a 2DEG area of $\sim 30 \times 30\text{nm}^2$ – well within the realm of current technology [39] – would keep ϱ unchanged. Optimization of donor depth might relax this size requirement [38]. A higher order analysis would require detailed investigation of the device specific interface and intrinsic contributions to the other scattering processes, and hence to the channel mobility and τ_0. Related to this concern, the mobility of the 2DEG channel can be improved – e.g., by using hydrogen passivation to mitigate surface roughness at the oxide interface [35] – to increase ϱ. We conservatively estimate a factor of 10 increase in $\Delta I/I_0$ from such improvements. The saturation parameter $s \sim 1$ for large enough microwave powers in the recent measurements [26] and so does not present an area for improvement. Finally, an avenue for significant improvement in signal is to increase the conduction electron polarization, P_c^0, which is currently $\sim 0.1 - 1\%$. This polarization is roughly proportional to the applied static magnetic field, and therefore a factor of 10 improvement is possible by operating at $B = 3T$. Additionally, spin injection techniques can be employed to achieve $P_c^0 > 10\%$ (e.g. [40, 41]), resulting in a 100-fold improvement in $\Delta I/I_0$. Hence, by improvements in device scaling and channel mobility, and by incorporating spin injection, we estimate a realistic, improved current differential of $\Delta I/I_0 \sim 10^{-4}$. Given this $\Delta I/I_0$ and a probe current of $I_0 \sim 1\mu A$, to achieve an signal-to-noise (SNR) of 10 through shot-noise limited detection we require a τ_m satisfying:

$$\left(\frac{\Delta I}{I_0}\right)\frac{I_0\tau_m}{e} > 10\sqrt{\frac{I_0\tau_m}{e}} \tag{3.10}$$

where the left hand side is the signal, the right hand side is the accumulated shot-noise multiplied by the SNR, and e is the fundamental unit of electric charge. Solving this yields a measurement integration time of $\tau_m \sim 10^{-3}$ s.

In order to complete the measurement analysis we need to address the second issue identified above and confirm that the state of the nuclear spin does not flip within the measurement time – i.e. the measurement time τ_m has to be shorter then the nuclear spin flip time T_1. Once the 2DEG current is switched on (the 2DEG current is off during the electron-nucleus state transfer) the dynamics of the donor electron due to scattering and microwave driving will contribute to the decoherence of the nuclear spin. Donor nuclear spin relaxation is not well characterized under these conditions but we expect that in large magnetic fields the donor electron dynamics contributes primarily only to dephasing of the

nuclear state. This can be made precise by performing perturbation theory on Eq. (3.6) in the parameter A/Δ, where $\Delta \equiv \omega_e - \omega_n = B(g_e\mu_B - g_n\mu_n)$. In the *detuned* regime where $A/\Delta \ll 1$, the effective Hamiltonian describing the coupled systems is

$$H \approx H_{\text{eff}} = \frac{1}{2}\omega_e\sigma_z^e - \frac{1}{2}\omega_n\sigma_z^n + A\sigma_z^e\sigma_z^n + \frac{A^2}{\Delta}(\sigma_z^e - \sigma_z^n) \tag{3.11}$$

This effective Hamiltonian is of course also the justification for the secular approximation made earlier (see below Eq. (3.6)). Therefore we see that to first order in A/Δ the donor electron can only dephase the nuclear spin and that direct contributions to nuclear spin T_1 through the hyperfine interaction are small. Secondary mechanisms such as phonon-assisted cross relaxation can, in principle, also contribute to the nuclear T_1 during electron driving. However, these contributions were shown to be very small by Feher and Gere who demonstrated that the electron-nuclear cross relaxation time, T_x, under electron driving conditions is on the order of hours [42]. Given this extremely long cross-relaxation time and the equivalently long nuclear T_1 in a static electron environment at low temperatures [42, 43], we conclude that the nuclear T_1 in the presence of electron driving will be comfortably larger than $\tau_m \sim 10^{-3}$ s. Indeed, this has very recently been confirmed by explicit measurements of nuclear T_1 under the conditions of electron driving [44]. This analysis implies that once the measurement collapses onto a nuclear basis state, the nuclear spin state does indeed effectively remain there and therefore the EDMR measurement satisfies the QND requirement on the qubit state.

3.5 Conclusion

By utilizing resonant pulse gates and 2DEG-mediated-EDMR readout, we have proposed a realistic scheme for measuring the spin state of a single donor electron in silicon. By making use of the hyperfine coupled donor nuclear spin, the readout scheme provides a single shot measurement that is both projective and QND. The QND aspect also makes this technique an effective method for initializing the state of the nuclear spin.

We have analyzed the measurement procedure, the factors which influence the signal-to-noise ratio, and the experimental apparatus to arrive at realistic modifications/improvements that can be made to current 2DEG-based EDMR apparati [26] so that a single nuclear spin can be measured. One concern is that at the required transistor size of $\sim 30 \times 30\text{nm}^2$, the MOSFET device will no longer act as a 2DEG but rather more like a quantum dot. However, while the proposed MOSFET is small, there are no physical tunnel barriers between the transistor island and the source and drain leads (no intentional confinement). Important Coulomb blockade effects have indeed been observed in such MOSFETs, but only in a biasing regime with very low source-drain biases [45, 39, 46] and in sub-threshold currents at very low gate biases [47]. In these studies, confinement was found to be induced by impurity potentials [45, 46]. It has also been shown that one can easily tune in and out of the Coulomb blockade regime by suitable gate and source-drain biasing [45] (and see also Ref.

[46] for related results with tunable tunnel barriers). In the experiments we envision, which are closely related to and inspired by recent demonstrations of spin dependent transport in micron-scale devices [25, 26], the source-drain as well as the gate bias can be tuned over a large range of voltages, from the very low values needed to study interesting and important Coulomb blockade effects in a low current regime (< 50 nA) to higher bias values desirable for electrical detection of magnetic resonance through scattering of conduction electrons off neutral donors [25, 26] with much higher channel currents (~ 1 μA). Such MOSFET devices can be operated as 2DEGs and away from the quantum dot regime, by imposing large enough source-drain and gate biases.

Finally, we note that the fact that the measurement is facilitated by the nucleus of the donor atom intimates a hybrid donor qubit where quantum operations are carried out on the electron spin and the state is transfered to the nucleus for measurement and storage (advantageous due to the longer relaxation times). Although the above analysis was done with the example of a phosphorous donor, it applies equally well to other donors, such as antimony [20, 26], and some paramagnetic centers [48]. One merely has to isolate two (dipole-transition allowed) nuclear spin levels to serve as qubit basis states and transfer the electron state to these nuclear states with resonant pulses as outlined here.

3.6 Appendix

Here we detail the procedure used in the main text for calculating the amount of time taken for n scattering events. We assume a MOSFET channel of width $W = 30$nm probed with a current of 1μA. This means that there are $I/e \approx 1.6 \times 10^{13}$ electrons per second crossing the channel. However, to gain a more accurate estimate of the number of electrons interrogating the donor electron per second we scale this total number by a ratio of the scattering length to the width of the channel: $\sqrt{\sigma}/W$. A crude estimate of the scattering cross section, σ, can be obtained using the singlet and triplet scattering lengths given in Ref. [28]: $a_s = 6.167\mathring{A}$ and $a_t = 2.33\mathring{A}$. These scattering lengths are for three dimensional electron-hydrogen scattering in bulk-doped systems, however we consider them sufficient for an order of magnitude estimate of the number of interrogating electrons per second. In terms of these scattering lengths, the scattering cross section for the 2DEG interacting with an ensemble of donors is [25]:

$$\sigma = 2\pi \left[(a_s^2 + 3a_t^2) - (a_s^2 - a_t^2) P_c^0 P_i^0 \right] \tag{3.12}$$

where P_c^0 and P_i^0 are the conduction band and impurity polarization. Since $P_c^0 \ll 1$ we will approximate the above expression as:

$$\sigma \approx 2\pi (a_s^2 + 3a_t^2) = 341.3 \ \mathring{A}^2 \tag{3.13}$$

For a single donor, Eq. (3.12) should be modified to take into account the time-dependent "polarization" of the single donor spin (see Sec. 3.4). However, since this quantity drops out

in the final approximate expression for the cross section, Eq. (3.13), we will not be concerned with this modification.

Using this estimate of the average scattering cross section, the number of electrons interrogating the donor per second is given by

$$n_e \approx 1.6 \times 10^{13} \cdot \frac{18.5 \times 10^{-10}}{30 \times 10^{-9}} \approx 1 \times 10^{12} \tag{3.14}$$

Therefore we estimate that the time taken for n scattering events is $n \times 10^{-12}$ seconds.

Chapter 4

Quantum multi-parameter estimation

4.1 Introduction

The tremendous allure of quantum information processing has fueled recent progress in the experimental and theoretical understanding of physical systems operating in regimes where classical physics fails to hold. The precise control and characterization of physical systems demanded by quantum information processors, e.g. for performing high-fidelity quantum gates, has extended our mastery of optical, gas phase, and condensed phase physical systems.

Typically, as the precision to which one must characterize a physical system increases, the sophistication of the techniques used to study the system must also increase. Recent *tour-de-force* experiments have fully characterized quantum systems of small dimension by performing exhaustive process tomography (e.g. [49, 50]). Such exhaustive tomography requires resources that scale exponentially with the dimension of the system being studied and so is infeasible for systems much larger than those already characterized in this manner. Consequently, many techniques for approximate characterization of large dimensional quantum systems have been formulated in recent years [51, 52, 53, 54].

In many situations one is not completely ignorant about the dynamical system being studied. An experimentalist may have partial knowledge of the system through information from system preparation or prior characterization studies. In such cases the system characterization often becomes a problem of parameter estimation, and an important question arises: how does one design an experiment to identify the unknown parameters of the dynamical process most efficiently, or even optimally with respect to some metric? Experiment design for optimal parameter estimation in quantum systems is a natural extension of the equivalent classical design problem; one typically attempts to rapidly reduce the variance in the unknown parameters by performing as few experiments as possible. The goal of experiment design is to identify the input states to probe the dynamical process with and the measurements to perform on the outputs, so that the variance in the unknown parameters can be decreased as quickly as possible (with the number of experiments performed). Analytical and numerical methods for optimal experiment design have been widely explored for one parameter quantum processes (e.g. [55, 56, 57, 58, 59, 60]), but very few analytic optimality results exist for the multi-parameter case (for exceptions, see Refs. [61, 62]), in part because of difficulties in optimizing over noncompatibile (noncommuting) quantum observables [63]. Numerical approaches to optimal experiment design for quantum tomography (when all parameters of the quantum process or state are unknown) and Hamiltonian parameter estimation using convex optimization were first proposed in Ref. [64], and applied to experiments in Refs. [65, 66]. The method follows from the optimal experiment design approach described in Ch. 7.5 of Ref. [67]. Recently, a similar numerical approach to multi-parameter quantum process estimation, using convex optimization, was formulated [68] and we shall further refer to it below. Experimentally motivated techniques for multi-parameter estimation have also been proposed [69, 70], but the optimality and asymptotic performance of these are unknown.

In this chapter we examine this problem of optimal multi-parameter estimation for quantum processes when there are constraints on the possible input probe states and on

the possible measurements. The constraints on the input states and spin measurements result from experimental limitations on the types of input states (measurements) that can be realistically prepared (performed). We consider a concrete example motivated by an experimental platform for quantum information processing: donors in semiconductors with electrical control and measurement [17, 71, 72]. We solve the problem of precisely identifying the coupling between two electron spin qubits that interact through a combination of exchange and dipole-dipole interactions by a preparation of input states and measurement of electron spins after a suitable interaction period. Note that precise knowledge of the qubit-qubit interaction is crucial for the execution of two-qubit gates which typically work by transforming this interaction into the desired gate by single qubit manipulation pulses [73]. We apply a recently re-formulated numerical approach to optimal experiment design for multi-parameter quantum estimation [64] which also incorporates available experimental configurations into a convex optimization [68]. This formulation allows us to efficiently identify the optimal characterization experiment and estimate the number of experimental runs necessary to achieve a desired accuracy in the estimated parameters.

In section 4.2 we provide background on the quantum parameter estimation problem and recap the formulation of the multi-parameter constrained estimation problem as a convex optimization from Refs. [68, 64]. Section 4.3 presents the experiment design framework in full generality and sketches an algorithm for optimally estimating a set of unknown parameters of a quantum process. Section 4.4 introduces the example we explicitly solve: two coupled electron spin qubits. We summarize the experimental capabilities of this implementation of quantum information processing and give a detailed description of the coupling dynamics. Then in section 4.5 we apply our experiment design framework to formulate the optimal estimation scheme for identifying the unknown parameters under the given constraints.

4.2 Parameter estimation

Suppose a sequence of data that is independent and identically distributed (*iid*) is drawn from a distribution that is parametrized by one or several unknown quantities. For instance, the distribution could be Gaussian with unknown mean and variance. The parameter estimation problem is to estimate the value(s) of the unknown quantities from the sample data.

A central task of parameter estimation is the construction of an estimator, $T_{\boldsymbol{\theta}}(\boldsymbol{X})$, which maps the sampled data, $\boldsymbol{X}$, to an estimate, $\hat{\boldsymbol{\theta}}$, of the parameters. In what follows, we will assume the use of unbiased estimators,

$$\langle \hat{\boldsymbol{\theta}} \rangle \equiv \langle T_{\boldsymbol{\theta}}(\boldsymbol{X}) \rangle = \boldsymbol{\theta}.$$

The generalization to biased estimators is well known, but needlessly complicates our discussion.

However, some probability distributions are more easily estimable than others. Take for example a Dirac-delta distribution centered at x_0, so that the probability density function is $p_{x_0}(x) = \delta(x - x_0)$. The parameter to be estimated in this case is x_0 and only a single

measurement is required. On the other hand, accurately estimating the mean of a large-variance Gaussian distribution requires many samples. Estimability of a parameter is thus a property of the probability distribution and is independent of the estimator used. This idea is encapsulated by the Cramér-Rao bound, which places a lower limit on the variance of any single-parameter estimator [74],

$$\operatorname{var} T_\theta(\boldsymbol{X}) \geq \frac{1}{NF(\theta)}.$$

Here, N is the number of samples and $F(\theta)$ is the Fisher Information, defined as a functional of the probability distribution,

$$F(\theta) \equiv \left\langle \left(\frac{d}{d\theta} \ln p_\theta(x) \right)^2 \right\rangle$$

where we use the shorthand $p_\theta(x)$ for the conditional distribution $p(x|\theta)$ and $\langle f(x) \rangle = \int f(x)\, p(x|\theta)\, dx$ (or $\sum_i f(x_i) p(x_i|\theta)$ if the probability distribution is discrete) is the expectation value of $f(x)$. Note that the Fisher information is a function of the true value (not the estimate) of the parameter. Intuitively it represents the amount of "information" about the parameter in the conditional probability distribution for the data.

In the multi-parameter case the generalized Cramér-Rao inequality bounds the covariance matrix of the (now vector-valued) estimator [74],

$$\operatorname{cov}_\theta T_{\boldsymbol{\theta}}(\boldsymbol{X}) \geq \frac{\mathcal{I}(\boldsymbol{\theta})^{-1}}{N} \tag{4.1}$$

where $\mathcal{I}(\boldsymbol{\theta})$ is the Fisher information matrix:

$$\mathcal{I}(\boldsymbol{\theta}) \equiv \left\langle \left(\boldsymbol{\nabla}_{\boldsymbol{\theta}} \ln p_{\boldsymbol{\theta}}(x)\right) \left(\boldsymbol{\nabla}_{\boldsymbol{\theta}} \ln p_{\boldsymbol{\theta}}(x)\right)^{\mathsf{T}} \right\rangle$$

We have used the notation $\boldsymbol{\nabla}_{\boldsymbol{\theta}} f(\boldsymbol{\theta}) = \left(\frac{d}{d\theta_1} f, \frac{d}{d\theta_2} f, \ldots, \frac{d}{d\theta_n} f \right)^{\mathsf{T}}$.

The Cramér-Rao inequality provides a bound on how well we can do when estimating the parameter(s) from the data. A derivation of this bound in the single parameter case is given in Appendix 4.7.1 of this chapter. While the actual variance in the parameter estimate is dependent on the particular estimator used, there exist estimators that are known to saturate this bound asymptotically (in the limit of large N) [74]. An example, that we shall employ below, is the the maximum likelihood estimator (MLE). The MLE is defined as

$$T^{ML}_{\boldsymbol{\theta}}(\boldsymbol{X}) = \arg\max_{\boldsymbol{\theta}} p_{\boldsymbol{\theta}}(\boldsymbol{X}) \tag{4.2}$$

4.3 Optimal experiment design for quantum parameter estimation

Up to this point we have discussed the mathematics of parameter estimation. The physics of a particular problem becomes important only in calculating the probability distributions (and their derivatives). Quantum mechanics provides the tools with which these distributions can be obtained.

We begin by defining an experiment, $\mathcal{E}$, as a choice of the initial state, ρ_0; evolution time, t; and a positive operator valued measure (POVM), $\mathbf{M} = \{M_i\}$ [75]. The POVM, also known as a generalized measurement, satisfies $\sum_i M_i = 1$ and $M_i \geq 0$, $0 \leq i \leq n_{out}$. Each M_i corresponds to a possible outcome from applying the measurement $\mathbf{M}$. Through the application of the Born rule, each experiment determines a parametrized family of discrete probability distributions,

$$\begin{aligned} p^{\mathcal{E}}_{\boldsymbol{\theta}}(i) &= \mathrm{Tr}\left(M_i\left(U_{\boldsymbol{\theta}}(t)\rho_0 U_{\boldsymbol{\theta}}(t)^\dagger\right)\right) \\ &= \mathrm{Tr}\left(M_i \rho_{\boldsymbol{\theta}}(t)\right) \end{aligned} \tag{4.3}$$

Here $U_{\boldsymbol{\theta}}(t) = \mathcal{T}\exp\left(-i\int_0^t H_{\boldsymbol{\theta}}(t')dt'/\hbar\right)$ is the unitary time evolution operator and $H_{\boldsymbol{\theta}}(t)$ is the Hamiltonian whose parameters, $\boldsymbol{\theta}$, we wish to estimate. $p^{\mathcal{E}}_{\boldsymbol{\theta}}(i)$ is the probability, given a fixed experiment $\mathcal{E} = \{\rho_0, \mathbf{M}, t\}$ and assuming the parameter takes the value $\boldsymbol{\theta}$, that we get the measurement result i. From this probability distribution one can calculate the Fisher information matrix associated to this experiment:

$$\mathcal{I}^{\mathcal{E}}(\boldsymbol{\theta}) = \sum_i \frac{\left(\boldsymbol{\nabla}_{\boldsymbol{\theta}} p^{\mathcal{E}}_{\boldsymbol{\theta}}(i)\right)\left(\boldsymbol{\nabla}_{\boldsymbol{\theta}} p^{\mathcal{E}}_{\boldsymbol{\theta}}(i)\right)^{\mathsf{T}}}{p^{\mathcal{E}}_{\boldsymbol{\theta}}(i)} \tag{4.4}$$

Inserting this quantity into Eq. (4.1) gives a lower bound on the variance of our estimate. We will restrict our discussion to closed-system (i.e. Hamiltonian) evolution and, for the sake of clarity, to finite-dimensional Hilbert spaces. The generalization to non-unitary processes is straightforward; the most difficult step being the calculation of $\boldsymbol{\nabla}_{\boldsymbol{\theta}} p$, which must often be performed numerically.

It is often the case that an experimentalist has access to a number of different initial conditions and measurement bases. We would like to answer the question: which of these initial conditions and measurements should the experimentalist use in order to best estimate the unknown parameters in the quantum process? In order words, we would like to design our experiment so that we sample the quantum process in a manner that produces the most information about the unknown parameters. Formally. suppose we are given a menu of possible experiments, $\{\mathcal{E}\}$ and each time we sample our quantum process, an experiment, $\mathcal{E} = \{\rho_0^{\mathcal{E}}, \mathbf{M}^{\mathcal{E}}, t^{\mathcal{E}}\}$, is chosen with probability $\lambda_{\mathcal{E}}$ (so $\sum_{\mathcal{E}} \lambda_{\mathcal{E}} = 1$). The result of that measurement is governed by the probability distribution, $p^{\mathcal{E}}_{\boldsymbol{\theta}}(\cdot)$, and so is associated with its own Fisher matrix – i.e. Eq. (4.4). The probability of any particular measurement result must now

be scaled by the probability that a particular experiment will be performed. So the Fisher matrix for the combined experimental scheme defined by $\mathcal{E}$ and $\lambda_{\mathcal{E}}$ is

$$\begin{aligned}\mathcal{I}(\boldsymbol{\theta}) &= \sum_{\mathcal{E},i} \frac{\left(\boldsymbol{\nabla}_{\boldsymbol{\theta}} \lambda_{\mathcal{E}} p^{\mathcal{E}}_{\boldsymbol{\theta}}(i)\right)\left(\boldsymbol{\nabla}_{\boldsymbol{\theta}} \lambda_{\mathcal{E}} p^{\mathcal{E}}_{\boldsymbol{\theta}}(i)\right)^{\mathsf{T}}}{\lambda_{\mathcal{E}} p^{\mathcal{E}}_{\boldsymbol{\theta}}(i)} \\ &= \sum_{\mathcal{E}} \lambda_{\mathcal{E}} \mathcal{I}^{\mathcal{E}}(\boldsymbol{\theta})\end{aligned}$$

It is now natural to ask, given a menu of experiments, what choice of $\boldsymbol{\lambda}$ minimizes $\mathrm{Tr}\,(\mathcal{I}^{-1}(\boldsymbol{\theta})) = \mathrm{Tr}\,\left(\sum_{\mathcal{E}} \lambda_{\mathcal{E}} \mathcal{I}^{\mathcal{E}}(\boldsymbol{\theta})\right)^{-1}$ and thereby provides the best upper bound on the average of the estimate variance across all parameters through Eq. (4.1). This optimization problem, known as *A-optimal experiment design*, can be written as [67]:

$$\begin{aligned}\text{minimize} \quad & \mathrm{Tr}\left[\sum_{\mathcal{E}} \lambda_{\mathcal{E}} \mathcal{I}^{\mathcal{E}}(\boldsymbol{\theta})\right]^{-1} \\ \text{subject to} \quad & \lambda_i \geq 0, \quad \sum_i \lambda_i = 1\end{aligned}$$

Note that the optimization parameter is the vector of probabilities $\lambda_{\mathcal{E}}$. This optimization is difficult because the cost function is not linear or convex in the optimization parameter. However, through the use of the Schur complement (see Appendix 4.7.2), it can be reformulated as the convex optimization problem:

$$\begin{aligned}\text{minimize} \quad & \mathrm{Tr}\, Q \\ \text{subject to} \quad & \begin{pmatrix} Q & I \\ I & F \end{pmatrix} \succeq 0, \; F = \sum_{\mathcal{E}} \lambda_{\mathcal{E}} \mathcal{I}^{\mathcal{E}}(\boldsymbol{\theta}), \\ & \lambda_i \geq 0, \quad \sum_i \lambda_i = 1. \end{aligned} \tag{4.5}$$

To make this problem tractable, the menu of experiments can be chosen as a discretization of the continuous space of all possible experiments. The exact nature of the discretization must be determined for each problem individually, but, in general, a finer grained discretization produces a larger optimization problem. A coarser graining will result in a smaller optimization problem, but one whose solution will more poorly approximate the true achievable lower bound on the variance given by Eq. (4.1). In practice, one will discretize the space of initial states, the space of POVMs, and time. Given n_ρ initial states, n_M POVM's, and n_t times, we have $\tilde{n} = n_\rho \times n_M \times n_t$ experiments, and thus an optimization vector, $\lambda_{\mathcal{E}}$, of length $\tilde{n}$.

This procedure for framing the optimal estimation problem as a convex optimization over a discrete space space of experiments is extremely powerful. Experimental constraints can be used to limit the menu of possible experiments and the optimal distribution can be

found quickly, even for large problems. Such a restriction to exclude unfeasible experiments is very difficult to incorporate into a continuous optimization technique. From the convex structure of the optimization, we also gain insight into the expected results. By the complementary slackness theorem [67], we expect only a small subset of the possible experiments to contribute to the optimal distribution. This expectation is borne out in the example presented in section 4.5.

Given this formulation for identifying the optimal experiment, we now detail the entire optimal experiment design procedure:

1. **Guess parameters.** We always need an initial estimate of the unknown parameters with which to begin. This assumed value of the parameters, $\boldsymbol{\theta}_p$, can be based on prior knowledge about the quantum process, other studies, or even an educated guess.

2. **Enumerate possible experiments.** The menu of possible experiments $\mathcal{E}$ is dictated by experimental constraints.

3. **Calculate Fisher matrices.** For each experiment on the menu from step 2, the probability distribution for the outcome data, $p^{\mathcal{E}}_{\boldsymbol{\theta}_p}(i)$, and associated Fisher matrices $\mathcal{I}^{\mathcal{E}}(\boldsymbol{\theta}_p)$, must be calculated using the assumed value of the parameters.

4. **Perform optimization.** The optimization specified by Eq. (4.5) must be performed to obtain an optimal probability distribution of experiments, $\lambda_{\mathcal{E}}$.

5. **Perform experiments.** The unknown quantum process should be probed with experiments distributed according to $\lambda_{\mathcal{E}}$. That is, if a total of N samples are taken, $\lceil \lambda_{\mathcal{E}} / N \rceil$ of them should be using experiment $\mathcal{E}$.

6. **Estimate parameter(s).** Use the collected data to estimate the parameters using an estimator of choice. This results in the refined parameter estimate, $\boldsymbol{\theta}_e$. If the maximum likelihood estimator is used, we can readily form the likelihood function since the N samples are independent – the likelihood function will be a multinomial distribution:

$$p_{\boldsymbol{\theta}}(\boldsymbol{X}) \propto N! \prod_{\mathcal{E}} \frac{1}{n_{\mathcal{E}}!} \prod_{i=1}^{n^{\mathcal{E}}_{out}} \left(\lambda_{\mathcal{E}} p^{\mathcal{E}}_{\boldsymbol{\theta}}(i) \right)^{n^{\mathcal{E}}_i} \tag{4.6}$$

   where $n_{\mathcal{E}}$ is the number of times experiment $\mathcal{E}$ was performed ($n_{\mathcal{E}} = \lceil \lambda_{\mathcal{E}} / N \rceil$, and $n^{\mathcal{E}}_i$ is the number of times result i (corresponding to POVM element $M^{\mathcal{E}}_i$) is obtained. The $\boldsymbol{\theta}$ that maximizes this likelihood function is $\boldsymbol{\theta}_e$, the maximum likelihood estimate.

7. **Repeat if necessary.** This procedure can be repeated, with $\boldsymbol{\theta}_p$ in step 1 replaced by $\boldsymbol{\theta}_e$ from step 6. The decision of whether or not to repeat the procedure can be based on a number of factors: (i) experimental resources, (ii) desired accuracy: if $\boldsymbol{\theta}_e$ is very different from $\boldsymbol{\theta}_p$ then repeating the steps is likely to be helpful. Such an adaptive procedure will converge on the true value of the parameter(s), $\boldsymbol{\theta}_t$, through repetition.

We now illustrate this procedure by treating a specific example of constrained multi-parameter estimation that is very relevant to quantum computing: the identification of coupling parameters in a multi-qubit system.

4.4 Dipole- and exchange-coupled qubits

Donors in silicon have been of increasing interest in the quantum computing community since the seminal paper by Kane in Ref. [17]. Most donor based quantum computing schemes use the spins of electrons bound to donors to encode qubits. Single qubit readout for this implementation is an active area of research, but electrically detected magnetic resonance techniques [22, 21, 26, 72, 76] are showing potential for delivering high-quality single qubit measurements. In order to execute high-fidelity quantum gates, accurate knowledge of the coupling Hamiltonian between two donor-bound electron spins is required. Given exact knowledge of the location of the donors in the substrate, this coupling could in principle be computed theoretically. However, donors in silicon devices are subject to uncertainty in location that is only magnified by subsequent annealing processes. Hence it is highly likely that it will be necessary to characterize the qubit couplings for each device separately and therefore an efficient (and preferably optimal) method of doing this characterization is highly desirable. As we will demonstrate in the next section our constrained parameter estimation scheme is well suited to this task because it is numerically efficient and can handle realistic experimental constraints. Before applying our technique we present some details about the physical system.

Two electrons bound to donors implanted in silicon will interact through a combination of the dipole and exchange interactions. The spin Hamiltonian governing dipole coupling between two qubits is

$$H_{\mathrm{d}} = \sum_{i,j} \gamma_1 \gamma_2 \left\langle \frac{3\hat{r}_i \hat{r}_j - \delta_{ij}}{4\pi \left|\boldsymbol{r}\right|^3} + \frac{2}{3}\delta_{ij}\delta^3(\boldsymbol{r}) \right\rangle \sigma_i^{(1)} \otimes \sigma_j^{(2)}$$

where γ_i is proportionality factor relating the magnetic dipole moment operator to the Pauli matrices:

$$\hat{\boldsymbol{\mu}} = \gamma_i \boldsymbol{\sigma^i},$$

$\boldsymbol{r} = \boldsymbol{r}_2 - \boldsymbol{r}_1$ is the vector connecting the two qubits, and $\langle \hat{\mathcal{O}} \rangle = \langle \boldsymbol{\Psi} | \, \hat{\mathcal{O}} \, | \boldsymbol{\Psi} \rangle$ is the expectation value of $\hat{\mathcal{O}}$ over the two electron spatial wavefunction, $\Psi(\boldsymbol{r_1}, \boldsymbol{r_2})$.

The exchange Hamiltonian, a consequence of the Coulomb interaction applied to identical spin-$\frac{1}{2}$ particles, is

$$H_{\mathrm{e}} = J \boldsymbol{\sigma}^{(1)} \cdot \boldsymbol{\sigma}^{(2)}.$$

Here, J is the magnitude of the exchange interaction, calculable from the localized, single-qubit wavefunctions, ϕ, ψ, by:

$$J = e^2 \iint \frac{\phi^*(\boldsymbol{r}_1)\psi^*(\boldsymbol{r}_2)\phi(\boldsymbol{r}_2)\psi(\boldsymbol{r}_1)}{|\boldsymbol{r}|} \mathrm{d}\boldsymbol{r_1} \mathrm{d}\boldsymbol{r_2}$$

The qubits are also subjected to a magnetic field, $B = B_0\hat{\boldsymbol{z}}$, leading to the Hamiltonian,

$$H_0 = -\gamma_1 B_0 \sigma_z^{(1)} - \gamma_2 B_0 \sigma_z^{(2)}.$$

The dipole moments, γ_i, of the two qubits may be different due to local inhomogeneities in the substrate or the magnetic field. This results in each qubit, even without any dipole or exchange interaction, having a distinct resonance frequency $\omega_i = \gamma_i B_0/\hbar$, with difference $\Delta\omega = \omega_2 - \omega_1 << \omega_1, \omega_2$.

The static interaction, H_0, is presumed to be much larger than either H_d or H_e, so it is helpful to work in the interaction picture (also known as the rotating frame) [16]. The effective Hamiltonian in the interaction picture is:

$$H_\mathrm{I} = \sum_{i,j} \left(\gamma_1\gamma_2 \left\langle \frac{\delta_{ij} - 3\hat{r}_i\hat{r}_j}{|\boldsymbol{r}|^3} - \frac{8\pi}{3}\delta^3(\boldsymbol{r})\delta_{ij} \right\rangle + J\delta_{ij} \right) \tilde{\sigma}_i^{(1)} \otimes \tilde{\sigma}_j^{(2)} \tag{4.7}$$

Where, $\tilde{\sigma}_i^{(\alpha)} = e^{-i\omega_\alpha \sigma_z^{(\alpha)} t} \sigma_i^{(\alpha)} e^{i\omega_\alpha \sigma_z^{(\alpha)} t}$, is the i^{th} Pauli matrix in the rotating frame of the α^{th} qubit. These are, explicitly,

$$\begin{aligned}
\tilde{\sigma}_x^{(\alpha)} &= e^{-2i\omega_\alpha t}\sigma_+^{(\alpha)} \;+\; e^{2i\omega_\alpha t}\sigma_-^{(\alpha)} \\
\tilde{\sigma}_y^{(\alpha)} &= -i\, e^{-2i\omega_\alpha t}\sigma_+^{(\alpha)} \;+\; i\, e^{2i\omega_\alpha t}\sigma_-^{(\alpha)} \\
\tilde{\sigma}_z^{(\alpha)} &= \sigma_z^{(\alpha)}
\end{aligned}$$

Substituting these into (4.7), we find many terms proportional to $e^{\pm 2i(\omega_1+\omega_2)t}$. These terms are *very* rapidly oscillating and will average to zero in a short time. We take the rotating wave approximation and neglect these terms, keeping only those that rotate no faster than $e^{\pm 2i\Delta\omega t}$. This leaves us with:

$$H_\mathrm{I} \approx \hbar G \sigma_z^{(1)} \otimes \sigma_z^{(2)} + \left(\hbar F e^{2i\Delta\omega t} \sigma_+^{(1)} \otimes \sigma_-^{(2)} + \text{h.c.} \right)$$

where,

$$\begin{aligned}
\hbar F &= 2J - \gamma_1\gamma_2 \left\langle \frac{(1 - 3\hat{r}_z^2)}{|\boldsymbol{r}|^3} + \frac{16\pi}{3}\delta^3(\boldsymbol{r}) \right\rangle \\
\hbar G &= \;\; J + \gamma_1\gamma_2 \left\langle \frac{(1 - 3\hat{r}_z^2)}{|\boldsymbol{r}|^3} - \frac{8\pi}{3}\delta^3(\boldsymbol{r}) \right\rangle
\end{aligned}$$

In the basis $\{|\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle\}$, this Hamiltonian can be expressed in matrix form as:

$$H_\mathrm{I} = \hbar \begin{pmatrix} G & 0 & 0 & 0 \\ 0 & -G & F\, e^{2i\Delta\omega t} & 0 \\ 0 & F\, e^{-2i\Delta\omega t} & -G & 0 \\ 0 & 0 & 0 & G \end{pmatrix}$$

The unitary evolution operator, $U_I(t) = \mathcal{T} \exp\left(-i \int_0^t H(t')dt'/\hbar\right)$ generated by this Hamiltonian is found to be:

$$U_I(t) = \begin{pmatrix} e^{-iGt} & 0 & 0 & 0 \\ 0 & \chi_- & \Xi & 0 \\ 0 & \Xi & \chi_+ & 0 \\ 0 & 0 & 0 & e^{-iGt} \end{pmatrix} \tag{4.8}$$

where we have defined the following,

$$\Omega = \sqrt{F^2 + \Delta\omega^2},$$

$$\chi_\pm = e^{-i(\Delta\omega - G)t} \left(\cos(\Omega t) \pm i\Delta\omega \sin(\Omega t)/\Omega\right),$$

$$\Xi = -iFe^{-i(\Delta\omega - G)t} \sin(\Omega t)/\Omega.$$

From this expression for the time evolution operator, the Fisher information matrices can be computed through Eq. (4.3).

