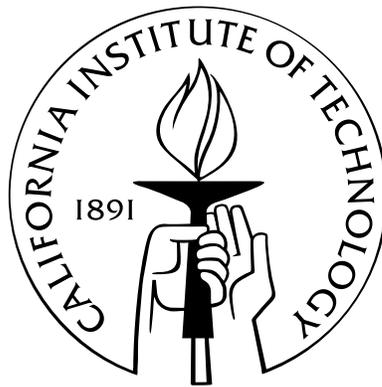# Continuous Quantum Measurement
# of Cold Alkali-Atom Spins

Thesis by

John K. Stockton

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2007

(Defended December 11, 2006, Adviser: Prof. Hideo Mabuchi)

ii

All electronic versions of this thesis (original and corrected) and relevant supplemental

information can be found at

`http://qwiki.caltech.edu/wiki/John_K._Stockton/Thesis`.

Acknowledgements

Throughout my time at Caltech, Hideo Mabuchi has been a great research advisor, teacher, and friend. Somehow Hideo manages to maintain professional excellence by the usual measures, while never compromising his kindness, creativity, nor the intellectual freedom he allows all of his students. He has given me the opportunity to experience what will be one of the best scientific and personal periods of my life, and I will always be appreciative.

One of Hideo's primary skills is convincing high-quality people to join his group. I will always fondly recall the initial few years working with Mike Armen, John Au, Andy Berglund, and Ben Lev, who have been and will be great friends and collaborators. I am extremely indebted to Andrew Doherty for teaching me most of the quantum optics I know during so many unscheduled intrusions to his paper-covered office. Thanks to JM Geremia for contributing to much of the experimental and theoretical work described in this thesis. I am grateful to Ramon Van Handel for keeping me rigorous. I would thank Kevin McHale, but not until we get to writing that basketball paper together. Thanks to Anthony Miller for keeping the group keen and continuing what we have begun. Sorry to everyone else in the lab for grouping you like this, but thanks to Luc Bouten, Andre Conjusteau, Nicole Czakon, Nathan Hodas, Asa Hopkins, Joe Kerckhoff, Chungsok Lee, Tim McGarvey, Alexandra Olaya-Castro, Ben Rahn, Gopal Sarma, Andrew Silberfarb, Jen Sokol, Felipe Caycedo Soler, Jon Williams, and Naoki Yamamoto. You were all extremely helpful and friendly labmates. Sheri, thanks for your daily banter and doing all the real work. I will never forget Moutai, foosball, and our perfect record $\hat{a}$ team.

Thanks also to all of the other Caltech folk I am fortunate enough to call friends: Dave Bacon, Paul Barclay, Parsa Bonderson, Megan Eckart, Chris Lee, Nathan Lundblad, Paige Randall, Kartik Srinivasan, Devin Sullivan, Ian Swanson, and Joe Zadeh. Kevin Birnbaum in particular has been a great roommate and collaborator. It was also nice to have a few non-Caltech friends to hang out with every once in a while, including James Hong, Kim

Keith, Amit Sinha, and Craig Sipes. And, of course, thanks to Ernie for feeding me nearly every day.

At its best, science is a friendly and collaborative endeavor, and I have been lucky to work in a community with so many brilliant and generous minds. Jeff Kimble and his lab members have been invaluable neighbors, especially Joe Buck, Jimbo Chou, Theresa Lynn, and Jason McKeever. Howard Wiseman has been an indispensable theory collaborator, supplying essentially every core idea for our experiments. Salman Habib, Poul Jessen, Ivan Deutsch, Ari Tuchman, Mark Kasevich, Klaus Mølmer, Steven Van Enk, Guifre Vidal, Eugene Polzik, John Preskill, Masahiro Yanagisawa, Ken Libbrecht, Dmitry Budker, and their students have all been a continuous source of inspiration and advice.

I am indebted to the John and Fannie Hertz Foundation, which has not only given me the financial means to pursue my passion independently, but has also demonstrated a genuine interest in their students through regular community events.

Not for the last time, I thank Wajma for loving me, and her family for giving me a feeling of home I never thought I would have in LA. Finally, I acknowledge my parents (Joan and Steve), their parents (Alice, Andy, Myrtle, and Doyle), my brothers (Mark and Matt), and the rest of my family. They have given me the love and support I have needed to become who I am today, and enough humor to keep me grounded throughout. I will always be grateful.

Abstract

The field of quantum metrology concerns the physical measurement of sensors with a precision comparable to fundamental limits set by quantum mechanics. It is possible to outperform naïve interpretations of these limits by using entangled states of the sensor system. One example is that of a spin-squeezed state, in which the uncertainty of one variable is decreased at the expense of another while still obeying Heisenberg's uncertainty principle, improving rotation sensitivity along a chosen axis. These states are potentially useful in devices including atomic clocks, inertial sensors, and magnetometers.

Any model of a quantum metrology device must respect the fact that physical measurements are not passive, as imagined classically, but necessarily invasive. Far from being a negative feature, well-understood quantum measurement can conditionally drive a system into desirable entangled states, including spin-squeezed states. Furthermore, the fundamental randomness of this process can, in principle, be removed with real-time feedback control, motivating an adaptation of classical feedback concepts to the quantum realm.

In this thesis, I describe these ideas in the context of one experimental example. A laser-cooled cloud of cesium spins is polarized along one axis via optical pumping and, subsequently, a linearly polarized far-off resonant probe beam traverses the sample. Due to the interaction Hamiltonian, the optical polarization rotates by an amount nominally proportional to one spin component of the collective spin state, enacting a weak, continuous, nondemolition measurement of that collective variable. This optical Faraday rotation is then measured with a polarimeter and the inherently noisy result used to condition the collective atomic state via a quantum filter, or stochastic master equation. Ideally, this process is capable of producing spin-squeezed states via the measurement itself.

The details of this measurement are investigated in depth, including a derivation of the nonideal polarizability Hamiltonian, an analysis of the projection process with control, and a derivation of the magnetometry sensitivity. Experimentally, we demonstrate continuous

measurement of the collective spin state with a large single-shot signal-to-noise ratio and verify many predictions of the model. Finally, we describe attempts to observe the atomic projection noise, which would infer the preparation of spin-squeezed states.

Contents

ix

List of Figures

Chapter 1

Introduction

Besides the ability to write successful grant proposals, the most important skill of a scientist is the ability to shamelessly ask simple questions until it becomes clear what is, and what is not, known about a particular subject. For me, the appeal of physics is the fact that seemingly simple questions, however well disguised in details, continue to both confuse and excite those at the leading edge of research.

Prior to graduate school, I became interested with the degree of confusion that basic questions concerning the role of measurement in quantum mechanics seemed to elicit from people whom I thought should know better. It seemed that the confusion was greatest the further removed the discussion was from a particular experiment. Without the context of a real problem, conversations about quantum measurement and physical uncertainty would become fruitless and, even worse, philosophical. Thus, I was pushed towards increasingly applied questions of the theory, grounded in technological relevance. I became less interested nontestable philosophical questions and even in testable questions that aim to confirm the theory, many of which have been examined in landmark experiments over the last century. Rather, I was happy to assume the theory and use it to do something of practical use.

With this attitude, I ended up joining Prof. Hideo Mabuchi's group at Caltech primarily for the opportunity to perform experiments motivated by real, technical quantum measurement problems. Hideo's program was especially appealing for the fact that it aimed to unify perspectives of quantum measurement to engineering-minded modes of thought from classical estimation and control theories. Further, it seemed a fun goal to control quantum systems by interpreting the noisy measurement record and using feedback to adjust an actuator in real time, despite the presence of inherent uncertainty. If this could be done successfully, we would not have to convince others that we understood the measurement

process because we would, in effect, be convincing the system itself.

Although these principles are applicable to many technologies, my interests have also become focused on the subfield of *quantum metrology,* which concerns the use of quantum systems as sensors. Like quantum communication, the goals in this field are generally more near-term than those of quantum computation. In this thesis, I investigate one quantum metrology problem in depth, as I describe the continuous optical measurement of a cold atomic spin cloud near fundamental noise limits, with immediate applications to magnetometry. As specific as this treatment is, hopefully it will be of use to researchers seeking a practical implementation of general quantum measurement concepts.

1.1   Getting to the Point

After several related projects, I eventually came to focus on an research problem that essentially amounts to, "How well can one measure a magnetic field with a compass needle?" This is related to the question of, "How well can one measure the direction that a compass needle points?" In this case, the compass needle of interest consists of a collection of laser-cooled atomic spins all polarized initially along a known direction. (For a discussion of how gas phase atomic spins differ dynamically from an actual compass needle, where the spins are fixed to an inertial mass, see [1].) An unknown field can be inferred from a measurement of the magnetic field induced Larmor precession of the spins. In practice, the measurement is inherently uncertain because quantum mechanics imposes a limit upon how well one can measure the direction that a given number of spins point in space. This seemingly specific scenario is actually representative of the general field of quantum metrology, each application of which essentially concerns the measurement of rotation angles in the presence of unavoidable system noise. Many useful devices fall under this basic category, including magnetometers, atomic clocks, and inertial gravity sensors.

Now, to get more specific about how well a collection of spins can point, consider a *coherent spin-state* (CSS), which has all spins aligned along one direction. For a cloud of N spin-1/2 atoms aligned along the x-axis, this state can be represented as

$$|\Psi\rangle = |\uparrow_1 \uparrow_2 \cdots \uparrow_N\rangle_x \,, \tag{1.1}$$
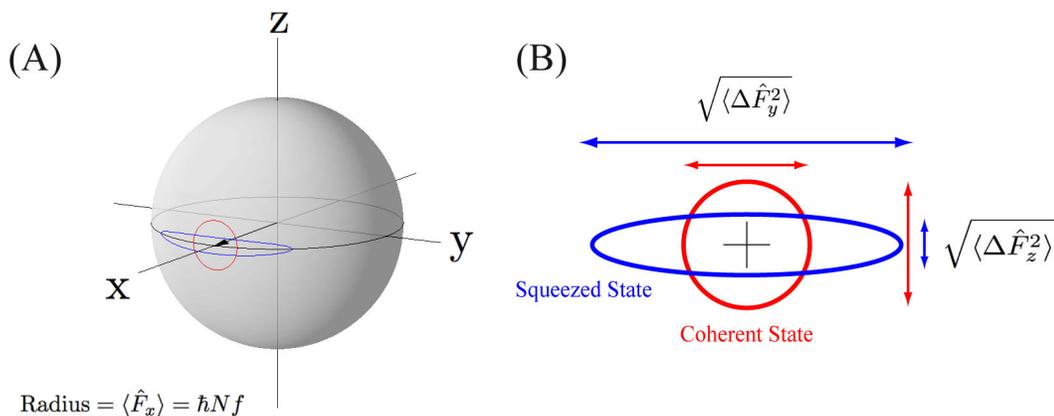
Figure 1.1: (A) Graphical representation of the spin-polarized atomic sample as a classical magnetization vector with transverse quantum uncertainty. (B) Schematic of the transverse quantum uncertainties for coherent spin-states and spin-squeezed states.

such that the expectation values of the collective spin operators are $\langle \hat{F}_{\mathrm{x}} \rangle = \hbar N/2$ and $\langle \hat{F}_{\mathrm{y}} \rangle = \langle \hat{F}_{\mathrm{z}} \rangle = 0$. From the commutation relations of the spin operators, we have the Heisenberg uncertainty relation

$$\langle \Delta \hat{F}_y^2 \rangle \langle \Delta \hat{F}_z^2 \rangle \geq \frac{\hbar^2 \langle \hat{F}_x \rangle^2}{4}, \tag{1.2}$$

which, for the aligned spin-state, gives

$$\langle \Delta \hat{F}_y^2 \rangle = \langle \Delta \hat{F}_z^2 \rangle = \frac{\hbar^2 N}{4}. \tag{1.3}$$

As depicted in figure 1.1, these uncertainties indicate how fuzzy the tip of the atomic compass needle is for a given number of atoms. Thus the "sharpness" of the state can be given by the size of the uncertainty disk divided by the length of the spin vector

$$\frac{\sqrt{\langle \Delta \hat{F}_z^2 \rangle}}{\langle \hat{F}_{\mathrm{x}} \rangle} = \frac{1}{\sqrt{N}}. \tag{1.4}$$

As a result, we see that with more atoms one can measure rotations of the spin-state more precisely. Often, we only care about the rotation of the atoms in one direction. In this case, the performance can be improved beyond the above limit by squeezing the distribution along one direction and antisqueezing it along the irrelevant direction such that the equality of the

4

above uncertainty relation is maintained but $\langle \Delta \hat{F}_z^2 \rangle < \langle \Delta \hat{F}_y^2 \rangle$. These states are known as *spin-squeezed states* (SSS) and were introduced in the context of improving the performance of atomic spectroscopy applications by Kitagawa and Ueda [2, 3] and the Wineland group at NIST [4, 5]. (The idea of spin-squeezed states detailed by Kitagawa and Ueda in [2] was actually first introduced in their earlier paper regarding fermionic interferometry [3].) These ideas adapted the well-established idea of optical squeezing to the atomic domain. The cartoon in figure 1.1 depicts the uncertainty distribution difference between coherent and spin-squeezed states. One can show that individual spin-1/2 particles comprising a spin-squeezed state are entangled with each other [6], lending some credence to the oft-made statement that entanglement is a useful resource for quantum information tasks. (However, when it comes to similar tasks, such as communicating the direction of a reference frame, there remain many interesting questions about the entanglement and symmetry properties of the most useful states [7].)

Given the possibility of spin-squeezed states, we now begin to address the issues of how to create and use them for practical tasks. For example, given an entangled state that is rotating due to a field, what is the best possible way to estimate that field using measurements of the state? How should this state be physically observed during or after the field induced rotation? Also, how does one actually produce an entangled spin-squeezed state? Remarkably, one way to create an entangled SSS from the coherent state above is through a collective measurement of the spin-state. Thus measurement can be used as both a way of reading the state and shaping it.

The fact that measurement of a collective variable can randomly prepare an entangled spin-state is easy to demonstrate. Consider the following simple example with two spin-1/2 particles both aligned along the x-direction, but with the total state represented in the z-basis

$$
\begin{aligned}
|\Psi\rangle_0 &= |\uparrow\rangle_x |\uparrow\rangle_x \\
&= \frac{1}{\sqrt{2}} (|\uparrow\rangle_z + |\downarrow\rangle_z) \frac{1}{\sqrt{2}} (|\uparrow\rangle_z + |\downarrow\rangle_z) \\
&= \frac{1}{2} |\uparrow\rangle_z |\uparrow\rangle_z + \frac{1}{\sqrt{2}} \left(\frac{1}{\sqrt{2}} (|\uparrow\rangle_z |\downarrow\rangle_z + |\downarrow\rangle_z |\uparrow\rangle_z) \right) + \frac{1}{2} |\downarrow\rangle_z |\downarrow\rangle_z. \quad (1.5)
\end{aligned}
$$

Now if we perform a measurement of the collective spin in the z-direction, $\hat{F}_z$, we will get one

of three results: $+1$, -1, or 0. The measurement will return the value $+1$ with probability $1/4$, in which case the postmeasurement state will be $|\Psi\rangle_{c,+1} = |\uparrow\rangle_z|\uparrow\rangle_z$. Similarly, the measurement will return the value -1 with probability $1/4$, with a postmeasurement state of $|\Psi\rangle_{c,-1} = |\downarrow\rangle_z|\downarrow\rangle_z$. Both of these states are unentangled and equal to the initial state rotated by 90 degrees. Finally, the measurement will return the value 0 with probability $1/2$ and give a postmeasurement state of

$$|\Psi\rangle_{c,0} = \frac{1}{\sqrt{2}} \left(|\uparrow\rangle_z|\downarrow\rangle_z + |\downarrow\rangle_z|\uparrow\rangle_z\right). \tag{1.6}$$

This state (the $m = 0$ state of the spin-triplet) is clearly entangled because it cannot be factored into a product of states for the two spins. Also notice that due to the symmetry of the initial state and the symmetry of the measurement, all of the (triplet) postmeasurement states are also symmetric. In chapter 4, we consider how to take advantage of this symmetry to compute entanglement measures for large numbers of spins.

This example represents one of the core ideas of this thesis, but the experiment we consider has several key differences. First, the experimental measurement is weak and continuous rather than strong and discrete. To infer the spin direction, an optical probe beam traverses and dispersively interacts with the atoms, and the resulting Faraday rotation of the optical polarization is measured with a polarimeter. Instead of imposing the full projection at once, this measurement allows us to acquire information slowly such that the state diffusively and gradually projects onto an entangled subspace. Second, we consider many atoms (typically billions) so that the entanglement is mostly related to spin-squeezing in the Gaussian limit. Finally, we use realistic atoms with many sublevels (cesium), which affects both our model of the atom-light interaction and our representation of the collective states. A depiction of the Faraday rotation measurement we employ in the lab is shown in figure 1.2.

As opposed to the discrete example above, with continuous measurement the induced projection is dilated in time. This allows us to roughly observe where the state is projecting and, if we do not like it, apply feedback control to guide the projection into a more desirable location. Feedback can also enhance the measurement of magnetic fields. In chapter 10, we discuss how a field may be estimated concurrently with the squeezing via measurement and also how feedback can make this process robust to model uncertainties.
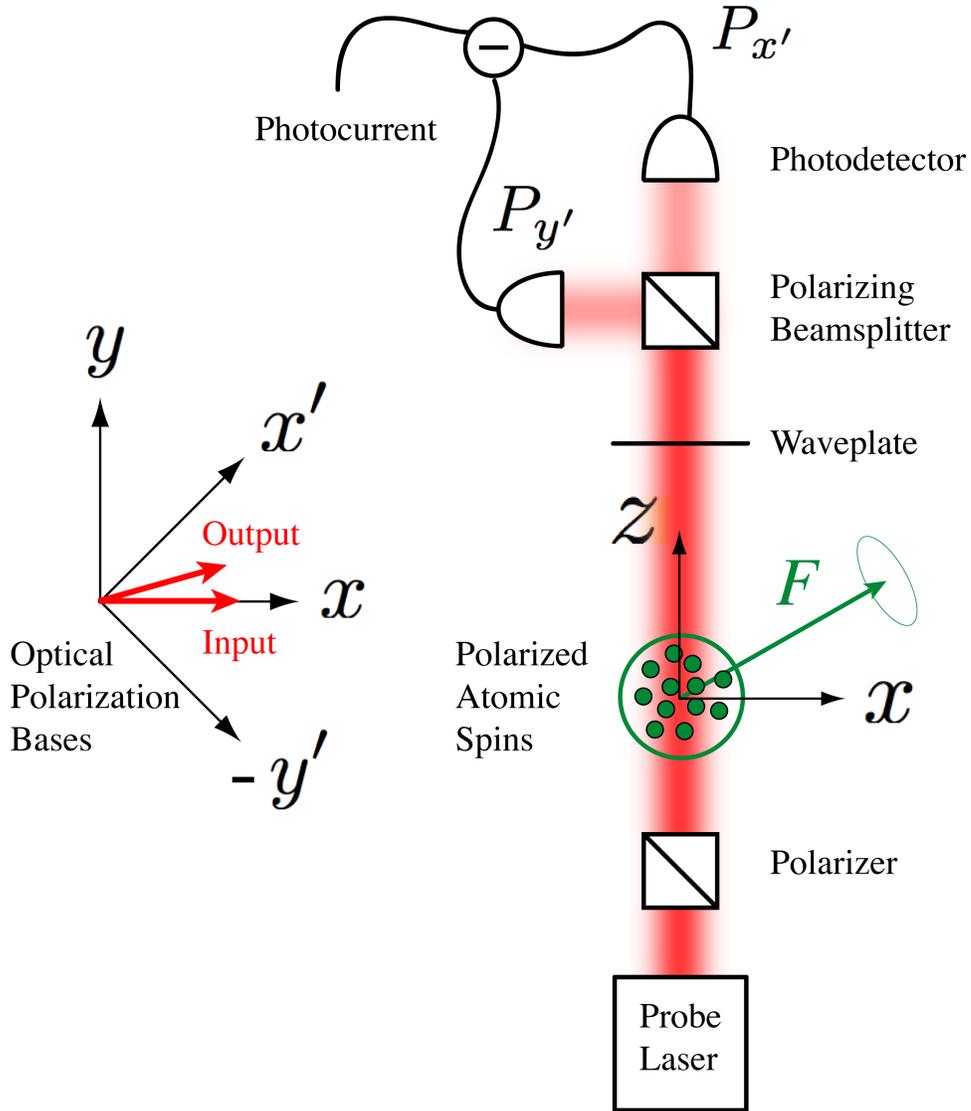
Figure 1.2: A basic schematic of our experimental apparatus. A polarized beam of off-resonant light traverses the cold atomic cloud and the resulting Faraday rotation is measured with a polarimeter.

Although it is these general ideas of quantum measurement and feedback control that motivate us, a large part of this thesis is devoted to the accurate modeling of the atom-light interaction. The true interaction consists of a component that leads to the idealized spin-squeezing measurement (a collective measurement of $\hat{F}_z$ via the vector component of the polarizability Hamiltonian), but also a complicating component (the tensor component of the polarizability Hamiltonian) that results from the fact that the cesium spin we use is $f = 4$ and not $1/2$. Depending on the context, we either assume the simplified ideal Hamiltonian, or the realistic full version.

Of course, this is largely an experimental thesis, with the experimental goal of confirming the model and observing the atomic projection noise that implies the preparation of spin-squeezed states. This is a challenging endeavor that necessitates the preparation of cold atomic ensembles with large optical depth. This pursuit and attempts to hammer down technical sources of noise with modulation schemes are detailed in depth only after the theoretical foundation has been set.

1.2 Personal Lab History

I decided to attend Caltech primarily to work in Hideo's group and began immediately upon my arrival in the summer of 2000. I worked on several different projects before beginning to research what turned out to be the core of this thesis. Here I describe the diffusive trajectory of my research career at Caltech.

1.2.1 Initial Projects

I began to ramp up to speed theoretically by considering a paper investigating potential bound spin-states of an experimentally relevant magnetic potential. Due to concerns of single-valuedness, we discovered that this paper's results were questionable and composed a comment to that paper contained in appendix F. Around the same time there was interest in our group about extending models of quantum position measurement [8] to classically chaotic systems with phase-space islands of stability and ergodicity [9].

I soon ventured into the lab and started working on a field-programmable gate array (FPGA) system, which was viewed as a potentially key component in the quantum control experiments we planned to perform. In general a quantum controller must take a classical

measurement result, use quantum estimation theory to map that result to an appropriate feedback response, and deliver that response all within the coherence time of the physical system. Typically such a quantum controller must be capable of nonlinear calculations and impose a minimal delay on the feedback signal. It is also desirable to be able to reprogram the controller algorithm over reasonably fast timescales. Many programmable logic devices approximately fit this bill, and in [10] we discuss and demonstrate our particular FPGA system, locking a Fabry-Perot cavity as an example application.

Concurrently with the general FPGA effort, I worked with Mike Armen and John Au on implementing a proposal by Howard Wiseman to measure the phase of a light pulse using an adaptive homodyne procedure, where the measurement is optimized by mapping the measured photocurrent via the FPGA to the local oscillator phase in real time [11]. In the resulting paper [12], we showed that for a certain set of photon numbers in a pulse, this measurement procedure was capable of beating the standard heterodyne limit. More on this experiment can be found in [13]. While developing the FPGA system, I also collaborated with Kevin Birnbaum in the Kimble group on implementing an FPGA algorithm to radially cool atoms falling through their cavity [14].

In the adaptive phase measurement experiment, we were measuring coherent states of the optical field, and because these states are free of correlation between photons, the back-action and conditional quantum dynamics are relatively trivial. After demonstrating the effectiveness of the adaptive measurement, we were thus tempted to enrich the experiment by instead measuring correlated photon states, or squeezed optical states, emanating from an optical parametric oscillator (OPO). After some consideration of this potential experiment, we ultimately decided that the coherence timescales were too short to perform a measurement and feedback experiment effectively with available technology. Nevertheless, a theoretical analysis of the conditional measurement of the squeezed optical state within an OPO is considered in appendix C following another paper by Wiseman.

1.2.2 MOT Building

Around that time, Wiseman's group published yet another proposal involving the continuous measurement and control of an atomic spin ensemble, with the goal of producing spin-squeezed states deterministically [15, 16]. That idea provided the main motivation for this thesis and throughout this work we adapt the concept to our experiment, with particular

concern for modeling the actual atom-light interaction for an optical probe beam interacting with a cloud of multilevel cesium atoms.

The proposal also gave me the opportunity to build an atomic physics experiment from scratch, which provided a substantial educational incentive. Towards the end of 2001, I began to build the experiment including the vacuum and magnetic subsystems, owing much to the advice of Ben Lev. With substantial help from Andrew Berglund, who helped build the trapping diode lasers, we achieved a magneto-optical trap (MOT) in February of 2002. The basic experimental apparatus is described in detail in chapter 11.

1.2.3 Theoretical Offshoots

Throughout the experimental effort, we have spent a considerable amount of time looking into various theoretical aspects of the experiment, writing papers as a means of teaching ourselves the relevant physics. First, because we were attempting to create spin-squeezed states, which are internally entangled, we began to ask ourselves how to describe the entanglement when billions of particles are involved. These questions eventually led to the paper [17], summarized in chapter 4, wherein we investigate the entanglement between subsamples of an ensemble that is symmetric with respect to exchange of particles. In this pursuit we discovered some highly entangled families of states living in this space and also derived ways of numerically calculating entanglement measures for large ensembles by taking advantage of symmetry constraints.

Because spin-squeezed states are supposed to be useful for improving the precision of rotation measurements, we also became interested in how to measure magnetic fields with continuous measurement of the atomic spin-state. As opposed to the usual sequential procedure where the state is initialized, rotated, then projectively measured, we investigated a procedure where the state is squeezed via continuous measurement but rotated by the unknown field concurrently. The resulting *quantum parameter estimation* analysis, essentially relying on Kalman filters and Bayesian estimation to perform magnetometry, is described in [18] and was largely inspired by [19]. The addition of feedback to make the procedure robust to classical parameter uncertainties is analyzed in [20]. These results are summarized in chapter 10.

If spontaneous emission can be avoided (which it can never be at long timescales), the continuous measurement equation (stochastic Schrodinger equation, or SSE) of [15] de-

scribes a continuous version of a measurement that projects the atoms onto an eigenstate of the collective angular momentum, i.e., a Dicke state. In [21], we investigate this equation theoretically and numerically demonstrate that, just as feedback can enhance the preparation of spin-squeezed states, feedback can also make the preparation of particular Dicke states deterministic. In effect, whether or not the state is Gaussian, feedback can be used to direct the state diffusion caused by measurement. These ideas, summarized in chapter 9, were made more analytically precise for few spin systems, mainly by Ramon Van Handel, in [22]. More recently, Ramon and I have written a more general and mathematical review of these ideas in [23].

We have also spent a great deal of time modeling the atom-light interaction physics, much of which is presented here for the first time. The core ideas are presented linearly in chapters 5–8 although, historically, the order in which we understood the physics was out of phase with our experimental findings. Much of the point of this thesis is to tell the combined theoretical and experimental story in logical order.

1.2.4   Experimental Continuous Measurement

In the beginning of 2003, we began to observe the spin-state of the atom cloud via Faraday rotation of the probe beam. This allowed us to observe both the Larmor precession (due to a constant magnetic field) and decay dynamics of the spin-state (due to spontaneous emission among other processes) over millisecond timescales. Around the same time, Poul Jessen's group performed similar experiments and described certain effects in terms of the tensor polarizability interaction Hamiltonian [24]. Subsequently, they, along with the Polzik group, started to investigate the non-trivial effects of the tensor Hamiltonian on the measurement dynamics [25, 26].

During this time I was working on the theoretical papers above and also worrying about controlling the magnetic fields in our experiment (see appendix A and chapter 11). Toward the end of 2003 and beginning of 2004, with JM Geremia leading the data-taking effort, we published one paper on experimental spin-squeezing [27] and one on "sub-shotnoise" magnetometry [28]. Although the theoretical ideas in those papers are valid, the experimental results are now considered questionable and inconsistent with later findings, in part due to the lack of consideration of the tensor terms. More on these papers can be found in section 14.4.

Our own effort to characterize the effect of the tensor terms eventually resulted in reference [29]. This effort was also flawed initially, in both the experimental and theoretical results, but was subsequently corrected before publication. The results of the theoretical corrections can be found in chapter 6. In the subsequent chapter 7, we consider an unconditional master equation derivation based upon the entire tensor Hamiltonian, which presents surprising results. Around this time we also started to appreciate the positive side of the tensor Hamiltonian, which allows us to measure the pumping efficiency of the quantum state and also align the optical polarization to the atomic state. Without the nonlinear (quadratic) terms of the full Hamiltonian, these practical tasks would be impossible. In chapter 13, we include experimental results from the corrected version of [29] that confirm the tensor model. We also present several new measurements including spectra of the tensor Hamiltonian coefficients and a measurement of the the optical pumping efficiency using the tensor terms.

Much of my subsequent time in lab was spent developing techniques to characterize and increase the optical depth of the ensemble. I also began to develop a modulated version of the experiment that avoids the low-frequency noise that plagued previous versions. During this time, Anthony Miller and Mike Armen helped out substantially while also developing the next generation version of the experiment. They contributed to the construction of both a new photodetector and a low-noise battery charger to supply the magnetic field drivers. Attempts at observing atomic projection noise using these new techniques and equipment are described in chapter 14.

1.2.5 Qwiki

In 2005, with help from Kartik Srinivasan, Anthony Miller, Kevin McHale and several other lab members, I launched a group website titled Qwiki (`http://qwiki.caltech.edu`). This website is a "quantum wiki" where anyone can collaboratively contribute technical information. Thus far the site has been useful for both our group and our small quantum measurement and control community, and it was featured in the Web Watch section of the July 2006 edition of *Physics Today*.

All electronic versions of this thesis (original and corrected) and relevant supplemental information can be found at `http://qwiki.caltech.edu/wiki/John_K._Stockton/Thesis`.

1.3 Chronological History of Concepts

This thesis represents a point of intersection for several historical strands of thought in physics, which cannot be fully unentangled. Much of what inspires us is the evolution of ideas in the general field of quantum measurement and open quantum systems. But, by way of our experiment, we restrict ourselves to the atomic and optical realm and we modestly expand on centuries of work on the interaction between light and atoms. With regard to our measurement, this leads us into the history of Faraday rotation, while for preparing our sample we use techniques from the the field of atom cooling and trapping. Because of the precision with which we can model and measure these systems, we naturally intersect with the disciplines of metrology and the continuing quest to make better magnetometers, clocks, and other practical devices. Here I review some of the key physics milestones in these separate fields to give our work a sense of historical context.

The particular physics of our experiment can be traced back more than a century. In 1845, Michael Faraday discovered that linearly polarized light could be rotated by sending it through a material medium with a magnetic field applied along the propagation direction of the beam. Subsequently, in 1898, Macaluso and Corbino discovered this Faraday effect in a gas of alkali atoms and investigated its resonant behavior near atomic transitions [30]. The field of magneto-optics was enabled many decades later by the discovery of optical pumping [31] and, soon after, the rotation of off-resonant light in an optically pumped vapor cell was investigated [32]. Historically, vapor cells with thermal distributions of atoms have been used as the system of study in Faraday rotation measurements, and state-of-the-art experiments with thermal atoms from magnetometry to fundamental physics continue to be performed in the groups of Budker, Romalis, and Kitching, among others [33, 34, 35]. It is worth noting that there have been several quantum measurement and entanglement experiments, similar to ours, performed with thermal ensembles of atoms [36, 37].

While thermal ensembles are adequate for some work, a large number of experiments rely on the revolutionary advances made in the field of laser cooling and trapping during the 1980s and 1990s [38]. The ability to produce large numbers of confined cold atoms continues to enable a whole discipline, including the production of better atomic clocks.

In the early 1990s ideas of spin-squeezed states were introduced [4, 2, 5, 3] with the aim of improving the standard Ramsey interferometry scheme of atomic clocks. In the

ensuing decade, spin-squeezing related effects have been experimentally reported via several methods, including absorption by cold atoms of squeezed light [39] and interactions between trapped ions [40].

In parallel with these developments of the past few decades, the formalism of quantum measurement has matured and become more experimentally relevant. In particular, the notion of quantum nondemolition (QND) measurement [41] has provided a path for the creation and verification of spin-squeezed states. QND measurements of collective spin systems via Faraday rotation was suggested in several works [42, 43]. A few years later, spin-squeezing was reported using QND measurement of a hot thermal ensemble of atoms [37]. This experiment uses the same technique as ours, but because it uses hot, rather than cold, atoms its methods of producing spin-squeezing may not be as useful for metrology applications. In parallel, the Polzik group has reported the entanglement of two spatially separated cells of atoms with similar techniques (i.e., two-mode spin-squeezed states) [36].

Our work was inspired by this work and a paper by the Wiseman group suggesting the use of feedback to enhance the QND state preparation process [15, 16]. To distinguish ourselves from previous work, and for technical reasons, we decided to use cold atoms. Over the last few years, continuous QND measurement of cold atomic ensembles has been reported in our own work and also in others, including the Jessen and Takahashi groups [44, 24, 25]. Currently several groups are working on implementing QND measurement on the pseudospin associated with the clock transition [45, 46, 47].

1.4 Thesis Organization

This thesis is intended to be read primarily by beginning graduate students with an interest in experimental quantum measurement and control. The first few chapters outline general concepts and definitions related to quantum measurement and control, the middle chapters discuss the theoretical modeling of our particular experiment, and the final several chapters discuss the experimental implementation of these ideas.

Specifically, the chapters of this thesis are as follows. Chapter 2 is meant to be a general quantum measurement primer that reviews the basic literature and introduces general concepts from open quantum systems theory that motivate much of the work we do.

In chapter 3, I introduce the optical and atomic states and terminology that will be used

throughout this thesis, with an emphasis on optical polarization operators and collective spin-states. Chapter 4 is devoted to the question of entanglement in collective spin-states, but is largely independent of the rest of the thesis, because the entanglement physics is mostly referred to indirectly.

Chapter 5, chapter 6, and chapter 7 are devoted to theoretically modeling the atom-light interaction for a multilevel alkali atom. I begin to address the interaction Hamiltonian in chapter 5, where the polarizability form of the Hamiltonian is derived from the dipole Hamiltonian by adiabatic elimination of the excited atomic states under the assumption that the probe is sufficiently far-off resonance. Chapter 6 is devoted to recasting the polarizability Hamiltonian into a more physically intuitive form in terms of optical Stokes operators and atomic spin operators. This Hamiltonian is then used to derive an unconditional master equation for the collective spin-state of the probed atomic cloud in chapter 7.

The stochastic master equation (SME) describing the conditional evolution of the collective spin-state due to a polarimetry measurement of the probe light is introduced in chapter 8. A full derivation is not given, but justification for the physical applicability of the SME is provided. Chapter 9 and chapter 10 are detailed replicas of theoretical papers detailing the theory of the SME and magnetometry respectively. This idealized SME is analyzed in full numerical detail in chapter 9, where it is shown that feedback can lead to deterministic entangled state preparation. The focus turns to metrology in chapter 10, where we use the short-time Gaussian limit of this SME and discuss the application of *quantum parameter estimation* techniques to magnetometry.

The experimental apparatus is introduced in full detail in chapter 11. Before turning to the experimental data, we discuss the relevant atomic physics of trapping and cooling in chapter 12. Semiclassical experimental results that support the theoretical analysis of the interaction Hamiltonian are presented in chapter 13. The search for quantum fluctuations indicative of spin-squeezed state preparation are then detailed in chapter 14. In particular, a modulation scheme for avoiding low frequency noise sources is presented. At last, I suggest several potential theoretical and experimental directions to pursue in chapter 15.

The appendices consist of several research projects that either supplement the body of this thesis or did not fit coherently in the primary story-line. Included are an experimental primer for managing magnetic fields in the lab (appendix A), tutorial examples of continuous measurement and control (appendix B and appendix C), an adaptation of magnetic

resonance physics to our experiment (appendix D), an analysis of spin-squeezing with decay (appendix E), and a discussion of bound states inn magnetic traps (appendix F).

Chapter 2

General Quantum Measurement

2.1   Introduction

The field of quantum optics has historically been the domain where experiments with quantum limited measurements were most accessible, with laser-cooled atoms or ions trapped and isolated from the environment. As a result it was in this field that many of the concepts describing quantum measurement and control were developed. Now that many other types of systems (superconducting qubits, quantum dots, etc.) are coming into play, an increasing proportion of the larger physics community is becoming interested in open quantum system methods.

Frequently, beginning students are not necessarily given the appropriate background to readily acquire the most useful aspects of quantum measurement and control, and need to be referred to relevant background material. In this chapter, I first provide a brief literature review that places this thesis in the context of the various fields from which it draws. After this brief review, I go through a common and simple measurement formalism in order to logically introduce physically intuitive ideas and terminology that might later be clouded by mathematical or experimental details particular to this thesis.

2.2   Review of Literature

The best way to learn about a person quickly is to look at their bookshelf. This section is meant to be a guided glimpse of our lab bookshelves, which will hopefully give the student reader a number of places to get a feel for our group's work. Of course, this is just a small personalized sample of appropriate reading, which is meant to be suggestive but not

complete, and surely many great references are left out. For more on the rest of the lab's reading recommendations, see `http://minty.caltech.edu/reading.php`.

One of our lab's basic tenets is that to understand quantum measurement, estimation, and control theory, it helps a great amount to first understand their classical counterparts. Unfortunately, most undergraduates are only given the most cursory version of the probability and estimation theory, prior to being introduced to the noncommutative weirdness of quantum measurement. An entertaining reference on probability theory from a Bayesian perspective is given in a book by Ed Jaynes [48]. For an introduction to classical control and estimation theory from an engineering perspective I refer the reader to a book by Jacobs [49]. A more sophisticated treatment of control theory can be found in works by Doyle and co-workers [50, 51], which investigate notions of optimal and robust control. Stochastic processes and differential equations are ubiquitous in the context of noisy measurement and are treated mathematically in [52, 53, 54].

While these classical references provide an extreme amount of insight into general estimation and control theory, they mostly neglect the constraints opposed by quantum mechanical estimation including noncommutativity of observables and fundamental uncertainty. These constraints are simply put by Caves, Fuchs, and Schack [55]:

> *In the classical world, maximal information about a physical system is complete in the sense of providing definite answers for all possible questions that can be asked of the system. In the quantum world, maximal information is not complete and cannot be completed.*

Our pragmatic Bayesian perspective of quantum measurement is very much the same as this reference: any quantum measurement scenario is a game of successive measurements, where the goal is to best predict a measurement result given a history of previous measurement results. In this sense, the quantum state that is updated via measurement and conditioning is a state of knowledge that tells us how to place our future bets. A nice "translation" of concepts from classical to quantum measurement and control is provided in [56].

A long-time reference used by our group that describes the conditioning of a quantum state via successive measurements is the thesis of our collaborator Howard Wiseman [57]. Howard is currently updating this work in an anticipated book [58]. For a shorter review article, see [59, 60] for a simple tutorial description of continuous measurement, or [23] for

a more rigorous mathematical approach.

There are several other quite useful open quantum systems books in print, including [61] (although the second edition is not in print), [41] an older but physically intuitive work with an analysis of quantum oscillator measurement limits, and [62], which is more modern and in the spirit of this thesis.

Of course, when studying all of these references it helps to have as many perspectives from related fields as possible. For discussions of measurement with respect to quantum information/computing applications the reader is referred to [63, 64]. There are also many great atomic physics books available relevant to our experiments [65, 1, 38, 66]. Finally, the reader should have a solid understanding of cavity-QED [67] and traditional quantum optics [68, 69, 61]. These references are useful for understanding the open systems trajectory formalism, but many do not discuss conditional measurement and control as much as more recent works.

2.3   Measurement Operator Formalism

Here we describe a simple measurement formalism adapted heavily from [57] and [63] in order to introduce basic concepts of quantum measurement. The measurement operator formalism discussed here is more suited to discrete time measurement events, rather than the experimental measurement scheme of this thesis, but the general concepts essentially transfer. An example using this formalism is presented in appendix B, where we analyze an atom-cavity system with direct photon detection at the output of both sides of the cavity.

Typically, an experimentalist does not *directly* measure the system of interest. The system, through a chain of interactions with larger and larger systems, imprints information about itself on an apparatus of a scale similar to our own. In some cases, this mapping may be trivial. In others, the connection may not be so obvious. To deduce this mapping, we must model the measurement apparatus to a sufficient level of accuracy. The more of the apparatus we describe quantum mechanically the better our predictions will be. Clearly, the choice of this "Heisenberg cut along the Von Neumann chain" must be pragmatically chosen. In practice, we essentially always make this cut at the optical photodiode interface where the probe beam is converted into a classical current.

Here we formally describe the measurement apparatus as the environment (or bath) and

consider our measurements as being projections of the environment state. These do not, in general, correspond to projections of the system state. Thus describing the environment quantum mechanically will fundamentally alter our predictions from a description where we chose measurement to correspond with direct projections on the system.

The basic measurement scheme is displayed in figure 2.1. The world is divided into several parts, including the system of interest (S) and the environment (E). In the instance of our experiment, we imagine the system to be the spin-state of a collection of atoms, and the environment to be the optical state of a laser beam by which we indirectly measure the spin-state. The "classical world" consists of the detector (with efficiency η) and the observer who estimates the spin-state and actuates the system via a controller. There also exist unobserved environment modes through which the observer can lose information (e.g., spontaneous emission). We have also considered the possibility of an unknown Hamiltonian (with parameter λ) acting on the system, under which case the observer performs quantum parameter estimation to update a classical distribution $P_\lambda(t)$. Due to physical interactions, we graphically indicate in figure 2.1 the presence of correlations between the system and the outgoing environment modes. There may also be correlations within the system due to either direct interaction, or via the observer conditioning on the environment detection (e.g., spin-squeezing).

The total state (system plus environment) is represented as $\hat{\rho}_T$, while the subcomponents are represented by tracing out the other so that the environment state is $\hat{\rho}_E = \text{Tr}_S(\hat{\rho}_T)$ and the system state is $\hat{\rho} = \hat{\rho}_S = \text{Tr}_E(\hat{\rho}_T)$. In a single timestep, the joint state evolves according to the Hamiltonian $\hat{H}_T = \hat{H}_S + \hat{H}_E + \hat{H}_{SE}$ where $\hat{H}_S$ evolves the state, $\hat{H}_E$ evolves the environment, and $\hat{H}_{SE}$ lets the system interact with the environment, allowing a measurement of the environment to give information about the system. In the time dt, the joint state then evolves via the propagator $\hat{U}(dt) = \exp[-i\hat{H}_T dt]$.

2.3.1 Indirect Measurement

Now we describe the process of inferring the state of the system after allowing the system to interact with the environment and measuring the environment. The logic of the reduction is as follows (following section 2.2 of [57]). We describe both the system and environment as pure states for simplicity. The environment is prepared in the state $|\psi_E\rangle$. The total state, which we assume to be initially separable, evolves into the generally entangled form
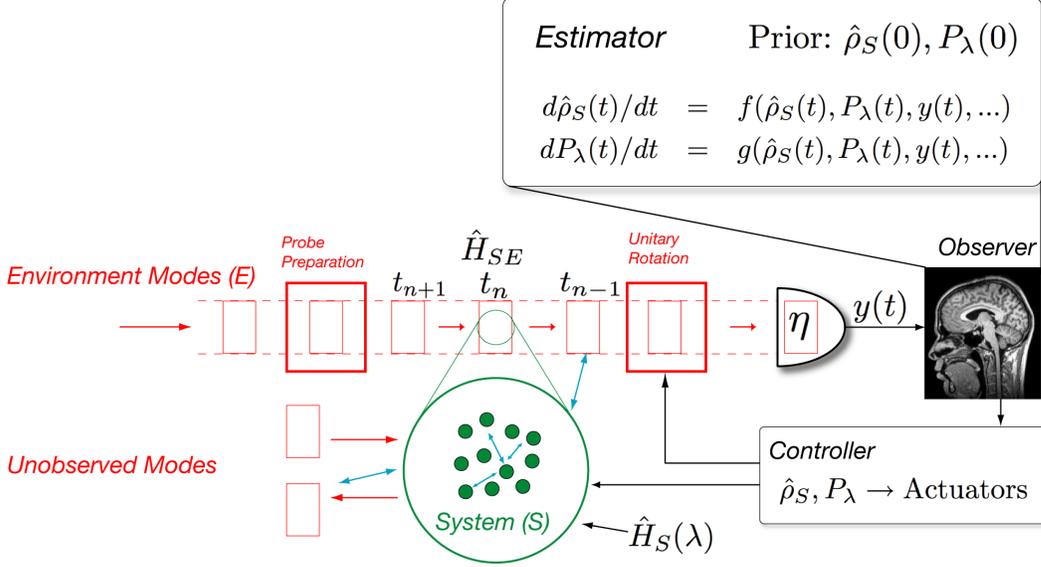
Figure 2.1: A general quantum measurement and feedback control schematic as described in the text.

$|\psi_T(t + dt)\rangle = \hat{U}(dt)|\psi_E\rangle|\psi_S(t)\rangle$. Through measurement, the environment is randomly projected into a particular state α via the projector $\hat{P}_\alpha = 1_S \otimes |\psi_{E,\alpha}\rangle\langle\psi_{E,\alpha}|$. So the resulting unnormalized total state is

$$|\psi_{T,\alpha}(t + dt)\rangle = \hat{P}_\alpha|\psi_T(t + dt)\rangle = |\psi_{E,\alpha}\rangle|\psi_{S,\alpha}(t + dt)\rangle \tag{2.1}$$

where

$$|\psi_{S,\alpha}(t + dt)\rangle = \hat{\Omega}_\alpha(dt)|\psi_S(t)\rangle \tag{2.2}$$

$$\hat{\Omega}_\alpha(dt) = \langle\psi_{E,\alpha}|\hat{U}(dt)|\psi_E\rangle. \tag{2.3}$$

Notice that the measurement operator $\hat{\Omega}_\alpha(dt)$, which acts on only the system subspace, is not necessarily a projector. Also note that this reduction is only useful if the environment has certain dynamical properties. If we are to use the same measurement operators for every timestep, the environment must quickly return to its prepared state (equilibrium approximation) and never return the measured information to the system (Markovian approximation). Of course, we may explicitly return information to the system via Hamiltonian feedback, but here we are considering the open-loop dynamics alone.

2.3.2 Dynamics of Continuous Measurement

Now we assume the existence of measurement operators $\hat{\Omega}_\alpha(dt)$ based on an adequate model of the system-environment interaction and work out the dynamics. The continuous measurement operators are usually specified as follows, with the $\{\hat{L}_i\}$ being the so-called *jump operators*

$$\hat{\Omega}_i(dt) = \sqrt{dt}\,\hat{L}_i \tag{2.4}$$

$$\hat{\Omega}_0(dt) = 1 - i\hat{H}_{\text{eff}}dt = 1 - i(\hat{H}_S - i\hat{K})dt \tag{2.5}$$

with $i = 1, 2, 3, ...$ and $\hat{K} = \frac{1}{2}\sum_{i>0}\hat{L}_i^\dagger\hat{L}_i$ from the Kraus normalization condition.

2.3.2.1 Conditional Evolution

Let us define an effect $\hat{\Upsilon}_i$ to be associated with a measurement operator through $\hat{\Upsilon}_i = \hat{\Omega}_i^\dagger\hat{\Omega}_i$. In the simulation of conditional evolution, we imagine flipping a coin weighted according to $P_i = \text{Tr}(\hat{\Upsilon}_i\hat{\rho})$ to get a random result i on every timestep dt. In the experiment, we let nature flip the coin for us, and we measure for time period dt to get a result i. Note that one possible measurement result is the *null measurement* of $i = 0$. In any case, given result i, the evolution rule is to apply the appropriate jump operator by

$$\hat{\rho}(t + dt) = \frac{\hat{\Omega}_i(dt)\hat{\rho}(t)\hat{\Omega}_i^\dagger(dt)}{P_i}. \tag{2.6}$$

The random state evolution described by this formalism is referred to as a *quantum trajectory*. Given a model of the environment and the interaction this conditional rule for updating the quantum state is called many things, but here we refer to it as the stochastic master equation (SME), the stochasticity coming from the inherent randomness of the detection in the environment. Note that if the detection of the environment is perfectly efficient, then the state changes but remains pure because we have not lost any information.

It is also important to point out that this equation is a experimental filter meant to be used in the lab: the inputs are the initial state (of both the system and environment), the model of the interaction, *and* the measurement record, and the output is the updated system state. The filtering procedure is shown schematically in figure 2.1. Here the dependence on the measurement record is somewhat implicit, but later it will be made more clear. This

point is worth making because some physicists use trajectories as merely a computational tool to evolve the unconditional master equation, discussed below, and do not associate the trajectories with any particular measurement. (The advantage is that with the pure state trajectories one works with the state vector, which is smaller than the state matrix used in the unconditional master equation evolution.)

Also note that the update rule above is very discrete in nature, as we imagine applying the rule every dt. This is fine if the number of detection events expected in the time dt is very small (see appendix B), however often we use another high power beam as a local oscillator in our measurement scheme. In this case, the number of detection events in any dt is much larger and we use a different description as discussed below, incorporating Wiener white noise increments dW to represent the randomness of the measurement [52]. This limit is mathematically distinct, but the underlying physical intuition remains the same as above.

2.3.2.2   Unconditional Evolution

In the case where the measurement record is completely ignored, but the system is still interacting with the environment, our best estimate of the system state must be given by the average of all possibilities

$$\hat{\rho}(t+dt) = \sum_i \hat{\Omega}_i(dt)\hat{\rho}(t)\hat{\Omega}_i^\dagger(dt) \tag{2.7}$$

or

$$\frac{d\hat{\rho}}{dt} = \hat{\mathcal{L}}\hat{\rho} \tag{2.8}$$

where $\hat{\mathcal{L}}$ is the Liouvillian (i.e., Lindbladian). This is the unconditional master equation. Note that, by definition, the average of the conditional trajectories should reproduce this unconditional behavior.

2.3.3   Transforming the Measurement

Formally, one can unitarily rearrange the measurement operators without changing the unconditional evolution, thus creating what is called a different *unravelling* of the master

equation. Thus the transformation

$$\hat{\Omega}'_i = \sum_j \hat{U}_{i,j} \hat{\Omega}_j \tag{2.9}$$

results in no change of the unconditional evolution

$$\hat{\rho}(t + dt) = \sum_j \hat{\Omega}_j(dt)\hat{\rho}(t)\hat{\Omega}_j^\dagger(dt) = \sum_i \hat{\Omega}'_i(dt)\hat{\rho}(t)\hat{\Omega}_i'^\dagger(dt) \tag{2.10}$$

where $\sum_r \hat{U}_{r,s}\hat{U}_{r,q}^* = \delta_{s,q}$. Whether or not a particular unravelling is physically realizable is another question. For the detection of a laser beam, different unravellings correspond to different means of optical detection. Direct detection results in conditional evolution with jump-like behavior. But by adding a local oscillator to the output field one can perform a homodyne measurement, which in the large oscillator limit results in diffusive motion of the conditional state. The switch from direct detection to homodyne (or heterodyne) corresponds to a unitary rearrangement of the direct detection measurement operators. This is depicted by the unitary rotation element in figure 2.1. In the case of polarimetry as discussed in section 8.1 different unravellings can be achieved through different settings of the wave-plates prior to the polarimeter. In appendix B, a more explicit example is given.

The concept of unravelling is a very intuitive and important: in essence, what we know about a system depends on how we look at it. Depending on our measurement and control objective, the choice of unravelling will be critical and should be optimized [70].

2.4 The Continuous Limit

As discussed in chapter 8 and chapter 9, the stochastic master equation describing the conditional evolution of our collective spin-state (system) due to the detection of a far-off resonant probe beam (environment) is represented in the formalism of continuous stochastic measurement. The formal derivation of the SME is beyond the scope of this chapter, but here we introduce it and show its qualitative similarities to the above analysis.

The SME describing the conditional evolution is [15, 16]

$$d\hat{\rho}(t) = -i[\hat{H}_S, \hat{\rho}(t)]dt + \mathcal{D}[\sqrt{M}\hat{F}_z]\hat{\rho}(t)dt$$
$$+\sqrt{\eta}\mathcal{H}[\sqrt{M}\hat{F}_z]\hat{\rho}(t)\left(2\sqrt{M\eta}[y(t)dt - \langle\hat{F}_z\rangle dt]\right) \quad (2.11)$$

where $\hat{H}_S$ is the system Hamiltonian (possibly magnetic field rotations), M is the probe parameter dependent measurement rate, $\mathcal{D}$ and $\mathcal{H}$ are superoperators, η is the quantum efficiency of the measurement, and the photocurrent is represented as

$$y(t)dt = \langle\hat{F}_z\rangle(t)dt + dW(t)/2\sqrt{M\eta}. \quad (2.12)$$

The stochastic quantity $dW(t) \equiv 2\sqrt{M\eta}(y(t)dt - \langle\hat{F}_z\rangle(t)dt)$ is a Wiener increment and $dW(t)/dt$ is a Gaussian white noise associated with the shotnoise of the homodyne local oscillator. See [53, 52] for an introduction to stochastic differential equations (SDE).

There are several features to notice about this equation. First, if $M = 0$ we recover the usual Heisenberg equation of motion for the state. If $M \neq 0$, but the quantum efficiency of the detector $\eta = 0$, then we get the unconditional evolution of the state due to the measurement. If $M \neq 0$ and $\eta \neq 0$, then the update rule is an explicit function of the photocurrent $y(t)$. This conditional part of the equation is reminiscent of classical estimators that take the difference between what the photocurrent is and what the photocurrent is expected to be, to get an *innovation* by which the system is updated.

This equation is analyzed extensively in chapter 9, where it is seen that this particular form of the measurement enacts a projection of the state onto the eigenstates of $\hat{F}_z$ at long times. This form of the SME is quite general and represents the canonical example of a continuous measurement. It is discussed from a more mathematical perspective in [71, 72, 23].

2.5   Modeling the Interaction

Clearly the derivation of the filtering equations is wholly dependent on our model of the interaction between the system and the environment. This includes the Hamiltonian $\hat{H}_{SE}$, which, for our experiment, we spend a great deal of space analyzing and simplifying in Chapters 4–7 of this thesis. In practice it is always a challenging task to produce a model

that captures all of the salient features of a physical experiment, while neglecting the superfluous details. For example, in our experiment we often neglect the spatial motion of the atoms and the fact that they collectively emit into a full three-dimensional mode and it is our job as physicists to justify these approximations.

Even under the ideal circumstances often considered in this thesis, the exponentially large Hilbert space used to describe a large number of spins is often prohibitive and model reduction needs to be used to reduce the size of the state description. One approach is to use the symmetry, and assume the state is invariant to spin exchange, as discussed in chapter 4. Another approach is to assume the collective spin-state is in a Gaussian state and work with only a few moments of the overall distribution. This is discussed in chapter 3 and implied throughout much of this thesis. The Gaussian description often allows us to make direct connections between results of classical and quantum control as in the recent work [70].

2.6   QND Measurement

Often in this thesis the term quantum nondemolition, or QND, measurement is used casually. Here we clarify its definition to avoid confusion. The notion of a QND measurement was introduced in the context of force detection by Braginsky and others [41]. In a simple and discrete definition, a QND measurement is one that will project the system into an eigenstate of the observable such that subsequent measurements of the same observable will return the same result. One example of a measurement that is *not* QND is photodetection at the output of a cavity, because energy is lost from the cavity. In principle, a Stern-Gerlach apparatus could be QND if the spin carrier were not destroyed, but practically it often is, removing the possibility of subsequent measurement. Clearly, one must distinguish between notions of demolition and destruction.

After a completely projective QND measurement, the observed variable is known completely while complementary observables become less uncertain, hence the inherent uncertainty of the state is physically shunted to other regions of Hilbert space by the measurement. Also, if the system Hamiltonian does not commute with the measured observable, the measurement will not be QND. For example, if a spin eigenstate of $\hat{F}_z$ is prepared then a magnetic field (along x for example) rotates this eigenstate into another basis, then the

next measurement result may not be the same as the first.

The notion of QND measurement is already continuous in a sense because it concerns repeated measurements, but one can also discuss a single continuous QND measurement where the projection onto an eigenstate is gradual as described by equation (2.11), which is analyzed in chapter 9. A more mathematical, quantum filtering perspective of the non-demolition property was introduced by Belavkin, whose early work on quantum filtering continues to be reproduced in more physical contexts decades later [73, 74, 75]. In that context, nondemolition means that a system observable commutes with all prior observations. This definition turns out to imply the fact that, given all prior observations, estimating the system observable is a classical statistical inference problem [23]. This is the basis for the classical model of the quantum measurement used in chapter 10.

QND measurements have been performed on optical quantum states [76, 77], and the idea of continuous QND measurement for spin-squeezing was introduced in [42, 43]. In this thesis, we mostly discuss a continuous QND measurement of the collective $\hat{F}_z$ variable of a spin ensemble, which is measured via an interaction Hamiltonian proportional to that variable. In particular, we investigate the long-time limit of the measurement in chapter 9, which produces the results of a discrete projective QND measurement. There are several places that the QND feature of the measurement is practically compromised, however. First, the tensor terms introduced in chapter 6 complicate the interaction Hamiltonian such that it is not simply represented in terms of collective variables. Second, when any magnetic field is applied in the xy-plane, the measurement is no longer QND because $\hat{F}_z$ does not then commute with the magnetic field dependent Larmor precession Hamiltonian (although a z-field is acceptable). In chapter 14 and appendix D, we discuss applying an x-field that destroys the QND nature of a continuous measurement. Because of this, we can then modulate the measurement strength in time to effectively restore the QND nature of the measurement.

2.7   Quantum Parameter Estimation

Quantum metrology typically involves the estimation of a classical parameter by the detection of a quantum system that evolves due to that parameter. Traditionally, there has been significant interest in detecting classical forces on quantum harmonic oscillators [78, 41].