4.5 Optimal estimation for dipole- and exchange-coupled qubits

In the model described above, the parameters to be estimated are F, G, and $\Delta\omega$. To simplify the presentation and for ease of visualization, we assume here that $\Delta\omega$ has been found through standard resonance techniques, and focus our attention on the two remaining parameters. The general technique is of course valid for any number of parameters (within computational constraints). We choose $\Delta\omega = 1$ and take as the true parameter values, $\boldsymbol{\theta}_t = (F_t = 1.1, G_t = 0.9)$. We work in units where these parameters are dimensionless.

Realistic experimental constraints for the optimization are that the initial states be easy to prepare and the measurements be experimentally accessible. This is satisfied by assuming that all initial states and POVMs are separable. Both the initial state and the POVM set can be specified by the choice of a Bloch-vector for each of the two qubits. To discretize the space of initial states and POVMs, the Bloch vectors are chosen from among the 26 unit-norm vectors, $\boldsymbol{v}$, of the form

$$\boldsymbol{v} = (\alpha\hat{\boldsymbol{x}} + \beta\hat{\boldsymbol{y}} + \gamma\hat{\boldsymbol{z}})/\sqrt{\alpha^2 + \beta^2 + \gamma^2}$$

where $\alpha, \beta, \gamma \in \{\pm 1, 0\}$ are not all zero. These vectors are illustrated in Fig. (4.1)

Given a choice of two Bloch vectors $\left\langle \boldsymbol{\sigma}_0^{(1)} \right\rangle$ and $\left\langle \boldsymbol{\sigma}_0^{(2)} \right\rangle$, the density matrix which describes the resulting initial state is

$$\rho_0 = \frac{1}{4}\left(\boldsymbol{I} + \left\langle \boldsymbol{\sigma}_0^{(1)} \right\rangle \cdot \boldsymbol{\sigma}^{(1)}\right) \otimes \left(\boldsymbol{I} + \left\langle \boldsymbol{\sigma}_0^{(2)} \right\rangle \cdot \boldsymbol{\sigma}^{(2)}\right)$$

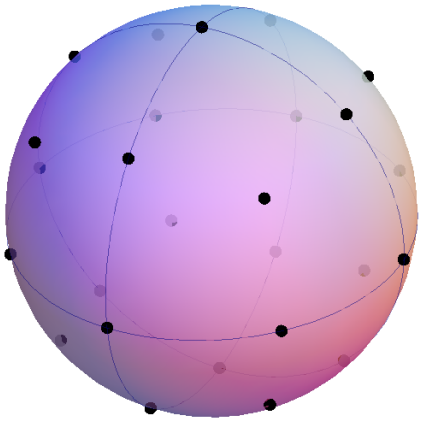


Figure 4.1: (Color online) The single qubit states and measurements used to probe the process represented on the Bloch sphere.

Here $\boldsymbol{\sigma} \equiv (\sigma_x, \sigma_y, \sigma_z)$ is a vector formed from the three non-trivial Pauli matrices. Similarly, given a choice of two Bloch vectors $\left\langle \boldsymbol{\sigma}_M^{(1)} \right\rangle$ and $\left\langle \boldsymbol{\sigma}_M^{(2)} \right\rangle$ the corresponding POVM elements, which we choose in this case projective quantum measurements, are

$$
\begin{aligned}
M_1 &= \frac{1}{4}\left(\boldsymbol{I} + \left\langle \boldsymbol{\sigma}_M^{(1)} \right\rangle \cdot \boldsymbol{\sigma}^{(1)}\right) \otimes \left(\boldsymbol{I} + \left\langle \boldsymbol{\sigma}_M^{(2)} \right\rangle \cdot \boldsymbol{\sigma}^{(2)}\right) \\
M_2 &= \frac{1}{4}\left(\boldsymbol{I} + \left\langle \boldsymbol{\sigma}_M^{(1)} \right\rangle \cdot \boldsymbol{\sigma}^{(1)}\right) \otimes \left(\boldsymbol{I} - \left\langle \boldsymbol{\sigma}_M^{(2)} \right\rangle \cdot \boldsymbol{\sigma}^{(2)}\right) \\
M_3 &= \frac{1}{4}\left(\boldsymbol{I} - \left\langle \boldsymbol{\sigma}_M^{(1)} \right\rangle \cdot \boldsymbol{\sigma}^{(1)}\right) \otimes \left(\boldsymbol{I} + \left\langle \boldsymbol{\sigma}_M^{(2)} \right\rangle \cdot \boldsymbol{\sigma}^{(2)}\right) \\
M_4 &= \frac{1}{4}\left(\boldsymbol{I} - \left\langle \boldsymbol{\sigma}_M^{(1)} \right\rangle \cdot \boldsymbol{\sigma}^{(1)}\right) \otimes \left(\boldsymbol{I} - \left\langle \boldsymbol{\sigma}_M^{(2)} \right\rangle \cdot \boldsymbol{\sigma}^{(2)}\right)
\end{aligned}
$$

These are projectors onto anti-podal points (along the axes defined by $\left\langle \boldsymbol{\sigma}_0^{(1)} \right\rangle$ and $\left\langle \boldsymbol{\sigma}_0^{(2)} \right\rangle$) on the Bloch spheres of the two qubits. The set of initial states and POVMs are explicitly enumerated in Appendix 4.7.3. For simplicity, we will fix the duration of each experiment in the menu to $t = 1$. Therefore, we have $n_\rho = 26^2, n_M = 13^2, n_t = 1$.

The Fisher matrices are calculated using an initial guess $\boldsymbol{\theta}_p = (F_p = 1, G_p = 1)$ and the optimal experiment is then identified using the convex optimization defined in section 4.3. This optimization over $\tilde{n} \equiv n_\rho n_M n_t = 114244$ experiments takes < 3 minutes on an average, consumer-grade desktop computer. The result of this optimization is an experimental configuration with only two elements of $\boldsymbol{\lambda}_\mathcal{E}^o > 0$ ($\boldsymbol{\lambda}_\mathcal{E}^o$ is the probability distribution describing the optimal configuration). This means that the optimal process probe need only sample using two experimental configurations out of the 114244 possible ones. These optimal experimental configurations are shown in Appendix 4.7.3. The non-zero elements of $\boldsymbol{\lambda}_\mathcal{E}^o$ are 0.8 and 0.2, which means 4/5 of the process probes should be performed with one experimental

configuration and the remaining 1/5 with the other. The Fisher information matrix of the optimal configuration is:

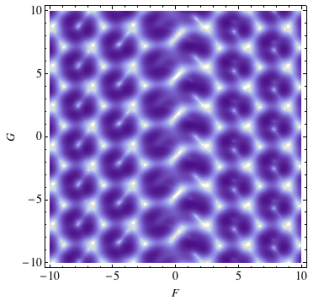
$$\mathcal{I}^o(\boldsymbol{\theta}_p) = \begin{pmatrix} 1.8853 & -0.18431 \\ -0.18431 & 3.3578 \end{pmatrix}$$

The results of the experiments were simulated using $\boldsymbol{\theta}_t$, the actual value of the parameters and the unitary transformation given by Eq. (4.8). We sampled the process $N = 200$ times with initial states and POVMs dictated by the optimal distribution, $\boldsymbol{\lambda}^o_{\mathcal{E}}$, and the resulting data was used to estimate the parameters using the maximum likelihood estimator, Eq. (4.2). The likelihood function, Eq. (4.6), is plotted for a large range of the parameters F and G in Fig. 4.2(a). Finding the maximum over this surface yields $\boldsymbol{\theta}^o_e = (F^o_e = 1.10028, G^o_e = 0.8845)$. The estimate of the parameters is extremely close to the real values given by $\boldsymbol{\theta}_t$. In addition, we can bound the variance of this estimate using the Cramér-Rao bound:

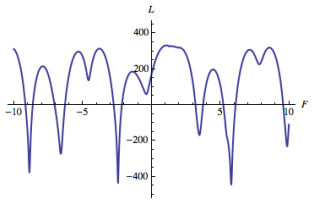
$$\mathrm{Var}[F^o_e] + \mathrm{Var}[G^o_e] \geq \frac{\mathrm{Tr}\left(\mathcal{I}^o(\boldsymbol{\theta}_p)^{-1}\right)}{200} = 0.0042$$

As noted earlier, we know that as the number of samples, N, increases this bound will be saturated. Thus the estimation error is controlled and well-known. In figures 4.2(b) and 4.2(c) we show cross sections across the likelihood function at the estimated values F_e and G_e. These cross sections show how estimation performance is non-uniform for F and G. While the value of F is fairly well resolved, the likelihood function is highly periodic in G. This periodicity reflects the periodic manner in which G enters the evolution unitary, Eq. (4.8). We can break the periodicity of the likelihood function by varying the time for which the quantum process is probed. Figs. 4.2(d)–4.2(f) show that likelihood function and its cross sections when the optimal configuration (for $t = 1$) is used to probe the process for times $t = 1, 1.1, 1.4$. Again, a total of $N = 200$ samples were taken of the process. Periodicity in G is largely absent in 4.2(f), and furthermore, that the likelihood function in 4.2(d) has a dominant central peak around the true values of F and G. This technique of probing a quantum process for varied times is essential when estimating parameters in unitary processes because of the potential for parameters to appear in a periodic manner in unitary maps. We note that since the probe time, t, is actually a parameter of the process it should also be optimized over when identifying the optimal experimental configuration. However, we have not here included this step in the optimization in the interest of keeping the search space of the optimization small enough to explore within ≈ 3 minutes on our simulation computer.

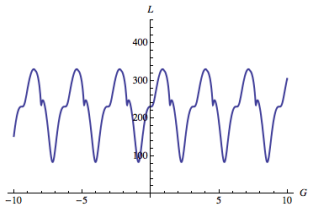
To further evaluate the optimal experiment design we compare its performance against a sub-optimal estimation strategy. The sub-optimal strategy we choose is a discrete set of initial preparations and measurements all aligned along the principal Bloch sphere axes (x, y, z). The 12 possible experimental configurations for this sub-optimal strategy are listed explicitly in Appendix 4.7.3. This is a reasonable naive strategy, and we again collected $N = 200$ samples with experiments distributed uniformly among the 12 possible configurations.



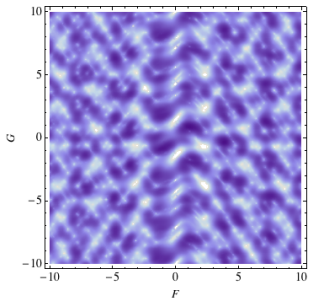
(a) Log likelihood function attained using optimal configuration for a single probe time.



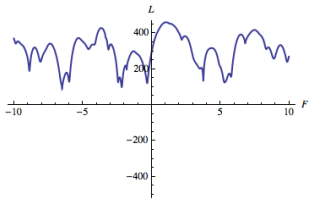
(b) Cross section of the log likelihood function in (a) at the value $G = G_e$.



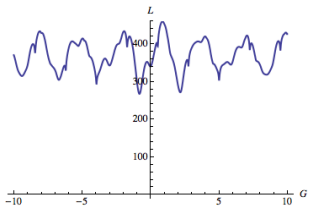
(c) Cross section of the log likelihood function in (a) at the value $F = F_e$.



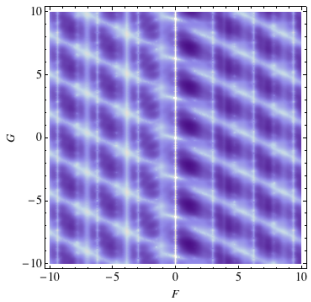
(d) Log likelihood function attained using optimal configuration for multiple probe times.



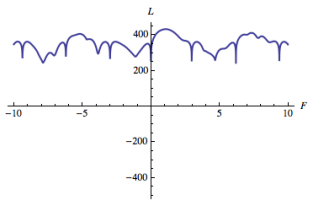
(e) Cross section of the log likelihood function in (d) at the value $G = G_e$.



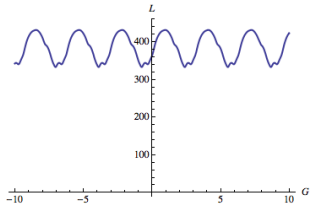
(f) Cross section of the log likelihood function in (d) at the value $F = F_e$.



(g) Log likelihood function attained using sub-optimal configuration for a single probe time.



(h) Cross section of the log likelihood function in (g) at the value $G = G_e$.



(i) Cross section of the log likelihood function in (g) at the value $F = F_e$.

Figure 4.2: (Color online) The logarithm of the likelihood function, Eq. (4.6), for a set of simulated data. Darker areas indicate a larger likelihood function. Sub-figures (a)–(c) show the likelihood function attained using the optimal experiment design. Sub-figures (d)–(f) show the likelihood function attained using the optimal experiment design when the quantum process is probed for different times to break the periodicity of the likelihood in G. Figures (g)–(i) show the likelihood function attained using a sub-optimal configuration of experiments to probe the quantum process. In sub-figures (a), (d) and (g) the regions of white along the $F = 0$ axis are where the likelihood function is zero and hence its log diverges.

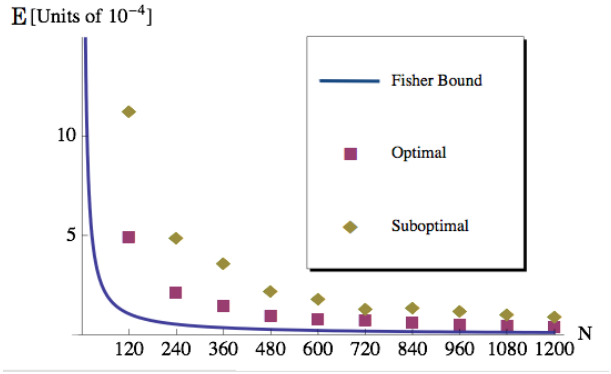


Figure 4.3: (Color online) Plot of the mean squared error (MSE) of the MLE estimator for the optimal (red squares) and suboptimal (yellow diamonds) configurations. Also shown (solid blue line) is the Fisher bound for the mean squared error of *any* estimator as given by the optimal experiment.

The resulting likelihood function is shown in Fig. 4.2(g), and cross sections of it in Figs. 4.2(h) and 4.2(i). Taking the maximum over this likelihood surface yields an estimate of the parameters: $\boldsymbol{\theta}_e^{so} = (F_e^{so} = 1.085, G_e^{so} = 0.969)$. This is clearly a poorer estimate of the true parameters. We can also calculate the Fisher information matrix for this suboptimal strategy:

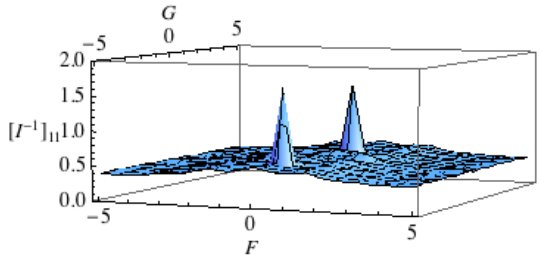
$$\mathcal{I}^{so}(\boldsymbol{\theta}_p) = \begin{pmatrix} 0.5417 & 0.1662 \\ 0.1662 & 0.8562 \end{pmatrix}$$

This Fisher matrix results in the following bound on the combined estimation variance:

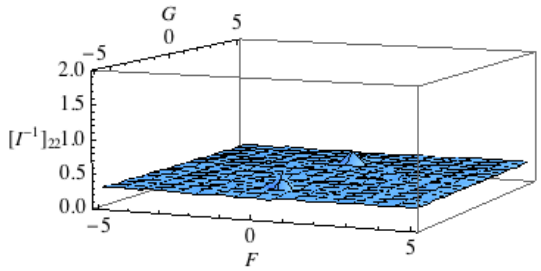
$$\mathrm{Var}[F_e^{so}] + \mathrm{Var}[G_e^{so}] \geq \frac{\mathrm{Tr}\left(\mathcal{I}^{so}(\boldsymbol{\theta}_p)^{-1}\right)}{200} = 0.016$$

The poorer estimate and the larger bound on the variance for the suboptimal configuration are clear indications of the superiority of the optimal experiment design. Furthermore, the number of experimental configurations required to produce a precise estimate of $\boldsymbol{\theta}$ is vastly smaller for the optimal design. In Fig. 4.3, we plot the mean squared error of the maximum likelihood estimate as a function of the number of experiments performed, N. While the MLE for both the optimal and sub-optimal configurations approaches the Fisher information bound (provided by the optimal configuration) as $N \to \infty$, the optimal configuration more rapidly approaches this bound. Furthermore, the mean squared error of the MLE is lower for the optimally configured experiments for all N. To achieve the same mean squared error, one must perform roughly twice as many experiments with the suboptimal configuration as are required with the optimal configuration for this particular set of guessed and actual parameters.

To quantify the estimability of the the parameters in this example, we plot the diagonal elements of the inverse of the Fisher information matrix as of function of the parameters, F



(a) $(1,1)$ element of $\mathcal{I}^{-1}(F,G)$



(b) $(2,2)$ element of $\mathcal{I}^{-1}(F,G)$

Figure 4.4: (Color online) The diagonal elements of the optimal inverse Fisher information matrix over a range of values for the unknown parameters, F and G.

and G in Fig. 4.4. Note that the optimal (constrained) probe configuration has been determined for each (F,G) in the plot since this determines the ultimate estimation performance limit. Fig. 4.4(a) shows element $[\mathcal{I}^{-1}]_{11}(F,G)$ and Fig. 4.4(b) shows element $[\mathcal{I}^{-1}]_{22}(F,G)$. As the values of F and G are changed these elements of the inverse Fisher information matrix remain fairly constant apart from a few notable excursions. This implies that the parameters are nearly equally well estimable across all possible values. If the optimal probe configuration is utilized the single sample variance bound is ~ 0.5 for F and G. If the parameters happen to lie on one of the few peaks of $[\mathcal{I}^{-1}]_{11}(F,G)$, or $[\mathcal{I}^{-1}]_{22}(F,G)$ then they are slightly more difficult to estimate (*i.e.*, a larger number of process probes, N, will be necessary to reduce the estimate variance) for any possible experimental configuration. However, note that neither $[\mathcal{I}^{-1}]_{11}(F,G)$ nor $[\mathcal{I}^{-1}]_{22}(F,G)$ diverge for any value of (F,G), and hence the parameters are always estimable.

4.5.1 Robustness of estimation procedure

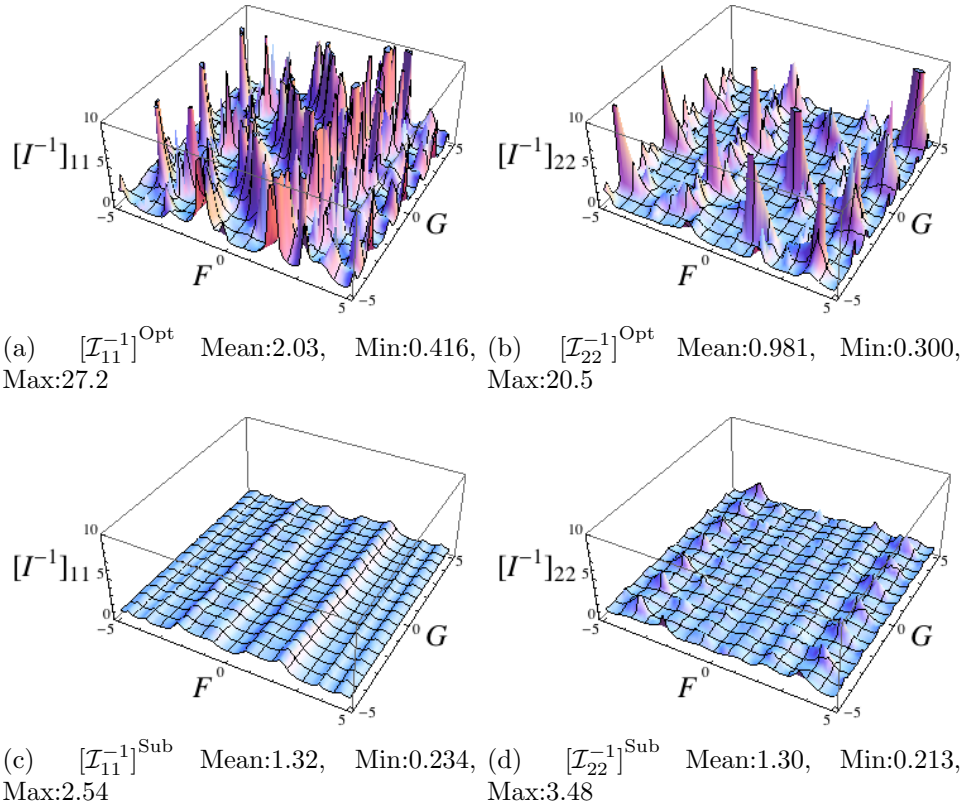
Finally, we turn to the issue of the robustness of the optimal experimental configuration identified by our method. To evaluate robustness, we calculate the inverse Fisher information matrix as a function of the parameters F and G for a fixed experimental configuration (the configuration that is optimal for $(F,G) = (1,1)$). For comparison, we also calculate the inverse Fisher information matrix as a function of the parameters for the fixed sub-optimal process probe configuration used above. The diagonal elements of these matrices are shown in

Fig. 4.5. These figures clearly show that the optimal configuration is much more sensitive to parameter variations than the sub-optimal configuration. In fact, the single sample variance bound for the optimal experiment is quite large at some points. This is a consequence of the small number of finely tuned experimental configurations utilized by the optimal experiment. On the other hand, the large number of experimental configurations exploited by the suboptimal experiment allows for a moderate performance for almost all (F, G). That the optimal experiment is rather sensitive to the accuracy of the initial guess emphasizes the importance of going to the adaptive strategy mentioned in section 4.3. That is, as better estimates of the parameters are produced, the process probes should be adapted to be the optimal configurations for the current guess for the parameters. We expect that this lack of robustness of the optimal experiment will be present for the vast majority of parameter estimation problems and is not a special feature of the example considered here. The cost of finely tuning the process probes to optimally estimate the parameters based on an initial guess is that these probes become less adept at identifying values of parameters too far from the initial guess.

Another important point governing the success of the optimization procedure deals with the experimental ability to accurately prepare and measure the qubit states. In a real experiment, single qubit operations cannot be performed perfectly, and as such will always include a small error. The state prepared under such a noisy operation will be a mixed state that is proximate to the desired target state. Such inaccuracies in preparation and measurement can be easily incorporated into our procedure by replacing the probe state (POVM measurement) constellations with the corresponding achievable mixed states (averaged POVMs). Appendix 4.7.4 analyzes the specific case of small, random gate error in a single-qubit, single-parameter estimation problem. The presence of such error is shown to increase the number of experiments required by an amount proportional to a certain measure of the error.

4.6 Conclusion

The precise estimation of quantum processes is a key ingredient in the engineering of robust quantum information processing devices. For example, to construct two-qubit gates for a quantum computer the interaction between qubits must be precisely known. This estimation task is an increasingly demanding one as the scale of the quantum process being estimated increases. Thus it is essential to have experimental techniques that use minimal resources, but are also accurate. In this work we have demonstrated a method for designing the optimal experiments for multi-parameter quantum process estimation. Particular advantages of the method are that it can tackle multi-parameter estimation, it naturally incorporates realistic experimental constraints, and that the numerical optimization it involves can be implemented efficiently. To demonstrate our approach we have applied it to the estimation of parameters dictating the coupling of two donor electron qubits in silicon. We found the optimal experimental configuration among a very large candidate set ($> 10^5$ experiments)



(a) $[\mathcal{I}_{11}^{-1}]^{\text{Opt}}$ Mean:2.03, Min:0.416, Max:27.2

(b) $[\mathcal{I}_{22}^{-1}]^{\text{Opt}}$ Mean:0.981, Min:0.300, Max:20.5

(c) $[\mathcal{I}_{11}^{-1}]^{\text{Sub}}$ Mean:1.32, Min:0.234, Max:2.54

(d) $[\mathcal{I}_{22}^{-1}]^{\text{Sub}}$ Mean:1.30, Min:0.213, Max:3.48

Figure 4.5: (Color online) The diagonal elements of the inverse Fisher Information matrix over a range of values for the unknown parameters, F and G for fixed process probe configurations. Subfigures (a) and (b) show the diagonal elements of the inverse Fisher matrix when the optimal experimental configuration for the guess $(F, G) = (1, 1)$ is used, and subfigures (c) and (d) show these matrix elements when the suboptimal experimental configuration identified in the text is used. As evident from the large deviations in (a) and (b), the small number of experiments used in the optimal configuration reduces the robustness of the procedure to errors in the initial guess.

and simulated the parameter estimation using this optimal configuration. The results show that the our method can drastically reduce the number of experiments required to perform parameter estimation for quantum processes. We compared the optimal configuration found by our method with a sub-optimal approach and quantified the performance improvement of the optimal configuration. We also found that while the optimal experiments designed by our procedure – which are based on an initial guess of the parameters – perform very well, they are very sensitive to variations in the actual values of the parameters, and hence lack robustness. However, the general algorithm outlined in 4.3 takes this into account by specifying a recipe for adapting the estimation procedure as data about the values of the parameters is obtained, and hence is capable of compensating for this lack of robustness in the results of the optimization.

A useful extension of this work is to investigate the feasibility of including a robustness measure directly into the cost function of the optimization. It remains to be seen if this can be done while maintaining the optimization's convexity. Robust estimation procedures have been addressed in the context of classical control theory [77] and the extension of these results to quantum models would increase the practicality and appeal of optimal estimation in the quantum setting.

While we illustrated the method here with the example of electron qubits in silicon, the general technique of optimal experiment design for parameter estimation outlined in section 4.3 is applicable to a wide array of physical systems. An interesting avenue for further research would be to apply this method to identify the Hamiltonians governing small dipole-coupled spin clusters such as those probed in recent experiments with diamond [78, 79].

Finally, although the numerical optimization required to find the optimal experimental configuration is convex, and therefore efficient, in the process of applying our technique to the example detailed above we noticed that current optimization libraries were unable to handle a very large ($\tilde{n} > 150,000$) search space. Therefore a possible extension of this work is to use the inherent structure in the parameter estimation problem to form a smaller optimization program, or possibly to iteratively identify the optimal solution.

4.7 Appendices

4.7.1 Derivation of the Cramér-Rao bound for a single-parameter

In this appendix we provide a derivation of the Cramér-Rao bound on the variance of a single-parameter estimator. We denote by $E_\theta(X)$ the function which estimates the parameter, θ from a set of data, X. Furthermore, we define the *score function* as the derivative of the log-likelihood,

$$S = \frac{\partial}{\partial\theta} \ln f(X;\theta)$$

The likelihood function is the probability of getting the results, X, given that the parameter, θ. We begin by calculating the covariance of the score function and the estimator,

$$\mathrm{cov}_\theta\left(S, E_\theta(x)\right) = \langle V E_\theta(X)\rangle - \langle S\rangle \langle E_\theta(X)\rangle$$

However, the expectation value of the score is zero,

$$\begin{aligned}
\langle S \rangle &= \left\langle \frac{\partial}{\partial \theta} \ln f(X;\theta) \right\rangle \\
&= \int dX f(X;\theta) \frac{\partial}{\partial \theta} \ln f(X;\theta) \\
&= \int dX f(X;\theta) \frac{\frac{\partial}{\partial \theta} f(X;\theta)}{f(X;\theta)} \\
&= \int dX \frac{\partial}{\partial \theta} f(X;\theta) \\
&= \frac{\partial}{\partial \theta} \int dX f(X;\theta) \\
&= \frac{\partial}{\partial \theta} 1 \\
&= 0
\end{aligned}$$

We can now write the above covariance as

$$\mathrm{cov}_\theta \left(S, E_\theta(x) \right) = \langle S E_\theta(X) \rangle$$

Expanding this,

$$\begin{aligned}
\mathrm{cov}_\theta \left(S, E_\theta(x) \right) &= \langle S E_\theta(X) \rangle \\
&= \int dX f(X;\theta) E_\theta(X) \frac{\frac{\partial}{\partial \theta} f(X;\theta)}{f(X;\theta)} \\
&= \int dX E_\theta(X) \frac{\partial}{\partial \theta} f(X;\theta) \\
&= \frac{\partial}{\partial \theta} \int dX E_\theta(X) f(X;\theta) \\
&= \frac{\partial}{\partial \theta} \langle E_\theta(X) \rangle
\end{aligned}$$

If the estimator is unbiased, then its expectation value should be

$$\langle E_\theta(X) \rangle = \theta.$$

Thus the above covariance is

$$\mathrm{cov}_\theta \left(S, E_\theta(X) \right) = \frac{\partial}{\partial \theta} \theta = 1$$

We can use the Cauchy-Schwartz inequality to relate the covariance of the score and the estimator as,

$$\sqrt{\mathrm{var}\left(S\right) \mathrm{var}\left(E_\theta(X)\right)} \geq \left| \mathrm{cov}_\theta \left(S, E_\theta(X) \right) \right| = 1$$

Squaring and rearranging this, we have

$$\text{var}\left(E_\theta(X)\right) \geq \frac{1}{\text{var}\left(S\right)}$$

And we define the *Fisher Information* as the variance of the score,

$$F(\theta) = \text{var}\left(S\right) = \int dX f(X;\theta) \left(\frac{\partial}{\partial\theta} \ln f(X;\theta)\right)^2 = \int dX \frac{\left(\frac{\partial}{\partial\theta} f(X;\theta)\right)^2}{f(X;\theta)}.$$

Thus we recover the single-parameter Cramér-Rao bound,

$$\text{var}\left(E_\theta(X)\right) \geq \frac{1}{\text{var}\left(S\right)}$$

4.7.2 Schur complement

Consider a nonsingular block matrix,

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

where A is an invertible submatrix. The *Schur complement* of M with respect to A is defined as,

$$M/A \equiv D - CA^{-1}B.$$

A principal theorem in the study of the Schur complement [80] says that

$$M \geq 0 \iff \text{both } A \geq 0 \text{ and } M/A \geq 0.$$

Where $M \geq 0$ means that M is positive semi-definite. Now consider the minimization problem

$$\begin{aligned} \text{minimize} \quad & \text{Tr}\, F^{-1} \\ \text{subject to} \quad & \text{some constraints} \end{aligned}$$

As stated in the main text, this objective function is not convex. We then propose a matrix $Q \geq F^{-1}$ which is an upper-bound on the matrix F^{-1}. This definition implies that

$$Q - F^{-1} = Q - IF^{-1}I \geq 0.$$

And, because F is positive semi-definite, F^{-1} is as well, implying $Q \geq 0$. So, by the above theorem,

$$\begin{pmatrix} Q & I \\ I & F \end{pmatrix} \geq 0.$$

So we construct the following optimization problem:

$$\begin{aligned} \text{minimize} \quad & \operatorname{Tr} Q \\ \text{subject to} \quad & \text{same constraints}, \\ & \begin{pmatrix} Q & I \\ I & F \end{pmatrix} \geq 0. \end{aligned}$$

The (convex) matrix positivity constraint enforces the (non-convex) constraint $Q \geq F^{-1}$, leaving us with a convex optimization problem (assuming the remaining constraints are also convex).

4.7.3 Probe constellation and optimal experiments from section 4.5

The 26 single qubit initial states available to probe the quantum process in the example presented in section 4.5 are given in table 4.1. Explicitly, for each set of Bloch angles, the initial state is: $|\psi(\phi,\theta)\rangle = \cos(\phi/2)\,|0\rangle_z + e^{i\theta}\sin(\phi/2)\,|1\rangle_z$ where $|i\rangle_z$ are σ_z eigenstates. The 13 single qubit POVMs assumed available in section 4.5 are defined by the first 13 angles in table 4.1, and each POVM has two projector elements given by: $|\psi(\phi,\theta)\rangle\langle\psi(\phi,\theta)|$ and $\boldsymbol{I} - |\psi(\phi,\theta)\rangle\langle\psi(\phi,\theta)|$.

The sub-optimal estimation strategy used in section 4.5 used a fixed set of state preparations and measurements to probe the quantum process. There were 12 possible experimental configurations and they are explicitly enumerated in table 4.2. The initial states and POVM elements are defined explicitly in terms of these Bloch sphere angles by the same procedure outline above.

Table 4.1: Bloch sphere angles for the 26 initial states in section 4.5. ϕ is the polar angle and $\chi \equiv \cos^{-1}(1/\sqrt{3})$. Antipodal points are equivalent when choosing POVM's, leading to 13 inequivalent, single-qubit measurement bases.

ϕ	θ
0	0
$\pi/4$	$\{0, \pi/2, \pi, 3\pi/2\}$
χ	$\{\pi/4, 3\pi/4, 5\pi/4, 7\pi/4\}$
$\pi/2$	$\{0, \pi/4, \pi/2, 3\pi/4, \pi, 5\pi/4, 3\pi/2, 7\pi/4\}$
$\pi - \chi$	$\{\pi/4, 3\pi/4, 5\pi/4, 7\pi/4\}$
$3\pi/4$	$\{0, \pi/2, \pi, 3\pi/2\}$
π	0

The optimal experimental design from the $\tilde{n} = 114244$ possible configurations (defined by all possible combinations of initial state and POVM parameters from Table 4.1) consists

Table 4.2: Bloch sphere angles (ϕ, θ) for the 12 experimental configurations used by the sub-optimal estimation strategy in section 4.5. ϕ is the polar angle, and Q1 and Q2 refer to qubit 1 and qubit 2.

Init. state Q1	Init. state Q2	POVM Q1	POVM Q2
$(0, 0)$	$(0, 0)$	$(0, 0)$	$(0, 0)$
$(0, 0)$	$(\pi, 0)$	$(0, 0)$	$(0, 0)$
$(\pi/2, 0)$	$(-\pi/2, 0)$	$(0, 0)$	$(0, 0)$
$(\pi/2, 0)$	$(0, 0)$	$(0, 0)$	$(0, 0)$
$(0, 0)$	$(0, 0)$	$(\pi/2, \pi/2)$	$(\pi/2, \pi/2)$
$(0, 0)$	$(\pi, 0)$	$(\pi/2, \pi/2)$	$(\pi/2, \pi/2)$
$(\pi/2, 0)$	$(-\pi/2, 0)$	$(\pi/2, \pi/2)$	$(\pi/2, \pi/2)$
$(\pi/2, 0)$	$(0, 0)$	$(\pi/2, \pi/2)$	$(\pi/2, \pi/2)$
$(0, 0)$	$(0, 0)$	$(\pi/2, 0)$	$(\pi/2, 0)$
$(0, 0)$	$(\pi, 0)$	$(\pi/2, 0)$	$(\pi/2, 0)$
$(\pi/2, 0)$	$(-\pi/2, 0)$	$(\pi/2, 0)$	$(\pi/2, 0)$
$(\pi/2, 0)$	$(0, 0)$	$(\pi/2, 0)$	$(\pi/2, 0)$

of only two experiments. These are given in Table 4.3, and graphical representations of the initial states and POVM axes are given in Fig. 4.6.

Given below are the Fisher information matrices for the experimental configurations chosen by the optimization procedure. Following these is the inverse of the total Fisher matrix, $\mathcal{I}^{\rm opt}(\boldsymbol{\theta}_p) = \sum_{\mathcal{E}} \lambda_{\mathcal{E}} \mathcal{I}^{\rm opt}_{\lambda_{\mathcal{E}}}(\boldsymbol{\theta})$.

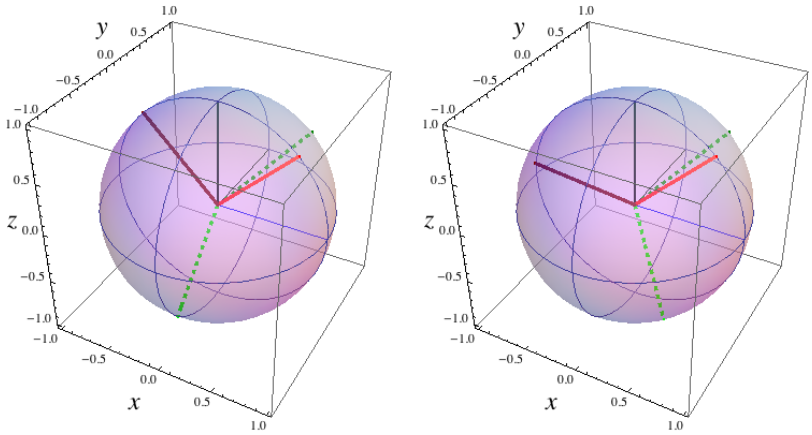
$$\mathcal{I}^{\rm opt}_{0.2} = \begin{pmatrix} 2.03 & -0.034 \\ -0.034 & 2.82 \end{pmatrix} \qquad \left[\mathcal{I}^{\rm opt}_{0.2}\right]^{-1} = \begin{pmatrix} 0.49 & 0.0059 \\ 0.0059 & 0.35 \end{pmatrix}$$

$$\mathcal{I}^{\rm opt}_{0.8} = \begin{pmatrix} 1.85 & -0.22 \\ -0.22 & 3.49 \end{pmatrix} \qquad \left[\mathcal{I}^{\rm opt}_{0.8}\right]^{-1} = \begin{pmatrix} 0.54 & 0.035 \\ 0.035 & 0.29 \end{pmatrix}$$

$$\left[0.2 \times \mathcal{I}^{\rm opt}_{0.2} + 0.8 \times \mathcal{I}^{\rm opt}_{0.8}\right]^{-1} = \begin{pmatrix} 0.53 & 0.029 \\ 0.029 & 0.30 \end{pmatrix}$$

Table 4.3: Bloch sphere angles (ϕ, θ) and relative weights in $\boldsymbol{\lambda}^o_{\mathcal{E}}$ for the two experimental configurations that are optimal for the estimation problem of section 4.5. ϕ is the polar angle, and Q1 and Q2 refer respectively to qubit 1 and qubit 2.

$\boldsymbol{\lambda}^o_{\mathcal{E}}$	Init. state Q1	Init. state Q2	POVM Q1	POVM Q2
0.2	$(3\pi/4, 3\pi/2)$	$(\chi, \pi/4)$	$(\pi/4, 0)$	$(\pi/4, \pi)$
0.8	$(\pi - \chi, 7\pi/4)$	$(\chi, \pi/4)$	$(\pi/4, 0)$	$(\chi, 5\pi/4)$



(a) $\boldsymbol{\lambda}^o_{\mathcal{E}} = 0.2$ experiment (b) $\boldsymbol{\lambda}^o_{\mathcal{E}} = 0.8$ experiment

Figure 4.6: (Color online) Bloch sphere representations of the initial states and POVM axes for the two experiments of the optimal configuration. The green (dotted) lines are Bloch vectors for the initial states of each qubit, and the red (solid) lines define the axes whose antipodal points define the projectors of the optimal POVM for each qubit.

4.7.4 Robustness to gate errors

When a pure state is acted upon by a noisy gate, the result is a mixed state. This mixed state can be represented by a Bloch vector which terminates on the interior of the Bloch sphere. Though the details depend on the error model for the gate, imperfections in preparation and measurement can easily be taken into account by our formalism. One simply optimizes over a discretized set of imperfectly prepared input states and imperfect measurements. Note that even though all inputs states and measurements treated in the example were pure states and projective measurement, respectively, our formalism is not restricted to optimizing over such states and POVMs. Specifically, imperfections in state preparation can be taken into account by considering the result of these imperfections on the Bloch vector, $\vec{v}$, of the state, defined through it's relation to the single-qubit density matrix.

If we assume that our target state is pure, then $\vec{v}_f$ is of unit-norm and the density matrix is:

$$\rho = \frac{1}{2}\left(I + \vec{v}_f \cdot \vec{\sigma}\right)$$

Random error in preparation of the initial state corresponds to the creation of a mixed state. If this error is assumed to be such that the final state is instead created with some finite probability density surrounding the target state, the effect on the Bloch vector is that it is contracted by some factor, $\vec{v}'_f = (1-\epsilon)\vec{v}_f$. The details of the error govern the magnitude of ϵ. (Of course, there are errors which do not just shrink the Bloch vector, but also rotate it. As long as these errors are well characterized, then a similar analysis may be performed.)

The density matrix is then,

$$\rho' = \frac{1}{2}\left(I + (1-\epsilon)\vec{v}_f \cdot \vec{\sigma}\right)$$

In the case where there exists only a single parameter, θ, the Fisher information takes the form:

$$\mathcal{I} = \sum_i \frac{\left(\frac{\partial}{\partial\theta} p_i(\theta)\right)^2}{p_i(\theta)}$$

The probability, $p_i(\theta)$, as given by the Born rule for a POVM element, M_i, is

$$\begin{aligned} p_i(\theta) &= \mathrm{Tr}\left(M_i \rho'\right) \\ &= \frac{1}{2}\mathrm{Tr}\left(M_i(I + (1-\epsilon)\vec{v}_f \cdot \vec{\sigma})\right) \\ &= \frac{1}{2}\left(\mathrm{Tr}\left(M_i\right) + (1-\epsilon)\mathrm{Tr}\left(\vec{v}_f \cdot \vec{\sigma} M_i\right)\right) \end{aligned}$$

Then the Fisher information becomes, in terms of the Bloch vector,

$$\begin{aligned} \mathcal{I} &= \frac{1}{2}\sum_i \frac{\left((1-\epsilon)\frac{d}{d\theta}\mathrm{Tr}\left(\vec{v}_f \cdot \vec{\sigma} M_i\right)\right)^2}{\mathrm{Tr}\left(M_i\right) + (1-\epsilon)\mathrm{Tr}\left(\vec{v}_f \cdot \vec{\sigma} M_i\right)} \\ &\approx (1-\epsilon)^2 \frac{1}{2}\sum_i \frac{\left(\frac{d}{d\theta}\mathrm{Tr}\left(\vec{v}_f \cdot \vec{\sigma} M_i\right)\right)^2}{\mathrm{Tr}\left(M_i\right) + \mathrm{Tr}\left(\vec{v}_f \cdot \vec{\sigma} M_i\right)} \\ &= (1-\epsilon)^2 \mathcal{I}_0 \end{aligned}$$

Here, $\mathcal{I}_0$, is the Fisher Information achieved without the presence of gate error. The estimator error, $\mathrm{var}\, E_\theta$, is bounded by the Cramer-Rao inequality,

$$\mathrm{var}\, E_\theta \geq \frac{\mathcal{I}^{-1}}{N} = \frac{1}{(1-\epsilon)^2}\frac{\mathcal{I}}{N}.$$

So to achieve the same bound on the estimator variance as is found with perfect gates, one must increase the number of measurements from N to $N' \approx N(1+2\epsilon)$. If there are similar POVM errors as well, then a nearly identical calculation shows that $N' \approx N(1+4\epsilon)$.

Part III

Noise in quantum systems

Chapter 5

Noise simulation and mitigation by pulse sequences

5.1 Introduction

Coherent control of the interactions between quantum bits (qubits) and their environment is essential for any physical system to be useful for quantum computation. Exercising control of coherence by application of carefully designed pulse sequences is a standard tool in NMR and ESR, where powerful sets of tools have been developed for protection against qubit dephasing. Such coherent control provides a complementary approach to the protection of quantum information by encoding, whether with active quantum error correction codes or passive encoding into decoherence free ('noiseless') subspaces and subsystems. In the quantum information processing community, the application of coherent control ideas to preservation of qubit coherence has prompted an explosion of work in the field of dynamical decoupling [81, 82] which builds on the classic spin echo techniques from NMR. An alternative approach is the design of numerically optimized control pulse sequences, which have the advantage of flexibility and ready applicability to both quantum memory and protection of arbitrary quantum gates against dephasing and leakage errors [14, 83, 84].

In this chapter we consider the design of optimal pulse sequences for control of decoherence during single qubit operations when the qubit is coupled to source of Markovian noise that possesses an arbitrary noise spectrum. The Hamiltonian governing the evolution of the qubit is taken to be

$$H(t) = \frac{1}{2}\left(a_x(t)\sigma_x + a_y(t)\sigma_y + \eta(t)\sigma_z\right), \tag{5.1}$$

where we are working in a rotating frame so that the qubit energy level splitting is, on average, zero. Here, $a_x(t)$ and $a_y(t)$ are bounded-amplitude control fields, while $\eta(t)$ is a classical stochastic process. In the absence of control, this Hamiltonian generates dephasing of the qubit, at a rate that depends on both the amplitude distribution and the temporal correlation function of the stochastic process $\eta(t)$, as discussed in Chapters 2 and 6. Previous work in our group has modeled this stochastic process as a multi-state Markovian fluctuator, the dynamics of which may be tuned to mimic a particular noise source [14] The use of a multi-state Markovian fluctuator allows the evolution under the Hamiltonian, Eq. (5.1) to be efficiently solved through use of a deterministic master equation [85]. In the present work, we extend this approach from the $1/\omega^\alpha$ spectral noise sources for which analytic representations of the fluctuator dynamics could be found [14], to representation of a broad variety of Markovian spectral noise distributions by making use of convex optimization techniques. We then use gradient ascent methods as in Ref. [14] to derive control sequences for single qubit operations. Following established literature convention, these pulse sequences will be generically refered to as "GRAPE sequences" (i.e., gradient-ascent pulse engineering sequences) [86]. We provide examples of two single-qubit operations: i) the identity, which is equivalent to extension of qubit coherence, i.e., to quantum memory; ii) the Hadamard gate. We analyze the robustness of these pulse sequences for protection against the combined effects of $1/\omega$ noise and zero frequency noise (the latter is equivalent to a resonance frequency error) and compare with the corresponding performance of standard dynamical decoupling sequences, in particular with Carr–Purcell sequences. We find that the numerically optimized control

sequences improve on the dynamical decoupling sequences over a broad range of zero frequency noise offsets, resulting in considerably greater robustness in addition to improved decoherence mitigation. Finally, we demonstrate the applicability of the method for current day experiments by making explicit application to the protection of coherence for dopant spin qubits in silicon using realistic estimates of spectral noise and control pulse capabilities. The results indicate that gates with errors less than 10^{-5} can be designed and implemented with current technology. This is well below current estimates of the fault-tolerant threshold [87].

5.2 Simulated noise model

Simulating the noisy evolution of a quantum system is a notoriously difficult task. In this chapter, evolution under classical noise $\eta(t)$ is simulated using an extension of a method first developed in [85, 14] using a deterministic master equation to evaluate the evolution under a stochastic fluctuator whose dynamics are chosen to best approximate a target noise spectrum. Prior to the development of this method, one had to rely on Monte Carlo simulations of stochastic dynamics. This technique is computationally expensive and therefore did not allow for the fast optimization of noise mitigating pulse sequences.

To illustrate our method, we consider the model Hamiltonian Eq. (5.1). We shall restrict the possible values of $\eta(t)$ to any of N possible noise states, allowing Markovian fluctuations between states. We denote the amplitude of noise state k as η_k. The probability of the fluctuator to be in noise state k is given by $p_k(t)$ at time, t. These noise amplitudes and occupation probabilities will be represented as vectors, $\vec{\eta}$ and $\vec{p}$. Transitions between noise states are governed by a rate matrix, Γ,

$$\frac{d}{dt}\vec{p}(t) = \Gamma\,\vec{p}(t).$$

To conserve probability, the transition rate matrix must satisfy $\sum_k \Gamma_{kj} = 0$. This constraint implies that the vector $\vec{p}_{\rm s} = (1/N, 1/N, \ldots, 1/N)$ is a stationary probability vector, and is an eigenvector of Γ with zero eigenvalue. We shall limit our study of the rate matrices to those satisfying the additional requirement, $\Gamma = \Gamma^{\rm T}$. This condition makes the forward and backward transition rates between any two noise states to be equal, enforcing time-reversal invariance on the fluctuator dynamics.

The noise source may be further characterized by its temporal correlation function

$$\begin{aligned} C(t) &= \langle \eta(t)\eta(0)\rangle \\ &= \sum_{i,j} P(\eta_j, t|\eta_i, 0)\eta_j P(\eta_i(0))\eta_i \\ &= \frac{1}{N}\sum_{i,j} \eta_i \left[e^{\Gamma|t|}\right]_{ij} \eta_j, \end{aligned}$$

where $P(\eta_i(0)) \equiv p_i(0)$ denotes the probability of the fluctuator being in state i at time $t = 0$ and $P(\eta_j, t|\eta_i, 0)$ the conditional probability of it being in state j at time t, given state i at $t = 0$. Here we have chosen as the initial noise probabilities the stationary vector $\vec{P}(t) = \vec{p}_s = (1/N, 1/N, \ldots, 1/N)$. Because Γ is a symmetric matrix, it can be diagonalized by an orthogonal matrix, $\Gamma = V^\dagger \Lambda V$, so that

$$C(t) = \frac{1}{N} \vec{\eta}^\dagger V^\dagger e^{\Lambda|t|} V \vec{\eta} = \vec{b}^\dagger e^{\Lambda|t|} \vec{b}, \tag{5.2}$$

where we have defined the transformed noise amplitude vector $\vec{b} = V\vec{\eta}/\sqrt{N}$ and $\Lambda = \mathrm{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_N\}$ is the diagonal matrix of eigenvalues of Γ. For convenience, the eigenvalues are ordered $i > j \Rightarrow \lambda_i < \lambda_j$. The power spectrum, $S(\omega) = \int C(t) e^{-i\omega t} dt$, of the multistate fluctuator is a sum of zero-mean Lorentzian distributions:

$$S\left(\omega; \vec{\lambda}, \vec{b}\right) = \sum_j \frac{-2\, b_j^2 \lambda_j}{\lambda_j^2 + \omega^2} \tag{5.3}$$

Ref. [14] derived an analytic form of Γ and V that generates noise with a $1/\omega^\alpha$, $0 < \alpha < 2$, power spectrum. As noted there, numerical optimization may result in a more accurate representation.

As with all Markovian processes, the form of Eq. (5.3) is, in accordance with Doob's theorem [88], a sum of Lorentzians. This form constrains the possible target spectra to those which are monotonically decreasing and which never decay faster than $1/\omega^2$. We have found that by proper choice of $\vec{\lambda}$ and $\vec{b}$, this spectrum may indeed be brought arbitrarily close to a given target spectrum, $S_t(\omega)$, (chosen with the above constrains in mind) over a finite specified range of frequencies, $\omega \in [\omega_{\min}, \omega_{\max}]$. The choice of $\vec{\lambda}$ and $\vec{b}$ is made by a numerical optimization that minimizes the deviation of Eq. (5.3) from the target spectrum. In particular, we carry out the following optimization:

$$\begin{aligned} \underset{\vec{\lambda},\vec{b}}{\text{minimize}} \quad & \int_0^\infty W(\omega) \left(S\left(\omega; \vec{\lambda}, \vec{b}\right) - S_t(\omega) \right)^2 d\omega \\ \text{subject to} \quad & b_i \geq 0, \;\; \lambda_i \leq 0. \end{aligned}$$

Since i) ω can span many orders of magnitude, and ii) analytic representations of power spectra often diverge at $\omega = 0$, we have incorporated here a weighting function, $W(\omega)$, into the usual L^2 distance measure. In particular, we have set $W(\omega) = 1/\omega$ for $\omega \in [\omega_{\min}, \omega_{\max}]$ and $W(\omega) = 0$ otherwise. This weight function is uniformly distributed in $\log \omega$, preventing higher frequencies from dominating the integral. Restricting ourselves to the range $\lambda_i \leq 0$ is physically realistic, since positive eigenvalues would not conserve probability and would cause the correlation function Eq. (5.2) to diverge at long times.

The results of this optimization are the two vectors $\vec{b}_{\text{opt}}$ and $\vec{\lambda}_{\text{opt}}$. Recall that the constraints on Γ imply the existence of a stationary probability vector, $\vec{p}_{\text{s}}$ with eigenvalue $\lambda_1 = 0$. This implies that one component of $\vec{b}_{\text{opt}}$, say b_1, can be taken to be a free parameter and

may be chosen to make the arithmetic mean of $\vec{\eta} = \sqrt{N}V^\dagger \vec{b}$ equal to zero, guaranteeing the existence of the stationary solution $\vec{p}_s = (1/N, 1/N, \ldots, 1/N)$. We note that it is convenient to further make a restriction to $b_i \geq 0$ during the numerical optimization, because the power spectrum depends only on b_i^2. However, following the optimization, we may subsequently adjust the signs of all components $b_{i\neq 1}$ so that $|b_1|$ is as small as possible, consistent with the existence of the stationary probability vector.

It now remains to construct a valid transition rate matrix Γ with eigenvalues given by $\vec{\lambda}_{\rm opt}$. This is again done by a numerical optimization, namely

$$\begin{aligned} \underset{\Gamma}{\text{minimize}} \quad & \left(\text{eigs}(\Gamma) - \vec{\lambda}_{\rm opt}\right)^2 \\ \text{subject to} \quad & \Gamma = \Gamma^{\mathsf{T}}, \;\; \Gamma_{i\neq j} \geq 0 \\ & \sum_j \Gamma_{ij} = 0, \end{aligned} \tag{5.4}$$

with eigs(Γ) the vector of eigenvalues of Γ.

Both of the above optimizations can performed very quickly using standard convex optimization tools [89, 90], though the optimization in Eq. (5.4) may be accelerated via the methods given in Appendix. 5.6.1. Our numerical approach is quite general and may be readily applied to generate other spectral distributions (subject to the constraints mentioned above). In the remainder of this paper we shall employ a multi-state Markovian fluctuator representation of $1/\omega$ noise together with a zero frequency component η_{os} that describes possible sources of inhomogeneous dephasing. This combination is experimentally relevant to a broad range of physical qubits.

To illustrate the efficiency and flexibility of this generation of arbitrary noise spectral densities by numerical optimization of a multi-state Markovian fluctuator, we applied the approach to calculation of a 4-state Markovian fluctuator representation of a target spectrum $S_t \propto 1/\omega + \eta_{\rm os}^2 \delta(\omega)$. The result is

$$\vec{\eta} = \eta_{\rm os} + \epsilon(-0.875, 1.36, -1.36, 0.875) \tag{5.5}$$

$$\Gamma = \Gamma_m \begin{pmatrix} -7.69 & 7.64 & 0.0322 & 0.0123 \\ 7.64 & -8.41 & 0.694 & 0.0694 \\ 0.0322 & 0.694 & -0.730 & 0.00437 \\ 0.0123 & 0.0694 & 0.00437 & -0.0861 \end{pmatrix}.$$

with $\eta_{\rm os}$ the constant noise offset responsible for zero frequency noise, ϵ the scaling of the noise amplitude (represented as a fraction of the maximum control amplitude) and Γ_m a constant that tunes the range of frequencies over which the fluctuator best approximates S_t. The resulting power spectrum for parameter set $\epsilon = 10^{-3}$, $\Gamma_m = 1/30$ and $\eta_{\rm os} = 0$ is shown for a range of finite frequency ω in Fig. 5.1. For comparison we also show the corresponding approximation to the $1/\omega$ power spectrum derived from a 32-state Markovian fluctuator with the analytic form of Ref. [14]. It is evident that the numerically optimized 4-state Markovian

fluctuator provides a significantly improved fit relative to the analytic approximation, as well as a significantly greater range of representation. Such enhanced accuracy, together with the considerable increase in efficiency and greater flexibility, illustrated here by the addition of the zero frequency noise component in the fit (see Eq. (5.5)), render this numerical optimization approach to generation of arbitrary spectral noise densities extremely attractive.

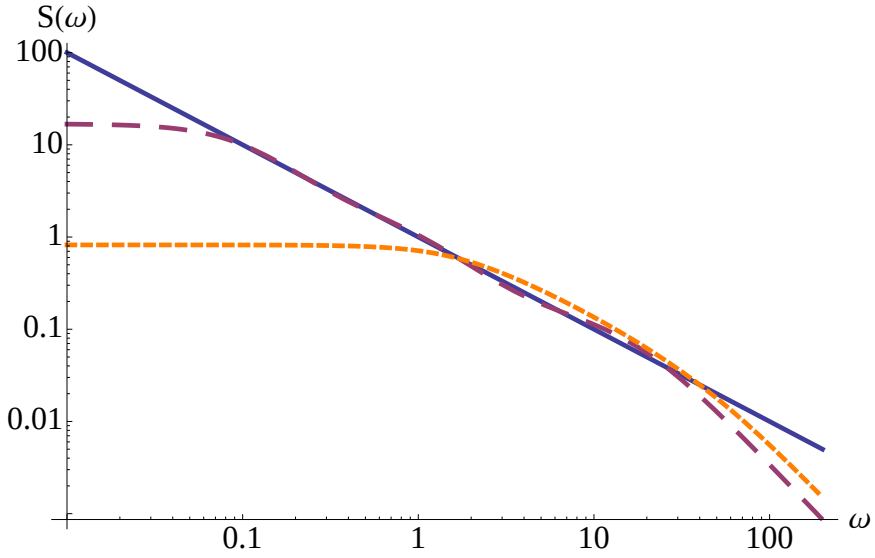


Figure 5.1: (Color online) Numerically optimized noise power spectral density (red, dashed line) with optimization constructed to match the target noise spectrum $S_t \propto 1/\omega + \eta_{os}^2 \delta(\omega)$ (blue, solid line) over two decades of frequency ω. Also shown is the fit obtained with the analytic representation of Ref. [14] using 32 noise states (yellow, dotted line).

Qubit evolution with noise

The evolution of this one qubit system under classical dephasing noise $\eta(t)$ is exactly solvable through the use of conditional density matrices, $\rho_k(t)$, as described in [85]. We outline here a slightly modified version of this approach. Defining $\rho_k(t)$ as the density operator of the system conditioned on the environment being in the state k, the total density operator of the system is given by the sum of the conditional density operators weighted by the probability of occupation of the associated noise state:

$$\rho(t) = \sum_{j=1}^{N} p_k(t)\rho_k(t).$$

We choose as the initial probability vector $\vec{p}(0)$ the stationary probability, $\vec{p}_s$. For a single qubit, we can parameterize the conditional density matries $\rho_k(t)$ by their Bloch vectors,

$\vec{\zeta}_k(t)$,

$$\rho_k(t) = \frac{1}{2}\left(\mathcal{I} + \vec{\zeta}_k(t)\cdot\vec{\sigma}\right),$$

where $\mathcal{I}$ is the identity operator and $\vec{\sigma}$ is the vector of Pauli spin-1/2 matrices. The resulting stochastic Liouville equation for the conditional density matrices can be transformed to yield the dynamics of the conditional Bloch vectors, which are given by

$$\frac{d}{dt}\vec{\zeta}_k(t) = M_k(t)\vec{\zeta}_k(t) + \sum_j \Gamma_{kj}\vec{\zeta}_j(t). \tag{5.6}$$

Here $M_k \in \mathfrak{so}(3)$ is the generator of Bloch vector rotations,

$$M_k(t) = \begin{pmatrix} 0 & -\eta_k & a_y(t) \\ \eta_k & 0 & -a_z(t) \\ -a_y(t) & a_z(t) & 0 \end{pmatrix}.$$

The second term in Eq. (5.6) describes the effect of the noise switching on the conditional Bloch vectors. We have thus arrived at a set of N coupled matrix differential equations for the evolution of the N conditional Bloch vectors. These can be solved by treating the set of Bloch vectors as a single vector composed by stacking the conditional Bloch vectors to get a single $3N$-dimensional vector, $\vec{Z}(t) = \bigoplus_{k=1}^{N}\vec{\zeta}_k(t)$. The equation of motion for $\vec{Z}(t)$ can be straightforwardly derived from Eq. (5.6) and is given by

$$\frac{d}{dt}\vec{Z}(t) = \left(\bigoplus_{k=1}^{N} M_k(t) + \Gamma\otimes\mathcal{I}\right)\vec{Z}(t) \equiv \mathcal{L}(t)\vec{Z}(t).$$

This is solved formally in the usual way, namely as $\vec{Z}(t) = \underset{\leftarrow}{T}\exp\left(\int_0^t \mathcal{L}(t')dt'\right)\vec{Z}(0)$, where the symbol, $\underset{\leftarrow}{T}$, is the usual Dyson time-ordering operator. This time-ordered integral becomes a time-ordered product if we restrict the control functions, $a_x(t)$ and $a_y(t)$, to those that are piecewise-constant in time. The Lindblad operators, $\mathcal{L}(t)$, are also then piecewise constant, taking values $\mathcal{L}_i$ for times δ_i. For future convenience, we divide the control functions into $2N_\mathrm{p}$ subintervals, where subinterval i will in general take nonzero amplitude for i even (and be called a "pulse"), and will take zero amplitude for i odd, corresponding to a quiescent time between pulses. Thus, N_p is understood to mean the number of pulses in the control pulse sequence. Each control has time duration δ_i and the total time for a pulse sequence is equal to $\tau = \sum_i \delta_i$. With $2N_\mathrm{p}$ control function values, the corresponding Bloch vector dynamics are given by

$$\vec{Z}(\tau) = \left(\underset{\leftarrow}{T}\prod_{i=1}^{2N_\mathrm{p}}\exp\left(\mathcal{L}_i\,\delta_i\right)\right)\vec{Z}(0).$$

Calculating the evolution of a given initial state is then a matter of matrix multiplication. Because the probability vector $\vec{p}(t) = \vec{p}_s$, all noise states are equally probable and the relation

between $\vec{Z}(t)$ and the Bloch vector, $\vec{\zeta}(t)$, is given by

$$\vec{\zeta}(t) = \frac{1}{N}\begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & 1 & 0 & \cdots \\ 0 & 0 & 1 & 0 & 0 & 1 & \cdots \end{pmatrix} \vec{Z}(t) \equiv \frac{1}{N}\mathcal{I}_N \vec{Z}(t).$$

The inverse relation is simply $\vec{Z}(0) = \mathcal{I}_N^\dagger \vec{\zeta}(0)$ and the final Bloch vector is then $\vec{\zeta}(\tau) = \mathcal{E}\vec{\zeta}(0)$, where

$$\mathcal{E} \equiv \frac{1}{N}\mathcal{I}_N \cdot \left(\underset{\leftarrow}{T} \prod_{i=1}^{N_\mathrm{p}} \exp\left(\mathcal{L}_i \, \delta t_i\right) \right) \cdot \mathcal{I}_N^\dagger. \tag{5.7}$$

Using the above expression allows one to compute the evolution of a quantum system under the action of both noise and control by simple matrix multiplication.

Numerically optimized control

The control functions must now be chosen to realize a desired target operation on the Bloch vector, $\vec{\zeta} \rightarrow G\vec{\zeta}$. We choose the operator fidelity to measure the efficacy of these control functions. We define the fidelity function as [15],

$$\begin{aligned} \Phi_G[a_x(t), a_y(t)] &= \frac{1}{2}\left(1 + \vec{\zeta}(\tau) \cdot G\,\vec{\zeta}(0)\right) \\ &= \frac{1}{2}\left(1 + \mathcal{E}\,\vec{\zeta}(0) \cdot G\,\vec{\zeta}(0)\right). \end{aligned}$$

Note that the fidelity is a functional of the control fields, $a_{x/y}(t)$. From the perspective of quantum information, no state is any more important that any other, so we would ideally like our pulse sequences to maximize the worst-case fidelity over all possible initial states. However, the minimization over initial states to find the worst-case fidelity is too expensive a computation to yield a useful cost function. Therefore we use instead as our cost function the average case fidelity and compare this with the worst case fidelity obtained from the optimized pulse sequence at the end of the computation in order to ascertain the range of errors. Thus we

$$\begin{aligned} \underset{a_x(t), a_y(t)}{\text{maximize}} \quad & \left\langle\, \Phi_G[a_x(t), a_y(t)] \,\right\rangle_{\vec{\zeta}(0)} \\ \text{subject to} \quad & a_x(t)^2 + a_y(t)^2 \leq 1, \end{aligned}$$

where the notation, $\langle\cdot\rangle_{\vec{\zeta}}$, implies an average taken over the surface of the Bloch sphere. The above average can be evaluated as

$$\begin{aligned}\langle\, \Phi_G[a_x(t), a_y(t)]\, \rangle_{\vec{\zeta}(0)} &= \frac{1}{2} + \frac{1}{2}\sum_{i,j,k} \langle \mathcal{E}_{ij} G_{ik} \zeta_j(0)\zeta_k(0)\rangle_{\vec{\zeta}(0)} \\ &= \frac{1}{2} + \frac{1}{2}\sum_{i,k} \mathcal{E}_{ik} G_{ik} \left\langle \zeta_k(0)^2 \right\rangle_{\vec{\zeta}(0)} \\ &= \frac{1}{2} + \frac{1}{6}\,\mathrm{Tr}\big(\mathcal{E}G^{\mathsf{T}}\big)\,.\end{aligned}$$

We have used the result that the average over the surface of a unit sphere is given by $\langle\zeta_i\zeta_j\rangle_{\vec{\zeta}} = \delta_{ij}/3$. Finally, as discussed above, we demand that the optimized pulse sequences be insensitive to zero-frequency noise. This is achieved by including a constant offset value, η_{os}, to the noise vector, as in Eq. 5.5. We systematically analyze the effect of this additional zero frequency noise by choosing the offset to take values within the range $|\eta_{\mathrm{os}}| \leq 10\epsilon$. The optimization problem then becomes

$$\begin{aligned}\underset{a_x(t),a_y(t)}{\text{maximize}} \quad & \min_{\eta_{\mathrm{os}}} \ \mathrm{Tr}\big(\mathcal{E}G^{\mathsf{T}}\big) \\ \text{subject to} \quad & a_x(t)^2 + a_y(t)^2 \leq 1.\end{aligned} \tag{5.8}$$

It is in general possible to find analytic gradients of $\Phi_G[a_x(t), a_y(t)]$ in terms of the pulse sequence parameters by straightforward methods of [91] when η_{os} is fixed. However, because the objective function has the form of a minimum value over some range of η_{os}, the objective function is not in general differentiable everywhere. We therefore use the solver [89, 90] which employs finite difference approximations to the gradient (which may be undefined in certain regions). A finite difference minimization approach requires many more function evaluations than an explicit gradient calculation, greatly increasing optimization time.