These ideas were discussed more recently in the context of more modern techniques in [19]. Following on this paper, we adapted the techniques to the measurement of a magnetic field with our system in a set of theoretical papers [18, 20] that are summarized in chapter 10. This work was compared to traditional magnetometers by the Budker group in [79] and extended by the Mølmer group in several papers, including [80]. For a general view of how entanglement enhances metrology see [81].

The basic idea of quantum parameter estimation is that system Hamiltonian is dependent on a particular classical parameter, say λ such that $\hat{H}_S(\lambda)$. In addition to evolving the system state $\hat{\rho}(t)$ however we must simultaneously evolve a classical parameter distribution $P_\lambda(t)$ using Bayesian methods as discussed in the references above. The uncertainties characterizing either distribution will be coupled via the evolution. In chapter 10, we make this coupling explicit and analyze the limits that atomic spin uncertainty place on the magnetic field uncertainty, and vice versa.

As quantum metrology devices improve, the use of these rigorous parameter estimation techniques will become as common in the quantum domain as they are in the classical metrology.

2.8   Quantum Feedback Control

In relation to this thesis, there is an apt quote, attributed to Norbert Wiener, that states

> *In the absence of uncertainty, there is no need for feedback.*

In a certain sense, this implies that feedback control is more fundamental in the inherently uncertain world of quantum mechanics than in the classical realm, where idealized levels of certainty and precision can make open-loop control all-powerful and feedback control pointless.

The application of feedback control concepts to measured quantum systems is discussed in many recent references, including [56, 23]. In the context of this thesis, the term *quantum feedback control* is meant to imply something very specific: during the course of a continuous measurement as described above, quantum feedback involves the real time mapping of the measurement record $y(t)$ to the Hamiltonian $\hat{H}_S$ to achieve a predefined control goal. This is to distinguish the use of the term from other meanings often found in the literature.

The use of feedback is helpful in the quantum domain for many of the same reasons it is used in the classical world. For example, there are many proposed applications of quantum control for cooling external, motional degrees of systems, from atoms in a cavity to nanomechanical oscillators. During the course of this thesis, feedback is used in two separate but related contexts. In chapter 9, which is based on ideas from [16, 21, 22], we discuss how feedback can be used to direct the state diffusion from the SME above, allowing the *deterministic* preparation of entangled states. Essentially, the feedback allows us to remove the randomness of the measurement but retain the useful entangling dynamics the conditioning induces. In some cases, the same states could be prepared with a conditional postmeasurement feedback step. However, in many cases the continuous measurement with concurrent feedback is necessary for the preparation of certain states. Second, as discussed in chapter 10, feedback can make the quantum parameter estimation procedure *robust* to classical model uncertainties in other parameters describing the model. For example, without feedback the magnetometry procedure we describe needs to know the total atom number used in the experiment, but with feedback we need to know much less information about the atom number to achieve the same performance.

The magnetometry example is representative of the larger pursuit to analyze robustness and risk-sensitivity in the context of quantum estimation and control. For more on these general ideas see [82, 83].

Chapter 3

Optical and Atomic States

In this chapter, we introduce the notation used throughout this thesis to represent the optical and atomic states and operators common to our experiment. We begin by reviewing the basic scheme of the experiment, with atomic spins measured via Faraday rotation of a probe beam and subsequent polarimetry. Next, we describe the optical states used, with an emphasis on the quantum Stokes parameter description of optical polarization. After this, we introduce the atomic spin-states, starting with a description of the individual cesium spin and moving on to the definition of spin-squeezed states and other collective states. Here, we state and prove several facts related to spin-squeezing, including the proof that spin-squeezed states with spin-1/2 particles are necessarily entangled. After describing the relevant states and operators, we begin to analyze the entanglement properties of collective atomic states in the next chapter.

3.1   Experimental Scenario

The experimental apparatus described throughout this work is represented in figure 1.2. A beam of off-resonant probe light is polarized and sent through the atomic ensemble. As described in later chapters, the interaction between the atoms and light is somewhat complicated but for now we assume that the atoms act as a generalized waveplate that rotates the polarization on the Stokes sphere, which includes the possibility of creating circularly polarized light. The way in which this rotation happens depends on the internal state of the atoms, and the noise imparted to the light in turn depends on the uncertainty of the collective atomic state describing the ensemble. Motivated by this scenario, we now describe the optical polarization operators and states, and subsequently the relevant

collective atomic states.

3.2 Optical States

We begin by describing the electric field representation of the probe beam with two transverse polarization modes, and then we derive several polarization operators and some of their properties. The coherent state of the probe beam is specified and related to the polarimeter detection described in subsequent chapters.

3.2.1 Electric Field Representations

The classical electric field describing a single spatial mode of light with two orthogonal polarization components $\sigma = 1, 2$ is denoted

$$\mathbf{E} = \mathcal{E} \sum_{\sigma=1,2} \epsilon_\sigma \mathbf{e}_\sigma \exp[i\kappa z - i\omega t] + \text{c.c.} \tag{3.1}$$

We always consider the beams to be propagating along the z-direction. The power in this beam is then given by

$$P/A = \frac{\epsilon_0 c \overline{\mathbf{E}^2}}{2} = \epsilon_0 c \mathcal{E}^2. \tag{3.2}$$

The quantized version of this field [68] is defined as

$$\hat{\mathbf{E}} = \sqrt{\hbar g} \sum_{\sigma=1,2} \hat{a}_\sigma \mathbf{e}_\sigma \exp[i\kappa z - i\omega t] + \text{h.c.} \tag{3.3}$$

where $g = \omega/(2\epsilon_0 V)$ and V is the volume of the mode. (Note that, in this thesis, I use a notation where vectors are given boldface and operators are given hats.) The $\hat{a}_\sigma$ are the annihilation operators satisfying the usual commutation relation $[\hat{a}_i, \hat{a}_j^\dagger] = \delta_{ij}$. Unless otherwise noted, we will ignore the spatial dependence and only consider $\hat{\mathbf{E}}$ at the spatial location $z = 0$. However, this simplification would need to be removed in the consideration of spatially extended scattering media imparting significant polarization rotation (on the order of radians).

In quantum optics, it is often convenient to move into the frame rotating at the optical frequency to remove the time dependence of a particular operator. Explicitly, we transform

the field operator as

$$\hat{\mathbf{E}}_R \;=\; \exp[i\omega t(\hat{a}_1^\dagger \hat{a}_1 + \hat{a}_2^\dagger \hat{a}_2)]\hat{\mathbf{E}}\exp[-i\omega t(\hat{a}_1^\dagger \hat{a}_1 + \hat{a}_2^\dagger \hat{a}_2)] \tag{3.4}$$

$$=\; \hat{\mathbf{E}}^{(-)} + \hat{\mathbf{E}}^{(+)}. \tag{3.5}$$

where

$$\hat{\mathbf{E}}^{(-)} \;=\; \sqrt{\hbar g}\left[\hat{a}_1^\dagger \mathbf{e}_1^* + \hat{a}_2^\dagger \mathbf{e}_2^*\right] \tag{3.6}$$

$$\hat{\mathbf{E}}^{(+)} \;=\; \sqrt{\hbar g}\left[\hat{a}_1 \mathbf{e}_1 + \hat{a}_2 \mathbf{e}_2\right]. \tag{3.7}$$

From now on we will assume the rotating frame is being used and remove the R subscript from $\hat{\mathbf{E}}$.

3.2.2 Polarization Basis

Now we discuss ways of representing the polarization of the optical beam propagating along the z-direction. We start in the basis given by $\mathbf{e}_x$ and $\mathbf{e}_y$. In general one can represent the polarization of the light in an elliptical superposition of these elements [68], but we choose to limit ourselves to three distinct basis pairs. We denote by $\mathbf{e}_{x'}$ and $\mathbf{e}_{y'}$ the basis pair that is also real, but at 45 degrees relative to $\mathbf{e}_x$ and $\mathbf{e}_y$. The circular basis is then defined as the complex superposition of the real basis and denoted by $\mathbf{e}_+$ and $\mathbf{e}_-$. More specifically, these vectors are related to each other through the following transformations:

$$\mathbf{e}_x \;=\; \tfrac{1}{\sqrt{2}}(\mathbf{e}_{x'} - \mathbf{e}_{y'}) \;=\; \frac{1}{\sqrt{2}}(-\mathbf{e}_+ + \mathbf{e}_-) \tag{3.8}$$

$$\mathbf{e}_y \;=\; \tfrac{1}{\sqrt{2}}(\mathbf{e}_{x'} + \mathbf{e}_{y'}) \;=\; \frac{i}{\sqrt{2}}(\mathbf{e}_+ + \mathbf{e}_-) \tag{3.9}$$

$$\mathbf{e}_{x'} \;=\; \tfrac{1}{\sqrt{2}}(\mathbf{e}_x + \mathbf{e}_y) \;=\; \frac{1}{2}((-1+i)\mathbf{e}_+ + (1+i)\mathbf{e}_-) \tag{3.10}$$

$$\mathbf{e}_{y'} \;=\; \tfrac{1}{\sqrt{2}}(-\mathbf{e}_x + \mathbf{e}_y) \;=\; \frac{1}{2}((1+i)\mathbf{e}_+ + (-1+i)\mathbf{e}_-) \tag{3.11}$$

$$\mathbf{e}_+ \;=\; \tfrac{1}{\sqrt{2}}(-\mathbf{e}_x - i\mathbf{e}_y) \;=\; \frac{1}{2}((-1-i))\mathbf{e}_{x'} + (1-i)\mathbf{e}_{y'}) \tag{3.12}$$

$$\mathbf{e}_- \;=\; \tfrac{1}{\sqrt{2}}(\mathbf{e}_x - i\mathbf{e}_y) \;=\; \frac{1}{2}((1-i)\mathbf{e}_{x'} + (-1-i)\mathbf{e}_{y'}). \tag{3.13}$$

3.2.3 Spherical Basis and Tensors

The basis $\mathbf{e}_{x'}$ and $\mathbf{e}_{y'}$ is a trivial real rotation of the original basis $\mathbf{e}_x$ and $\mathbf{e}_y$, however the circular basis is somewhat less intuitive, so we review some of its properties here. When the circular basis is combined with the propagation vector along z, we get the spherical basis

$$\mathbf{e}_+ = -(\mathbf{e}_x + i\mathbf{e}_y)/\sqrt{2} \tag{3.14}$$

$$\mathbf{e}_- = (\mathbf{e}_x - i\mathbf{e}_y)/\sqrt{2} \tag{3.15}$$

$$\mathbf{e}_z = \mathbf{e}_z, \tag{3.16}$$

which is often used in atomic physics due to its symmetry properties [84]. When making the associations $+\,(q = +1)$, $-\,(q = -1)$, and $z\,(q = 0)$ the elements of the spherical basis have the properties

$$\mathbf{e}_q^* = \mathbf{e}_{-q}(-1)^q \tag{3.17}$$

$$\mathbf{e}_q \cdot \mathbf{e}_{q'}^* = \delta_{q,q'}. \tag{3.18}$$

For an arbitrary vector $\mathbf{A}$ we have $A_q = \mathbf{e}_q \cdot \mathbf{A}$ so that $\mathbf{A} = \sum_q A_q \mathbf{e}_q^* = \sum_q (-1)^q A_q \mathbf{e}_{-q}$. Due to the element $\mathbf{e}_z$ this basis describes any field for any propagation direction and not just a beam propagating along z.

3.2.4 Stokes Representation

Given these basis pairs, we define the Stokes operators (also known as Schwinger boson operators) as

$$\hat{S}_x = \frac{1}{2}(\hat{a}_y^\dagger \hat{a}_y - \hat{a}_x^\dagger \hat{a}_x) \tag{3.19}$$

$$\hat{S}_y = \frac{1}{2}(\hat{a}_{y'}^\dagger \hat{a}_{y'} - \hat{a}_{x'}^\dagger \hat{a}_{x'}) \tag{3.20}$$

$$\hat{S}_z = \frac{1}{2}(\hat{a}_+^\dagger \hat{a}_+ - \hat{a}_-^\dagger \hat{a}_-) \tag{3.21}$$

$$\hat{S}_0 = \frac{1}{2}(\hat{a}_y^\dagger \hat{a}_y + \hat{a}_x^\dagger \hat{a}_x). \tag{3.22}$$

Clearly, $\hat{S}_x$ measures in one linearly polarized basis, $\hat{S}_y$ measures in the other linearly polarized basis, and $\hat{S}_z$ measures in the circular basis. Each of these components can be

measured in a polarimeter using at most two waveplates and a polarizing beamsplitter as described below. Note that if two polarizations are perpendicular in real space (e.g., along $\mathbf{e}_x$ and $\mathbf{e}_y$), they will be represented by vectors that are pointing opposite directions on the Stokes sphere. Using the above transformations of basis, one can show that these operators can be defined in terms of the other basis pairs as

$$
\hat{S}_x = \tfrac{1}{2}(\hat{a}_\mathrm{y}^\dagger \hat{a}_\mathrm{y} - \hat{a}_\mathrm{x}^\dagger \hat{a}_\mathrm{x}) = \frac{1}{2}(\hat{a}_\mathrm{y'}^\dagger \hat{a}_\mathrm{x'} + \hat{a}_\mathrm{x'}^\dagger \hat{a}_\mathrm{y'}) = \frac{1}{2}(\hat{a}_+^\dagger \hat{a}_- + \hat{a}_-^\dagger \hat{a}_+) \tag{3.23}
$$

$$
\hat{S}_y = \tfrac{1}{2}(-\hat{a}_\mathrm{y}^\dagger \hat{a}_\mathrm{x} - \hat{a}_\mathrm{x}^\dagger \hat{a}_\mathrm{y}) = \frac{1}{2}(\hat{a}_\mathrm{y'}^\dagger \hat{a}_\mathrm{y'} - \hat{a}_\mathrm{x'}^\dagger \hat{a}_\mathrm{x'}) = \frac{i}{2}(\hat{a}_-^\dagger \hat{a}_+ - \hat{a}_+^\dagger \hat{a}_-) \tag{3.24}
$$

$$
\hat{S}_z = \tfrac{i}{2}(\hat{a}_\mathrm{y}^\dagger \hat{a}_\mathrm{x} - \hat{a}_\mathrm{x}^\dagger \hat{a}_\mathrm{y}) = \frac{i}{2}(\hat{a}_\mathrm{y'}^\dagger \hat{a}_\mathrm{x'} - \hat{a}_\mathrm{x'}^\dagger \hat{a}_\mathrm{y'}) = \frac{1}{2}(\hat{a}_+^\dagger \hat{a}_+ - \hat{a}_-^\dagger \hat{a}_-) \tag{3.25}
$$

$$
\hat{S}_0 = \tfrac{1}{2}(\hat{a}_\mathrm{y}^\dagger \hat{a}_\mathrm{y} + \hat{a}_\mathrm{x}^\dagger \hat{a}_x) = \frac{1}{2}(\hat{a}_\mathrm{y'}^\dagger \hat{a}_\mathrm{y'} + \hat{a}_\mathrm{x'}^\dagger \hat{a}_\mathrm{x'}) = \frac{1}{2}(\hat{a}_+^\dagger \hat{a}_+ + \hat{a}_-^\dagger \hat{a}_-). \tag{3.26}
$$

Using $[\hat{a}_i, \hat{a}_j^\dagger] = \delta_{ij}$ one can show that the Stokes operators have the same commutation relations as angular momentum operators

$$
\left[\hat{S}_x, \hat{S}_y \right] = i\hat{S}_z \tag{3.27}
$$

$$
\left[\hat{S}_y, \hat{S}_z \right] = i\hat{S}_x \tag{3.28}
$$

$$
\left[\hat{S}_z, \hat{S}_x \right] = i\hat{S}_y \tag{3.29}
$$

$$
\left[\hat{S}_i, \hat{S}_0 \right] = 0. \tag{3.30}
$$

All of the usual angular momentum relations immediately follow. Other means of describing the polarization state, e.g., the Jones vector, often assume further constraints and can be found in [85].

3.2.5 Arbitrary Vector Operator Rotations

Imagine a polarized beam of light propagating through, or reflecting off, some material. If the beam is not attenuated, then the material will rotate the polarization of the beam about some, possibly arbitrary, direction on the Stokes sphere. Thus we can represent the unitary behavior of the evolution (whether it be due to a waveplate, a mirror, a piece of glass, cold atoms, etc.) simply by a rotation vector. In this section, we are interested in evaluating the operation of rotating a Stokes vector about an arbitrary direction by an arbitrary angle.

Consider the rotation of the vector operator

$$\hat{\mathbf{S}} = \left[\hat{S}_x, \hat{S}_y, \hat{S}_z\right] \tag{3.31}$$

in Cartesian coordinates about an arbitrary direction $\mathbf{n} = [\gamma_x, \gamma_y, \gamma_z]/\gamma$ by the angle $\gamma = \sqrt{\gamma_x^2 + \gamma_y^2 + \gamma_z^2}$. Using the rotation vector $\boldsymbol{\gamma} = \gamma\mathbf{n} = [\gamma_x, \gamma_y, \gamma_z]$, this rotation can be represented in the Heisenberg picture as

$$\hat{S}_i' = \hat{U}\hat{S}_i\hat{U}^\dagger \tag{3.32}$$

where

$$\hat{U} = \exp[-i\hat{\mathbf{S}} \cdot \boldsymbol{\gamma}] = \exp[-i(\gamma_x\hat{S}_x + \gamma_z\hat{S}_z + \gamma_z\hat{S}_z)]. \tag{3.33}$$

The $\hat{S}_i'$ can be derived explicitly using the following equation for the arbitrary rotation of any vector

$$\hat{S}_i' = (\hat{\mathbf{S}} \cdot \mathbf{i}) \cos\gamma + (\mathbf{n} \cdot \mathbf{i})(\mathbf{n} \cdot \hat{\mathbf{S}})(1 - \cos\gamma) + \left((\mathbf{n} \times \mathbf{i}) \cdot \hat{\mathbf{S}}\right) \sin\gamma. \tag{3.34}$$

(See, for example, eq. (3.16) of `http://www.phys.vt.edu/~mizutani/quantum/rotations.pdf`.) Expanding and rearranging terms we get

$$
\begin{aligned}
\hat{S}'_x &= \hat{S}_x \left(\frac{\gamma_x^2}{\gamma^2}(1 - \cos\gamma) + \cos\gamma \right) \\
&\quad + \hat{S}_y \left(\frac{\gamma_x\gamma_y}{\gamma^2}(1 - \cos\gamma) + \frac{\gamma_z}{\gamma}\sin\gamma \right) \\
&\quad + \hat{S}_z \left(\frac{\gamma_x\gamma_z}{\gamma^2}(1 - \cos\gamma) - \frac{\gamma_y}{\gamma}\sin\gamma \right) \tag{3.35}
\end{aligned}
$$

$$
\begin{aligned}
\hat{S}'_y &= \hat{S}_x \left(\frac{\gamma_y\gamma_x}{\gamma^2}(1 - \cos\gamma) - \frac{\gamma_z}{\gamma}\sin\gamma \right) \\
&\quad + \hat{S}_y \left(\frac{\gamma_y^2}{\gamma^2}(1 - \cos\gamma) + \cos\gamma \right) \\
&\quad + \hat{S}_z \left(\frac{\gamma_y\gamma_z}{\gamma^2}(1 - \cos\gamma) + \frac{\gamma_x}{\gamma}\sin\gamma \right) \tag{3.36}
\end{aligned}
$$

$$
\begin{aligned}
\hat{S}'_z &= \hat{S}_x \left(\frac{\gamma_z\gamma_x}{\gamma^2}(1 - \cos\gamma) + \frac{\gamma_y}{\gamma}\sin\gamma \right) \\
&\quad + \hat{S}_y \left(\frac{\gamma_z\gamma_y}{\gamma^2}(1 - \cos\gamma) - \frac{\gamma_x}{\gamma}\sin\gamma \right) \\
&\quad + \hat{S}_z \left(\frac{\gamma_z^2}{\gamma^2}(1 - \cos\gamma) + \cos\gamma \right). \tag{3.37}
\end{aligned}
$$

As an example, a half-waveplate with its primary axis rotated by an angle ϕ from the x-axis will correspond to a rotation vector of

$$
\boldsymbol{\gamma} = \pi[\cos(2\phi), \sin(2\phi), 0]. \tag{3.38}
$$

Similarly a quarter-waveplate with its primary axis rotated ϕ from the x-axis will correspond to a rotation vector of

$$
\boldsymbol{\gamma} = \frac{\pi}{2}[\cos(2\phi), \sin(2\phi), 0]. \tag{3.39}
$$

3.2.6   Fock States and Coherent States

We work with the convention that Fock states are represented as $|n\rangle$ where n are integers and coherent states are represented as $|\alpha\rangle$ where α are complex numbers. The type of state should usually be clear by context so distinguishing subscripts are not used.

Fock states of a particular mode have the properties [68]

$$\hat{a}|n\rangle = \sqrt{n}|n-1\rangle \tag{3.40}$$

$$\hat{a}^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle \tag{3.41}$$

$$\hat{a}^\dagger\hat{a}|n\rangle = n|n\rangle. \tag{3.42}$$

For the same mode, we define the coherent state [68] in terms of the Fock states as

$$|\alpha\rangle = \exp\left(-|\alpha|^2/2\right)\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}|n\rangle \tag{3.43}$$

and it quickly follows from the above that we have

$$\hat{a}|\alpha\rangle = \alpha|\alpha\rangle \tag{3.44}$$

$$\langle\alpha|\hat{a}^\dagger = \langle\alpha|\alpha^* \tag{3.45}$$

$$\langle\alpha|\hat{a}^\dagger\hat{a}|\alpha\rangle = |\alpha|^2. \tag{3.46}$$

The power represented in a beam sliced into modes of length cdt (where c is the speed of light) is

$$P = \langle\hat{a}^\dagger\hat{a}\rangle\hbar\omega/dt = |\alpha|^2\hbar\omega/dt \tag{3.47}$$

For computational reasons we will sometimes take the time slice dt small such that $\alpha \ll 1$ and expand the coherent state to first-order

$$|\alpha\rangle \approx (1-|\alpha|^2/2)(|0\rangle + \alpha|1\rangle). \tag{3.48}$$

For an arbitrarily polarized beam propagating along one direction we will represent the optical state as the tensor product

$$|\Phi\rangle = |\alpha_1\rangle_1 \otimes |\alpha_2\rangle_2 \tag{3.49}$$

where 1 and 2 label orthogonal polarization modes. One can then use the representations of the $\hat{S}_i$ operators and the properties of the coherent state above to calculate $\langle\hat{S}_i\rangle$ in terms

of the α_i. The power of this beam is then simply the sum of the powers in each mode

$$P = \langle \hat{a}_x^\dagger \hat{a}_x + \hat{a}_y^\dagger \hat{a}_y \rangle \hbar\omega/dt \tag{3.50}$$

$$= 2\langle \hat{S}_0 \rangle \hbar\omega/dt \tag{3.51}$$

$$= (|\alpha_x|^2 + |\alpha_y|^2)\hbar\omega/dt. \tag{3.52}$$

Throughout this thesis we focus mostly on using coherent states of probe light to create spin-squeezed states. It turns out that one can also discuss the creation of *polarization squeezed light* by multipassing the probe beam through the unsqueezed atomic ensemble [86], although we do not consider this effect here.

3.2.7 Polarimetry

One can measure any of the above bases or any linear combination thereof (simultaneously with $\hat{S}_0$) with a fixed combination of a quarter-waveplate, a half-waveplate, a polarizing beamsplitter (PBS), and two independent detectors. The sum current (optical power) will be $\hat{S}_0$ and the difference will be $\hat{S}_i$ where i is the measurement basis. The observable $\hat{S}_x$ is measured with just a PBS with axes along the directions x and y. The observable $\hat{S}_y$ is measured with a half-waveplate prior to the PBS with primary axis at 45 degrees relative to the PBS (x-y) axes (see figure 1.2). The observable $\hat{S}_z$ is measured with a quarter-waveplate prior to the PBS with primary axis also at 45 degrees relative to the PBS (x-y) axes. We discuss the theory of polarimeters more in section 8.1 and the experimental implementation in section 11.6.

3.3   Atomic Spin-States

Turning from light to atoms, we now introduce the language used to describe a collective atomic spin-state. For the sake of simplicity we assume that the collective states are symmetric with respect to particle exchange and we also ignore the external degrees of freedom associated with position. Throughout this analysis we consider the use of actual spin-particles, where the angular momentum operators represent actual angular momentum. This is in contrast to a treatment of a pseudospin defined with respect to an arbitrary set of levels, e.g., the clock transition. We use the convention that lower case operators are

used with individual atoms and capital letters are associated with collective variables.

3.3.1 Cesium

In our experiments, our alkali atom of choice is cesium. Throughout this thesis, we exclusively refer to the only stable isotope of cesium: ^{133}Cs with 55 protons and 78 neutrons. The choice of cesium over other alkali atoms that can be trapped and cooled is mostly historical: other Caltech groups have used it and we have available equipment and diode lasers at the appropriate wavelength. Here I summarize a few of the properties of cesium, while leaving the details to other works, most notably [84]. The cesium atom consists of a single valence electron with spin $s = 1/2$ and a nucleus of spin $i = 7/2$. Throughout most of this thesis we consider the D2 optical transition between the ground state with orbital momentum $l = 0$ and the excited state with $l = 1$ and $j = l + s = 3/2$ (as opposed to the D1 line with $j = l - s = 1/2$). The hyperfine levels corresponding to this transition are displayed in figure 3.1. This fine structure splitting corresponds to a wavelength of 852 nm, which is the coarse setting for all of our trapping and probing lasers. Now consider the total angular momentum

$$\hat{\mathbf{f}} = \hat{\mathbf{s}} \otimes \hat{\mathbb{1}}_{l \otimes i} + \hat{\mathbb{1}}_s \otimes \hat{\mathbf{l}} \otimes \mathbb{1}_i + \hat{\mathbb{1}}_{s \otimes l} \hat{\mathbf{i}} \tag{3.53}$$

where $\hat{\mathbf{s}}$, $\hat{\mathbf{l}}$, and $\hat{\mathbf{i}}$ are respectively the electron spin, orbital angular momentum, and the nuclear spin. For the ground state we have two possible spins $f = 7/2 \pm 1/2 = 3, 4$ (each with $2f + 1$ magnetic sublevels), these ground states have a hyperfine splitting of exactly 9.192631770 GHz, which defines the unit of time. For the D2 line there are four excited levels $f' = 2, 3, 4, 5$ (each with $2f' + 1$ sublevels). Thus to simulate the entire ground and excited state manifold, one needs to keep track of $(2 \cdot 3 + 1) + (2 \cdot 4 + 1) = 16$ ground state sublevels and $(2 \cdot 2 + 1) + (2 \cdot 3 + 1) + (2 \cdot 4 + 1) + (2 \cdot 5 + 1) = 32$ excited state sublevels.

We represent the internal state of the atom in terms of the (Zeeman degenerate) atomic hyperfine ground states, $|f, m\rangle$, and excited states, $|f', m'\rangle$. Here f and f' are the total spin quantum numbers for the ground and excited hyperfine levels respectively while m and m' are their projections on the z-axis. That is to say $|f, m\rangle$ are eigenstates of the total atomic angular momentum defined above. The quantum numbers, f, and m, are defined in the

usual manner,

$$\hat{\mathbf{f}}^2|f,m\rangle = \hbar^2 f(f+1)|f,m\rangle \tag{3.54}$$

$$\hat{f}_z|f,m\rangle = \hbar m|f,m\rangle. \tag{3.55}$$

The projector onto the ground state f is given by

$$\hat{P}_f = \sum_{m_f} |f,m_f\rangle\langle f,m_f| \tag{3.56}$$

and the projector onto the excited state f' is given by

$$\hat{P}_{f'} = \sum_{m'_f} |f',m'_f\rangle\langle f',m'_f|. \tag{3.57}$$

Summing over all ground state hyperfine levels gives the ground state projector

$$\hat{P}_g = \sum_f \hat{P}_f \tag{3.58}$$

and summing over all excited state hyperfine levels gives the excited state projector

$$\hat{P}_e = \sum_{f'} \hat{P}_{f'} \tag{3.59}$$

and adding these two together gives the identity $\hat{I} = \hat{P}_e + \hat{P}_g$.

We leave a discussion of the fine and hyperfine energy level splittings to chapter 5. The spectrum and many other properties of cesium can be found in [84].

3.3.2   Ground State Spin Operators

In most of our experiments, the atomic population is primarily among the sublevels of one of the ground states, usually $f = 4$. The excited states are populated transiently via the probing process as discussed in subsequent chapters, and this will eventually lead to state decay to both ground state levels. However, at any one time all of the spins are in one level, therefore we restrict the following analysis to a single ground state spin of size f. The

Cesium D$_2$ Transition
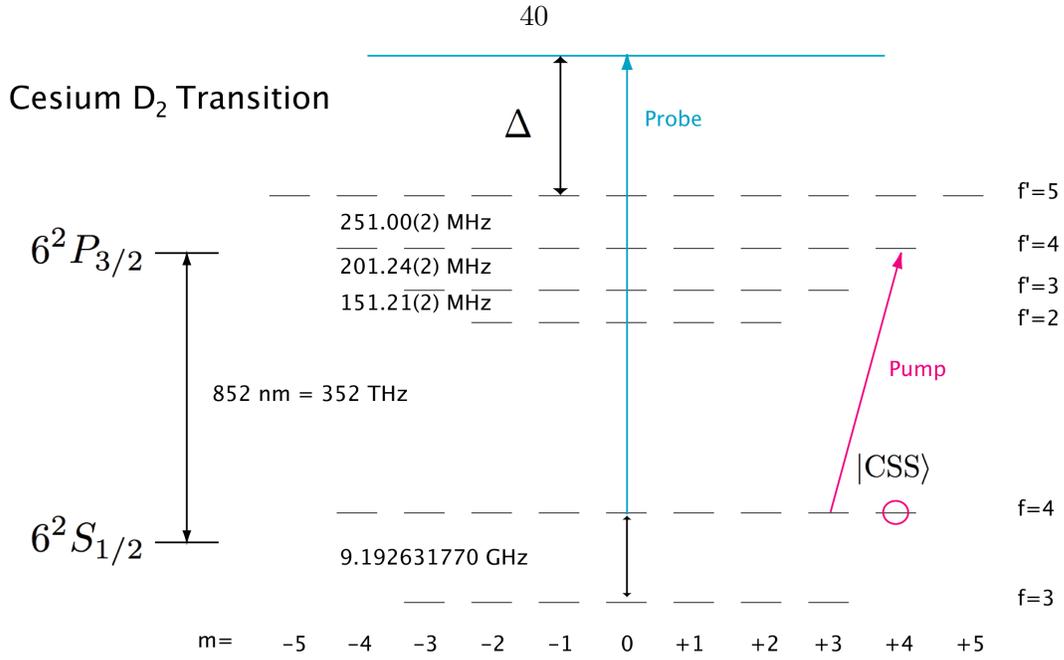
Figure 3.1: Cesium D_2 transition levels as described in the text. The probe and pump laser configurations are also shown.

individual spin vector operator for an atom labeled i is denoted

$$\hat{\mathbf{f}}^{(i)} = [\hat{f}_x^{(i)}, \hat{f}_y^{(i)}, \hat{f}_z^{(i)}] \tag{3.60}$$

and these spin operators obey the usual commutation relations

$$\left[\hat{f}_x, \hat{f}_y\right] = i\hbar\hat{f}_z \tag{3.61}$$

$$\left[\hat{f}_z, \hat{f}_x\right] = i\hbar\hat{f}_y \tag{3.62}$$

$$\left[\hat{f}_y, \hat{f}_z\right] = i\hbar\hat{f}_x. \tag{3.63}$$

In terms of these individual operators, we now define the collective spin vector operator as the sum of all of the individual operators

$$\hat{\mathbf{F}} = \sum_i^N \hat{\mathbf{f}}^{(i)} = [\hat{F}_x, \hat{F}_y, \hat{F}_z] \tag{3.64}$$

and these operators satisfy the same commutation relations as above.

In fact, as is clear from above, the "individual spin" is individual from the atomic

perspective, but also collective because it is composed of the electron spin, orbital and nuclear spin degrees of freedom. Thus, for larger than spin-1/2 states, like the $f = 4$ used here, one can speak of spin-squeezing an individual atom, which corresponds to entangling the atom's inner components.

3.3.3 Spin Uncertainty Relations

Now we translate the commutation relations to uncertainty relations via the usual analysis. From [87], for two operators $\hat{A}$ and $\hat{B}$ we have the general uncertainty relation

$$\langle \Delta \hat{A}^2 \rangle \langle \Delta \hat{B}^2 \rangle \geq |\langle \Delta \hat{A} \Delta \hat{B} \rangle|^2, \tag{3.65}$$

which can be rewritten as

$$\langle \Delta \hat{A}^2 \rangle \langle \Delta \hat{B}^2 \rangle \geq \frac{1}{4} |\langle [\hat{A}, \hat{B}] \rangle|^2 + \frac{1}{4} |\langle \{ \Delta \hat{A}, \Delta \hat{B} \} \rangle|^2. \tag{3.66}$$

Usually the cross-correlation terms (diagonals of the covariance matrix) can be ignored, but we include them here for the sake of completeness. In the case of the collective spin operators, this leads to the uncertainty relations

$$\langle \Delta \hat{F}_x^2 \rangle \langle \Delta \hat{F}_y^2 \rangle \geq \frac{\hbar^2}{4} |\langle \hat{F}_z \rangle|^2 + \frac{1}{4} |\langle \{ \Delta \hat{F}_x, \Delta \hat{F}_y \} \rangle|^2 \tag{3.67}$$

$$\langle \Delta \hat{F}_y^2 \rangle \langle \Delta \hat{F}_z^2 \rangle \geq \frac{\hbar^2}{4} |\langle \hat{F}_x \rangle|^2 + \frac{1}{4} |\langle \{ \Delta \hat{F}_y, \Delta \hat{F}_z \} \rangle|^2 \tag{3.68}$$

$$\langle \Delta \hat{F}_z^2 \rangle \langle \Delta \hat{F}_x^2 \rangle \geq \frac{\hbar^2}{4} |\langle \hat{F}_y \rangle|^2 + \frac{1}{4} |\langle \{ \Delta \hat{F}_z, \Delta \hat{F}_x \} \rangle|^2. \tag{3.69}$$

Below we use these relations to get a sense of the moment distribution on the Bloch sphere of several representative states, including spin-squeezed states.

3.3.4 Symmetric States and Moments

3.3.4.1 Dicke States

The Dicke states [88, 68], $|F, m\rangle$ are simply the eigenstates of $\hat{F}_z$ whether or not it is a collective operator

$$\hat{F}_z |F, m\rangle = \hbar m |F, m\rangle. \tag{3.70}$$

In this section we will calculate some moments of these states for future reference. The collective spin raising and lowering operators are useful for this purpose and are defined as

$$\hat{F}_\pm = \hat{F}_x \pm i\hat{F}_y \tag{3.71}$$

with the inverse being

$$\hat{F}_x = (\hat{F}_+ + \hat{F}_-)/2 \tag{3.72}$$

$$\hat{F}_y = -i(\hat{F}_+ - \hat{F}_-)/2. \tag{3.73}$$

The property that makes $\hat{F}_\pm$ so useful is that it acts as a raising/lowering operator

$$\hat{F}_\pm|F,m\rangle = \hbar\sqrt{F(F+1) - m(m\pm1)}|F,m\pm1\rangle, \tag{3.74}$$

which can be used in expectation value calculations. Now the first-order moments of the Dicke states are

$$\langle F,m|\hat{F}_z|F,m\rangle = \hbar m \tag{3.75}$$

$$\langle F,m|\hat{F}_x|F,m\rangle = \langle F,m|\hat{F}_y|F,m\rangle = 0 \tag{3.76}$$

and the second-order moments are

$$\langle F,m|\hat{F}_z^2|F,m\rangle = \hbar^2 m^2 \tag{3.77}$$

$$\langle F,m|\hat{F}_x^2|F,m\rangle = \langle F,m|\hat{F}_y^2|F,m\rangle = \frac{\hbar^2}{2}[F(F+1) - m^2]. \tag{3.78}$$

Here $\langle \hat{F}_x^2\rangle$ is calculated by expanding $\hat{F}_x$ into the raising/lowering operators and using their properties to simplify.

From these moments, we see that we can visualize each Dicke state m as having a perfectly determined z-component $\hbar m$ and a completely undetermined transverse component while still respecting the extent of the total sphere. The $m = 0$ state lives along the equator of the sphere, while the $m = \pm F$ states represent the polar caps, which still have finite lateral extent. The single atom Dicke states corresponding to the sublevels of the $f = 4$ ground state of cesium are depicted visually in figure 3.2.

Dicke States:

$|f = 4, m\rangle_z, m = -f \cdots f$

$\text{Radius} = \hbar\sqrt{f(f+1)}$

Z

y

X
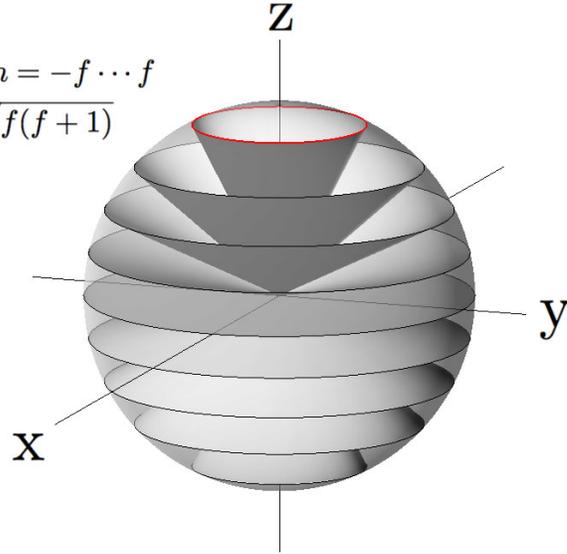
Figure 3.2: Graphical depiction of Dicke States, $|f, m\rangle_z$, for the cesium $f = 4$ ground state. The highlighted uppermost state is a single-atom coherent spin-state.

We define these extended Dicke states with $m = F$ as the *coherent spin-states* (CSS), which obey

$$\langle F, F|\hat{F}_z|F, F\rangle = \hbar F \qquad (3.79)$$

$$\langle F, F|\hat{F}_x^2|F, F\rangle = \langle F, F|\hat{F}_y^2|F, F\rangle = \frac{\hbar^2}{2}F \qquad (3.80)$$

$$\langle\{\Delta\hat{F}_x, \Delta\hat{F}_y\}\rangle = 0. \qquad (3.81)$$

The CSS is a minimum uncertainty state because it satisfies the equality of the uncertainty inequality

$$\langle\Delta\hat{F}_x^2\rangle\langle\Delta\hat{F}_y^2\rangle = \frac{\hbar^2}{4}|\langle\hat{F}_z\rangle|^2 + \frac{1}{4}|\langle\{\Delta\hat{F}_x, \Delta\hat{F}_y\}\rangle|^2. \qquad (3.82)$$

As opposed to all other Dicke states, the CSS is an unentangled (separable) state, with all spins pointing in the same direction, i.e., $|\text{CSS}\rangle = |F, F\rangle = |\uparrow \cdots \uparrow\rangle$. The entanglement of non-CSS Dicke States is simply seen by the two spin example in chapter 1 and discussed more thoroughly in chapter 4.

As discussed in the introduction, the CSS can be turned into a spin-squeezed state if the x variance can be reduced (squeezed) while the y variance is increased (antisqueezed)

to maintain the validity of the uncertainty relation. Below we discuss the definition of spin-squeezing in more detail.

The two single atom states that are easiest to prepare via optical pumping are the individual CSS $|f, f\rangle$, and the individual $m = 0$ state $|f, m\rangle$. From the above relations we can see that the difference in variance one expects to see, for $f = 4$, is

$$\langle f, f | \hat{f}_x^2 | f, f \rangle = \frac{\hbar^2}{2} f = \hbar^2 2 \tag{3.83}$$

$$\langle f, 0 | \hat{f}_x^2 | f, 0 \rangle = \frac{\hbar^2}{2} f(f+1) = \hbar^2 10. \tag{3.84}$$

Another important reference state is the completely mixed state defined by incoherently populating all of the sublevels equally

$$\hat{\rho} = \sum_m \frac{1}{2f+1} |f, m\rangle\langle f, m|. \tag{3.85}$$

With this state, one can then show that the mean square of any moment is

$$\langle \hat{f}_i^2 \rangle = \text{Tr}[\hat{f}_i^2 \hat{\rho}] = \frac{2\hbar^2}{2f+1} \sum_{m=1}^{f} m^2 \tag{3.86}$$

and for $f = 4$ this gives

$$\langle \hat{f}_i^2 \rangle = \text{Tr}[\hat{f}_i^2 \hat{\rho}] = \hbar^2 20/3 = \hbar^2 6.67. \tag{3.87}$$

From above, we see that this variance is intermediate between that of a coherent state and the $m = 0$ state as expected.

3.3.4.2 Gaussian States

The Hilbert space describing an ensemble of N spins of size f is ridiculously large, and without any constraints the state vector would consist of $(2f + 1)^N$ complex elements. When we deal with an ensemble of more than a billion spins, there is no hope of using the full state space in our description. One option is to restrict the state description to the symmetric subspace, as discussed in the next chapter, which only has size $(2Nf+1)$, however this is still a potentially large number. Fortunately, we can use a Gaussian description of

the state under certain circumstances where we only have to keep track of a few moments of the distribution. In general, this kind of simplification of the Hilbert space description is a kind of "model reduction," used to make simulations and estimation more efficient by removing unnecessary parts of the full model.

For example, consider the collective CSS just described with a large number of spins. If this state is aligned along z and we represent it in a perpendicular basis, x or y, then the populations among the collective sublevels will be distributed in a Gaussian manner, centered at zero with a variance given by $\langle F, F | \hat{F}_x^2 | F, F \rangle = \frac{\hbar^2}{2} F$. Thus the width of this distribution grows as the square root of the number as the collective Bloch sphere grows linearly with the number. The more atoms we have, the more localized the state is on the sphere, and the "flatter" the space on which it lives. Formally, these ideas can be made precise via the Holstein-Primakoff transformation [89], which is commonly used in the condensed matter physics literature and makes it possible to derive the Gaussian approximation as an expansion in $1/N$.

Many other papers, such as [39, 86], make use of this approximation and divide by the "classical" length of the vector to create effective "position" and "momentum" operators

$$\frac{\hat{F}_x}{\sqrt{\langle \hat{F}_z \rangle}} \rightarrow \hat{X} \tag{3.88}$$

$$\frac{\hat{F}_y}{\sqrt{\langle \hat{F}_z \rangle}} \rightarrow \hat{P}, \tag{3.89}$$

which then follows the usual commutation and uncertainty relationship. In this thesis, we make this analogy in appendix D to simulate a quantum harmonic oscillator with a field along the same direction as the CSS.

These concepts are also discussed in the context of measurement and control in both chapter 9 and chapter 10. In section 9.7, I describe how to expand the Gaussian description naturally in terms of *cumulants*, which stochastically evolve due to the quantum measurement in an interdependent way.

So far we have only discussed the Dicke states and the CSS, which is an unentangled Dicke state. As seen in figure 1.1, the spin-squeezed state is merely a perturbation of the CSS for small degrees of squeezing and is also a Gaussian state. Thus, for much of the work described in this thesis where the spin-squeezing is relatively small, the Gaussian

approximation will be used.

3.3.5 Spin-Squeezed States

The notion of spin-squeezing has been introduced via the uncertainty relations in chapter 1 and the cartoon in figure 1.1. In this section, we technically define a spin-squeezed state and analyze some of its properties.

For a collective spin-state polarized in the x-direction ($\langle \hat{F}_y \rangle = \langle \hat{F}_z \rangle = 0$) and with minimum transverse uncertainty in the z-direction, the spin-squeezing parameter ξ is defined as

$$\xi^2 = \frac{2F \langle \Delta \hat{F}_z^2 \rangle}{\langle \hat{F}_x \rangle^2}. \tag{3.90}$$

If a state has $\xi^2 < 1$, then that state is referred to as a *spin-squeezed state* (SSS). Notice that the coherent spin-state satisfies $\xi^2 = 1$. The spin-squeezing parameter was introduced in the context of improving the performance of atomic clocks in [4, 2, 5]. In the next section, we prove that spin-squeezed states (composed of spin-1/2 particles) are necessarily entangled.

3.3.5.1 Spin-Squeezing Implies Entanglement (Spin-1/2)

Following [6], we now show that any state composed of spin-1/2 particles with $\xi^2 < 1$ is nonseparable (entangled). First we define some of the spin-operators for a spin-1/2 particle. The Pauli angular momentum operators for the individual spins are

$$\hat{f}_x = \frac{1}{2}(|0\rangle\langle 1| + |1\rangle\langle 0|) \tag{3.91}$$

$$\hat{f}_y = \frac{i}{2}(|0\rangle\langle 1| - |1\rangle\langle 0|) \tag{3.92}$$

$$\hat{f}_z = \frac{1}{2}(|1\rangle\langle 1| - |0\rangle\langle 0|). \tag{3.93}$$

The joint/collective angular momentum variables are the sums of the microscopic variables: $\hat{F}_i = \sum_{n=1}^{N} \hat{f}_i^{(n)}$. The total angular momentum vector will be denoted $\hat{\mathbf{F}} = [\hat{F}_x, \hat{F}_y, \hat{F}_z]$.

With the definition $\hat{F}_{ij} = \sum_n^N |i\rangle_n \langle j|_n$, $i,j = 0,1$ we see that

$$\hat{F}_x = \frac{1}{2}(\hat{F}_{01} + \hat{F}_{10}) \tag{3.94}$$

$$\hat{F}_y = \frac{i}{2}(\hat{F}_{10} - \hat{F}_{01}) \tag{3.95}$$

$$\hat{F}_z = \frac{1}{2}(\hat{F}_{00} - \hat{F}_{11}) \tag{3.96}$$

$$N = \hat{F}_{00} + \hat{F}_{11}. \tag{3.97}$$

Now, to begin the proof, we use the fact that any separable state can be written in the form

$$\hat{\rho} = \sum_k P_k \hat{\rho}_k^{(1)} \otimes \cdots \otimes \hat{\rho}_k^{(N)}. \tag{3.98}$$

We can then rewrite the variance for a separable state into the useful form

$$\langle \hat{F}_z^2 \rangle = \mathrm{Tr}\left(\hat{\rho}\hat{F}_z^2\right) \tag{3.99}$$

$$= \mathrm{Tr}\left(\sum_k P_k \hat{\rho}_k^{(1)} \otimes \cdots \otimes \hat{\rho}_k^{(N)} \left(\sum_{n=1}^N \hat{f}_z^{(n)}\right)^2\right) \tag{3.100}$$

$$= \sum_k P_k \left\langle \sum_{n=1}^N \hat{f}_z^{(n)2} + \sum_{n \neq m}^N \hat{f}_z^{(n)} \hat{f}_z^{(m)} \right\rangle_k \tag{3.101}$$

$$= \sum_k P_k \left(\frac{N}{4} + \left\langle \sum_{n \neq m}^N \hat{f}_z^{(n)} \hat{f}_z^{(m)} \right\rangle_k \right) \tag{3.102}$$

$$= \frac{N}{4} + \sum_k P_k \left(\langle \hat{F}_z \rangle_k^2 - \sum_{n=1}^N \langle \hat{f}_z^{(n)} \rangle_k^2 \right). \tag{3.103}$$

The relations $\sum_k P_k = 1$, $\hat{f}_i^2 = \frac{1}{4}$, and $\langle \sum_{n \neq m}^N \hat{f}_z^{(n)} \hat{f}_z^{(m)} \rangle = \sum_{n \neq m}^N \langle \hat{f}_z^{(n)} \rangle \langle \hat{f}_z^{(m)} \rangle = \langle \hat{F}_z \rangle^2 - \sum_{n=1}^N \langle \hat{f}_z^{(n)} \rangle^2$ have all been used, and the last equation follows from separability. So the variance is now

$$\langle \Delta \hat{F}_z^2 \rangle = \langle \hat{F}_z^2 \rangle - \langle \hat{F}_z \rangle^2 \tag{3.104}$$

$$= \frac{N}{4} + \sum_k P_k \left(\langle \hat{F}_z \rangle_k^2 - \sum_{n=1}^N \langle \hat{f}_z^{(n)} \rangle_k^2 \right) - \langle \hat{F}_z \rangle^2. \tag{3.105}$$

This is equality 1.

Another condition $\langle \hat{f}_x^{(n)} \rangle_k^2 + \langle \hat{f}_y^{(n)} \rangle_k^2 + \langle \hat{f}_z^{(n)} \rangle_k^2 \leq \frac{1}{4}$ implies that

$$\frac{N}{4} - \sum_k P_k \sum_n \langle \hat{f}_z^{(n)} \rangle_k^2 \geq \sum_k P_k \sum_n \left(\langle \hat{f}_x^{(n)} \rangle_k^2 + \langle \hat{f}_y^{(n)} \rangle_k^2 \right). \tag{3.106}$$

This is inequality 1.

Now we use the Cauchy-Schwarz inequality to derive two more useful inequalities. First notice that

$$
\begin{aligned}
\langle \hat{F}_i \rangle^2 &= \left(\sum_k P_k \langle \hat{F}_i \rangle_k \right)^2 & \text{(3.107)} \\
&= \left(\sum_k \left(\sqrt{P_k} \right) \left(\sqrt{P_k} \langle \hat{F}_i \rangle_k \right) \right)^2 & \text{(3.108)} \\
&\leq \left(\sum_k P_k \right) \left(\sum_k P_k \langle \hat{F}_i \rangle_k^2 \right) & \text{(3.109)} \\
&= \sum_k P_k \langle \hat{F}_i \rangle_k^2. & \text{(3.110)}
\end{aligned}
$$

This is inequality 2. Now we continue this inequality and use Cauchy-Schwarz again to get

$$
\begin{aligned}
\langle \hat{F}_i \rangle^2 &\leq \sum_k P_k \langle \hat{F}_i \rangle_k^2 & \text{(3.111)} \\
&= \sum_k P_k \left(\sum_n \langle \hat{f}_i^{(n)} \rangle_k \right)^2 & \text{(3.112)} \\
&= \sum_k P_k \left(\sum_n 1 \times \langle \hat{f}_i^{(n)} \rangle_k \right)^2 & \text{(3.113)} \\
&\leq \sum_k P_k \left(\sum_n 1 \right) \left(\sum_n \langle \hat{f}_i^{(n)} \rangle_k^2 \right) & \text{(3.114)} \\
&= \sum_k P_k N \sum_n \langle \hat{f}_i^{(n)} \rangle_k^2. & \text{(3.115)}
\end{aligned}
$$

This is inequality 3.

Finally, we combine these relations to derive an inequality for the ξ^2 of a separable state:

$$\xi^2 = \frac{N\langle\Delta\hat{F}_z^2\rangle}{\langle\hat{F}_x\rangle^2 + \langle\hat{F}_y\rangle^2} \tag{3.116}$$

$$= \frac{N\left(\frac{N}{4} + \sum_k P_k\left(\langle\hat{F}_z\rangle_k^2 - \sum_{n=1}^N\langle\hat{f}_z^{(n)}\rangle_k^2\right) - \langle\hat{F}_z\rangle^2\right)}{\langle\hat{F}_x\rangle^2 + \langle\hat{F}_y\rangle^2} \tag{3.117}$$

$$\geq \frac{N\left(\frac{N}{4} + \sum_k P_k\left(\langle\hat{F}_z\rangle_k^2 - \sum_{n=1}^N\langle\hat{f}_z^{(n)}\rangle_k^2\right) - \langle\hat{F}_z\rangle^2\right)}{\sum_k P_k N \sum_n(\langle\hat{f}_x^{(n)}\rangle_k^2 + \langle\hat{f}_y^{(n)}\rangle_k^2)} \tag{3.118}$$

$$\geq \frac{\frac{N}{4} + \sum_k P_k\left(\langle\hat{F}_z\rangle_k^2 - \sum_{n=1}^N\langle\hat{f}_z^{(n)}\rangle_k^2\right) - \sum_k P_k\langle\hat{F}_z\rangle_k^2}{\sum_k P_k \sum_n\left(\langle\hat{f}_x^{(n)}\rangle_k^2 + \langle\hat{f}_y^{(n)}\rangle_k^2\right)} \tag{3.119}$$

$$\geq \frac{\frac{N}{4} - \sum_k P_k \sum_{n=1}^N\langle\hat{f}_z^{(n)}\rangle_k^2}{\sum_k P_k \sum_n\left(\langle\hat{f}_x^{(n)}\rangle_k^2 + \langle\hat{f}_y^{(n)}\rangle_k^2\right)} \tag{3.120}$$

$$\geq 1 \tag{3.121}$$

where equality 1 is used in 3.117, inequality 3 is used in 3.118, inequality 2 is used in 3.119, and inequality 1 is used in 3.121. Thus separability implies the absence of spin-squeezing. Conversely, spin-squeezing implies entanglement.

3.3.5.2 Heisenberg Limit of Spin-Squeezing

Optical squeezed states can in principle be infinitely squeezed, although this takes an infinite amount of energy. For a fixed number of atoms, spin-squeezed states cannot be infinitely squeezed because of the fact that the collective Bloch sphere is finite in extent, and if one squeezes too much then the distribution begins to "wrap" around the sphere.

We can describe this quantitatively as follows. Assume the sample is polarized along x: $\langle\hat{F}_x\rangle = F = \frac{N}{2}$. By standard angular momentum commutation relations we know $\langle\Delta\hat{F}_y^2\rangle\langle\Delta\hat{F}_z^2\rangle \geq \frac{\langle\hat{F}_x\rangle^2}{4}$. So $\langle\Delta\hat{F}_z^2\rangle \geq \frac{F^2}{4\langle\Delta\hat{F}_y^2\rangle}$. But because the variance cannot be bigger than the entire Bloch sphere we have $\langle\Delta\hat{F}_y^2\rangle \leq \hat{F}^2$, which implies $\langle\Delta\hat{F}_z^2\rangle \geq \frac{1}{4}$. This gives us $\xi^2 = \frac{N\langle\Delta\hat{F}_z^2\rangle}{\langle\hat{F}_x\rangle^2 + \langle\hat{F}_y\rangle^2} \geq \frac{N/4}{F^2} = \frac{1}{N}$. This is the so called *Heisenberg limit* of spin-squeezing.

3.3.5.3 Spin-Squeezing Under Particle Loss

Here we analytically investigate the spin-squeezing parameter as particles are removed. We find that the spin-squeezing parameter of a state with N_r spins remaining $(\xi_{N_r}^2)$ is dependent

on the initial squeezing parameter (ξ_N^2) and polarization of the state with all spins remaining in the following way

$$\xi_{N_r}^2 = \xi_1^2 + (\xi_N^2 - \xi_1^2)\frac{N_r - 1}{N - 1} \tag{3.122}$$

where $\xi_1^2 \equiv N^2/(4\langle \hat{F}_z \rangle_N^2)$. So the loss of squeezing is only linear in the number of particles that are traced out. Hence the entanglement is rather robust, especially as compared to the extreme fragility of cat states, described below, to particle loss. The idea of entangled states that are robust to particle loss is also discussed in chapter 4.

3.3.5.4   Squeezing Generation

As mentioned in the introduction, there are many ways to imagine creating a spin-squeezed state. Numerically, the easiest way to spin-squeeze a coherent state is to apply either a twisting Hamiltonian ($\hat{H} = \hat{F}_x^2$) or a countertwisting Hamiltonian ($\hat{H} = -i(\hat{F}_+^2 - \hat{F}_-^2)$) [2]. The latter is designed to ensure that the squeezed axis stays constant in time and does not "twist" out of alignment.

On the experimental side, one can either apply controlled Hamiltonians to entangle the internal degrees of freedom as has been achieved for several ions in a trap [40]. One could also create the correlations in the atoms via absorption of a beam of squeezed light [39]. Finally, one could use QND measurement to prepare spin-squeezed states conditionally [16, 37, 36] as we mostly discuss in this thesis.

3.3.6   Cat States

As we have mentioned, the Hilbert space describing the collective spin-states is a large place. So far we have only described a very small but useful subsample of the vast number of potential states. Other states in the symmetric subspace we have yet to consider include the GHZ state

$$|\text{GHZ}\rangle = (|\uparrow\uparrow \cdots \uparrow\rangle + |\downarrow\downarrow \cdots \downarrow\rangle)/\sqrt{2}, \tag{3.123}$$

which is alternately known as a cat state, or an EPR state. One can also prepare similar coherent superpositions of CSS states but with a smaller angle between the two [90]. The entanglement properties of these states and the states mentioned previously are analyzed in the next chapter.

Chapter 4

Entanglement of Collective Spin-States

This chapter provides a highly abbreviated summary of [17], which discusses the efficient calculation of entanglement measures for symmetric spin-states. For the sake of brevity, I left out most of the details and proper references, with the aim of getting across the few core ideas. Although this work is interesting from a theoretical entanglement perspective, it is largely independent of the rest of this thesis and can safely be skipped or read separately.

4.1 Introduction

In [17], we investigate entanglement restricted to symmetric spin-states. The collective state size without restrictions is exponentially large in the atom number. With the symmetry restriction, on the other hand, the state grows only linearly with number. We show that, just as the state can be simplified, so can the operations needed to calculate entanglement measures for bipartite splits of symmetric states. As a result, we can efficiently calculate entanglement measures for ensemble numbers up to $N \approx 10^3$.

Using this machinery, we can numerically investigate the entanglement properties of certain parametrized families of symmetric states. This has allowed us to infer certain properties that can later be proven analytically. When investigating a particular state we can ask questions of both large scale entanglement (e.g., between halves of the ensemble) and small scale entanglement (e.g., between two spins removed from the rest of the ensemble). The small scale entanglement also tells us how robust the original state was to particle loss. Finally, these tools have allowed us to investigate how entanglement is dynamically

generated via, for example, a spin-squeezing Hamiltonian.

4.2   Computing Entanglement

The question of how to quantify entanglement is notoriously subtle [91], but the basic idea under ideal circumstances is very simple. Given a collective pure state, $|\Psi\rangle$, and a partition for the system, $\{A, B\}$, the entropy of entanglement is defined as,

$$E(|\Psi\rangle, \{A, B\}) \equiv S(\hat{\rho}_A) = S(\hat{\rho}_B) \tag{4.1}$$

where the von Neumann entropy is $S(\hat{\rho}) = -\mathrm{Tr}(\hat{\rho} \log_2 \hat{\rho})$ and $\hat{\rho}_A = \mathrm{Tr}_B(|\Psi\rangle\langle\Psi|)$. Any entropy that results from performing a partial trace on the system must be a consequence of initial entanglement provided that the initial state is pure. For product states, $|\Psi\rangle = |\Psi\rangle_A \otimes |\Psi\rangle_B$, the entropy is zero since the single eigenvalue for each of the pure states $\hat{\rho}_A$ and $\hat{\rho}_B$ is one. The maximum entropy of entanglement given a partition with dimensions, $\dim(A) = d_A$ and $\dim(B) = d_B$, with $d_A \leq d_B$, is $\log_2(d_A)$. A state that achieves this maximum is,

$$|\Psi\rangle = |0\rangle_A \otimes |0\rangle_B + |1\rangle_A \otimes |1\rangle_B + \cdots + |d_A - 1\rangle_A \otimes |d_A - 1\rangle_B. \tag{4.2}$$

For initial pure states, the entropy as a measure of entanglement is very intuitive: the more two parts of a single system are entangled, the more information is lost from one when the other is lost. This is the same as with any meaningful relationship in life, quantum mechanical, personal, or otherwise.

The entropy of entanglement has the attractive feature that it is straightforward to compute; it requires only performing a partial trace, $\hat{\rho}_A = \mathrm{Tr}_B(\hat{\rho})$, then computing eigenvalues of the result. The drawback of the entropy is that it only qualifies as an entanglement measure for initially pure states. One can define a set of constraints that all measures of entanglement, i.e., entanglement monotones, must satisfy [92]. However, in the case of initial pure states, all entanglement measures must be equivalent to the entropy as defined above.

4.3   Computing Entanglement in the Symmetric Subspace

For simplicity we consider the entanglement within a collective state of N spin-1/2 particles. When working with the symmetric subspace, it is necessary to convert between the large full, $\mathbb{C}_2^{\otimes N}$, and small symmetric, $\mathbb{C}_{N+1}$, basis representations of the state. In order to provide a systematic means for changing bases, it is convenient to define a symmetry operator, $\hat{S}_N : \mathbb{C}_2^{\otimes N} \to \mathbb{C}_{N+1}$, whose action on the density operator in the 2^N-dimensional basis,

$$\widetilde{\hat{\rho}}_N = \hat{S}_N \hat{\rho}_N \hat{S}_N^\dagger \tag{4.3}$$

projects the state into $\mathbb{S}_N$ expressed in an $(N+1)$-dimensional basis. We have adopted the notation that $\widetilde{\hat{\rho}}_N$ is the symmetric density matrix represented in $\mathbb{C}_{N+1}$.

$\hat{S}_N$ is an $(N+1) \times 2^N$-dimensional matrix that transfers between the bases. It should be noted that $\hat{S}_N$ is not a permutation operator, but rather a projector. Therefore, it is only appropriate to operate on symmetric states with $\hat{S}_N$ as,

$$\hat{S}_N \hat{S}_N^\dagger \;=\; \mathbb{1}_{\text{sym}} \tag{4.4}$$

$$\hat{S}_N^\dagger \hat{S}_N \;\neq\; \mathbb{1}_{\text{full}} \tag{4.5}$$

where $\mathbb{1}_{\text{sym}}$ is the identity in the $(N+1)$-dimensional symmetric basis and $\mathbb{1}_{\text{full}}$ is the identity in the 2^N-dimensional full basis. Consequently, $\hat{S}_N^\dagger \hat{S}_N \hat{\rho}_N \hat{S}_N^\dagger \hat{S}_N = \hat{\rho}_N$ only if $\hat{\rho}_N$ is symmetric. Acting on a nonsymmetric state with $\hat{S}_N$ and $\hat{S}_N^\dagger$ results in a loss of information, as the nonsymmetric components of that state are lost in the projection onto $\mathbb{S}_N$.

For the purpose of making a bipartite split, $\{A, B\}$, the essential property of the symmetric subspace is that it can be expressed as a tensor product of smaller symmetric spaces. However, the tensor product of arbitrary symmetric states is not necessarily symmetric,

$$\mathbb{S}_N \subset \mathbb{S}_{N-k} \otimes \mathbb{S}_k \tag{4.6}$$

where the partition $\{A, B\}$ has been denoted by the number of spins in each subsystem, $\{N-k, k\}$. $\mathbb{S}_{N-k} \otimes \mathbb{S}_k$ is larger than $\mathbb{S}_N$. The structure of valid symmetric products is

given by the relation [93],

$$|m, N\rangle = \sum_{p=0}^{k} |m - p, N - k\rangle \otimes |p, k\rangle \tag{4.7}$$

in terms of constituent symmetric states expressed in the large basis.

The above equations raise the point that the N particle symmetric space, $\mathbb{S}_N$, is smaller than the product space, $\mathbb{S}_{N-k} \otimes \mathbb{S}_k$. Therefore, the entanglement of states in $\mathbb{S}_N$ will generally be more restricted than those in the tensor product space. While, it is straightforward to identify the maximal entanglement bounds for states in, $\mathbb{S}_{N-k} \otimes \mathbb{S}_k$, the same is not true for $\mathbb{S}_N$. Therefore, it is convenient to use the product space entanglement bounds as an upper limit, albeit an overestimate, for the scaling of states in $\mathbb{S}_N$.

To calculate the entanglement between different partitions, we need to perform the partial trace. In particular we would like to derive an expression for

$$\widetilde{\widehat{\rho}}_{N-k} = \mathrm{Tr}_k[\widetilde{\widehat{\rho}}_N] \tag{4.8}$$

that avoids expressing any of the density matrices (in any intermediate step) in their large bases. Staying within the smaller representation will clearly allow calculations of entanglement measures to be considerably more efficient.

Using a formalism involving the projection operators above, we show in [17] that this is possible, resulting in an expression relating components of the combined and traced out density matrices:

$$\langle \widetilde{a} | \widetilde{\widehat{\rho}}_{N-k} | \widetilde{b} \rangle = \sum_{j=0}^{k} \langle \widetilde{a+j} | \widetilde{\widehat{\rho}}_N | \widetilde{b+j} \rangle C_{k,j}^{-2} \frac{C_{N,a+j} C_{N,b+j}}{C_{N-k,a} C_{N-k,b}} \tag{4.9}$$

where we use

$$C_{N,m} = \binom{N}{m}^{-\frac{1}{2}} = \left[\frac{N!}{m!(N-m)!} \right]^{-\frac{1}{2}}. \tag{4.10}$$

The final state then resides in $\mathbb{C}_{N-k+1}$. Here the $|\widetilde{n}\rangle$ state is just a Dicke state, as presented in the previous chapter (see figure 3.2), expressed in the notation of [17].

Other operations besides the partial trace, such as the partial transpose, that are necessary to compute certain entanglement measures can also be represented in a similarly
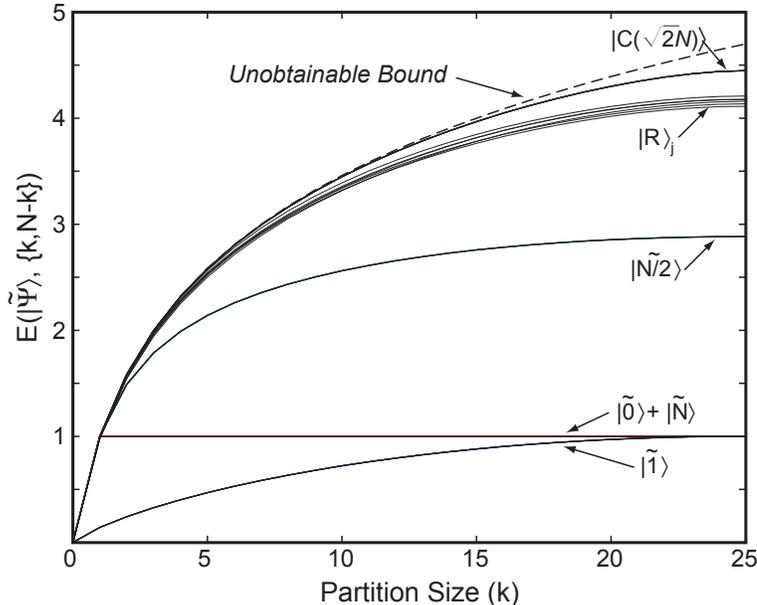
Figure 4.1: Entropy of entanglement for representative symmetric states with $N = 50$ particles as a function of the dimension of the bipartite split, $\{k, N - k\}$, where $k = 1, \ldots, \lfloor N/2 \rfloor$. The *unobtainable bound* $\log_2(k+1)$ is the entropy that could be achieved by a nonsymmetric product of the two symmetric subsystems, $\{A, B\}$. Several representative states nearly achieve this maximum.

efficient way.

4.4 Entanglement of Reference States

After deriving the set of functions that facilitate efficient numerical calculation of entanglement measures via symmetry, the rest of [17] concentrates on using these relations to investigate the entanglement properties of certain families of states for large numbers (several hundred spins, which is essentially the asymptotic regime).

For example, in figure 4.1 we plot the entanglement as a function of partition size for a fixed number of total spins. The maximal bound is provided by the state of equation (4.2), which does not live in the overall symmetric subspace. We plot the entanglement of several states including the Dicke states $|\tilde{n}\rangle$, the GHZ (cat) state $(|\tilde{0}\rangle + |\tilde{N}\rangle)/\sqrt{2}$, random states with a coherent sum of randomly populated Dicke states, and the comb state $|C(\sqrt{2N})\rangle$ defined in [17], which nearly attains the maximal bound.

One can also consider only the even split point $E(\rho, \{N/2, N/2\})$, which is rightmost in figure 4.1, and vary the total number to investigate the scaling. This is shown in figure
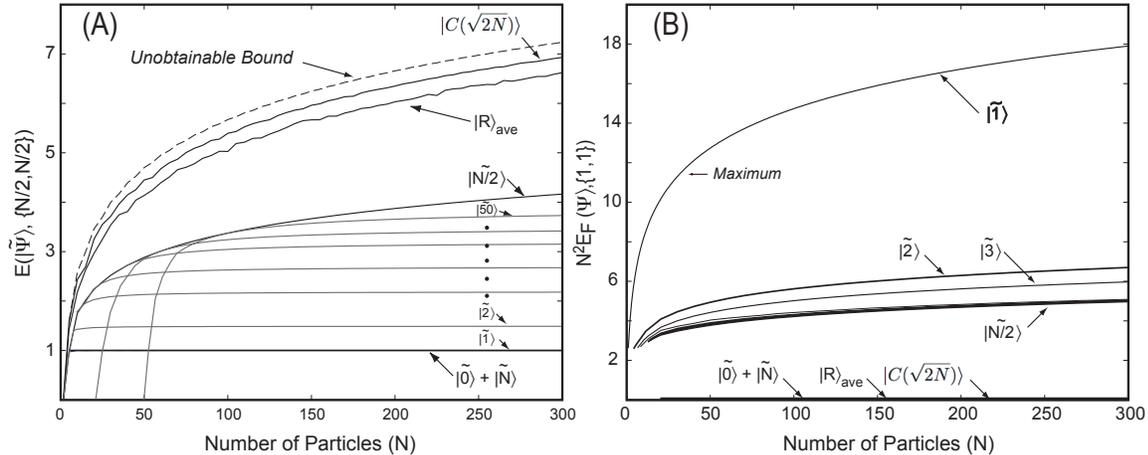
Figure 4.2: (A) Plot of the even split entropy of entanglement, $E(\rho, \{N/2, N/2\})$, for representative states as a function of the number of particles, N (which is also equal to the entanglement of formation and distillation). Note that the average entropy of 25 random states, $|R\rangle$, as well as the entropy of $|C(\sqrt{2N})\rangle$, nearly attain the *unobtainable bound* $\log_2(\lfloor N/2 \rfloor + 1)$. (B) A plot of the two-particle entanglement of formation, $E_F(\rho, \{1,1\})$, as a function of the number of particles, N. The W state, $|\widetilde{1}\rangle$, maximizes this entanglement measure, which quantifies robustness to particle loss.

4.2A. While the Dicke states do not scale with the upper bound, it appears that both the random states and comb states do scale. For the comb states, this numerical observation was turned into analytical proof in the appendix of [17]. The fact that an overall symmetric state could have an entanglement that scaled with this upper bound was not known prior to this paper and is a testament to the power of the numerical tools developed here.

The quantity $E(\rho, \{N/2, N/2\})$ can be taken as a notion of "large-scale" entanglement in the ensemble. By contrast, consider $E_F(\rho, \{1,1\})$, which represents the entanglement between two individual spins with the remaining $N - 2$ traced out. (Because the state of two-particles removed from the rest is mixed we use the entanglement of formation rather than the entropy [17].) We plot the scaled version of this quantity for various states in figure 4.2B, where it is seen that the Dicke states have more pairwise entanglement than those states with more large scale entanglement. This indicates some sort of trade-off between large and small scale entanglement. This trade-off is plotted in figure 4.3C for a particular number of spins using the same representative states.
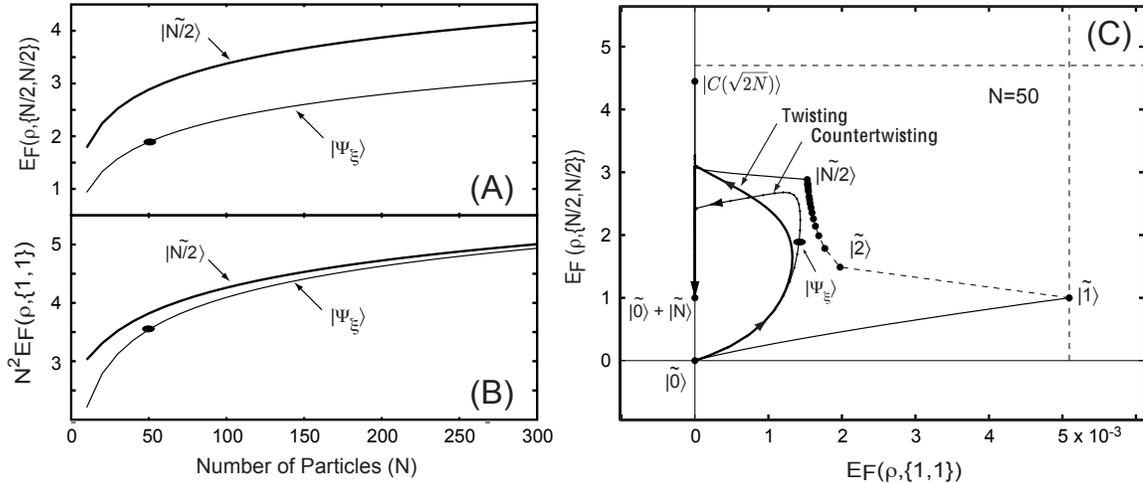
Figure 4.3: (A) A plot of the even split entanglement of formation (and entropy), $E_F(\rho, \{N/2, N/2\})$, for a system of N spin-1/2 particles evolved under a countertwisting spin-squeezing Hamiltonian. The state $|\Psi_\xi\rangle$ minimizes the squeezing parameter, ξ^2. (B) A similar plot using the scaled entanglement of formation, $N^2 E_F(\rho, \{1, 1\})$, for a system with all but two particles removed. (C) The time evolution of states evolving under both the countertwisting Hamiltonian ($H_{ct} = (\hat{F}_+^2 - \hat{F}_-^2)/i$) and twisting Hamiltonian ($H_t = \hat{F}_x^2$) in the space of extremal split entanglement.

4.5 Spin-Squeezing and the Dynamics of Entanglement

One can now plot these entanglement measures with reference to spin-squeezed states as generated by, for example, the countertwisting Hamiltonian $H_{ct} = (\hat{F}_+^2 - \hat{F}_-^2)/i$. In figure 4.3C, the path of an evolving spin-squeezed state is plotted alongside the reference states previously discussed. The scalings in number of the optimal spin-squeezed point (denoted by the dot) for large and small scale entanglement are plotted in figure 4.3A/B.

By looking at these plots it is clear that the spin-squeezed states are particularly robust to particle loss (as discussed in section 3.3.5.3), like the Dicke states, and unlike the more even-split entangled states (random, comb, GHZ, etc.). However, the spin-squeezed states also have a fair amount of large-scale even-split entanglement. Given the observed trade-off between small and large scale entanglement, it appears that the spin-squeezed states represent a convenient compromise.

Another interesting venue of investigation using these tools is to investigate how the entanglement grows on different scales when being dynamically generated. In figure 4.4, the generation of squeezing is plotted along with the (normalized) small and large scale entanglement. One can see that the small scale entanglement peaks first, giving way to the
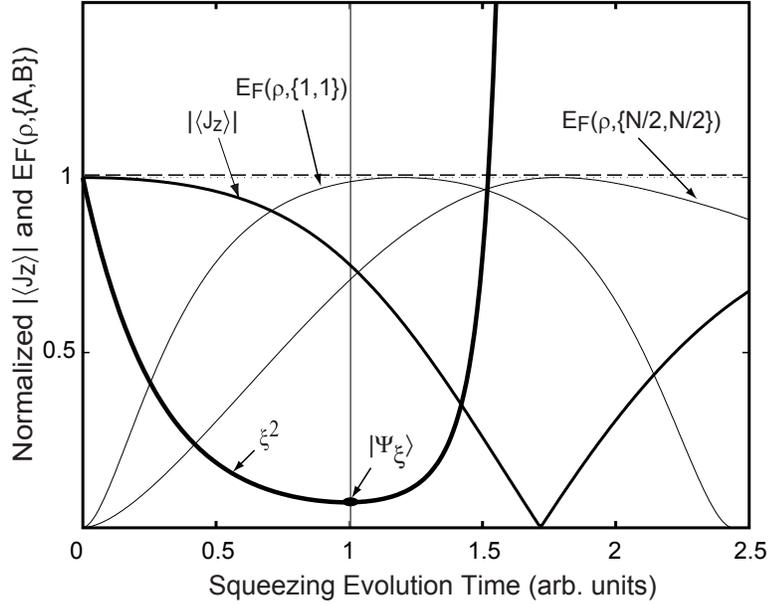
Figure 4.4: Spin-squeezing evolution for a system of $N = 50$ spin-$1/2$ particles evolving by the countertwisting Hamiltonian as measured by the squeezing parameter, ξ^2. The time is scaled such that maximal spin-squeezing occurs at $t = 1$. The mean or $\hat{F}_z$ (i.e., J_z) and the entanglements of formation are all independently normalized by their own maximum in the time period shown. Notice that the small scale correlations, $E_F(\rho, \{1, 1\})$, peak before the large scale correlations, $E_F(\rho, \{N/2, N/2\})$, as the squeezing evolves.

larger scale entanglement at subsequent times. This ordering might be expected, as one intuitively expects the pairwise correlations to grow first.

Another observation is that for small times, the state gets progressively more entangled in the sense of majorization [94]. In other words, the eigenvalues of $\mathrm{Tr}_k(\rho(t + dt))$ are more disordered than the eigenvalues of $\mathrm{Tr}_k(\rho(t))$ for all $k \leq N/2$ and small t. Thus, despite certain ordering difficulties with various entropies, the entanglement of any split is strictly increasing initially.

The paper [17] on which this chapter is based can be extended in many directions. First, one can use the tools developed to discover properties of entanglement and its generation in symmetric states, as we have only begun. Second, one can use the same idea as this paper to extend the efficient calculation of entanglement measures to other classes of states, besides the symmetric examples discussed here.

Chapter 5

The Polarizability Hamiltonian and Spontaneous Emission from Adiabatic Elimination

Now that we have described the relevant optical and atomic states, we are equipped to begin describing the interaction of the optical probe beam with the atomic spin-state. In this chapter, we derive the master equation describing the unconditional evolution of a single alkali atom when driven by a far-off resonant beam. Although our goal in subsequent chapters will be to describe the effect of the atoms on the light (i.e., Faraday rotation), here we start simple and consider the quantum evolution of the atoms in the presence of a classical light field. We start with the dipole Hamiltonian and then follow an *adiabatic elimination* procedure to remove the excited states, thus representing the dynamics in terms of the ground states alone. One goal of this chapter is to provide a more expansive treatment of the adiabatic elimination procedure than is typically given in the literature. However, the primary goal is to efficiently describe the nontrivial, state-dependent dispersion and absorption dynamics of the multilevel single atom system.

The resulting master equation consists of the polarizability Hamiltonian (including all of the hyperfine levels) and also the spontaneous emission terms, corresponding to weighted jumps between ground state sublevels. As expected, the derived spontaneous emission terms are state dependent for the multilevel atom and we end this chapter with an analysis of this dependence for common experimental configurations. This is crucial because many simplified models of spontaneous emission neglect the potentially important state dependence.

In the next chapter, we further simplify the polarizability Hamiltonian in terms of

intuitive atomic angular momentum operators and optical Stokes operators.

5.1 Master Equation Derivation

We begin by stating the relevant Hamiltonians, most importantly the dipole Hamiltonian describing the atom-light interaction, and then introduce the full master equation, which adds spontaneous emission from the multiple excited levels. From this point, we follow an adiabatic elimination procedure that gives the master equation in terms of only the ground states.

5.1.1 Hamiltonians

The Hamiltonian for the system, with a quantum atom and classical light, is given by

$$\hat{H} = \hat{H}_B + \hat{H}_g + \hat{H}_e + \hat{H}_i. \tag{5.1}$$

Larmor precession due to a magnetic field is given by

$$\hat{H}_B = \sum_{f,m_f} \frac{\mu_B g_f}{\hbar} \hat{\mathbf{f}} \cdot \mathbf{B}. \tag{5.2}$$

For the cesium $f = 4$ ground state the above prefactor gives a Larmor frequency of $\gamma_{B4} = \mu_B g_{f=4}/h = 350$ kHz/G [84]. We work with Hamiltonian and energy units of rad/s and divide by $\hbar$ unless otherwise noted. The energies of the different hyperfine levels are given by

$$\hat{H}_g = \sum_{f,m_f} \omega_f |f, m_f\rangle\langle f, m_f|, \tag{5.3}$$

$$\hat{H}_e = \sum_{f',m_{f'}} \omega_{f'} |f', m_{f'}\rangle\langle f', m_{f'}|, \tag{5.4}$$

and the dipole Hamiltonian describing the interaction the atom with the field (and the coupling of the ground and excited levels) is given by

$$\hat{H}_i = -\hat{\mathbf{d}} \cdot \mathbf{E}/\hbar. \tag{5.5}$$

The dipole operator and the classical field can be broken up into their rotating and counter-rotating components as

$$
\begin{aligned}
\hat{\mathbf{d}} &= \hat{P}_g \hat{\mathbf{d}} \hat{P}_e + \hat{P}_e \hat{\mathbf{d}} \hat{P}_g \\
&= \hat{\mathbf{d}}^- + \hat{\mathbf{d}}^+ \tag{5.6} \\
\mathbf{E} &= \sum_\sigma \mathcal{E} \epsilon_\sigma \mathbf{e}_{-\sigma} \exp[-i\omega t] + \text{h.c.} \\
&= \mathbf{E}^+ + \mathbf{E}^-. \tag{5.7}
\end{aligned}
$$

The rotating wave approximation (RWA) is best seen after going into the rotating frame but the simple result just amounts to leaving out the two sum terms rotating at optical frequencies and keeping the two difference terms, which rotate at the much smaller detunings. This leads to

$$
\begin{aligned}
\hat{H}_i &= -\frac{1}{\hbar} \hat{\mathbf{d}}^+ \cdot \mathbf{E}^- + \text{h.c.} \\
&= -\frac{1}{\hbar} \sum_{f,f',\sigma} \mathcal{E} \exp[i\omega t] \epsilon_{-\sigma} \hat{P}_g (\hat{\mathbf{d}} \cdot \mathbf{e}_{-\sigma}) \hat{P}_e + \text{h.c.} \\
&= -\frac{1}{\hbar} \sum_{f,f',\sigma} \sum_{m_f,m_{f'}} \mathcal{E} \exp[i\omega t] \epsilon_{-\sigma} \langle f||\hat{\mathbf{d}}||f'\rangle \langle f, m_f | f', m_{f'}; 1, \sigma\rangle |f, m_f\rangle \langle f', m_{f'}| + \text{h.c.} \\
&= -\frac{\Omega}{2} \exp[i\omega t] \sum_{f,f',\sigma} \epsilon_{-\sigma} \hat{A}_{f,f',\sigma} + \text{h.c.} \tag{5.8}
\end{aligned}
$$

Here we have defined the Rabi frequency as

$$
\Omega = 2\mathcal{E}\langle j = 1/2||\hat{\mathbf{d}}||j' = 3/2\rangle / \hbar \tag{5.9}
$$

where j is the spin-angular momentum plus the orbital angular momentum as discussed in chapter 3. We have also defined the jump operators as

$$
\hat{A}_{f,f',\sigma} = \sum_{m_f,m_{f'}} \langle f, m_f | f', m_{f'}; 1, \sigma\rangle \frac{\langle f||\hat{\mathbf{d}}||f'\rangle}{\langle j = 1/2||\hat{\mathbf{d}}||j' = 3/2\rangle} |f, m_f\rangle \langle f', m_{f'}|. \tag{5.10}
$$

See reference [84] for finding the ratio of dipole moments in terms of six-j and three-j symbols. Another indispensable reference for Clebsch-Gordan coefficients and quantum angular momentum is [95]. In Matlab, we can conveniently create the numerical jump

operators using

$$[\hat{A}_{f,f',-1}, \hat{A}_{f,f',0}, \hat{A}_{f,f',+1}] \;\; = \;\; \mathtt{murelf(f,f',j,j',i)} \tag{5.11}$$

with $j = 1/2$, $j' = 3/2$, and $i = 7/2$. Note that we have the useful identities

$$\sum_{f,\sigma} \hat{A}^{\dagger}_{f,f',\sigma} \hat{A}_{f,f',\sigma} \;\; = \;\; \hat{P}_{f'} \tag{5.12}$$

$$\sum_{f,f',\sigma} \hat{A}^{\dagger}_{f,f',\sigma} \hat{A}_{f,f',\sigma} \;\; = \;\; \hat{P}_e, \tag{5.13}$$

which are used below for simplifying expressions.

5.1.2 Full Master Equation

The unconditional master equation is given by

$$\dot{\hat{\rho}} \;\; = \;\; -i[\hat{H}, \hat{\rho}] + \Gamma \sum_{f,f',\sigma} \mathcal{D}[\hat{A}_{f,f',\sigma}]\hat{\rho} \tag{5.14}$$

$$= \;\; -i[\hat{H}, \hat{\rho}] + \Gamma \sum_{f,f',\sigma} \hat{A}_{f,f',\sigma} \hat{\rho} \hat{A}^{\dagger}_{f,f',\sigma} - \frac{\Gamma}{2}(\hat{P}_e \hat{\rho} + \hat{\rho} \hat{P}_e) \tag{5.15}$$

where we have used the superoperator definition

$$\mathcal{D}[\hat{c}]\hat{\rho} \;\; = \;\; \hat{c}\hat{\rho}\hat{c}^{\dagger} - (\hat{c}^{\dagger}\hat{c}\hat{\rho} + \hat{\rho}\hat{c}^{\dagger}\hat{c})/2 \tag{5.16}$$

along with the convenient identity of equation (5.13). Here we have simply asserted without derivation what the spontaneous emission decay terms are that correspond to the coupling of the atom with all of the three-dimensional free-space vacuum modes, following the treatment of [96]. We leave for future work how to derive a conditional master equation that would provide our description for the atomic state if we were to measure some fraction of the scattered light. Still, such a treatment should become the same as the case we consider here under the limit that the measured fraction of scattered light goes to zero.

5.1.3 Rotating Frame

Now our goal is to remove the explicit time dependence from the Hamiltonian, so we move into a frame rotating at the probe frequency ω. This is done by transforming all operators according to

$$\hat{X} \rightarrow \exp[i\hat{P}_e \omega t]\hat{X}\exp[-i\hat{P}_e \omega t]. \tag{5.17}$$

Notice that the only effect this has is to remove the time dependence from the Hamiltonian and simultaneously turn the excited energy levels into detunings via the term derived from the transformation of $\dot{\hat{\rho}}$ itself. The dynamics are still given by equation (5.15), but with the Hamiltonian now given by the following (without bothering to change notation)

$$
\begin{align}
\hat{H} &= \hat{H}_B + \hat{H}_g + \hat{H}_e + \hat{H}_i \tag{5.18}\\
\hat{H}_B &= \sum_{f,m_f} \frac{\mu_B g_f}{\hbar} \mathbf{B} \cdot \hat{\mathbf{f}} \tag{5.19}\\
\hat{H}_g &= -\Delta_3 \hat{P}_3 \tag{5.20}\\
\hat{H}_e &= -\sum_{f'} \Delta_{f'} \hat{P}_{f'} \tag{5.21}\\
\hat{H}_i &= -\frac{\Omega}{2} \sum_{f,f',\sigma} \epsilon_{-\sigma}\hat{A}_{f,f',\sigma} + \text{h.c.} \tag{5.22}
\end{align}
$$

(notice that we chose the $f = 4$ state to be the zero energy reference) with the detunings defined as

$$
\begin{align}
\Delta_{f'} &= \omega - (\omega_{f'} - \omega_4) \tag{5.23}\\
\Delta_3 &= \omega_4 - \omega_3 = 2\pi \cdot 9.192631770 \text{ GHz.} \tag{5.24}
\end{align}
$$

The last detuning is exact because this quantity currently defines the unit of time.

5.1.4 Adiabatic Elimination of Excited Levels

There are two reasons for not wanting to integrate the above master equation numerically. First, there are terms proportional to the detunings, which can be quite large, meaning that the numerical integration would have to be done with unreasonably small timesteps.

Second, the total Hilbert space is somewhat big with 48 levels as opposed to the 16 levels describing only the ground states.

Fortunately, with a far-off resonant probe beam, the procedure of adiabatic elimination [96, 97, 98, 52] allows us to eliminate the excited state levels in our description while retaining their effect on the physics. The approximation assumes that there is little population in the excited states, i.e., the saturation parameter

$$s = \frac{\Omega^2}{2(\Delta^2 + \Gamma^2/4)} \ll 1 \tag{5.25}$$

should be small for the detuning under consideration.

Many references exist that partially explain the process of adiabatic elimination, but most are incomplete, therefore we will attempt to be more forthcoming in our presentation. The following procedure is by no means rigorous, but at least every step is spelled out. We proceed as follows, ignoring the Hamiltonian $\hat{H}_B$ until the end. First, we decompose $\hat{\rho}$ into four parts

$$
\begin{align}
\hat{\rho} &= (\hat{P}_g + \hat{P}_e)\hat{\rho}(\hat{P}_g + \hat{P}_e) \tag{5.26} \\
&= \hat{P}_g\hat{\rho}\hat{P}_g + \hat{P}_g\hat{\rho}\hat{P}_e + \hat{P}_e\hat{\rho}\hat{P}_g + \hat{P}_e\hat{\rho}\hat{P}_e \tag{5.27} \\
&= \hat{\rho}_{gg} + \hat{\rho}_{ge} + \hat{\rho}_{eg} + \hat{\rho}_{ee}. \tag{5.28}
\end{align}
$$

Now we find $\dot{\hat{\rho}}_{ij}$ for each term, by judicious insertion of the identity $\hat{P}_e + \hat{P}_g$ into the above master equation

$$
\begin{align}
\dot{\hat{\rho}}_{gg} &= -i[\hat{H}_g, \hat{\rho}_{gg}] + i\hat{\rho}_{ge}\hat{H}_i\hat{P}_g - i\hat{P}_g\hat{H}_i\hat{\rho}_{eg} + \Gamma \sum_{f,f',\sigma} \hat{A}_{f,f',\sigma}\hat{\rho}_{ee}\hat{A}^\dagger_{f,f',\sigma} \tag{5.29} \\
\dot{\hat{\rho}}_{ge} &= i\hat{\rho}_{gg}\hat{H}_i\hat{P}_e + i\hat{\rho}_{ge}\hat{H}_e - i\hat{H}_g\hat{\rho}_{ge} - i\hat{P}_g\hat{H}_i\hat{\rho}_{ee} - \Gamma\hat{\rho}_{ge}/2 \tag{5.30} \\
\dot{\hat{\rho}}_{ee} &= -i\hat{P}_e\hat{H}_i\hat{\rho}_{ge} + i\hat{\rho}_{eg}\hat{H}_i\hat{P}_e - i[\hat{H}_e, \hat{\rho}_{ee}] - \Gamma\hat{\rho}_{ee} \tag{5.31}
\end{align}
$$

where $\dot{\hat{\rho}}_{ge} = (\dot{\hat{\rho}}_{eg})^\dagger$.

According to the weak-coupling approximation, $s \ll 1$, the order of magnitude of each of these parts is given by $\langle \hat{\rho}_{ee} \rangle \sim s$, $\langle \hat{\rho}_{ge} \rangle \sim \sqrt{s}$, and $\langle \hat{\rho}_{gg} \rangle \sim 1$. Using these approximations

in the decay terms of equations (5.29, 5.30, 5.31) (the last term of each equation) we get

$$\langle \dot{\hat{\rho}}_{gg} \rangle / \langle \hat{\rho}_{gg} \rangle \sim -\Gamma s \tag{5.32}$$

$$\langle \dot{\hat{\rho}}_{eg} \rangle / \langle \hat{\rho}_{eg} \rangle \sim -\Gamma/2 \tag{5.33}$$

$$\langle \dot{\hat{\rho}}_{ee} \rangle / \langle \hat{\rho}_{ee} \rangle \sim -\Gamma. \tag{5.34}$$

Thus, we see that the decay rate for the ground states $\hat{\rho}_{gg}$ is much smaller than the decay rate for the cross $\hat{\rho}_{eg}$ and excited $\hat{\rho}_{ee}$ components by a factor of s. In other words, the cross and excited components reach equilibrium on a timescale much faster than the ground state component. As a result, we can set $\dot{\hat{\rho}}_{eg} = 0$ and $\dot{\hat{\rho}}_{ee} = 0$, solve for $\hat{\rho}_{ee}$ in terms of $\hat{\rho}_{eg}$, solve $\hat{\rho}_{eg}$ in terms of $\hat{\rho}_{gg}$, and substitute these results into equation (5.29) to get the evolution of $\hat{\rho}_{gg}$ in terms of itself. We say that $\hat{\rho}_{ee}$ "adiabatically follows" $\hat{\rho}_{eg}$, and in turn $\hat{\rho}_{eg}$ "adiabatically follows" $\hat{\rho}_{gg}$. The entire process is thus referred to as "adiabatic elimination" and represents the core idea of this chapter.

There is one more step necessary to make the above procedure work. First, note that the Hamiltonians are approximately of orders $\hat{H}_g \sim \hat{H}_e \sim \Delta$ and $\hat{H}_i \sim \sqrt{s}\Delta$. So all of the Hamiltonian (nondecay) terms in equation (5.30) are of order $\sqrt{s}\Delta$ except for the $\hat{\rho}_{ee}$ term, which is of order $s\sqrt{s}\Delta$. Consequently, the latter can be ignored under the limit $s \ll 1$.

We now follow this adiabatic elimination procedure in detail to produce the desired ground-state-only master equation. As mentioned above, we first set $\dot{\hat{\rho}}_{ee} = 0$ and solve for $\hat{\rho}_{ee}$ in terms of $\hat{\rho}_{eg}$ and $\hat{\rho}_{ge}$. When doing this we run into an equation that looks like

$$\hat{\mathcal{L}}\hat{\rho}_{ee} \equiv -\Gamma\hat{\rho}_{ee} - i[\hat{H}_e, \hat{\rho}_{ee}] = i\hat{P}_e\hat{H}_i\hat{\rho}_{ge} - i\hat{\rho}_{eg}\hat{H}_i\hat{P}_e \tag{5.35}$$

$$\hat{\rho}_{ee} = \hat{\mathcal{L}}^{-1}(i\hat{P}_e\hat{H}_i\hat{\rho}_{ge} - i\hat{\rho}_{eg}\hat{H}_i\hat{P}_e). \tag{5.36}$$

Thus we need to somehow invert the superoperator $\hat{\mathcal{L}}$. References [98, 52] give us a clue how to do this. Using the definition of the superoperator inversion (on arbitrary $\hat{X}$) we

have

$$\hat{\mathcal{L}}^{-1}\hat{X} \;=\; -\int_0^\infty \exp[\hat{\mathcal{L}}t]\hat{X}\,dt \tag{5.37}$$

$$=\; -\int_0^\infty \exp[-\Gamma t]\exp[-i\hat{H}_e t]\hat{X}\exp[i\hat{H}_e t]dt \tag{5.38}$$

$$=\; -\int_0^\infty \exp[-\Gamma t]\left(1+\sum_{f'}(\exp[i\Delta_{f'}t]-1)\hat{P}_{f'}\right)\hat{X}$$

$$\times\left(1+\sum_{f'}(\exp[-i\Delta_{f'}t]-1)\hat{P}_{f'}\right)dt \tag{5.39}$$

$$=\; -\frac{\hat{X}}{\Gamma}-\sum_{f'}\left(\left(\frac{1}{\Gamma+i\Delta_{f'}}-\frac{1}{\Gamma}\right)\hat{X}\hat{P}_{f'}+h.c.\right)$$

$$-\sum_{f_1',f_2'}\left(\frac{1}{i(\Delta_{f_1'}-\Delta_{f_2'})+\Gamma}+\frac{1}{i\Delta_{f_1'}-\Gamma}+\frac{1}{-i\Delta_{f_2'}-\Gamma}+\frac{1}{\Gamma}\right)P_{f_1'}\hat{X}\hat{P}_{f_2'} \tag{5.40}$$

where we have used the definition of $\hat{H}_e$ and the fact that $\hat{P}^2=\hat{P}$ for the projectors. Actually, what we care about in the end is only the simplification of equation (5.29) and we do not need to worry about the full inversion of the superoperator. We can rewrite the only term in equation (5.29) where $\hat{\rho}_{ee}$ appears as

$$\Gamma\sum_{f,f',\sigma}\hat{A}_{f,f',\sigma}\hat{\rho}_{ee}\hat{A}_{f,f',\sigma}^\dagger \;=\; \Gamma\sum_{f,f',\sigma}\hat{A}_{f,f',\sigma}\left(\hat{P}_{f'}\hat{\rho}_{ee}\hat{P}_{f'}\right)\hat{A}_{f,f',\sigma}^\dagger$$

$$=\; \Gamma\sum_{f,f',\sigma}\hat{A}_{f,f',\sigma}\left(\hat{P}_{f'}\hat{\mathcal{L}}^{-1}(i\hat{P}_e\hat{H}_i\hat{\rho}_{ge}-i\hat{\rho}_{eg}\hat{H}_i\hat{P}_e)\hat{P}_{f'}\right)\hat{A}_{f,f',\sigma}^\dagger$$

$$=\; \Gamma\sum_{f,f',\sigma}\hat{A}_{f,f',\sigma}\left(\hat{P}_{f'}\left(-\frac{1}{\Gamma}\right)(i\hat{P}_e\hat{H}_i\hat{\rho}_{ge}-i\hat{\rho}_{eg}\hat{H}_i\hat{P}_e)\hat{P}_{f'}\right)\hat{A}_{f,f',\sigma}^\dagger$$

$$=\; -\sum_{f,f',\sigma}\hat{A}_{f,f',\sigma}\left(\hat{P}_{f'}(i\hat{P}_e\hat{H}_i\hat{\rho}_{ge}-i\hat{\rho}_{eg}\hat{H}_i\hat{P}_e)\hat{P}_{f'}\right)\hat{A}_{f,f',\sigma}^\dagger \tag{5.41}$$

where we have used equation (5.36) in the second line and

$$P_{f'}(\hat{\mathcal{L}}^{-1}\hat{X})P_{f'}=P_{f'}\left(-\frac{\hat{X}}{\Gamma}\right)P_{f'} \tag{5.42}$$

in the third line, which is evident from equation (5.40). In other words, we do not care about the coherences between different excited states and we are only concerned with the block diagonal components.

Next, as justified above, we set the $\hat{\rho}_{ee}$ term in equation (5.30) to zero, set $\dot{\hat{\rho}}_{ge}$ to zero, and solve for $\hat{\rho}_{ge}$ in terms of $\hat{\rho}_{gg}$. We again use the inversion definition for a new superoperator, but this time we do not have to use any projectors to simplify the solution. This gives

$$\hat{\rho}_{ge} = i\sum_{f'}\left[\frac{\hat{P}_4}{\Gamma/2 + i\Delta_{f'}} + \frac{\hat{P}_3}{\Gamma/2 + i(\Delta_{f'} - \Delta_3)}\right]\hat{\rho}_{gg}\hat{H}_i\hat{P}_{f'} \qquad (5.43)$$

and we say that $\hat{\rho}_{ge}$ adiabatically follows $\hat{\rho}_{gg}$.

Finally, we use all of these results to solve for $\dot{\hat{\rho}}_{gg}$ in terms of itself in equation (5.29). This results in what we have been seeking, a self-contained master equation involving only the ground states:

$$\begin{aligned}
\dot{\hat{\rho}}_{gg} = {}& -i[\hat{H}_g + \hat{H}_B, \hat{\rho}_{gg}] \\
& + i\hat{P}_4\hat{\rho}_{gg}\sum_{f'}\frac{\hat{H}_i\hat{P}_{f'}\hat{H}_i}{\Delta_{f'} - i\Gamma/2} + \text{h.c.} \\
& + i\hat{P}_3\hat{\rho}_{gg}\sum_{f'}\frac{\hat{H}_i\hat{P}_{f'}\hat{H}_i}{(\Delta_{f'} - \Delta_3) - i\Gamma/2} + \text{h.c.} \\
& + \sum_{f,f',\sigma}\hat{A}_{f,f',\sigma}\hat{H}_i\frac{\hat{P}_4}{i\Delta_{f'} + \Gamma/2}\hat{\rho}_{gg}\hat{H}_i\hat{A}^\dagger_{f,f',\sigma} + \text{h.c.} \\
& + \sum_{f,f',\sigma}\hat{A}_{f,f',\sigma}\hat{H}_i\frac{\hat{P}_3}{i(\Delta_{f'} - \Delta_3) + \Gamma/2}\hat{\rho}_{gg}\hat{H}_i\hat{A}^\dagger_{f,f',\sigma} + \text{h.c.} \qquad (5.44)
\end{aligned}$$

The key result here is that we have kept the state-dependent spontaneous emission terms (in the last two lines). Other treatments have found only the effective Hamiltonian for $f = 4$ (represented here in the second line) and treated spontaneous emission in a nonrigorous way [25, 29, 26]. Note that, whereas the $\hat{A}$ operators were the jump operators on the full Hilbert space connecting excited states to ground states, now the effective jump operators are of the form $\hat{A}\hat{H}_i$ and serve to mix the ground states amongst themselves. We also see that, despite our approximations, the norm of the state is preserved, $\text{Tr}[\dot{\hat{\rho}}_{gg}] = 0$. This can be seen by using both the cyclic property of the trace and the jump-operator identity of equation (5.12). Finally, notice that in the limit $\Delta_{5'} \gg \Delta_3$ (i.e., the detuning is much larger than the hyperfine ground state splitting) the probe does not distinguish ground states and we can group terms to get the same results as reference [96] for $j = 1/2$ to $j' = 3/2$ transition (for stationary atoms) as we should.

The second and third lines of the above master equation represent the *polarizability Hamiltonian* and describe the coherent, dispersive interaction of the beam with the atoms. In the next chapter, we start from this point to analyze the evolution of the optical polarization due to the atomic spin-state. In the remaining part of this chapter, however, we continue to analyze the state dependence of the incoherent spontaneous emission terms.

5.2 Applications of the Master Equation

The master equation of equation (5.44) is useful for many purposes. Quite generally, it compactly describes the coherent and incoherent evolution of a multilevel atomic state under arbitrary optical manipulation (in the dipole limit and low saturation parameter limits). The coherent polarizability Hamiltonian component and the state-dependent spontaneous emission are considered in detail throughout this thesis. It can also be used to investigate optical pumping dynamics for a small saturation parameter. Finally, for spatially dependent fields, it can be used to succinctly describe optical forces on a real multilevel alkali atom.

Equation (5.44) is relatively easy to integrate because the Hilbert space is smaller without the excited states and the timescales involved are relatively large. As long as we keep the Larmor precession rate and the optical pumping rate ($\propto \Gamma s$) small, the numerical timesteps can be made much larger than what would be necessary in simulating the full master equation. We ignore the Hamiltonian $\hat{H}_g$ as we assume the $f = 3$ level does not couple coherently to the $f = 4$ level and only serves as a population "sink," or atom loss channel.

The creation of a numerical integration routine is facilitated by the Matlab routines within the *Quantum Optics Toolbox* [99], which provides angular momentum operator functions (e.g., `murelf`) and efficient integration routines with sparse matrices. In this section, we demonstrate a few examples using this code. Each simulation shown takes only a few seconds on a reasonably equipped computer.

In the case of an optically thin ensemble, but still with many atoms, the simulated value of $\langle \hat{f}_z \rangle$ represents approximately what the photocurrent would look like without noise because the Faraday rotation would be proportional to the z-component of the collective spin as seen in the next chapter. For an optically thick ensemble, where the probe polarization changes significantly within the cloud, then the situation is more complicated and we refer the reader to reference [26].
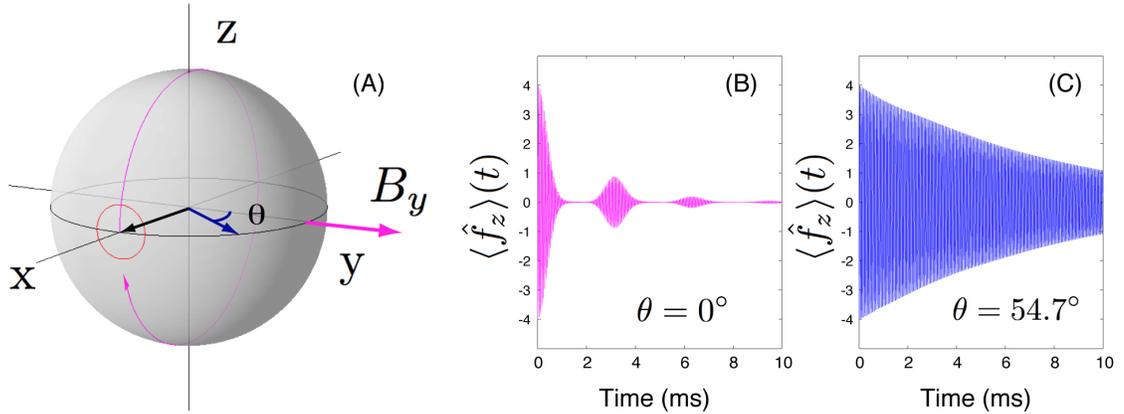
Figure 5.1: (A) Schematic showing the relative orientation of the spin, magnetic field, and optical polarization. (B) Simulation of $\hat{f}_z$ for an initial state along the x-axis, magnetic field along the y-axis, probe propagation along the z-axis, probe polarization along the y-axis. (C) Simulation for an initial state along x-axis, magnetic field along the y-axis, probe propagation along the z-axis, probe polarization 54.7 degrees from the y-axis. The probe detuning used is 500 MHz.

When considering the following simulation configurations, keep in mind that we continue to consider only the evolution of a single atom, but that evolution consists of many parts, including the polarizability Hamiltonian (which can cause coherent "dephasing") and the incoherent spontaneous emission. We do not consider the collective evolution of the spin-state until chapter 7.

5.2.1 Decay with Larmor Precession

First we consider Larmor precession dynamics and produce results similar to those from [25]. Figure 5.1A shows the relative orientation of the spin, magnetic field, and optical polarization directions. Figure 5.1B depicts a situation with an initial state polarized along the x-axis, a DC magnetic field along the y-axis, and a probe beam propagating along the z-axis with linear optical polarization along the y-axis. The decay and revivals, or echoes, are from the tensor Hamiltonian part of the master equation, which is shown to be nonlinear in the angular momentum operators in the next chapter. Without spontaneous emission the revival would be complete. Figure 5.1C depicts the same situation, but with the probe polarization aligned along a "magic" angle (54.7 degrees from the y-axis) where the nonlinear Hamiltonian term disappears and a more exponential decay is observed. The

rotating frame analysis used to derive this magic angle effect is presented in section 6.5.

5.2.2 Decay without Larmor Precession

Now we consider the evolution under configurations where either there is no magnetic field or the field and spin-state remain parallel, hence dynamics at the Larmor frequency are not observed. We consider two cases, in one the linear optical polarization is parallel to the initial spin-state, and in the other it is perpendicular.

The simulated evolution under the four possible configurations (two optical polarizations, with and without field) is presented in figure 5.2. For parallel polarizations, the field and no field cases are not any different because the evolution due to the field does not break the symmetry of the no field case. For perpendicular polarizations, however, there is a large difference between the field and no field case. Without a field the length of the vector dephases quickly (blue curve in figure 5.2A) and displays echoes similar to the effects demonstrated in the last section. Not surprisingly the perpendicular uncertainties (in the yz-plane) also behave dramatically and asymmetrically in figure 5.2B–C although there is no single-atom spin-squeezing in this configuration. In contrast, the parallel spin case shows simple exponential decay and the uncertainty variances initially grow in time (this is purely a result of the decay and spontaneous emission, *not* the tensor Hamiltonian). When a field is applied to the spins in the perpendicular field case, the dephasing is eliminated and the resulting spontaneous emission time becomes longer than the parallel field case (as demonstrated below).

figure 5.2B displays the loss of population to the $f = 3$ ground state where it is clear that the spontaneous emission rate is state (or optical polarization) dependent. A more extreme example of the state dependent pumping could be demonstrated by putting the atoms in a state that was dark to the probe light, e.g., the state $|4, 4\rangle$ with σ_+ probe light.

Now we go about calculating the expected spontaneous emission decay rates. In analyzing the decay, we can consider many quantities, including the moment $\hat{f}_z$ or the population among all of the $f = 4$ sublevels. Here we consider the damping of $\hat{f}_z$. Assume we start with the atom aligned along the $+z$ direction and consider the quantity $d\langle\hat{f}_z\rangle/dt$ at time zero calculated from equation (5.44). If we convert this derivative into a scattering rate by
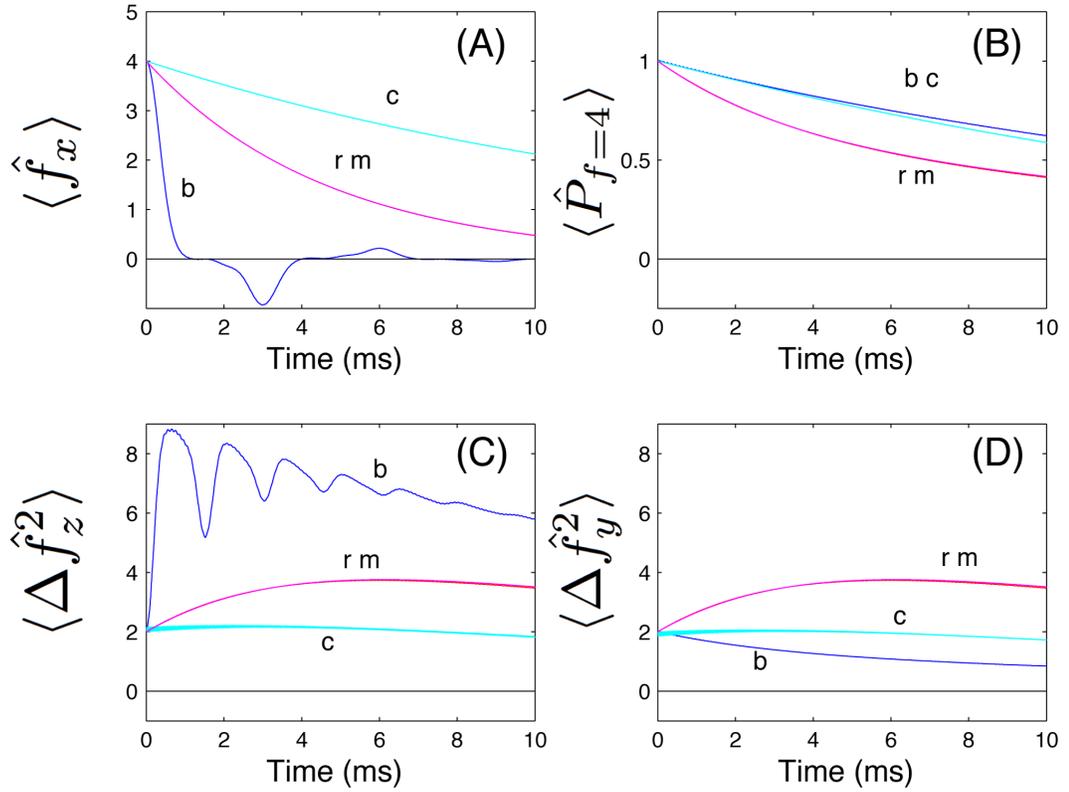
Figure 5.2: As described in the text, simulations of decay and decoherence via linearly polarized probe light for an initial coherent spin-state. The blue (b) and cyan (c) curves represent optical polarizations (along y) perpendicular to the initial spin direction (along x). The red (r) and magenta (m) curves represent parallel optical polarization (along x). The blue and red curves are without any field, and the cyan and magenta curves are with a field along the spin direction x. The moments $\langle \hat{f}_z \rangle$ and $\langle \hat{f}_y \rangle$ are zero. The probe detuning used is 500 MHz.

assuming initial exponential decay

$$\frac{d\langle \hat{f}_z \rangle}{dt}\bigg|_{t=0} = -\frac{\langle \hat{f}_z \rangle}{\tau_{sc}}\bigg|_{t=0} \tag{5.45}$$

we get the decay rate

$$
\begin{aligned}
\tau_{sc}^{-1} &= \frac{I\sigma_0}{2\hbar\omega}\left[\frac{\Gamma^2/4}{\Gamma^2/4 + \Delta_{45}^2}\left(|\epsilon_0|^2 0.16 + |\epsilon_-|^2 0.04 + |\epsilon_+|^2 0\right)\right.\\
&\quad + \frac{\Gamma^2/4}{\Gamma^2/4 + \Delta_{44}^2}\left(|\epsilon_0|^2 0.83 + |\epsilon_-|^2 0.41\right)\\
&\quad + \left.\frac{\Gamma^2/4}{\Gamma^2/4 + \Delta_{43}^2}\left(|\epsilon_-|^2 0.60\right)\right]. \tag{5.46}
\end{aligned}
$$

Clearly, we see that this expression respects some expected selection rules by, for example, not allowing positive circular polarization to produce any decay. This is because such light either does not couple to a level (4–3' or 4–4') or it couples to a cycling transition that only returns the population to where it started (4–5'). For an optical polarization parallel to the spins we use: $|\epsilon_+|^2 = |\epsilon_+|^2 = 0$ and $|\epsilon_0|^2 = 1$. Under the simulation, we see that the Hamiltonian terms alone do not cause decay so the calculated exponential from above works with or without a holding field. For an optical polarization perpendicular to the spins we use: $|\epsilon_+|^2 = |\epsilon_+|^2 = 1/2$ and $|\epsilon_0|^2 = 0$. Under the simulation, we see that there is substantial Hamiltonian "dephasing" (with echoes) of the moment, but applying a holding field makes equation (5.46) a valid expression at all times.

When the field is applied in figure 5.2, the coherent dephasing is removed and the resulting decay curves are exponential with time constants given by equation (5.46). Even though the above expression was only derived for the first timestep, it is seen by simulation to remain valid at all times with such a holding field. Both equation (5.46) and the simulations of figure 5.2 show the feature that spontaneous emission occurs faster for parallel polarizations than perpendicular, which can also be expressed in terms of selection rules and Clebsch-Gordan coefficients.

Again, we have said nothing about the consequences of this decay or non-QND Hamiltonian evolution on the collective spin-squeezed state preparation and we save this discussion for chapter 7 where both single-atom and collective effects are discussed. In chapter 8, we also refer back to equation (5.46) to provide an estimate of the decay time in relation to the

time that it takes to spin-squeeze the atoms via the measurement. The potential for a less ad hoc combination of spin-squeezing in the presence of decay is discussed in both chapter 15 and appendix E.

Chapter 6

The Irreducible Representation of the Polarizability Hamiltonian

In the last chapter, we demonstrated that adiabatic elimination of the dipole Hamiltonian describing the atom-light interaction produces the polarizability form of the Hamiltonian in terms of only the ground states. In this chapter, we rewrite the polarizability Hamiltonian in terms of the angular momentum operators for the ground state, rather than the dipole operators, and the Stokes operators describing the polarization of the probe beam. There is an irreducible decomposition of the polarizability tensor into purely scalar, vector, and tensor terms, and the physics of each corresponding Hamiltonian is discussed. The scalar term determines the state-independent light shift, which is important in, for example, the trapping of atoms with an optical lattice. The vector term leads to the QND Hamiltonian we desire and provides the basis for much of this work. The tensor term arises because the spin has more than two sublevels and is important at detunings comparable to the hyperfine splittings. After introducing these terms, and investigating their spectral properties, we describe the semiclassical evolution of the probe polarization due to the atomic state via the Hamiltonian components. We then consider the evolution of the spin-state due to the tensor term in special, experimentally relevant, configurations. While the tensor term complicates the ideal description of the quantum measurement, it also proves technically useful because of its nonlinearity. We end by showing how the tensor term provides us with techniques for aligning the optical and atomic polarizations, as well as measuring the degree of pumping of the atomic state.

Much of this chapter is adapted from [29], however many sections are new, including section 6.3, section 6.5, section 6.7, and section 6.8.