The optimization is performed by undertaking a sampling over the allowed parameter space. We begin by randomly selecting an initial point in the space, and applying numerical optimization techniques to find a locally optimal value of the objective function. We repeat this process many times, each time obtaining a value for Φ_G. After some fixed number of initial conditions are sampled (typically thousands), the pulse sequence obtaining the greatest value of the operator fidelity, Φ_G, is selected as the optimal sequence.

It is important to note that the dimension of the parameter space for optimization is $4N_{\mathrm{p}}$. A pulse sequence contains N_{p} pulses, each of which can be characterized by four parameters: the amplitude, duration, phase angle of the control fields, and quiescent period before the next pulse. To thoroughly sample the enlarging parameter space, the number of initial points sampled for optimization should grow as $n_s^{4N_{\mathrm{p}}}$, where n_s is the number of initial samples taken for $N_{\mathrm{p}} = 1$.

An interesting point is raised by Fig. 5.2, which shows that the worst-case error of a numerically optimized Hadamard gate as a function of N_{p}. If the computational effort is allowed to grow exponentially with N_{p} (and therefore linearly with the size of the parameter

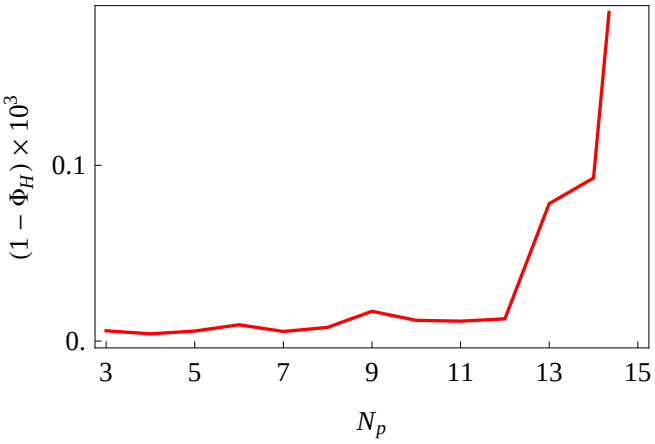


Figure 5.2: Worst-case error, $(1-\Phi_H)$, over the Bloch sphere for a Hadamard gate generated by GRAPE as a function of the number of pulses in the pulse sequence up to $N_p = 15$ for fixed computational effort. See text for disucssion.

space), we can expect the gate error to decrease monotonically with N_p. With exponential computational resources, one is able to sample all of the parameter space defining the pulse sequences. And because the set of $N_\mathrm{p} = n$ pulse sequences is a strict subset of the set of $N_\mathrm{p} = n+1$ pulse sequences, such a search should yield sequences that, at the very worst, do not decrease in efficacy. However, when the number of initial sample points does not grow exponentially with N_p, the performance of GRAPE will suffer because it becomes exceedingly unlikely to find good optima of the objective function as N_p becomes large. This exponential scaling in the number of sample points required places an upper limit on the largest values of N_p for which the GRAPE approach will be useful. In the examples studied in this work, we find that for gates other than the identity, this scaling restricts values of N_p to a maximum of 6-10. Thus in Fig. 5.2 we see that the error begins to rise after $N_\mathrm{p} = 10$. For non-trivial gates this is not a disadvantage since it is in any case advantageous to implement the gate in as short a time as possible to avoid decoherence. The exception to this is the identity gate, which one wishes to realize over as long a time as possible when implementing a quantum memory. In this case we find that good identity gates may be found for $N_\mathrm{p} \sim 30$.

5.3 Results

To demonstrate the power and flexibility of the improved multi-state Markovian fluctuator approach to coherent control of qubit dephasing, we apply it to studying magnetic interface noise experienced by a phosphorus dopant atom implanted in the channel of a silicon MOSFET. de Sousa has proposed a model of this noise as caused by dangling bonds located at the interface between the crystalline Si and the amorphous oxide [92]. These defects, known as P_b centers, are associated with oxygen vacancies in the oxide and have the structure

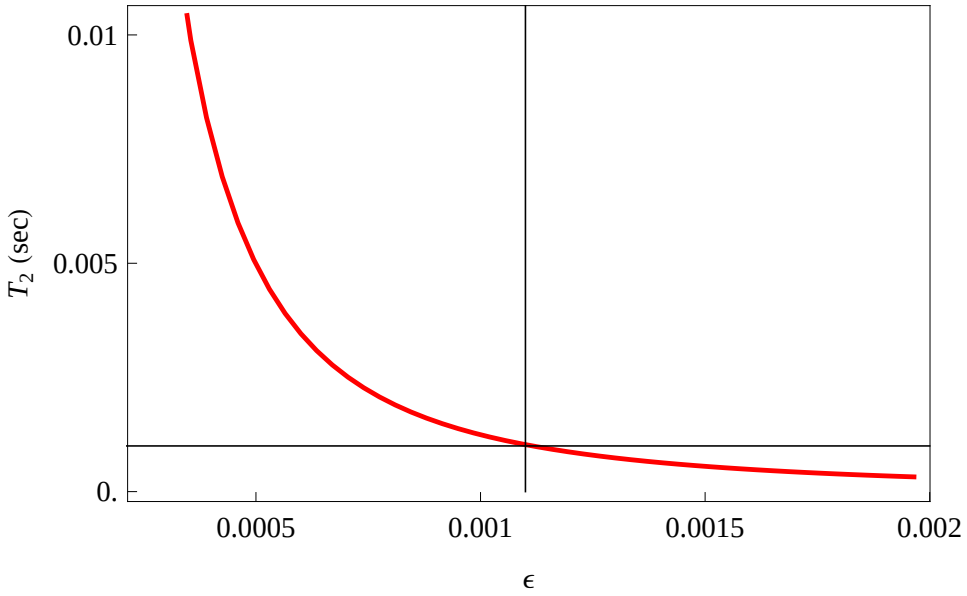


Figure 5.3: (Color online) Calculated qubit T_2 times as a function of the noise parameter ϵ. Results are presented for qubit T_2 with the application of a Carr-Purcell sequence with a 1% duty cycle. The intersecting horizontal and vertical lines indicate the value of ϵ at which the electron T_2 equals 1ms.

$\cdot\text{Si} \equiv \text{Si}_3$. The lone electron in the dangling bond can provide a thermally switching paramagnetic contribution to the magnetic environment experienced by the donor that causes magnetic field noise. This noise then acts to dephase the electron spin qubit defined on the phosphorus dopant. By modeling the P_b center spin flips as coupled to tunneling two-level systems in the oxide, de Sousa has shown that the resulting magnetic field noise possesses an approximate power spectral density $S(\omega) \propto 1/\omega$ [92]. Recent work by Paik et al. has provided evidence in favor of de Sousa's model [93].

We can empirically estimate the proportionality constant, or equivalently the noise strength ϵ (Eq. (5.5)) by comparing the calculated T_2 time of an electron spin qubit with the experimental values extracted for phosphorus donors implanted in Si MOSFET devices. Donor electron T_2 times have been shown to be several milliseconds in isotopically purified silicon [94, 95].

We can determine an effective T_2 time for a pulse sequence in our model as follows. We initialize a qubit in the +1 eigenstate of σ_x, and apply a quantum memory pulse sequence repeated several times. If one measures only at the conclusion of each repetition of the pulse sequence, the quantity $\langle \sigma_x \rangle$ will decay approximately exponentially in time, with a time constant that we define to be T_2.

We have set the value of the parameter ϵ to be that which yields $T_2 \approx 1$ms under a Carr-Purcell pulse sequence operated with a 1% duty cycle. As demonstrated in Figure 5.3, this condition is satisfied with $\epsilon = 0.0011$, which we have rounded to $\epsilon = 10^{-3}$.

In the remainder of this section we explore the power of numerically optimized pulse sequences obtained with the improved multistate Markovian fluctuator, for two target unitary operations subject to this interface-induced $1/\omega$ noise with an additional constant noise offset that allows for heterogeneous dephasing. The first is quantum memory, i.e., the preservation of coherence of an arbitrary quantum state, while the second is a single-qubit Hadamard transformation. We find that excellent performance of GRAPE for both operations can be attained even in the presence of additional zero-frequency (constant) noise.

Consistent with the application to experimentally accessible phosphorus dopants silicon devices, we construct here pulse sequences that may be implemented by current signal generators and that are thus subject to limitations on the on/off ratio. Consequently, we enforce a 50% duty cycle on the GRAPE sequences, i.e., each pulse is followed by a brief quiescent period and where the total quiescent time must be at least half of the total pulse length, or, $\tau/2$.

5.3.1 Quantum memory

We begin this section with a discussion of coherence maintaining operations, known generally as *quantum memory gates*. When designing such pulses, one must make pulse design decisions based on the specifics of the experiment in question. Consider a particular experiment which requires that coherence be maintained for a certain time, t. Ideally, one would design a pulse sequence itself having total length t, as well. In principle, such a pulse sequence would have much more flexibility than a sequence of duration t/n repeated n times. However, the computational effort scales exponentially in the number of pulses, as discussed above, and long sequences may be difficult to find which match or exceed the performance of repeated short sequences. Which choice is made will depend strongly on the computational resources available to the pulse designer.

Here we present a numerical solution of Eq. (5.8) with G the identity matrix, total pulse sequence time $\tau = 30\tau_\pi$ and total number of pulses $N_\mathrm{p} = 30$. The particular optimal solution found under these constraints is shown in Fig. 5.4. Here τ_π refers to the time required to perform a full π rotation of the qubit at maximum control amplitude. This value of N_p was chosen because it was the largest value for which we were able to obtain results in a reasonable amount of computer time (see discussions above). To compare with these numerically optimized pulse sequences we construct an equivalent length Carr-Purcell (CP) decoupling pulse sequence, defined by

$$w - \pi_x - w - w - \pi_x - w$$

repeated 7 times, where π_x denotes a π-pulse about the x-axis, and w denotes a quiescent period of $\tau_w = \frac{4}{7}\tau_\pi$.

Fig. 5.5 shows the error (defined as 1 - Φ_I) as a function of the zero frequency noise η_os for a numerically generated pulse sequence that is optimized over all values of η_{os} (red line), in addition to optimization against the $1/\omega$ noise. The blue and green lines show

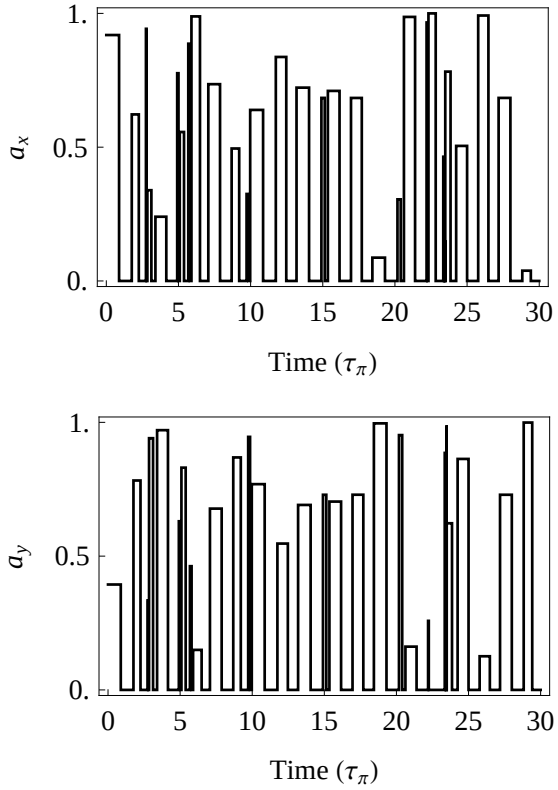


Figure 5.4: (a) x-axis control function and (b) y-axis control function for implementing quantum memory operations.

the corresponding fidelities obtained with the Carr–Purcell sequence of equivalent duration specified above, using finite amplitude (dot-dashed blue line) and infinite amplitude (dotted green line) pulses.

Infinite amplitude Carr–Purcell pulses are capable of refocusing arbitrarily large zero-frequency noise, resulting in a constant error as a function of $\eta_{\rm os}$ whose value can be taken as a measure of the uncorrected error due to the $1/\omega$ noise component. Unlike the ideal, infinite amplitude pulse sequence, a Carr–Purcell sequence with finite amplitude pulses does not allow complete Bloch sphere rotations, which prevents the exact refocusing of zero-frequency noise, resulting in a significant rise in error as the zero-frequency noise magnitude $|\eta_{os}|$ increases. Note that the lack of time reversal symmetry possessed by the optimized pulses (unlike the Carr–Purcell sequence) results in an asymmetry with respect to $\eta_{\rm os}$, as illustrated in Fig. 5.5.

The numerically optimized sequence shows improved performance relative to these Carr-Purcell decoupling sequences in two respects. First, it performs better than the Carr–Purcell pulse sequences for zero and small $|\eta_{os}|$, due to the greater flexibility of the numerical optimization in developing protection against the $1/\omega$ noise component. Thus, at $\eta_{\rm os} = 0$, the error obtained with the numerically optimized pulse sequence is 2.88×10^{-5}, compared to 3.26×10^{-5} with the finite amplitude Carr–Purcell pulse sequence. However, a far more dramatic difference is the greater robustness against the magnitude of zero-frequency noise. The numerically optimized pulse sequence is seen to show very small error over a broad range of η_{os}, attesting to the power of the numerical approach to mitigate combined decoherence effects deriving from very different noise sources.

5.3.2 Hadamard gate

Our second target operation is the Hadamard gate,

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \tag{5.9}$$

a common single qubit operation in quantum algorithms. The optimization considerations for implementing such a single qubit rotation with numerically optimized pulse sequences are similar to those for generating sequences to protect the identity gate. However, in contrast to the situation for quantum memory, here we are interested in maximizing fidelity and robustness to a constant noise offset, rather than in the maintenance of coherence over a long time. Thus the optimal pulse sequences for protection of the Hadamard gate are considerably shorter than the sequences derived above for protection of quantum memory.

Using the same cost function as Eq. (5.8) and evaluating Φ_H for the H operation, we were able to obtain high-fidelity pulse sequences with $\tau = 6\tau_\pi$ and $N_{\rm p} = 6$. Fig. 5.6 shows the resulting pulse sequence when optimization is made for the case of zero offset noise, $\eta_{os} = 0$.

This pulse sequence results in a worst case error for the Hadamard gate of 8.27×10^{-6} at $\eta_{os} = 0$. Thus it is evident that gate operations can readily be corrected at similar or

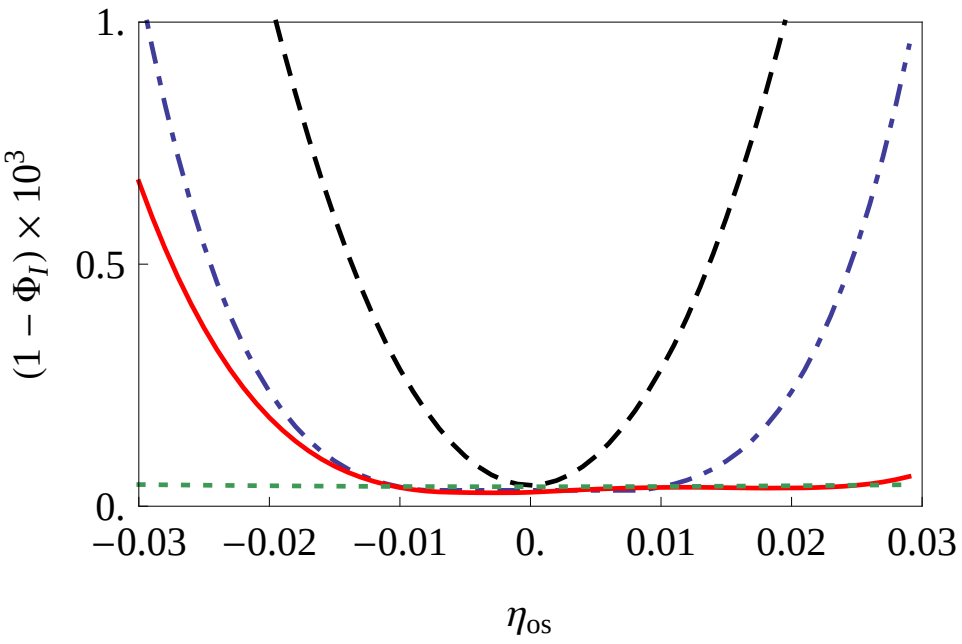


Figure 5.5: (Color online) Error $(1-\Phi_I)$ in quantum memory of a qubit subject to dephasing noise with spectral density $S_t \propto 1/\omega + \eta_{\rm os}^2\delta(\omega)$ under various control pulse sequences of duration $30\tau_\pi$, shown as a function of offset noise values η_{os}. The solid red line represents the error for a sequence that is optimized over a range of offset noise. Green dotted line: error obtained with infinite-amplitude Carr–Purcell sequence. Blue dot-dashed line: error obtained with finite amplitude Carr-Purcell sequence. Also shown as the black, dashed line is a pulse sequence obtained without regard to robustness over $\eta_{\rm os}$.

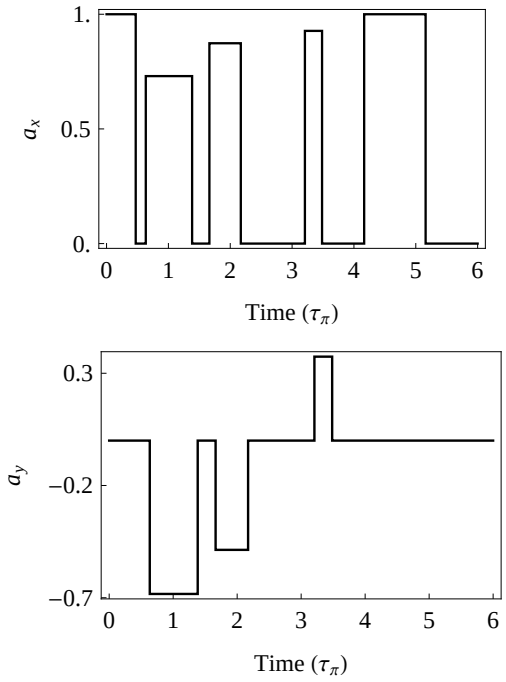


Figure 5.6: Two-dimensional control function producing a high fidelity H rotation in $T = 6\pi$ for $\eta_{os} = 0$. This pulse sequence results in a worst case fidelity of $\Phi_H = 1 - 8.27 \times 10^{-6}$ and exhibits a strong robustness to the value of constant offset noise η_{os}. Panel (a) shows the x-axis control function and panel (b) the y-axis control.

better levels than quantum memory, using shorter pulse sequences. Fig. 5.7 shows the worst case error for a numerically optimized pulse sequence optimized over a range of $\eta_{\rm os}$ values, as before, as a function of the noise offset η_{os}. Comparison with the results obtained with a single pulse sequence that is optimized only for $\eta_{os} = 0$ shows again the enhanced robustness afforded by the numerical optimization approach.

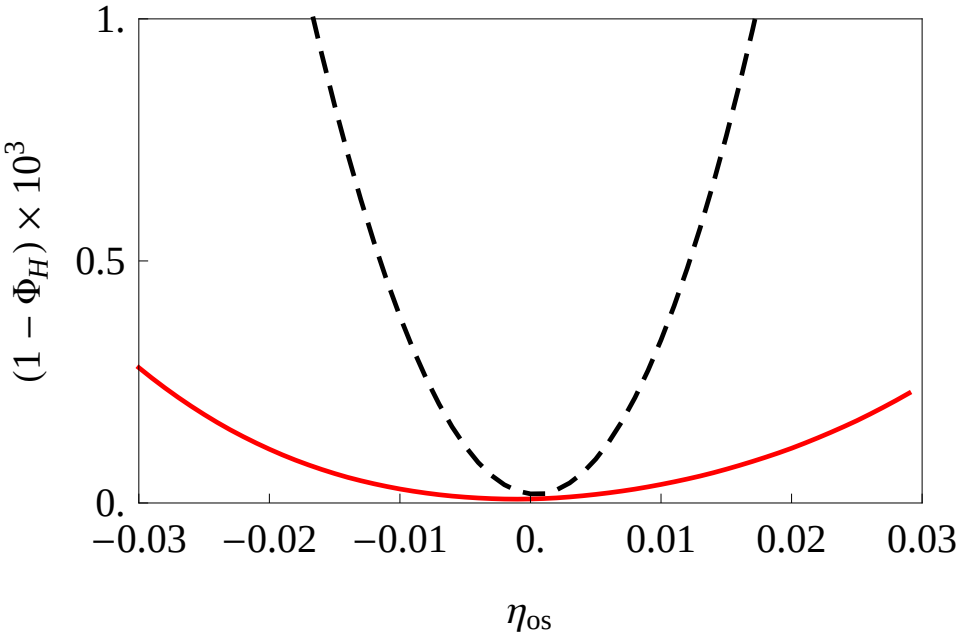


Figure 5.7: (Color online) Error (1-Φ_H) of a H rotation under numerically optimized pulse sequences mitigating against noise with spectral density $S_t \propto 1/\omega + \eta_{\rm os}^2\delta(\omega)$ as a function of $\eta_{\rm os}$. The solid, red line shows the result of optimizing with respect to $\eta_{\rm os}$ as well as the $1/\omega$ noise. The dashed, black line shows the considerably less robust result of using a pulse sequence that is optimized only at a single value of η ($\eta_{\rm os} = 0$).

5.4 Summary

We have expanded the techniques of [14, 83, 85] to develop a general numerical method for simulating noise sources deriving from a broad variety of Markovian power spectra. The method employs a new numerical approach to generation of the noise spectrum that can greatly reduce the number of noise states required to simulate a noise source with a given spectral density. We illustrated this with the example of a four state simulation of a $1/\omega$ spectrum over two decades of frequency ω, which is significantly more efficient than the constructive method employed previously in Ref. [83]. This numerical representation of

Markovian noise was then used in the context of numerical generation of control pulse sequences to mitigate the effects of this noise on a single qubit. Here we have extended the methods of [83] to allow control pulses to be performed along both x- and y-axes of the qubit, rather than along a single axis. Using numerical generation of the noise allows spectral densities from different sources of dephasing to be combined, giving rise to considerable additional flexibility and robustness in the decoherence mitigation. This was illustrated by generation of pulse sequences designed to minimize decoherence in the presence of both homogeneous dephasing characterized by $1/\omega$ spectral density and a source of heterogeneous dephasing, characterized by a zero frequency noise offset η_{os}. The numerical optimization approach allows the pulse sequences to be simultaneously optimized with respect to the parameter $\eta_{\rm os}$ and the $1/\omega$ noise. This introduces an unprecedented robustness to decoherence mitigation with realistic bounded amplitude controls in the presence of resonance frequency errors and inhomogeneous broadening. In particular, the performance of the numerically optimized pulse sequences over a range of $\eta_{\rm os}$ values was seen to be considerably superior to the corresponding performance of a standard dynamical decoupling pulse sequence with bounded amplitudes.

As a demonstration of the power and flexibility of these numerical methods for noise mitigation, we have explicitly studied the protection of quantum memory and the protection of the Hadamard gate. To ground the derived pulse sequences to a physical system, we took estimates of noise strength that are appropriate to the situation of dephasing noise acting on phosphorus donors in silicon and implemented the numerical optimization subject to realistic constraints of duty cycle and pulse amplitude limitations. The remarkable robustness of the optimal pulse sequences with respect to the constant noise offset, showing worst case gate errors of order $10^{-6} - 10^{-5}$ over a range of noise offsets, is encouraging for application of these pulse sequences to current experiments with spin qubits in semiconductors [94].

5.5 Acknowledgements

This work has been supported by the National Security Agency under MOD713100A. DJG also thanks UC LEADS for financial support. The authors would like to thank Thomas Schenkel for many useful discussions.

5.6 Appendix

5.6.1 Optimizing fluctuator dynamics

The multistate markovian fluctuator is a powerful tool for simulating Markovian noise. However, one must rely on a two-step numerical optimization to construct a fluctuator with a power spectrum similar to the desired spectrum. As discussed above, the first step is to determine the values for b_i and λ_i which make the power spectrum, Eq. (5.3), match the desired spectrum. The second step is to come up with a proper transition rate matrix, Γ,

with the λ_i's as eigenvalues. This step requires the solution to the following optimization problem,

$$\begin{aligned} \underset{\Gamma}{\text{minimize}} \quad & \left(\text{eigs}(\Gamma) - \vec{\lambda}_{\text{opt}}\right)^2 \\ \text{subject to} \quad & \Gamma = \Gamma^{\mathsf{T}},\ \Gamma_{i\neq j} \geq 0 \\ & \sum_j \Gamma_{ij} = 0. \end{aligned}$$

In order to make the spectrum match the desired spectrum over several decades of frequency space, the transition rates must span several orders of magnitude. Experience has shown that this causes problems for the optimization routines. Instead, we developed a method inspired by the Gershgorin circle theorem [96].

The circle theorem describes the eigenvalues of an $N \times N$ complex matrix, A, with entries, a_{ij}. We construct the N Gershgorin disks, $D_i(a_{ii}, R_i)$, as the disks in the complex plane centered at a_{ii} with radius $R_i = \sum_{j\neq i} |a_{ij}|$, the sum of the absolute values of the i^{th} row of A. The theorem says that each eigenvalue of A lies within one of the Gershgorin disks. If all of the disks are disjoint, then each disk contains a single eigenvalue. If the disks overlap, however, this may not be the case, an issue that arrises in our application. Because of the limits placed on the entries of Γ, the radii $R_i = |a_{ii}|$, so all of the Gershgorin disks will overlap, as shown in Fig. 5.8. Therefore, the theorem cannot be directly applied in the construction of transition rate matrices.

However, we have discovered a surprisingly good *ad hoc* method for constructing this matrix which is inspired by the circle theorem. Essentially, we construct a set of pseudorandom matrices such that the Gershgorin disks are centered on the target eigenvalues. The procedure follows.

Given the optimized eigenvalues, $\{\lambda_i\}$, we sort them in ascending order and divide each by the maximum eigenvalue, so the set becomes:

$$\lambda \to \left(0, \tilde{\lambda}_1, \tilde{\lambda}_2, \ldots, -1\right)$$

We then form the difference set,

$$\delta = \left(\left|\tilde{\lambda}_1 - 0\right|, \left|\tilde{\lambda}_2 - \tilde{\lambda}_1\right|, \left|\tilde{\lambda}_3 - \tilde{\lambda}_2\right|, \ldots, 1\right)$$

We now construct a large number (10,000) of matrices, Γ_k, as follows. First construct the strictly upper triangular matrix,

$$\Gamma_k = \begin{pmatrix} 0 & \delta_1\mathcal{R} & \delta_2\mathcal{R} & \delta_3\mathcal{R} & \delta_4\mathcal{R} \\ 0 & 0 & \delta_2\mathcal{R} & \delta_3\mathcal{R} & \delta_4\mathcal{R} \\ 0 & 0 & 0 & \delta_3\mathcal{R} & \delta_4\mathcal{R} \\ 0 & 0 & 0 & 0 & \delta_4\mathcal{R} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

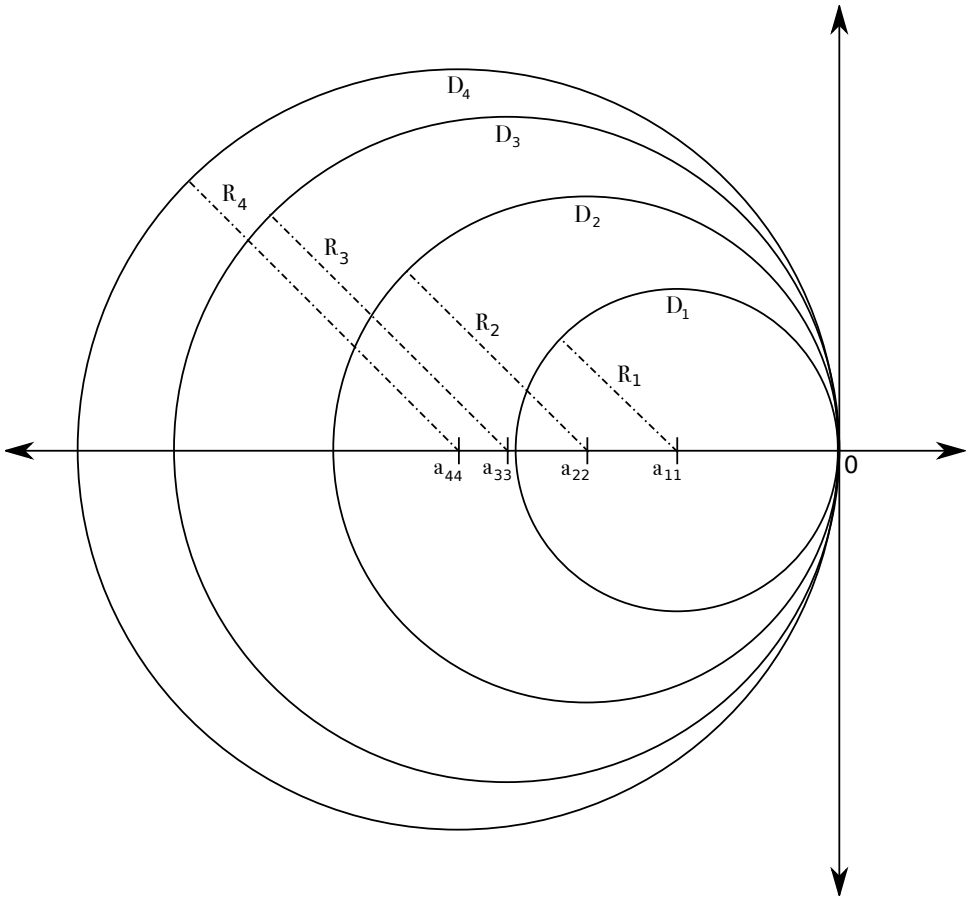


Figure 5.8: Cartoon illustration of the Gershgorin circle theorem as applied to transition rate matrices. All entries of such matrices must be positive except for those along the diagonal, the matrices are symmetric, and all rows and columns sum to zero. As a result, the Gershgorin disks in the complex plane overlap maximally.

Here, each instance of $\mathcal{R}$ is a random number drawn from a uniform distribution on the interval $[0, 1]$. We symmetrize this matrix,

$$\Gamma_k \rightarrow \Gamma_k + {\Gamma_k}^{\mathsf{T}}$$

The diagonal elements are then added so that the rows sum to zero:

$$\Gamma_k \rightarrow \Gamma_k - \mathrm{diag}\{\Gamma_k \cdot \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}\}$$

Once the matrix is constructed, we calculate the eigenvalues. In general, the circle theorem says that these eigenvalues will be complex and will therefore generate unitary evolution in addition to the purely dephasing evolution we desire. To avoid such issues, we discard matrices whose eigenvalues do not satisfy $|\arg(\lambda) - \pi| < \epsilon$. Of those that remain, we choose the one that minimizes the least-squares distance to the target eigenvalues. The closest fit is designated, Γ_{opt}.

Having found a q-matrix that performs well, we still have to construct find the noise amplitudes, $\{\eta_i\}$, so that the noise is unbiased, *ie.*, $\langle\eta\rangle = \sum_i \eta_i/N = 0$. To do this, we diagonalize $\Gamma_{\mathrm{opt}} = S^{\mathsf{T}}\Lambda S$. Then

$$\vec{\eta} = S^{\mathsf{T}}\vec{\chi}$$

The first element of $\vec{\chi}$, χ_0, is free because it only affects zero-frequency noise. The remaining elements of $\vec{\chi}$ can be chosen to be positive or negative, as only their square enters into the spectral-density. The ambiguity is resolved by demanding that $\sum_i \eta_i = 0$ and choosing the signs so that χ_0 is as small as possible, thereby minimizing the zero-frequency noise.

Chapter 6

Qubits as spectrometers of dephasing noise

6.1 Introduction

A key step in the design of a quantum information processing device is gaining a quantitative understanding of the decoherence-inducing noise processes present in the system under study. Knowledge of the statistical properties of this noise both informs and constrains the theoretical models of the system, further aiding in the design process. Some technologies allow for direct measurement of the noise. In the superconducting flux qubit, for example, this is as simple as measuring the stray magnetic field with a SQUID. Frequently, however, the noise acting on a quantum system is inaccessible to such direct measurements and one is left to describe the noise only indirectly through its effects on measurable quantities, such as a qubit decoherence rate.

In 2002 Schoelkopf et al. showed that a single qubit is a tremendously valuable resource for measuring the power spectrum of bit-flip noise [97]. In many implementations, however, dephasing is the dominant decoherence mechanism [98] and the techniques described in [97] are not applicable. In this chapter we demonstrate that a controllable single qubit *may* be used as a spectrometer of dephasing noise. We present a procedure for estimating the correlation function of such noise, and we show that this procedure may be enhanced through the use of dynamical decoupling pulses sequences. These pulse sequences, consisting of repeated π-pulses, can be used to extend the coherence of a qubit by several orders of magnitude [99], thus widening the range over which the correlation function may be estimated. While our technique may be adapted to estimate the noise power spectral density as in Appendix. 6.7.3, we choose to work in the correlation function picture because our reconstruction is local in the time-domain.