6.1 Deriving the Irreducible Representation

6.1.1 Hamiltonian Approximation

We begin again with the single-particle dipole Hamiltonian $H = -\hat{\mathbf{d}} \cdot \hat{\mathbf{E}}$. The dipole operator $\hat{\mathbf{d}} = e\hat{\mathbf{r}}_e$ can be split into its raising and lowering components

$$\hat{\mathbf{d}} = \hat{\mathbf{d}}^{(-)} + \hat{\mathbf{d}}^{(+)} \tag{6.1}$$

$$\hat{\mathbf{d}}^{(-)} = \sum_{f,f'} \hat{P}_f \hat{\mathbf{d}} \hat{P}_{f'} \tag{6.2}$$

$$\hat{\mathbf{d}}^{(+)} = \sum_{f,f'} \hat{P}_{f'} \hat{\mathbf{d}} \hat{P}_f \tag{6.3}$$

and the electric field operator can be split into rotating and counter-rotating terms

$$\hat{\mathbf{E}} = \hat{\mathbf{E}}^{(-)} + \hat{\mathbf{E}}^{(+)} \tag{6.4}$$

$$\hat{\mathbf{E}}^{(-)} = \sqrt{\hbar g} \left[\hat{a}_-^\dagger \vec{\mathbf{e}}_-^* + \hat{a}_+^\dagger \vec{\mathbf{e}}_+^* \right] \tag{6.5}$$

$$\hat{\mathbf{E}}^{(+)} = \sqrt{\hbar g} \left[\hat{a}_- \vec{\mathbf{e}}_- + \hat{a}_+ \vec{\mathbf{e}}_+ \right]. \tag{6.6}$$

Assuming that the beam interacts weakly with the atom (i.e., the saturation parameter is small), we can use the rotating wave approximation and one of many available perturbation expansion techniques to express this Hamiltonian only in terms of the atomic ground states. In the last chapter, an adiabatic elimination procedure resulted in equation (5.44), which contains both coherent dynamics (a polarizability Hamiltonian) and incoherent dynamics (spontaneous emission). The coherent polarizability Hamiltonian component can be expressed as [65, 100, 31],

$$\hat{H} = \sum_{f,f'} \hat{\mathbf{E}}^{(-)} \cdot \frac{\hat{\boldsymbol{\alpha}}_{f,f'}}{\hbar \Delta_{f,f'}} \cdot \hat{\mathbf{E}}^{(+)} \tag{6.7}$$

where we have assumed that $\Delta_{f,f'} \gg \Gamma$ in the denominator. In this far detuned limit (where the saturation parameter is also small), the beam interacts dispersively with the atoms while absorption and reemission via spontaneous emission is temporarily ignored. The atomic polarizability between a particular ground state (f) and excited state (f') is

defined as

$$\hat{\boldsymbol{\alpha}}_{f,f'} \;=\; \hat{P}_f \hat{\mathbf{d}} \hat{P}_{f'} \hat{\mathbf{d}}^\dagger \hat{P}_f \tag{6.8}$$

$$= \sum_m \sum_{m'} \sum_{m''} |f, m''\rangle\langle f, m''|\hat{\mathbf{d}}|f', m'\rangle\langle f', m'|\hat{\mathbf{d}}^\dagger|f, m\rangle\langle f, m|. \tag{6.9}$$

This expanded expression involves dipole operator matrix elements of the form, $\langle f', m'|\hat{d}_q|f, m\rangle$ where $|f, m\rangle$ is a Zeeman sublevel in the ground state hyperfine manifold, $|f', m'\rangle$ is a virtual state in the excited hyperfine manifold, and $q = 0, \pm 1$ labels the helicity of the electromagnetic field.

This Hamiltonian has a satisfying physical interpretation as a scattering interaction: the atom is first brought from its ground state to a virtual excited state via the raising operator, $\hat{\mathbf{d}}^\dagger$, by annihilating a photon from the probe field through $\hat{\mathbf{E}}^{(+)}$. Then, the temporarily excited atom returns to a (potentially different) ground state by emitting a photon into a (potentially different) scattered probe mode via $\hat{\mathbf{d}}$ and $\hat{\mathbf{E}}^{(-)}$.

The central operator in the scattering Hamiltonian,

$$\hat{\boldsymbol{\alpha}}_{f,f'} = \hat{P}_f \hat{\mathbf{d}} \hat{P}_{f'} \hat{\mathbf{d}}^\dagger \hat{P}_f, \tag{6.10}$$

commonly called the *atomic polarizability tensor*, is a dyad involving vector operators [101, 32]. Thus $\hat{\boldsymbol{\alpha}}_{f,f'}$ is a rank-2 spherical tensor that can be decomposed into irreducible components,

$$\hat{\boldsymbol{\alpha}}_{f,f'} = \hat{\boldsymbol{\alpha}}_{f,f'}^{(0)} + \hat{\boldsymbol{\alpha}}_{f,f'}^{(1)} + \hat{\boldsymbol{\alpha}}_{f,f'}^{(2)}. \tag{6.11}$$

The scattering Hamiltonian similarly decomposes into irreducible spherical tensor operators,

$$\hat{H} = \hat{H}^{(0)} + \hat{H}^{(1)} + \hat{H}^{(2)} \tag{6.12}$$

where

$$\hat{H}^{(j)} = \sum_{f,f'} \hat{\mathbf{E}}^{(-)} \cdot \frac{\hat{\boldsymbol{\alpha}}_{f,f'}^{(j)}}{\hbar \Delta_{f,f'}} \cdot \hat{\mathbf{E}}^{(+)}. \tag{6.13}$$

The $\hat{H}^{(0)}$ is a scalar contribution, $\hat{H}^{(1)}$ transforms as a vector, and $\hat{H}^{(2)}$ transforms as a rank-2 symmetric tensor in the group representation theory of $SO(3)$. Were the atomic system composed of spin-1/2 particles, it would be possible to neglect the rank-2 Hamiltonian [100]

(as will become explicit), however, we cannot do so for higher spin alkali atoms [25, 26, 102].

The above notation is complete, but we temporarily work with only one particular f, f' combination and remove the subscripts with the simplifying notation change

$$\hat{P}_f \hat{\mathbf{d}}^{(-)} \hat{P}_{f'} \to \hat{\mathbf{d}} \tag{6.14}$$

$$\hat{P}_{f'} \hat{\mathbf{d}}^{(+)} \hat{P}_f \to \hat{\mathbf{d}}^\dagger \tag{6.15}$$

$$\hat{\boldsymbol{\alpha}}_{f,f'} \to \hat{\boldsymbol{\alpha}}. \tag{6.16}$$

However, when the complete Hamiltonian is considered, the summation over all possible f, f' combinations is reestablished.

6.1.2   Matrix Element Decomposition

In order to work with the above expressions, it is advantageous to simplify the dipole matrix elements as far as possible. By employing the Wigner-Eckart theorem, the angular dependence of the matrix element, $\langle f', m' | \hat{\mathbf{d}} | f, m \rangle$ can be factored into the product of a Clebsch-Gordan coefficient and a reduced matrix element,

$$\langle f, m | \hat{d}_q | f', m' \rangle = \langle f, m | 1, q; f', m - q \rangle \langle f || \hat{\mathbf{d}} || f' \rangle. \tag{6.17}$$

Since the dipole operator acts only on electronic degrees of freedom, it is further possible to factor out the nuclear spin degrees of freedom via the explicit coupling,

$$\langle f || \hat{\mathbf{d}} || f' \rangle \;\; = \;\; (-1)^{f'+j+i+1} \sqrt{(2f'+1)(2j+1)} \begin{Bmatrix} 1 & j & j' \\ i & f' & f \end{Bmatrix} \langle j || \hat{\mathbf{d}}_e || j' \rangle \tag{6.18}$$

where i is the nuclear spin quantum number, j and j' are the ground and excited state fine structure quantum numbers, and $\hat{\mathbf{d}}_e$ is the dipole operator with respect to the electronic degrees of freedom.

6.1.3 Tensor Decomposition

From equation (17.89) of reference [103] we see that we can form an irreducible tensor, $\hat{Z}_m^{(j)}$, from a linear combination of tensor operators $\hat{U}_q^{(\kappa)}$ and $\hat{V}_{q'}^{(\kappa')}$ via the definition

$$\hat{Z}_m^{(j)} = \sum_{q,q'} \hat{U}_q^{(\kappa)} \hat{V}_{q'}^{(\kappa')} \langle \kappa, q; \kappa', q'|j, m\rangle \tag{6.19}$$

where $\langle \kappa, q; \kappa', q'|j, m\rangle$ are Clebsch-Gordan coefficients. This expression can then be inverted using

$$\hat{U}_q^{(\kappa)} \hat{V}_{q'}^{(\kappa')} = \sum_{j,m} \hat{Z}_m^{(j)} \langle \kappa, q; \kappa', q'|j, m\rangle. \tag{6.20}$$

We now specialize to the case where $\hat{\mathbf{U}} = \hat{\mathbf{d}}$ and $\hat{\mathbf{V}} = \hat{\mathbf{d}}^\dagger$. For this particular case, use the notation $\hat{Z}_m^{(j)} = \hat{T}_m^{(j)}$. Because we are creating a dyad (with two vectors), we have $\kappa = \kappa' = 1$. Inserting these above gives the definition

$$\hat{T}_m^{(j)} = \sum_{q,q'} \hat{d}_q \hat{d}_{q'}^\dagger \langle 1, q; 1, q'|j, m\rangle \tag{6.21}$$

and the inverse

$$\hat{d}_q \hat{d}_{q'}^\dagger = \sum_{j,m} \hat{T}_m^{(j)} \langle 1, q; 1, q'|j, m\rangle. \tag{6.22}$$

We can use this latter expression to write the polarizability as

$$\hat{\boldsymbol{\alpha}} = \hat{\mathbf{d}}\hat{\mathbf{d}}^\dagger \tag{6.23}$$

$$= \sum_{q,q'} \vec{\mathbf{e}}_q^* \vec{\mathbf{e}}_{q'}^* \hat{d}_q \hat{d}_{q'}^\dagger \tag{6.24}$$

$$= \sum_{j,m} \sum_{q,q'} \vec{\mathbf{e}}_q^* \vec{\mathbf{e}}_{q'}^* \hat{T}_m^{(j)} \langle 1, q; 1, q'|j, m\rangle \tag{6.25}$$

$$= \hat{\boldsymbol{\alpha}}^{(0)} \oplus \hat{\boldsymbol{\alpha}}^{(1)} \oplus \hat{\boldsymbol{\alpha}}^{(2)} \tag{6.26}$$

where

$$\hat{\boldsymbol{\alpha}}^{(j)} = \sum_{m=-j}^{j} \hat{T}_m^{(j)} \sum_{q,q'} \vec{\mathbf{e}}_q^* \vec{\mathbf{e}}_{q'}^* \langle 1, q; 1, q'|j, m\rangle. \tag{6.27}$$

Filling in these Clebsch-Gordan coefficients explicitly, we get

$$
\hat{\boldsymbol{\alpha}}^{(0)} = \hat{T}_0^{(0)}\left[-\frac{1}{\sqrt{3}}\vec{\mathbf{e}}_0^*\vec{\mathbf{e}}_0^* + \frac{1}{\sqrt{3}}\vec{\mathbf{e}}_+^*\vec{\mathbf{e}}_-^* + \frac{1}{\sqrt{3}}\vec{\mathbf{e}}_-^*\vec{\mathbf{e}}_+^*\right] \tag{6.28}
$$

$$
\begin{aligned}
\hat{\boldsymbol{\alpha}}^{(1)} = {}& \hat{T}_0^{(1)}\left[\frac{1}{\sqrt{2}}\vec{\mathbf{e}}_+^*\vec{\mathbf{e}}_-^* - \frac{1}{\sqrt{2}}\vec{\mathbf{e}}_-^*\vec{\mathbf{e}}_+^*\right] \\
&+\hat{T}_{+1}^{(1)}\left[-\frac{1}{\sqrt{2}}\vec{\mathbf{e}}_0^*\vec{\mathbf{e}}_+^* + \frac{1}{\sqrt{2}}\vec{\mathbf{e}}_+^*\vec{\mathbf{e}}_0^*\right] \\
&+\hat{T}_{-1}^{(1)}\left[\frac{1}{\sqrt{2}}\vec{\mathbf{e}}_0^*\vec{\mathbf{e}}_-^* - \frac{1}{\sqrt{2}}\vec{\mathbf{e}}_-^*\vec{\mathbf{e}}_0^*\right]
\end{aligned} \tag{6.29}
$$

$$
\begin{aligned}
\hat{\boldsymbol{\alpha}}^{(2)} = {}& \hat{T}_0^{(2)}\left[\frac{2}{\sqrt{6}}\vec{\mathbf{e}}_0^*\vec{\mathbf{e}}_0^* + \frac{1}{\sqrt{6}}\vec{\mathbf{e}}_+^*\vec{\mathbf{e}}_-^* + \frac{1}{\sqrt{6}}\vec{\mathbf{e}}_-^*\vec{\mathbf{e}}_+^*\right] \\
&+\hat{T}_{+1}^{(2)}\left[\frac{1}{\sqrt{2}}\vec{\mathbf{e}}_0^*\vec{\mathbf{e}}_+^* + \frac{1}{\sqrt{2}}\vec{\mathbf{e}}_+^*\vec{\mathbf{e}}_0^*\right] \\
&+\hat{T}_{-1}^{(2)}\left[\frac{1}{\sqrt{2}}\vec{\mathbf{e}}_0^*\vec{\mathbf{e}}_-^* + \frac{1}{\sqrt{2}}\vec{\mathbf{e}}_-^*\vec{\mathbf{e}}_0^*\right] \\
&+\hat{T}_{+2}^{(2)}\left[\vec{\mathbf{e}}_+^*\vec{\mathbf{e}}_+^*\right] \\
&+\hat{T}_{-2}^{(2)}\left[\vec{\mathbf{e}}_-^*\vec{\mathbf{e}}_-^*\right].
\end{aligned} \tag{6.30}
$$

Furthermore, using the definition of $\hat{T}_m^{(j)}$ and filling in the Clebsch-Gordan coefficients explicitly, we get

$$
\begin{aligned}
\hat{T}_0^{(0)} &= -\frac{1}{\sqrt{3}}\left(\hat{d}_0\hat{d}_0^\dagger - \hat{d}_+\hat{d}_-^\dagger - \hat{d}_-\hat{d}_+^\dagger\right) \tag{6.31}\\
\hat{T}_0^{(1)} &= \frac{1}{\sqrt{2}}\left(\hat{d}_+\hat{d}_-^\dagger - \hat{d}_-\hat{d}_+^\dagger\right) \\
\hat{T}_{+1}^{(1)} &= \frac{1}{\sqrt{2}}\left(-\hat{d}_0\hat{d}_+^\dagger + \hat{d}_+\hat{d}_0^\dagger\right) \\
\hat{T}_{-1}^{(1)} &= \frac{1}{\sqrt{2}}\left(\hat{d}_0\hat{d}_-^\dagger - \hat{d}_-\hat{d}_0^\dagger\right) \\
\hat{T}_0^{(2)} &= \frac{1}{\sqrt{6}}\left(\hat{d}_+\hat{d}_-^\dagger + 2\hat{d}_0\hat{d}_0^\dagger + \hat{d}_-\hat{d}_+^\dagger\right) \\
\hat{T}_{+1}^{(2)} &= \frac{1}{\sqrt{2}}\left(\hat{d}_0\hat{d}_+^\dagger + \hat{d}_+\hat{d}_0^\dagger\right) \\
\hat{T}_{-1}^{(2)} &= \frac{1}{\sqrt{2}}\left(\hat{d}_0\hat{d}_-^\dagger + \hat{d}_-\hat{d}_0^\dagger\right) \\
\hat{T}_{+2}^{(2)} &= \hat{d}_+\hat{d}_+^\dagger \\
\hat{T}_{-2}^{(2)} &= \hat{d}_-\hat{d}_-^\dagger.
\end{aligned}
$$

Note that several standard references (including references [103, 87]) contain an error in the

prefactor of the $j = 0$ term and in the sign of the $j = 1$ terms. However, the fundamental definitions of $\hat{T}_m^{(j)}$ and its inverse above are valid.

Using recursion relations for the Clebsch-Gordan coefficients we can recast the tensor operators in terms of more intuitive $\hat{f}$ operators [26, 95]

$$
\begin{aligned}
\hat{T}_0^{(0)} &= -\alpha_{f,f'}^{(0)} \hat{\mathbb{1}}_f / \sqrt{3} \\
\hat{T}_0^{(1)} &= +\alpha_{f,f'}^{(1)} \hat{f}_z / \sqrt{2} \\
\hat{T}_{+1}^{(1)} &= +\alpha_{f,f'}^{(1)} \hat{f}_+ / \sqrt{2} \\
\hat{T}_{-1}^{(1)} &= +\alpha_{f,f'}^{(1)} \hat{f}_- / \sqrt{2} \\
\hat{T}_0^{(2)} &= -\alpha_{f,f'}^{(2)} \left(3\hat{f}_z^2 - f(f+1)\hat{\mathbb{1}}_f \right) / \sqrt{6} \\
\hat{T}_{+1}^{(2)} &= -\alpha_{f,f'}^{(2)} \sqrt{2}\hat{f}_+ \left(\hat{f}_z + \hat{\mathbb{1}}_f/2 \right) \\
\hat{T}_{-1}^{(2)} &= -\alpha_{f,f'}^{(2)} \sqrt{2}\hat{f}_- \left(\hat{f}_z - \hat{\mathbb{1}}_f/2 \right) \\
\hat{T}_{+2}^{(2)} &= -\alpha_{f,f'}^{(2)} \hat{f}_+^2 \\
\hat{T}_{-2}^{(2)} &= -\alpha_{f,f'}^{(2)} \hat{f}_-^2 \,.
\end{aligned}
\tag{6.32}
$$

Here we have defined

$$
\begin{aligned}
\alpha_{f,f'}^{(0)} &= \alpha_f^{f'} \Big((2f-1)\delta_{f-1}^{f'} + (2f+1)\delta_f^{f'} \\
&\quad + (2f+3)\delta_{f+1}^{f'} \Big)
\end{aligned}
\tag{6.33}
$$

$$
\begin{aligned}
\alpha_{f,f'}^{(1)} &= \alpha_f^{f'} \left(-\frac{2f-1}{f}\delta_{f-1}^{f'} - \frac{2f+1}{f(f+1)}\delta_f^{f'} \right. \\
&\quad \left. + \frac{2f+3}{f+1}\delta_{f+1}^{f'} \right)
\end{aligned}
\tag{6.34}
$$

$$
\begin{aligned}
\alpha_{f,f'}^{(2)} &= \alpha_f^{f'} \left(\frac{1}{f}\delta_{f-1}^{f'} - \frac{2f+1}{f(f+1)}\delta_f^{f'} \right. \\
&\quad \left. + \frac{1}{f+1}\delta_{f+1}^{f'} \right).
\end{aligned}
\tag{6.35}
$$

These definitions have been chosen to make the each of the quantities

$$
\sum_{f'} \frac{\alpha_{f,f'}^{(j)}}{\alpha_0 \Delta_{f,f'}} > 0
\tag{6.36}
$$

for $\Delta_{f,f'} \gg 0$ for each term j. We have defined the polarizability constants

$$\alpha_f^{f'} = \alpha_0 \frac{(2j'+1)^2}{(2j+1)^2} \left| \left\{ \begin{array}{ccc} 1 & j & j' \\ i & f' & f \end{array} \right\} \right|^2 \tag{6.37}$$

and

$$\begin{aligned} \alpha_0 &= \frac{3\epsilon_0 \hbar \Gamma \lambda_0^3}{8\pi^2} \\ &= \left| \langle j || \hat{\mathbf{d}} || j' \rangle \right|^2 \frac{(2j+1)}{(2j'+1)}, \end{aligned} \tag{6.38}$$

which involves the atomic the spontaneous emission rate, Γ, and transition wavelength, λ_0.

6.1.4 The Irreducible Hamiltonian

Now, to complete the derivation, insert equations (6.32) into the polarizability components of equations (6.28), then insert this and the definition of the electric field, equation (6.4), into the Hamiltonian, equations (6.12, 6.13). Expanding, using the properties of the spherical dot product, and the Stokes component definitions from chapter 3, summing over the f', and defining the detuning dependent prefactors (for $i = 0, 1, 2$)

$$\tilde{\alpha}_{i,f} = \sum_{f'} \frac{\alpha_{f,f'}^{(i)}}{\Delta_{f,f'}} \tag{6.39}$$

we get the final form of the Hamiltonians:

Scalar Hamiltonian

$$\hat{H}^{(0)} = g\tilde{\alpha}_{0,f} \frac{2}{3} \hat{S}_0 \hat{\mathbb{1}}_f \tag{6.40}$$

Vector Hamiltonian

$$\hat{H}^{(1)} = g\tilde{\alpha}_{1,f} \hat{S}_z \hat{f}_z \tag{6.41}$$

Tensor Hamiltonian

$$
\begin{aligned}
\hat{H}^{(2)} \;=\; & g\tilde{\alpha}_{2,f}\left(\hat{S}_{\mathrm{x}}\left(\hat{f}_{\mathrm{x}}^2 - \hat{f}_{\mathrm{y}}^2\right)\right. \\
& + \hat{S}_{\mathrm{y}}\left(\hat{f}_{\mathrm{x}}\hat{f}_{\mathrm{y}} + \hat{f}_{\mathrm{y}}\hat{f}_{\mathrm{x}}\right) \\
& \left. + \hat{S}_0\left(3\hat{f}_{\mathrm{z}}^2 - f(f+1)\hat{\mathbb{1}}_f\right)/3\right)
\end{aligned}
\tag{6.42}
$$

6.2 Interpreting the Irreducible Hamiltonian

6.2.1 The Scalar Hamiltonian

This rank-0 Hamiltonian couples the atomic identity operator $\hat{\mathbb{1}}_f$ to the field mode number operator and can be interpreted as an atomic state-independent light shift. It therefore affects both polarization modes of the probe field in an equivalent manner and will not influence the measurement process since it does not provide any state-dependent information. However, this Hamiltonian would be important if the measurement was meant to distinguish between populations across hyperfine states (e.g., $f = 3$ and $f = 4$ using homodyne detection) instead of across the sublevel populations within one hyperfine state (using polarimetry, as discussed here). This term is also of importance if the Hamiltonian is being considered as a spatially dependent potential for the atoms (e.g., in an optical lattice).

6.2.2 The Vector Hamiltonian

The rank-1 Hamiltonian can be interpreted as causing a differential phase shift on the two circular polarization modes by an amount that is proportional to the z-component of the atomic angular momentum. Thus the vector Hamiltonian leads to optical activity in the atomic sample and produces the familiar Faraday rotation effect often used to address continuous measurement of collective spin [42, 15, 16, 104, 24]. This is the QND Hamiltonian that enables much of the idealized physics of this thesis, as it is linear in the spin-operator and the collective Hamiltonian resulting from summing individual Hamiltonians will be proportional to $\hat{F}_{\mathrm{z}}$.

6.2.3 The Tensor Hamiltonian

The rank-2 Hamiltonian couples spin coordinates to the elliptical components of the probe laser field and produces a second-order light shift proportional to the atomic quadrupole moment. These terms vanish for $f = 1/2$ (as can be seen by evaluating the operators within the parentheses above) but are nonzero for any higher spin number. The symmetry of the tensor term can be seen by noting the identity $\hat{f}_x \hat{f}_y + \hat{f}_y \hat{f}_x = \hat{f}_{x'}^2 - \hat{f}_{y'}^2$. For a linearly polarized input beam, the tensor term leads to an elliptically polarized scattered probe field [32, 26]. The rank-2 interaction potentially limits the validity of any collective measurement scheme based on the vector Hamiltonian alone. However, in the next section we analyze the spectral properties of the prefactors on all of these Hamiltonians and see that the tensor term vanishes for large detunings.

6.3 Coefficient Spectra

As we have seen, it is convenient to break down the Hamiltonian interaction into the scalar, vector, and tensor components, with the respective detuning dependent prefactors

$$\tilde{\alpha}_{0,f} = \sum_{f'} \frac{\alpha_{0ff'}}{\Delta_{ff'}} \tag{6.43}$$

$$\tilde{\alpha}_{1,f} = \sum_{f'} \frac{\alpha_{1ff'}}{\Delta_{ff'}} \tag{6.44}$$

$$\tilde{\alpha}_{2,f} = \sum_{f'} \frac{\alpha_{2ff'}}{\Delta_{ff'}}. \tag{6.45}$$

In this section, we investigate the spectral properties of these prefactors. The full spectrum of these quantities for the $f = 3$ and $f = 4$ lines are shown in figure 6.1 and figure 6.2 respectively. Experimental results presenting the vector and tensor spectra for the $f = 4$ line are presented in section 13.5. In figure 6.3 and figure 6.4 it is shown that the tensor term scales as $1/\Delta^2$ at detunings larger than the excited state hyper-fine splittings.

Notice that, because of the fact that the $\alpha_{iff'}$ can be either positive or negative, there are certain noninfinite frequencies at which each of the prefactors $\tilde{\alpha}_{i,f}$ can be zero, where the other terms are nonzero. These specific frequencies may be of interest because of the fact that one can remove one type of evolution while retaining the others. The character of each type of interaction is, by definition, fundamentally different and choosing a frequency
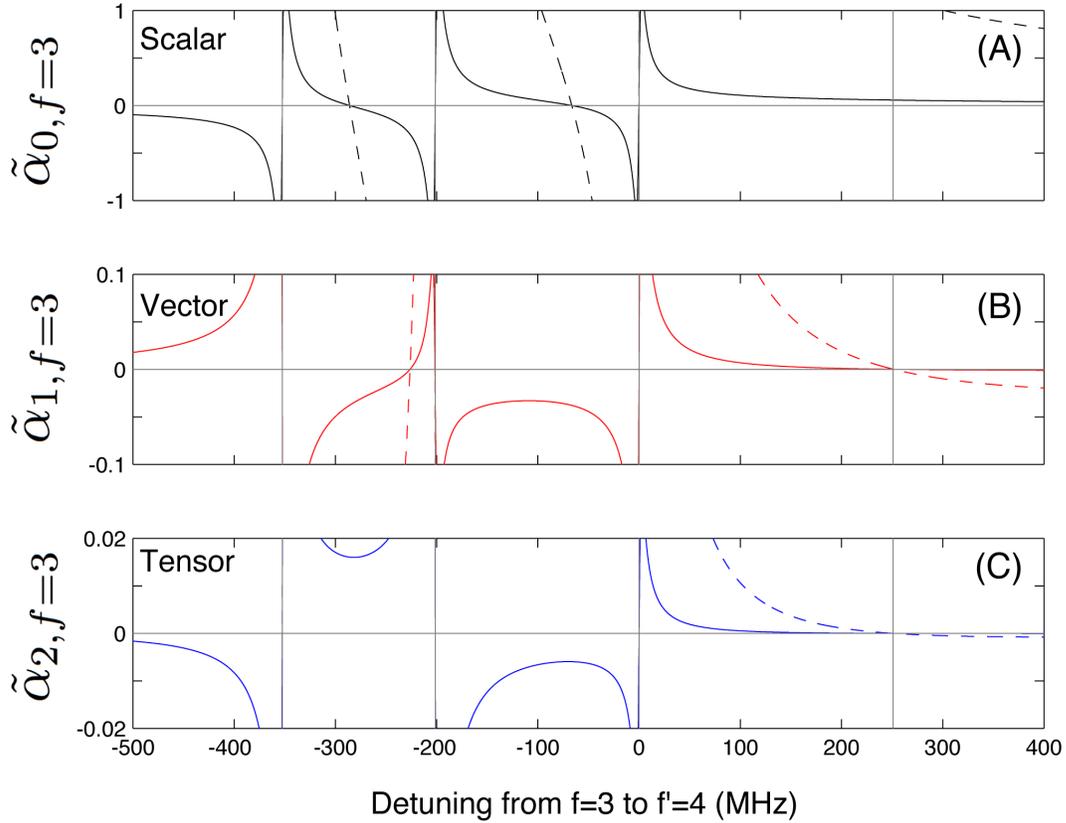
Figure 6.1: Plotted, in order, are $\tilde{\alpha}_{0,f=3}$ (A), $\tilde{\alpha}_{1,f=3}$ (B), and $\tilde{\alpha}_{2,f=3}$ (C). The green curve is a magnified version ($\times 20$) of the red curve to better see the zero-crossings. Looking at the axes of each plot it is clear that the scalar term is approximately an order of magnitude greater than the vector term, which, in turn, is approximately an order of magnitude greater than the tensor term. Notice that the vector and tensor terms cross zero close to the forbidden transition $f = 3$ to $f' = 5$.

to amplify or suppress one type may be of practical consequence.

Define $\Delta_{i,f,n}$ as the nth detuning (with reference to one particular transition) that the i prefactor in the Hamiltonian is zero: $\tilde{\alpha}_{i,f}(\Delta_{i,f,n}) = 0$. We introduce n because there can be many points at which the function attains the value of zero.
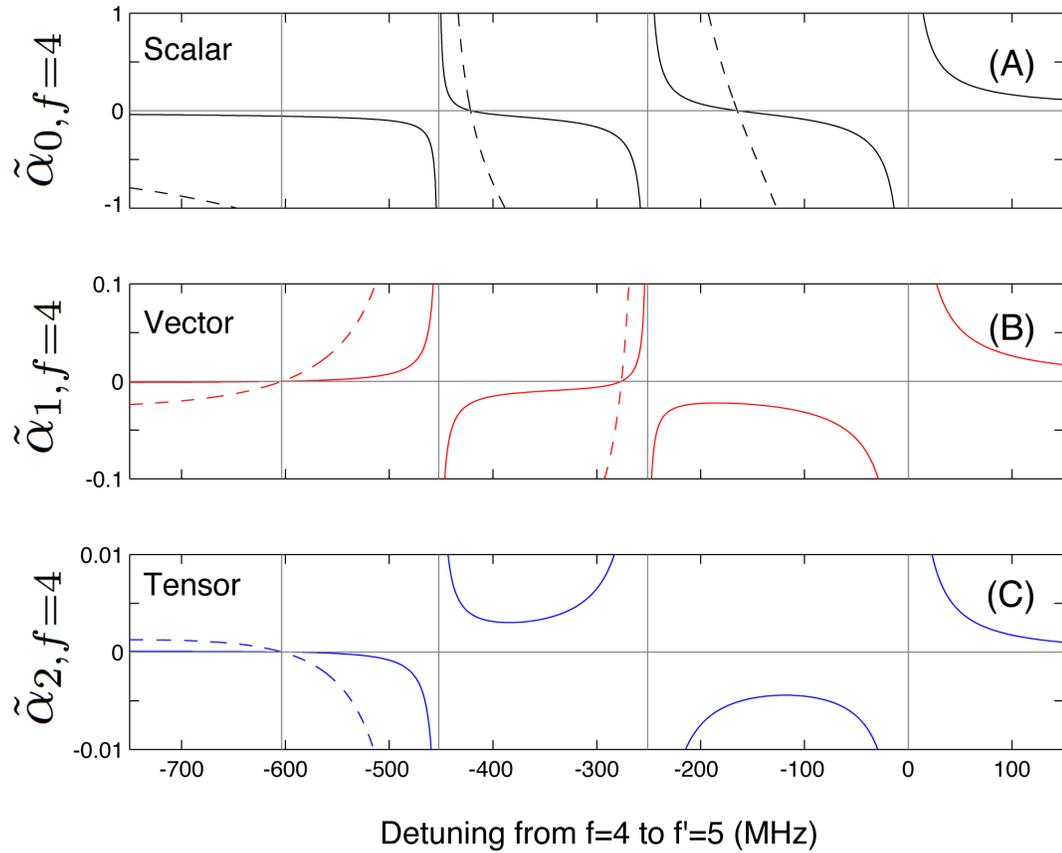
Figure 6.2: Plotted, in order, are $\tilde{\alpha}_{0,f=4}$ (A), $\tilde{\alpha}_{1,f=4}$ (B), and $\tilde{\alpha}_{2,f=4}$ (C). The green curve is a magnified version ($\times 20$) of the red curve to better see the zero-crossings. Looking at the axes of each plot it is clear that the scalar term is approximately an order of magnitude greater than the vector term, which, in turn, is approximately an order of magnitude greater than the tensor term. Notice that the vector and tensor terms cross zero close to the forbidden transition $f = 4$ to $f' = 2$.
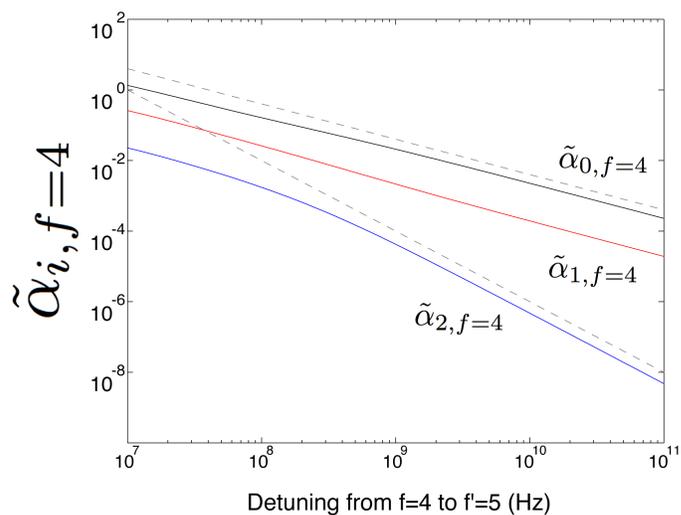
Figure 6.3: The top curve is $\propto 1/\Delta$ for reference. Below this in order are $\tilde{\alpha}_{0,f=4}$, $\tilde{\alpha}_{1,f=4}$, and $\tilde{\alpha}_{2,f=4}$ plotted on a log scale. Notice that between 100 MHz and 1 GHz the tensor term starts to approach zero faster at a rate $\propto 1/\Delta^2$.
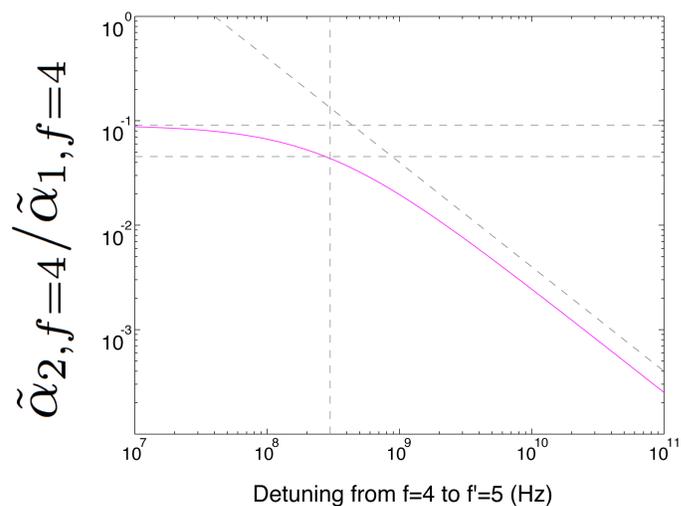
Figure 6.4: The ratio $\tilde{\alpha}_{2,f=4}/\tilde{\alpha}_{1,f=4}$ is plotted in log space. The close to resonance value is 0.09 and scales down as $1/\Delta$ at large detunings. The 3 dB point where the ratio is $0.09/2$ is approximately 300 MHz.

6.3.1 D2: f = 4 Zeros

Consider the D2 $f = 4$ line of cesium and define the detuning as that from the $f = 4$ to $f' = 5$ transition,

$$\Delta_{45} \equiv \Delta \tag{6.46}$$

$$\Delta_{44} = \Delta_{H45} + \Delta \tag{6.47}$$

$$\Delta_{43} = \Delta_{H35} + \Delta \tag{6.48}$$

where we have defined the hyperfine splittings as

$$\Delta_{H45} = 251.002 \text{ MHz} \tag{6.49}$$

$$\Delta_{H35} = 251.002 + 201.242 = 452.244 \text{ MHz} \tag{6.50}$$

$$\Delta_{H25} = 251.002 + 201.242 + 151.212 = 603.456 \text{ MHz} \tag{6.51}$$

$$\Delta_{i=0,f=4,n=1} = -421.5 \text{ MHz} \tag{6.52}$$

$$\Delta_{i=0,f=4,n=2} = -164.6 \text{ MHz} \tag{6.53}$$

$$\Delta_{i=1,f=4,n=1} = -602.9 \text{ MHz} \tag{6.54}$$

$$\Delta_{i=1,f=4,n=2} = -276.1 \text{ MHz} \tag{6.55}$$

$$\Delta_{i=2,f=4,n=1} = -603.4 \text{ MHz.} \tag{6.56}$$

Notice that $\Delta_{i=1,f=4,n=1} \approx \Delta_{i=2,f=4,n=1} \approx \Delta_{H25}$ to within 1 MHz. This is of interest because at this frequency the light shift becomes essentially polarization independent. It is also of interest that this frequency occurs near the frequency of the forbidden $f = 4$ to $f' = 2$ transition. The intuitive reason for this "coincidence" is left as an exercise for the reader (i.e., I am sure there is a reason, but I have not worked it out).

6.3.2   D2: f = 3 Zeros

Now consider the D2 $f = 3$ line of cesium and define the detuning as that from the $f = 3$ to $f' = 4$ transition

$$
\begin{align}
\Delta_{34} &\equiv \Delta \tag{6.57}\\
\Delta_{33} &= \Delta_{H34} + \Delta \tag{6.58}\\
\Delta_{32} &= \Delta_{H24} + \Delta \tag{6.59}
\end{align}
$$

where we have defined the hyperfine splittings as

$$
\begin{align}
\Delta_{H45} &= 251.002 \text{ MHz} \tag{6.60}\\
\Delta_{H34} &= 201.242 \text{ MHz} \tag{6.61}\\
\Delta_{H24} &= 201.242 + 151.212 = 352.454 \text{ MHz} \tag{6.62}
\end{align}
$$

$$
\begin{align}
\Delta_{i=0,f=3,n=1} &= -285.9 \text{ MHz} \tag{6.63}\\
\Delta_{i=0,f=3,n=2} &= -66.4 \text{ MHz} \tag{6.64}\\
\Delta_{i=1,f=3,n=1} &= -226.4 \text{ MHz} \tag{6.65}\\
\Delta_{i=1,f=3,n=2} &= 251.7 \text{ MHz} \tag{6.66}\\
\Delta_{i=2,f=3,n=1} &= 251.0 \text{ MHz.} \tag{6.67}
\end{align}
$$

Similarly to the last section, notice that $\Delta_{i=1,f=3,n=2} \approx \Delta_{i=2,f=4,n=1} \approx \Delta_{H45}$ to within 1 MHz. Again the polarization dependence of the light shift can be removed by using a frequency near the forbidden transition ($f = 3$ to $f' = 5$).

6.4   Semiclassical Evolution of the Probe State

We can greatly simplify the dynamics by eliminating atomic evolution due to the probe beam and only considering the evolution of the probe beam due to the atomic state. Under this semiclassical approximation, we replace all atomic operators with their expectation values with respect to an assumed fixed spin-state. (This is the opposite of the semiclassical situation often considered in atom-light interactions where the atomic system is considered
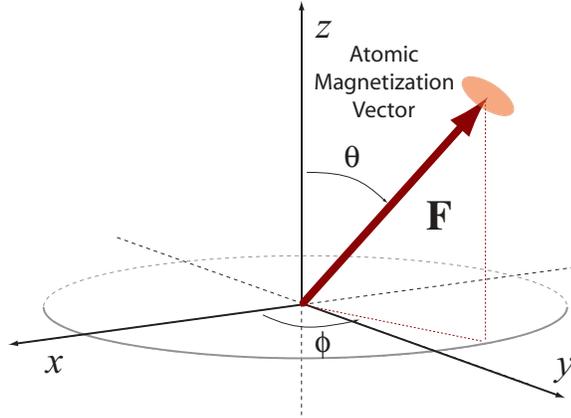
Figure 6.5: Definition of the spherical coordinate angles used to describe the orientation of the collective atomic magnetization vector, **F**, relative to the fixed laboratory cartesian coordinate system. The polarization vector of the input probe light resides in the xy-plane and forms an angle, ϕ_p, with respect to the laboratory x-axis.

quantum mechanically while the optical beam is made classical.) For a large ensemble of atoms and small interaction times, fixing the atomic state will accurately reproduce the mean behavior of the measured photocurrent corresponding to one of the Stokes vector components. This is confirmed experimentally in chapter 13, where the atomic state is fixed and adiabatically positioned with a magnetic holding field. The holding field serves to both position the atomic state and protect it from the influence of the probe light, such that the analysis of this section remains valid even for long interaction times or large optical depth clouds. Ultimately, however, probe induced decoherence will dominate all interactions. In chapter 8, we then reconsider the full analysis including the atomic quantum noise (related to spin-squeezing) for a particular alignment of the collective spin-state.

We approximate the N-atom Hamiltonian, $\hat{H}_N$, by replacing the single-atom operators with their expectation value taken with respect to an optically pumped spin pointing with direction θ and φ given in spherical coordinates (Fig. 6.5). In other words, for an individual atom operator $\hat{O}_f$, we take

$$\hat{O}_f \rightarrow \langle \hat{O}_f \rangle = \langle \Psi(\theta, \varphi) | \hat{O}_f | \Psi(\theta, \varphi) \rangle \tag{6.68}$$

where $|\Psi(\theta, \varphi)\rangle = \exp[-i\hat{f}_z\varphi] \exp[-i\hat{f}_y\theta] |f, f\rangle_z$.

The relevant operators from the Hamiltonian decomposition are given by

$$\langle \hat{f}_z \rangle = f \cos\theta \tag{6.69}$$

$$\langle \hat{f}_x^2 - \hat{f}_y^2 \rangle = f(f-1/2)\sin^2\theta \cos 2\varphi \tag{6.70}$$

$$\langle \hat{f}_x \hat{f}_y + \hat{f}_y \hat{f}_x \rangle = f(f-1/2)\sin^2\theta \sin 2\varphi. \tag{6.71}$$

Within the semiclassical approximation, we obtain an effective scattering interaction Hamiltonian that only involves operators on the probe field Hilbert space. Ignoring all terms proportional to $\hat{S}_0$ (because it commutes with each term of the semiclassical Hamiltonian) we have

$$\begin{aligned} \tilde{H} &= \tilde{H}_N^{(1)} + \tilde{H}_N^{(2)} \\ &= (\gamma_x \hat{S}_x + \gamma_y \hat{S}_y + \gamma_z \hat{S}_z)\frac{\hbar}{\delta t}, \end{aligned} \tag{6.72}$$

which leads to a rotation of the Stokes vector $\hat{\mathbf{S}}$ about a vector $\boldsymbol{\gamma} = [\gamma_x, \gamma_y, \gamma_z]$ according to the evolution operator

$$\begin{aligned} \tilde{U}_{\delta t} &= \exp\left[-i\tilde{H}\frac{\delta t}{\hbar}\right] \tag{6.73} \\ &= \exp\left[-i(\gamma_x \hat{S}_x + \gamma_y \hat{S}_y + \gamma_z \hat{S}_z)\right] \tag{6.74} \end{aligned}$$

where $\delta t = L/c$ is the interaction (transit) time of the discrete spatial modes of the probe beam across the atomic cloud of length L. The rotation vector $\boldsymbol{\gamma}$ is defined by

$$\gamma_x = \gamma_0 f(f-1/2)\sin^2\theta\cos 2\varphi \sum_{f'} \frac{\alpha_{f,f'}^{(2)}}{\alpha_0 \Delta_{f,f'}} \tag{6.75}$$

$$\gamma_y = \gamma_0 f(f-1/2)\sin^2\theta\sin 2\varphi \sum_{f'} \frac{\alpha_{f,f'}^{(2)}}{\alpha_0 \Delta_{f,f'}} \tag{6.76}$$

$$\gamma_z = \gamma_0 f\cos\theta \sum_{f'} \frac{\alpha_{f,f'}^{(1)}}{\alpha_0 \Delta_{f,f'}}. \tag{6.77}$$

Here we have normalized by the state-independent polarizability constant

$$\alpha_0 = \frac{3\epsilon_0\hbar\Gamma\lambda_0^3}{8\pi^2} \tag{6.78}$$

$$= \left|\langle j||\hat{\mathbf{d}}||j'\rangle\right|^2 \frac{(2j+1)}{(2j'+1)} \tag{6.79}$$

such that $\alpha_{f,f'}^{(j)}/\alpha_0$ is dimensionless. The rotation strength is represented by

$$\gamma_0 = \frac{Ng\delta t\alpha_0}{\hbar} \tag{6.80}$$

where we have used the field coefficient $g = \omega_0/(2\epsilon_0 V)$, the atomic resonance frequency ω_0, and the interaction volume (the volume of the atomic sample) $V = AL$. The optical depth takes on an important role in the spin-squeezing discussion of chapter 8.

From an experimental standpoint, it is useful to note that γ_0 is directly related to the on-resonance *optical depth*, OD, of the atomic sample and the decay rate Γ via,

$$\gamma_0 = \left(\frac{\Gamma}{4}\right)\text{OD} \tag{6.81}$$

where

$$\text{OD} = N\frac{\sigma_0}{A}, \quad \sigma_0 = \frac{3\lambda_0^2}{2\pi}. \tag{6.82}$$

The quantity, σ_0, is the resonant atomic scattering cross section and $A = \pi r^2$ is the cross sectional area of the atomic sample.

In section 3.2.5, the equations for a general rotation of $\hat{\mathbf{S}}$ about $\boldsymbol{\gamma}$ are given. Here we specialize to the case where the input beam is linearly polarized in the x-direction such that $\langle \hat{S}_y \rangle = \langle \hat{S}_z \rangle = 0$. The output expectation values are then given by

$$\langle \hat{S}_x' \rangle = \langle \hat{S}_x \rangle \left(\cos\gamma + \frac{\gamma_x^2}{\gamma^2}(1 - \cos\gamma) \right) \tag{6.83}$$

$$\langle \hat{S}_y' \rangle = \langle \hat{S}_x \rangle \left(-\frac{\gamma_z}{\gamma}\sin\gamma + \frac{\gamma_y\gamma_x}{\gamma^2}(1 - \cos\gamma) \right) \tag{6.84}$$

$$\langle \hat{S}_z' \rangle = \langle \hat{S}_x \rangle \left(\frac{\gamma_y}{\gamma}\sin\gamma + \frac{\gamma_z\gamma_x}{\gamma^2}(1 - \cos\gamma) \right). \tag{6.85}$$

Taking the total rotation angle small ($\gamma \ll 1$) this becomes (to second-order in γ)

$$
\begin{align}
\langle \hat{S}'_x \rangle &\approx \langle \hat{S}_x \rangle \left(1 - \gamma_z^2/2 - \gamma_y^2/2 \right) \tag{6.86} \\
\langle \hat{S}'_y \rangle &\approx \langle \hat{S}_x \rangle \left(-\gamma_z + \frac{\gamma_y \gamma_x}{2} \right) \tag{6.87} \\
\langle \hat{S}'_z \rangle &\approx \langle \hat{S}_x \rangle \left(\gamma_y + \frac{\gamma_z \gamma_x}{2} \right). \tag{6.88}
\end{align}
$$

In this semiclassical approximation, we have completely neglected any evolution of the atomic state due to the probe beam. We demonstrate in the next section that the above model agrees well with experimental data when the spin-state is fixed with a magnetic holding field.

6.5 Tensor Suppression at the Magic Angle

Now we will write the tensor Hamiltonian in another form to arrive at a more intuitive form of its dynamics. From before, we have

$$
\begin{align}
\hat{H}^{(2)} = g\tilde{\alpha}_{2,f} \Big(&\hat{S}_{\mathrm{x}} \left(\hat{f}_{\mathrm{x}}^2 - \hat{f}_{\mathrm{y}}^2 \right) \\
&+ \hat{S}_{\mathrm{y}} \left(\hat{f}_{\mathrm{x}} \hat{f}_{\mathrm{y}} + \hat{f}_{\mathrm{y}} \hat{f}_{\mathrm{x}} \right) \\
&+ \hat{S}_0 \left(3\hat{f}_{\mathrm{z}}^2 - f(f+1)\hat{\mathbb{1}}_f \right)/3 \Big). \tag{6.89}
\end{align}
$$

Now let us assume the light to be linearly polarized along the unit vector $\boldsymbol{\epsilon}_p = \cos\theta \mathbf{e}_x + \sin\theta \mathbf{e}_y$ such that we have $\langle \hat{S}_x \rangle = N_p(\sin^2\theta - \cos^2\theta)/2$, $\langle \hat{S}_y \rangle = -N_p \sin\theta \cos\theta$, and $\langle \hat{S}_z \rangle = 0$ where N_p is the number of photons in the mode. Using standard identities it can be shown that

$$
\hat{H}^{(2)} = g\tilde{\alpha}_{2,f} N_p \left(-(\boldsymbol{\epsilon}_p \cdot \hat{\mathbf{f}})^2 + f(f+1)\hat{\mathbb{1}}_f/3 \right). \tag{6.90}
$$

Now we assume a large magnetic field along the y-direction and go into the rotating frame about y, using the transformation

$$
\hat{\tilde{O}} = \exp[-i\omega t \hat{f}_{\mathrm{y}}] \hat{O} \exp[+i\omega t \hat{f}_{\mathrm{y}}]. \tag{6.91}
$$

If we *time-average* over the timescale given by the frequency ω, then we get (ignoring identity terms)

$$\hat{\bar{H}} \quad \propto \quad \hat{f}_{\mathrm{y}}^2 \left(\sin^2(\theta) - \cos^2(\theta)/2 \right) . \tag{6.92}$$

This time-averaged Hamiltonian clearly vanishes at the so-called magic angle $\theta = \arctan(1/\sqrt{2}) = 35.3$ degrees relative to the x-axis, or 54.7 degrees relative to the field along the y-axis [25]. The Larmor precession dynamics at the magic angle were discussed earlier in section 5.2.1.

6.6 Tensor Suppression with Parallel Polarizations

In the last section, we considered nulling the effect of the tensor Hamiltonian *on average* in the presence of a magnetic field. Here we show another way to avoid the effects of the tensor Hamiltonian, but without a field, by simply putting the spin-state at a fixed point of the Hamiltonian evolution. In the case that the linear optical polarization is also along x ($\langle \hat{S}_x \rangle = -\langle \hat{S}_0 \rangle$), it turns out that a spin-state along x does not evolve at all due to the semiclassical tensor Hamiltonian (taking the expectation of the optical operators before evolving). This can be easily seen because the Hamiltonian (ignoring terms that obviously do not cause evolution) becomes proportional to $\hat{f}_{\mathrm{z}}^2 - (\hat{f}_{\mathrm{x}}^2 - \hat{f}_{\mathrm{y}}^2) = \hat{\mathbf{f}}^2 - 2\hat{f}_{\mathrm{x}}^2$ of which the x-polarized state is obviously an eigenstate. In contrast, for perpendicularly polarized light along y ($\langle \hat{S}_x \rangle = +\langle \hat{S}_0 \rangle$) the x-polarized state is not an eigenstate and the state evolves in a nonlinear way. The dynamics for both of these situations is demonstrated in figure 5.2. Actually, in that simulation the uncertainties in the parallel configuration grow in time, but this is purely a spontaneous emission effect and not a result of the tensor Hamiltonian.

Despite the fact that the spin-state does not evolve due to the semiclassical tensor Hamiltonian in the parallel configuration, it is not necessarily a good idea to try spin-squeezing close to resonance for several reasons. First, any static or dynamic misalignments will cause the tensor Hamiltonian to introduce non-QND effects. Second, non-single-atom collective effects (as shown in chapter 7) can lead to nontrivial evolution of the collective uncertainty, even in the parallel configuration.

6.7 Tensor-Driven Oscillations

Here we expand on the last section, but include a field along the spin direction. This is relevant for the stroboscopic scheme discussed in chapter 14. Consider a single spin aligned along x-direction, in the state $\hat{\rho}_x$, and apply only the tensor Hamiltonian above, and a field along x. The relevant expectation values (using the Quantum Optics Toolbox for the case $f = 4$) are

$$\text{Tr}(\hat{\rho}_x[\hat{\gamma}_i, \hat{f}_x]) = 0 \tag{6.93}$$

$$\text{Tr}(\hat{\rho}_x[\hat{\gamma}_i, \hat{f}_y]) = -4i\delta_{i,z}\tilde{\alpha}_{1,f}\frac{gdt}{\hbar} \tag{6.94}$$

$$\text{Tr}(\hat{\rho}_x[\hat{\gamma}_i, \hat{f}_z]) = 28i\delta_{i,y}\tilde{\alpha}_{2,f}\frac{gdt}{\hbar}, \tag{6.95}$$

which give the $t = 0$ derivatives as

$$\frac{d\langle\hat{f}_x\rangle}{dt} = 0 \tag{6.96}$$

$$\frac{d\langle\hat{f}_y\rangle}{dt} = 4\langle\hat{S}_z\rangle\tilde{\alpha}_{1,f}\frac{gdt}{\hbar} \tag{6.97}$$

$$\frac{d\langle\hat{f}_z\rangle}{dt} = -28\langle\hat{S}_y\rangle\tilde{\alpha}_{2,f}\frac{gdt}{\hbar}. \tag{6.98}$$

By simulation (including Lindblad terms) one can see that both $\langle\hat{f}_y\rangle$ and $\langle\hat{f}_z\rangle$ remain zero for all time if $\langle\hat{S}_y\rangle = \langle\hat{S}_z\rangle = 0$.

In the case that $\langle\hat{S}_y\rangle$ is nonzero, then a small oscillation between $\langle\hat{f}_y\rangle$ and $\langle\hat{f}_z\rangle$ is started. On timescales large compared to the Larmor frequency, the average of this effect is indeed zero (according to the rotating wave approximation of the previous section). However, in our scheme where we expect the quantum noise to be oscillating at the Larmor frequency, this nonzero semiclassical amplitude is a problem because classical multiplicative number fluctuations will cause a signal that acts like (and masks) the quantum noise signal. In the search for the quantum noise, the situation is a bit more complicated because you want to be strobing the probe and filtering appropriately, but the same general considerations apply.

Although this effect can be negative for these technical reasons, the effect can also be used to align (orthogonalize) the polarization of the probe beam with the polarization of

the atomic state. In fact this is how we fine tune this alignment in the lab and experimental results along these lines are presented in section 13.1.3. The nonlinearity of the tensor term is also useful in other practical tasks. For example, one can determine the degree of optical pumping via the tensor dynamics as discussed in the next section.

6.8 Tensor Pumping Tomography

The previous section showed that the tensor Hamiltonian can be useful for aligning the atomic and optical polarizations. In this section, we show that, as a direct consequence of the nonlinearity in atomic spin operators, the tensor term can also be used to measure the atomic pumping efficiency.

Optical pumping is the incoherent process [31] by which an atom absorbs angular momentum from an optical beam, thus aligning the spin of the atom. We discuss simulations of this process for a real multilevel cesium atom in section 12.5.9. When it comes to measuring the spin-state of a single atom, there are various state tomography procedures available. In section 12.5.5, we discuss a standard Stern-Gerlach technique by which the internal sublevels are separated with a sufficiently strong magnetic field gradient and measured via fluorescence. This technique is used for state tomography in experiments similar to ours, for example [105]. There are many reasons for not wanting to use this technique however, including experimental complexity and a limited measurement resolution. In principle, one can use the continuous measurement as discussed in this work to determine the initial spin-state. Extended notions on tomography with continuous quantum measurement and control are discussed in [106, 107].

Mostly we are interested in a very restricted version of the general tomography problem and we would like to know with what efficiency we have prepared the particular coherent spin-state $|f = 4, m_f = 4\rangle$. Here we present a simple example showing how we can do this by aligning the spin-state along a certain direction, e.g., with a adiabatic holding field, sending in a beam polarized along x, and measuring a particular output Stokes component. By following this procedure for two different configurations, the pumping efficiency can be determined as follows.

Consider first the atomic alignment with $\theta = \phi = 0$ and measured Stokes component $\langle \hat{S}'_y \rangle$. Referring to equation (6.77) and making the replacement $f \to \eta f$ where η is repre-

sentative of the pumping efficiency, we have

$$\gamma_{z,1} \approx \gamma_0 \eta f \sum_{f'} \frac{\alpha_{f,f'}^{(1)}}{\alpha_0 \Delta_{f,f'}} \tag{6.99}$$

with the $\gamma_{x,1} = \gamma_{y,1} = 0$. Then by specializing equation (6.84) to this case we see that

$$\langle \hat{S}'_{y,1} \rangle \approx -\langle \hat{S}_x \rangle \gamma_{z,1}. \tag{6.100}$$

Next, consider the state with $\theta = \pi/2$ and $\phi = \pi/4$ such that

$$\gamma_{y,2} \approx \gamma_0 \eta^2 f(f - 1/2) \sum_{f'} \frac{\alpha_{f,f'}^{(2)}}{\alpha_0 \Delta_{f,f'}} \tag{6.101}$$

with the others again zero. The measurement of the circular polarization then gives

$$\langle \hat{S}'_{z,2} \rangle \approx \langle \hat{S}_x \rangle \gamma_{y,2}. \tag{6.102}$$

Now by considering the measured *ratio* of the measurements we have $\langle \hat{S}'_{z,2} \rangle / \langle \hat{S}'_{y,1} \rangle \propto \eta$, with some model dependent terms canceling in the ratio. Clearly, this scheme uses the fact that one interaction is quadratic while the other is only linear.