6.2 Model

To illustrate our procedure, we consider a single qubit subjected to a classical, dephasing noise source as described by the Hamiltonian,

$$H(t) = \frac{1}{2}\vec{a}(t) \cdot \vec{\sigma} + \frac{1}{2}\left(\eta_0 + \eta(t)\right)\sigma_z. \tag{6.1}$$

Such a Hamiltonian could arise, for instance, for the spin degree of freedom of an electron in a fluctuating magnetic field. Here $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ are the Pauli matrices and $\vec{a}(t)$ is a control field. For later convenience, we have separated the second term into a constant offset field, η_0, and a zero-mean stochastic process, $\eta(t)$. We assume that this qubit can be initialized into an arbitrary pure state and measured in any basis. We will additionally assume the noise process to be wide-sense stationary, allowing us to write the correlation function as a function of only a single time,

$$C(t) = \langle(\eta_0 + \eta(t))(\eta_0 + \eta(0))\rangle = \eta_0^2 + C_\eta(t).$$

Here we have separated the full correlation function into its offset contribution, η_0^2, and stochastic contribution, $C_\eta(t) = \langle\eta(t)\eta(0)\rangle$.

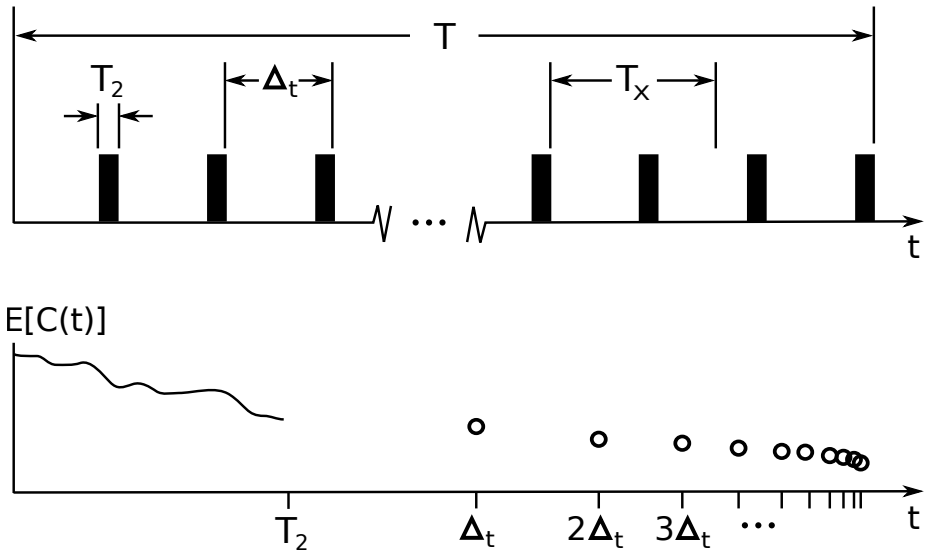


Figure 6.1: (a) Illustration of timescales involved in noise correlation function measurements. Times runs along the horizontal axis and individual measurements are indicated by the tall, black rectangles. Each measurement lasts a time approximately equal to T_2. Initialization steps cause a delay between measurements, Δ_t. Measurements are repeated until time, T. Also illustrated is T_X, one of the many timescales at which correlations are inaccessible to measurement. (b) Cartoon reconstruction of correlation function. A continuous estimate is made by inversion of dynamical decoupling sequences for times shorter than T_2. The correlation is also estimated at times equal to integer multiples of the measurement time, Δ_t. Time axis is scaled logarithmically for clarity.

Experimental constraints introduce several timescales which will limit what may be learned about the statistical properties of the noise. The longest of these timescales is the total length of time, T, that the experiment is performed. Correlations long compared to this time cannot affect any measurements performed and therefore cannot be investigated. Additionally, the state of the qubit must be reinitialized between each measurement, a process which will lead to a maximum repetition rate of one experiment per time, Δ_t. Finally, through the application of control, the coherence of the qubit may be extended to a maximum time, $T_2 \ll \Delta_t$, beyond which the coherence has decayed to a point where it is no longer measurably different from zero.

Taken together, these limitations provide a natural separation of the problem into short- and long- timescales. In the following sections we will construct a method for estimating the correlation function on times short compared to the maximum coherence time, T_2, and a second method for estimating correlations at times long compared to the measurement separation time, Δ_t. Correlations on timescales intermediate between T_2 and Δ_t are inaccessible to measurement, as illustrated in Fig. 6.1.

6.3 Long-time correlations

To estimate the correlation function on timescales longer than the inter-measurement time, Δ_t, we propose a rapid sequence of free evolution measurements in which the qubit is

1. initialized in the +1 eigenstate of σ_x,
2. allowed to evolve for a short time δ_t,
3. measured in the basis of eigenstates of σ_y.

We will assume that the j^{th} run of this experiment begins at time, $t_j = j\Delta_t$. During step (2) above, the qubit acquires a relative phase,

$$\phi_j = \int_{t_j}^{t_j+\delta_t} (\eta_0 + \eta(s))\, ds$$

By measuring in the σ_y basis, the measurement probabilities are,

$$P_y^{\pm}(j) = \frac{1}{2}\left(1 \mp \sin\phi_j\right) \simeq \frac{1}{2}\left(1 \mp \phi_j\right), \tag{6.2}$$

Here we have taken a small angle approximation, which is valid if the total evolution time is chosen sufficiently small. Notice that if we were to measure in σ_x, the measurement probabilities would then depend quadratically on the phase, $P_x^+(j) = \cos^2\phi_j$ and $P_x^-(j) = \sin^2\phi_j$. These probabilities are independent of the sign of the acquired phase and are therefore not useful for determining the long correlations [1]. The probabilities in Eq. (6.2) are approximately linear in the accumulated phase and may be used to estimate the noise correlation function as follows.

Taking a large number, N, of repetitions of the above procedure will yield N measurement results. We put these results in a vector, $\vec{r}$, with $r_j = \pm 1$ the result of the j^{th} experiment. As shown in Appendix 6.7.1, this result vector may be used to estimate the full noise correlation function at times $t_k = k\Delta_t$

$$E[C(k\Delta_t)] = \frac{1}{\delta_t^2(N-k)} \sum_i^{N-k} r_i r_{i+k}. \tag{6.3}$$

The expected error of this estimate will scale inversely with number of experiments performed for each interval,

$$\text{var}\left(E[C(t_k)]\right) \propto \frac{1}{N-k}.$$

So for a given k, this error can be reduced by increasing N, the total number of experiments.

[1] Repeated measurements in σ_x are able to give information about the zero-time correlation, $C(0) = \langle \eta^2 \rangle$. By making a series of σ_x measurements and recording the results in a vector, $\vec{r}$, the variance of η is $\langle \eta^2 \rangle = (1 - \langle r \rangle)/(2\delta_t^2)$

6.4 Short-time correlations

Noise correlations on timescales between $t = 0$ and $t = T_2$ cannot be investigated by the methods discussed in the previous section, as these timescales are shorter than the minimum time between measurements. Instead we propose to bootstrap the techniques of dynamical decoupling (DD) which allow us to measure the overlap integral of the correlation function with a filter function that is defined simply in terms of the applied control field.

To see this, we begin by transforming the Hamiltonian (6.1) into an interaction picture which removes the explicit dependence on the control fields,

$$H_T(t) = \frac{1}{2}\left(\eta_0 + \eta(t)\right) U_a^\dagger(t)\sigma_z U_a(t).$$

Here $U_a(t) = e^{i\int_0^t \vec{a}(s)\cdot\vec{\sigma}ds/2}$ the unitary operator due to only to control. Dynamical decoupling procedures typically limit the control fields to π-pulses polarized along σ_x. The Hamiltonian then remains proportional to σ_z,

$$H_T(t) = \frac{1}{2}y(t)\left(\eta_0 + \eta(t)\right)\sigma_z \tag{6.4}$$

Eq. 6.4 introduces the pulse function, $y(t)$, defined as

$$y(t) = \begin{cases} 1 & \text{after even number of } \pi\text{-pulses,} \\ -1 & \text{after odd number of } \pi\text{-pulses.} \end{cases}$$

The pulse function describes that, from the point of view of the qubit, each π-pulse acts to flip the sign of the noise. When averaged over all possible noise trajectories, $\eta(t)$, evolution under this Hamiltonian results in dephasing, which may be quantified by the decay of the expectation value of the coherence, $\sigma_+ = (\sigma_x + i\sigma_y)/2$.

$$\begin{aligned}
\langle\langle\sigma_+(t)\rangle\rangle &= \left\langle \mathrm{Tr}\,\left(e^{i\int_0^t H_T(s)ds}\,\sigma_+\, e^{-i\int_0^t H_T(s)ds}\,\rho_0\right)\right\rangle \\
&= \left\langle \exp\left(i\int_0^t \left(\eta_0+\eta(s)\right)y(s)\,ds\right)\right\rangle \mathrm{Tr}\left(\sigma_+\rho_0\right) \\
&= \exp\left(-\sum_l \chi^{(l)}(t)\right)\mathrm{Tr}\left(\sigma_+\rho_0\right)
\end{aligned} \tag{6.5}$$

We have indicated averages over random variables with the single expectation value, $\langle\cdot\rangle$, while the double expectation, $\langle\langle\cdot\rangle\rangle$, represents both the quantum average as well as a stochastic average over all consistent trajectories of the noise term. The expression above relies on a cumulant expansion [100]. For our purposes, we will consider experiments in which the qubit is initialized into the $+1$ eigenstate of σ_x, a pulse sequence is applied, and the coherence is measured at time, τ. As such, $\mathrm{Tr}\left(\sigma_+\rho_0\right) = 1$, and we will drop the explicit dependance on t

from the cumulant expansion, assuming that it is evaluated at $t = \tau$. The first term in this series, the 0^{th}-order cumulant, is

$$\chi^{(0)} = -i\eta_0 \int_0^\tau y(t_1) dt_1.$$

This term, which is purely imaginary, represents the coherent precession of the qubit due to the offset field. The 0-order cumulant vanishes if we select any of the myriad ırefocusing pulse sequences for which $\int_0^\tau y(s)ds = 0$. The next term in the expansion vanishes,

$$\chi^{(1)} = -i \int_0^\tau \langle \eta(t_1) \rangle \, y(t_1) dt_1 = 0,$$

because the stochastic term, $\eta(t)$, has zero mean by construction. In fact, all odd-order cumulants are purely imaginary and will vanish provided the unconditional probability, $P(\eta(t))$, is symmetric. All even-order cumulants, $l = 2$ and greater, are purely real, resulting in decay of the coherence and are responsible for dephasing of the qubit. The $l = 2$ cumulant is the dominant decoherence causing term in the expansion (6.5),

$$\chi^{(2)}(t) = \int_0^\tau dt_1 \int_0^\tau dt_2 \, \langle \eta(t_1)\eta(t_2) \rangle \, y(t_1) y(t_2) \tag{6.6}$$

The assumption of wide-sense stationarity we made earlier allows us to relate the above expression to the stochastic part of the noise correlation function, $C_\eta(t_2 - t_1) = \langle \eta(t_1)\eta(t_2) \rangle$. By changing variables from t_1, t_2 to $u = t_2 - t_1, v = t_2 + t_1$, Eq. (6.6) may be expressed as a single-variate integral over the stochastic part of the correlation function,

$$\chi^{(2)} = \int_0^\tau du \, C_\eta(u) \mathcal{F}(u). \tag{6.7}$$

Eq. (6.7) is known as the coherence integral and defines the correlation filter function (CFF),

$$\mathcal{F}(u) = \int_u^{2\tau - u} dv \, y\left(\frac{v+u}{2}\right) y\left(\frac{v-u}{2}\right).$$

The CFF specifies the regions of the correlation function contribute to dephasing under a particular pulse sequence and is only defined for $u \in [0, \tau]$. Example correlation filter functions are shown in Fig. 6.2.

We note that above discussion is usually given in terms of the spectral density [101], the Fourier transform of the correlation function [102]. However, formulation in terms of the time-domain correlation function allows for consistency with our method of elucidating long-time correlations, as given in the previous section.

By taking the absolute value of Eq. 6.5, we can remove the dependance on the 0-order cumulant (because it is purely imaginary),

$$|\langle \langle \sigma_+(\tau) \rangle \rangle| \simeq \left| \exp\left(-\chi^{(0)} - \chi^{(2)}\right) \right| = \exp\left(-\chi^{(2)}\right) \tag{6.8}$$

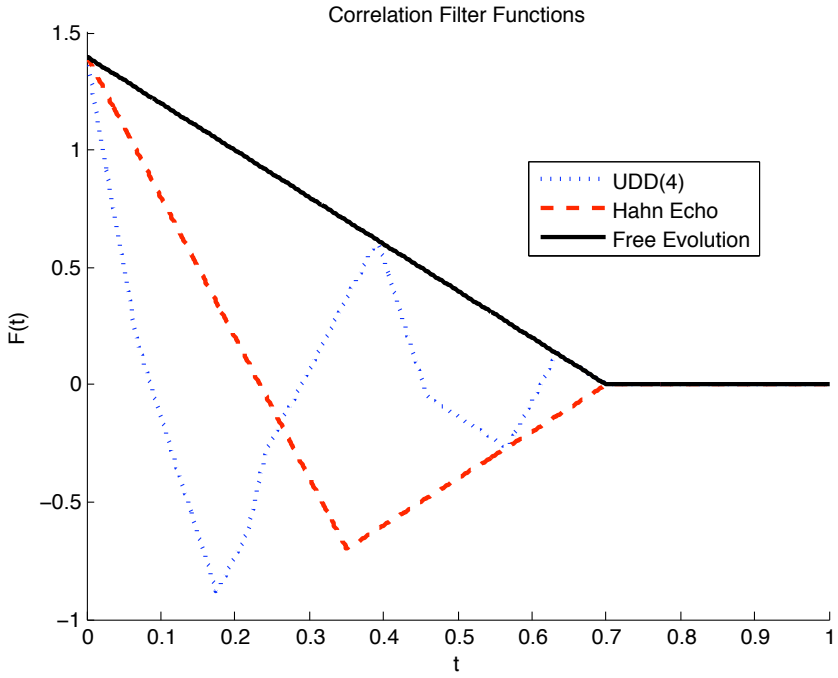


Figure 6.2: (Color online) Example correlation filter functions. i) (black, solid) Free evolution. ii) (red, dashed) Hahn echo. iii) (blue, dotted) 4 pulse Uhrig [82] sequence, UDD(4)

The above equation relates the CFF (calculable from the pulse sequence), the coherence (measurable), and the correlation function (unknown). To reconstruct the correlation function on short-timescales, one must measure the coherence decay for a number of pulse sequences. With each pulse sequence is associated a filter function, $\mathcal{F}_i(t)$, and coherence integral, $\chi_i^{(2)} = \int_0^{\tau_i} C_\eta(t)\mathcal{F}_i(t)dt$. We see from Eq. 6.5, Here $|\langle\langle\sigma_+\rangle\rangle_i|$ is the norm of the coherence measured after the pulse sequence associated with filter function $\mathcal{F}_i(t)$. Because the norm of the coherence is an experimentally accessible quantity, Eq. 6.8 allows us to calculate $\chi_i^{(2)}$ (after repeating the measurement sufficiently many times to gather good statistics). By repeating this for a large number of pulse sequences, we will be left with a set of measured coherence integrals and filter functions and can then use the theory of underdetermined least-squares [103] to regress an estimate, $E[C_\eta(t)]$, of the short-time correlation function as

$$E[C_\eta(t)] = \sum_{ij} \chi_i^{(2)} \mathbb{F}^+_{ij} \mathcal{F}_j(t) \tag{6.9}$$

Here, $\mathbb{F}_{ij} = \int_0^\infty \mathcal{F}_i(t)\mathcal{F}_j(t)dt$, is the filter function overlap matrix. Because the set of equations is under determined (we are trying to reconstruct a continuous function by measuring a finite set of real numbers), there are in general an infinite number of possible correlation functions which reproduce the measured coherences. Our estimate uses the Moore-Penrose pseudoinverse, $\mathbb{F}^+$, which yields the solution with minimal Euclidean norm [103]. A derivation of Eq. 6.9 is given in Appendix. 6.7.5.

Choosing pulse sequences

The particular set of pulse sequences selected will drastically affect the quality of the correlation function estimate, both by dictating the range of time over which the correlation function may be measured and affecting the accuracy of the estimate within that range. For instance, limiting oneself to a series of free evolution experiments will only allow for an estimate the correlation function at very short times. Decoupling sequences greatly extend the average coherence time, facilitating a concomitant extension of the region of the correlation function one may estimate with this procedure.

To quantify this, suppose we perturb the correlation function of the noise, $C_\eta(t) \rightarrow C_\eta(t) + \lambda_\tau \delta(t-\tau)$. This perturbation affects the correlation function estimate by,

$$E[C_\eta(t)] \rightarrow E[C_\eta(t)] + \lambda_\tau \sum_{i,j} \mathcal{F}_i(t) \mathbb{F}^+_{ij} \mathcal{F}_j(s)$$

Taking the variation with λ, we have a measure of the effect of the perturbation,

$$\frac{\delta E[C_\eta(t)]}{\delta \lambda_\tau} = \mathcal{F}_i(\tau) \mathbb{F}^+_{ij} \mathcal{F}_j(t)$$

Squaring this quantity and integrating over t provides us with a scalar measure of the sensitivity of our reconstruction to variation of the correlation function at $t = \tau$,

$$Q(\tau) = \sum_{i,j} \int_0^\infty \left(\mathcal{F}_i(\tau)\mathbb{F}^+_{ij}\mathcal{F}_j(t)\right)^2 dt \tag{6.10}$$

This function depends only on the filter functions, $\mathcal{F}_i(t)$ and the overlap matrix, $\mathbb{F}$. Examination of the quality function for various sets of filter functions has allowed us to see that the estimated correlation function becomes unreliable at times for which $Q(t) < \max(Q(t))/5$.

6.5 Numerical simulations

To demonstrate its efficacy, we apply our procedure to a single qubit dephasing under the action of two mutually uncorrelated random telegraph (RT) fluctuators, using Monte Carlo techniques to simulate a a statistically consistent noise trajectory for each measurement. Each fluctuator is capable of existing in either of two states, $\pm\eta_i$, and will stochastically transition from one state to the other at rate, γ_i. To capture both short- and long-time correlations, we choose a fast, low amplitude fluctuator and a slow, high amplitude fluctuator, $\eta_1 = 1, \gamma_1 = 10, \eta_2 = 10, \gamma_2 = 0.01$, and we have set the offset field, $\eta_0 = 0$. The resulting noise correlation function is $C(t) = \eta_1^2 e^{-2\gamma_1|t|} + \eta_2^2 e^{-2\gamma_2|t|}$, as derived in Appendix. 6.7.2.

We simulated a series of $N = 10000$ measurements, of which the first $N_l = 5000$ were FID experiments to measure long-time correlations, while the last $N_s = 5000$ were DD experiments to measure the short-time correlations. We assumed that the j^{th} experiment began at time $t_j = j\Delta_t$, with $\Delta_t = 4$. For each DD measurement, we simulated a noise trajectory of length $t_{\text{DD}} = 1$. The free induction decay time, $t_{\text{FID}} = 0.04$, was chosen for consistency with the small angle approximation used in Eq. (6.2). The states of the RT fluctuators at the beginning of each trace are determined from their values at the end of the previous trace and the statistics of the noise process: after a time t, the probability of a single RT fluctuator switching an odd number of times is, $(1 - e^{-2\gamma t})/2$. The pulse sequences chosen to investigate the short-time coherences are are given in Table. 6.1. Because of their demonstrated success [104] in extending coherence, we used the Uhrig decoupling sequence [82] as the basis for the experiment.

The results are shown in Fig. 6.3. The reconstructed correlation demonstrates remarkable overlap with the analytically computed spectrum.

6.6 Discussion

By inverting the traditional thinking of dynamical decoupling, we have demonstrated that a single qubit can be an incredibly valuable resource for studying dephasing noise. This noise, pervasive in many quantum systems, is still relatively poorly understood. And because our

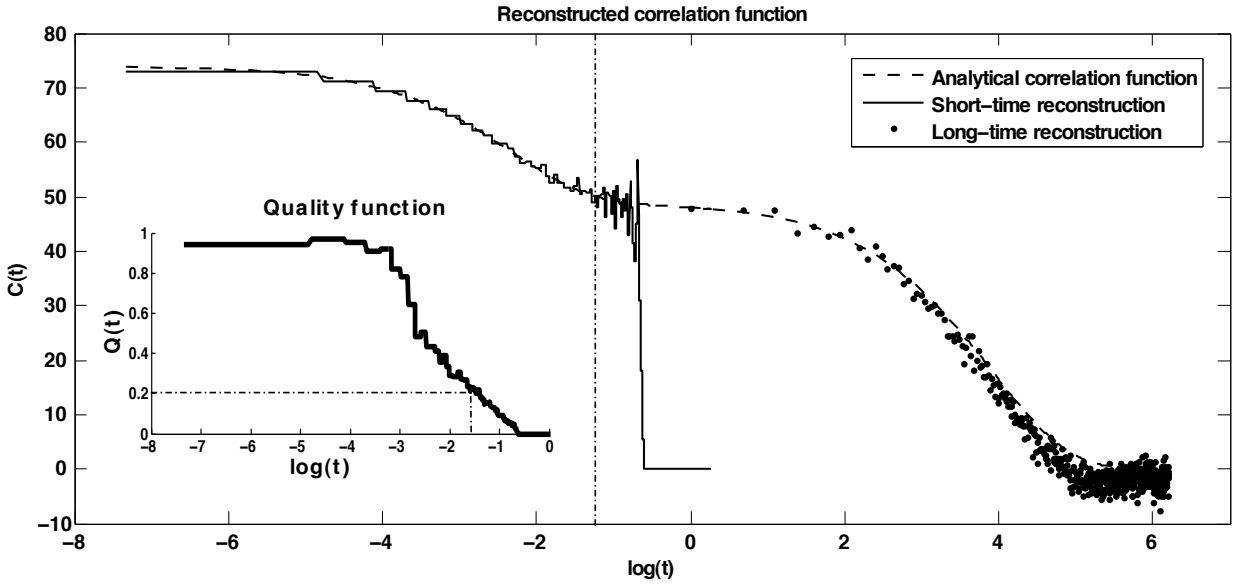


Figure 6.3: Combined reconstruction of the correlation function of two mutually uncorrelated RT fluctuators at both short- and long-times. The dashed line is the analytical correlation function, the solid line is the short-time reconstruction and the black dots represent the long-time correlations. Inset is the quality function for the short-time reconstruction. Dashed-dotted lines demarcate low quality regions. The short-time reconstruction is unreliable at times for which the quality functions, $Q(t) < 0.2$, corresponding to $\log(t) > -1.3$. This unreliable portion is separated by the dot-dashed line in the main figure.

Pulse Sequence	Range	Divisions	Repetitions
FID(1)	0.1-0.5	10	100
UDD(2)	0.1-0.5	10	100
UDD(3)	0.1-0.6	10	100
UDD(4)	0.1-0.7	10	100
UDD(5)	0.1-0.9	10	100

Table 6.1: Pulse sequences used to measure the short-time correlations of the noise discussed in Sec. 6.5. For each pulse sequence and each time step, one hundred noise trajectories are simulated and the coherence decay is calculated. Uhrig DD (UDD) sequences [82] are used for their ability to extend coherence.

method is general, it may be applied to any system for which dephasing is the dominant decoherence mechanism.

As shown in [98], donor qubits in silicon near an oxide interface demonstrate a marked increase in coherence time as the distance from the oxide is increased. Theoretical models of the noise process causing this decoherence suggest that the presence of fluctuating dangling bonds at the interface is responsible for decoherence. However, these models require a dangling bond density which is inconsistent with the measured density. The ability to measure the statistical properties of this noise could aid greatly in the construction of new theoretical models to describe the interplay of donor spins and interfaces.

6.7 Appendices

6.7.1 Estimate of long-time correlation function from free evolution measurements

The correlation function at long-times is sampled by a series of FID measurements, yielding a results vector, $\vec{r}$. We define the correlator of this result vector,

$$C_k = \frac{1}{N-k}\sum_{i}^{N-k} r_i r_{i+k}. \tag{6.11}$$

The expected value of this correlator may be calculated from the measurement probabilities given in Eq. (6.2), as

$$\begin{aligned}\langle C_k\rangle &\simeq \frac{1}{N-k}\sum_{i}\langle r_i r_{i+k}\rangle &(6.12)\\ &\simeq \sum_{i=1}^{N-k}\sum_{m,n=\pm 1}\left\langle \frac{((-1)^m+\phi_i)\,((-1)^n+\phi_{i+k})}{4(N-k)}\right\rangle \\ &= \frac{1}{N-k}\sum_{i=1}^{N-k}\langle\phi_i\phi_{i+k}\rangle &(6.13)\end{aligned}$$

The covariance of the acquired phases may be simplified as,

$$\begin{aligned}\langle\phi_i\phi_{i+k}\rangle &= \int_{t_i}^{t_i+\delta_t} dt_1 \int_{t_{i+k}}^{t_{i+k}+\delta_t} dt_2\,\langle(\eta_0+\eta(t_1))\,(\eta_0+\eta(t_2))\rangle \\ &= \delta_t^2\eta_0^2 + \int_{t_i}^{t_i+\delta_t} dt_1 \int_{t_{i+k}}^{t_{i+k}+\delta_t} dt_2\,\langle\eta(t_1)\eta(t_2)\rangle \\ &= \delta_t^2\eta_0^2 + \int_{k\Delta-\delta_t}^{k\Delta+\delta_t} C_\eta(u) f_k(u) du &(6.14)\end{aligned}$$

In the last equality we have changed variables in the integral from t_1 and t_2 to $v = t_2 + t_1$ and $u = t_2 - t_1$ and then integrated over v. The filter function, $f(u)$, that appears in integral is defined as,

$$F_k(u) = \begin{cases} \sqrt{2}(u - k\Delta_t + \delta_t) & u \in [k\Delta_t - \delta_t, k\Delta_t] \\ \sqrt{2}(1 - u + k\Delta_t) & u \in [k\Delta_t, k\Delta_t + \delta_t] \\ 0 & \text{otherwise}\end{cases}$$

However, from small δ_t, we can assume that $C_\eta(t)$ is constant over the range $t \in [k\Delta_t - \delta_t, k\Delta_t + \delta_t]$ and so comes out of the integral. We can now rewrite the correlation function as

$$\langle\phi_i\phi_{i+k}\rangle = \delta_t^2\eta_0^2 + \delta_t^2 C_\eta(k\Delta_t)$$

Combining this with Eq. 6.13, we see that

$$\langle C_k \rangle = \delta_t^2 \eta_0^2 + \delta_t^2 C_\eta(k\Delta_t)$$

Because the best estimate of $\langle C_k \rangle$ is the sample correlation, C_k, given in Eq. 6.11, we are left with

$$\delta_t^2 \eta_0^2 + \delta_t^2 C_\eta(k\Delta_t) \simeq \frac{1}{N-k} \sum_i^{N-k} r_i r_{i+k}$$

This may be solved for the full correlation function, $C(k\Delta_t)$, as

$$C(k\Delta_t) \simeq \frac{1}{\delta_t^2 (N-k)} \sum_i^{N-k} r_i r_{i+k}$$

In the very-long-time limit we expect that the stochastic part becomes completely uncorrelated,

$$\lim_{t\to\infty} C_\eta(t) \equiv \lim_{t\to\infty} \langle \eta(t)\eta(0) \rangle = 0.$$

The stochastic part of the correlation function may thus be recovered from the full correlation function by subtracting

$$C_\eta(t) = C(t) - \lim_{t\to\infty} C(t)$$

6.7.2 Calculation of the correlation function of a random telegraph fluctuator

The random telegraph fluctuator is defined as a classical stochastic process taking one of two values, $\pm\eta$, with a flipping rate, γ. The probability of being in the state $\pm\eta$ at time t is given by $p_\pm(t)$. Defining the vector $\vec{p}(t) = (p_+(t), p_-(t))$, the time evolution of the probabilities may be written as

$$\frac{d\vec{p}(t)}{dt} = \begin{pmatrix} -\gamma & \gamma \\ \gamma & -\gamma \end{pmatrix} \vec{p}(t) \equiv \Gamma \vec{p}(t)$$

This equation defines the transition rate matrix, Γ. From this, the time evolution of the probability vector may be solved as,

$$\vec{p}(t) = e^{\Gamma t} \vec{p}(0).$$

The correlation function for the noise may then be calculated to be

$$\begin{aligned} C(t) &= \sum_{i,j} \eta_i P(\eta_i, t|\eta_j, 0) \eta_j P(\eta_j, 0) \\ &= \frac{1}{2} \vec{\eta} \cdot e^{\Gamma t} \cdot \vec{\eta}, \end{aligned}$$

which, for the transition rate matrix given above, simplifies to,

$$C(t) = \eta^2 e^{-2\gamma|t|}.$$

Because the covariance vanishes, the correlation function of multiple, uncorrelated RT fluctuators is simply the sum of the correlation function for each fluctuator individually.

$$C_{ij}(t) = \langle \eta_i(t)\eta_j(0)\rangle = \langle \eta_i(t)\rangle \langle \eta_j(0)\rangle = 0$$

Alternative calculation of RT correlation function

In an infinitesimal time, δ_t, the transition probabilities in the i^{th} fluctuator are approximately linear in the rate,

$$\begin{aligned} P_{\text{flip}} &= \gamma\delta_t \\ P_{\text{no flip}} &= 1 - \gamma\delta_t \end{aligned}$$

Therefore, the probability of n transitions in a time interval, Δ_t is given as

$$\begin{aligned} P_n(\Delta_t) &= \lim_{N\to\infty} \left(1 - \frac{\gamma\Delta_t}{N}\right)^{N-n} \left(\frac{\gamma\Delta_t}{N}\right)^n \frac{N!}{n!(N-n)!} \\ &= e^{-\gamma\Delta_t} \frac{(\gamma\Delta_t)^n}{n!} \end{aligned}$$

The last term in the first expression above is a combinatorial factor. From this we see that the probability of an even number of transitions (which would leave the state unchanged) is

$$P_e(\Delta_t) = \sum_{n \text{ even}} P_n(\Delta_t) = \frac{1}{2}\left(1 + e^{-2\gamma\Delta_t}\right)$$

While the probability of an odd number of transitions is

$$P_o(\Delta_t) = \sum_{n \text{ odd}} P_n(\Delta_t) = \frac{1}{2}\left(1 - e^{-2\gamma\Delta_t}\right)$$

The correlation function is then

$$\begin{aligned} C(t) &= \sum_{i,j} \eta_i P(\eta_i @ t | \eta_j @ 0) \eta_j P(\eta_j @ 0) \\ &= \frac{\eta^2}{2}(2P_e(t) - 2P_o(t)) \\ &= \eta^2 e^{-2\gamma|t|} \end{aligned}$$

6.7.3 Frequency domain filter functions

The discussion in the main text for the short-time correlations may be instead expressed in terms of the spectral density. Recall the $l = 2$ cumulant is

$$\begin{aligned}\chi^{(2)} &= \left\langle \int_0^t \eta(t_1)y(t_1)dt_1 \int_0^t \eta(t_2)y(t_2)dt_2 \right\rangle \\ &= \int_0^t dt_1 \int_0^t dt_2 \left\langle \eta(t_1 - t_2)\eta(0) \right\rangle y(t_1)y(t_2)\end{aligned}$$

Using the Wiener-Khintchine theorem,

$$C(t) = \int \frac{d\omega}{2\pi} e^{i\omega t} S(\omega),$$

can rewrite the correlation function in terms of the power spectrum. This gives

$$\begin{aligned}\chi^{(2)} &= \int \frac{d\omega}{2\pi} S(\omega) \left| \int_0^t e^{i\omega\tau} y(\tau) d\tau \right|^2 \\ &\equiv \int \frac{d\omega}{2\pi} S(\omega) F(\omega; t) \end{aligned} \tag{6.15}$$

The last line above defines the filter function as the square of the Fourier transformed pulse function,

$$\begin{aligned}F(\omega; t) &= \left| \int_0^t e^{i\omega\tau} y(\tau)\, d\tau \right|^2 \\ &= \left| 1 + (-1)^{N+1} e^{i\omega t} + 2\sum_{j=1}^{N} (-1)^j e^{i\omega\Delta_j t} \right|^2\end{aligned}$$

The filter functions indicate the range of frequencies of the noise power spectrum which contribute to dephasing. Example filter functions for the Carr-Purcell sequence are shown in Fig. 6.4. Furthermore, the effects of nonzero pulse widths may be included to first-order through the modification,

$$\begin{aligned}y(t) &= (-1)^{\int_0^t a(t')dt'/\pi} \\ &= \begin{cases} 1 & \text{after even number of } \pi\text{-pulses,} \\ 0 & \text{during application of } \pi\text{-pulses,} \\ -1 & \text{after odd number of } \pi\text{-pulses.} \end{cases}\end{aligned}$$

Including this modification to the pulse function changes the the filter function expression to,

$$F(\omega; t) = \left| 1 + (-1)^{N+1} e^{i\omega t} + 2\sum_{j=1}^{N} (-1)^j e^{i\omega\Delta_j t} \cos\left(\omega t_\pi/2\right) \right|^2$$

We now illustrate our procedure in the spectral density formalism. We suppose an approximately Ohmic power spectrum with a high-frequency cutoff and an additional high frequency peak, as shown in Fig. 6.5. The filter functions chosen are 59 different Carr-Purcell (CP(10)) trains with total durations, τ, chosen to such that the their most sensitive frequencies are uniformly distributed from $\omega = 1$ to $\omega = 30$. The reconstruction in terms

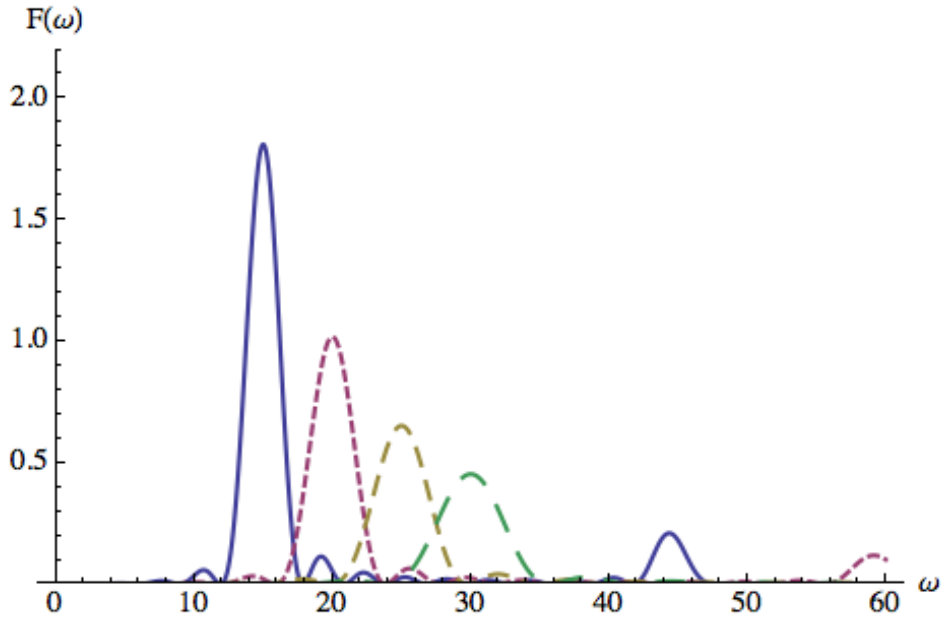


Figure 6.4: Representative filter functions of those used to generate Fig. 6.5. Function maxima are uniformly distributed over $\omega = 1$ to $\omega = 30$ in increments of 0.5 in calculation, though for clarity only 4 of the 59 filters are shown.

of these filter functions is also shown in Fig. 6.5, and is in close agreement to the actual spectrum. We have additionally simulated what would happen in the case of errors in the measurement. If there is a 5% error in measurement, you see the results as the yellow dashed line in Fig. 6.5. This corresponds roughly to a 1% error in the measurement of $\langle \sigma_+(t) \rangle$. An increase in the number of filters used will increase the tolerance of the estimator to measurement error.