Of course, mapping the measured ratio to the pumping efficiency will depend on the model being used for the state population. Our goal is only to prepare a measurement that is approximately correct near unity pumping efficiency. Consider two models of the state distribution. In one, given a "temperature parameter" t we populate the states according to

$$\rho = \sum_{i=-f}^{f} t^m |f = 4, m = i\rangle\langle f = 4, m = i| \tag{6.103}$$

then normalize. In another model, we allow only two levels to have any population and define the state as

$$\rho = t|f = 4, m = 4\rangle\langle f = 4, m = 4| + (1 - t)|f = 4, m = 3\rangle\langle f = 4, m = 3|. \tag{6.104}$$

As seen in figure 6.6, the models give nearly the same result for the measured ratio for nearly perfect preparation efficiencies greater than 90%. Experimental measurements using
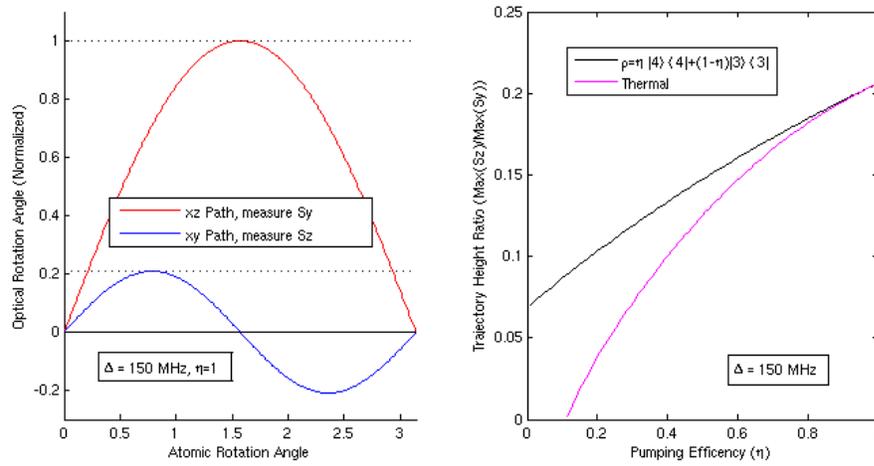
Figure 6.6: Measurement of pumping efficiency as described in the text. On the left, the upper plot corresponds to dragging the spins in the xz-plane and measuring $\langle \hat{S}'_y \rangle$. The lower plot corresponds to dragging the spins in the xy-plane and measuring $\langle \hat{S}'_z \rangle$. The measurements referred to in the text are the peaks of these curves. On the right the ratio is mapped to the pumping efficiency via two different models.

this technique are presented in section 13.6.

Chapter 7

The Unconditional Collective Master Equation

Thus far we have considered only the semiclassical descriptions of the atom-light interaction. In chapter 5, we used a classical description of the light to derive a single atom master equation with spontaneous emission. After that, we discussed the decomposition of the polarizability Hamiltonian in chapter 6, where we treated the light quantum mechanically, but only considered one of two semiclassical limits where either the optical moments or spin moments were considered as classical.

In this chapter, we start from the fully quantum Hamiltonian of the last chapter, with quantum light and quantum atoms, and derive the unconditional master equation for the collective state of the atoms. Under these circumstances, we can begin using some of the quantum measurement language that was introduced in chapter 2. In the case that the Hamiltonian only consists of the QND vector term, this will result in no entanglement unconditionally, only antisqueezing. On the other hand, the conditional solution will produce spin-squeezing as discussed in the next chapter. However, if we include the tensor terms, *and* the assumption that all of the atoms interact with the same optical mode symmetrically, we will see that collective spin-squeezed states can be prepared via the unconditional master equation described here.

The latter assumption is a significant one, which will not be completely true under current experimental circumstances, but we continue with it for pedagogical reasons. A more realistic, and complicated, examination of the effect of the tensor term in a spatially extended atomic cloud is analyzed in [26]. For simplicity, we also continue to neglect the spontaneous emission, and other three-dimensional effects, discussed in chapter 5.

7.1 Collective Master Equation Derivation

Now we derive the unconditional master equation describing a collection of cesium atoms being probed by an arbitrarily polarized beam with detuning comparable to the hyperfine splitting between the excited states of the atom, such that the tensor Hamiltonian terms are significant. We neglect spontaneous emission and assume symmetry of the cloud so all atoms are sitting at one point and are completely forward scattering.

We begin with the Hamiltonian as derived in the previous chapter

$$\hat{H} = \frac{\hbar}{dt}(\hat{\gamma}_0 \hat{S}_0 + \hat{\gamma}_x \hat{S}_x + \hat{\gamma}_y \hat{S}_y + \hat{\gamma}_z \hat{S}_z) \tag{7.1}$$

where the $\hat{\gamma}_i$ are atomic operators and the $\hat{S}_i$ are the Stokes operators representing the polarization of the probe beam. The $\hat{\gamma}_i$ are defined as

$$\hat{\gamma}_0 = c_0 \hat{I} + c_2 \sum_i \hat{f}_{z,i}^2 \tag{7.2}$$

$$\hat{\gamma}_x = c_2 \sum_i (\hat{f}_{x,i}^2 - \hat{f}_{y,i}^2) \tag{7.3}$$

$$\hat{\gamma}_y = c_2 \sum_i (\hat{f}_{x,i}\hat{f}_{y,i} + \hat{f}_{y,i}\hat{f}_{x,i}) \tag{7.4}$$

$$\hat{\gamma}_z = c_1 \sum_i \hat{f}_{z,i} = c_1 \hat{F}_z \tag{7.5}$$

with the detuning dependent coefficients

$$c_0 = \frac{dt}{\hbar}g \sum_{f'} \left(\frac{\alpha_{0ff'}}{\Delta_{ff'}} \frac{2}{3} - \frac{\alpha_{2ff'}}{\Delta_{ff'}} \frac{f(f+1)}{3} \right) \tag{7.6}$$

$$c_1 = \frac{dt}{\hbar}g \sum_{f'} \frac{\alpha_{1ff'}}{\Delta_{ff'}} \tag{7.7}$$

$$c_2 = \frac{dt}{\hbar}g \sum_{f'} \frac{\alpha_{2ff'}}{\Delta_{ff'}} \tag{7.8}$$

where we have removed the f subscripts for convenience. Unless otherwise stated, we assume the $f = 4$ ground state of cesium.

Now we go through a standard master equation derivation assuming each of the polarization modes to be sparsely populated in a time-slice dt, then expanding to second-order in the coherent state amplitude of each of the polarization modes [104]. This includes the

quantum effects of the light. We start at time t_i, just as a slice of laser light of width dt, with state $\hat{\rho}_{L,i}$ begins to pass over the collective atomic sample with initial state $\hat{\rho}_A(t_i)$. At the beginning of each timestep $\hat{\rho}_{L,i}$ for a new i is always the same, while $\hat{\rho}_A(t_i)$ depends on past interactions. The total system state, $\hat{\rho}_T(t_i)$ at this time is separable because the incoming optical state has yet to interact with the atoms

$$\hat{\rho}_T(t_i) = \hat{\rho}_{L,i} \otimes \hat{\rho}_A(t_i). \tag{7.9}$$

After interacting for a time dt, via the Hamiltonian $\hat{H}$ the total system state before the next timestep is then,

$$\hat{\rho}_T(t_i + dt) = \hat{U}\hat{\rho}_T(t_i)\hat{U}^\dagger \tag{7.10}$$

where $\hat{U} = \exp[-i\hat{H}dt/\hbar]$. At this stage the joint atom-light system is potentially entangled as the traveling light mode leaves the atoms and ceases to interact. Now we can either measure this light mode or ignore it. In the next chapter we consider measuring it but for now we simply ignore it. Tracing over the light mode after the dt interaction gives the atomic state at the next timestep $t_{i+1} = t_i + dt$ as

$$
\begin{aligned}
\hat{\rho}_A(t_{i+1}) &= \mathrm{Tr}_L[\hat{\rho}_T(t_{i+1})] \tag{7.11} \\
&= \mathrm{Tr}_L[\hat{U}\hat{\rho}_{L,i} \otimes \hat{\rho}_A(t_i)\hat{U}^\dagger]. \tag{7.12}
\end{aligned}
$$

Henceforth we assume that the initial optical state is the same on every timestep t_i and simplify this equation to derive the unconditional master equation describing the atomic spins. First we define the initial optical state (on every timestep) as

$$|\Phi\rangle_L = |\alpha_x\rangle_x \otimes |\alpha_y\rangle_y \tag{7.13}$$

where the $|\alpha_i\rangle_i$ are coherent states in the i polarization basis as defined in chapter 3. We leave the polarization of the state in its most general form. This is a key generalization beyond [104] because, with a vector interaction alone, the angle of linear polarization does not matter, but with the tensor interaction considered here the resulting dynamics will be highly polarization dependent.

Using the definition of the coherent state, and taking α_i small because the timestep dt

is small, we expand to first-order in α_i to get

$$|\Phi\rangle_L \approx \left((1 - |\alpha_x|^2/2)|0\rangle_x + \alpha_x|1\rangle_x\right) \otimes \left((1 - |\alpha_y|^2/2)|0\rangle_y + \alpha_y|1\rangle_y\right) \tag{7.14}$$

$$= (1 - |\alpha_x|^2/2 - |\alpha_y|^2/2)|0\rangle_x|0\rangle_y + \alpha_x|1\rangle_x|0\rangle_y + \alpha_y|0\rangle_x|0\rangle_y. \tag{7.15}$$

The $|\alpha_i|^2$ terms are necessary to maintain normalization. Now we expand the propagator to second-order in γ (which is the same as taking dt small) to get

$$\hat{U} = \exp[-i\hat{H}dt/\hbar] \tag{7.16}$$

$$= \exp[-i(\hat{\gamma}_0\hat{S}_0 + \hat{\gamma}_x\hat{S}_x + \hat{\gamma}_y\hat{S}_y + \hat{\gamma}_z\hat{S}_z)] \tag{7.17}$$

$$\approx 1 - i(\hat{\gamma}_0\hat{S}_0 + \hat{\gamma}_x\hat{S}_x + \hat{\gamma}_y\hat{S}_y + \hat{\gamma}_z\hat{S}_z)$$
$$- \frac{1}{2}(\hat{\gamma}_0\hat{S}_0 + \hat{\gamma}_x\hat{S}_x + \hat{\gamma}_y\hat{S}_y + \hat{\gamma}_z\hat{S}_z)^2. \tag{7.18}$$

Combining these two expansions and using the $\hat{S}_i$ and the properties of coherent states from chapter 3, we can show that

$$\hat{U}|\Phi\rangle_L \approx \left[(1 - |\alpha_x|^2/2 - |\alpha_y|^2/2)|0\rangle_x|0\rangle_y + \alpha_x|1\rangle_x|0\rangle_y + \alpha_y|0\rangle_x|0\rangle_y\right] \tag{7.19}$$

$$-i\hat{\gamma}_0\left[\frac{\alpha_x}{2}|1\rangle_x|0\rangle_y + \frac{\alpha_y}{2}|0\rangle_x|1\rangle_y\right] \tag{7.20}$$

$$-i\hat{\gamma}_x\left[\frac{-\alpha_x}{2}|1\rangle_x|0\rangle_y + \frac{\alpha_y}{2}|0\rangle_x|1\rangle_y\right] \tag{7.21}$$

$$-i\hat{\gamma}_y\left[\frac{-\alpha_{x'}}{2}|1\rangle_{x'}|0\rangle_{y'} + \frac{\alpha_{y'}}{2}|0\rangle_{x'}|1\rangle_{y'}\right] \tag{7.22}$$

$$-i\hat{\gamma}_z\left[\frac{-\alpha_-}{2}|1\rangle_-|0\rangle_+ + \frac{\alpha_+}{2}|0\rangle_-|1\rangle_+\right] \tag{7.23}$$

$$+\frac{1}{2}\hat{\gamma}_0^2\left[\frac{\alpha_x}{4}|1\rangle_x|0\rangle_y + \frac{\alpha_y}{4}|0\rangle_x|1\rangle_y\right] \tag{7.24}$$

$$+\frac{1}{2}\hat{\gamma}_x^2\left[\frac{\alpha_x}{4}|1\rangle_x|0\rangle_y + \frac{\alpha_y}{4}|0\rangle_x|1\rangle_y\right] \tag{7.25}$$

$$+\frac{1}{2}\hat{\gamma}_y^2\left[\frac{\alpha_{x'}}{4}|1\rangle_{x'}|0\rangle_{y'} + \frac{\alpha_{y'}}{4}|0\rangle_{x'}|1\rangle_{y'}\right] \tag{7.26}$$

$$+\frac{1}{2}\hat{\gamma}_z^2\left[\frac{\alpha_-}{4}|1\rangle_-|0\rangle_+ + \frac{\alpha_+}{4}|0\rangle_-|1\rangle_+\right] \tag{7.27}$$

$$+12 \text{ more terms.} \tag{7.28}$$

Now we insert this into equation (7.12) and expand while minding the fact that the $\hat{\gamma}_i$ terms do not necessarily commute. After some amount of work, we arrive at the following

master equation for the atoms alone (without the A subscript):

$$
\begin{aligned}
\hat{\rho}(t + dt) &= \hat{\rho}(t) - \frac{i}{\hbar}[\langle \hat{H} \rangle_L, \hat{\rho}]dt \\
&\quad + \sum_{i=0,x,y,z} dt \frac{\langle \hat{S}_0 \rangle}{2dt} \mathcal{D}[\hat{\gamma}_i] \hat{\rho} \\
&\quad + \sum_{i=x,y,z} dt \frac{\langle \hat{S}_i \rangle}{2dt} (\hat{\gamma}_i \hat{\rho} \hat{\gamma}_0 + \hat{\gamma}_0 \hat{\rho} \hat{\gamma}_i \\
&\quad - \frac{1}{2}((\hat{\gamma}_0 \hat{\gamma}_i + \hat{\gamma}_i \hat{\gamma}_0)\hat{\rho} + \hat{\rho}(\hat{\gamma}_0 \hat{\gamma}_i + \hat{\gamma}_i \hat{\gamma}_0)) \\
&\quad + i(\hat{\gamma}_{i+2} \hat{\rho} \hat{\gamma}_{i+1} - \hat{\gamma}_{i+1} \hat{\rho} \hat{\gamma}_{i+2}) + \frac{i}{2}\{[\hat{\gamma}_{i+2}, \hat{\gamma}_{i+1}], \hat{\rho}\})
\end{aligned} \tag{7.29}
$$

where the semiclassical Hamiltonian (with expectations taken over the light) is defined as

$$
\langle H \rangle_L = \frac{\hbar}{dt} \left(\hat{\gamma}_0 \langle \hat{S}_0 \rangle + \hat{\gamma}_x \langle \hat{S}_x \rangle + \hat{\gamma}_y \langle \hat{S}_y \rangle + \hat{\gamma}_z \langle \hat{S}_z \rangle \right). \tag{7.30}
$$

The master equation of equation (7.29) is the primary result of this chapter. Note that, because the $\hat{\gamma}_i$ do not commute, there are several cross terms that can evolve the state in nontrivial ways. We examine the effect of these cross-terms on the collective state for various configurations in the next section.

Under the limit of large detuning, the master equation becomes considerably simplified. In this case, $c_2/c_1 \to 0$ as $\Delta \gg 300$ MHz $\approx \Delta_{HFS}$. In other words, the tensor terms vanish in the limit that the detuning is larger than the hyperfine splitting as discussed in chapter 6. We can further simplify by ignoring the global phase shifting $\hat{\gamma}_0$ terms and assume linearly polarized input light, to get

$$
\hat{H} = \frac{\hbar}{dt} \hat{\gamma}_z \hat{S}_z \tag{7.31}
$$

and

$$
\hat{\rho}(t + dt) = \hat{\rho}(t) + dt \frac{\langle \hat{S}_0 \rangle}{2dt} \mathcal{D}[\hat{\gamma}_z] \hat{\rho}. \tag{7.32}
$$

This is a familiar unconditional master equation that unconditionally antisqueezes the initial coherent spin-state along x. The conditional version of this master equation produces squeezing as discussed in chapter 8 and chapter 9, where a more intuitive notation is used for the prefactors.

A major point of this chapter is that typical experiments in this field often operate at detunings that are small or comparable to the hyperfine splittings for purely technical reasons. In this case, the above "large detuning" simplification of the master equation does not hold and the dynamics are expected to be complicated by the nonnegligible tensor terms. Many experiments including our own have been performed at detunings around 1 GHz where the tensor terms cannot be fully neglected [37, 36, 29].

7.2 Moment Evolution

In this section, we consider the evolution of various spin moments due to the full master equation, equation (7.29). We cannot analytically evolve the state beyond the first timestep because the master equation populates more than just the fully symmetric subspace. Regardless, we show that unconditional spin-squeezing does occur in this first timestep under certain circumstances. While analyzing the collective uncertainty moments in this section, keep in mind that the expressions contain both incoherently added single-atom effects (e.g., from figure 5.2) and truly collective effects resulting from the atoms residing at the same location.

For initial coherent spin-state (CSS) aligned along the x-direction, the time zero derivatives of the collective variances are given by

$$
\begin{aligned}
\langle \dot{\hat{F}}_x \rangle|_{t=0} &= \frac{\langle \hat{S}_0 \rangle}{2dt}(-126Nc_2^2 - 2Nc_1^2) \\
&\quad + \frac{\langle \hat{S}_x \rangle}{2dt}(-28Nc_2^2 + 28Nc_1c_2) \qquad (7.33) \\
\langle \dot{\hat{F}}_y^2 \rangle|_{t=0} &= \frac{\langle \hat{S}_0 \rangle}{2dt}(98Nc_2^2 + (16N^2 + 30N)c_1^2) \\
&\quad + \frac{\langle \hat{S}_x \rangle}{2dt}(-112Nc_1c_2 - 112N(N-1)c_1c_2) \qquad (7.34) \\
\langle \dot{\hat{F}}_z^2 \rangle|_{t=0} &= \frac{\langle \hat{S}_0 \rangle}{2dt}((-616N + 784N^2)c_2^2) \\
&\quad + \frac{\langle \hat{S}_x \rangle}{2dt}(168Nc_2^2 - 112N(N-1)c_1c_2). \qquad (7.35)
\end{aligned}
$$

The equation for $\langle \dot{\hat{F}}_y^2 \rangle|_{t=0}$ is intentionally left unsimplified for the sake of comparison to the

equation for $\langle \dot{\hat{F}}_z^2 \rangle|_{t=0}$. We also have the mean moment

$$
\begin{aligned}
\langle \dot{\hat{F}}_z \rangle|_{t=0} &= \frac{\langle \hat{S}_y \rangle}{2dt}(28Nc_2) \\
&\quad + \frac{\langle \hat{S}_z \rangle}{2dt}(490Nc_2^2 - 392N(N-1)c_2^2),
\end{aligned} \tag{7.36}
$$

which is consistent with section 6.7. Now we consider the following specializations of these equations using a large number $(N \gg 1)$.

7.2.1 Case 1: Near Resonance, Perpendicular Polarizations

Consider a detuning close to resonance $(c_1/c_2 = 11)$, and with an optical polarization along the y-axis $(\langle \hat{S}_x \rangle = +\langle \hat{S}_0 \rangle)$. Even though the expressions below contain only the vector coefficient c_1 for the sake of comparison, they are tensor in origin because we have assumed $c_2 = c_1/11$. This gives the following moment derivatives.

$$
\langle \dot{\hat{F}}_x \rangle|_{t=0} = \frac{\langle \hat{S}_0 \rangle}{2dt}c_1^2 N(-0.73) \tag{7.37}
$$

$$
\langle \dot{\hat{F}}_y^2 \rangle|_{t=0} = \frac{\langle \hat{S}_0 \rangle}{2dt}c_1^2 N^2(5.8) \tag{7.38}
$$

$$
\langle \dot{\hat{F}}_z^2 \rangle|_{t=0} = \frac{\langle \hat{S}_0 \rangle}{2dt}c_1^2 N^2(-3.7). \tag{7.39}
$$

7.2.2 Case 2: Near Resonance, Parallel Polarizations

Again consider a detuning close to resonance $(c_1/c_2 = 11)$, but with an optical polarization along the x-axis $(\langle \hat{S}_x \rangle = -\langle \hat{S}_0 \rangle)$. Even though the expressions below contain only the vector coefficient c_1 for the sake of comparison, they are tensor in origin because we have assumed $c_2 = c_1/11$. This gives the following moment derivatives.

$$
\langle \dot{\hat{F}}_x \rangle|_{t=0} = \frac{\langle \hat{S}_0 \rangle}{2dt}c_1^2 N(-5.36) \tag{7.40}
$$

$$
\langle \dot{\hat{F}}_y^2 \rangle|_{t=0} = \frac{\langle \hat{S}_0 \rangle}{2dt}c_1^2 N^2(26.2) \tag{7.41}
$$

$$
\langle \dot{\hat{F}}_z^2 \rangle|_{t=0} = \frac{\langle \hat{S}_0 \rangle}{2dt}c_1^2 N^2(16.6). \tag{7.42}
$$

7.2.3 Case 3: Far-Off Resonance, Polarization Independent

Finally, consider a detuning far from resonance ($c_1/c_2 = \infty$), in which case the polarization angle does not matter. Because we are far-off resonance, the tensor Hamiltonian is of no consequence. The moment derivatives are then

$$\langle \dot{\hat{F}}_x \rangle|_{t=0} = \frac{\langle \hat{S}_0 \rangle}{2dt} c_1^2 N(-2) \tag{7.43}$$

$$\langle \dot{\hat{F}}_y^2 \rangle|_{t=0} = \frac{\langle \hat{S}_0 \rangle}{2dt} c_1^2 N^2(16) \tag{7.44}$$

$$\langle \dot{\hat{F}}_z^2 \rangle|_{t=0} = \frac{\langle \hat{S}_0 \rangle}{2dt} c_1^2 N^2(0). \tag{7.45}$$

The second line indicates the expected unconditional antisqueezing.

7.3 Unconditional Squeezing via Collective Tensor Terms

At large detunings (case 3), the initial behavior of the antisqueezed component is:

$$\langle \hat{F}_y^2 \rangle|_t = \langle \hat{F}_y^2 \rangle|_{t=0} \left(1 + \frac{t}{\tau_{ss}} \frac{f}{4} \mathrm{OD} \right) \tag{7.46}$$

where we define the *spin-squeezing* time τ_{ss} as

$$\tau_{ss}^{-1} = \frac{\langle \dot{\hat{F}}_y^2 \rangle|_{t=0}}{\langle \hat{F}_y^2 \rangle|_{t=0}} \tag{7.47}$$

$$= \frac{2I\sigma_0}{\hbar\omega} \left(\frac{\Gamma}{4} \sum_{f'} \frac{\alpha_{f,f'}^{(1)}}{\alpha_0 \Delta_{f,f'}} \right)^2. \tag{7.48}$$

From the analysis of chapter 8 we expect that the conditional value of the z-component will evolve at a similar rate, but be squeezed rather than antisqueezed

$$\langle \hat{F}_z^2 \rangle|_t = \langle \hat{F}_z^2 \rangle|_{t=0} \frac{1}{1 + \frac{t}{\tau_{ss}} \frac{f}{4} \mathrm{OD}\eta}. \tag{7.49}$$

With these assumptions, for unity detection efficiency, $\eta = 1$, the minimum uncertainty product remains initially constant

$$\langle \hat{F}_y^2 \rangle|_t \langle \hat{F}_z^2 \rangle|_t = \langle \hat{F}_y^2 \rangle|_{t=0} \langle \hat{F}_z^2 \rangle|_{t=0}. \tag{7.50}$$

As discussed in chapter 3, the spin-squeezing parameter is defined as

$$\xi^2 = \frac{2F\langle \Delta \hat{F}_z^2 \rangle}{\langle F_x \rangle^2} \tag{7.51}$$

and its rate of change at time zero is given by

$$\frac{d(\xi^2)}{dt} = \frac{d}{dt}\frac{2F\langle \Delta \hat{F}_z^2 \rangle}{\langle \hat{F}_x \rangle^2} \tag{7.52}$$

$$= 2F\left(\frac{d\langle \Delta \hat{F}_z^2 \rangle/dt}{\langle \hat{F}_x \rangle^2} - 2\frac{\langle \Delta \hat{F}_z^2 \rangle}{\langle \hat{F}_x \rangle^3} d\langle \hat{F}_x \rangle/dt \right). \tag{7.53}$$

For Case 3 (with conditional squeezing) the evolution of the spin-squeezing parameter is given by

$$\frac{d(\xi^2)}{dt}\Big|_0 = \frac{\langle \hat{S}_0 \rangle}{dt}\frac{c_1^2}{f}(-16N + 2) \tag{7.54}$$

in accordance with the results of the next chapter.

Now consider the evolution of the squeezing parameter due to the unconditional evolution in case 1 and 2. In case 2 (unconditional, close to resonance, parallel light), all of the perpendicular variances grow initially, thus the state gets antisqueezed in all directions. However, in case 1 (unconditional, close to resonance, perpendicular light), we get a reduction of the z variance ($\langle \dot{\hat{F}}_z^2 \rangle|_{t=0} < 0$) indicating the preparation of a spin-squeezed state. Here, the evolution of the spin-squeezing parameter in the initial timestep is given by

$$\frac{d(\xi^2)}{dt}\Big|_0 = \frac{\langle \hat{S}_0 \rangle}{dt}\frac{c_1^2}{f}(-3.7N + 7.2). \tag{7.55}$$

The question then becomes, where does this unconditional spin-squeezing come from? First, we note that this is not the result of adding "individually squeezed spins." This is mentioned because we are dealing with spins larger than spin-1/2 and the notion of a individually squeezed spin is indeed possible. Further, the single atom quadratic tensor Hamiltonian is comparable the twisting Hamiltonians used to create spin-squeezing as discussed in chapter 3. However, when calculating the above expression, we see that the squeezing is generated by cross-terms in the master equation equation (7.29), and is thus a truly collective effect. The question is then whether the assumptions leading to this anal-

ysis, namely that all of the atoms symmetrically couple to the same optical mode, is valid or not. A more realistic consideration of a beam propagating through a spatially extended atomic sample, where the assumption is not valid, is considered in [26].

Interestingly, we see that the rate of squeezing for the unconditional case is a factor of $0.23 = 3.7/16$ smaller than what is expected in the conditional case discussed in the next chapter.

Chapter 8

Polarimetry and Quantum Measurement

In previous chapters, we have described general quantum measurement concepts (Chapters 1–2) and unconditional descriptions of the probe-atom interaction without measurement (Chapters 3–7). Now we begin to put the two together to consider how the measurement of the probe beam allows us to conditionally evolve the collective quantum spin-state of the atom cloud. A proper description of this quantum estimation process will allow us to understand the back-action of the measurement and conditional preparation of spin-squeezed states.

We begin by describing some of the basic concepts behind the detection of optical polarization rotations with an emphasis on signal-to-noise ratios ($\mathcal{SNR}$). The concept of shotnoise in a polarimeter is introduced prior to the conditioning section because it is the fundamental source of noise that the quantum filter estimating the atomic state must fight against. Then we describe the derivation of the stochastic master equation telling us how to conditionally update the collective state of the atoms via the inherently noisy measurement record.

8.1 Polarimetry

In section 3.2.7, we discussed how the Stokes components of an optical beam can be measured with a polarimeter. Here we discuss in more detail the noise processes associated with polarimetry to better understand the nature of the conditional dynamics described later.

The input stage of a generalized polarimeter consists of a set of optical waveplates (one

quarter-waveplate and one half-waveplate), which together can enact a general rotation on the Stokes sphere as described in section 3.2.5. After passing through the waveplates, the beam is split with a polarizing beamsplitter (PBS) and each beam is then detected by a photodiode. The photodiode currents are subtracted and amplified to give the output of the polarimeter, which is proportional to the measured Stokes component. See section 11.6 for a description of the actual components used in the experiment.

8.1.1   Optical Shotnoise

Consider a photodiode absorbing a continuous power P. The power leads directly to a mean photocurrent of value

$$I = RP. \tag{8.1}$$

The responsivity of a photodiode is defined as

$$R = \frac{\eta e}{\hbar \omega} \, [\text{A/W}] \tag{8.2}$$

where η is the quantum efficiency of the detector, e is the electron charge, and $\hbar\omega$ is the single photon energy. If the quantum efficiency is equal to unity, one can think of each photon as being converted into an electron that contributes to the current. In a measurement of the current, there will be random fluctuations in the signal due to the random uncorrelated arrival times of the electrons, described by a Poisson process. In a given time τ, one can show via the properties of the Poisson process (see `http://qwiki.caltech.edu/wiki/` `Shot_Noise` or [108]) that the RMS fluctuation is related to the mean current by

$$i_{rms}^2 = \frac{eI}{\tau}. \tag{8.3}$$

For an averaging filter with time constant τ, the frequency interval is defined as $\Delta f = 1/2\tau$ (note the factor of two) [108]. The optical shotnoise floor is then

$$
\begin{aligned}
i_{rms} &= \sqrt{2eI} \left[\text{A}/\sqrt{\text{Hz}} \right] \sqrt{\Delta f} \tag{8.4} \\
&= \sqrt{2e^2 \eta P/\hbar\omega} \left[\text{A}/\sqrt{\text{Hz}} \right] \sqrt{\Delta f}. \tag{8.5}
\end{aligned}
$$

This is consistent with the noise one expects to get from performing photon detection on an optical coherent state [109, 68]. The goal in the design of the photodiode detector is to get the optical shotnoise floor significantly larger than the electronic noise floor of the detector for the power the experiment demands. See chapter 11 for a discussion of the detector used in our experiment. Note that when representing the optical shotnoise as a Wiener process we use

$$I(t)dt = Idt + \sqrt{eI}dW(t) \tag{8.6}$$

where I is a constant representing the mean current. This expression reproduces the relations above when calculating the RMS current using Itô's rule [52].

8.1.2 Polarimeter Unravellings

After the the probe beam interacts with the atoms, one can imagine measuring the probe beam in many different ways, each of which leads to a different *unravelling* of the conditional dynamics as discussed in chapter 2. Note that here we consider the case of an input linearly polarized probe beam of constant power and only change the measurement basis, although we could also imagine adaptively changing the probe beam state for some purpose.

As discussed in section 3.2.7, one can measure any Stokes component with a polarimeter (two input waveplates, a PBS, and a subtracting photodetector). The standard balanced polarimetry measurement is where the linearly polarized probe beam is placed at 45 degrees with respect to the axes of a polarizing beamsplitter (PBS) and half of the light is put on each detector. Any rotations of the linear polarized light that keep it linear will then be measured by the difference output of the photodetectors. In terms of Stokes components, we say that the input state maximizes $\hat{S}_x$ and we measure $\hat{S}_y$.

As another example, the polarimeter can be operated in a completely "imbalanced" configuration, with the input state at $\hat{S}_x$ and measuring $\hat{S}_x$. In this case the analyzer PBS is oriented at 45 degrees to the balanced configuration and all of the noninteracted light goes to one of the detectors. The other detector with less light can then be monitored, possibly in photon-counting configuration, for a signal from the atoms. Note that in this configuration, the detector cannot possibly distinguish which way (cw or ccw) the linear polarization was rotated when a nonzero signal is detected. As a result of this indistinguishability, the conditional dynamics under this unravelling must respect this lack of knowledge, leading to

the possibility of conditional cat state production [90].

As a final possible detection scheme, we mention that the rotations of a linearly polarized probe beam can considered as coming from a differential phase shift between the circular basis components composing the linearly polarized light. Thus one could also do a measurement by physically separating these two components (e.g., with a quarter-wave plate and a PBS) and then doing a homodyne measurement on each of the components. The signal-to-noise ratio from this scheme is compared to the balanced polarimetry scheme in the next section.

8.1.3 $\mathcal{SNR}$ Comparisons

Consider two separate schemes for measuring the polarization rotation of a single beam. In both, we initialize the input beam polarization along $\hat{S}_y$ on the Stokes sphere such that we can represent its electric field as

$$\mathbf{E}_i = \frac{E_0}{\sqrt{2}}(\mathbf{e}_x + \mathbf{e}_y). \tag{8.7}$$

We temporarily ignore some constant prefactors and represent the power simply as $P = |\mathbf{E}_i|^2 = E_0^2$. Now suppose that something rotates the polarization by a small angle $\phi \ll 1$ such that the output field is

$$\mathbf{E} = \frac{E_0}{\sqrt{2}}(\mathbf{e}_x(1 + \phi) + \mathbf{e}_y(1 - \phi)). \tag{8.8}$$

With this output light, we now consider two different measurement schemes and compare their signal-to-noise ratios.

8.1.3.1 Balanced Polarimetry

In the balanced polarimetry configuration, we measure the two powers P_x and P_y and then subtract to get the signal

$$P_x = \frac{E_0^2}{2}(1 + \phi)^2 \tag{8.9}$$

$$P_y = \frac{E_0^2}{2}(1 - \phi)^2 \tag{8.10}$$

$$\Delta P_P = P_x - P_y = 2P\phi. \tag{8.11}$$

We then represent the optical shotnoise as being the square root of total power measured (again ignoring constant prefactors)

$$\zeta_P = \sqrt{P} = |E_0|.$$ (8.12)

8.1.3.2 Double Homodyne

Now we decompose the output light into its circular components

$$\mathbf{E} = E_+\mathbf{e}_+ + E_-\mathbf{e}_-,$$ (8.13)

which can be represented as

$$E_\pm = \frac{E_0}{2}(i(1-\phi) \mp (1+\phi)).$$ (8.14)

After splitting these components apart, with the quarter-wave plate and PBS, we perform a homodyne measurement on each of the components, i.e., we mix one component with a strong local oscillator E_{LO} at a 50/50 beamsplitter and measure the power difference of the outputs. After mixing the local oscillator with E_- we have

$$E_A = (E_{LO} + E_-)/\sqrt{2}$$ (8.15)

$$E_B = (E_{LO} - E_-)/\sqrt{2}$$ (8.16)

and

$$\Delta P_- = P_A - P_B = E_{LO}E_0(1+\phi),$$ (8.17)

which gives the total signal. When we take the $P_{LO} \gg P$ limit we have for the overall shotnoise

$$\zeta_{P_-} = E_{LO}.$$ (8.18)

We also get the same signal from the other homodyne setup (measuring +)

$$\Delta P_+ = E_{LO}E_0(1-\phi)$$ (8.19)

and the same noise

$$\zeta_{P_+} = E_{LO} \tag{8.20}$$

assuming we used the same power local oscillator. To extract the total signal we subtract the two homodyne outputs to get

$$\Delta P_H = \Delta P_- - \Delta P_+ = E_{LO}E_0 2\phi \tag{8.21}$$

and add the noises in quadrature to get

$$\zeta_H = \sqrt{2}|E_0|. \tag{8.22}$$

8.1.3.3 Polarimetry versus Double Homodyne

Comparing the signal-to-noise ratios of the above two analyses we get that the polarimetry signal-to-noise is a factor of $\sqrt{2}$ greater than the double homodyne measurement

$$\mathcal{SNR}_P = \Delta P_P/\zeta_P = 2|E_0|\phi \tag{8.23}$$

$$\mathcal{SNR}_H = \Delta P_H/\zeta_H = \sqrt{2}|E_0|\phi. \tag{8.24}$$

This difference of $\sqrt{2}$ can be explained by the fact that the local oscillator adds unnecessary noise to the measurement. If squeezed local oscillators were used, the $\mathcal{SNR}$ of the polarimetry measurement could be recovered. Also note that if heterodyne measurements were performed on the $\pm$ state, then these would be another $\sqrt{2}$ worse than the homodyne performance [11, 12].

8.1.4 Polarimeter "Amplification"

We end this section by considering a possibly convenient trick, which can be used to reduce the total power detected by the polarimeter while keeping the same signal-to-noise ratio. A major source of technical noise in a balanced polarimeter is offset fluctuations, which are proportional to the total power into the polarimeter. These can be caused, for example, by imperfect balancing. Here we briefly analyze an "amplification" scheme that involves selectively damping the polarization component that does not get measured, as discussed in [110].

Consider a polarimeter that measures the power along axes x and y and with input light of power P aligned mostly along $\mathbf{e}_{x'} = (\mathbf{e}_x + \mathbf{e}_y)/\sqrt{2}$. Define θ as the angle of the state away from $\mathbf{e}_{x'}$. The polarimeter output is then

$$P_x - P_y = P\sin(2\theta) + \zeta(P) \tag{8.25}$$

where the shotnoise of power P is represented as $\zeta(\hat{P}) = \sqrt{\hbar\omega P}\,dW/dt$. Now imagine selectively damping the $\mathbf{e}_{x'}$-component by a factor of $\chi < 1$ such that

$$\mathbf{E} = E(\cos(\theta)\mathbf{e}_{x'} + \sin\theta\,\mathbf{e}_{y'}) \rightarrow \tag{8.26}$$

$$\tilde{\mathbf{E}} = E(\chi\cos(\theta)\mathbf{e}_{x'} + \sin(\theta)\mathbf{e}_{y'}). \tag{8.27}$$

Now the new polarimeter output is

$$\tilde{P}_x - \tilde{P}_y = \chi P\sin(2\theta) + \zeta(\tilde{P}) \tag{8.28}$$

where $\tilde{P} = P(\chi^2\cos^2(\theta) + \sin^2(\theta))$. Comparing the signal-to-noise ratios assuming $\theta \ll 1$ we have

$$\mathcal{SNR} = \frac{P_x - P_y}{\zeta(P)} \approx 2\theta\sqrt{\frac{P}{2\hbar\omega}} \tag{8.29}$$

$$\widetilde{\mathcal{SNR}} = \frac{\tilde{P}_x - \tilde{P}_y}{\zeta(\tilde{P})} \approx 2\theta\sqrt{\frac{P}{2\hbar\omega}}\sqrt{\frac{\chi^2}{\chi^2 + \theta^2(1 - \chi^2)}}. \tag{8.30}$$

So if we have $\chi^2 \gg \theta^2$ then we have $\mathcal{SNR} \approx \widetilde{\mathcal{SNR}}$. Given that the polarization rotation is going to be small, this technique can be practically useful if, for example, the polarimeter subtraction is not perfect and there are unwanted power fluctuations. This erroneous signal would then go down as χ^2 while the desired signal only goes down as χ. This effect can also be useful because in practice detectors often become nonlinear if they absorb too much power.

One large caveat to this approach is that the electronic noise floor does not get decreased at all, thus the relative noise-equivalent-power (NEP) of the detector will go up by a factor of $1/\chi^2$ and the detector becomes less optical shotnoise limited compared to the power of the total probe light. If there is enough of an initial cushion between optical and the electronic

noise floors, then this should not be a problem, but in practice the electronic noise floor is never going to be zero.

8.2 Conditional Dynamics

Given the Hamiltonian interaction and a subsequent detection of the probe via the polarimeter, how do we describe the conditional evolution of the atomic spin-state? We discussed the general problem of conditional evolution in chapter 2 and an outline of the needed treatment from a more mathematical perspective is provided in [23]. Here we begin by discussing the stochastic master equation derived in slightly different context. Then we discuss the measurement signal-to-noise ratio in our experiment and relate that to the master equation parameterization. From this we physically derive the rate one expects to prepare spin-squeezed states via QND measurement with real alkali atoms. We end this chapter by considering the functional dependence of the spin-squeezing and decay timescales on detuning.

8.2.1 The SME with Homodyne Detection

One of the fundamental ideas presented in this thesis was introduced in a paper by Thomsen, Mancini, and Wiseman [15, 16]. In that paper, they derived the stochastic master equation (SME) describing the conditional state of an atomic ensemble within a cavity, when the output of the cavity is monitored with homodyne detection. This model contains the basic idea for the rest of this thesis, yet it is different from our experiment in several respects. The model considers the output of a cavity (whose internal state is adiabatically eliminated) while our experiment is currently performed in free-space. Furthermore, the atom-light interaction Hamiltonian (and level scheme) is different and leads to an analysis of homodyne detection as opposed to our balanced polarimeter detection. As a result of these differences, the derivation of the stochastic master equation for our experiment is expected to be quite different and, as yet, this analysis has not been performed.

As introduced briefly in section 2.4, the derivation of [15, 16] results in the stochastic

master equation

$$d\hat{\rho}(t) = -i[\hat{H}_S, \hat{\rho}(t)]dt + \mathcal{D}[\sqrt{M}\hat{F}_z]\hat{\rho}(t)dt$$

$$+\sqrt{\eta}\mathcal{H}[\sqrt{M}\hat{F}_z]\hat{\rho}(t)\left(2\sqrt{M\eta}[y(t)dt - \langle\hat{F}_z\rangle dt]\right) \tag{8.31}$$

where $\hat{H}_S$ is the system Hamiltonian, η is the quantum detection efficiency, and M is the measurement strength, which is physically derived in the next section for our experimental scheme, but left general here. The superoperators, $\mathcal{D}$ and $\mathcal{H}$, are defined as

$$\mathcal{D}[\hat{c}]\hat{\rho} \equiv \hat{c}\hat{\rho}\hat{c}^\dagger - (\hat{c}^\dagger\hat{c}\hat{\rho} + \hat{\rho}\hat{c}^\dagger\hat{c})/2 \tag{8.32}$$

$$\mathcal{H}[\hat{c}]\hat{\rho} \equiv \hat{c}\hat{\rho} + \hat{\rho}\hat{c}^\dagger - \mathrm{Tr}[(\hat{c} + \hat{c}^\dagger)\hat{\rho}]\hat{\rho}, \tag{8.33}$$

and the photocurrent is represented as

$$y(t)dt = \langle\hat{F}_z\rangle(t)dt + dW(t)/2\sqrt{M\eta}. \tag{8.34}$$

The stochastic quantity $dW(t) \equiv 2\sqrt{M\eta}(y(t)dt - \langle\hat{F}_z\rangle(t)dt)$ is a Wiener increment and $dW(t)/dt$ is a Gaussian white noise associated with the shotnoise of the homodyne local oscillator.

As mentioned in section 2.4, there are several intuitive features to notice about this equation. First, if $M = 0$ we recover the usual closed-system equation of motion for the state. If $M \neq 0$, but the quantum efficiency of the detector $\eta = 0$, then we get the unconditional evolution of the state due to the measurement. If $M \neq 0$ and $\eta \neq 0$, then the update rule is an explicit function of the photocurrent $y(t)$. This conditional part of the equation is reminiscent of classical estimators that take the difference between what the photocurrent is and what the photocurrent is expected to be, to get an *innovation* $dW(t)$ by which the system is updated. In fact, this stochastic master equation (or stochastic Schrodinger equation, SSE) is one of the most general mathematical examples of continuous quantum measurement and is discussed in many different contexts [71, 72, 23]. We discuss its properties in more detail in chapter 9.

8.2.2 Deriving the Measurement Strength from the Polarimeter SNR

The SME above is the filter that essentially describes our polarimetry experiment (far-off resonance) although we have not derived it in this context. We avoid the derivation, and identify the physical measurement strength with our experimental parameters, by using results from previous chapters. In the end, it is the form of M we are interested in, but we are also interested in the validity of the SME as a model. There are two reasons that it will not be an exact model. First, if the detuning is not large enough, the tensor terms will destroy the description, as discussed in the unconditional case in chapter 7. Second, we only consider the SME as a valid description for times less than the spontaneous emission time as derived in chapter 5. We end this chapter with a discussion related to these crudely modeled, but real, effects, although in the next few chapters we largely ignore them for pedagogical reasons.

(Update: Led by Luc Bouten, our group has recently finished a paper in which the polarimetry stochastic master equation is derived and shown to be as assumed in this thesis [111]. The paper uses a formal quantum stochastic differential equation (QSDE) formalism and also shows that the scheme is equivalent to performing homodyne detection on the y channel when the probe beam is initially polarized along x. This imbalanced homodyne detection case is as opposed to the balanced homodyne case considered earlier in this chapter.)

In chapter 7, equation (7.32), we showed that the (ignoring the tensor Hamiltonian) the unconditional master equation is simply

$$d\hat{\rho} = \frac{\langle \hat{S}_0 \rangle}{2dt} \mathcal{D}[\hat{\gamma}_z]\hat{\rho}dt \qquad (8.35)$$

$$= \mathcal{D}[\sqrt{M}\hat{F}_z]\hat{\rho}dt. \qquad (8.36)$$

Here, we have now physically identified the measurement strength M in terms of our experimental parameters as

$$M = \frac{1}{8\hbar^2} \tau_{ss}^{-1} \left(\frac{\sigma_0}{A} \right) \qquad (8.37)$$

where we have used the definition, from equation (7.48), of the spin-squeezing time as

$$\tau_{ss}^{-1} = \frac{2I\sigma_0}{\hbar\omega}\left(\frac{\Gamma}{4}\sum_{f'}\frac{\alpha_{f,f'}^{(1)}}{\alpha_0\Delta_{f,f'}}\right)^2. \tag{8.38}$$

Thus, for the $\eta = 0$ case we can identify this measurement strength as having the same function as the measurement strength in equation (8.31). Notice that the measurement strength is a rate when $\hbar$ is assumed unitless. Also notice that the ratio of areas, the cross section divided by the beam area, $\sigma_0/A \propto \lambda^2/A$ can never be greater than one in free-space due to the diffraction limit.

Rather than this indirect approach for deriving the physical measurement strength, we can also identify it through a direct expression for the physical photocurrent in units of power. As shown in section 6.4, the measurement of $\hat{S}_y$ will result in equation (6.87) when atoms are present. Now we wish to include measurement noise in order to calculate the signal-to-noise ratio. It is readily shown that all terms not linear in $\hat{F}_z$ vanish in equation (6.87) provided that the state is aligned along x with $\theta = \pi/2$ and $\phi = 0$. That is, a pure Faraday rotation Hamiltonian is recovered when the atomic magnetization vector is oriented along the x-axis. However, rotating $\hat{F}$ in the xy-plane results in elliptically polarized scattered probe light, and moving out of this plane results in nonlinear atomic dephasing due to scattering terms, which are quadratic in the single-particle spin operators, $\hat{f}_z$. These adverse effects are avoided for the experimental geometry where $\hat{F}$ is collinear with the x-axis.

Taking the input probe field to be in an x-polarized optical coherent state, and considering the small γ limit, equation (6.87) leads to a semiclassical photocurrent (with units of optical power) of the form,

$$y_t = \eta\sqrt{S}\langle\hat{F}_z\rangle + \sqrt{\eta}\,\zeta_t, \tag{8.39}$$

where we have made the substitution, $\hbar N f \cos\theta \to \langle\hat{F}_z\rangle$ (see equation (6.69)), and included the photodetector quantum efficiency, η. We have introduced the *scattering strength S* defined as

$$S = \frac{1}{\hbar^2}\left[I_p\sigma_0\left(\frac{\Gamma}{4}\right)\sum_{f'}\frac{\alpha_{f,f'}^{(1)}}{\alpha_0\Delta_{f,f'}}\right]^2, \tag{8.40}$$

which depends upon the probe intensity, $I_p = P/A$, determined by the coherent state amplitude, $P = 2\hbar\omega|\beta|^2$ and cross sectional area, $A = \pi r^2$ (for a mode-matched probe laser). It is useful to note that the scattering strength has units of $W^2/\hbar^2$ (power squared per $\hbar^2$) and characterizes the degree of coupling between the atoms and the probe field; $\sqrt{S}$ quantifies the polarimeter optical power imbalance per unit spin (as $\hat{F}_z$ has units of $\hbar$). As discussed above, we have represented the optical shotnoise by $\zeta_t = \sqrt{\hbar\omega P}dW_t/dt$ such that $\Delta\zeta^2 = \hbar\omega P$. To convert the above power units to photocurrent units, we simply use the conversion $e/\hbar\omega$, or the responsivity if the detection efficiency has not been included.

Our expressions are similar to previous results [15, 104, 45, 24]. However, our specific expressions for γ_x, γ_y and γ_z, and hence M, account for the detailed hyperfine structure of the atomic excited states, including the fact that the oscillator strengths and signs of the contributions from different participating excited states are not equal, and doing so is required for quantitative agreement between theory and experiment.

Finally, we can identify the physical photocurrent equation (8.39) with the photocurrent from equation (8.34). The measurement strength is then seen to be given by

$$M = \frac{S}{4\Delta\zeta^2},$$ (8.41)

which again matches the above expression for M, in equation (8.37), when expanded.

Now that we have identified the measurement rate, we spend a considerable amount of time in the next two chapters investigating the SME under ideal circumstances without spontaneous emission or tensor complications. There it will be clear exactly how the SME conditionally projects an initial coherent spin-state into a spin-squeezed state. In the next section, we consider the expected degree of squeezing from a much simpler perspective.

8.2.3   Squeezing by Averaging

Now consider a measurement of $\hat{F}_z$ as described above (and displayed in figure 14.1). In the small time limit where probe induced decoherence can be neglected, the full quantum filter describing this measurement is equivalent a classical model in which $\hat{F}_z$ is simply a random constant on every trial drawn from a distribution with variance equal to the quantum variance of $\langle\Delta\hat{F}_z^2\rangle_0$ [20]. Then the generally complicated full quantum filter [23] is equivalent to linear regression, or fitting a constant to the noisy measurement record in real

time. In essence, the optimal filter serves to average away the optical shotnoise to reveal the underlying value of $\hat{F}_z$. Under these statistical assumptions, with the signal-to-noise ratios of the previous section, the quantum uncertainty at small times is given by

$$\langle \Delta \hat{F}_z^2 \rangle_\tau = \frac{\langle \Delta \hat{F}_z^2 \rangle_0}{1 + 4M\eta \langle \Delta \hat{F}_z^2 \rangle_0 \tau}. \tag{8.42}$$

This can be shown either with the full quantum filter or by using the equivalent classical model combined with Bayesian estimation (from which a Kalman filter or linear regression can be derived). In the filtering terminology, the above expression is seen to be the solution to a Riccati equation as discussed in chapter 10 [49].

We now define the signal-to-noise ratio as

$$\text{SNR}^2 \equiv 4M\eta \langle \Delta \hat{F}_z^2 \rangle_0 \tau \tag{8.43}$$

and express the degree of squeezing (ignoring decay of the $\hat{F}_x$) as

$$W \equiv \frac{\langle \Delta \hat{F}_z^2 \rangle_t}{\langle \Delta \hat{F}_z^2 \rangle_0} \tag{8.44}$$

$$= \frac{1}{1 + \text{SNR}^2}. \tag{8.45}$$

Using $\langle \Delta \hat{F}_z^2 \rangle_0 = \hbar^2 N f / 2$, we can express the signal-to-noise ratio as

$$\text{SNR}^2 = \eta \text{OD} \frac{f}{4} \frac{\tau}{\tau_{ss}} \tag{8.46}$$

where the optical depth is $\text{OD} = N\sigma_0/A = \rho L \sigma_0$, if A is the area and L the length of the cloud. (This SNR is implied throughout the rest of this thesis, as opposed to the $\mathcal{SNR}$ definitions used earlier.)

To get the most spin-squeezing in an experiment, we want to make the SNR as large as possible. The quantum efficiency η is typically of order unity for good photodiodes, thus one cannot gain orders of magnitude by considering better detectors. The optical depth should be made as large as possible, but typically cold atom densities and cloud sizes are limited by technical considerations as discussing in chapter 12. Finally, one wants to average for a long time to make the fit to a random offset buried better resolved, thus the ratio τ/τ_{ss} should be as large as possible. Unfortunately, the averaging time is limited both by

technical considerations (e.g., atoms falling out of the trap) and fundamental considerations, because the averaging time needs to be less than the spontaneous emission time, $\tau \ll \tau_{sc}$ for this entire story to remain valid. With a cavity, the measurement time τ_{ss} can be made much smaller than the τ_{sc} and more squeezing can be obtained, but in free-space these two quantities are fixed relative to each other as discussed in the next section.

In appendix E, we consider a filtering model with spontaneous emission, where the ideal filter is *not* equivalent to the averaging filter discussed here. There we analyze exactly how poorly the nonoptimal averaging filter performs in the presence of decay.

8.2.4 Squeezing versus Decay Timescales

Ignoring the tensor Hamiltonian temporarily, it is interesting to consider the ratio of the timescale τ_{sc} from equation (5.46) describing the incoherent decay and the timescale τ_{ss} from above, which is related to the rate at which spin-squeezing occurs. The larger the ratio τ_{sc}/τ_{ss} is, the more squeezing one expects to get. In the following, we assume either an optical polarization parallel or perpendicular to the atomic spin alignment. In figure 8.1, these ratios are plotted as a function of detuning.

First consider the parallel polarization case. When plotting the timescale ratio we see that close to resonance (less than 300 MHz as discussed in chapter 6) the ratio for the cesium D2 line is nearly an order of magnitude higher than far-off resonance. (The perpendicular polarization case displays similar behavior but exhibits a higher overall ratio.) The reason for this behavior is that all of the excited level contributions to τ_{sc} are the same sign (positive), while the contributions to τ_{ss} are alternating in sign. So naïvely one wants to be close to the 4–5' transition, where the negative effect of the 4–4' transition on the measurement strength can be neglected, giving a higher expected ratio of times and hence more spin-squeezing. However, this logic tells us to be close to resonance where all of the tensor terms are large and, as we have seen in chapter 7, the simple spin-squeezing story gets considerably more complicated. Thus one has two options: either go to a large detuning and take the lower ratio of timescales for the sake of simplicity, or try to work close to resonance and either understand or cancel the tensor terms in some way, e.g., [112]. This strategy is discussed more in chapter 14.

In the next chapter, we numerically and analytically investigate the behavior of the ideal SME of equation (8.31). For convenience, we temporarily neglect the spontaneous emission
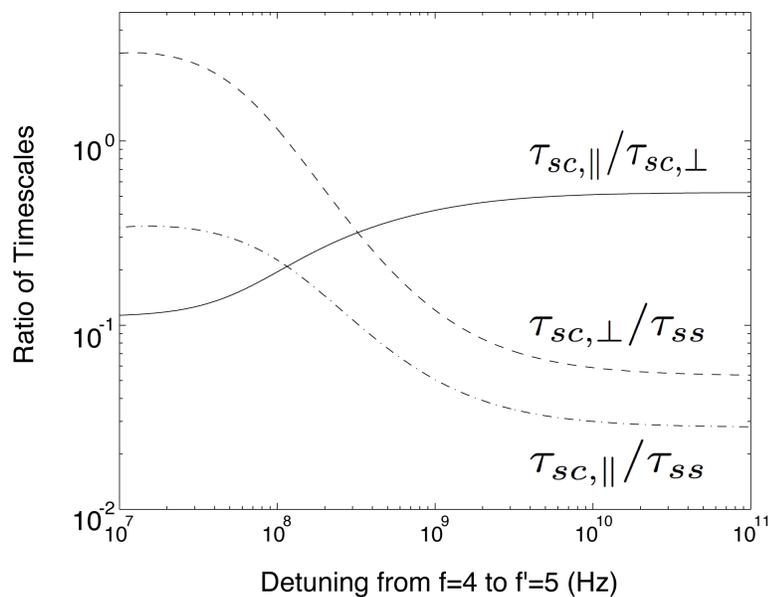
Figure 8.1: Comparisons of the of the spontaneous emission timescales (perpendicular polarizations $\tau_{sc,\perp}$, parallel polarizations $\tau_{sc,\parallel}$) and the spin-squeezing timescale (τ_{ss}) as a function of detuning from the $f = 4$ to $f' = 5$ transition.

and tensor effects.

Chapter 9

Eigenstate Preparation with Measurement and Control

In this chapter, we start with the idealized stochastic master equation of equation (8.31) and analyze it in full detail, with the addition of feedback to make the preparation of $\hat{F}_{\mathrm{z}}$ eigenstates deterministic. This chapter is adapted directly from [21] and is somewhat repetitive of previous chapters in order to be self-contained. However, the two sections section 9.7 and section 9.8 are new and present original analytic results. The idea of this chapter is made considerably more rigorous, albeit for a spin system of only one spin, in [22] where notions of stochastic stability with state-based control laws are introduced. That work and the review [23] also discuss the *constructive* derivation of adequate control laws rather than the intuitive process used in this chapter.

9.1 Abstract

We characterize the long-time projective behavior of the stochastic master equation describing a continuous, collective spin measurement of an atomic ensemble both analytically and numerically. By adding state-based feedback, we show that it is possible to prepare highly entangled Dicke states deterministically.

9.2 Introduction

It has long been recognized that measurement can be used as a *nondeterministic* means of preparing quantum states that are otherwise difficult to obtain. With projective measurements that are truly discrete in time, the only way an experimentalist can direct the

outcome of the measurement is by preparing the initial state to make the desired result most probable. Generally, it is impossible to make this probability equal to one, as the measurement will, with some nonzero probability, result in other undesirable states. If the experimentalist can afford to be patient, then accepting a low efficiency is not a problem, but this is not always the case. In recent years, a theory of continuous quantum measurement has been developed that fundamentally changes the nature of state preparation via measurement [113]. When a measurement and the corresponding acquisition of information are sufficiently gradual, there exists a window of opportunity for the experimentalist to affect the outcome of the measurement by using feedback control [114]. In this chapter, we demonstrate that it is possible to deterministically prepare highly entangled Dicke states [88, 68] of an atomic spin ensemble by adding state-based feedback to a continuous projective measurement.

It has been shown that models of quantum state reduction exist that exhibit the usual rules of projective measurement except the state reduction occurs in a continuous, stochastic manner [71]. These models are not without physical relevance as they are the same as those derived to describe the conditional evolution of atomic spin-states under continuous quantum nondemolition (QND) measurement [15, 16, 115, 116, 117, 118]. By measuring the collective angular momentum operator, $\hat{F}_z$, of an initially polarized coherent spin-state via the phase shift of an off-resonant probe beam, conditional spin-squeezed states can in principle be produced. These states are of considerable interest for applications in quantum information processing and precision metrology [5, 45].

In these models, the reduction in variance that initially leads to conditional spin-squeezing is the precursor of the projection onto a random eigenstate of $\hat{F}_z$ at longer times. Figure 9.1 demonstrates the projection process for a single numerically simulated measurement trajectory.[1]. Like spin-squeezed states, these Dicke states offer potential for quantum information applications because of their unique entanglement properties [17]. Although the experimental difficulties in obtaining these states via QND measurement or other experimental methods [120, 121, 122] are considerable, the details of the continuous projective process that leads to them are of fundamental interest.