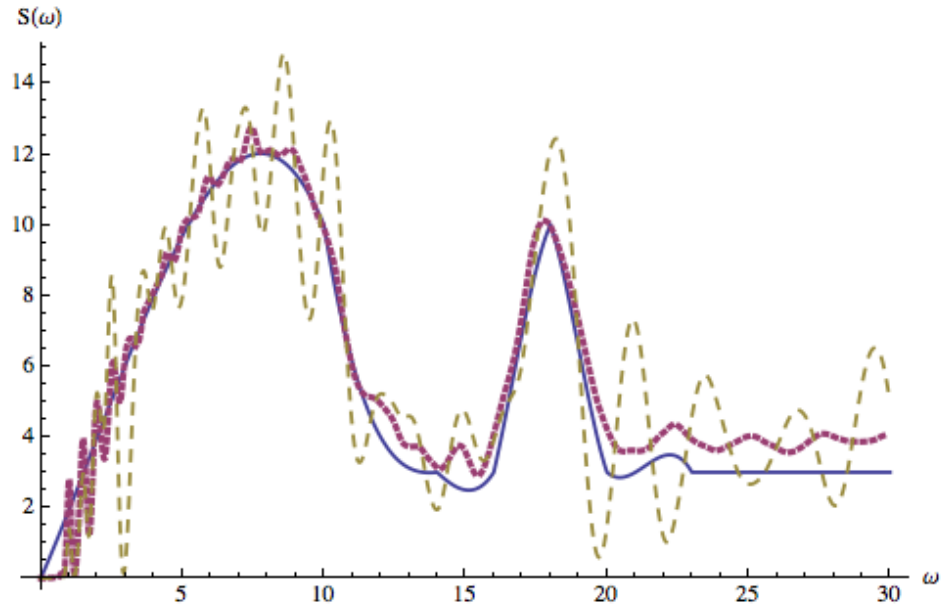


Figure 6.5: The trial power spectrum (blue, solid), the estimate of the noise power spectrum (red, dotted) with perfect measurements, and a sample estimate (yellow, dashed) of the power spectrum in the case of 5% random error on the measurements $\vec{\chi}$.

6.7.4 Cumulant expansion of echo envelope

The echo envelope decays as,

$$\begin{aligned}\langle\langle\sigma_+(t)\rangle\rangle &= \left\langle \exp\left(i\int_0^t (\eta_0+\eta(s))\,y(s)ds\right)\right\rangle \\ &= \exp\left(i\Omega\int_0^t y(s)ds\right)\left\langle \exp\left(i\int_0^t \eta(s)y(s)ds\right)\right\rangle \\ &= e^{\chi_0}\left\langle \sum_n \frac{1}{n!}\left(i\int_0^t \eta(s)y(s)ds\right)^n\right\rangle \end{aligned} \tag{6.16}$$

This expectation value is difficult to evaluate, so we seek a cumulant expansion of the form

$$\langle\langle\sigma_+(t)\rangle\rangle = e^{\chi_0}e^{\sum_{m=1}^{\infty}\lambda^m\chi_m}$$

Here λ is a parameter that will be used only to keep track of orders in the expansion. The notation here is, unfortunately, different that that which appears in the main body. This is

to avoid cluttered notation.

$$\begin{aligned} e^{\sum_{m=1}^{\infty}\lambda^m\chi_m} &= 1+\sum_m \lambda^m\chi_m + \frac{1}{2}\sum_{mn}\lambda^{m+n}\chi_m\chi_n \\ &\quad + \frac{1}{6}\sum_{mnp}\lambda^{m+n+p}\chi_m\chi_n\chi_p + \cdots \\ &= 1+\lambda\left(\chi_1\right) + \lambda^2\left(\chi_2+\frac{1}{2}\chi_1^2\right) \\ &\quad + \lambda^3\left(\chi_3+\frac{1}{2}\left(\chi_1\chi_2+\chi_2\chi_1\right)+\frac{1}{6}\chi_1^3\right)+\cdots \end{aligned} \tag{6.17}$$

We now match terms of order λ^n in eq.(6.17) to terms of order $\eta(t)^n$ in eq.(6.16). Solving for the cumulants, we have:

$$\begin{aligned} \chi_0 &= i\Omega\int_0^t y(s) \\ \chi_1 &= \left\langle i\int_0^t ds\;\eta(s)y(s)\right\rangle \\ \chi_2 &= \frac{1}{2}\left\langle\left(i\int_0^t ds\;\eta(s)y(s)\right)^2\right\rangle - \frac{1}{2}\chi_1^2 \\ \chi_3 &= \frac{1}{6}\left\langle\left(i\int_0^t ds\;\eta(s)y(s)\right)^3\right\rangle - \chi_1\chi_2 - \frac{1}{6}\chi_1^3 \\ \chi_4 &= \frac{1}{24}\left\langle\left(i\int_0^t ds\;\eta(s)y(s)\right)^4\right\rangle - \chi_1\chi_3 - \frac{1}{2}\left(\chi_2^2+\chi_1^2\chi_2\right) - \frac{1}{24}\chi_1^4 \end{aligned}$$

These results simplify greatly in the case of Gaussian noise. According to Wick's theorem, all odd-order cumulants will vanish, while the even order terms may be calculated, for example,

as

$$\begin{aligned}
\chi_4 &= \frac{1}{24}\left\langle \left(i\int_0^t ds\eta(s)y(s)\right)^4 \right\rangle - \frac{1}{2}\chi_2^2 \\
&= \frac{1}{24}\iiiint_0^t dt^4 \left\langle \eta(t_1)\eta(t_2)\eta(t_3)\eta(t_4)\right\rangle y(t_1)y(t_2)y(t_3)y(t_4) - \frac{1}{2}\chi_2^2 \\
&= \frac{1}{24}\left(\iint dt^2 \left\langle \eta(t_1)\eta(t_2)\right\rangle y(t_1)y(t_2) \iint dt^2 \left\langle \eta(t_3)\eta(t_4)\right\rangle y(t_3)y(t_4) \right. \\
&\qquad + \iint dt^2 \left\langle \eta(t_1)\eta(t_3)\right\rangle y(t_1)y(t_3) \iint dt^2 \left\langle \eta(t_2)\eta(t_4)\right\rangle y(t_2)y(t_4) \\
&\qquad \left. + \iint dt^2 \left\langle \eta(t_1)\eta(t_4)\right\rangle y(t_1)y(t_4) \iint dt^2 \left\langle \eta(t_2)\eta(t_3)\right\rangle y(t_2)y(t_3) \right) - \frac{1}{2}\chi_2^2 \\
&= \frac{3}{24}\left(\iint dt^2 \left\langle \eta(t_1)\eta(t_4)\right\rangle y(t_1)y(t_4) \right)^2 - \frac{1}{2}\chi_2^2 \\
&= -\chi_2^2
\end{aligned}$$

Where we have used that $\chi_1 = \chi_3 = 0$. Inclusion of higher order moments is not necessary until $\chi_2 \lesssim 1$. This corresponds to an echo decay of approximately $\langle \sigma_+(t)\rangle = 1/\sqrt{e} \approx 0.6$.

6.7.5 Estimation of correlation function by CFFs

Suppose we have chosen a large number of pulse sequences and constructed their associated filter functions, $\mathcal{F}_i(t)$. As described in the main text we have experimental access to the coherence integrals,

$$\chi_i^{(2)} = \int_0^\infty C_\eta(t)\mathcal{F}_i(t)\, dt.$$

Because the filter functions are known in terms of the applied pulse sequences, we can use this integral to describe correlation function, $C_\eta(t)$. If our set of filter functions were orthonormal, it would be a trivial task to expand the correlation function as a weighted sum of correlation filter functions, much like a Fourier series expansion. However, this is not the case, so we instead construct a new set of orthonormal functions. This can be done via the Gram-Schmidt orthogonalization procedure to yield the set,

$$f_i(t) = \sum_j c_{ij}\mathcal{F}_j(t). \tag{6.18}$$

Properly normalized, these functions, $f_i(t)$, are orthogonal under the inner product,

$$\langle f_i, f_j\rangle = \int_0^\infty f_i(t)f_j(t)\, dt = \delta_{ij}$$

We can now expand the stochastic part of correlation function, $C_\eta(t)$, as

$$\begin{aligned} C_\eta(t) &\simeq \sum_i \langle f_i(t), C_\eta(t)\rangle \, f_i(t) \\ &= \sum_{i,j} c_{ij} \langle \mathcal{F}_j(t), C_\eta(t)\rangle \, f_i(t) \\ &= \sum_{i,j} c_{ij} \chi_j^{(2)} \sum_k c_{ik} \mathcal{F}_k(t) \\ &= \vec{\chi}^{\mathsf{T}} \cdot \mathbf{c}^{\mathsf{T}} \mathbf{c} \cdot \vec{\mathcal{F}}(t) \end{aligned} \tag{6.19}$$

Here, $\vec{\chi}$ and $\vec{\mathcal{F}}$ are the vectors of measurement outcomes and filter functions, respectively, and $\mathbf{c}$ is the matrix of expansion coefficients from (6.18). We can determine the matrix $\mathbf{c}^{\mathsf{T}}\mathbf{c}$ by examining the orthogonalized filter functions,

$$\begin{aligned} \langle f_i(t), f_j(t)\rangle &= \sum_{m,n} \langle c_{im}\mathcal{F}_m(t), c_{jn}\mathcal{F}_n(t)\rangle \\ &= \sum_{m,n} c_{im}c_{jn} \langle \mathcal{F}_m(t), \mathcal{F}_n(t)\rangle \\ &= \delta_{ij}. \end{aligned}$$

These last two lines may be cast as a matrix equation, $\mathbb{1} = \mathbf{c}\mathbb{F}\mathbf{c}^{\mathsf{T}}$. Here $\mathbb{F}_{ij} = \langle \mathcal{F}_i, \mathcal{F}_j\rangle$ is the filter overlap matrix and $\mathbb{1}$ is the identity matrix. This implies, $\mathbf{c}^{\mathsf{T}}\mathbf{c} = \mathbb{F}^{-1}$, and we can now rewrite (6.19) as,

$$C_\eta(t) \simeq \vec{\chi}^{\mathsf{T}} \cdot \mathbb{F}^{-1} \cdot \vec{\mathcal{F}}(t) \tag{6.20}$$

In reality one will only be able to perform a finite number of experiments, so this expansion is only approximate (much as a finite Fourier-expansion is only an approximation of the expanded function). We point out that we are able to express our estimate power spectrum entirely in terms of the filter functions, their overlap integrals, and experimentally accessible parameters. It is not necessary to explicitly construct the orthogonalized filter functions defined in (6.18), and we avoid the numerical errors commonly associated [105] with the Gram-Schmidt procedure.

Part IV

Quantum simulation

Chapter 7

Quantum simulation in neutral-atom optical lattices

7.1 Introduction

Among the most exciting aspects of quantum simulation is the possibility of generating and studying exotic quantum phases. This includes systems whose ground states posses topological order which can be used to robustly store and process quantum information. The Hamiltonians governing these phases frequently require more-than-2-body interactions that are hard or even impossible to realize naturally. This difficulty has spurred much theoretical and experimental effort in the artificial engineering of Hamiltonians, particularly for trapped ions [106] and trapped neutral atoms [107]. Many proposals have been made for the generation of 2-body Hamiltonians using static emulation schemes and some experimental realizations have appeared [108, 109, 110]. A small number of specific proposals have appeared for generating n-body interactions [111], but systematic efforts to generate n-body interactions have focused on static emulation.

We present here an alternative, dynamic emulation approach to systematic generation of n-body interactions that is based on sequences of control pulses which individually realize 1- and 2-body operations on internal atomic levels. We show that this stroboscopic realization of the Hamiltonian can be implemented simultaneously with a dissipative thermalization or cooling protocol to stabilize the system from the effects of imperfect quantum operations and environmental noise. In the zero temperature limit, the thermalization protocol can be viewed as replacing algorithmic quantum error correcting schemes in an equivalent quantum circuit model with a driven dissipative procedure to remove excitations and return the system to the ground state. We note that the resource requirements for this thermalization protocol are different, and scale differently with system size, than algorithmic error correction, and may be more accessible to experiment in the foreseeable future.

We illustrate the approach here with stroboscopic generation of the 4-body toric code Hamiltonian, which constitutes one of the simplest exactly solvable models with a ground state topological phase [112]:

$$H_0^{TC} = -J_e \sum_v \prod_{j \in v} \sigma_j^z - J_m \sum_p \prod_{j \in p} \sigma_j^x, \tag{7.1}$$

where σ_j denotes a Pauli operator on the links of a square lattice and v/p denote the vertex/plaquette of the lattice. The ground state of this model possesses topological order, and therefore has anyonic quasiparticle excitations and, on a lattice with periodic boundary conditions, an emergent topological degeneracy. Quantum information can be encoded in this ground state degeneracy and manipulated with controlled creation and braiding of anyons [112, 113]. In a finite sized system [114, 115], the topological order of the ground state and gap to excited states protects against decoherence and loss of quantum information due to noise *provided the system is coupled to a low temperature bath.* Our analysis below will provide a scheme for generating both H_0^{TC} and an effective low temperature bath, realizing the topological protection characteristic of the toric code.

The physical context for our analysis is a set of ~ 250 individual ^{133}Cs atoms trapped at the sites of an addressable simple cubic optical lattice [116]. The orbital degrees of freedom

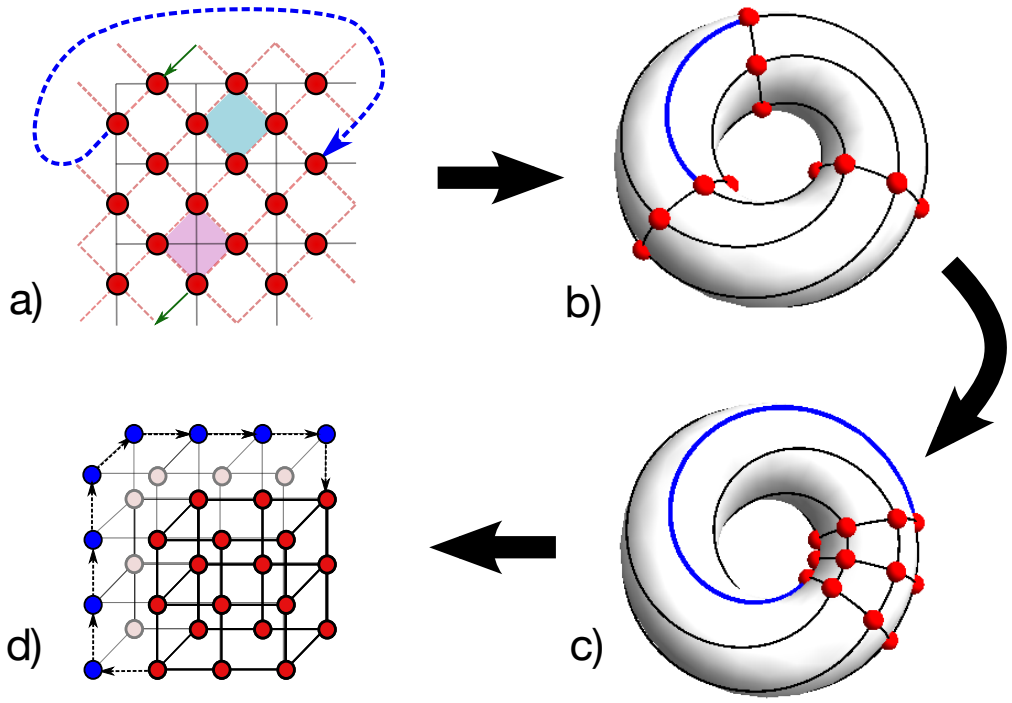


Figure 7.1: (Color online) One of several possible embeddings of the toric code geometry onto a three dimensional array of atoms. Twisted-periodic boundary conditions can be imposed by SWAP-gate shuttling along auxilliary sites, indicated by blue lines. Bold lines connect logical nearest-neighbors. [121]

are frozen on the time scales relevant to our analysis and we need consider only internal atomic degrees of freedom. Two hyperfine levels (e.g., $|F, m_F\rangle = |4,4\rangle, |3,3\rangle$) define a 2-level pseudospin system. We realize H_0^{TC} in the interaction representation defined by the pseudospin energies. Auxiliary internal levels are used to realize 1-spin and 2-spin quantum operations, using optical frequency Raman pulses to generate arbitrary single-spin operations and excitation of one atom to a Rydberg state, e.g., the $n \approx 80$ state, to generate controlled-phase gates, CPHASE [117, 118]. To achieve thermalization or cooling, the Hamiltonian H_0^{TC} is supplemented by coupling the primary system spins to an ancillary set of pseudospins that will be dissipatively controlled to simulate a thermal reservoir. Since the pseudospins are localized at the sites of a cubic lattice, one can choose to either realize H_0^{TC} on a single plane using a surface code [119, 120] or in a three-dimensional cubic array with toroidal boundary conditions realized by SWAP operations.

7.2 Effective Hamiltonian evolution

In what follows, we assume the ability to perform both Rydberg-induced CPHASE gates between atoms in neighboring sites and arbitrary 1-body rotations, $\exp(-i\theta\sigma_j)$, on individual atoms, where θ is a variable phase angle. Sequences of these one and 2-body unitary operations can be chosen to generate effective *n-body* interactions through high-order terms in the Magnus expansion [122], allowing stroboscopic simulation of a broad class of Hamiltonians.

Consider the operator sequence, $U_n U_{n-1} \cdots U_2 U_1$, where the gates U_j impose either the 1- or 2-body gates described above. Effective interactions are found through:

$$\begin{aligned} H_{\text{eff}}(t) &\equiv \frac{i\hbar}{t} \ln(U_n U_{n-1} \cdots U_2 U_1) \\ &= \sum_j \frac{i\hbar}{t} \ln U_j - \sum_{j<k} \frac{\hbar}{2t} [\ln U_j, \ln U_k] + \mathcal{O}\left(||\ln U||^3\right). \end{aligned}$$

Consider now simulation of the 4-body interactions in H_0^{TC}. We use the notation $U_j(\phi) \equiv e^{-i\phi\Sigma_j}$ and define $\Sigma_1 = \sigma^z\sigma^y\sigma^0\sigma^0$, $\Sigma_2 = \sigma^0\sigma^x\sigma^y\sigma^0$, and $\Sigma_3 = \sigma^0\sigma^0\sigma^x\sigma^z$, where σ^0 is the identity operator. For simplicity, it is assumed that each $U_j(\alpha)$ takes a time τ to execute. We construct the operator sequence,

$$U_{123}(\alpha,\beta,\gamma) = U_{12}(\alpha,\beta)U_3(\gamma)U_{12}^\dagger(\alpha,\beta)U_3^\dagger(\gamma), \tag{7.2}$$

where $U_{12}(\alpha,\beta) = U_2(\beta)U_1(\alpha)U_2^\dagger(\beta)U_1^\dagger(\alpha)$. This sequence acts over a time 10τ to generate the following effective Hamiltonian at a single vertex, v:

$$\begin{aligned} H_{\text{eff}}^{zzzz} =& J_e \left[\sigma^z\sigma^z\sigma^z\sigma^z\right]_v + \frac{\chi}{\alpha\gamma}\left(\alpha\left[\sigma^0\sigma^x\sigma^z\sigma^z\right]_v + \gamma\left[\sigma^z\sigma^z\sigma^y\sigma^0\right]_v\right) \\ &+ \chi\left(\left[\sigma^0\sigma^x\sigma^y\sigma^0\right]_v - 2\beta/\gamma\left[\sigma^0\sigma^0\sigma^x\sigma^z\right]_v\right) + \mathcal{O}\left(\phi^6\right) \end{aligned}$$

where $[O]_{v/p}$ denotes the application of the (up to) four-body operator O to the spins meeting at a vertex v or surrounding a plaquette p,

$$J_e = \frac{\chi}{\alpha\gamma}\left(1 - \frac{2}{3}\left(\alpha^2+\beta^2+\gamma^2\right)\right) + \mathcal{O}(\phi^6)$$

and

$$\chi \equiv \frac{2\hbar}{5\tau}\alpha^2\beta\gamma^2.$$

We have additionally used the notation $|\alpha| = |\beta| = |\gamma| \equiv \phi$. By repeating the operator sequence a second time with sign reversals $\alpha \to -\alpha$ and $\gamma \to -\gamma$, we cancel the fourth order terms in ϕ giving $U_{123}(-\alpha,\beta,-\gamma)U_{123}(\alpha,\beta,\gamma)$, acting for a time 20τ to generate the effective Hamiltonian

$$H_{\text{eff}}^{zzzz} = J_e\left[\sigma^z\sigma^z\sigma^z\sigma^z\right]_v + \chi\left[\sigma^0\sigma^x\sigma^y\sigma^0\right]_v + \mathcal{O}\left(\phi^6\right). \tag{7.3}$$

The sequence $U_{123}(-\alpha,\beta,-\gamma)U_{123}(\alpha,\beta,\gamma)$ is specifically designed to cancel the lowest-order (ϕ^4) perturbation terms without affecting the gap. The remaining ϕ^5 term is a 2-body perturbation to H_0^{TC}. Repeating this sequence with appropriate sign reversals will cancel these higher order terms. However, the ground state subspace of H_0^{TC} is robust to these remaining perturbations (see below). A shorter operator sequence may then be preferable as it is less likely to cause gate errors. The effective plaquette operator, H_{eff}^{xxxx}, can be generated by cyclic permutation of the Pauli operators in the above expressions for Σ_1, Σ_2 and Σ_3. More details on the Magnus expansion approach are considered in Appendix. 7.7.1.

Simulation of H_0^{TC} then requires application of the pulse sequence to all vertices and plaquettes conforming to a two dimensional square lattice with periodic boundary conditions. Vertex and plaquette terms may be applied serially as:

$$\exp\left(-iH_{\text{eff}}^{xxxx}t/\hbar\right)\exp\left(-iH_{\text{eff}}^{zzzz}t/\hbar\right) \approx \exp -i\left(H_{\text{eff}}^{xxxx}+H_{\text{eff}}^{zzzz}\right)t/\hbar.$$

Because only the perturbation terms fail to commute, the truncation error in the above expression occurs at orders larger than ϕ^7. For 18 pseudospins, representing a 3×3 system with toroidal boundary conditions, a completely serial implementation yields a stroboscopic cycle time of 720μs using the estimate $\tau \sim 500$ns [123] and the minimal count of one CPHASE and four 1-spin gates to realize all $U_j(\alpha)$ [124]. This serial stroboscopic cycle time scales approximately linearly with the number of pseudospins in the lattice. This may be reduced in larger lattices by implementing some vertex and plaquette operators in parallel.

7.3 Simulated thermalization

The pseudospin subspace of the system defined by the internal states of the trapped atoms is an open quantum system that will interact with the external environment through the controlled quantum operations in the above pulse sequences and uncontrolled noisy interactions. Unlike a solid state system, for which system-environment interactions generally lead to thermal equilibration, noise in the optical lattice system will not drive the simulation subspace to a state that is thermal under the simulated Hamiltonian [125]. Additionally, noise in the above sequence of control gates will add entropy and effectively heat the system. Using the linear entropy to estimate the entropy production ΔS resulting from imperfect gate operation yields an estimate $\Delta \text{S} \sim \text{EPG}$, where EPG is the error per gate [121]. Quantum circuit models are usually supplemented by error correction schemes to effectively remove entropy from the system. We take a different approach here, stroboscopically constructing an effective system-reservoir interaction to control system entropy and relax the system to the ground state or a thermal state.

To maintain the simulated system at a thermal steady state we add a "thermalization" interaction H_{sr} of the system pseudospins with a set of ancillary pseudospins. In the optical lattice system, these ancillary pseudospins, which may be a 2^{nd} species of atom, are trapped in an offset, intercalated optical lattice, such that each ancillary atom is adjacent to a system

atom. Consider a Hamiltonian with local n-body interactions of the form:

$$H_0 = -\sum_{\nu} J_\nu \sum_{\mathcal{N}} h^\nu_{\mathcal{N}}. \tag{7.4}$$

In Eq. (7.4), $h^\nu_{\mathcal{N}}$ is an n-body operator involving a neighborhood of pseudospins, $\mathcal{N}$, including pseudospin i, with eigenvalues ± 1, ν labels the type of interaction, and J_ν is a constant. Additionally we define the pseudospin flip operator Σ^ν_i such that $\Sigma^\nu_i |h^\nu_{\mathcal{N}} = \pm 1\rangle = |h^\nu_{\mathcal{N}} = \mp 1\rangle$ when $i \in \mathcal{N}$. When all $[h^\nu_{\mathcal{N}}, h^\mu_{\mathcal{N}'}] = 0$, as is the case for Eq. (7.1), we can define:

$$\begin{aligned} E^\dagger_{i,\nu} &= \frac{1}{4}\Sigma^\nu_i \left(\mathbb{1} + h^\nu_{\mathcal{N}}\right)\left(\mathbb{1} + h^\nu_{\mathcal{N}'}\right) \\ T_{i,\nu} &= \frac{1}{4}\Sigma^\nu_i \left(\mathbb{1} - h^\nu_{\mathcal{N}}\right)\left(\mathbb{1} + h^\nu_{\mathcal{N}'}\right). \end{aligned} \tag{7.5}$$

with $i \in \mathcal{N}$ and $i \in \mathcal{N}'$, and where $E^\dagger_{i,\nu}$ and $T_{i,\nu}$ are $(2n-1)$-body interactions; $E^\dagger_{i,\nu}$ creates a pair of excitations about i and $T_{i,\nu}$ translates an excitation about i. The energy gap for creation of a pair of excitations of type ν is $\Delta_\nu = 4J_\nu$.

A route to guaranteeing the thermal equilibration of this system is for it to evolve under the Lindblad master equation $\dot{\rho} = -i/\hbar\,[H_0, \rho] + L\,[\rho]$, where ρ is the density matrix and $L\,[\rho]$ is the superoperator $L\,[\rho] = \sum_\omega \left(2c_\omega \rho c^\dagger_\omega - c^\dagger_\omega c_\omega \rho - \rho c^\dagger_\omega c_\omega\right)$, with $\{c_\omega, c^\dagger_\omega\}$ the Lindblad operators. With the choice

$$\{c_{i,\nu}\} = \left\{\sqrt{\frac{1-p}{2}\lambda^*}E_{i,\nu}, \sqrt{\frac{p}{2}\lambda^*}E^\dagger_{i,\nu}, \sqrt{\frac{\gamma^*}{4}}T_{i,\nu}, \sqrt{\frac{\gamma^*}{4}}T^\dagger_{i,\nu}\right\}, \tag{7.6}$$

the Lindblad master equation describes equilibration with a bath of temperature $T = -\Delta/\ln(p)$. The unique stationary state of the system is then the thermal state under H_0 with temperature T. λ^* and γ^* are relaxation rates, and their values dictate the thermalization time. For simplicity we have set $\Delta_\nu = \Delta$.

To generate evolution under such a master equation, we introduce a set of non-interacting ancillary pseudospins that independently undergo strong dissipation. Each local neighborhood of the system interacts locally with a single thermal ancillary pseudospin $\mathsf{T}_{i,\nu}$ and a single maximally mixed ancillary pseudospin $\mathsf{M}_{i,\nu}$ for each type of local excitation via:

$$H_{\text{sr}} = g\sum_{\nu,i}\left(E^\dagger_{i,\nu}\sigma^-_{\mathsf{T}_{i,\nu}} + T_{i,\nu}\sigma^-_{\mathsf{M}_{i,\nu}} + h.c.\right) \tag{7.7}$$

The master equation of the system and ancilla pseudospins combined is of the above Lindblad form with Lindblad operators:

$$\{c_{i,\nu}\} = \left\{\sqrt{\frac{1-p}{2}\lambda}\sigma^-_{\mathsf{T}_{i,\nu}}, \sqrt{\frac{p}{2}\lambda}\sigma^+_{\mathsf{T}_{i,\nu}}, \sqrt{\frac{\gamma}{4}}\sigma^+_{\mathsf{M}_{i,\nu}}, \sqrt{\frac{\gamma}{4}}\sigma^-_{\mathsf{M}_{i,\nu}}\right\},$$

where λ and γ define the relaxation rates of the individual ancillary pseudospins. With this choice of $H_{\text{sys-res}}$ it can be shown [121] that for $g \ll \lambda$ the system pseudospins evolve under a renormalized Lindblad master equation with c_ω given by Eq. (7.6) and $\lambda^* = 4\,(g/\hbar)^2\,\lambda$, thus leaving the thermal state of H_0 as the unique stationary state of the coupled system.

For $p = 0$, the effective system-reservoir interaction cools the system towards the ground state, and the Lindblad operators can be reduced to the n-body terms $\{\sqrt{\lambda^*}(E_{i,\nu} + T_{i,\nu}), \sqrt{\lambda^*}(E_{i,\nu} + T_{i,\nu}^\dagger)\}$ [126, 127]. In this limit the ancillary pseudospins become an effective low temperature bath with a cooling rate $\Gamma_{\text{c}} \sim \lambda^*$ and heating rate determined by gate errors and any environmental noise. Competition between these rates leads to a minimum reachable temperature for the system, which can be estimated as $T_{\min} \sim \Delta/\ln(\Gamma_{\text{c}}/\Gamma_{\text{e}})$, where $\Gamma_{\text{e}} \sim \text{EPG} \times \Omega$, with EPG and Ω the error rate and frequency of application of $U_j(\alpha)$, respectively.

The Lindblad master equation, with operators given by Eq. (7.6), generates a *unitary* system-reservior interaction but *nonunitary* reservoir relaxation. Stroboscopic simulation of the system-reservoir interaction, Eq. (7.7), is performed in a manner analogous to the H_0^{TC} simulation described above. Phase angles are chosen in the 1- and 2-body gates to generate an effective static interaction strength g over the time t_{sr} between applications of H_{sr}, such that $gt_{sr}/\hbar < \pi/2$. Nonunitary evolution of the reservoir is generated by encoding the reservoir as two levels of a Λ-system. The pseudospin states are the ground state $|0\rangle$ and the meta-stable state $|1\rangle$. State $|2\rangle$ is chosen to have fast spontaneous emission to $|0\rangle$, with rate Γ_{20}. This spontaneous emission is the decoherence mechanism required to generate the nonunitary Lindblad evolution. The ancillary pseudospin levels can be placed in a thermal state via the following procedure: i) π-pulse on the $|1\rangle \rightarrow |2\rangle$ transition. ii) Wait for decay to ground state, $|0\rangle$. iii) π-pulse on $|0\rangle \rightarrow |1\rangle$ transition. iv) θ-pulse on the $|1\rangle \rightarrow |2\rangle$ transition. v) Wait for decay, which now yields the final pseudospin state, $\rho = \text{diag}\left\{\sin^2(\theta), \cos^2(\theta)\right\}$, corresponding to an effective temperature $T_{\text{eff}} = \Delta/\left(2\ln(\cot\theta)\right)$. The above stroboscopic procedure generates $\lambda^* \approx g^2 t_{sr}/\hbar^2$ in Eq. (7.6). The procedure can be simplified in the limit of cooling towards zero temperature by eliminating steps iii-v, when it becomes similar to the optical pumping scheme employed in measurement of qubit states for trapped ions [128]. This thermalization or cooling procedure is then repeated and interleaved with the stroboscopic application of H_0.