Whenever the measurement is sufficiently slow, an experimentalist may steer the result

---

[1]All numerical simulations shown were performed using the parameters $\{N = 10, M = 1, T = 5, dt = 0.001\}$ The stochastic integrator used the norm-preserving, nonlinear SSE of equation (9.5) and a weak second-order derivative-free predictor-corrector structure as can be found in [119].
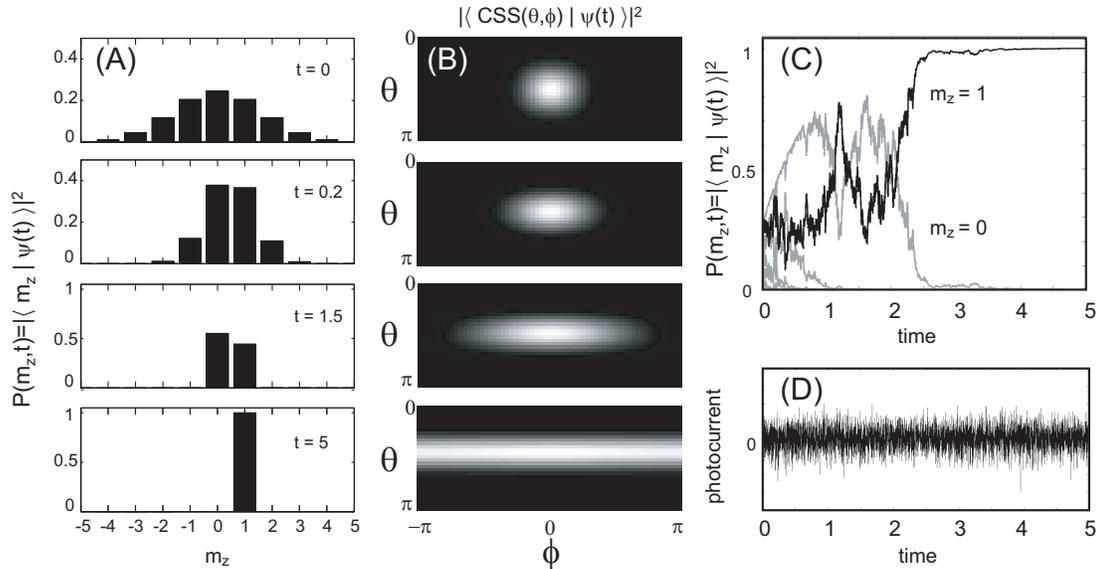
Figure 9.1: The results of a single numerical simulation [36] of the SSE, equation (9.5), with $M = 1$, $\eta = 1$, and $N = 10$ spins initially aligned along the x-axis. (A) In a quantization axis perpendicular to the polarization, the level distribution of a coherent spin-state (CSS) is Gaussian for large N. Under conditional measurement the state evolves at short times into a spin-squeezed state and, eventually, into a random eigenstate of $\hat{F}_z$. (B) A map of the state's angular distribution on the Bloch sphere in spherical coordinates. The uncertainty in the transverse direction to the measurement axis grows until there is no information about the perpendicular component direction. (C) At long times, the population is at most divided among two levels that compete to be the final winner, which in this case appears to be $m = 1$. (D) All of the state information is obtained by properly filtering the noisy photocurrent.

by feeding back the measurement results in real time to a Hamiltonian parameter. Indeed, the measurement process, as a state preparation process, can be made deterministic with the use of feedback control [15, 16]. This is just one example of the growing confluence of quantum measurement with classical estimation and control theory [56, 75]. Other applications of quantum feedback include parameter estimation, metrology, and quantum error correction [18, 20, 123, 12, 124].

In this chapter, we focus on the long-time limit of the QND measurement and feedback process. Just as spin-squeezed states can be deterministically prepared at short times, we numerically demonstrate that individual Dicke states can be deterministically prepared at long times with the use of state-based feedback [125]. While our proposed feedback laws are nonoptimal, they demonstrate the adequacy of intuitive controllers with finite gain for directing the diffusion of the quantum state towards desirable regions of Hilbert space with

unity efficiency. This is in contrast to other proposed schemes using measurement to prepare Dicke states probabilistically [120, 121]. A more systematic approach utilizing stochastic notions of stability and convergence in the continuous measurement and control of a single spin is presented in reference [22].

This chapter is organized as follows. In section 9.3, we introduce the stochastic master equation that represents the rule for updating the system state in real time via the incoming measurement record. Here we discuss the various representations of the dynamics in both the short and long-time limits. Section 9.4 describes the probabilistic preparation of Dicke states via observation alone. The numerical demonstration of the open-loop projection process reveals statistical features that clarify the details of the projection. Feedback is added to the procedure in section 9.5, where we show that state-based control allows one to prepare the same Dicke state deterministically on every measurement. Finally, in section 10.9, we discuss future directions and imminent challenges regarding quantum state preparation via measurement and control.

9.3   Representations of the Conditional Evolution

The physical system we will consider is an ensemble of N spin-1/2 particles contained within a cavity and interacting with a far-off resonant single field mode. We will denote the conditional state of the spin ensemble as $\rho(t)$ and the homodyne measurement record of the output as $y(t)$. The stochastic master equation (SME) describing the conditional evolution is [15, 16]

$$d\rho(t) = -i[H(t), \rho(t)]dt + \mathcal{D}[\sqrt{M}\hat{F}_z]\rho(t)dt$$
$$+\sqrt{\eta}\mathcal{H}[\sqrt{M}\hat{F}_z]\rho(t)\left(2\sqrt{M\eta}[y(t)dt - \langle\hat{F}_z\rangle dt]\right) \tag{9.1}$$

where $H(t) = \gamma\hat{F}_y b(t)$ is the control Hamiltonian that we will allow ourselves (without feedback $b(t) = 0$), γ is the gyromagnetic ratio, M is the probe parameter dependent measurement rate, and

$$\mathcal{D}[\hat{c}]\rho \equiv \hat{c}\rho\hat{c}^\dagger - (\hat{c}^\dagger\hat{c}\rho + \rho\hat{c}^\dagger\hat{c})/2 \tag{9.2}$$

$$\mathcal{H}[\hat{c}]\rho \equiv \hat{c}\rho + \rho\hat{c}^\dagger - \text{Tr}[(\hat{c} + \hat{c}^\dagger)\rho]\rho. \tag{9.3}$$

The (scaled) difference photocurrent is represented as

$$y(t)dt = \langle \hat{F}_z \rangle(t)dt + dW(t)/2\sqrt{M\eta}. \tag{9.4}$$

The stochastic quantity $dW(t) \equiv 2\sqrt{M\eta}(y(t)dt - \langle \hat{F}_z \rangle(t)dt)$ is a Wiener increment and $dW(t)/dt$ is a Gaussian white noise that can be identified with the shotnoise of the homodyne local oscillator. (See [53, 52] for an introduction to stochastic differential equations (SDE).) The sensitivity of the photodetection per $\sqrt{\text{Hz}}$ is represented by $1/2\sqrt{M\eta}$, where the quantity $\eta \in [0,1]$ represents the quantum efficiency of the detection. If $\eta = 0$, we are essentially ignoring the measurement result and the conditional SME becomes a deterministic unconditional master equation. If $\eta = 1$, the detectors are maximally efficient. In this latter case, the conditioned state will remain pure for the entire measurement, thus we can use a state vector description, and the SME can be replaced with a stochastic Schrodinger equation (SSE)

$$
\begin{aligned}
d|\psi(t)\rangle &= (-iH(t) - M(\hat{F}_z - \langle \hat{F}_z \rangle(t))^2/2)|\psi(t)\rangle dt \\
&\quad + \sqrt{M}(\hat{F}_z - \langle \hat{F}_z \rangle(t))|\psi(t)\rangle dW(t).
\end{aligned} \tag{9.5}
$$

This SSE was considered in [71] where the motivation was more abstract and less concerned with the experimental filtering perspective presented here. We emphasize that the SME/SSE is physically derived and is an explicit function of a measured photocurrent variable $y(t)$, through which the randomness enters. The states are considered as states of knowledge and, in practice, an experimentalist updates the description of the system, $\rho(t)$ (figure 9.1A-C), as the measurement results, $y(t)$ (figure 9.1D), arrive in time.

The stochastic master equation of equation (B.43) describes only the dispersive part of the atom-field interaction. Physically, however, any dispersive phase shift must be accompanied by some degree of decohering absorption and spontaneous emission from the auxiliary excited state level(s). Generally, the dispersive SME will be valid until some time, at which point spontaneous emission catches up to destroy the validity of the above description. The resulting cutoff time will impose a limit to the amount of observable squeezing or projection.

In free-space measurements, e.g., free-space Faraday rotation [29, 37, 24, 104], the effects of spontaneous emission make this cutoff time relatively short. By surrounding the atomic

cloud with a cavity however, spontaneous emission can be suppressed and the validity of the SME correspondingly extended.

For a cavity with decay rate κ, N atoms with a decay rate γ, and an atom-cavity coupling constant g, the requirement to see any spin-squeezing is only $g^2/\kappa\gamma > 1/N$. On the edge of the strong-coupling regime, with $g^2/\kappa\gamma \approx 1$, spin-variances can be further decreased from initial values $\langle\Delta\hat{F}_z^2\rangle \propto N$ to levels $\propto \sqrt{N}$ [16, 116, 117]. (In free-space, it is in principle possible to achieve this degree of squeezing with a maximally focused probe beam, but one can do no better because of the diffraction limit.) To further reduce the uncertainty to the point where $\langle\Delta\hat{F}_z^2\rangle \propto 1$ (i.e., the Heisenberg limit of spin-squeezing) the cavity needs to be in the *very* strong coupling regime with $g^2/\kappa\gamma > N$. If one requires that a single $\hat{F}_z$ eigenstate becomes resolvable ($\langle\Delta\hat{F}_z^2\rangle \ll 1$), the cavity coupling requirements become even more stringent depending on the degree of projection desired.

While there are currently few experimental systems even in the strong coupling regime, we expect this very strong coupling regime to eventually be reached for moderate numbers of atoms. With this attitude we continue to focus on the long-time limit of the pure dispersive SME in the interest of understanding the idealized limits of continuous projective measurement. For a more complete discussion of the realistic physical limits of this type of QND measurement, see references [15, 16, 115, 116, 117, 118].

9.3.1 Hilbert Space, Coherent Spin-States, and Dicke States

Under certain idealizations, we can considerably reduce the size of the Hilbert space needed to describe the conditionally measured ensemble. Throughout this chapter, the initial state $\hat{\rho}(0)$ will be made equal to a coherent spin-state (CSS) polarized along an arbitrary direction [68]. For example, a CSS pointing along the z-axis is denoted $|\uparrow_1\uparrow_2\cdots\uparrow_N\rangle_z$ and all others can be prepared by rotating this state with the angular momentum operators $\hat{F}_i$, with $i \in \{x, y, z\}$. A CSS, typically obtained via a dissipative optical pumping process, is an eigenstate of $\hat{\mathbf{F}}^2$ with maximal eigenvalue $F(F+1)$, where $F = N/2$. Because the SME works under the QND approximation of negligible absorption (i.e., the large detuning dispersive limit), no angular momentum will be exchanged between the probe beam and the ensemble. The only other allowed dynamics possible are rotations of the angular momentum induced by applied magnetic fields, thus the state will maintain maximal $\langle\hat{\mathbf{F}}^2\rangle$ over the course of the measurement.

The Dicke states are defined [68] as the states $|l, m\rangle$ that are simultaneous eigenstates of both $\hat{\mathbf{F}}^2$ and $\hat{F}_z$

$$\hat{F}_z|l, m\rangle = m|l, m\rangle \tag{9.6}$$

$$\hat{\mathbf{F}}^2|l, m\rangle = l(l+1)|l, m\rangle \tag{9.7}$$

where

$$|m| \leq l \leq F = N/2. \tag{9.8}$$

Under the above approximations, we can neglect any state with $l \neq F$. We then shorten the labeling of our complete basis from $|F, m\rangle$ to $|m\rangle$ so that

$$\hat{F}_z|m\rangle = m|m\rangle \tag{9.9}$$

$$\hat{\mathbf{F}}^2|m\rangle = F(F+1)|m\rangle \tag{9.10}$$

where $m \in \{-N/2, -N/2 + 1, ..., N/2 - 1, N/2\}$. We earlier discussed the properties of Dicke states in chapter 3 and displayed them graphically in figure 3.2.

When the physical evolution is such that the $|m\rangle$ states remain complete, we can limit ourselves to a density matrix of size $(N + 1) \times (N + 1)$ rather than the full size $2^N \times 2^N$. This reduced space is referred to as the symmetric subspace, as its states are invariant to particle exchange [10, 126]. For the case of two spins, the symmetric subspace contains the triplet states, but not the singlet. States contained within the symmetric subspace can be described as a pseudospin of size $F = N/2$.

In the z-basis, the extremal values of m, $\pm N/2$, are simply the coherent spin-states pointing along the z-axis

$$|m = +N/2\rangle = |\uparrow_1\uparrow_2 \cdots \uparrow_N\rangle \tag{9.11}$$

$$|m = -N/2\rangle = |\downarrow_1\downarrow_2 \cdots \downarrow_N\rangle. \tag{9.12}$$

In terms of the constituent spins, these states are obviously unentangled. In contrast, consider the state with $m = 0$ (for N even)

$$|m = 0\rangle = C \Sigma_i P_i(|\uparrow_1 \cdots \uparrow_{N/2}\downarrow_{N/2+1} \cdots \downarrow_N\rangle) \tag{9.13}$$

where the P_i represent all permutations of the spins and C is a normalization constant. This state is highly entangled in a way that is robust to particle loss [17]. Even though the expectation values $\langle \hat{F}_i \rangle$ vanish for this state, it still has maximal $\mathbf{F}^2$ eigenvalue. Loosely, this state represents a state of knowledge where the length of the spin vector is known and the z-component is known to be zero, but the direction of the spin vector in the xy-plane is completely indeterminate. Similarly, the entangled states with $0 < |m| < N/2$ can be imagined as living on cones aligned along the z-axis with projection m. The loss of pointing angle information from the measurement process is diagrammed in figure 9.1B.

Along with their unique entanglement and uncertainty properties, Dicke states are also of interest for the important role they play in descriptions of collective radiation processes [68] and for their potential role in quantum information processing tasks [127, 121, 120].

9.3.2   Short-Time Limit

Even when working within the symmetric subspace, for a large number of spins the size of $\hat{\rho}(t)$ may be too unwieldy for computational efficiency. Because it is often desirable to update our state description in real time (e.g., for optimal feedback procedures), finding simple but sufficient descriptors is of considerable importance.

We can derive a reduced model by employing a moment expansion for the observable of interest. Extracting the conditional expectation values of the first two moments of $\hat{F}_z$ from the SME gives the following scalar stochastic differential equations:

$$
\begin{aligned}
d\langle \hat{F}_z \rangle(t) &= \gamma \langle \hat{F}_x \rangle(t) b(t)\, dt \\
&\quad + 2\sqrt{M\eta} \langle \Delta \hat{F}_z^2 \rangle(t)\, dW(t) \qquad (9.14) \\
d\langle \Delta \hat{F}_z^2 \rangle(t) &= -4M\eta \langle \Delta \hat{F}_z^2 \rangle^2(t)\, dt \\
&\quad - i\gamma \langle [\Delta \hat{F}_z^2, \hat{F}_y] \rangle(t) b(t)\, dt \\
&\quad + 2\sqrt{M\eta} \langle \Delta \hat{F}_z^3 \rangle(t) dW(t). \qquad (9.15)
\end{aligned}
$$

Note that these equations are not closed because higher-order moments couple to them.

At short times, $t \ll 1/\eta M$, we can make this set of equations closed with the following approximations. If the spins are initially fully polarized along x then, by using the evolution equation for the x-component, we can show $\langle \hat{F}_x \rangle(t) \approx F \exp[-Mt/2]$. Making the Gaussian

approximation at short times, the third-order terms $\langle \Delta \hat{F}_z^3 \rangle$ and $-i\gamma \langle [\Delta \hat{F}_z^2, \hat{F}_y] \rangle (t) b(t)$ can be neglected. The Holstein-Primakoff transformation makes it possible to derive this Gaussian approximation as an expansion in $1/F$ [89]. Both of the removed terms can be shown to be approximately $1/F\sqrt{F}$ smaller than the retained nonlinear term. Thus we can approximate the optimal solution with

$$
\begin{aligned}
d\langle \hat{F}_z \rangle_s(t) &= \gamma F \exp[-Mt/2] b(t)\, dt \\
&\quad + 2\sqrt{M\eta} \langle \Delta \hat{F}_z^2 \rangle_s(t)\, dW_s(t) \quad\quad (9.16) \\
d\langle \Delta \hat{F}_z^2 \rangle_s(t) &= -4M\eta \langle \Delta \hat{F}_z^2 \rangle_s^2(t)\, dt \quad\quad\quad\quad\quad (9.17)
\end{aligned}
$$

where the s subscript denotes the short-time solution and $dW_s(t) \equiv 2\sqrt{M\eta}[y(t)dt - \langle \hat{F}_z \rangle_s(t)dt]$. Also $b(t)$ is assumed to be of a form that keeps the total state nearly pointing along x. The differential equation for the variance $\langle \Delta \hat{F}_z^2 \rangle_s(t)$ is now deterministic. It can be solved to give

$$
\langle \Delta \hat{F}_z^2 \rangle_s(t) = \frac{\langle \Delta \hat{F}_z^2 \rangle(0)}{1 + 4M\eta \langle \Delta \hat{F}_z^2 \rangle(0)t}. \quad\quad (9.18)
$$

The deterministically shrinking value of $\langle \Delta \hat{F}_z^2 \rangle_s(t)$ represents the squeezing about the initially fluctuating value of $\langle \hat{F}_z \rangle_s(t)$ as shown in the first two frames of figure 9.1A-B. If feedback is added, then the value of $\langle \hat{F}_z \rangle_s(t)$ can be zeroed via Larmor precession due to a control field along y and the same centered spin-squeezed state can be prepared on every trial [15, 16].

The resulting spin-squeezed states can be used in subsequent precision measurements [5, 45]. It is also worth pointing out that a precision measurement can be performed *during* the production of the conditional spin-squeezing. For example, we have shown that the by properly estimating both the spin-state and an unknown classical field simultaneously with continuous measurement and Kalman filtering techniques, the field estimation can be improved over conventional limits by the presence of the simultaneous squeezing [18, 20].

9.3.3 Long-Time Limit

The approximations made in the previous section are no longer valid at times $t \gg 1/\eta M$. The third-order terms become nonnegligible at long times, hence the variance becomes stochastic. Subsequently, other high-order moments couple into the problem and we are

forced to consider the stochastic differential equation for each. Eventually, any finite num-
bered moment description is no longer useful and it initially appears that we must resort
back to the full symmetric density matrix and the SME, equation (B.43), as our primary
description.

Fortunately, we can take another approach and describe the state in terms of other
sufficient statistics. Without a field, the only statistic of the photocurrent needed to describe
the state at time t is its integral, $\int_0^t y(s)\,ds$ (see section 9.6 or [128]). Knowing that the
state is only a function of this variable and the initial state (prior information) makes the
experimental design of a real time estimator experimentally convenient. For example, we
could use an analog integrator to create this sufficient statistic from the raw photocurrent,
then feed it into a possibly nonlinear device (like an FPGA [10]) to perform the estimation.

With the integrated photocurrent and the initial state

$$|\psi(0)\rangle = \sum_{m=-F}^{F} c_m |m\rangle \tag{9.19}$$

we can calculate (see section 9.6) the conditional expectation value of any power of $\hat{F}_z$ with
the expression

$$\begin{aligned}
\mathrm{Tr}[\hat{F}_z^k \tilde{\tilde{\rho}}(t)] &= \sum_{m=-F}^{F} m^k |c_m|^2 \exp[-2M\eta m^2 t \\
&\quad + 4mM\eta \int_0^t y(s)\,ds]
\end{aligned} \tag{9.20}$$

where $\tilde{\tilde{\rho}}(t)$ is the unnormalized density matrix, and setting $k = 0$ represents its trace, so

$$\langle \hat{F}_z^k \rangle(t) = \mathrm{Tr}[\hat{F}_z^k \tilde{\tilde{\rho}}(t)] / \mathrm{Tr}[\hat{F}_z^0 \tilde{\tilde{\rho}}(t)]. \tag{9.21}$$

Consider the case when the system starts in the x-polarized spin-coherent state. To very
good approximation (with reasonably large F) we can write for this state in the z-basis

$$|c_m|^2 \propto \exp\left[-\frac{m^2}{F}\right]. \tag{9.22}$$

Using these coefficients, we now have the rule for mapping the photocurrent to the expec-

tation of $\hat{F}_z$

$$\langle\hat{F}_z\rangle(t) = \text{Tr}[\hat{F}_z^1\tilde{\bar{\rho}}(t)]/\text{Tr}[\hat{F}_z^0\tilde{\bar{\rho}}(t)]. \tag{9.23}$$

Other than the minor approximation of the initial coefficients, using this estimate is essentially the same as using solution to the full SME, so we do not give it a new subscript.

To simplify further, we can change the sums to integrals, giving

$$\text{Tr}[\hat{F}_z^k\tilde{\bar{\rho}}(t)] \simeq \int_{-F}^{F} m^k e^{-Am^2+2Bm}dm \tag{9.24}$$

with

$$A = \frac{1}{F} + 2M\eta t \qquad B = 2M\eta\int_0^t y(s)\,ds. \tag{9.25}$$

This approximation produces an estimate

$$\langle\hat{F}_z\rangle_i(t) = \frac{\int_{-F}^{F} m e^{-Am^2+2Bm}dm}{\int_{-F}^{F} e^{-Am^2+2Bm}dm} \tag{9.26}$$

that performs suboptimally when the distribution of states becomes very narrow at long times. Interestingly, the integral approximation here numerically appears to give the same estimate as the one derived previously for short times when no field is present, i.e.,

$$\langle\hat{F}_z\rangle_i(t) = \langle\hat{F}_z\rangle_s(t). \tag{9.27}$$

This is not entirely surprising as both of these estimators ignore the discreteness of the Dicke levels. Also, at long times, it turns out that both of these estimates appear to be numerically equivalent to the simplest of all estimates: averaging the photocurrent. In other words, one simple and intuitive approximation to the optimal $\langle\hat{F}_z\rangle(t)$ would be

$$\langle\hat{F}_z\rangle_a(t) = \frac{\int_0^t y(s)ds}{t}, \tag{9.28}$$

which is an estimate one might guess from the form of the photocurrent, equation (9.4). From simulation, it appears that this estimate is the same as both $\langle\hat{F}_z\rangle_i(t)$ and $\langle\hat{F}_z\rangle_s(t)$ for $t \gg 1/\eta M$. Despite the nonoptimality of these simple estimators, they perform well enough to resolve the discretization of the Dicke levels at long times.

Unfortunately, the addition of a feedback field makes these simplified estimators inad-
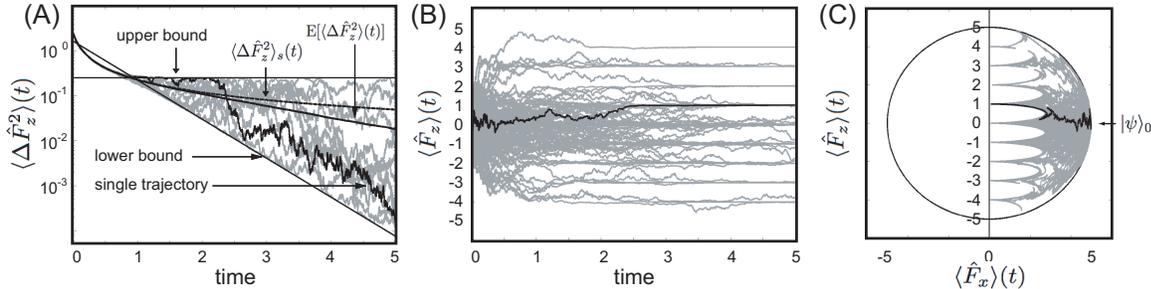
Figure 9.2: Many open-loop moment trajectories [36] of the SSE, equation (9.5). The trajectory of figure 9.1 is darkened. (A) At short times, the evolution of the variance (shown on a log scale) is deterministic and given by $\langle \Delta \hat{F}_z^2 \rangle_s(t)$. At long times, the variances become stochastic but bounded (above by $1/4$ and below by $\exp[-2(Mt-1)]/4$). The average of all $10,000$ trajectories (only 10 are shown) gives $\mathrm{E}[\langle \Delta \hat{F}_z^2 \rangle(t)]$. (B) The projective nature of the measurement is made clear by the evolution of 100 trajectories of $\langle \hat{F}_z \rangle(t)$. The distribution of the final results is given by the first histogram of figure 9.1A. (C) The evolution of the 100 trajectories all starting in an x-polarized CSS. When $\eta = 1$, certain regions of Hilbert space are forbidden by the evolution.

equate at long times, and deriving simple reduced models with a field present is difficult, thus forcing us to use the full SME in our state-based controller. Despite this difficulty, during our subsequent feedback analysis we assume sufficient control bandwidth that the SME can be evolved by the observer in real time.

9.4 Measurement of Evolution without Feedback

In this section, our goal is to describe how the estimates of the last section probabilistically evolve at long times into Dicke states via observation alone. First, we discuss steady-state and statistical properties of the SME, equation (B.43). Then, we examine the unconditional dynamical solution with $\eta = 0$, which gives the average state preparation behavior when $\eta \neq 0$. We then consider in detail how individual trajectories behave when $\eta \neq 0$. Finally, we discuss the performance of the nonoptimal estimators relative to the optimal projective estimator.

9.4.1 Steady-States of the SME and Martingale Properties

The fact that the SME eventually prepares eigenstates of $\hat{F}_z$ is rather intuitive from a projection postulate perspective because $\hat{F}_z$ is the quantity being measured. If we insert

the pure Dicke state, $\hat{\rho} = |m\rangle\langle m|$, into the SME with no Hamiltonian (or only a field along z), we find that it is a steady-state, $d\hat{\rho} = 0$, no matter what happens with the subsequent measurement record. Of course, this does not yet prove that the state will eventually be obtained, as we have not discussed the stability of attractors in stochastic systems.

Without a field present, the SME has several convenient properties. First of all, from the evolution equation for the variance notice that the variance is a stochastic process that decreases on average. In fact it is a supermartingale, in that for times $s \leq t$ we have

$$\mathrm{E}_s[\langle\Delta\hat{F}_z^2\rangle(t)] \leq \langle\Delta\hat{F}_z^2\rangle(s) \tag{9.29}$$

where the notation $\mathrm{E}[x(t)]$ denotes the average of the stochastic variable $x(t)$ at time t and the s subscript represents conditional expectation given a particular stochastic trajectory up to the time s. Additionally, it can be shown [71] that the average variance obeys the equation

$$\mathrm{E}[\langle\Delta\hat{F}_z^2\rangle(t)] = \frac{\langle\Delta\hat{F}_z^2\rangle(0)}{1 + 4M\eta\langle\Delta\hat{F}_z^2\rangle(0)(t + \xi(t))} \tag{9.30}$$

where

$$\begin{aligned}
\xi(t) &= \int_0^t \frac{\mathrm{E}[(\langle\Delta\hat{F}_z^2\rangle(s) - \mathrm{E}[\langle\Delta\hat{F}_z^2\rangle(s)])^2]}{\mathrm{E}[\langle\Delta\hat{F}_z^2\rangle(s)]^2}\, ds \\
&\geq 0.
\end{aligned} \tag{9.31}$$

A more explicit solution of $\xi(t)$ is not necessarily needed as its positivity ensures that $\langle\Delta\hat{F}_z^2\rangle(t)$ stochastically approaches zero. This implies that a Dicke state is eventually prepared. The numerical simulation of figures 9.1 and 9.2 demonstrates this behavior for an initially x polarized state. As expected, $\mathrm{E}[\langle\Delta\hat{F}_z^2\rangle(t)]$ in figure 9.2A appears to be less than the short-time solution $\langle\Delta\hat{F}_z^2\rangle_s(t)$, equation (9.18), at long times.

Other useful properties of the stochastic evolution are evident from the moment equations. For example, we can show that

$$d\langle\hat{F}_z^n\rangle = 2\sqrt{M\eta}(\langle\hat{F}_z^{n+1}\rangle - \langle\hat{F}_z^n\rangle\langle\hat{F}_z\rangle)dW(t) \tag{9.32}$$

for integer n, hence

$$d\mathrm{E}[\langle\hat{F}_z^n\rangle] = 0 \tag{9.33}$$

and for times $s \leq t$ we have the martingale condition

$$E_s[\langle \hat{F}_z^n \rangle(t)] = \langle \hat{F}_z^n \rangle(s). \tag{9.34}$$

This equation for $n = 1$ gives us the useful identity

$$E[\langle \hat{F}_z \rangle(t)\langle \hat{F}_z \rangle(s)] = E[\langle \hat{F}_z \rangle(s)^2] \tag{9.35}$$

for $s \leq t$. Also, we can rewrite the expression for $n = 2$ as

$$E_s[\langle \hat{F}_z \rangle(t)^2 + \langle \Delta \hat{F}_z^2 \rangle(t)] = \langle \hat{F}_z \rangle(s)^2 + \langle \Delta \hat{F}_z^2 \rangle(s). \tag{9.36}$$

This implies a sort of conservation of uncertainty as the diffusion in the mean, shown in figure 9.1B, makes up for the decreasing value of the variance.

9.4.2 Zero Measurement Efficiency

It is insightful to examine the behavior of the master equation with $\eta = 0$, which corresponds to ignoring the measurement results and turns the SME equation (B.43) into a deterministic unconditional master equation. We continue to consider only those initial states that are polarized. This is because these states are experimentally accessible (via optical pumping) and provide some degree of selectivity for the final prepared state. To see this, let us consider a spin-1/2 ensemble polarized in the xz-plane, making angle θ with the positive z-axis, such that

$$\begin{aligned}
\langle \hat{F}_x \rangle(0) &= \sin(\theta)N/2 \\
\langle \hat{F}_y \rangle(0) &= 0 \\
\langle \hat{F}_z \rangle(0) &= \cos(\theta)N/2 \\
\langle \Delta \hat{F}_x^2 \rangle(0) &= \cos^2(\theta)N/4 \\
\langle \Delta \hat{F}_y^2 \rangle(0) &= N/4 \\
\langle \Delta \hat{F}_z^2 \rangle(0) &= \sin^2(\theta)N/4.
\end{aligned} \tag{9.37}$$

Solving the unconditional moment equations, and labeling them with u subscripts, we get

$$
\begin{aligned}
\langle \hat{F}_x \rangle_u(t) &= \sin(\theta)\exp(-Mt/2)N/2 & (9.38)\\
\langle \hat{F}_y \rangle_u(t) &= 0 \\
\langle \hat{F}_z \rangle_u(t) &= \cos(\theta)N/2 \\
\langle \Delta \hat{F}_x^2 \rangle_u(t) &= \sin^2(\theta)[N^2 - N - 2N^2\exp(-Mt) \\
&\qquad + (N^2 - N)\exp(-2Mt)]/8 + N/4 \\
&\to \sin^2(\theta)(N^2 - N)/8 + N/4 \\
\langle \Delta \hat{F}_y^2 \rangle_u(t) &= \sin^2(\theta)[N^2 - N \\
&\qquad + (N - N^2)\exp(-2Mt)]/8 + N/4 \\
&\to \sin^2(\theta)(N^2 - N)/8 + N/4 \\
\langle \Delta \hat{F}_z^2 \rangle_u(t) &= \sin^2(\theta)N/4.
\end{aligned}
$$

Note that, because the unconditional solutions represent the average of the conditional solution, i.e., $\hat{\rho}_u(t) = \mathrm{E}[\hat{\rho}(t)]$, we have

$$
\mathrm{E}[\langle \hat{F}_z \rangle(t)] = \langle \hat{F}_z \rangle_u(t) = \langle \hat{F}_z \rangle(0) = \cos(\theta)N/2. \tag{9.39}
$$

This also follows from the martingale condition for $\langle \hat{F}_z \rangle(t)$. From the martingale condition for $\langle \hat{F}_z^2 \rangle(t)$ we get

$$
\begin{aligned}
\mathrm{E}[(\langle \hat{F}_z \rangle(t) &- \mathrm{E}[\langle \hat{F}_z \rangle(t)])^2] \\
&= \langle \Delta \hat{F}_z^2 \rangle(0) - \mathrm{E}[\langle \Delta \hat{F}_z^2 \rangle(t)] & (9.40)\\
&\to \langle \Delta \hat{F}_z^2 \rangle(0) = \sin^2(\theta)N/4.
\end{aligned}
$$

Thus, when $0 < \eta \le 1$, we expect the final random conditional Dicke state on a given trial to fall within the initial z distribution. Given θ, the distribution will have spread $|\sin(\theta)|\sqrt{N}/2$ about the value $\cos(\theta)N/2$. Although the final state is generally random, starting with a polarized state clearly gives us some degree of selectivity for the final Dicke state because $\sqrt{N} \ll N$.

9.4.3   Nonzero Measurement Efficiency

When $\eta \neq 0$, the measurement record is used to condition the state, and we can determine which Dicke state the system diffuses into. Given the task of preparing the state $|m_d\rangle$, the above analysis suggests the following experimental procedure. First, polarize the ensemble (via optical pumping) into an unentangled coherent state along any direction. Then rotate the spin vector (with a magnetic field) so that the z-component is approximately equal to m_d. Finally, continuously measure z until a time $t \gg 1/\eta M$. The final estimate will be a random Dicke state in the neighborhood of m_d. When the trial is repeated, the final states will make up a distribution described by the initial moments of $\hat{F}_z$ ($\langle \hat{F}_z \rangle(0), \langle \Delta \hat{F}_z^2 \rangle(0), ...$). To reduce the effects of stray field fluctuations and gradients, a strong holding field could be applied along the z-axis. Because this Hamiltonian commutes with the observable $\hat{F}_z$, the final open-loop measurement results would be unchanged.

This process (with zero field) is shown schematically in figure 9.1 for $m_d = 0$ where the initial state is polarized along x. Because $\langle \hat{F}_z \rangle(0) = 0$, the final state with the highest probability is the entangled Dicke state $m_d = 0$. In contrast, if $\langle \hat{F}_z \rangle(0) = F$ the state would start in an unentangled CSS polarized along z and would not subsequently evolve.

One way of characterizing how close the state is to a Dicke state is through the variance, $\langle \Delta \hat{F}_z^2 \rangle(t)$. Figure 9.2A displays many trajectories for the variance as a function of time. For times $t \ll 1/\eta M$ the variance is approximately deterministic and obeys the short-time solution of equation (9.18). During this period, the mean $\langle \hat{F}_x \rangle(t)$ is decreasing at rate $M/2$. Before this mean has completely disappeared, a conditional spin-squeezed state is created. However, for larger times the mean and the variance stochastically approach zero, and the state, while still entangled, no longer satisfies the spin-squeezing criterion [17].

There are several features to notice about the approach to a Dicke state that are evident in figures 9.1 and 9.2. The variance at time $t = 1/\eta M$ is already of order unity. Thus, at this point, only a few neighboring m levels contain any population, as can be seen in figure 9.1C. Also, it can be numerically shown that, for x-polarized initial states, the diffusion of the variance at long times $t \gg 1/\eta M$ is bounded above and below by

$$\exp[-2(\eta M t - 1)]/4 < \langle \Delta \hat{F}_z^2 \rangle(t) \leq 1/4, \tag{9.41}$$

which is evident from figure 9.2A. These facts indicate that the population is divided among

at most two levels at long times which "compete" to be the final winner. If we assume that only two neighboring levels are occupied and apply the SSE (with $\eta = 1$), the probability, p, to be in one level obeys the stochastic equation

$$dp = -2Mp(1-p)\, dW(t) \tag{9.42}$$

and the variance takes the form $\langle \Delta \hat{F}_z^2 \rangle(t) = p(1-p)$. As simple as it looks, this stochastic differential equation (SDE) is not analytically solvable [53, 52]. The maximum variance is $1/4$ and it can be shown that, for $p \equiv 1 - \epsilon$, with ϵ small, the lower bound is of the exponential form stated above, so the two-level assumption seems to be a good one. The fact that occupied Hilbert space becomes small at long times is also evident in figure 9.2C, where the allowed states are seen to be excluded from certain regions when $\eta = 1$. The arclike boundaries of the forbidden space are where the two level competition occurs. This few-level SDE is discussed more fully in section 9.8.

In practice, an experimentalist does not always have an infinite amount of time to prepare a state. Eventually spontaneous emission and other decoherence effects will destroy the dispersive QND approximation that the present analysis is based upon. Suppose our task was to prepare a Dicke state with, on average, a desired uncertainty, $\langle \Delta \hat{F}_z^2 \rangle_d \ll 1$, such that one level was distinguishable from the next. From equation (9.30), we see that the time that it would take to do this on average is given by

$$t_d = \left[\frac{1}{\langle \Delta \hat{F}_z^2 \rangle_d} - \frac{1}{\langle \Delta \hat{F}_z^2 \rangle(0)} \right] /4M\eta. \tag{9.43}$$

Thus if $\langle \Delta \hat{F}_z^2 \rangle_d \ll 1$ is our goal, then t_d is how long the state must remain coherent. The larger $\langle \Delta \hat{F}_z^2 \rangle(0)$ is the more entangled the final states are likely to be ($m \approx 0$) [17], hence, by equation (9.43), the longer it takes to prepare the state for a given $\langle \Delta \hat{F}_z^2 \rangle_d$. Hence, we arrive at the intuitively satisfying conclusion that conditional measurement produces entangled states more slowly than unentangled states. Of course, equation (9.43) is an average performance limit. In a best-case scenario, the variance would attain the lower bound of equation (9.41) where the state reduction happens exponentially fast.

9.4.4 Performance of Suboptimal Estimators

Now we consider the performance of the suboptimal estimators discussed previously, in particular the current average $\langle \hat{F}_z \rangle_a(t)$ of equation (9.28). It makes sense to associate the overall "error" of this estimator, denoted V_a, to be the average squared distance of the estimator from the optimal estimator *plus* the average uncertainty of the optimal estimator itself, $\mathrm{E}[\langle \Delta \hat{F}_z^2 \rangle(t)]$. Using the martingale properties of the optimal estimate and the definition of the photocurrent gives this quantity as

$$
\begin{aligned}
V_a &\equiv \mathrm{E}[(\langle \hat{F}_z \rangle_a(t) - \langle \hat{F}_z \rangle(t))^2] + \mathrm{E}[\langle \Delta \hat{F}_z^2 \rangle(t)] \\
&= \frac{1}{4M\eta t}.
\end{aligned}
\tag{9.44}
$$

This is just the error in estimating a constant masked by additive white noise with the same signal-to-noise ratio [20]. The optimal estimator is better than this suboptimal estimator at long times only through the quantity $\xi(t)$, equation (9.31).

In the open-loop experimental procedure described at the beginning of the last section, the above observation indicates that we can replace the optimal estimator with the photocurrent average and still resolve the projective behavior (given sufficient elimination of extraneous noise). The price paid for the simplicity of the averaging estimator is that it converges more slowly and it only works when a field is not present (hence without control).

In appendix E, we analyze the performance of the averaging estimator when incoherent decay is added to the model.

9.5 Closed-Loop Evolution

The primary problem with the open-loop state preparation scheme (and other approaches [120, 121, 122]) is that it is probabilistic. For a single measurement, there exists some degree of control, by adjusting the initial angle of rotation θ, but the final state is a priori unpredictable within the variance of the initial state. In this section, we show that the state preparation can be made deterministic with the use of feedback. Just as the control scheme of [15, 16] produces deterministically centered spin-squeezed states, we present a simple feedback controller that will prepare the same desired Dicke state (particularly $m_d = 0$) on every measurement trial.

We choose to work with y-axis magnetic field actuator corresponding to the Hamiltonian, $H(t) = \gamma b(t) \hat{F}_y$. If the CSS initial state begins in the xz-plane this will ensure that the vector $\langle \mathbf{F} \rangle(t)$ remains in this plane. This actuator is natural for the control of spin-squeezed states at short times, where the linear moments of $\langle \mathbf{F} \rangle(t)$ are large and allow intuitive rotation of the spin vector. However, at long times the field will mostly be affecting nonlinear terms in the moment expansion and the dynamics are less intuitive as can be seen by the structure near the z-axis in figure 9.2C. Still, we continue to give ourselves only these rotations to work with as they are the most experimentally accessible actuation variable.

In principle, the fact that Dicke states can be prepared deterministically with feedback should not be surprising. Given the aforementioned characteristics of the noncontrolled measurement one could imagine preparing a particular state by *alternating* measurement and control periods. For example, an initial measurement (lasting for a time $\Delta t \ll 1/\eta M$) would determine the fluctuation of $\langle \hat{F}_z \rangle$ while the uncertainty $\langle \Delta \hat{F}_z^2 \rangle$ simultaneously decreased (on average). Then the measurement would be turned off and the state would be rotated with a control field to "zero" the conditional quantity $\langle \hat{F}_z \rangle - m_d$ (if preparing $|m_d\rangle$). The process of alternating measurement and control could then be repeated and would eventually clamp down on the desired state. Notice that, unlike the preparation of spin-squeezed states [15, 16], this procedure could not be performed with a *single* measurement and control cycle. In other words, if we measure for a time $t \gg 1/\eta M$, and prepare a probabilistic Dicke state, then a single post-measurement rotation cannot prepare a different desired Dicke state in the same basis.

With this intuitive picture in mind, now consider the continuous limit of this process, where the measurement and control are performed simultaneously. We wish to find a mapping from the photocurrent history to the control field that prepares our state of interest in a satisfactory manner on *every* trial. For simplicity, we work with $\eta = 1$ and use the SSE of equation (9.5) for all simulations [36]. In selecting a controller, we could choose one of several strategies, including either direct current feedback or a feedback rule based on the state (i.e., what has been called Markovian and Bayesian feedback, respectively [129, 125]). While direct current feedback possesses certain advantages, mainly simplicity that allows practical implementation, and is capable of working adequately at short times, any constant gain procedure would never prepare a Dicke state with confidence. If the current is directly fed back, a finite amount of noise will unnecessarily drive the system away from its target,
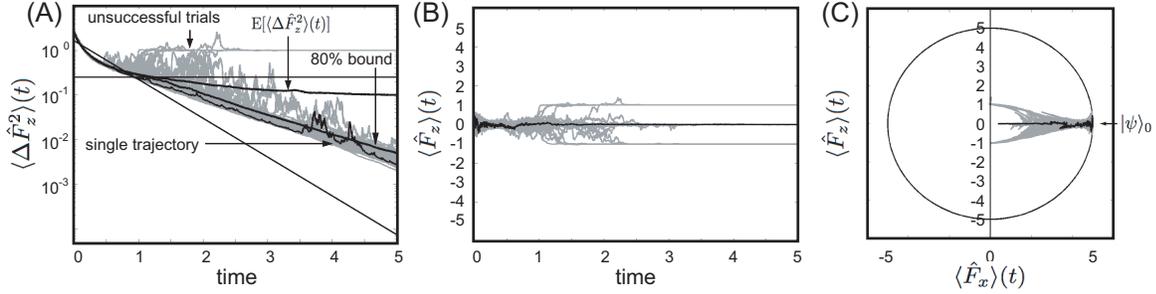
Figure 9.3: One hundred closed-loop moment trajectories [36] of the SSE with feedback law $b(t) = \lambda \langle \hat{F}_x \hat{F}_z + \hat{F}_z \hat{F}_x \rangle(t)/2$ and $\lambda = 10$ chosen from numerical considerations. (A)-(B) If the control is successful the quantity $\langle \hat{F}_z^2 \rangle(t)$ should go to zero on every trial. For this controller the number of successful trajectories is increased significantly (from 25 to 90 %), but the remaining fraction is attracted to neighboring fixed points, causing the mean $E[\langle \hat{F}_z^2 \rangle(t)]$ to saturate at a nonzero value. Although the successful fraction converges exponentially, the fastest converging trajectories are slower than in the open-loop case. This is evident in (C) as the converging trajectories have visibly not yet reached $\langle \hat{F}_x \rangle = 0$ at time $t = 5$.

even if the state starts there. Of course the gain could be ramped to zero in time, but unlike the short-time case, it is not clear how to tailor the gain intelligently.

Another alternative would be to prepare a spin-squeezed state with this approach and then turn off the feedback at some intermediate time. This would certainly enhance the probability of obtaining a certain Dicke state, but the process would remain probabilistic to some degree. For these reasons, we continue considering only state-based feedback, despite the fact that updating the state in real time is experimentally challenging.

9.5.1 Defining a Cost Function

A useful first step in the design of any controller is to define the quantity that the ideal controller should minimize: the cost function. For example, consider a state preparation application where the controller aims to produce the desired target state $|\psi_d\rangle$. In this case, one possible cost function is the quantity

$$U_f \equiv 1 - \langle \psi_d | \hat{\rho} | \psi_d \rangle \geq 0 \tag{9.45}$$

evaluated at the stopping time, which is zero iff the fidelity of the state with respect to the target is unity. In the current application, where we desire a final Dicke state $|m_d\rangle$ we wish

to minimize a different quantity

$$
\begin{aligned}
U &\equiv (\langle \hat{F}_z \rangle - m_d)^2 + \langle \Delta \hat{F}_z^2 \rangle \\
&= \Sigma_m \langle m | \hat{\rho} | m \rangle^2 (m - m_d)^2 \\
&\geq 0,
\end{aligned}
\tag{9.46}
$$

which is zero iff $\hat{\rho} = |m_d\rangle\langle m_d|$. Notice that U gives a higher penalty than U_f to states that are largely supported by Dicke states far removed from the target. In general, U will evolve stochastically and we may be more interested in the mean behavior, denoted $\mathrm{E}[U]$. In the uncontrolled case, it can be shown that this quantity remains constant, $\mathrm{E}[U(t)] = U(0)$. For the controlled case, we wish for $\mathrm{E}[U] \to 0$ as time advances, which, because $U \geq 0$, implies that every trajectory approaches the target state $|m_d\rangle$.

In general, the cost function could also include an integral of the quantity $U(t)$ instead of just the final value. As in classical control theory [20], it is also practical to include a function of $b(t)$ in the cost as a way of expressing our experimental feedback gain and bandwidth constraints. Analytically proceeding in this way by optimizing the average cost is too difficult for the current problem, but, with this perspective in mind, we proceed by proposing controllers according to related considerations.

9.5.2   Control Law 1

Now consider the average evolution of the above cost function, which is given by

$$
d\mathrm{E}[U(t)] = -2\gamma \mathrm{E}\left[b(t) \left(\frac{\langle \hat{F}_x \hat{F}_z + \hat{F}_z \hat{F}_x \rangle(t)}{2} - m_d \langle \hat{F}_x \rangle(t) \right) \right] dt.
\tag{9.47}
$$

Because we want the expectation of the cost function to continuously decrease, this derivative should be negative at all times. If we have full access to the density matrix, and minimal feedback delay, we could use the controller

$$
b_1(t) = \lambda \left(\frac{\langle \hat{F}_x \hat{F}_z + \hat{F}_z \hat{F}_x \rangle(t)}{2} - m_d \langle \hat{F}_x \rangle(t) \right)
\tag{9.48}
$$

where λ is a constant positive gain factor. This law guarantees that $d\mathrm{E}[U(t)] \leq 0$. Still, this does not yet prove that $U = 0$ is obtained because $d\mathrm{E}[U(t)] = 0$ for states other than
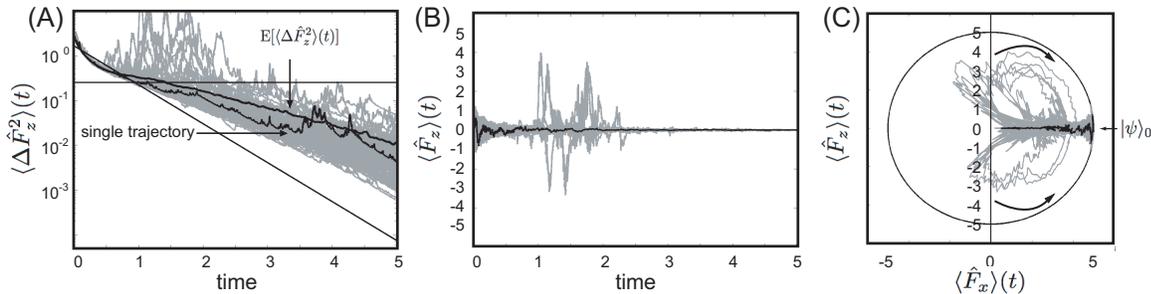
Figure 9.4: One hundred closed-loop moment trajectories [36] of the SSE with feedback law $b(t) = \lambda \langle \hat{F}_z \rangle(t)$ and $\lambda = 10$ chosen from numerical considerations. (A) The average over 10,000 trajectories suggests that with this control law the mean $\mathrm{E}[\langle \hat{F}_z^2 \rangle(t)]$ descends to zero exponentially and the target state is deterministically prepared. (B) Despite a number of early excursions, all 100 trajectories shown converge to the desired value of $m = 0$. (C) Those trajectories that do not descend to the goal directly (about 10 of 100) are recycled and rotated back into the attractive region of the target state. Again, the control slightly compromises the best-case convergence rate and the trajectories have a nonzero (but still decreasing) $\langle \hat{F}_x \rangle$ at $t = 5$.

the target state. Furthermore, even with this control law applied, all Dicke states *remain* fixed points.

Regardless of these issues, we proceed by analyzing the performance of this control law numerically with $m_d = 0$. In principle, the gain could be chosen arbitrarily large. Here we choose to work with a gain that is large enough to be effective but small enough to keep the numerical simulation results valid [36]. The choice of a limited gain is a necessity in both simulation and experiment, thus we wish to find a control law that works within this constraint. For the parameters used in our simulation, we use a gain of $\lambda = 10$, which produces the results shown in figure 9.3.

In figure 9.3A, we now plot the figure of merit for $m_d = 0$, $U(t) = \langle \hat{F}_z^2 \rangle(t)$. In open-loop configuration, only 25% of all trajectories are attracted to $m = 0$, whereas with this controller the percentage reaches 90%. Furthermore, most of these trajectories approach the state at an exponential rate close to M, as indicated by the curve under which 80% of the trajectories lie. Interestingly, this is at the expense of those trajectories that in open-loop approached the target state at an exponential rate of $2M$. There is a trade-off by which the control slightly compromises the convergence of the best-case trajectories.

Unfortunately, because all other Dicke states are still fixed points of the controlled SSE and the gain is finite, a small fraction (10%) of trajectories are attracted to those states

neighboring the target state. Thus this controller does not appear to deterministically prepare all trajectories into the target state and the mean $E[\langle \hat{F}_z^2 \rangle(t)]$ flattens at a level determined by the unsuccessful fraction of trials.

9.5.3 Control Law 2

The obvious solution to the above problem is to try a controller that ensures the target state is the *only* fixed point of the SME/SSE. In this section we propose the control law

$$b_2(t) = \lambda(\langle \hat{F}_z \rangle(t) - m_d) \tag{9.49}$$

for which the state $|m_d\rangle$ is the only fixed point. However, unlike $b_1(t)$ this controller lacks the x symmetry that ensures $dE[U(t)] \leq 0$. Also, while the symmetry of $b_1(t)$ will allow it to lock to both sides of the Bloch sphere, $b_2(t)$ will only lock to one side of the sphere.

Again, we proceed by numerically analyzing the performance of this controller for $m_d = 0$, with the results displayed in figure 9.4. The gain is chosen in the same manner as before, which leads to the same reasonable choice of $\lambda = 10$. In figure 9.4C the fundamental nature of the dynamics can be seen. Close to 90% of the trajectories are directly transported towards the target state, but the remaining fraction "misses" on the first pass. Instead of being attracted towards other fixed points though, this unsuccessful fraction is *recycled* and rotated back onto the positive x-axis where they can reattempt convergence onto the target state. These large excursions can be seen in figure 9.4A-B as well, but they do not appear to dominate the net flow. The average of $10,000$ trajectories gives a quantity $E[\langle \hat{F}_z^2 \rangle(t)]$, which appears to exponentially descend towards zero, implying that the state preparation has been made deterministic. As with the control of $b_1(t)$ there is again a trade-off: the trajectories that previously descended at the exponential rate of $2M$ converge more slowly, but still exponentially.

9.6 Solution of the SME without a Field

An explicit solution to the SME, equation (B.43), can easily be found in the case $H(t) = 0$. First, the SME is rewritten as

$$d\tilde{\rho}(t) = \mathcal{D}[\sqrt{M}\hat{F}_z]\tilde{\rho}(t)dt + 2M\eta\left(\hat{F}_z\tilde{\rho}(t) + \tilde{\rho}(t)\hat{F}_z\right)y(t)dt. \tag{9.50}$$

This equation, known as the *unnormalized* or *linear* SME, is equivalent to equation (B.43) with the identification

$$\hat{\rho}(t) = \tilde{\hat{\rho}}(t)/\mathrm{Tr}[\tilde{\hat{\rho}}(t)].\tag{9.51}$$

Introducing the notation

$$\mathcal{G}_1\tilde{\hat{\rho}} = \hat{F}_z\tilde{\hat{\rho}}\hat{F}_z$$
$$\mathcal{G}_2\tilde{\hat{\rho}} = \hat{F}_z^2\tilde{\hat{\rho}} + \tilde{\hat{\rho}}\hat{F}_z^2 \tag{9.52}$$
$$\mathcal{G}_3\tilde{\hat{\rho}} = \hat{F}_z\tilde{\hat{\rho}} + \tilde{\hat{\rho}}\hat{F}_z.$$

equation (9.50) can be written in the more suggestive form

$$d\tilde{\hat{\rho}}(t) = M(\mathcal{G}_1 - \tfrac{1}{2}\mathcal{G}_2)\tilde{\hat{\rho}}(t)dt + 2M\eta\,\mathcal{G}_3\tilde{\hat{\rho}}(t)\,y(t)dt.\tag{9.53}$$

Now note that equation (9.53) is a linear Itô stochastic differential equation (SDE) [53] for $\tilde{\hat{\rho}}(t)$, and moreover $\mathcal{G}_{1,2,3}$ all commute with each other in the sense that $\mathcal{G}_i\mathcal{G}_j\tilde{\hat{\rho}} = \mathcal{G}_j\mathcal{G}_i\tilde{\hat{\rho}}$. Such SDEs have a simple explicit solution [52]

$$\begin{aligned}\tilde{\hat{\rho}}(t) &= \exp[(M(1-\eta)\mathcal{G}_1 - M(1+\eta)\mathcal{G}_2/2)t \\ &\quad +2M\eta\,\mathcal{G}_3\int_0^t y(s)ds]\tilde{\hat{\rho}}(0)\end{aligned}\tag{9.54}$$

as is easily verified by taking the time derivative of this expression, where care must be taken to use Itô's rule for the stochastic term.

Now consider an initial pure state of the form

$$|\psi(0)\rangle = \sum_{m=-F}^{F} c_m|m\rangle.\tag{9.55}$$

The associated initial density matrix is then

$$\tilde{\hat{\rho}}(0) = |\psi(0)\rangle\langle\psi(0)| = \sum_{m,m'=-F}^{F} c_m c_{m'}^*|m\rangle\langle m'|.\tag{9.56}$$

Substituting into equation (9.54) gives

$$
\begin{aligned}
\tilde{\rho}(t) = \sum_{m,m'=-F}^{F} c_m c_{m'}^* \exp[&(M(1-\eta)mm' \\
&- \tfrac{1}{2}M(1+\eta)(m^2 + (m')^2))t \\
&+ 2M\eta\,(m+m') \int_0^t y(s)ds]\,|m\rangle\langle m'|.
\end{aligned}
\tag{9.57}
$$

Hence

$$
\begin{aligned}
\mathrm{Tr}[\hat{F}_z^k \tilde{\rho}(t)] \;=\; \sum_{m=-F}^{F} m^k |c_m|^2 \exp[&-2M\eta m^2 t \\
&+ 4mM\eta \int_0^t y(s)\,ds],
\end{aligned}
\tag{9.58}
$$

which is the result used in the text, equation (9.20).

9.7   Moment Evolution via Cumulants

Previously we have discussed the dynamics either in terms of the Gaussian description or with the full stochastic master equation. In this section, we discuss the Gaussian approximation in its most natural form, by deriving the equations of motion for the *cumulants* describing the distribution [68]. Because the cumulants are the proper description of any Gaussian distribution, their interdependent dynamical equations turn out to be *much* easier to derive from the SME than those for other distribution parameterizations. Once these equations are in hand, we can more plainly see when and how the Gaussian approximation of the SME breaks down at long times.

Again, we begin with the same master equation as before

$$
d\hat{\rho} = M\mathcal{D}[\hat{F}_z]\hat{\rho}\,dt + \sqrt{\eta M}\mathcal{H}[\hat{F}_z]\hat{\rho}\,dW.
\tag{9.59}
$$

Notice that for the above without a Hamiltonian we have

$$
\langle \mathcal{D}[\hat{F}_z]\hat{F}_z^n \rangle \;=\; 0
\tag{9.60}
$$

$$
\langle \mathcal{H}[\hat{F}_z]\hat{F}_z^n \rangle \;=\; 2(\langle \hat{F}_z^{n+1} \rangle - \langle \hat{F}_z^n \rangle\langle \hat{F}_z \rangle).
\tag{9.61}
$$

Now we make the following notation simplifications

$$2\sqrt{\eta M} \;=\; 1 \tag{9.62}$$

$$\hat{F}_z \;=\; \hat{x} \tag{9.63}$$

$$v_n \;=\; \langle \hat{x}^n \rangle \tag{9.64}$$

with $v_0 = 1$. It is easily seen from the master equation that the evolution of the *moments* v_n are given by

$$dv_1 \;=\; (v_2 - v_1 v_1)dW \tag{9.65}$$

$$dv_2 \;=\; (v_3 - v_2 v_1)dW \tag{9.66}$$

$$\vdots \tag{9.67}$$

$$dv_n \;=\; (v_{n+1} - v_n v_1)dW. \tag{9.68}$$

Now we are interested in finding the evolution of the cumulants that best describe the distribution [68]. First, we define the *central moments* as

$$\mu_n = \langle (\hat{x} - \langle \hat{x} \rangle)^n \rangle. \tag{9.69}$$

Then we define the corresponding *moment generating function* as

$$M = \sum_{n=0}^{\infty} \frac{v_n \xi^n}{n!} \tag{9.70}$$

such that the moments can be generated using

$$v_n = \frac{d^n M}{d\xi^n}\Big|_{\xi=0}. \tag{9.71}$$

Finally, we define the *cumulants* of the distribution by taking the logarithm of the moment generating function to produce the *cumulant generating function*

$$K \;=\; \ln[M] \tag{9.72}$$

$$\;=\; \sum_{n=1}^{\infty} \frac{\kappa_n \xi^n}{n!} \tag{9.73}$$

149

such that the cumulants can be generated using

$$\kappa_n = \frac{d^n K}{d\xi^n}\Big|_{\xi=0}. \tag{9.74}$$

From this expression, we find that the first few cumulants are

$$\kappa_1 = \langle \hat{x} \rangle \tag{9.75}$$

$$\kappa_2 = \mu_2 \tag{9.76}$$

$$\kappa_3 = \mu_3 \tag{9.77}$$

$$\kappa_4 = \mu_4 - 3\mu_2^2 \tag{9.78}$$

$$\vdots \quad . \tag{9.79}$$

To find the evolution equations for these cumulant terms $\dot{\kappa}_n$ in terms of each other we could naïvely use the above definitions along with Itô's rule, but after the third cumulant or so this becomes way too much work. Of course there is an easier way, which is why we defined them in the first place.

Differentiating dK while treating each v_n as an independent variable gives

$$dK = d\left(\ln[M]\right) \tag{9.80}$$

$$= \sum_{n=0}^{\infty} \frac{dK}{dv_n} dv_n + \frac{1}{2} \sum_{n,m=0}^{\infty} \frac{d^2 K}{dv_n dv_m} dv_n dv_m \tag{9.81}$$

$$= \sum_{n=0}^{\infty} \frac{1}{M} \frac{\xi^n}{n!} dv_n - \frac{1}{2} \sum_{n,m=0}^{\infty} \frac{1}{M^2} \frac{\xi^n}{n!} \frac{\xi^m}{m!} dv_n dv_m \tag{9.82}$$

$$= dW \sum_{n=0}^{\infty} \frac{\xi^n}{n!} (v_{n+1} - v_n v_1) \exp[-K]$$

$$\quad - \frac{dt}{2} \sum_{n,m=0}^{\infty} \frac{\xi^n}{n!} \frac{\xi^m}{m!} (v_{n+1} - v_n v_1)(v_{m+1} - v_m v_1) \exp[-2K] \tag{9.83}$$

$$= dW \exp[-K] \left(\frac{dM}{d\xi} - v_1 M \right) - \frac{dt}{2} \exp[-2K] \left(\frac{dM}{d\xi} - v_1 M \right)^2 \tag{9.84}$$

$$= dW \left(\frac{dK}{d\xi} - \kappa_1 \right) - \frac{dt}{2} \left(\frac{dK}{d\xi} - \kappa_1 \right)^2 \tag{9.85}$$

where we have used $dv_n = (v_{n+1} - v_n v_1)dW$ and $dW^2 = dt$, and $dM/d\xi = \exp[K]dK/d\xi$.

Now we use the notation $K^{(n)} = \frac{d^n K}{d\xi^n}$ to make the last line

$$dK = dW \left(K^{(1)} - \kappa_1 \right) - \frac{dt}{2} \left(K^{(1)} - \kappa_1 \right)^2 \tag{9.86}$$

leading to

$$
\begin{aligned}
d\kappa_n &= d\left(K^{(n)}|_{\xi=0} \right) && \text{(9.87)} \\
&= (dK)^{(n)}|_{\xi=0} && \text{(9.88)} \\
&= dW d\kappa_{n+1} - \frac{dt}{2}((K^{(1)} - \kappa_1)^2)^{(n)}|_{\xi=0}. && \text{(9.89)}
\end{aligned}
$$

Evaluating the latter quantity is just a matter of taking subsequent derivatives

(n)	$-\frac{1}{2}((K^{(1)} - \kappa_1)^2)^{(n)}$	(9.90)
0	$-\frac{1}{2}((K^{(1)} - \kappa_1)^2)$	(9.91)
1	$-(K^{(1)} - \kappa_1)K^{(2)}$	(9.92)
2	$-K^{(2)2} - (K^{(1)} - \kappa_1)K^{(3)}$	(9.93)
3	$-3K^{(2)}K^{(3)} - (K^{(1)} - \kappa_1)K^{(4)}$	(9.94)
4	$-3K^{(3)2} - 4K^{(2)}K^{(4)} - (K^{(1)} - \kappa_1)K^{(5)}$	(9.95)
5	$-10K^{(3)}K^{(4)} - 5K^{(2)}K^{(5)} - (K^{(1)} - \kappa_1)K^{(6)}$	(9.96)
6	$-15K^{(3)}K^{(5)} - 10K^{(4)2} - 6K^{(2)}K^{(6)} - (K^{(1)} - \kappa_1)K^{(7)}$	(9.97)
$\vdots$	$\vdots$	(9.98)

Finally we arrive at what we were seeking, the evolution of the cumulants in terms of each other

$$
\begin{aligned}
d\kappa_1 &= \kappa_2 dW && \text{(9.99)} \\
d\kappa_2 &= \kappa_3 dW - \kappa_2^2 dt && \text{(9.100)} \\
d\kappa_3 &= \kappa_4 dW - 3\kappa_2\kappa_3 dt && \text{(9.101)} \\
d\kappa_4 &= \kappa_5 dW + (-3\kappa_3^2 - 4\kappa_2\kappa_4)dt && \text{(9.102)} \\
\vdots &= \vdots && \text{(9.103)}
\end{aligned}
$$

and the initial conditions for N spin $1/2$ particles aligned along x are

$$\kappa_{1,0} = 0 \tag{9.104}$$

$$\kappa_{2,0} = \frac{N}{4} \tag{9.105}$$

$$\kappa_{3,0} = 0 \tag{9.106}$$

$$\kappa_{4,0} = \frac{-N}{8} \tag{9.107}$$

$$\kappa_{5,0} = 0 \tag{9.108}$$

$$\vdots = \vdots \tag{9.109}$$

Clearly, this is a much more efficient way of representing the dynamics of the distribution than using either the moments or the central moments. Note that the second-order cumulant equation $(d\kappa_2)$ contains a deterministic Riccati term $(-\kappa_2^2 dt)$ but also an initially small stochastic term $(\kappa_3 dW)$. For an ideal Gaussian initial distribution, all cumulants beyond the second are zero by definition. Thus, in this case, one can plainly see exactly how the slightly non-Gaussian nature of the initial spin-state (e.g., $\kappa_{4,0} = \frac{-N}{8} \neq 0$), feeds up the chain of cumulant equations of motion, eventually affecting the low-order cumulants directly.

9.8 Few Level Dynamics

In this section, I consider the evolution of the late time SSE dynamics when all but two levels have been removed from the competition as discussed briefly in section 9.4.3. The dynamics are comparable to the single spin measurements of [22], but not exactly because the final level can have a wide range of possible m values. In the end, we essentially derive the fastest asymptotic rate at which the SSE can project the state into one final value. The fact that the projection rate is bounded is intuitive physically and useful for understanding the late time dynamics of the projection.

Consider the initial state

$$|\Psi\rangle = c_1|m\rangle + c_2|m+1\rangle \tag{9.110}$$

and define the probability of being in $|m\rangle$ to be p such that

$$p = |c_1|^2 \tag{9.111}$$

$$(1-p) = |c_2|^2. \tag{9.112}$$

Setting $M = 1$, $\eta = 1$ in the SSE, we get the very simple SDE

$$dp = -2p(1-p)dW. \tag{9.113}$$

As simple as this SDE is, it does not have an analytic solution. It clearly describes competition between the two stable points of $p = 0$ and $p = 1$. If we continue to assume that the only population is between these two sublevels we can show that $\kappa_2 = p(1-p)$ and that

$$d\kappa_2 = -4\kappa_2^2 dt - 2\kappa_2(1-2p)dW \tag{9.114}$$

if we assume $(1-p) \ll 1$ this becomes

$$d\kappa_2 = -4\kappa_2^2 dt + 2\kappa_2 dW, \tag{9.115}$$

which actually has a solution (see page 77 of [53]) in terms of $W_t = \int_0^t dW_s$

$$\kappa_2 = \frac{\exp[-2t + W_t]}{\kappa_{2,0}^{-1} + 4\int_0^t \exp[-2s + 2W_s]ds}. \tag{9.116}$$

On average this implies that

$$\mathrm{E}[\kappa_2] \propto \exp[-2t] \tag{9.117}$$

and this is used to derive lower bound on curves in plots. In other words, this is the fastest that one of the eigenstates can be prepared. For another m attractor with $m \gg 1$ a similar analysis shows that the best-case decay rate is enhanced by a factor of m^2. As discussed in chapter 4, the higher values of m contain less large scale entanglement. Thus, when entanglement is considered as a resource, it is an intuitive fact that more entangled states take longer to prepare.

9.9 Conclusion

The purpose of this chapter is to demonstrate the fact that the process of continuous projective measurement can be made deterministic with a theoretically simple and intuitive state-based control law. In the context of an atomic spin ensemble, the resulting Dicke states are highly entangled and otherwise difficult to reliably produce from an initially unentangled state.

However, there is much work to be done in the general field of quantum state estimation and control, of which this is one example. In this pursuit, it is helpful to utilize and adapt methods from the developed fields of classical stochastic estimation and control theory. In [22], for example, the problem of this chapter is considered for a single spin with greater emphasis on technical notions of stochastic stability and convergence. Ultimately, we would like to discover constructive methods for deriving optimal control laws given a cost function and realistic actuators.

Even with an optimal control law in hand, there is no guarantee that experimental implementation will be possible. Any analysis should incorporate, among other constraints, nonunity detection efficiencies and finite controller resources (bandwidth, memory, etc). For experimental application of quantum feedback, the controller complexity needs to be reduced to the point where the delay is minimal compared to other dynamical timescales [10]. As in classical control, effective model reduction techniques are indispensable when it comes to implementation.

Despite these difficulties, the increasing number of physical systems that can be measured reliably at the quantum limit will surely hasten the effort to solve many of these technical challenges. By respecting the physical basis of measurement dynamics, experimentalists will be able to more efficiently use measurement itself, in tandem with more traditional techniques, to actuate quantum systems into desirable states.

Chapter 10

Magnetometry Theory

The general notion of quantum parameter estimation was introduced in section 2.7. Here, we discuss the measurement of magnetic fields with our apparatus in terms of quantum parameter estimation. Our initial theoretical magnetometry paper was [18] wherein we used Bayesian analysis with a simplified model of the measurement to derive the expected magnetometry sensitivity. This work is significantly extended in the more complete paper [20] where both open and closed-loop configurations are considered in detail. This chapter is adapted directly from this latter paper, which is titled *Robust quantum parameter estimation: Coherent magnetometry with feedback.* Again for the sake of keeping this chapter independent, some discussion is redundant with discussion in previous chapters. The results of these papers have been considered in the context of traditional magnetometry in [79].

10.1 Abstract

We describe the formalism for optimally estimating and controlling both the state of a spin ensemble and a scalar magnetic field with information obtained from a continuous quantum limited measurement of the spin precession due to the field. The full quantum parameter estimation model is reduced to a simplified equivalent representation to which classical estimation and control theory is applied. We consider both the tracking of static and fluctuating fields in the transient and steady-state regimes. By using feedback control, the field estimation can be made robust to uncertainty about the total spin number.

10.2   Introduction

As experimental methods for manipulating physical systems near their fundamental quantum limits improve [12, 29, 130, 131, 132], the need for quantum state and parameter estimation methods becomes critical. Integrating a modern perspective on quantum measurement theory with the extensive methodologies of classical estimation and control theory provides new insight into how the limits imposed by quantum mechanics affect our ability to measure and control physical systems [19, 133, 134, 75].

In this chapter, we illustrate the processes of state estimation and control for a continuously-observed, coherent spin ensemble (such as an optically pumped cloud of atoms) interacting with an external magnetic field. In the situation where the magnetic field is either zero or well-characterized, continuous measurement (e.g., via the dispersive phase shift or Faraday rotation of a far-off resonant probe beam) can produce a spin-squeezed state [2] conditioned on the measurement record [37]. Spin-squeezing indicates internal entanglement between the different particles in the ensemble [17] and promises to improve precision measurements [5]. When, however, the ambient magnetic environment is either unknown or changing in time, the external field can be estimated by observing Larmor precession in the measurement signal [18, 24, 34, 33], see figure B.1.

Here, we expand on our recent results [18] involving Heisenberg-limited magnetometry by demonstrating the advantages of including feedback control in the estimation process. Feedback is a ubiquitous concept in classical applications because it enables precision performance despite the presence of potentially large system uncertainty. Quantum optical experiments are evolving to the point where feedback can been used, for example, to stabilize atomic motion within optical lattices [131] and high finesse cavities [132]. In this work, demonstrate that, with feedback, an external magnetic field can be measured with high precision despite substantial ignorance of the size of the spin ensemble.

The chapter is organized as follows. In section 10.3, we provide a general introduction to quantum parameter estimation followed by a specialization to the case of a continuously measured spin ensemble in a magnetic field. By capitalizing on the Gaussian properties of both coherent and spin-squeezed states, we formulate the parameter estimation problem in such a way that techniques from classical estimation theory apply to the quantum system. section 10.4 presents basic filtering and control theory in a pedagogical manner with
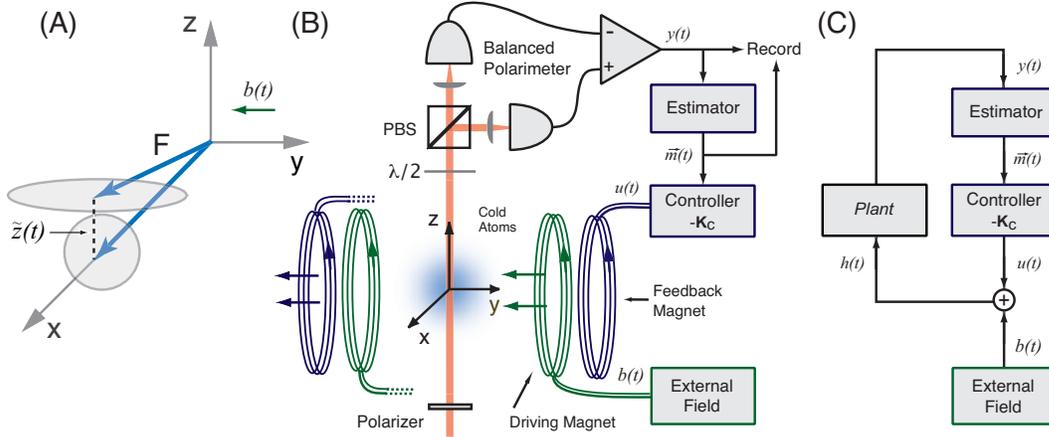
Figure 10.1: (A) A spin ensemble is initially prepared in a coherent state polarized along x, with symmetric variance in the y and z directions. Subsequently, a field along y causes the spin to rotate as the z-component is continuously measured. (B) Experimental schematic for the measurement process. A far-off resonant probe beam traverses the sample and measures the z-component of spin via Faraday rotation. The measurement strength could be improved by surrounding the ensemble with a cavity. (C) Experimental apparatus subsumed by the *Plant* block, which serves to map the total field to the photocurrent, $h(t) \rightarrow y(t)$.

the simplified spin model as an example. This theory is applied in section 10.5, where we simultaneously derive mutually dependent magnetometry and spin-squeezing limits in the ideal case where the observer is certain of the spin number. We consider the optimal measurement of both constant and fluctuating fields in the transient and steady-state regimes. Finally, we show in section 10.6 that the estimation can be made robust to uncertainty about the total spin number by using precision feedback control.

10.3   Quantum Parameter Estimation

First, we present a generic description of quantum parameter estimation [19, 133, 134, 75]. This involves describing the quantum system with a density matrix and our knowledge of the unknown parameter with a classical probability distribution. The objective of parameter estimation is then to utilize information gained about the system through measurement to conditionally update both the density matrix and the parameter distribution. After framing the general case, our particular example of a continuously measured spin ensemble is introduced.

10.3.1 General Problem

The following outline of the parameter estimation process could be generalized to treat a wide class of problems (discrete measurement, multiple parameters), but for simplicity, we will consider a continuously measured quantum system with scalar Hamiltonian parameter θ and measurement record $y(t)$.

Suppose first that the observer has full knowledge of the parameter θ. The proper description of the system would then be a density matrix $\hat{\rho}_\theta(t)$ conditioned on the measurement record $y(t)$. The first problem is to find a rule to update this density matrix with the knowledge obtained from the measurement. As in the problem of this chapter, this mapping may take the form of a stochastic master equation (SME). The SME is by definition a filter that maps the measurement record to an optimal estimate of the system state.

Now if we allow for uncertainty in θ, then a particularly intuitive choice for our new description of the system is

$$\hat{\rho}(t) \equiv \int_\theta \hat{\rho}_\theta(t) p(\theta, t) \, d\theta \tag{10.1}$$

where $p(\theta, t)$ is a probability distribution representing our knowledge of the system parameter. In addition to the rule for updating each $\hat{\rho}_\theta(t)$, we also need to find a rule for updating $p(\theta, t)$ according to the measurement record. By requiring internal consistency, it is possible to find a Bayesian rule for updating $p(\theta, t)$ [19]. These two update rules in principle solve the estimation problem completely.

Because evolving $\hat{\rho}(t)$ involves performing calculations with the full Hilbert space in question, which is often computationally expensive, it is desirable to find a reduced description of the system. Fortunately, it is often possible to find a closed set of dynamical equations for a small set of moments of $\hat{\rho}(t)$. For example, if c is an operator, then we can define the estimate moments

$$\langle \hat{c} \rangle(t) \equiv \mathrm{Tr}[\hat{\rho}(t)\hat{c}]$$
$$\langle \Delta \hat{c}^2 \rangle(t) \equiv \mathrm{Tr}[\hat{\rho}(t)(\hat{c} - \langle \hat{c} \rangle)^2]$$
$$\langle \theta \rangle(t) \equiv \int p(\theta, t)\theta d\theta$$
$$\langle \Delta \theta^2 \rangle(t) \equiv \int p(\theta, t)(\theta - \langle \theta \rangle)^2 d\theta$$

and derive their update rules from the full update rules, resulting in a set of $y(t)$-dependent differential equations. If those differential equations are closed, then this reduced description is adequate for the parameter estimation task at hand. This situation (with closure and Gaussian distributions) is to be expected when the system is approximately linear.

10.3.2 Continuously Measured Spin System

This approach can be applied directly to the problem of magnetometry considered in this chapter. The problem can be summarized by the situation illustrated in figure B.1: a spin ensemble of possibly unknown number is initially polarized along the x-axis (e.g., via optical pumping), an unknown possibly fluctuating scalar magnetic field b directed along the y-axis causes the spins to then rotate within the xz-plane, and the z-component of the collective spin is measured continuously. The measurement can, for example, be implemented as shown, where we observe the difference photocurrent, $y(t)$, in a polarimeter that measures the Faraday rotation of a linearly polarized far off resonant probe beam traveling along z [29, 24, 104]. The goal is to optimally estimate $b(t)$ via the measurement record and unbiased prior information. If a control field $u(t)$ is included, as it will be eventually, the total field is represented by $h(t) = b(t) + u(t)$.

In terms of our previous discussion, we have here the observable $\hat{c} = \sqrt{M}\hat{F}_z$, where M is the measurement rate (defined in terms of probe beam parameters), and the parameter $\theta = b$. When b is known, our state estimate evolves by the stochastic master equation [15]

$$d\hat{\rho}_b(t) = -i[H(b), \hat{\rho}_b(t)]dt + \mathcal{D}[\sqrt{M}\hat{F}_z]\hat{\rho}_b(t)dt$$
$$+\sqrt{\eta}\mathcal{H}[\sqrt{M}\hat{F}_z]\left(2\sqrt{M\eta}[y(t)dt - \langle\hat{F}_z\rangle_b dt]\right)\hat{\rho}_b(t) \tag{10.2}$$

where $\hat{H}(b) = \gamma\hat{F}_y b$, γ is the gyromagnetic ratio, and

$$\mathcal{D}[\hat{c}]\hat{\rho} \equiv \hat{c}\hat{\rho}\hat{c}^\dagger - (\hat{c}^\dagger\hat{c}\hat{\rho} + \hat{\rho}\hat{c}^\dagger\hat{c})/2$$
$$\mathcal{H}[\hat{c}]\hat{\rho} \equiv \hat{c}\hat{\rho} + \hat{\rho}\hat{c}^\dagger - \text{Tr}[(\hat{c} + \hat{c}^\dagger)\hat{\rho}]\hat{\rho}.$$

The stochastic quantity $2\sqrt{M\eta}[y(t)dt - \langle\hat{F}_z\rangle_b(t)dt] \equiv d\bar{W}(t)$ is a Wiener increment (Gaussian white noise with variance dt) by the optimality of the filter. The definition of the photocurrent may be scaled by any constant gain factor, as in [18], as long as the statistics

of the SME remain invariant. The sensitivity of the photodetection per $\sqrt{\text{Hz}}$ is represented by $1/2\sqrt{M\eta}$, where the quantity η represents the quantum efficiency of the detection. If $\eta = 0$, we are essentially ignoring the measurement result and the conditional SME becomes a deterministic unconditional master equation. If $\eta = 1$, the detectors are maximally efficient. Note that our initial state $\hat{\rho}(0) = \hat{\rho}_b(0)$ is made equal to a coherent state (polarized in x) and is representative of our prior information.

The stochastic master equation, equation (B.43), has previously been derived for homodyne detection of the output of a cavity with a single mode dispersively coupled to the collective atomic spin within the cavity [15]. The resulting form of the equation is, however, the most generic form of a continuous stochastic QND measurement and also applies under similar approximations to the free-space Faraday rotation measurement [104] diagrammed in figure B.1.

It can be shown that the unnormalized probability $\bar{p}(b, t)$ evolves according to [19]

$$d\bar{p}(b, t) = 4M\eta\langle\hat{F}_z\rangle_b(t)\bar{p}(b, t)y(t)dt. \tag{10.3}$$

The evolution equations (B.43, 10.3 together with equation (10.1) solve the problem completely, albeit in a computationally expensive way. Clearly, for large ensembles it would be advantageous to reduce the problem to a simpler description.

If we consider only the estimate moments $\langle\hat{F}_z\rangle(t)$, $\langle\Delta\hat{F}_z^2\rangle(t)$, $\langle b\rangle(t)$, and $\langle\Delta b^2\rangle(t)$ and derive their evolution with the above rules, it can be shown that the filtering equations for those variables are closed under certain approximations. First, the spin number F must be large enough that the distributions for $\hat{F}_y$ and $\hat{F}_z$ are approximately Gaussian for an x-polarized coherent state. Second, we only consider times $t \ll 1/M$ because the total spin becomes damped by the measurement at times comparable to the inverse of the measurement rate.

Although this approach is rigorous and fail-safe, the resulting filtering equations for the moments can be arrived at in a more direct manner as discussed in the appendix of [20]. Essentially, the full quantum mechanical mapping from $h(t)$ to $y(t)$ is equivalent to the mapping derived from a model that appears classical, and assumes an actual, but random, value for the z-component of spin. This correspondence generally holds for a stochastic master equation corresponding to an arbitrary linear quantum mechanical system

with continuous measurement of observables that are linear combinations of the canonical variables [70].

From this point on we will only consider the simplified Gaussian representation (used in the next section) since it allows us to apply established techniques from estimation and control theory. The replacement of the quantum mechanical model with a classical noise model is discussed more fully in the appendix. Throughout this treatment, we keep in mind the constraints that the original model imposed. As before, we assume F is large enough to maintain the Gaussian approximation and that time is small compared to the measurement induced damping rate, $t \ll 1/M$. Also, the description of our original problem demands that $\langle \Delta \hat{F}_z^2 \rangle (0) = F/2$ for a coherent state. [1] Hence our prior information for the initial value of the spin component will always be dictated by the structure of Hilbert space.

10.4   Optimal Estimation and Control

We now describe the dynamics of the simplified representation. Given a linear state-space model (L), a quadratic performance criterion (Q) and Gaussian noise (G), we show how to apply standard LQG analysis to optimize the estimation and control performance [49].

The system state we are trying to estimate is represented by

State

$$\mathbf{x}(t) \equiv \begin{bmatrix} z(t) \\ b(t) \end{bmatrix} \tag{10.4}$$

where $z(t)$ represents the small z-component of the collective angular momentum and $b(t)$ is a scalar field along the y-axis.

Our best guess of $\mathbf{x}(t)$, as we filter the measurement record, will be denoted

Estimate

$$\mathbf{m}(t) \equiv \begin{bmatrix} \tilde{z}(t) \\ \tilde{b}(t) \end{bmatrix}. \tag{10.5}$$

---

[1] We assume throughout the chapter that we have a system of N spin-1/2 particles, so for a polarized state along x, $\langle \hat{F}_x \rangle = F = N/2$ and $\sigma_{z0} = \langle \Delta \hat{F}_z^2 \rangle (0) = F/2 = N/4$. This is an arbitrary choice and our results are independent of any constituent spin value, apart from defining these moments. In [29], for example, we work with an ensemble of cesium atoms, each atom in a ground state of spin-4.

As stated in the appendix, we implicitly make the associations: $\tilde{z}(t) = \langle \hat{F}_z \rangle(t) = \text{Tr}[\hat{\rho}(t)\hat{F}_z]$ and $\tilde{b}(t) = \int p(b,t) b \, db$, although no further mention of $\hat{\rho}(t)$ or $p(b,t)$ will be made.

We assume the measurement induced damping of F to be negligible for short times ($F \exp[-Mt/2] \approx F$ if $t \ll 1/M$) and approximate the dynamics as

Dynamics

$$d\mathbf{x}(t) = \mathbf{A}\mathbf{x}(t)dt + \mathbf{B}u(t)dt + \begin{bmatrix} 0 \\ \sqrt{\sigma_{bF}} \end{bmatrix} dW_1 \tag{10.6}$$

$$\mathbf{A} \equiv \begin{bmatrix} 0 & \gamma F \\ 0 & -\gamma_b \end{bmatrix}$$

$$\mathbf{B} \equiv \begin{bmatrix} \gamma F \\ 0 \end{bmatrix}$$

$$\mathbf{\Sigma_0} \equiv \begin{bmatrix} \sigma_{z0} & 0 \\ 0 & \sigma_{b0} \end{bmatrix}$$

$$\mathbf{\Sigma_1} \equiv \begin{bmatrix} 0 & 0 \\ 0 & \sigma_{bF} \end{bmatrix}$$

where the initial value $\mathbf{x}(0)$ for each trial is drawn randomly from a Gaussian distribution of mean zero and covariance matrix $\mathbf{\Sigma_0}$. The initial field variance σ_{b0} is considered to be due to classical uncertainty, whereas the initial spin variance σ_{z0} is inherently nonzero due to the original quantum state description. Specifically, we impose $\sigma_{z0} = \langle \Delta \hat{F}_z^2 \rangle(0)$. The Wiener increment $dW_1(t)$ has a Gaussian distribution with mean zero and variance dt. $\mathbf{\Sigma_1}$ represents the covariance matrix of the last vector in equation (10.6).

We have given ourselves a magnetic field control input, $u(t)$, along the same axis, y, of the field to be measured, $b(t)$. We have allowed $b(t)$ to fluctuate via a damped diffusion (Ornstein-Uhlenbeck) process [52]

$$db(t) = -\gamma_b b(t) \, dt + \sqrt{\sigma_{bF}} \, dW_1. \tag{10.7}$$

The $b(t)$ fluctuations are represented in this particular way because Gaussian noise processes are amenable to LQG analysis. The variance of the field at any particular time is given by the expectation $\sigma_{bFree} \equiv \text{E}[b(t)^2] = \sigma_{bF}/2\gamma_b$. (Throughout the chapter we use the notation

E[$x(t)$] to represent the average of the generally stochastic variable $x(t)$ at the same point in time, over many trajectories.) The bandwidth of the field is determined by the frequency γ_b alone. When considering the measurement of fluctuating fields, a valid choice of prior might be $\sigma_{b0} = \sigma_{bFree}$, but we choose to let σ_{b0} remain independent. For constant fields, we set $\sigma_{bFree} = 0$, but $\sigma_{b0} \neq 0$.

Note that only the small angle limit of the spin motion is considered. Otherwise we would have to consider different components of the spin vector rotating into each other. The small angle approximation would be invalid if a field caused the spins to rotate excessively, but using adequate control ensures this will not happen. Hence, we use control for essentially two reasons in this chapter: first to keep our small angle approximation valid, and, second, to make our estimation process robust to our ignorance of F. The latter point will be discussed in section 10.6.

Our measurement of z is described by the process

Measurement

$$
\begin{aligned}
y(t)dt &= \mathbf{C}\mathbf{x}(t)dt + \sqrt{\sigma_M}dW_2(t) \\
\mathbf{C} &\equiv \begin{bmatrix} 1 & 0 \end{bmatrix} \\
\mathbf{\Sigma_2} &\equiv \sigma_M \equiv 1/4M\eta
\end{aligned}
\tag{10.8}
$$

where the measurement shotnoise is represented by the Wiener increment $dW_2(t)$ of variance dt. Again, $\sqrt{\sigma_M}$ represents the sensitivity of the measurement, M is the measurement rate (with unspecified physical definition in terms of probe parameters), and η is the quantum efficiency of the measurement. The increments dW_1 and dW_2 are uncorrelated.

Following [49], the optimal estimator for mapping $y(t)$ to $\mathbf{m}(t)$ takes the form

Estimator

$$
\begin{aligned}
d\mathbf{m}(t) &= \mathbf{A}\mathbf{m}(t)dt + \mathbf{B}u(t)dt \\
&\quad + \mathbf{K}_O(t)[y(t) - \mathbf{C}\mathbf{m}(t)]dt & (10.9)\\
\mathbf{m}(0) &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\
\mathbf{K}_O(t) &\equiv \boldsymbol{\Sigma}(t)\mathbf{C}^T\boldsymbol{\Sigma}_2^{-1} \\
\frac{d\boldsymbol{\Sigma}(t)}{dt} &= \boldsymbol{\Sigma}_1 + \mathbf{A}\boldsymbol{\Sigma}(t) + \boldsymbol{\Sigma}(t)\mathbf{A}^T \\
&\quad - \boldsymbol{\Sigma}(t)\mathbf{C}^T\boldsymbol{\Sigma}_2^{-1}\mathbf{C}\boldsymbol{\Sigma}(t) & (10.10)\\
\boldsymbol{\Sigma}(t) &\equiv \begin{bmatrix} \sigma_{zR}(t) & \sigma_{cR}(t) \\ \sigma_{cR}(t) & \sigma_{bR}(t) \end{bmatrix} & (10.11)\\
\boldsymbol{\Sigma}(0) &= \boldsymbol{\Sigma_0} \equiv \begin{bmatrix} \sigma_{z0} & 0 \\ 0 & \sigma_{b0} \end{bmatrix}. & (10.12)
\end{aligned}
$$

equation (10.9) is the Kalman filter that depends on the solution of the matrix Riccati equation (10.10). The Riccati equation gives the optimal observation gain $\mathbf{K}_O(t)$ for the filter. The estimator is designed to minimize the average quadratic estimation error for each variable: $\mathrm{E}[(z(t) - \tilde{z}(t))^2]$ and $\mathrm{E}[(b(t) - \tilde{b}(t))^2]$. If the model is correct, and we assume the observer chooses his prior information $\boldsymbol{\Sigma}(0)$ to match the actual variance of the initial data $\boldsymbol{\Sigma_0}$, then we have the self-consistent result:

$$
\begin{aligned}
\sigma_{zE}(t) &\equiv \mathrm{E}[(z(t) - \tilde{z}(t))^2] = \sigma_{zR}(t) \\
\sigma_{bE}(t) &\equiv \mathrm{E}[(b(t) - \tilde{b}(t))^2] = \sigma_{bR}(t).
\end{aligned}
$$

Hence, the Riccati equation solution represents both the observer gain *and* the expected performance of an optimal filter using that same gain.

Now consider the control problem, which is in many respects dual to the estimation problem. We would like to design a controller to map $y(t)$ to $u(t)$ in a manner that minimizes the quadratic cost function

Minimized Cost

$$
\begin{aligned}
H &= \int_0^T \left[\mathbf{x}^T(t)\mathbf{P}\mathbf{x}(t) + u(t)\mathbf{Q}u(t)\right] dt \\
&\quad + \mathbf{x}^T(T)\mathbf{P}_1\mathbf{x}(T) \tag{10.13} \\
\mathbf{P} &\equiv \begin{bmatrix} p & 0 \\ 0 & 0 \end{bmatrix} \\
\mathbf{Q} &\equiv q
\end{aligned}
$$

where $\mathbf{P}_1$ is the end-point cost. Only the ratio p/q ever appears, of course, so we define the parameter $\lambda \equiv \sqrt{p/q}$ and use it to represent the cost of control. By setting $\lambda \to \infty$, as we often choose to do in the subsequent analysis to simplify results, we are putting no cost on our control output. This is unrealistic because, for example, making λ arbitrarily large implies that we can apply transfer functions with finite gain at arbitrarily high frequencies, which is not experimentally possible. Despite this, we will often consider the limit $\lambda \to \infty$ to set bounds on achievable estimation and control performance. The optimal controller for minimizing equation (10.13) is

Controller

$$
\begin{aligned}
u(t) &= -\mathbf{K}_C(t)\mathbf{m}(t) \tag{10.14} \\
\mathbf{K}_C(t) &\equiv \mathbf{Q}^{-1}\mathbf{B}^T\mathbf{V}(T-t) \\
\frac{d\mathbf{V}(T)}{dT} &= \mathbf{P} + \mathbf{A}^T\mathbf{V}(T) + \mathbf{V}(T)\mathbf{A} \\
&\quad - \mathbf{V}(T)\mathbf{B}\mathbf{Q}^{-1}\mathbf{B}^T\mathbf{V}(T) \tag{10.15} \\
\mathbf{V}(T=0) &\equiv \mathbf{P}_1.
\end{aligned}
$$

Here $\mathbf{V}(T)$ is solved in reverse time T, which can be interpreted as the time left to go until the stopping point. Thus if $T \to \infty$, then we only need to use the steady-state of the $\mathbf{V}$ Riccati equation (10.15) to give the steady state controller gain $\mathbf{K}_C$ for all times. In this case, we can ignore the (reverse) initial condition P_1 because the controller is not designed to stop. Henceforth, we will make $\mathbf{K}_C$ equal to this constant steady-state value, such that the only time varying coefficients will come from $\mathbf{K}_O(t)$.

In principle, the above results give the entire solution to the ideal estimation and control problem. However, in the nonideal case where our knowledge of the system is incomplete,

e.g., F is unknown, our estimation performance will suffer. Notation is now introduced that produces trivial results in the ideal case, but is helpful otherwise. Our goal is to collect the above equations into a single structure that can be used to solve the nonideal problem. We define the *total* state of the system and estimator as

Total State

$$\theta(t) \equiv \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{m}(t) \end{bmatrix} = \begin{bmatrix} z(t) \\ b(t) \\ \tilde{z}(t) \\ \tilde{b}(t) \end{bmatrix}. \tag{10.16}$$

Consider the general case where the observer assumes the plant contains spin F', which may or may not be equal to the actual F. All design elements depending on F' instead of F are now labeled with a prime. Then it can be shown that the total state dynamics from the above estimator-controller architecture are a time-dependent Ornstein-Uhlenbeck process

Total State Dynamics

$$d\boldsymbol{\theta}(t) = \boldsymbol{\alpha}(t)\boldsymbol{\theta}(t)dt + \boldsymbol{\beta}(t)d\mathbf{W}(t) \tag{10.17}$$

$$\boldsymbol{\alpha}(t) \equiv \begin{bmatrix} \mathbf{A} & -\mathbf{B}\mathbf{K}'_C \\ \mathbf{K}'_O(t)\mathbf{C} & \mathbf{A}' - \mathbf{B}'\mathbf{K}'_C - \mathbf{K}'_O(t)\mathbf{C} \end{bmatrix}$$

$$\boldsymbol{\beta}(t) \equiv \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \sqrt{\sigma_{bF}} & 0 & 0 \\ 0 & 0 & \sqrt{\sigma_M}K'_{O1}(t) & 0 \\ 0 & 0 & \sqrt{\sigma_M}K'_{O2}(t) & 0 \end{bmatrix}$$

where the covariance matrix of $d\mathbf{W}$ is dt times the identity. Now the quantity of interest is the following covariance matrix

Total State Covariance

$$\boldsymbol{\Theta}(t) \equiv \mathrm{E}[\boldsymbol{\theta}(t)\boldsymbol{\theta}^T(t)]$$

$$\equiv \begin{bmatrix} \sigma_{zz} & \sigma_{zb} & \sigma_{z\tilde{z}} & \sigma_{z\tilde{b}} \\ \sigma_{zb} & \sigma_{bb} & \sigma_{b\tilde{z}} & \sigma_{b\tilde{b}} \\ \sigma_{z\tilde{z}} & \sigma_{b\tilde{z}} & \sigma_{\tilde{z}\tilde{z}} & \sigma_{\tilde{z}\tilde{b}} \\ \sigma_{z\tilde{b}} & \sigma_{b\tilde{b}} & \sigma_{\tilde{z}\tilde{b}} & \sigma_{\tilde{b}\tilde{b}} \end{bmatrix} \tag{10.18}$$

$$\sigma_{zz} \equiv \mathrm{E}[z(t)^2]$$

$$\sigma_{zb} \equiv \mathrm{E}[z(t)b(t)]$$

$$\vdots \equiv \ddots$$

It can be shown that this total covariance matrix obeys the following deterministic equations of motion,

Total State Covariance Dynamics

$$\frac{d\boldsymbol{\Theta}(t)}{dt} = \boldsymbol{\alpha}(t)\boldsymbol{\Theta}(t) + \boldsymbol{\Theta}(t)\boldsymbol{\alpha}^T(t) + \boldsymbol{\beta}(t)\boldsymbol{\beta}^T(t) \tag{10.19}$$

$$\boldsymbol{\Theta}(t) = \exp\left[-\int_0^t \boldsymbol{\alpha}(t')dt'\right]\boldsymbol{\Theta}_0 \exp\left[-\int_0^t \boldsymbol{\alpha}^T(t')dt'\right]$$
$$+ \int_0^t dt' \exp\left[-\int_{t'}^t \boldsymbol{\alpha}(s)ds\right]\boldsymbol{\beta}(t')\boldsymbol{\beta}^T(t')$$
$$\times \exp\left[-\int_{t'}^t \boldsymbol{\alpha}^T(s)ds\right] \tag{10.20}$$

$$\boldsymbol{\Theta}_0 = \begin{bmatrix} \sigma_{z0} & 0 & 0 & 0 \\ 0 & \sigma_{b0} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Equation (10.20) is the matrix form of the standard integrating factor solution for time-dependent scalar ordinary differential equations [52]. Whether we solve this problem numerically or analytically, the solution provides the following quantity that we ultimately care about,

Average Magnetometry Error

$$
\begin{aligned}
\sigma_{bE}(t) &\equiv \mathrm{E}[(\tilde{b}(t) - b(t))^2] \\
&= \mathrm{E}[b^2(t)] + \mathrm{E}[\tilde{b}^2(t)] - 2\mathrm{E}[b(t)\tilde{b}(t)] \\
&= \sigma_{bb}(t) + \sigma_{\tilde{b}\tilde{b}}(t) - 2\sigma_{b\tilde{b}}(t).
\end{aligned}
\tag{10.21}
$$

When all parameters are known (and $F' = F$), this total state description is unnecessary because $\sigma_{bE}(t) = \sigma_{bR}(t)$. This equality is by *design*. However, when the wrong parameters are assumed (e.g., $F' \neq F$) the equality does not hold $\sigma_{bE}(t) \neq \sigma_{bR}(t)$ and either equation (10.19) or equation (10.20) must be used to find $\sigma_{bE}(t)$. Before addressing this problem, we consider in detail the performance in the ideal case, where all system parameters are known by the observer, including F.

At this point, we have defined several variables. For clarity, let us review the meaning of several before continuing. Inputs to the problem include the field fluctuation strength σ_{bF}, equation (10.7), and the measurement sensitivity σ_M, equation (10.8). The prior information for the field is labeled σ_{b0}, equation (10.12). The solution to the Riccati equation is $\sigma_{bR}(t)$, equation (10.11), and is equal to the estimation variance $\sigma_{bE}(t)$, equation (10.21), when the estimator model is correct. In the next section, we additionally use σ_{bS}, equation (10.24), and $\sigma_{bT}(t)$, equation (10.25), to represent the steady-state and transient values of $\sigma_{bE}(t)$ respectively.

10.5 Optimal Performance: F Known

We start by observing qualitative characteristics of the b-estimation dynamics. Figure 10.2 shows the average estimation performance, $\sigma_{bR}(t)$, as a function of time for a realistic set of parameters. Notice that σ_{bR} is constant for small and large times, below t_1 and above t_2. If σ_{b0} is noninfinite then the curve is constant for small times, as it takes some time to begin improving the estimate from the prior. If σ_{b0} is infinite, then $t_1 = 0$ and the sloped transient portion extends towards infinity as $t \to 0$. At long times, σ_{bR} will become constant again, but only if the field is fluctuating ($\sigma_{bF} \neq 0$ and $\gamma_b \neq 0$). The performance saturates because one can track a field only so well if the field is changing and the signal-to-noise ratio is finite. If the field to be tracked is constant, then $t_2 = \infty$ and the sloped portion of
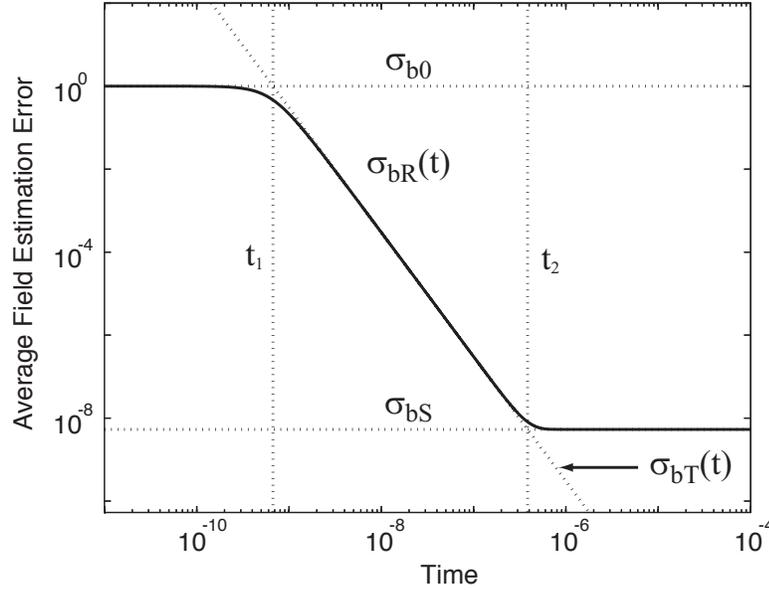
Figure 10.2: The Riccati equation solution gives the ideal field estimation performance. The parameters used here are $F = 10^6$, $\sigma_{z0} = F/2$ (for an ensemble of spin-1/2 particles), $\gamma = 10^6$, $M = 10^4$, $\sigma_{b0} = \sigma_{bFree} = 1$. (All quantities within figures are kept dimensionless, although expressions within the text may be interpreted as having dimension.) The solution starts at the free field fluctuation variance and saturates at σ_{bS}. The plot is not valid at times $t > 1/M$.

the curve extends to zero as $t \to \infty$ (given the approximations discussed in section 10.3.2). After the point where the performance saturates ($t \gg t_2$), all of the observer and control gains have become time independent and the filter can be described by a transfer function.

However, as will be shown, applying only this steady-state transfer function is nonoptimal in the transient regime ($t_1 \ll t \ll t_2$), because the time dependence of the gains is clearly crucial for optimal transient performance.

10.5.1 Steady-State Performance

We start by examining the steady-state performance of the filter. At large enough times (where we have yet to define large enough), $\mathbf{K}_O$ becomes constant and if we set $T \to \infty$ (ignoring the end-point cost), then $\mathbf{K}_C$ is always constant. Setting $d\mathbf{\Sigma}/dt = 0$ and $d\mathbf{V}/dt =$

0 we find:

$$\mathbf{K}_O(t) \;\rightarrow\; \begin{bmatrix} \sqrt{2\gamma F\sqrt{\frac{\sigma_{bF}}{\sigma_M}} + \gamma_b^2} - \gamma_b \\[2mm] \sqrt{\frac{\sigma_{bF}}{\sigma_M}} - \frac{\gamma_b}{\gamma F}\left(\sqrt{2\gamma F\sqrt{\frac{\sigma_{bF}}{\sigma_M}} + \gamma_b^2} - \gamma_b\right) \end{bmatrix}$$

$$\mathbf{K}_C(t) \;\rightarrow\; \begin{bmatrix} \lambda & 1/(1 + \frac{\gamma_b}{\gamma F \lambda}) \end{bmatrix}$$

where $\lambda = \sqrt{\frac{p}{q}}$.

Now assuming the gains to be constant, we can derive the three relevant transfer functions from $y(t)$ to $\mathbf{m}(t)$ ($\tilde{z}$ and $\tilde{b}$) and u. We proceed as follows. First, we express the estimates in terms of only themselves and the photocurrent

$$\begin{aligned}
\frac{d\mathbf{m}(t)}{dt} &= \mathbf{A}\mathbf{m}(t) + \mathbf{B}u(t) + \mathbf{K}_O(y(t) - \mathbf{C}\mathbf{m}(t)) \\
&= \mathbf{A}\mathbf{m}(t) + \mathbf{B}(-\mathbf{K}_C\mathbf{m}(t)) + \mathbf{K}_O(y(t) - \mathbf{C}\mathbf{m}(t)) \\
&= (\mathbf{A} - \mathbf{B}\mathbf{K}_C - \mathbf{K}_O\mathbf{C})\mathbf{m}(t) + \mathbf{K}_O y(t).
\end{aligned}$$

To get the transfer functions, we take the Laplace transform of the entire equation, use differential transform rules to give s factors (where $s = j\omega$, $j = \sqrt{-1}$), ignore initial condition factors, and rearrange terms. However, this process only gives meaningful transfer functions if the coefficients $\mathbf{K}_O$ and $\mathbf{K}_C$ are constant. Following this procedure, we have

$$\begin{aligned}
\mathbf{m}(s) &= (s\mathbf{I} - \mathbf{A} + \mathbf{B}\mathbf{K}_C + \mathbf{K}_O\mathbf{C})^{-1}\mathbf{K}_O y(s) \\
&= \mathbf{G}_m(s)y(s) \\
u(s) &= -\mathbf{K}_C\mathbf{m}(s) \\
&= -\mathbf{K}_C(s - \mathbf{A} + \mathbf{B}\mathbf{K}_C + \mathbf{K}_O\mathbf{C})^{-1}\mathbf{K}_O y(s) \\
&= G_u(s)y(s)
\end{aligned}$$

where

$$\mathbf{G}_m(s) = \begin{bmatrix} G_z(s) \\ G_b(s) \end{bmatrix}.$$

The three transfer functions ($G_z(s)$, $G_b(s)$, and $G_u(s)$) serve three different tasks. If

estimation is the concern, then $G_b(s)$ will perform optimally in steady-state. Notice that, while the Riccati solution is the same with and without control ($\mathbf{K}_C$ nonzero or zero), this transfer function is not the same in the two cases. So, even though the transfer functions are different, they give the same steady-state performance.

Let us now consider the controller transfer function $G_u(s)$ in more detail. We find the controller to be of the form

$$G_u(s) = G_{u,DC}\frac{1 + s/\omega_H}{1 + (1 + s/\omega_Q)s/\omega_L}. \tag{10.22}$$

Here each frequency ω represents a transition in the Bode plot of figure 10.3. A similar controller transfer function is derived via a different method in section 10.8.

If we are not constrained experimentally, we can make the approximations $\lambda^2 \gg \sqrt{\sqrt{\sigma_{bF}/\sigma_M}/2\gamma F}$ and $\gamma F \gg \gamma_b^2\sqrt{\sigma_M/\sigma_{bF}}$ giving

$$
\begin{aligned}
G_u(s) &\rightarrow G_{u,DC}\frac{1 + s/\omega_H}{1 + s/\omega_L} \\
\omega_L &\rightarrow \gamma_b \\
\omega_H &\rightarrow \sqrt{\frac{\gamma F}{2}\sqrt{\frac{\sigma_{bF}}{\sigma_M}}} \\
\omega_C &\rightarrow \sqrt{2\gamma F\sqrt{\frac{\sigma_{bF}}{\sigma_M}}} = 2\omega_H \\
\omega_Q &\rightarrow \lambda\gamma F \\
G_{u,DC} &\rightarrow -\frac{1}{\gamma_b}\sqrt{\frac{\sigma_{bF}}{\sigma_M}} \\
G_{u,AC} &\rightarrow G_{u,DC}\frac{\omega_L}{\omega_H} = -\sqrt{\frac{2}{\gamma F}\sqrt{\frac{\sigma_{bF}}{\sigma_M}}}
\end{aligned}
$$

where $G_{u,AC}$ is the gain at high frequencies ($\omega > \omega_H$) and we find the closing frequency ω_C from the condition $|P_z(j\omega_C)G_u(j\omega_C)| = 1$, with the plant transfer function being the normal integrator $P_z(s) = \gamma F/s$. Notice that the controller closes in the very beginning of the flat high frequency region (hence with adequate phase margin) because $\omega_C = 2\omega_H$.

Finally, consider the steady-state estimation performance. These are the same with and without control (hence λ-independent) and, under the simplifying assumption $\gamma F \gg$
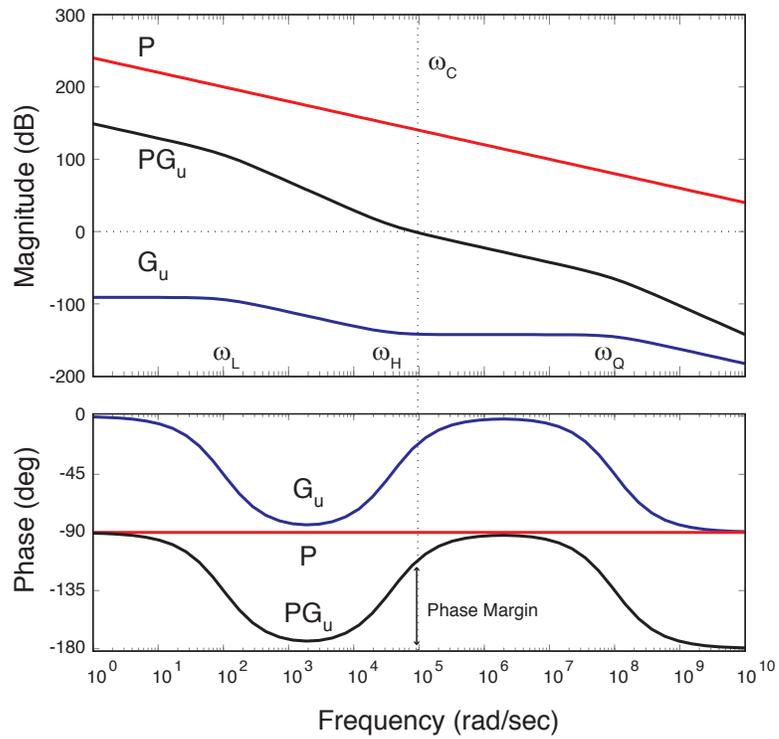
Figure 10.3: The Bode plot of $G_u(s)$, the transfer function of the filter in steady-state, for a typical parameter regime. Notice that the controller closes the plant with adequate phase margin to avoid closed-loop instability. At high frequencies the controller rolls off at ω_Q if $\lambda \neq \infty$.

Table 10.1: Field tracking error, $\sigma_{bR}(t)$, for different initial variances of b and z

	$\sigma_{b0} = 0$	σ_{b0}	$\sigma_{b0} \to \infty$
$\sigma_{z0} = 0$	0	$\dfrac{3\sigma_{b0}\sigma_M}{3\sigma_M + \gamma^2 F^2 \sigma_{b0} t^3}$	$\dfrac{3\sigma_M}{\gamma^2 F^2 t^3}$
σ_{z0}	0	$\dfrac{12\sigma_{b0}\sigma_M(\sigma_M + \sigma_{z0}t)}{12\sigma_M^2 + \gamma^2 F^2 \sigma_{b0}\sigma_{z0}t^4 + 4\sigma_M(3\sigma_{z0}t + \gamma^2 F^2 t^3 \sigma_{b0})}$	$\dfrac{12\sigma_M(\sigma_M + \sigma_{z0}t)}{\gamma^2 F^2 t^3 (4\sigma_M + \sigma_{z0}t)}$
$\sigma_{z0} \to \infty$	0	$\dfrac{12\sigma_{b0}\sigma_M}{12\sigma_M + \gamma^2 F^2 t^3 \sigma_{b0}}$	$\dfrac{12\sigma_M}{\gamma^2 F^2 t^3}$

$\gamma_b^2 \sqrt{\sigma_M/\sigma_{bF}}$, are given by

$$\sigma_{zR}(t) \quad \to \quad \sqrt{2\gamma F}\, \sigma_M^{3/4} \sigma_{bF}^{1/4} \equiv \sigma_{zS} \tag{10.23}$$

$$\sigma_{bR}(t) \quad \to \quad \sqrt{\frac{2}{\gamma F}}\, \sigma_{bF}^{3/4} \sigma_M^{1/4} \equiv \sigma_{bS}. \tag{10.24}$$

If the estimator reaches steady-state at $t \ll 1/M$, then the above variance σ_{zR} represents a limit to the amount of spin-squeezing possible in the presence of fluctuating fields.

Also the F scaling of the saturated field sensitivity $\sigma_{bR} \propto F^{-1/2}$ is not nearly as strong as the F scaling in the transient period $\sigma_{bR} \propto F^{-2}$. Next, we demonstrate this latter result as we move from the steady-state analysis to calculating the estimation performance during the transient period.

10.5.2 Transient Performance

We now consider the transient performance of the ideal filter: how quickly and how well the estimator-controller will *lock* onto the signal and achieve steady-state performance. In many control applications, the transient response is not of interest because the time it takes to acquire the lock is negligible compared to the long steady-state period of the system. However, in systems where the measurement induces continuous decay, this transient period can be a significant portion of the total lifetime of the experiment.

We will evaluate the transient performance of two different filters. First, we look at the ideal dynamic version, with time dependent observer gains derived from the Riccati equation. This limits to a transfer function at long times when the gains have become constant. Second, we numerically look at the case where the same steady-state transfer functions are used for the *entire* duration of the measurement. Because the gains are not adjusted smoothly, the small time performance of this estimator suffers. Of course, for long

Table 10.2: Spin tracking error, $\sigma_{zR}(t)$, for different initial variances of b and z

	$\sigma_{b0} = 0$	σ_{b0}	$\sigma_{b0} \to \infty$
$\sigma_{z0} = 0$	0	$\dfrac{3\gamma^2 F^2 \sigma_{b0}\sigma_M t^2}{3\sigma_M + \gamma^2 F^2 \sigma_{b0} t^3}$	$\dfrac{3\sigma_M}{t}$
σ_{z0}	$\dfrac{\sigma_M \sigma_{z0}}{\sigma_M + \sigma_{z0} t}$	$\dfrac{4\sigma_M(\gamma^2 F^2 \sigma_{b0}\sigma_{z0} t^3 + 3\sigma_M(\sigma_{z0} + \gamma^2 F^2 t^2 \sigma_{b0}))}{12\sigma_M^2 + \gamma^2 F^2 \sigma_{b0}\sigma_{z0} t^4 + 4\sigma_M(3\sigma_{z0}t + \gamma^2 F^2 t^3 \sigma_{b0})}$	$\dfrac{4\sigma_M(3\sigma_M + \sigma_{z0}t)}{t(4\sigma_M + \sigma_{z0}t)}$
$\sigma_{z0} \to \infty$	$\dfrac{\sigma_M}{t}$	$\dfrac{4\sigma_M(3\sigma_M + \gamma^2 F^2 t^3 \sigma_{b0})}{12\sigma_M t + \gamma^2 F^2 t^4 \sigma_{b0}}$	$\dfrac{4\sigma_M}{t}$

times the estimators are equivalent.

10.5.2.1 Dynamic Estimation and Control

Now consider the transient response of $\mathbf{\Sigma}(t)$ (giving $\mathbf{K}_O(t)$). We will continue to impose that $\mathbf{V}$ (thus $\mathbf{K}_C$) is constant because we are not interested in any particular stopping time.

The Riccati equation for $\mathbf{\Sigma}(t)$ (equation (10.10)) appears difficult to solve because it is nonlinear. Fortunately, it can be reduced to a much simpler linear problem. See section 10.7 for an outline of this method.

The solution to the fluctuating field problem ($\sigma_{bF} \neq 0$ and $\gamma_b \neq 0$) is represented in figure 10.2. This solution is simply the constant field solution ($\sigma_{bF} = 0$ and $\gamma_b = 0$) smoothly saturating at the steady-state value of equation (10.24) at time t_2. Thus, considering the long time behavior of the constant field solution will tell us about the transient behavior when measuring fluctuating fields. Because the analytic form for the constant field solution is simple, we consider only it and disregard the full analytic form of the fluctuating field solution.

The analytic form of $\mathbf{\Sigma}(t)$ is highly instructive. The general solutions to $\sigma_{bR}(t)$ and $\sigma_{zR}(t)$, with arbitrary prior information σ_{b0} and σ_{z0}, are presented in the central entries of Tables 10.1 and 10.2 respectively. The other entries of the tables represent the limits of these somewhat complicated expressions as the prior information assumes extremely large or small values. Here, we notice several interesting trade-offs.

First, the left hand column of Table 10.1 is zero because if a constant field is being measured, and we start with complete knowledge of the field ($\sigma_{b0} = 0$), then our job is completed trivially. Now notice that if σ_{b0} and σ_{z0} are both nonzero, then at long times we

have the lower-right entry of Table 10.1

$$\sigma_{bR}(t) = \frac{12\sigma_M}{\gamma^2 F^2 t^3} \equiv \sigma_{bT}(t). \tag{10.25}$$

This is the same result one gets when the estimation procedure is simply to perform a least-squares line fit to the noisy measurement curve for constant fields. (Note that all of these results are equivalent to the solutions of [18], but without F damping.) If it were physically possible to ensure $\sigma_{z0} = 0$, then our estimation would improve by a factor of four to the upper-right result. However, quantum mechanics imposes that this initial variance is nonzero (e.g., $\sigma_{z0} = F/2$ for a coherent state and less, but still nonzero, for a squeezed state), and the upper-right solution is unattainable.