7.4 Thermalization of the toric code

H_0^{TC} is of the form of Eq. (7.4), with two types of excitations, electric charges and magnetic vortices, ($\nu = \text{e}, \text{m}$) that reside on vertices and plaquettes, respectively, of the square lattice. The excitation operators are defined from (7.5) with $h_v^e = \prod_{j\in v}\sigma_j^z$, $h_p^m = \prod_{j\in p}\sigma_j^x$, and $\Sigma_i^{e,m} = \sigma_i^{x,z}$. Each link must interact with four ancillary pseudospins in the limit $T \ll \chi$ or $\chi \ll g$ to allow thermalization to the ground state or the thermal state of H_0^{TC}, respectively.

The stroboscopic generation of H_0^{TC} outlined above introduces truncation perturbations in the perturbative expansion, e.g, the second term in Eq. (7.3), which are distinct from

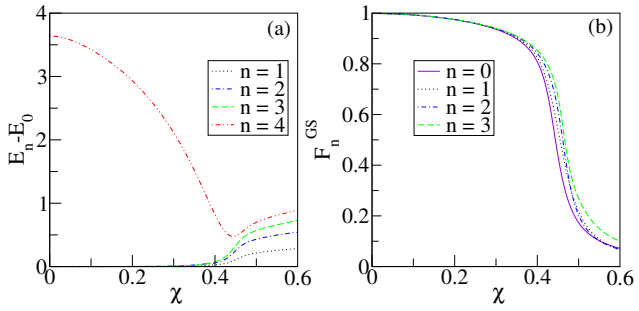


Figure 7.2: (a) Energy spectrum (in units of $J_e = J_m$) and (b) ground state fidelity vs. perturbation strength computed with exact diagonalization for the 18 site toric code with $h_z = 0.05$.

extrinsic errors due to experimental noise and gate inaccuracies. If sufficiently large, such truncation perturbations could drive the system away from the desired ground state phase . We now show, by direct calculation on a finite sized system accessible to current experiments [116], that the intrinsic perturbations can be kept sufficiently small.

Fig. 7.2(a) plots the gap of $H^{TC}_{\chi,h_z} = H^{TC}_0 - h_z \sum_i \sigma^z_i + \chi \sum_{\langle i,j\rangle} \sigma^x_i \sigma^y_i$ as a function of the strength of the perturbation for a 3×3 planar lattice with toroidal boundary conditions, realized by 18 spins (Fig. 7.1) . The Zeeman field is added here to fully split the ground state degeneracy and ensure robust characterization of the eigenstates of H even in the presence of small additional perturbations. We define the ground state fidelity as $F^{GS}_n = |\langle\Psi^0_n|\Psi_n(\chi,h_z)\rangle|$, where the $|\Psi^0_n\rangle$ are the degenerate ground states of H^{TC}_0 and $|\Psi_n(\chi,h_z)\rangle$ are the nearly degenerate ground states of H^{TC}_{χ,h_z}. Fig. 7.2(b) shows the ground state fidelity as a function of χ. This fidelity determines the robustness of topological operations that will be performed via string operators to measure or perform gates on the system (for a definition of these operators, see [112]). We see that for $|\chi| \lesssim 0.4$ the features of the topological phase persist, including the approximate four-fold degeneracy of the ground state and a finite gap to excitations. This corresponds to a maximum value of $\phi \sim 0.4$, which constrains the gate operations in the pulse sequences, Eq. (7.2). This robustness should increase with increasing lattice size, and is consistent with known stability of H^{TC}_0 to other perturbations [129].

Increasing ϕ increases J and therefore the gap of the H^{TC}_0; however it also increases χ/J, which reduce the gap of H^{TC}_{χ,h_z} and topological protection for large χ. We also note that $J \sim 1/N_G$ where N_G is the number of sequential gates used to simulate H. For larger lattices, some degree of parallelization, e.g., executing H^{xxxx}_{eff} for disjoint plaquettes simultaneously, is thus desirable to ensure that the gap does not decrease with the lattice size. Choosing $\chi = 0.2$, the gap achieved by a completely serial implementation is $\Delta \approx 0.6\ \mu\text{K}/N_{\text{sys}}$, where N_{sys} is the number of system atoms used. With the cooling sequence serially interleaved, $\Delta \approx 0.1\ \mu\text{K}/N_{\text{sys}}$ and $\lambda^* \sim 10^4\ \text{s}^{-1}/N_{\text{sys}}$ are achievable [121]. For a minimal system of 18 system atoms, this allows for an effective temperature $T_{\text{eff}} < \Delta$ to be reached with an error rate of EPG $\sim 10^{-4}$ or less [121].

7.5 Sources of errors

This scheme is designed to be robust against errors within the pseudospin subspace. The dominant source of residual error in the implementation discussed here is leakage from the Rydberg levels due to spontaneous emission and black body radiation. The latter may be effectively suppressed by working at low temperatures [130], and spontaneous emission is minimized by utilizing states with larger n (though sufficiently low that that effective atomic radius is significantly smaller than the interatomic spacing). With $n \lesssim 180$, we estimate that spointaneous emission errors can be reduced to $\sim 10^{-6}$ per gate, allowing for up to 10^3 stroboscopic cycles. Leakage errors are discussed in more detail in Appendix. 7.7.5.

7.6 Discussion

We have developed a formalism for the stroboscopic generation of n-body Hamiltonians using 1- and 2-body quantum operations together with a dissipative thermalization scheme. We have applied this here to the toric code Hamiltonian in the context of addressable optical lattice experiments [116]. Our implementation method also applies to a larger class of spin models, e.g, 2-body Ising models [121] and other experimental setups [131]. The dynamic generation both of a Hamiltonian possessing a topologically ordered ground state and of an effective thermalization mechanism offers the possibility of *robust* simulation of the ground state and of the creation and braiding of anyonic excitations [1] These are essential components required for the topologically protected storage and manipulation of quantum information.

We thank D. Weiss for useful discussions. This material is based upon work supported by DARPA under Award No. 3854-UCB-AFOSR-0041. During the preparation of this manuscript related results discussing ground state preparation were reported [132].

7.7 Appendices

7.7.1 Magnus expansion: Effective Hamiltonian formalism

After a time, t, a (possibly time-dependent) Hamiltonian, H, will generate the unitary operation,

$$U(t) = \underset{\leftarrow}{T} \exp\left(-i \int_0^T H(t')dt'/\hbar\right).$$

Here, $\underset{\leftarrow}{T}$, is the time ordering operator. Because this operator is difficult to work with, we would like to rewrite the propagator as

$$U(t) = \exp\left(-iH_{\text{eff}}t/\hbar\right).$$

[1] This differs from the proposal by Y. Han et al., in Phys. Rev. Lett. **98** (2007) to braid anyons in the absence of an applied Hamiltonian.

The *effective Hamiltonian*, H_{eff}, is expanded as a power-series in $\hbar$,

$$H_{\text{eff}} = \frac{1}{t}\sum_{n=1}^{\infty}\left(\frac{-i}{\hbar}\right)^{n-1}\Omega_n(t)$$

Following the method of Klarsfeld and Oteo [122], the Ω_n can be calculated throught a recursion relation,

$$\Omega_n = P_n - \sum_{k=2}^{n}\frac{1}{k!}Q_n^{(k)} \qquad (n \geq 2)\,,$$

where we define

$$P_n(t) = \int_0^t dt_1 \int_0^{t_1} dt_2 \cdots \int_0^{t_{n-1}} dt_n H(t_1)H(t_2)\cdots H(t_n),$$

and,

$$Q_n^{(k)} = \sum \Omega_{i_1}\cdots\Omega_{i_k} \qquad (i_1 + i_2 + \cdots + i_k = n)\,.$$

In terms of the operators P_i, the first several operators, Ω_j, are:

$$\begin{aligned}
\Omega_1 &= P_1 \\
\Omega_2 &= P_2 - \frac{1}{2}\Omega_1^2 \\
&= P_2 - \frac{1}{2}P_1^2 \\
\Omega_3 &= P_3 - \frac{1}{2}\left(\Omega_1\Omega_2 + \Omega_2\Omega_1\right) - \frac{1}{6}\Omega_1^3 \\
&= P_3 - \frac{1}{2}\left(P_1P_2 + P_2P_2\right) + \frac{1}{3}P_1^3
\end{aligned}$$

Simulating a $\sigma^z\sigma^z\sigma^z$ interaction

Consider the Hamiltonian, $H(t) = \{-H_1, -H_2, H_1, H_2\}_\tau$, where

$$H_1 = J_1\sigma^z\sigma^z\sigma^0 \qquad\qquad H_2 = J_2\sigma^0\sigma^y\sigma^z.$$

Calculating the P_i's, we have:

$$\begin{aligned}
P_1 &= \int_0^{4\tau} H(t)dt \\
&= 0 \\
P_2 &= \int_0^{4\tau} dt_1 \int_0^{t_1} dt_2 H(t_1)H(t_2) \\
&= -2i\tau^2 J_1J_2\,\sigma^z\sigma^z\sigma^z
\end{aligned}$$

This gives $\Omega_1 = 0$ and $\Omega_2 = -2i\tau^2 J_1 J_2\, \sigma^z\sigma^z\sigma^z$. So the effective Hamiltonian is, to second order,

$$H_{\text{eff}} = -\frac{\tau}{2\hbar} J_1 J_2\, \sigma^z\sigma^z\sigma^z + \mathcal{O}\left(\frac{\tau^2}{\hbar^2}\right)$$

The higher-oder terms can be found by the same method, yielding error terms which must also be accounted for. Carrying the expansion out further, we have:

$$\begin{aligned} H_{\text{eff}} = &-\frac{\tau}{2\hbar} J_1 J_2\, \sigma^z\sigma^z\sigma^z + \frac{\tau^2}{2\hbar^2} J_1 J_2 \left(J_1\ \sigma^0\sigma^y\sigma^z - J_2\ \sigma^z\sigma^x\sigma^0\right) + \frac{\tau^3}{3\hbar^3} J_1 J_2 \left(J_1^2 + J_2^2\right) \sigma^z\sigma^z\sigma^z + \\ &\frac{\tau^4}{6\hbar^4} J_1 J_2 \left(\left(2J_1^2 J_2 + J_2^3\right) \sigma^z\sigma^x\sigma^0 + \left(2J_2^2 J_1 + J_1^3\right) \sigma^0\sigma^y\sigma^z\right) + \mathcal{O}\left(\frac{\tau^5}{\hbar^5}\right) \end{aligned}$$

But notice the symmetry of the error terms here. By making the replacement $J_i \to -J_i$, the error terms (those not proportional to $\sigma^z\sigma^z\sigma^z$) switch sign. This implies that the Hamiltonian

$$H(t) = \{-H_1, -H_2, H_1, H_2, H_1, H_2, -H_1, -H_2, \}_\tau$$

may do better. Performing the expansion, we see that

$$H_{\text{eff}} = \left(-\frac{\tau}{2\hbar} J_1 J_2 + \frac{\tau^3}{3\hbar^3} J_1 J_2 \left(J_1^2 + J_2^2\right)\right) \sigma^z\sigma^z\sigma^z + \frac{\tau^4}{\hbar^4} J_1^2 J_2^2 \left(J_1\ \sigma^z\sigma^x\sigma^0 + J_2\, \sigma^0\sigma^y\sigma^z\right) + \mathcal{O}\left(\frac{\tau^5}{\hbar^5}\right)$$

The error in this effective Hamiltonian now appears to 5th order in $\tau/\hbar$. This process can be repeated until the error is suppressed as far into the expansion as needed.

7.7.2 Simulating the toric code

Consider the Hamiltonian,

$$H(t) = \{-H_1, -H_2, H_1, H_2, H_3, -H_2, -H_1, H_2, H_1, -H_3\}_\tau$$

Where

$$\begin{aligned} H_1 &= J_1\ \sigma^z\sigma^y\sigma^0\sigma^0 \\ H_2 &= J_2\ \sigma^0\sigma^x\sigma^y\sigma^0 \\ H_3 &= J_3\ \sigma^0\sigma^0\sigma^x\sigma^z \end{aligned}$$

By the method used above, this generates the effective Hamiltonian:

$$\begin{aligned} H_{\text{eff}} = &\frac{2J_1 J_2 J_3 \tau^2}{5\hbar^2} \sigma^z\sigma^z\sigma^z\sigma^z\sigma^z + \frac{2J_1 J_2 J_3 \tau^3}{5\hbar^3} \left(J_1\ \sigma^0\sigma^x\sigma^z\sigma^z + J_3\ \sigma^z\sigma^z\sigma^y\sigma^0\right) - \\ &\frac{2J_1 J_2 J_3 \tau^4}{5\hbar^4} \left(2J_1 J_2\, \sigma^0\sigma^0\sigma^x\sigma^z - J_1 J_3\, \sigma^0\sigma^x\sigma^y\sigma^0 + \frac{2}{3}\left(J_1^2 + J_2^2 + J_3^2\right) \sigma^z\sigma^z\sigma^z\sigma^z\right) + \mathcal{O}\left(\frac{\tau}{\hbar}\right)^5 \end{aligned}$$

which acts for a time, 10τ. Using the same error-canceling technique applied for the $\sigma^z\sigma^z\sigma^z$ interaction above, we can repeat this sequence with $J_1 \to -J_1$ and $J_3 \to -J_3$, leading to a Hamiltonian,

$$H(t) = \{ -H_1, -H_2, H_1, H_2, H_3, -H_2, -H_1, H_2, H_1, -H_3, \\ H_1, -H_2, -H_1, H_2, -H_3, -H_2, H_1, H_2, -H_1, H_3\}_\tau$$

The effective Hamiltonian generated by this is:

$$H_{\text{eff}} = \frac{2J_1J_2J_3\tau^2}{5\hbar^2}\sigma^z\sigma^z\sigma^z\sigma^z + \frac{2J_1J_2J_3\tau^4}{5\hbar^4}\left(J_1J_3\,\sigma^0\sigma^x\sigma^y\sigma^0 - \frac{2}{3}\left(J_1^2+J_2^2+J_3^2\right)\sigma^z\sigma^z\sigma^z\sigma^z\right) + \mathcal{O}\left(\frac{\tau}{\hbar}\right)^5$$

which acts for a time, 20τ. To cancel the $\sigma^x\sigma^y\sigma^0$ term, we can repeat the sequence again, this time with $H_1 \to -H_1$ and $H_2 \to -H_2$. This new sequence would generate an effective Hamiltonian which acts for a time 40τ.

The plaquette operator, $\sigma^x\sigma^x\sigma^x\sigma^x$, can be generated by cyclic permutation of the Pauli operators in the above expressions for H_1, H_2, H_3 and $\mathrm{H_{eff}}$. Simulation of the Toric Code, however requires not only the simulation of $\sigma^x\sigma^x\sigma^x\sigma^x$ and $\sigma^z\sigma^z\sigma^z\sigma^z$, but of the total Hamiltonian, $\sum_p \sigma^x\sigma^x\sigma^x\sigma^x + \sum_v \sigma^z\sigma^z\sigma^z\sigma^z$. Because the Hamiltonian is a sum of commuting terms, we can simulate the total Hamiltonian by applying the pulse sequences in serial. Schematically

$$e^{-i\sigma^x\sigma^x\sigma^x\sigma^xt/\hbar}e^{-i\sigma^z\sigma^z\sigma^z\sigma^zt/\hbar} = e^{-i(\sigma^x\sigma^x\sigma^x\sigma^x+\sigma^z\sigma^z\sigma^z\sigma^z)t/\hbar}$$

Because only the error terms fail to commute, the error in the above expression occurs at higher orders than the error in the simulation.

7.7.3 Physical implementation of necessary gates

The Weiss apparatus has the ability to perform controlled phase gates, $\mathbf{CPHASE_\phi}$, and arbitrary single qubit unitary operations. In a two-qubit Hilbert space with basis, $\{|\downarrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\uparrow\uparrow\rangle\}$, the $\mathbf{CPHASE_\phi}$ gate takes the form:

$$\mathbf{CPHASE_\phi} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & e^{i\phi} \end{pmatrix}$$

Note that the same gate is performed regardless of which qubit is acting the control.

The above simulation requires the ability to implement the two-qubit gate, $U(t) = \exp\left(\frac{-iJt}{\hbar}(\sigma_1 \otimes \sigma_2)\right)$. Here σ_i is a Pauli operator acting on the i^{th} qubit. To illustrate how this gate may be implemented with single qubit gates and $\mathbf{CPHASE_\phi}$ gates, we choose $(\sigma_1, \sigma_2) = (\sigma^x, \sigma^y)$. Consider the gates:

$$U = (\sigma^x \otimes \sigma^x) \cdot \mathbf{CPHASE_\phi} \cdot (\sigma^x \otimes \sigma^x) \cdot \mathbf{CPHASE_\phi}$$

Writing this out as a matrix,

$$U = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & e^{i\phi} \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & e^{i\phi} \end{pmatrix}$$
$$= \begin{pmatrix} e^{i\phi} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & e^{i\phi} \end{pmatrix}$$

Which, by a global phase change,

$$U \equiv e^{-i\phi/2} U = \begin{pmatrix} e^{i\phi/2} & 0 & 0 & 0 \\ 0 & e^{-i\phi/2} & 0 & 0 \\ 0 & 0 & e^{-i\phi/2} & 0 \\ 0 & 0 & 0 & e^{i\phi/2} \end{pmatrix} = e^{i\phi/2(\sigma^z \otimes \sigma^z)}$$

Expanding the RHS of the above expression,

$$U = e^{i\phi/2(\sigma^z \otimes \sigma^z)} = (I \otimes I)\cos\frac{\phi}{2} + i(\sigma^z \otimes \sigma^z)\sin\frac{\phi}{2}$$

But this can be rotated into a different basis using the Pauli identities:

$$e^{-i\frac{\pi}{4}\sigma^y} \sigma^z e^{i\frac{\pi}{4}\sigma^y} = \sigma^x$$
$$e^{i\frac{\pi}{4}\sigma^x} \sigma^z e^{-i\frac{\pi}{4}\sigma^x} = \sigma^y$$

So, for instance,

$$\begin{aligned} e^{i\frac{\pi}{4}(\sigma^x \otimes \sigma^0 - \sigma^0 \otimes \sigma^y)}\, U\, e^{-i\frac{\pi}{4}(\sigma^x \otimes \sigma^0 - \sigma^0 \otimes \sigma^y)} &= (I \otimes I)\cos\frac{\phi}{2} + i\left(\left(e^{i\frac{\pi}{4}\sigma^x} \sigma^z e^{-i\frac{\pi}{4}\sigma^x}\right) \otimes \left(e^{-i\frac{\pi}{4}\sigma^y} \sigma^z e^{i\frac{\pi}{4}\sigma^y}\right)\right)\sin\frac{\phi}{2} \\ &= (I \otimes I)\cos\frac{\phi}{2} + i\left(\sigma^y \otimes \sigma^x\right)\sin\frac{\phi}{2} \\ &= e^{i\phi/2(\sigma^y \otimes \sigma^x)} \end{aligned}$$

We then have the decomposition of the gate:

$$e^{i\phi(\sigma^y \otimes \sigma^x)} = e^{i\frac{\pi}{4}(\sigma^x \otimes \sigma^0)} \cdot e^{-i\frac{\pi}{4}(\sigma^0 \otimes \sigma^y)} \cdot (\sigma^x \otimes \sigma^x) \cdot \mathbf{CPHASE_{2\phi}} \cdot (\sigma^x \otimes \sigma^x) \cdot \mathbf{CPHASE_{2\phi}} \cdot e^{-i\frac{\pi}{4}(\sigma^x \otimes \sigma^0)} \cdot e^{i\frac{\pi}{4}(\sigma^0 \otimes \sigma^y)}$$

To implement this one physical gate, $e^{i\phi(\sigma^y \otimes \sigma^x)}$, we must execute 10 physical gates. It may be possible to reduce this by one or two gates, recognizing that:

$$\sigma^x \otimes \sigma^x = \left(\sigma^x \otimes \sigma^0\right) \cdot \left(\sigma^0 \otimes \sigma^x\right),$$

so that these gates can be combined with the single qubit rotations done later. We point out that this expansion in terms of primitive gates may be improved through use of Weyl chamber techniques as developed in [133].

7.7.4 Stroboscopic thermalization

To demonstrate the feasibility of simulating the thermalization of a quantum system, we consider the following model. We describe a system qubit by a generic density matrix,

$$\rho_s = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = a\,|\uparrow\rangle\langle\uparrow| + b\,|\uparrow\rangle\langle\downarrow| + c\,|\downarrow\rangle\langle\uparrow| + d\,|\downarrow\rangle\langle\downarrow|$$

We also assume the existence of a reservoir qubit which can be periodically initialized to the state,

$$\rho_r = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} = \alpha\,|+\rangle\langle+| + \beta\,|-\rangle\langle-|$$

This state corresponds to a (possibly negative) temperature:

$$T = \frac{\Delta}{k_b \ln \frac{\beta}{\alpha}}, \tag{7.8}$$

where Δ is the qubit splitting for both the reservoir and system qubits. These two systems are coupled by the interaction-picture Hamiltonian:

$$\mathcal{H}_{sr} = g\left(\sigma^+ s^- + \sigma^- s^+\right) = \frac{g}{2}\left(\sigma^x s^x + \sigma^y s^y\right)$$

Here g is assumed to be a real-valued coupling constant. The operators $\sigma^\pm$ are the raising/lowering operators on the system qubit,

$$\begin{array}{ll} \sigma^+\,|\uparrow\rangle = 0 & \sigma^-\,|\uparrow\rangle = |\downarrow\rangle \\ \sigma^+\,|\downarrow\rangle = |\uparrow\rangle & \sigma^-\,|\downarrow\rangle = 0, \end{array}$$

while the $s^\pm$ is the raising/lowering operator on the reservoir qubit,

$$\begin{array}{ll} s^+\,|+\rangle = 0 & s^-\,|+\rangle = |-\rangle \\ s^+\,|-\rangle = |+\rangle & s^-\,|-\rangle = 0. \end{array}$$

Expressed in terms of $\sigma^x, \sigma^y, s^x, s^y$, the Hamiltonian is a sum of two commuting terms, so the techniques described in the above sections allow for easy simulation of the corresponding time evolution. The eigenstates of $\mathcal{H}_{sr}$ are

$$\begin{array}{ll} |\psi_1\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow -\rangle + |\downarrow +\rangle\right) & |\psi_3\rangle = |\uparrow +\rangle \\ |\psi_2\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow -\rangle - |\downarrow +\rangle\right) & |\psi_4\rangle = |\downarrow -\rangle \end{array}$$

With associated eigenvalues, $E_1 = g, E_2 = -g, E_3 = 0, E_4 = 0$. Given an initial state, $\rho_{sr}(0) = \rho_s \otimes \rho_r$, The combined state of the system/reservior at a later time, t, is

$$\begin{aligned}
\rho_{sr}(t) &= e^{(-i\mathcal{H}_{sr}t/\hbar)}\rho_{sr}(0)e^{(i\mathcal{H}_{sr}t/\hbar)} \\
&= \frac{1}{2}(\alpha d + \beta a)\,|\psi_1\rangle\langle\psi_1| + \frac{1}{2}(\alpha d + \beta a)\,|\psi_2\rangle\langle\psi_2| + \alpha a\,|\psi_3\rangle\langle\psi_3| + \beta d\,|\psi_4\rangle\langle\psi_4| + \\
&\quad \frac{1}{2}(\beta a - \alpha d)\,|\psi_1\rangle\langle\psi_2|\,e^{-2ig/\hbar} + \frac{1}{2}(\beta a - \alpha d)\,|\psi_2\rangle\langle\psi_1|\,e^{2ig/\hbar} + \\
&\quad \frac{1}{\sqrt{2}}\alpha c\,|\psi_1\rangle\langle\psi_3|\,e^{-ig/\hbar} + \frac{1}{\sqrt{2}}\alpha b\,|\psi_3\rangle\langle\psi_1|\,e^{ig/\hbar} - \frac{1}{\sqrt{2}}\alpha c\,|\psi_2\rangle\langle\psi_3|\,e^{ig/\hbar} - \\
&\quad \frac{1}{\sqrt{2}}\alpha b\,|\psi_3\rangle\langle\psi_2|\,e^{-ig/\hbar} + \frac{1}{\sqrt{2}}\beta d\,|\psi_1\rangle\langle\psi_4|\,e^{-ig/\hbar} + \frac{1}{\sqrt{2}}\beta c\,|\psi_4\rangle\langle\psi_1|\,e^{ig/\hbar} + \\
&\quad \frac{1}{\sqrt{2}}\beta d\,|\psi_2\rangle\langle\psi_4|\,e^{ig/\hbar} + \frac{1}{\sqrt{2}}\beta c\,|\psi_4\rangle\langle\psi_2|\,e^{-ig/\hbar}
\end{aligned}$$

Tracing out the reservoir qubit, we see

$$\begin{aligned}
\rho_s(t) &= \mathrm{Tr}_r\rho(t) \\
&= \left(\alpha + \alpha d\sin(gt/\hbar)^2 + \beta a\cos(gt/\hbar)^2\right)|\uparrow\rangle\langle\uparrow| + b\cos(gt)\,|\uparrow\rangle\langle\downarrow| + \\
&\quad c\cos(gt)\,|\downarrow\rangle\langle\uparrow| + \left(\beta d + \beta a\sin(gt/\hbar)^2 + \alpha d\cos(gt/\hbar)^2\right)|\downarrow\rangle\langle\downarrow| \\
&= \begin{pmatrix} \alpha + \alpha d\sin(gt/\hbar)^2 + \beta a\cos(gt/\hbar)^2 & b\cos(gt/\hbar) \\ c\cos(gt/\hbar) & \beta d + \beta a\sin(gt/\hbar)^2 + \alpha d\cos(gt/\hbar)^2 \end{pmatrix}
\end{aligned}$$

We can then write this evolution as an operator on the *vectorized* system density matrix,

$$\rho_s = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \longrightarrow \tilde{\rho}_s = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}.$$

Then the evolution $\tilde{\rho}_s(t) = \tilde{L}(t)\tilde{\rho}_s(0)$ is described by the matrix,

$$\tilde{L}(t) = \begin{pmatrix} \alpha + \beta\cos^2(gt/\hbar) & 0 & 0 & \alpha\sin^2(gt/\hbar) \\ 0 & \cos(gt/\hbar) & 0 & 0 \\ 0 & 0 & \cos(gt/\hbar) & 0 \\ \beta\sin^2(gt/\hbar) & 0 & 0 & \beta + \alpha\cos^2(gt/\hbar) \end{pmatrix}. \tag{7.9}$$

Note that this evolution preserves $\mathrm{Tr}\,\rho(t) = 1$. This operator has a (non-normalized) eigensystem,

$$\begin{array}{ll}
\lambda_1 = 1 & \tilde{\rho}_1 = (\alpha, 0, 0, \beta) \\
\lambda_2 = \cos(gt/\hbar) & \tilde{\rho}_2 = (0, 1, 0, 0) \\
\lambda_3 = \cos(gt/\hbar) & \tilde{\rho}_3 = (0, 0, 1, 0) \\
\lambda_4 = \cos^2(gt/\hbar) & \tilde{\rho}_4 = (1, 0, 0, -1)
\end{array}$$

At almost all times, t, eigenvalues λ_2, λ_3 and λ_4 are *less than 1.* So if this procedure is repeated multiple times, the only eigenstate that survives is ρ_1. So, upon repeated application of this operation, any initial system state converges to

$$\rho_{\text{thermal}} = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}.$$

Explicitly,

$$\tilde{\rho}_s(0) = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \tilde{\rho}_1 + b\tilde{\rho}_2 + c\tilde{\rho}_3 + (a\beta - d\alpha)\,\tilde{\rho}_4.$$

Evolving this state in time,

$$\tilde{\rho}_s(t) = \tilde{L}(t)\tilde{\rho}_s(0) = \tilde{\rho}_1 + (b\cos(gt/\hbar))\,\tilde{\rho}_2 + (c\cos(gt/\hbar))\,\tilde{\rho}_3 + \left((a\beta - d\alpha)\cos^2(gt/\hbar)\right)\tilde{\rho}_4.$$

If this operation is applied N times,

$$\tilde{\rho}_s(Nt) = \tilde{L}^N(t)\rho_s(0) = \tilde{\rho}_1 + \left(b\cos^N(gt/\hbar)\right)\tilde{\rho}_2 + \left(c\cos^N(gt/\hbar)\right)\tilde{\rho}_3 + \left((a\beta - d\alpha)\cos^{2N}(gt/\hbar)\right)\tilde{\rho}_4.$$

Choosing $(gt/\hbar) < 1$ we see that,

$$\lim_{N\to\infty} \tilde{\rho}_s(Nt) = \tilde{\rho}_1 = \tilde{\rho}_{\text{thermal}}$$

Therefore, the asymptotic state of the system evolving under Eq. (7.9) is the thermal state at the temperature defined by Eq. (7.8).

7.7.5 Experimental details and error sources

The experiment run by Dave Weiss is a ^{1}33Cs neutral-atom optical trap. The qubit basis states are most likely $|F, m_F\rangle = \{|4,4\rangle, |3,3\rangle\}$, though it is possible that $|F, m_F\rangle = \{|4,1\rangle, |3,-1\rangle\}$ could be used as well. Gates will require the use of a Rydberg state, most likely the $n \approx 80$ state due to the availablily of lasers operating on the relevant transition.

The optical lattice will be formed with 3 pairs of lasers at 5 nm detuning with intrapair angle of 10 degrees. This makes for a lattice constant of approximately 5 microns. There are two Raman lasers near 852nm and 510nm. Plenty of power is available on the 852 nm laser. The 510 nm laser is at least 10mW, but up to 1W *may* be available. A microwave gate is available, which drives a π-pulse on the $|3,3\rangle \leftrightarrow |4,4\rangle$ transition in 30μs at a frequency of 9.2 GHz.

Trap loss due to collisions with background gas are roughly $(1/100)\,\text{s}^{-1}$. The following errors for single-qubit operations, as described in [134], have been calculated:

- Raman Scattering (blue-detuned lattice) rate across all atoms in trap

$$T_1 \frac{N_a}{n_a} \frac{\pi c \epsilon_0}{a\omega_L} \sqrt{\frac{k_b U_L}{2M}} \frac{\sigma}{\alpha} = 4.3 \times 10^{-11} s^{-1}$$

- Raman scattering (red-detuned lattice) rate across all atoms in trap

$$T_1 \frac{N_a}{n_a} \frac{2c\epsilon_0}{\hbar\omega_L} k_b U_L \frac{\sigma}{\alpha} = 4.5 \times 10^{-8} s^{-1}$$

- Neighbor atom errors (Raman gate) error per gate

$$\frac{2\pi^2}{3} \left(1 + \frac{a^2 {\lambda_R}^2}{\pi^2 {w_0}^4}\right)^{-2} \exp\left[-\frac{4a^2}{{w_0}^2}\right] = 1.16 \times 10^{-30}$$

- Spontaneous emission (Raman gate) error per gate

$$\frac{\pi}{2\left|\Delta_1\right|\tau} = 1.43 \times 10^{-6}$$

- AC Stark shift (Raman gate) error per gate

$$\frac{4}{3\pi^2} \frac{\hbar^2 a^2}{M {w_0}^4 k_b U_L} \left(\frac{610^{10} s^{-1}}{\Delta_1}\right)^2 = 1.07 \times 10^{-10}$$

- Atomic-motion-reduced pulse area (Raman gate) error per gate

$$\frac{1}{12} \frac{\hbar^2 a^2}{M {w_0}^4 k_b U_L} = 1.82 \times 10^{-5}$$

- Detuning Doppler shift (Raman gate) error per gate

$$0.98 \left(\frac{2\pi}{\lambda_R \Omega_R}\right)^2 \frac{\hbar}{M^{3/2} a} \left(k_b U_L\right)^{1/2} = 4.63 \times 10^{-10}$$

- Polarization effects (Raman gate) error per gate

$$\frac{\hbar {\lambda_R}^2 a}{\pi {w_0}^4} \frac{1}{\sqrt{2 M k_b U_L}} = 1.85 \times 10^{-3}$$

- Laser intensity noise (Raman gate) error per gate

$$\left(2.6 \times 10^{-6} m^2 s^{-1}\right) \frac{1}{\left|\Delta_1\right| {w_0}^2} = 5.75 \times 10^{-8}$$

As already noted above, leakage from excited Rydberg states incurred during two-qubit operations is the dominant source of error in our implementation. Rydberg state lifetimes are determined by both spontaneous emission and black body radiation and are sensitive to the choice of principal and angular momentum quantum numbers, n and l, respectively. Rates

of spontaneous emission decrease as n increases and as l increases. Losses due to interaction with black body radiation increase with n. However, this factor may be suppressed by working a low temperatures: for Cs, explicit calculations in Ref. [130] show that a decrease to 4K can very effectively remove black body radiation induced losses. Although imposing an additional experimental requirement, decreasing the temperature of the chamber to 4K is well within current day capabilities. Thus we neglect the losses induced by black body radiation.

The spontaneous emission losses can themselves be minimized by choosing the largest achievable quantum numbers, n, l. Since in the Rydberg blockade gate for addressable atoms (protocol B of [117]) only a single atom is excited to a Rydberg state, with the second, off-resonant atom having at most a small amount of Rydberg state amplitude (see Fig. 2 in [117] and Fig. 2 in [135], a bound on these parameters may be attained by demanding that the excited state Rydberg atom is small compared to the interatomic distance, $R \approx 5\mu$m. The effective size, $r(n, l)$, of a large-n Rydberg atom may be computed from the hydrogenic wavefunction of the outer electron (neglecting the quantum defect for large n values),

$$r(n, l) = \frac{1}{2}\left[3n^2 - l(l+1)\right] a_0, \tag{7.10}$$

where a_0 is the Bohr radius. Evaluating this for $R/r = 2$ and for a more conservative requirement $R/r = 10$ with $l = 0$ leads to the maximum n values 177 and 80 respectively (column 2 of Table 7.7.5). Using non-zero l values will allow somewhat larger n (e.g., $n = 90$ for $l = 89$ and $n = 217$ for $l = 176$) but for the current purpose of estimating the order of magnitude of the leakage error we shall take the more conservative, lower estimates of n, regardless of the angular momentum excitation.

Phenomenological fits to spontaneous emission rates for Rydberg atoms at low n and l values ($n \leq 26, l \leq 3$) are given in Gallagher's book on Rydberg atoms (Eq. (4.17) and Table 4.1 in [136]). For high n values where the effect of the quantum defect becomes small relative to n, we can estimate the spontaneous emission rates for non-zero l values using the high accuracy hydrogenic lifetime expressions developed recently by Horbatsch et al. [137]. We have computed lifetimes τ for several l states for each of the maximum n values, using [136] for $l = 0$ and [137] for $l > 0$ (column 4 in Table 7.7.5). These lifetimes were then used to calculate the initial state-averaged leakage error for the Rydberg gate in an addressable lattice (i.e., with the second protocol (B) of Jaksch et al [117]), using a modified form of Eqs. (26) and (28) in Saffman and Walker's recent review [135] in which we include only the errors due to leakage (first term in Eq. (26) of [135]) and not the blockade error (second term in Eq. (26) of [135]). We omit the blockade error since this gives rise to a phase shift which may be compensated for by the methods described in our manuscript. In this situation, the optimal Rabi frequency is then $\Omega_{opt} = \sqrt{7}B$ where B is the blockade energy shift (Eq. (24) of [135] with $N = 2$) and the resulting minimal leakage error per gate is $E_{leak} = \frac{\sqrt{7}\pi}{2B\tau}$. Evaluation of E_{leak} now requires an estimate for the blockade energy shift, which in turn depends on the strength of the Rydberg interaction. Saffman and Walker have measured a value of $B \sim 10$ MHz for $n = 90, l = 2$ Rydberg states of Rb separated

by more than 10 μm in a direct experimental measurement of the Rydberg blockade [138]. Noting that theoretical calculations for Rb in [139] show that the strength of the Rydberg interaction may be increased by several orders of magnitude in larger n states and at smaller distances (see Fig. 7 in [139]), we therefore use estimates of $B = 10 - 100$MHz to evaluate the leakage error. The resulting Rydberg gate leakage errors are listed in column 6 of Table 7.7.5. The gate errors obtained for the $n = 177, l = 5, 6$ states allow for $10^5 - 10^6$ CPHASE gates, corresponding to 350-3500 stroboscopic cycles using the estimates of cycle time given in our manuscript, while the $n = 80, l = 5, 6$ states allow for $10^4 - 10^5$ CPHASE gates, or 35-350 stroboscopic cycles. Operating at this level of robustness is sufficient to validate the approach.

Creation of the high principle quantum numbers n given in Table 7.7.5 requires stability and homogeneity of the electric fields, since Rydberg states are susceptible to ionization (see e.g., [135] for a discussion of this). However we note that Saffman and Walker have recently performed experiments with dipole trapped Rb atoms in Rydberg states with $n = 97$ [140], so $n \sim 100$ and the more conservative estimates of $n = 80$ are clearly feasible. We also note that additional gains in spontaneous emission lifetime and consequent leakage suppression are potentially possible by further increasing the angular momentum l of the Rydberg states. Chang has established an analytic bound on spontaneous emission lifetimes that is accurate to within 10% of experimental values and of the values of Horbatsch et al. [137] at large n [141] and which scales as $(l+1/2)^2$. This implies that significantly increased lifetimes will be obtained if it is experimentally possible to access states with $l > 10$. The direct, multiphoton excitation that would be required for this (rather than indirect stepwise excitation employed in [142]) has been demonstrated for higher l Rydberg states [143]. However the usefulness of higher l states for the Rydberg gate is complicated by the fact that the dipole moment is expected to be smaller so that B will then decrease. Detailed studies would be necessary to make any specific predictions beyond $l = 10$. Therefore our estimates here confined to low l values that can be achieved with current experimental technology and for which the dipole interactions are expected to be similar or larger than what has already been demonstrated.

Table 7.1: Values of n and l for ^{133}Cs atoms of spatial extent r in an optical lattice of intersite spacing $R = 5\mu$m, their spontaneous emission lifetimes τ and resulting leakage induced error per gate for the CPHASE Rydberg gate of Jaksch et al. [117] with their second protocol (protocol B) which is specific to addressable atoms. The error per gate E_{leak} was calculated with Eq. (26) of Ref. [135] using the values of τ given here, together with realistic estimates of the blockade parameter B.

R/r	n	l	τ	B	E_{leak}
2	177	0	7.9 msec	10 MHz	5.3 x 10^{-5}
2	177	1	1.0 msec	10 MHz	4.2 x 10^{-4}
2	177	2	2.9 msec	10 MHz	1.4 x 10^{-4}
2	177	5	14 msec	10 MHz	2.3 x 10^{-5}
2	177	6	20 msec	10 MHz	2.1 x 10^{-5}
2	177	0	7.9 msec	100 MHz	5.3 x 10^{-6}
2	177	1	1.0 msec	100 MHz	4.2 x 10^{-5}
2	177	2	2.9 msec	100 MHz	1.4 x 10^{-5}
2	177	5	14 msec	100 MHz	2.3 x 10^{-6}
2	177	6	20 msec	100 MHz	2.1 x 10^{-6}
10	80	0	0.73 msec	10 MHz	5.7 x 10^{-4}
10	80	1	0.094 msec	10 MHz	4.4 x 10^{-3}
10	80	2	0.27 msec	10 MHz	1.5 x 10^{-3}
10	80	5	1.3 msec	10 MHz	3.2 x 10^{-4}
10	80	6	1.8 msec	10 MHz	2.3 x 10^{-4}
10	80	0	0.73 msec	100 MHz	5.7 x 10^{-5}
10	80	1	0.094 msec	10 MHz	4.4 x 10^{-4}
10	80	2	0.27 msec	10 MHz	1.5 x 10^{-4}
10	80	5	1.3 msec	10 MHz	3.2 x 10^{-5}
10	80	6	1.8 msec	10 MHz	2.3 x 10^{-5}

Bibliography

[1] Richard Feynman. Simulating physics with computers. *International Journal of Theoretical Physics*, 21:467–488, 1982. 10.1007/BF02650179.

[2] Peter W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Review*, 41(2):303–332, 1999.

[3] Lieven M. K. Vandersypen, Matthias Steffen, Gregory Breyta, Costantino S. Yannoni, Mark H. Sherwood, and Isaac L. Chuang. Experimental realization of Shor's quantum factoring algorithm using nuclear magnetic resonance. *Nature*, 414:883–887, 2001.

[4] M. D. Barrett, J. Chiaverini, T. Schaetz, J. Britton, W.M. Itano, J.D. Jost, E. Knill, C. Langer, D. Leibfried, R. Ozeri, and D.J. Wineland. Deterministic quantum teleportation of atomic qubits. *Nature*, 429:737–739, 2004.

[5] J. Vala, K. B. Whaley, and D. S. Weiss. Quantum error correction of a qubit loss in an addressable atomic system. *Phys. Rev. A*, 72:052318, 2005.

[6] J.I. Cirac, H.J. Zoller, and H. Mabuchi. Quantum state transfer and entanglement distribution among distant nodes in a quantum network. *Phys. Rev. Lett.*, 78:3221–3224, 1997.

[7] Pieter Kok, W. J. Munro, Kae Nemoto, T. C. Ralph, Jonathan P. Dowling, and G. J. Milburn. Linear optical quantum computing with photonic qubits. *Rev. Mod. Phys.*, 79(1):135–174, Jan 2007.

[8] Mark A. Eriksson, Mark Friesen, Susan N. Coppersmith, Robert Joynt, Levente J. Klein, Keith Slinker, Charles Tahan, P. M. Mooney, J. O. Chu, and S. J. Koester. Spin-based quantum dot quantum computing in silicon. *Quantum Information Processing*, 3:133–146, 2004. 10.1007/s11128-004-2224-z.

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

[10] R. Hanson, O. Gywat, and D. D. Awschalom. Room-temperature manipulation and decoherence of a single spin in diamond. *Phys. Rev. B*, 74(16):161203, Oct 2006.

[11] Yuriy Makhlin, Gerd Schön, and Alexander Shnirman. Quantum-state engineering with josephson-junction devices. *Rev. Mod. Phys.*, 73(2):357–400, May 2001.

[12] P. M. Platzman and M. I. Dykman. Quantum Computing with Electrons Floating on Liquid Helium. *Science*, 284(5422):1967–1969, 1999.

[13] David Divincenzo. The physical implementation of quantum computation. arXiv:quant-ph/0002077, 2000.

[14] Pekko Kuopanportti, Mikko Mottonen, Ville Bergholm, Olli-Pentti Saira, Jun Zhang, and K. Birgitta Whaley. Suppression of $1/f^{\alpha}$ noise in one-qubit systems. *Physical Review A (Atomic, Molecular, and Optical Physics)*, 77(3):032334, 2008.

[15] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 1 edition, October 2000.

[16] C. P. Slichter. *Principles of Magnetic Resonance*. Springer, 1996.

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

[18] Rutger Vrijen, Eli Yablonovitch, Kang Wang, Hong Wen Jiang, Alex Balandin, Vwani Roychowdhury, Tal Mor, and David DiVincenzo. Electron-spin-resonance transistors for quantum computing in silicon-germanium heterostructures. *Phys. Rev. A*, 62:012306, 2000.

[19] C. D. Hill, L. C. L. Hollenberg, A. G. Fowler, C. J. Wellard, A. D. Greentree, and H.-S. Goan. Global control and fast solid-state donor electron spin quantum computing. *Phys. Rev. B*, 72:045350, 2005.

[20] T. Schenkel, J. A. Liddle, A. Persaud, A. M. Tyryshkin, S. A. Lyon, R. de Sousa, K. B. Whaley, J. Bokor, J. Shangkuan, and I. Chakarov. Electrical activation and electron spin coherence of ultralow dose antimony implants in silicon. *App. Phys. Lett.*, 88:112101, 2006.

[21] D. R. McCamey, H. Huebl, M. S. Brandt, W. D. Hutchison, J. C. McCallum, R. G. Clark, and A. R. Hamilton. Electrically detected magnetic resonance in ion-implanted si:p nanostructures. *App. Phys. Lett.*, 89:182115, 2006.

[22] Andre R. Stegner, Christoph Boehme, Hans Huebl, Martin Stutzmann, Klaus Lips, and Martin S. Brandt. Electrical detection of coherent 31p spin quantum states. *Nature Physics*, 2:835, 2006.

[23] J. M. Elzerman, R. Hanson, L. H. Willems van Beveren, B. Witkamp, L. M. K. Vandersypen, and L. P. Kouwenhoven. Single-shot read-out of an individual electron spin in a quantum dot. *Nature*, 430:431, 2004.

[24] M. Xiao, I. Martin, E. Yablonovitch, and H. W. Jiang. Electrical detection of the spin resonance of a single electron in a silicon field-effect transistor. *Nature*, 430:435, 2004.

[25] Ruby N. Ghosh and Robert H. Silsbee. Spin-spin scattering in a silicon two-dimensional electron gas. *Phys. Rev. B*, 46:12508, 1992.

[26] C. C. Lo, J. Bokor, T. Schenkel, A. M. Tyryshkin, and S. A. Lyon. Spin-dependent scattering off neutral antimony donors in 28si field-effect transistors. *App. Phys. Lett.*, 91:242106, 2007.

[27] Igor Zutic, Jaroslav Fabian, and S. Das Sarma. Spintronics: fundamentals and applications. *Rev. Mod. Phys.*, 76:323, 2004.

[28] A. Honig. Neutral-impurity scattering and impurity zeeman spectroscopy in semiconductors using highly spin-polarized carriers. *Phys. Rev. Lett.*, 17:186, 1966.

[29] J. Schmidt and I. Solomon. *C. R. Acad. Ser. B*, 263:169, 1966.

[30] L. H. Willems van Beveren, H. Huebl, D. R. McCamey, T. Duty, A. J. Ferguson, R. G. Clark, and M. S. Brandt. Broadband electrically detected magnetic resonance of phosphorus donors in a silicon field-effect transistor. *App. Phys. Lett.*, 93:072102, 2008.

[31] Vladimir B. Braginsky, Yuri I. Vorontsov, and Kip S. Thorne. Quantum Nondemolition Measurements. *Science*, 209(4456):547–557, 1980.

[32] Vladimir B. Braginsky and Farid Y. Khalili. *Quantum measurement.* Cambridge University Press, 1995.

[33] A. E. Glassgold. Spin exchange in collisions between atoms. *Phys. Rev.*, 132:2144, 1963.

[34] T. C. Ralph, S. D. Bartlett, J. L. O'Brien, G. J. Pryde, and H. M. Wiseman. Quantum nondemolition measurements for quantum information. *Phys. Rev. A*, 73:012113, 2006.

[35] K. Eng, R. N. McFarland, and B. E. Kane. Integer quantum hall effect on a six-valley hydrogen-passivated silicon (111) surface. *Phys. Rev. Lett.*, 99:016801, 2007.

[36] Rogerio de Sousa and S. Das Sarma. Electron spin coherence in semiconductors: considerations for a spin-based solid-state quantum computer architecture. *Phys. Rev. B*, 67:033301, 2003.

[37] John J. L. Morton, Alexei M. Tyryshkin, Richard M. Brown, Shyam Shankar, Brendon W. Lovett, Arzhang Ardavan, Thomas Schenkel, Eugene E. Haller, Joel W. Ager, and S. A. Lyon. Solid state quantum memory using the 31p nuclear spin. arXiv:0803.2021 [quant-ph], 2008.

[38] Rogerio de Sousa, C. C. Lo, and Jeff Bokor. Spin-dependent scattering in a silicon transistor. arXiv:0806.4638 [cond-mat.mes-hall], 2008.

[39] S. J. Park, J. A. Liddle, A. Persaud, F. I. Allen, T. Schenkel, and J. Bokor. Formation of 15nm scale coulomb blockade structures in silicon by electron beam lithography with a bilayer resist process. *J. Vac. Sci. Tech. B*, 22:3115, 2004.

[40] Ian Appelbaum, Biqin Huang, and Douwe J. Monsma. Electronic measurement and control of spin transport in silicon. *Nature*, 447:295, 2007.

[41] Berend T. Jonker, George Kioseoglou, Aubrey T. Hanbicki, Connie H. Li, and Phillip E. Thompson. Electrical spin-injection into silicon from a ferromagnetic metal/tunnel barrier contact. *Nature Physics*, 3:542, 2007.

[42] G. Feher and E. A. Gere. Electron spin resonance experiments on donors in silicon. ii. electron spin relaxation effects. *Phys. Rev.*, 114:1245, 1959.

[43] Alexei M. Tyryshkin, John J. L. Morton, Arzhang Ardavan, and S. A. Lyon. Davies electron-nuclear double resonance revisited: Enhanced sensitivity and nulcear spin relaxation. *J. Chem. Phys.*, 124:234508, 2006.

[44] S. A. Lyon and Alexei M. Tyryshkin. Private communication, 2008.

[45] M. Sanquer, M. Specht, L. Ghenim, S. Deleonibus, and G. Guegan. Coulomb blockade and transport spectroscopy in short si:mosfet. *Ann. Phys. (Leipzig)*, 8:743, 1999.

[46] A. Fujiwara, H. Inokawa, K. Yamazaki, H. Namatsu, Y. Takahashi, N. M. Zimmerman, and S. B. Martin. Single electron tunneling transistor with tunable barriers using silicon nanowire metal-oxide-semiconductor field-effect transistor. *App. Phys. Lett.*, 88:053121, 2006.

[47] H. Sellier, G. P. Lansbergen, J. Caro, S. Rogge, N. Collaert, I. Ferain, M. Jurczak, and S. Biesemans. Transport spectroscopy of a single dopant in a gated silicon nanowire. *Phys. Rev. Lett.*, 97:206805, 2006.

[48] G. D. Fuchs, V. V. Dobrovitski, R. Hanson, A. Batra, C. D. Weis, T. Schenkel, and D. D. Awschalom. Excited-state spectroscopy using single spin manipulation in diamond. *Phys. Rev. Lett.*, 101:117601, 2008.

[49] J. L. O'Brien, G. J. Pryde, A. Gilchrist, D. F. V. James, N. K. Langford, T. C. Ralph, and A. G. White. Quantum process tomography of a controlled-not gate. *Phys. Rev. Lett.*, 93:080502, 2004.

[50] Yaakov S. Weinstein, Timothy F. Havel, Joseph Emerson, Nicolas Boulant, Marcos Saraceno, Seth Lloyd, and David G. Cory. Quantum process tomography of the quantum fourier transform. *J. Chem. Phys.*, 121:6117, 2004.

[51] M. Mohseni and D. A. Lidar. Direct characterization of quantum dynamics. *Phys. Rev. Lett.*, 97:170501, 2006.

[52] Joseph Emerson, Marcus Silva, Osama Moussa, Colm Ryan, Martin Laforest, Jonathan Baugh, David G. Cory, and Raymond Laflamme. Symmetrized characterization of noisy quantum processes. *Science*, 317:1893, 2007.

[53] Mirko Lobino, Dmitry Korystov, Connor Kupchak, Eden Figueroa, Barry C. Sanders, and A. I. Lvovsky. Complete characterization of quantum-optical processes. *Science*, 322:563, 2008.

[54] R. L. Kosut. Quantum process tomography via l1-norm minimization. 2008.

[55] C. V. Helstrom. *Quantum Detection and Estimation Theory.* Academic Press, New York, 1976.

[56] A. S. Holevo. *Probabilistic and statistical aspects of Quantum Theory.* North-Holland, Amsterdam, 1982.

[57] S.L. Braunstein and C.M. Caves. Statistical distance and the geometry of quantum states. *Phys. Rev. Lett.*, 72:3439, 1994.

[58] A. Fujiwara. Quantum channel identification problem. *Phys. Rev. A*, 63:042304, 2001.

[59] Mohan Sarovar and G. J. Milburn. Optimal estimation of one-parameter quantum channels. *J. Phys. A*, 39:8487, 2006.

[60] Sergio Boixo, Steven T. Flammia, Carlton M. Caves, and JM Geremia. Generalized limits for single-parameter quantum estimation. *Phys. Rev. Lett.*, 98:090401, 2007.

[61] H. P. Yuen and M. Lax. Multiple-parameter quantum estimation and measurement of nonselfadjoint observables. *IEEE Trans. Inf. Theory*, 19:740, 1973.

[62] Akio Fujiwara and Imai Hiroshi. Quantum parameter estimation of a generalized pauli channel. *J. Phys. A*, 36:8093, 2003.

[63] A. Fujiwara and H. Nagaoka. Quantum fisher metric and estimation for pure state models. *Phys. Lett. A*, 201:119, 1995.

[64] R. L. Kosut, I. A. Walmsley, and H. Rabitz. Optimal experiment design for quantum state and process tomography and hamiltonian parameter estimation. *quant-ph/0411093*, Nov 2004.

[65] M. P. A. Branderhorst, I. A. Walmsley, and R. L. Kosut. Quantum process tomography of decoherence in diatomic molecules. In *European Conference on Lasers and Electro-Optics*, page 1, 2007.

[66] M. P. A. Branderhorst, I. A. Walmsley, R. L. Kosut, and H. Rabitz. Optimal experiment design for quantum state tomography of a molecular vibrational mode. *J. Phys. B*, 41:074004, 2008.

[67] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[68] R. L. Kosut. Quantum metrology subject to instrumentation constraints. arXiv:0803.4284 [quant-ph], 2008.

[69] Jared H. Cole, Sonia G. Schirmer, Andrew D. Greentree, Cameron J. Wellard, Daniel K. L. Oi, and Lloyd C. L. Hollenberg. Identifying an experimental two-state hamiltonian to arbitrary accuracy. *Phys. Rev. A*, 71:062312, 2005.

[70] Simon J. Devitt, Jared H. Cole, and Lloyd C. L. Hollenberg. Scheme for direct measurement of a general two-qubit hamiltonian. *Phys. Rev. A*, 73:052317, 2006.

[71] T. Schenkel, A. Persaud, S. J. Park, J. Nilsson, J. Bokor, J. A. Liddle, R. Keller, D. H. Schneider, D. W. Cheng, and D. E. Humphries. Solid state quantum computer development in silicon with single ion implantation. *J. App. Phys.*, 94:7017, 2003.

[72] Mohan Sarovar, Kevin C. Young, T. Schenkel, and K. Birgitta Whaley. Quantum nondemolition measurements of single spins in semiconductors. *Phys. Rev. B*, 78:245302, 2008.

[73] Jun Zhang, Jiri Vala, Shankar Sastry, and K.B. Whaley. Geometric theory of nonlocal two-qubit operations. *Phys. Rev. A*, 67:042313, 2003.

[74] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley-Interscience, 1991.

[75] Asher Peres. *Quantum theory: Concepts and Methods*. Springer, 1995.

[76] 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. *Phys. Rev. Lett.*, 101:207602, 2008.

[77] F. L. Lewis, L. Xie, and D. Popa. *Optimal and Robust Estimation: With an Introduction to Stochastic Control Theory*. John Wiley and Sons, 2007.

[78] P. Neumann, N. Mizuochi, F. Rempp, P. Hemmer, H. Watanabe, S. Yamasaki, V. Jacques, T. Gaebel, F. Jelezko, and J. Wrachtrup. Multipartite entanglement among single spins in diamond. *Science*, 320:1326, 2008.

[79] M. V. Gurudev Dutt, L. Childress, L. Jiang, E. Togan, J. Maze, F. Jelezko, A. S. Zibrov, P. R. Hemmer, and M. D. Lukin. Quantum register based on individual electronic and nuclear spin qubits in diamond. *Science*, 316:1312, 2007.

[80] Fuzhen Zhang. *The Schur Complement and its Applications*. Springer, 2005.

[81] Kaveh Khodjasteh and Daniel A. Lidar. Performance of deterministic dynamical decoupling schemes: Concatenated and periodic pulse sequences. *Physical Review A (Atomic, Molecular, and Optical Physics)*, 75(6):062310, 2007.

[82] Götz S. Uhrig. Keeping a quantum bit alive by optimized pi-pulse sequences. *Physical Review Letters*, 98(10):100504, 2007.

[83] Mikko Möttönen, Rogerio de Sousa, Jun Zhang, and K. Birgitta Whaley. High-fidelity one-qubit operations under random telegraph noise. *Physical Review A (Atomic, Molecular, and Optical Physics)*, 73(2):022332, 2006.

[84] P. Rebentrost and F. K. Wilhelm. Optimal control of a leaking qubit. *Phys. Rev. B*, 79(6):060507, Feb 2009.

[85] O.-P. Saira, V. Bergholm, T. Ojanen, and M. Möttönen. Equivalent qubit dynamics under classical and quantum noise. *Physical Review A (Atomic, Molecular, and Optical Physics)*, 75(1):012308, 2007.

[86] Navin Khaneja, Timo Reiss, Cindie Kehlet, Thomas Schulte-Herbrggen, and Steffen 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.

[87] Andrew M. Steane. Overhead and noise threshold of fault-tolerant quantum error correction. *Phys. Rev. A*, 68(4):042322, Oct 2003.

[88] J. L. Doob. The brownian movement and stochastic equations. *Ann. Math.*, 43:351 – 369, 1942.

[89] R. Andreani, E. G. Birgin, J. M. Martínez, and M. L. Schuverdt. On augmented lagrangian methods with general lower-level constraints. *SIAM Journal on Optimization*, 18(4):1286–1309, 2007.

[90] R. Andreani, E. G. Birgin, J. M. Martinez, and M. L. Schuverdt. Augmented lagrangian methods under the constant positive linear dependence constraint qualification. *Mathematical Programming*, 111(1-2):5–32, Jan 2008.

[91] I. Najfeld and T. F. Havel. Derivatives of the matrix exponential and their computation. *Advances in Applied Mathematics*, 16(3):321 – 375, 1995.

[92] Rogerio de Sousa. Dangling-bond spin relaxation and magnetic 1/f noise from the amorphous-semiconductor/oxide interface: Theory. *Physical Review B (Condensed Matter and Materials Physics)*, 76(24):245306, 2007.

[93] S.-Y. Paik, S.-Y. Lee, W. J. Baker, D. R. McCamey, and C. Boehme. $t1$ and $t2$ spin relaxation time limitations of phosphorous donor electrons near crystalline silicon to silicon dioxide interface defects. *Phys. Rev. B*, 81(7):075214, Feb 2010.

[94] T. Schenkel, J. A. Liddle, A. Persaud, A. M. Tyryshkin, S. A. Lyon, R. de Sousa, K. B. Whaley, J. Bokor, J. Shangkuan, and I. Chakarov. Electrical activation and electron spin coherence of ultralow dose antimony implants in silicon. *Applied Physics Letters*, 88(11):112101, 2006.

[95] A. M. Tyryshkin, S. A. Lyon, A. V. Astashkin, and A. M. Raitsimring. Electron spin relaxation times of phosphorus donors in silicon. *Phys. Rev. B*, 68(19):193207, Nov 2003.

[96] S Gerschgorin. über die Abgrenzung der Eigenwerte einer Matrix. *Izv. Akad. Nauk. USSR Otd. Fiz.-Mat. Nauk*, 7:749–754, 1931.

[97] R. J. Schoelkopf, A. A. Clerk, S. M. Girvin, K. W. Lehnert, and M. H. Devoret. Qubits as Spectrometers of Quantum Noise. *ArXiv Condensed Matter e-prints*, October 2002.

[98] A.M. Tyryshkin, S.A. Lyon, T. Schenkel, J. Bokor, J. Chu, W. Jantsch, F. Schffler, J.L. Truitt, S.N. Coppersmith, and M.A. Eriksson. Electron spin coherence in si. *Physica E: Low-dimensional Systems and Nanostructures*, 35(2):257 – 263, 2006. Proceedings of the 14th International Winterschool on New Developments in Solid State Physics -Charges and spins in nanostructures: basics and devices.

[99] Lorenza Viola, Emanuel Knill, and Seth Lloyd. Dynamical decoupling of open quantum systems. *Phys. Rev. Lett.*, 82(12):2417–2421, Mar 1999.

[100] Ryogo Kubo. Generalized cumulant expansion method. *Journal of the Physical Society of Japan*, 17(7):1100–1120, 1962.

[101] Thaddeus Ladd. Qubit relaxation and dynamical decoupling. 2009.

[102] Petre Stoica and Randolph L. Moses. *Introduction to spectral analysis*. Prentice Hall, 1997.

[103] Charles L. Lawson and Richard J. Hanson. *Solving least squares problems*. Prentice-Hall, Englewood Cliffs, NJ, 1974.

[104] Michael J. Biercuk, Hermann Uys, Aaron P. VanDevender, Nobuyasu Shiga, Wayne M. Itano, and John J. Bollinger. Experimental uhrig dynamical decoupling using trapped ions. *Phys. Rev. A*, 79(6):062324, Jun 2009.

[105] Å Björck. Numerics of gram-schmidt orthogonalization. *Linear Algebra and its Applications*, 197-198:297 – 316, 1994.

[106] D. Porras and J. I. Cirac. Effective quantum spin systems with trapped ions. *Phys. Rev. Lett.*, 92:207901, 2004.

[107] M. Lewensteina, A. Sanperab, V. Ahufingerb, B. Damskic, and U. Send, B. Sen. Ultracold atomic gases in optical lattices: mimicking condensed matter physics and beyond. *Adv. Phys.*, 56:243, 2007.

[108] M. Greiner, O. Mandel, T. Esslinger, T.W. Hnsch, and I. Bloch. Quantum phase transition from a superfluid to a mott insulator in a gas of ultracold atoms. *Nature*, 39:415, 2002.

[109] Thilo Stöferle, Henning Moritz, Christian Schori, Michael Köhl, and Tilman Esslinger. Transition from a strongly interacting 1d superfluid to a mott insulator. *Phys. Rev. Lett.*, 92(13):130403, Mar 2004.

[110] I. B. Spielman, W. D. Phillips, and J. V. Porto. Condensate fraction in a 2d bose gas measured across the mott-insulator transition. *Phys. Rev. Lett.*, 100(12):120402, Mar 2008.

[111] Büchler et al. Three-body interactions with cold polar molecules. *Nat. Phys.*, 3:726–731, 2007.

[112] A. Kitaev. Fault-tolerant quantum computation by anyons. *Ann. Phys.*, 303:2, 2003.

[113] Chetan Nayak, Steven H. Simon, Ady Stern, Michael Freedman, and Sankar Das Sarma. Non-abelian anyons and topological quantum computation. *Rev. Mod. Phys.*, 80(3):1083–1159, Sep 2008.

[114] Z. Nussinov and G. Ortiz. Autocorrelations and thermal fragility of anyonic loops in topologically quantum ordered systems. *Phys. Rev. B*, 77:064302, 2008.

[115] C. Castelnovo and C. Chamon. Entanglement and topological entropy of the toric code at finite temperature. *Phys. Rev. B*, 76:184442, 2007.

[116] K.D. Nelson, Xiao Li, and David S. Weiss. Imaging single atoms in a three-dimensional array. *Nature Phys.*, 3:556, 2007.

[117] D. Jaksch et al. Fast quantum gates for neutral atoms. *Phys. Rev. Lett.*, 85(10):2208–2211, Sep 2000.

[118] M. Cozzini, T. Calarco, A. Recati, and P. Zoller. Fast rydberg gates without dipole blockade via quantum control. *Optics Communications*, 264(2):375 – 384, 2006. Quantum Control of Light and Matter - In honor of the 70th birthday of Bruce Shore.

[119] S. B. Bravyi and A. Yu. Kitaev. Quantum codes on a lattice with boundary, 1998.

[120] Michael H. Freedman and David A. Meyer. Projective plane and planar quantum codes. *Foundations of Computational Mathematics*, 1:325–332, 2001. 10.1007/s102080010013.

[121] K. C. Young et al. (In preparation).

[122] S. Klarsfeld and J. A. Oteo. Recursive generation of higher-order terms in the magnus expansion. *Phys. Rev. A*, 39(7):3270–3273, Apr 1989.

[123] D. Weiss. 2009. private communication.

[124] J. Zhang et al. Exact two-qubit universal quantum circuit. *Phys. Rev. Lett.*, 91:027903, 2003.

[125] Kenneth R. Brown. Energy protection arguments fail in the interaction picture. *Phys. Rev. A*, 76(2):022327, Aug 2007.

[126] S. Diehl et al. Quantum states and phases in driven open quantum systems with cold atoms. *Nat. Phys.*, 4:878–883, 2008.

[127] B. Kraus, H. P. Büchler, S. Diehl, A. Kantian, A. Micheli, and P. Zoller. Preparation of entangled states by quantum markov processes. *Phys. Rev. A*, 78(4):042307, Oct 2008.

[128] D. J. Wineland et al. Double-resonance and optical-pumping experiments on electromagnetically confined, laser-cooled ions. *Opt. Lett.*, 5(6):245, 1980.

[129] S. Trebst et al. Breakdown of a topological phase: Quantum phase transition in a loop gas model with tension. *Phys. Rev. Lett.*, 98:070602, 2007.

[130] E. A. Hessels, D. M. Homan, and M. J. Cavagnero. Two-stage Rydberg charge exchange: An efficient method for production of antihydrogen. *Phys.Rev. A*, 57:1668, 1998.

[131] Waseem S. Bakr, Jonathon I. Gillen, Amy Peng, Simon Foelling, and Markus Greiner. A quantum gas microscope - detecting single atoms in a hubbard regime optical lattice. 2009.

[132] Hendrik Weimer, Markus Mller, Igor Lesanovsky, Peter Zoller, and Hans Peter Bchler. Digital coherent and dissipative quantum simulations with rydberg atoms. *Nature Phys.*, 6:382–388, 2010.

[133] Jun Zhang and K. Birgitta Whaley. Generation of quantum logic operations from physical hamiltonians. *Phys. Rev. A*, 71(5):052317, May 2005.

[134] T. R. Beals, J. Vala, and K. B. Whaley. Scalability of quantum computation with addressable optical lattices. *Phys. Rev. A*, 77(5):052309, May 2008.

[135] M. Saffman, T. G. Walker, and K. Molmer. arXiv:quant-ph/0909.4777.

[136] T. F. Gallagher. *Rydberg Atoms*. Cambridge Univ. Press, 1994.

[137] M. W. Horbatsch, M. Horbatsch, and E. A. Hessels. A universal formula for the accurate calculation of hydrogenic lifetimes. *J. Phys. B*, 38:1765, 2005.

[138] E. Urban, T. A. Johnson, T. Henage, L. Isenhower, D. D. Yavuz, T. G. Walker, and M. Saffman. Observation of rydberg blockade between two atoms. *Nature Physics*, 5:110, 2009.

[139] M. Saffman and T. G. Walker. Analysis of a quantum logic device based on dipole-dipole interactions of optically trapped Rydberg atoms. *Phys.Rev. A*, 72:022347, 2005.

[140] L. Isenhower et al. Demonstration of a neutral atom controlled-NOT quantum gate. arXiv:quant-ph/0907.5552.

[141] E. S. Chang. Radiative lifetime of hydrogenic and quasihydrogenic atoms. *Phys.Rev. A*, 31:495, 1985.

[142] J. Liang, M. Gross, P. Goy, and S. Haroche. Circular Rydberg-state spectroscopy. *Phys.Rev. A*, 33:4437, 1986.

[143] R. G. Hulet and D. Kleppner. Rydberg atoms in "circular" states. *Phys.Rev. Lett.*, 51:1430, 1983.