Now consider the dual problem of spin estimation performance $\sigma_{zR}(t)$ as represented in Table 10.2, where we can make analogous trade-off observations. If there is no field present, we set $\sigma_{b0} = 0$ and

$$\sigma_{zR}(t) = \frac{\sigma_{z0}\sigma_M}{\sigma_M + t\sigma_{z0}}. \tag{10.26}$$

When $\sigma_{zR}(t)$ is interpreted as the quantum variance $\langle \Delta \hat{F}_z^2 \rangle(t)$, this is the ideal (nondamped) conditional spin-squeezing result, which is valid at $t \ll 1/M$, before damping in F begins to take effect [15]. If we consider the solution for $t \gg 1/FM$, we have the lower-left entry of Table 10.2, $\sigma_{zR}(t) = \sigma_M/t$. However, if we must include constant field uncertainty in our estimation, then our estimate becomes the lower-right entry $\sigma_{zR}(t) = 4\sigma_M/t$, which is, again, a factor of four worse.

If our task is field estimation, intrinsic quantum mechanical uncertainty in z limits our performance just as, if our task is spin-squeezed state preparation, field uncertainty limits our performance.

10.5.2.2 Transfer Function Estimation and Control

Suppose that the controller did not have the capability to adjust the gains in time as it tracked a fluctuating field. One approach would then be to apply the steady-state transfer functions derived above for the *entire* measurement. While this approach performs optimally in steady-state, it approaches the steady-state in a nonoptimal manner compared

to the dynamic controller. figure 10.4 demonstrates this poor transient performance for tracking fluctuating fields of differing bandwidth. Notice that the performance only begins to improve around the time that the dynamic controller saturates.

Also notice that the transfer function $G_b(s)$ is dependent on whether or not the state is being controlled, i.e., whether or not λ is zero. The performance shown in figure 10.4 is for one particular value of λ, but others will give different estimation performances for short times. Still, all of the transfer functions generated from any value of λ will limit to the same performance at long times. Also, all of them will perform poorly compared to the dynamic approach during the transient time.

10.6   Robust Performance: F Unknown

Until this point, we have assumed the observer has complete knowledge of the system parameters, in particular the spin number F. We will now relax this assumption and consider the possibility that, for each measurement, the collective spin F is drawn randomly from a particular distribution. Although we will be ignorant of a given F, we may still possess knowledge about the distribution from which it is derived. For example, we may be certain that F never assumes a value below a minimal value F_{min} or above a maximal value F_{max}. This is a realistic experimental situation, as it is unusual to have particularly long tails on, for example, trapped atom number distributions. We do not explicitly consider the problem of F fluctuating during an individual measurement, although the subsequent analysis can clearly be extended to this problem.

Given a F distribution, one might imagine completely reoptimizing the estimator-controller with the full distribution information in mind. Our initial approach is more basic and in line with robust control theory: we design our filter as before, assuming a particular F', then analyze how well this filter performs on an ensemble with $F \neq F'$. With this information in mind, we can decide if estimator-controllers built with F' are robust, with and without control, given the bounds on F. We will find that, under certain conditions, using control makes our estimates robust to uncertainty about the total spin number.

The essential reason for this robustness is that when a control field is applied to zero the measured signal, that control field must be approximately equal to the field to be tracked. Because F is basically an effective gain, variations in F will affect the performance, but not
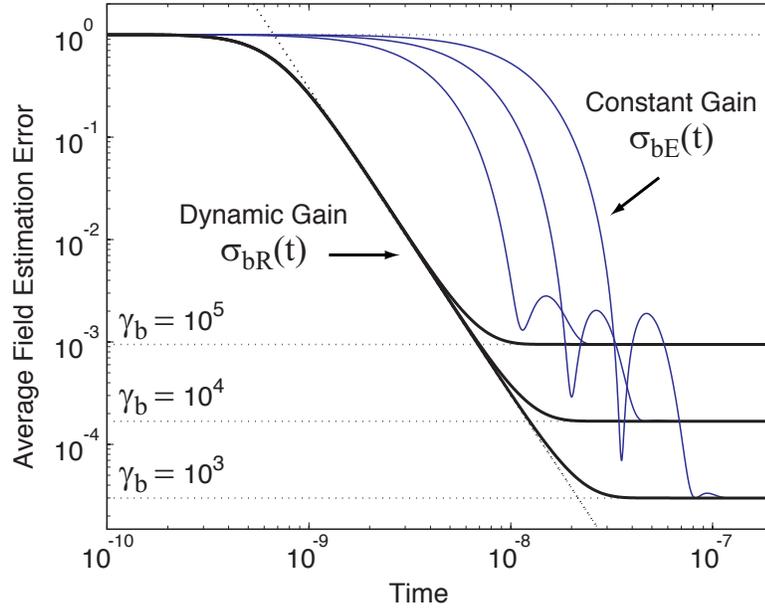
Figure 10.4: Estimation performance for estimators based on the dynamic gain solution of the Riccati equation, compared against estimators with constant estimation gain. The latter are the transfer function limits of the former, hence they have the same long-term performance. Three different bandwidth b processes are considered.

critically, so the error signal will still be approximately zero. If the applied signal is set to be the estimate, then the tracking error must also be approximately zero. (See section 10.8 for a robustness analysis along these lines in frequency space.)

Of course, this analysis assumes that we can apply fields with the same precision that we measure them. While the precision with which we can apply a field is experimentally limited, we here consider the ideal case of infinite precision. In this admittedly idealized problem, our estimation is limited by only the measurement noise and our knowledge of F.

First, to motivate this problem, we describe how poorly our estimator performs given ignorance about F without control.

10.6.1 Uncontrolled Ignorance

Let us consider the performance of our estimation procedure at estimating constant fields when $F' \neq F$. In general, this involves solving the complicated total covariance matrix equation (10.20). However, in the long-time limit ($t \gg 1/FM$) of estimating constant fields, the procedure amounts to simply fitting a line to the noisy measurement with a

least-squares estimate. Suppose we record an open-loop measurement that appears as a noisy sloped line for small angles of rotation due to the Larmor precession. Regardless of whether or not we know F, we can measure the slope of that line and estimate it to be $\tilde{m}$. If we knew F, we would know how to extract the field from the slope correctly: $\tilde{b} = \tilde{m}/\gamma F$. If we assumed the wrong spin number, $F' \neq F$, we would get the nonoptimal estimate: $\tilde{b}' = \tilde{m}/\gamma F' = \tilde{b}F/F'$.

First assume that this is a *systematic* error and F is unknown, but the same, on every trial. We assume that the constant field is drawn randomly from the σ_{b0} distribution for every trial. In this case, if we are wrong, then we are always wrong by the same factor. It can be shown that the error always saturates

$$\sigma_{bE} \rightarrow (1 - f)^2 \sigma_{b0}$$

where $f = F/F'$. Of course, because this error is systematic, the variance of the estimate does not saturate, only the error. This problem is analogous to ignorance of the constant electronic gains in the measurement and can also be calibrated away.

However, a significant problem arises when, on every trial, a constant b is drawn at random *and* F is drawn at random from a distribution, so the error is no longer systematic. In this case, we would not know whether to attribute the size of the measured slope to the size of F or to the size of b. Given the same b every trial, all possible measurement curves fan out over some angle due to the variation in F. After measuring the slope of an individual line to beyond this fan-out precision, it makes no sense to continue measuring.

We should also point out procedures for estimating fields in open-loop configuration, but *without* the small angle approximation. For constant large fields, we could observe many cycles before the spin damped significantly. By fitting the amplitude and frequency independently, or computing the Fourier transform, we could estimate the field somewhat independently of F, which only determines the amplitude. However, the point here is that b might not be large enough to give many cycles before the damping time or any other desired stopping time. In this case, we could not independently fit the amplitude and frequency because they appear as a product in the initial slope. Similar considerations apply for the case of fluctuating b and fluctuating F. See [135], for a complete analysis of Bayesian spectrum analysis with free induction decay examples.
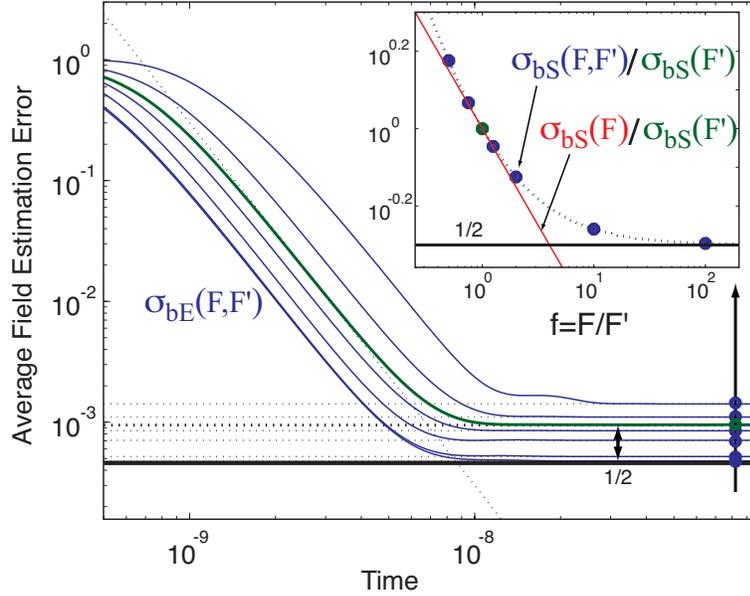
Figure 10.5: Steady-state estimation performance for estimator designed with $F' = 10^6$, and actual spin numbers: $F = F' \times [0.5, 0.75, 1, 1.25, 2, 10, 100]$. Other parameters: $\gamma = 10^6$, $M = 10^4$, $\gamma_b = 10^5$, $\sigma_{bFree} = 1$ (fluctuating field), $\lambda = 0.1$ (this is large enough to satisfy large-λ limits discussed in text). The inset compares the normalized robust estimation performance (curve) at a particular time, to the ideal performance (line) when F is known.

Fortunately, using precise control can make the estimation process relatively robust to such spin number fluctuations.

10.6.2 Controlled Ignorance: Steady-State Performance

We first analyze how the estimator designed with F' performs on a plant with F at tracking fluctuating fields with and without control. To determine this we calculate the steady-state of equation (10.19).

For the case of no control ($\lambda = 0$), we simplify the resulting expression by taking the same large F' approximation as before. This gives the steady-state uncontrolled error

$$\sigma_{bE} \quad \to \quad (1-f)^2 \frac{\sigma_{bF}}{2\gamma_b}$$
$$= (1-f)^2 \sigma_{bFree}$$

where $f = F/F'$. Because the variance of the fluctuating b is σ_{bFree}, the uncontrolled estimation performs worse than no estimation at all if $f > 2$.
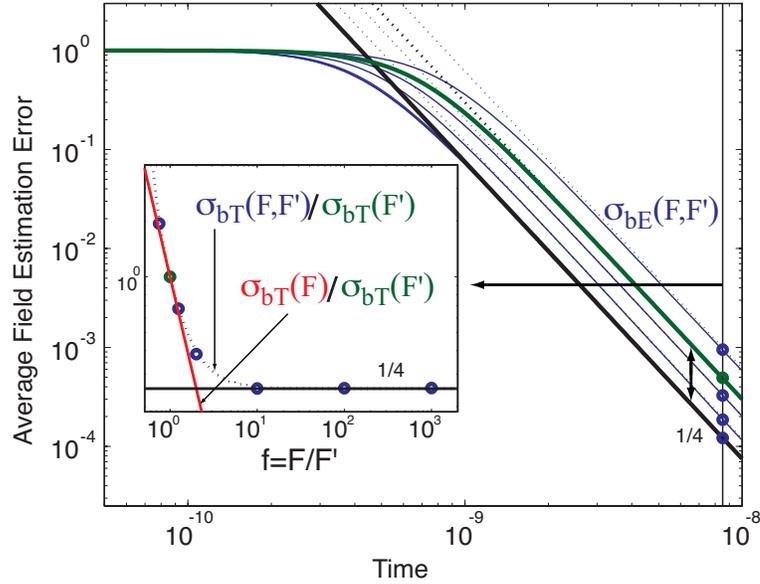
Figure 10.6: Transient estimation performance for controller designed with $F' = 10^6$, and actual spin numbers: $F = F' \times [0.75, 1, 1.25, 2, 10, 100, 1000]$. Other parameters: $\gamma = 10^6$, $M = 10^4$, $\gamma_b = 0$, $\sigma_{bFree} = 0$ (constant field), $\lambda = 1$. Note that this behavior is valid for $t < 1/M = 10^{-4}$. The inset compares the normalized robust estimation performance (curve) at a particular time, to the ideal performance (line) when F is known.

On the other hand, when we use precise control the performance improves dramatically. We again simplify the steady-state solution with the large F' and λ assumptions from before, giving

$$
\begin{aligned}
\sigma_{bS}(F, F') &\rightarrow \left(\frac{1+f}{2f}\right) \sqrt{\frac{2}{\gamma F'}} \sigma_{bF}^{3/4} \sigma_M^{1/4} \\
&= \left(\frac{1+f}{2f}\right) \sigma_{bS}(F')
\end{aligned}
$$

where $\sigma_{bS}(F, F')$ is the steady-state controlled error when a plant with F is controlled with a F' controller and $\sigma_{bS}(F')$ is the error when $F = F'$. One simple interpretation of this result is that if we set F' to be the minimum of the F distribution ($f > 1$) then we never do worse than $\sigma_{bS}(F')$ and we never do better than twice as well ($f \rightarrow \infty$). See figure 10.5 for a demonstration of this performance.

10.6.3 Controlled Ignorance: Transient Performance

Now consider measuring constant fields with the wrong assumed F'. Again, when control is not used, the error saturates at

$$\sigma_{bE} \to (1-f)^2 \sigma_{b0}.$$

When control is used, the transient performance again improves under certain conditions. The long-time transient solution of equation (10.19) is difficult to manage analytically, yet the behavior under certain limits is again simple. For large λ and F', and for $f > 1/2$, we numerically find the transient performance to be approximately

$$
\begin{aligned}
\sigma_{bT}(F, F') \quad &\to \quad \left(\frac{f^2 + 2}{4f^2 - 1}\right) \frac{12\sigma_M}{\gamma^2 F'^2 t^3} \\
&= \quad \left(\frac{f^2 + 2}{4f^2 - 1}\right) \sigma_{bT}(F')
\end{aligned}
\tag{10.27}
$$

where $\sigma_{bT}(F, F')$ is the transient controlled error when a plant with F is controlled with a F' controller and $\sigma_{bT}(F')$ is the error when $F = F'$. See figure 10.6 for a demonstration of this performance for realistic parameters. As $f \to \infty$ the f-dependent prefactor saturates at a value of $1/4$. However, as $f \to 1/2$ then the system takes longer to reach such a simple asymptotic form, and the solution of equation (10.27) becomes invalid.

Accordingly, one robust strategy would be the following. Suppose that the lower bound of the F-distribution was known and equal to F_{min}. Also assume that $\sigma_{bT}(F_{min})$ represents an *acceptable* level of performance. In this case, we could simply design our estimator based on $F' = F_{min}$ and we would be guaranteed at least the performance $\sigma_{bT}(F_{min})$ and at best the performance $\sigma_{bT}(F_{min})/4$.

This approach would be suitable for experimental situations because typical F distributions are narrow: the difference between F_{min} and F_{max} is rarely greater than an order of magnitude. Thus, the overall sacrifice in performance between the ideal case and the robust case would be small. The estimation performance still suffers because of our ignorance of F, but not nearly as much as in the uncontrolled case.

10.7 Riccati Equation Solution Method

The matrix Riccati equation is ubiquitous in optimal control. Here, following [136], we show how to reduce the nonlinear problem to a set of linear differential equations. Consider the generic Riccati Equation:

$$\frac{d\mathbf{V}(t)}{dt} = \mathbf{C} - \mathbf{D}\mathbf{V}(t) - \mathbf{V}(t)\mathbf{A} - \mathbf{V}(t)\mathbf{B}\mathbf{V}(t).$$

We propose the decomposition:

$$\mathbf{V}(t) = \mathbf{W}(t)\mathbf{U}^{-1}(t)$$

with the linear dynamics

$$\begin{bmatrix} \frac{d\mathbf{W}(t)}{dt} \\ \frac{d\mathbf{U}(t)}{dt} \end{bmatrix} = \begin{bmatrix} -\mathbf{D} & \mathbf{C} \\ \mathbf{B} & \mathbf{A} \end{bmatrix} \begin{bmatrix} \mathbf{W}(t) \\ \mathbf{U}(t) \end{bmatrix}.$$

It is straightforward to then show that this linearized solution is equivalent to the Riccati equation

$$
\begin{aligned}
\frac{d\mathbf{V}(t)}{dt} &= \frac{d\mathbf{W}(t)}{dt}\mathbf{U}^{-1} + \mathbf{W}(t)\frac{d\mathbf{U}^{-1}(t)}{dt} \\
&= \frac{d\mathbf{W}(t)}{dt}\mathbf{U}^{-1}(t) + \mathbf{W}(t)(-\mathbf{U}^{-1}(t)\frac{d\mathbf{U}(t)}{dt}\mathbf{U}^{-1}(t)) \\
&= (-\mathbf{D}\mathbf{W}(t) + \mathbf{C}\mathbf{U}(t))\mathbf{U}^{-1}(t) \\
&\quad -\mathbf{W}(t)\mathbf{U}^{-1}(t)(\mathbf{B}\mathbf{W}(t) + \mathbf{A}\mathbf{U}(t))\mathbf{U}^{-1}(t) \\
&= \mathbf{C} - \mathbf{D}\mathbf{V}(t) - \mathbf{V}(t)\mathbf{A} - \mathbf{V}(t)\mathbf{B}\mathbf{V}(t)
\end{aligned}
$$

where we have used the identity

$$\frac{d\mathbf{U}^{-1}(t)}{dt} = -\mathbf{U}^{-1}(t)\frac{d\mathbf{U}(t)}{dt}\mathbf{U}^{-1}(t).$$

Thus the proposed solution works and the problem can be solved with a linear set of differential equations.
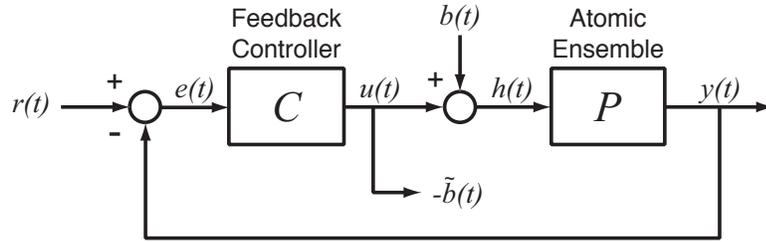
Figure 10.7: Spin control system with plant transfer function $P(s) = \gamma F/s$. The reference signal, $r(t)$, is usually zero. The error signal is $e(t)$. The control output is $u(t)$. The external field to be tracked is $b(t)$. The total field is $h(t) = b(t) + u(t)$ and $\tilde{b}(t)$ is the field estimate.

10.8   Robust Control in Frequency Space

Here we apply traditional frequency-space robust control methods [50, 51] to the classical version of our system. This analysis is different from the treatment in the body of the chapter in several respects. First, we assume nothing about the noise sources (bandwidth, strength, etc.). Also, this approach is meant for steady-state situations, with the resulting estimator-controller being a constant gain transfer function. The performance criterion we present here is only loosely related to the more complete estimation description above. Despite these differences, this analysis gives a very similar design procedure for the steady-state situation.

We proceed as follows with the control system shown in figure 10.7, where we label $h(t) = u(t) + b(t)$ as the total field. Consider the usual spin system but ignore noise sources and assume we can measure $z(t)$ directly, so that $z(t) = y(t)$. For small angles of rotation, the transfer function from $h(t)$ to $y(t)$ is an integrator

$$
\begin{aligned}
\frac{dy(t)}{dt} &= \frac{dz(t)}{dt} = \gamma F h(t) \\
s y(s) &= \gamma F h(s) \\
y(s) &= P(s) h(s) \\
P(s) &= \gamma F/s.
\end{aligned}
$$

Now we define the performance criterion. First notice that the transfer function from

the field to be measured $b(t)$ to the total field $h(t)$ is $S(s)$ where

$$
\begin{aligned}
h(s) &= S(s)b(s) \\
S(s) &= \frac{1}{1 + P(s)C(s)}.
\end{aligned}
$$

(Also notice that this represents the transfer function from the reference to the error signal $e(s) = S(s)r(s)$.) Because our field estimate will be $\tilde{b}(t) = -u(t)$, we desire $h(t)$ to be significantly suppressed. Thus we would like $S(s)$ to be small in magnitude (controller gain $|C(s)|$ large) in the frequency range of interest. However, because the gain $|C(s)|$ must physically decrease to zero at high frequencies we must close the feedback loop with adequate phase margin to keep the closed-loop system stable. This is what makes the design of $C(s)$ nontrivial.

Proceeding, we now define a function $W_1(s)$ that represents the degree of suppression we desire at the frequency $s = j\omega$. So our controller $C(s)$ should satisfy the following performance criterion

$$
\|W_1(s)S(s)\|_\infty < 1.
$$

Thus the larger $W_1(s)$ becomes, the more precision we desire at the frequency s. We choose the following performance function

$$
W_1(s) = \frac{W_{10}}{1 + s/\omega_1}
$$

such that ω_1 is the frequency below which we desire suppression $1/W_{10}$.

Because our knowledge of F is imperfect, we need to consider all plant transfer functions in the range

$$
P = \frac{\gamma}{s}\{F_{min} \to F_{max}\}.
$$

Our goal is now to find a $C(s)$ that can satisfy the performance condition for any plant in this family. We choose our nominal controller as

$$
C_0(s) = \frac{\omega_C}{\gamma F'}.
$$

So if $F = F'$ then the system closes at ω_C (i.e., $|P(i\omega_C)C_0(i\omega_C)| = 1$), whereas in general the system will close at $\omega_{CR} = \omega_C \frac{F}{F'}$. We choose this controller because $P(s)C(s)$ should be an integrator ($\propto 1/s$) near the closing frequency for optimal phase margin and closed-loop stability.

Next we insert this solution into the performance condition. We make the simplifying assumption $\omega_1 \ll \omega_C \frac{F}{F'}$ (we will check this later to be self-consistent). Then the optimum of the function is obvious and the condition of equation (10.28) becomes

$$\omega_1 W_{10} < \omega_{CR} = \omega_C \frac{F}{F'}.$$

We want this condition to be satisfied for all possible spin numbers, so we must have

$$\omega_1 W_{10} = \min[\omega_{CR}] = \omega_C \frac{F_{min}}{F'}. \tag{10.28}$$

Experimentally, we are forced to roll off the controller at some high frequency that we shall call ω_Q. Electronics can only be so fast. Of course, we never want to close above this frequency because the phase margin would become too small, so this determines the maximum F that the controller can reliably handle

$$\omega_Q = \max[\omega_{CR}] = \omega_C \frac{F_{max}}{F'}. \tag{10.29}$$

Combining equations (10.28, 10.29 we find our fundamental trade-off

$$\omega_1 W_{10} = \omega_Q \frac{F_{min}}{F_{max}}, \tag{10.30}$$

which is the basic result of this section. Given experimental constraints (such as F_{min}, F_{max}, and ω_Q), it tells us what performance to expect ($1/W_{10}$ suppression) below a chosen frequency ω_1.

From equation (10.30), we recognize that the controller gain at the closing frequency needs to be

$$|C|_C = \frac{\omega_C}{\gamma F'} = \frac{\omega_1 W_{10}}{\gamma F_{min}} = \frac{\omega_Q}{\gamma F_{max}}.$$

In the final analysis, we do not need to use F' and ω_C to parametrize the controller, only

the trade-off and the gain. Also, notice that now we can express $\min[\omega_{CR}] = \omega_1 W_{10}$.

To check our previous assumption

$$
\begin{aligned}
\omega_1 &\ll \omega_C \frac{F}{F'} \\
&= \omega_1 W_{10} \frac{F}{F_{min}},
\end{aligned}
$$

which is true if $W_{10} \gg 1$.

Finally, the system will never close below the frequency $\min[\omega_{CR}]$ so we should increase the gain below a frequency ω_H, which we might as well set equal to $\min[\omega_{CR}]$. This improves the performance above and beyond the criterion above. Of course we will be forced to level off the gain at some even lower frequency ω_L because infinite DC gain (a real integrator) is unreasonable. So the final controller can be expressed as

$$
C(s) = |C|_C \frac{1}{1 + s/\omega_Q} \frac{\omega_H(1 + s/\omega_H)}{\omega_L(1 + s/\omega_L)}
$$

with the frequencies obeying the order

$$
\begin{aligned}
\omega_L &< \\
\omega_H &= \min[\omega_{CR}] = \omega_1 W_{10} < \\
\omega_{CR} &= \frac{F}{F_{min}} \omega_1 W_{10} < \\
\omega_Q &= \max[\omega_{CR}] = \frac{F_{max}}{F_{min}} \omega_1 W_{10}.
\end{aligned}
$$

Notice that the controller now looks like the steady-state transfer function in figure 10.3 derived from the steady-state of the full dynamic filter. (The notation is the same to make this correspondence clear). Here ω_Q was simply stated, whereas there it was a function of λ that went to infinity as $\lambda \to \infty$. Here the high gain due to ω_L and ω_H was added manually, whereas before it came from the design procedure directly.

10.9 Conclusion

The analysis of this chapter contained several key steps that should be emphasized. Our first goal was to outline the proper approach to quantum parameter estimation. The second

was to demonstrate that reduced representations of the full filtering problem are relevant and convenient because, if a simple representation can be found, then existing classical estimation and control methods can be readily applied. The characteristic that led to this simple description was the approximately Gaussian nature of the problem. Next, we attempted to present basic classical filtering and control methodology in a self-contained, pedagogical format. The results emphasized the inherent trade-offs in simultaneous estimation of distinct, but dynamically coupled, system parameters. Because these methods are potentially critical in any field involving optimal estimation, we consider the full exposition of this elementary example to be a useful resource for future analogous work.

We have also demonstrated the general principle that precision feedback control can make estimation robust to the uncertainty of system parameters. Despite the need to assume that the controller produced a precise cancellation field, this approach deserves further investigation because of its inherent ability to precisely track broadband field signals. It is anticipated that these techniques will become more pervasive in the experimental community as quantum systems are refined to levels approaching their fundamental limits of performance.

Chapter 11

Experimental Apparatus

All previous chapters have motivated the design of an experiment where the collective spin of a cloud of laser-cooled cesium atoms is measured via the polarization rotation of an off-resonant probe beam. In this chapter, I describe the components of the apparatus built for this purpose, which was used to acquire the experimental data discussed in subsequent chapters.

figure 11.1 displays the general schematic of the experiment, as has been described previously in less detail. When walking through the door of the lab, one sees figure 11.2: an optical table with diode lasers in the foreground, a magnetically shielded vacuum chamber in the background, and a mess of optics everywhere else. When the magnetic shield is removed from around the vacuum chamber, one can see the quartz cell where the atoms are trapped and cooled, figure 11.3.

11.1 Vacuum Chamber

Compared to most atomic physics labs, our UHV vacuum chamber system is relatively simple and essentially consists of only a quartz cell, a cesium source, and an ion pump. Here we describe the parts of the system and their specifications. For more information on vacuum system procedures (cleaning, baking, assembly, etc.), refer to Kevin Birnbaum's thesis [14].

The vacuum system components are connected via Varian tubing and UHV compatible ConFlat flanges (CFF). We use a 40 L/s ion pump (Varian StarCell VacIon Pump 40, MiniVac Controller) to evacuate the system in steady-state and routinely achieve 10^{-10} Torr pressures, although we typically operate at a cesium pressure 10-100 times greater.
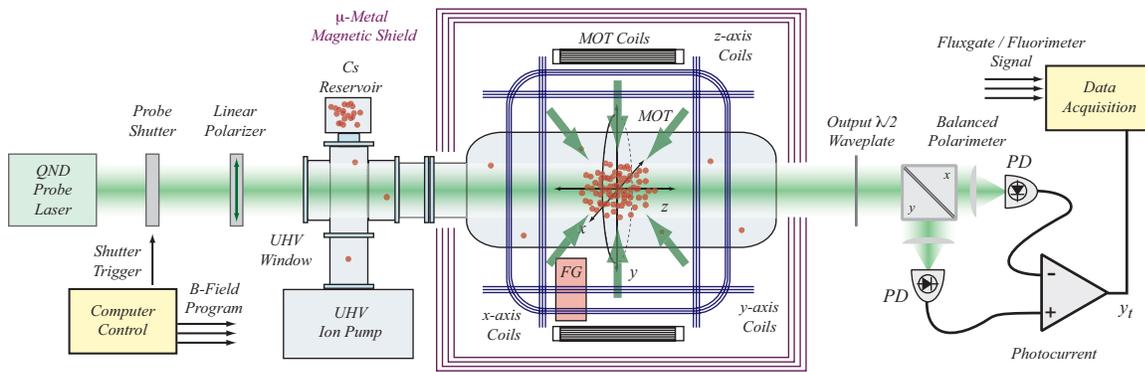
Figure 11.1: Schematic of our experimental apparatus in which collective spin angular momentum of a cloud of laser-cooled cesium atoms is measured by polarimetric detection of a scattered off-resonant probe laser. Ambient magnetic field fluctuations are suppressed by magnetic shielding and can be monitored with a fluxgate magnetometer (FG) situated nearby the atomic sample. Components not shown include the optical pumping laser (aligned along the laboratory x-axis) and external trim coils used to zero ambient magnetic fields and their first-order gradients.
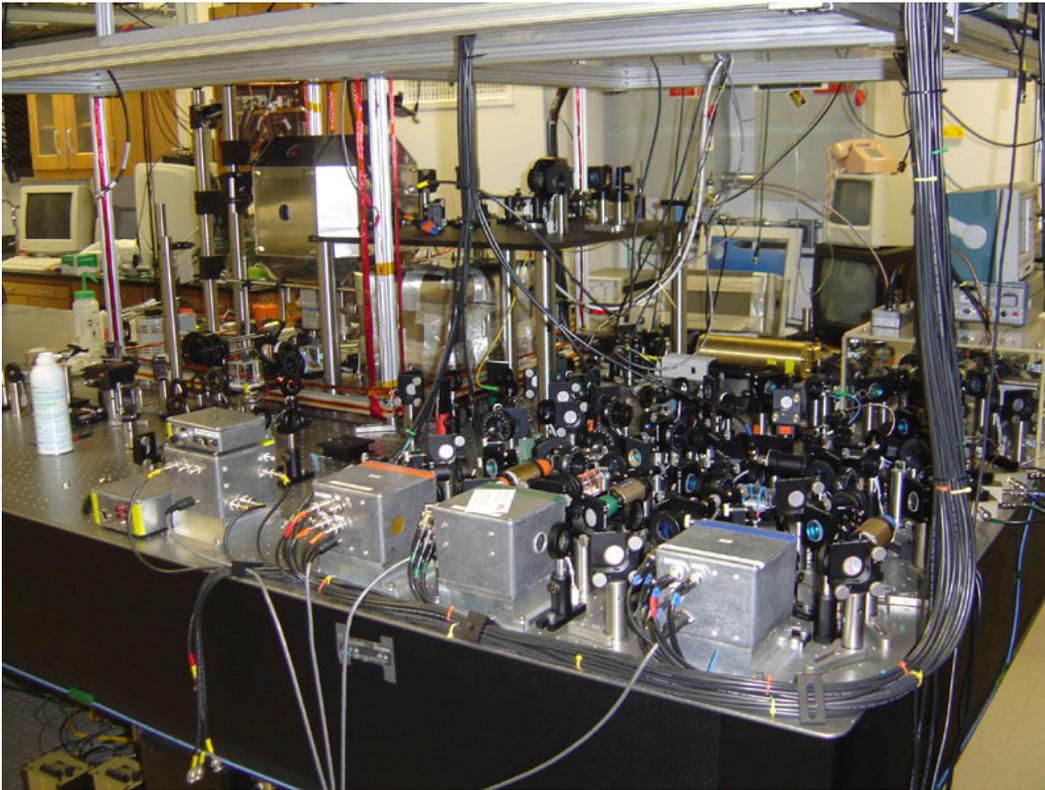
Figure 11.2: The view from inside the lab.
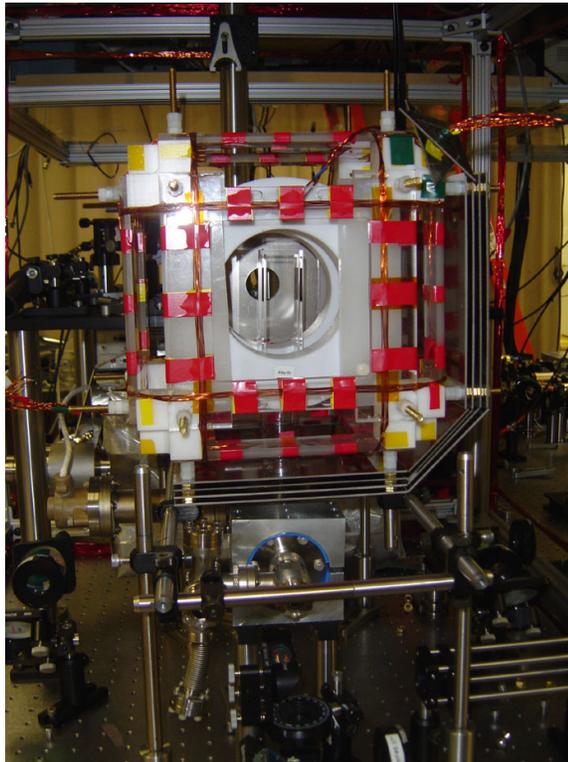
Figure 11.3: A view of the vacuum cell and magnetic cube with the magnetic shielding partially disassembled.

The ion pump was ordered with an extra side port (CFF 2.75"), to which a right angle valve (VAT 57 series, 57132-GE02) is attached. The system is rough pumped during baking through this port with our mobile turbo pump (Varian Minuteman Turbo V250).

The ion pump is a significant source of magnetic fields and, more importantly, magnetic field gradients (see appendix A). To reduce the effect of these fields at the atom site within the cell, we have surrounded the pump with a two-layer magnetic shield (Co-netic AA from Magnetic Shield Corp.), with thinner Co-netic foil corners. This shielding reduces the outer fields by more than an order of magnitude.

The pressure is measured with either a pressure gauge (Varian UHV 24p Nude Ionization Gauge, SenTorr Controller), the ion pump current (calibrated to pressure via the manual), or via the loading time of the MOT (see section 12.5.3). The gauge and ion pump current readings are somewhat unreliable and may read a higher value than the actual pressure if too much cesium accumulates internally, forming filament-like emitters that artificially increase the current. These deposits can be removed by operating the pump at a higher than normal potential (hi-potting) with, for example, a neon sign transformer.

For most experiments, our cesium source has consisted of a broken cesium ampoule (from Aldrich) within a bellows tube connected to the cell via a right-angle Varian valve (mini-CFF, 0.75"). In the past, we have attempted to use cesium dispenser (getter) wires from SAES (via Litton Industries). However, with these cesium sources the MOT loading is much more efficient if a direct line of sight can be established between the MOT and the dispenser. Because this was not the case in our experiment and our pressure constraints were not severe, we resorted to the standard ampoule source.

We use a cell rather than a metal chamber to avoid having magnetic or eddy current sources close to the atomic sample. Initially, the experiment was set up with a pyrex optical cell from Technical Glass (5" x 2" x 2", 2.75" CF flange). Eventually, this was replaced with a higher quality custom quartz cell from Starna Cells (Spectrosil Far UV Quartz, 5"x1.5"x1.5", 2.75" CF flange). The stainless steel CFF to quartz transition (Steel 304, to Pyrex, to Quartz GE214) was purchased from Larson Electronic Glass and then sent to Optiglass in the UK, which proceeded to assemble the transition to the body of the cell.

The cell is mounted in the vertical direction from a four-way cross. To one side of the cross is the valved cesium source, to the other the ion pump, and below the cell is an antireflection (AR) coated window from (MDC, 2.75" CFF) to allow for vertical probing

of the atom cloud. The cell itself is uncoated, meaning that both trapping beams and probing beams will suffer reflection from the air-quartz interface going in and out of the cell. To remedy this, one option is to place AR coated coverslips against the external faces of the cell with index-matching fluid in between to reduce internal reflections. Although the inner vacuum-quartz interface still leads to some reflection of the optical beams, the coverslips improve the loss from the outer quartz-air transition. We ordered a large supply of coverslips, with one side having a broadband antireflection (BBAR) coating, from ZC&R Coatings for Optics for this purpose. In practice, we have only used the slips for the top face of the cell (where the probe beam exits) and not the side faces for the trapping beams. The coating is specified at less than 0.5% reflectance between 800 and 1200 nm and less than 1% between 0 and 50 degrees incidence angle. The index matching fluid was purchased from Cargille Laboratories Immersion Liquid 06350, (n=1.45256 at 850 nm and n=1.44970 at 1064 nm).

During baking, care must be taken to ensure minimal temperature gradients from the steel section of the chamber to the cell. A special cage (with layering of aluminum foil to heater tape to more foil) is used to surround the cell at a distance of a few inches during the bake. The most temperature sensitive regions of the cell are the coated window (max 200-250°C), the steel to quartz transition (consisting of Pyrex/Borosilicate, 300°C, hard max 450°C), and the optical table (max 80°C). Other components of the system can typically withstand up to 400°C temperatures. During the bake, we reach maximum temperatures of around 220°C (via Variacs connected to heater tape), which is above the supposed temperature needed to remove thin films of water from the steel [14]. A slight temperature gradient (tens of degrees) is kept across the system with hotter temperatures away from the pump.

11.2   Field Control

In appendix A, I discuss the details of general "magnetic field management" in the lab, including a discussion of field constraints in our experiment, typical noise sources, shielding, coils, and field drivers. Here I only describe the different coils that exist in the lab.

To zero the DC field (primarily from the earth) and field gradients, we use a large (1 m) set of Helmholtz coils wrapped around a frame surrounding the experiment and supply

each side of each coil with an independent current. By driving a set of coils asymmetrically, (diagonal) field gradients can be nulled. On each axis there exists one many turn set of Helmholtz coils for coarse zeroing and a single turn coil for fine zeroing.

To apply fast fields of interest in the experiment we constructed a "magnetic cube" (pictured in figure 11.3) and placed it about the vacuum cell. The design was partially borrowed from the group of Poul Jessen and involves a 10" cubic plastic frame, with significant optical access. Wound around the cube are three Helmholtz coil pairs (of optimal spacing for uniformity [137]), one for each spatial direction. Within the cube, and supported by it, are the circular anti-Helmholtz MOT coils mounted on a teflon support. The cube itself structure consists of eight 2" Nylon small tapped corner cubes and six 8" square acrylic frames, which form the faces that connect the corners.

A four layer octagonal magnetic shield with optical access (for MOT, pump, probe beams, and fluorescence detection) was then placed on top of the cube and supported by brass rods emanating from the nylon corner cubes. The shield consists of Co-netic AA precut sheets from the Magnetic Shield Corp. This shield provides a reduction factor of approximately 30 at low frequencies.

11.3 Lasers and Optics

The experiment uses four homemade diode lasers, three for trapping and one for the probe light. The trapping lasers include a MOT master, a MOT slave, and a repump laser. The master output is locked to the 4–5' crossover transition and is used to injection-lock the MOT slave, which then provides trapping light. The repump provides 3–4' light to return atoms from the otherwise dark $f = 3$ state to the $f = 4$ cooling transitions. The probe slave is seeded by a New Focus Vortex laser as discussed below.

All of the homemade diode lasers (including the diode housing, temperature controller and current controller) are built from the standard Kimble and Mabuchi lab designs, which were developed in large part by Joe Buck [138] and based on earlier work from the atomic physics community [139, 140]. The frequency of the MOT master and the repump laser are locked using via a saturation absorption spectroscopy error signal [141]. Because our experiment was the second atomic physics apparatus in the Mabuchi lab, the diode laser setup is very similar to Ben Lev's experiment, which was the first, and more information

can be found in his thesis [142].

After the locking setup, the repump light is simply expanded and sent to the MOT described below. The output of the master is double passed through an acousto-optic modulator (Intra-Action, AOM-402AF3) whose frequency can be used to set the detuning of the eventual probe. This beam is then seeded into the slave laser, which provides close to 120 mW of total power. This light is then passed through a switching AOM (Intra-Action, AOM-402AF3) with a constant frequency shift after which about 90 mW of total power is available. The amplitude of the trapping light is then controlled via a voltage controlled attenuator (VCA), which attenuates the driving signal of this final AOM.

The extra 10 mW of master light that is not used to seed the slave is used for several purposes. Both the pumping beam and the depumping beam are derived from the master light and AOM shifted (Intra-Action, ATM-1251A2) to approximately the 4–4' transition. The pumping light is then fiber coupled to the table at the height of the cell, expanded, and sent through the cell horizontally. Just before passing through the cell the pump beam is circularly polarized to pump the spins along the axis defined by its propagation direction. Some of the master light is also used as a reference to measure the frequency of the Vortex laser probe beam via heterodyne mixing.

For turning off certain beams, we have created shutters by modifying acoustical speakers to raise and lower a blocking flag based on a design from [143]. Alternatively, we have used Oriel shutters (Oriel 76992), which give much more vibration to the table, but are generally more reliable.

11.4 Magneto-Optic Trap (MOT)

Assuming enough power is available to keep the MOT beam intensity at near saturation levels, making the beams larger is the easiest way to trap more atoms. As discussed in section 12.5.2, the MOT loading rate goes as the fourth power of the beam width. After the trapping beam has passed through the switching AOM it is expanded to "fill out" the last lens of the telescope to provide a 1" collimated beam. This beam is then either sent straight through 1" optics, reflected at 45 degrees with 2" New Focus Mirrors, or divided with 2" Polarizing Beam Splitters (Melles-Griot, 03PBB017). After the initial expansion the beam is split with the large PBS and VLOC waveplates into three trapping beams, of

full width 1" and power 30 mW. Each individual beam is sent through a quarter-waveplate, through the cell, through another quarter-waveplate, then retroreflected off of a 1" mirror to form the optical molasses (or MOT with anti-Helmholtz coil). Other beams such as the repump and depump are combined with the trapping beams on the large PBS.

11.5   Fluorimeter and Camera

As one means of measuring the atomic number we use a typical "fluorimeter" to detect the fluorescent light emanating from the MOT and back out the expected number with knowledge of the MOT beam intensity and detuning (see section 12.5.6). The fluorimeter consists of a pair of lenses, outside the magnetic shield, to image the MOT onto a large detector (ThorLabs, PDA55, 3.6 mm). We observe the MOT in steady-state and the thermal expansion of the cloud after the atoms have been released with a high performance CCD camera (Cohu 4912). When doing temperature measurements, the atoms are loaded, released, then after some delay time fluorescent beams (trapping beams) are turned on and an images are acquired via a National Instruments image acquisition board described below. Then the fit radius of the cloud as a function of time is used to extract the temperature (as in section 12.5.4).

11.6   Probe Laser and Polarimeter

The probe laser beam is derived from a New Focus Vortex External-Cavity Diode Laser (Model 6017, with Vortex 6000 Laser Controller), which offers 17 mW at 852 nm with a short term (50 ms) linewidth of less than 300 kHz. After being reshaped and passed through an optical isolator, some of the beam is picked off and mixed with light from the locked "master" diode laser to give a heterodyne signal (using a New Focus 1601, 1 GHz, detector), which is then sent to a frequency counter (Tektronix CMC251) to give the relative frequency. The remaining light is then used to optically seed another homemade diode laser to give an output of approximately 80 mW. This more powerful output is then sent through an isolator and split with a PBS. A small amount of light is then heterodyne mixed with some of the seed light to check that the slave is indeed seeded (using a New Focus 1801 Detector). The rest of the probe light is then sent to an AOM (Intra-Action,

AOM40R) at 40 MHz, which serves as the switch for turning the probe measurement on and off. The output of the AOM is then expanded and sent through a high quality AR coated polarizer (10^{-6} Karl Lambrecht Corp., MGT25E5, Glan-Thompson Prism Polarizer) with 10^{-6} extinction ratio. Because the light must pass off of several mirrors before entering the vacuum chamber, two waveplates (one quarter-wave and one half-wave, VLOC zero-order) are used to compensate the polarization rotation effects of the input mirrors. The waveplates are adjusted to linearize the optical polarization along the direction of the pumped atomic spin-state using tensor Hamiltonian induced oscillations as an "error signal" as discussed in section 6.7.

After the beam has passed through the vacuum chamber it is reflected off of three mirrors. Because these mirrors also induce a polarization rotation their effect needs to be compensated using a set of three subsequent waveplates. Unlike with the input light, these waveplates approximately compensate a *general rotation* (not specific to a single state) as described in section 3.2.5. The compensated light is then sent into the final polarimeter setup consisting of a pair of waveplates (one quarter-wave and one half-wave), a high quality AR coated polarizing beam splitter (10^{-6} extinction, Karl Lambrecht Corp., MSBTS-8-45, Beamsplitting Glan-Thompson Prism Polarizer), and a homemade subtracting photodetector. The arrangement of waveplates here typically consists of just a half-waveplate that zeros the output of the photodetector without atoms. To measure a general Stokes component though, both waveplates are necessary.

At times it is advantageous to strobe the measurement strength (probe power) sinusoidally. Because the analog output boards have limited bandwidth and resolution, it does not make sense to use these to modulate the power directly. Instead one can combine two signals of different frequency and use this to drive the AOM controlling the probe power. The probe power will then be modulated at the beatnote of the RF signals. For example, to modulate the power at 500 kHz, a 40.5 MHz signal and a 40 MHz signal are combined using a splitter (MiniCircuits ZSC-2-1) in reverse as an adder, and the sum is sent to an amplifier, which drives the AOM. By using two high quality synchronized signal generators (Agilent E4400B 1GHz ESG Signal Generator) as the source of the two high-frequency signals, which are in turn synchronized with the computer trigger, we can have adequate phase and power control to perform fast stroboscopic measurement.

The detector used in this experiment must satisfy several constraints. First, the detector

must have a large enough bandwidth to record the evolution of the spin dynamics. Second, the detector should be optical shotnoise limited for all frequencies from DC to the maximal bandwidth. The detector should also have a reasonably high quantum efficiency. As seen in the previous analysis of chapter 8, the signal-to-noise ratio, hence the degree of expected conditional spin-squeezing is critically dependent on the collection efficiency of the detector. Finally, the detector should be able to cleanly absorb the amount of power needed for the experiment. Typically, excess noise will be added to the detection if an intensity greater than some threshold is used. Thus, when larger powers are used (greater than a mW) it is necessary for the photodiodes to be larger to reduce the accepted intensity. Alternatively, one could use the trick discussed in section 8.1.4, but we have not resorted to that.

There are two sets of polarimeter detectors that have been used in the experiment. The first used a pair of smaller photodiodes (Hamamatsu S5973, 0.4 mm, $R = 0.47$ A/W, or QE= 68%) and was shotnoise limited for powers above 1 μW from DC to 1 MHz. This detector starts to go nonlinear above around 150 μW. The second detector, built with the help of Mike Armen, uses larger photodiodes (Hamamatsu S3071, 5 mm, $R = 0.56$ A/W, or QE $= 81$%) and is shotnoise limited above 100 μW out to approximately 10 MHz. Because of the larger diode size, this detector goes nonlinear at a much higher total power of 25 mW. Both detectors use a standard single transimpedance amp design [144, 145] in subtracting configuration with an Burr Brown OPA655 amplifier (now replaced by OPA656). More information on general photodetector design can also be found in [13]. The photodetector board and many of our controller boards were constructed using a T-Tech Quick Circuit 5000 circuit mill and Protel software.

Note that there are several techniques used to enhance the signal-to-noise of polarimeters in the presence of technical noise using, for example, frequency modulation [146], polarization modulation with a faraday rotator and feedback [147], and selective polarization component removal [110, 148] as discussed in section 8.1.4.

11.7 Computer Control

The computer control system uses multiple input and output boards, all running off of one PC computer, to coordinate the entire experiment. Here is a summary of these boards and their use:

- *Fast Analog Input Board:* Gage CompuScope 14050. Use: acquiring polarimeter output. Specs: 2 Channel, 14 bit, 50 MSample/s, 8 MS onboard memory, one PCI slot.

- *Slow (but precise) Analog Input Board:* National Instruments PCI-4474. Use: reading fluorimeter and 3 fluxgate channels. Specs: 4 Channel, 24 bit, 102.4 kSample/s, 45 kHz alias-free bandwidth.

- *Analog output board:* National Instruments PCI-6713. Use: MOT intensity, detuning, magnetic fields. Specs: 8 Channels, 12 bit, 1 Msample/s, 8 digital I/O lines in addition.

- *Digital output (and input) board:* National Instruments PCI-DIO-32HS. Use: 32 I/O channels, up to 13.3 MB/s pattern I/O.

- *Image Acquisition Board:* National Instruments PCI-1407. Use: Imaging. Specs: Single channel, analog, monochrome.

Instead of using LabView or a similar data acquisition environment we decided to construct our own within Matlab by using "mex" commands to access general functions written in C that control the boards. JM Geremia put significant effort into making this system a general purpose environment, providing a concise set of simple but general functions that bridge the gap between Matlab and the hardware. I then had the much easier task of designing the experimental control from the user-friendly Matlab side. The code is generally set up such that the outputs are held at steady values for several seconds while the MOT loads from vapor, after which the atoms are released and a high time resolution series of commands dictates the subsequent experiment over the course of several milliseconds. To reduce powerline related noise, the computer system is put into a wait-mode until it receives a line-triggered pulse from the output of an function generator (Stanford Research Systems DS345).

Chapter 12

Preparing and Characterizing Cold, Optically Thick Atomic Clouds

12.1 Overview

Many experiments in atomic physics require a large number of cold atoms as a resource. The goals of individual experiments are usually different but often overlapping. The primary goal in the free-space spin-squeezing experiment, described in section 8.2.3, is to maximize the optical depth OD $= N\sigma_0/A$ in a trap with reasonably cold temperatures for reasonably long times. Maximizing this parameter is similar to, but not the same as, the usual atomic physics goal of maximizing phase space density. There were a surge of papers on atom-trapping and cooling in the early 1990s aimed at increasing the phase space density to levels where a Bose-Einstein condensate (BEC) could be achieved. In this chapter we reference them heavily, keeping in mind our modified goal.

Large OD cold atomic samples are difficult to prepare and also somewhat tricky to characterize. For example, multiple scattering of photons within the atomic cloud can lead to nontrivial collective effects. Furthermore, the high degree of absorption for probing and trapping beams also lead to complications. The goal of this chapter is to review what has been achieved in terms of large OD cold atomic samples and also lay out the basic theory for preparing and characterizing these systems. Because the full theory for even the simplest MOT is surprisingly complicated, we keep in mind that all models used are not exact and merely a rough guide for scalings and intuition.

We begin by briefly reviewing the constraints of this experiment compared to other

similar experiments. Next, we discuss the basic theory of MOTs and how this theory breaks down at larger atom numbers. We then review what has been achieved in other atom-trapping experiments, in terms of number, density, optical depth, and so on, with attention paid to the differences in each experiment's ultimate goal. The details of experimental techniques to measure and predict various parameters of interest (fluorescence number, background pressures, OD, atom temperature) are then discussed. Finally, we consider the effects of various forces on the atoms within the probe beam after the initial trapping stage, including the effects of temperature, gravity, and the mechanical forces of the pump and the probe beams.

12.2   Experimental Goals

The requirement for a BEC [149] is that the phase space density

$$\rho \;=\; n\Lambda^3 \tag{12.1}$$

$$\;=\; n\hbar^3 \left(\frac{2\pi}{M_{Cs}k_B T}\right)^{3/2} \tag{12.2}$$

be greater than order unity, where Λ is the deBroglie wavelength and M_{Cs} is the mass of one cesium atom. For typical cesium MOT density of $n = 10^{11}$ cm^{-3} the atomic spacing is $n^{-1/3} = 2.2$ μm, and a typical temperature of $T = 10$ μK gives a wavelength $\Lambda = 48$ nm. Thus, the typical phase space density is around $\rho = 10^{-5}$, which is a long way from unity. Typically, the BEC people start from around this point and use a combination of techniques, primarily evaporative cooling, that trade off the loss of atoms for the reduction of temperature to attain higher values of ρ. Typical BECs have $n \approx 10^{14}$ cm^{-3}, $N \approx 10^5$, and sub-μK temperatures [150].

In contrast to this typical pursuit, our experimental needs are also somewhat different:

- As discussed previously, in chapter 8, we require a large optical depth (OD) to maximize the Faraday measurement signal-to-noise ratio. We can express the optical depth as

$$\mathrm{OD} = \sigma_0 N/A = \sigma_0 n^{2/3} N^{1/3} = \sigma_0 n L \tag{12.3}$$

 where the trap volume is $V = L^3 = A^{3/2}$, and the optical cross section is $\sigma_0 = 3\lambda_0^2/2\pi = 3.47 \times 10^{-13}$ m^2 (for the cesium D2 line with $\lambda_0 = 852$ nm). We see

that the density requirement is slightly more important than the atomic number, but ideally both would be large.

- The shape of the cloud must be such that it ensures strong forward scattering. In essence, the cloud of atoms can be viewed as an array of antennas that emit into some spatial profile. Although typically we do not resort to cloud shaping to enhance this profile, it is an important characteristic as investigated in [115, 118]. One can also think of the atomic profile as determining the "distinguishability" of which atom interacted with each part of the probe beam, which is important for the preparation of entanglement through symmetry. This issue is discussed more in section 15.1.2.

- The atoms must be relatively well localized for purely technical reasons. Typically, there exist magnetic field gradients (both ambient and due to applied fields). For a given gradient, smaller samples see less field variation and consequently less atomic dephasing. Also, with smaller samples, one can use smaller field coil sources (hence lower inductance and faster switching times). Finally, given a constraint on the probe intensity of the experiment, it helps to have a smaller sample when the power of the probe beam is limited.

- Although we do not require extremely cold temperatures (by laser-cooling standards), we do require that the atoms be relatively cold for technical reasons. First, it is desirable to avoid dephasing collisions with atoms or cell walls, and also to avoid Doppler shifts. It is also necessary for the atoms to remain within beam for length of probe measurement. However, when the spatial profile is inhomogenous, whether in a cavity or in free-space, it may actually be desirable for the atoms to have a nonzero temperature to 'symmetrize' the measurement such that each atom (on-average) interacts with the same number of photons. In some relevant theoretical work, particularly with room-temperature vapor cells, finite temperature is used to simplify theoretical calculations [118].

12.3 Basic MOT Regimes

The dynamics of a MOT are rather complicated and have been modeled in several papers to varying levels of sophistication [149, 151, 152, 153, 154]. In this section, we present a

brief analysis of the different regimes of a MOT, following closely the discussion of [149].

12.3.1 Temperature Limited

For small atom numbers, a completely single-atom picture of the MOT is adequate. The usual trapping and cooling occurs, but because the atoms do not interact the size of the MOT is completely independent of atom number. The radius of the MOT can be obtained by matching the thermal energy to the potential energy of the trap, with larger temperatures corresponding to larger radii. This kind of MOT is not typical because of its small size with numbers $N < 10^4$. It usually takes effort to make a MOT this bad.

12.3.2 Multiple Scattering

As the number of atoms increases, an individual atom becomes more likely to absorb a photon emitted by another trapped atom. This interaction leads to a repulsion between atoms that naturally follows a inverse distance squared Gauss' law scaling. Because of this effect the density of the cloud becomes independent of the number of atoms in the cloud (or the radius goes as the cube root of N). This regime typically applies for atom numbers in the range of $10^4 < N < 10^6$.

The spatial distribution of the MOT will become non-Gaussian and uniform past some point, but this can come long after the onset of this regime since a Gaussian distribution can still have number independent density. The multiple scattering can also lead to excess heating and, under certain circumstances can contribute to nontrivial spatial structures as described in section 12.3.5.

12.3.3 Two Component

Under precise modeling, it can be shown that the confining potential of the magneto-optic system consists of two regions [149]. Close to the center of the trap the restoring force is linear in the distance, but farther away the force gets weaker and levels off. The turn-around occurs at approximately the point where the Zeeman shift from the trapping coil is comparable to the light shift from the trapping beams. For a large number of atoms, the inner region is filled and the outer less dense region become populated. Thus a MOT in this regime will appear as a dense central core surrounded by a less-dense "fluff." This

two-component regime occurs typically for atom numbers $N > 10^6$.

12.3.4 Optically Thick

A final regime occurs in which the absorption of the trapping beams becomes a significant effect. At our MOT settings, the OD *at* the trapping detuning is approaching order unity. This is especially important if the MOT is double passed such that not only is a single beam significantly depleted as it goes through, but the returning beam is also depleted to start, causing a significant asymmetry in the trapping dynamics.

12.3.5 Ring MOT

After achieving an experimental MOT with a relatively large atom number, we noticed that with slight misalignments of the trapping beams, many interesting "ring-shaped" MOTs could be created. Such ring MOTs have been observed and modeled since the early 1990s [155, 156, 157, 158, 159, 160].

It turns out to be quite easy to model the behavior by simply displacing four of the six beams laterally in a plane such that a vortex force is created. This single-atom model describes the basic effect, but the experiment showed additional features that this model does not describe. First, under certain configurations, the MOT would consist of both a ring and a central core (Saturn-like). Phase diagrams of the different trapping regimes are detailed in [157]. Second, unlike the simulations above, a strong clumping effect was observed in the ring. Instead of the ring consisting of a continuous distribution of atoms moving along the ring, there was instead a clump of atoms moving in the ring-path. When the ring was observed continuously, the orbital frequency was fast enough (≈ 100 Hz) that the ring appeared continuous. However, when the image acquisition timing was synchronized with the orbital frequency a clump approximately a quarter of the orbital path could be observed distinctly. This clumping motion could also be observed via the fluorescence signal.

Several explanations have been proposed for either effect, including interatom repulsion via fluorescence. In the case of the Saturn MOT however, removing the "core" with a pumping beam had no visible effect on the outer ring. The clumping could have been caused by periodic (power line) field noise, although the frequencies were not necessarily commensurate with the expected 60/120 Hz. Because the frequency is mostly determined

by the vortex force of the simulations, it is possible that field fluctuations need only "seed" the formation of the clump in some sense.

Most of these behaviors are likely due to a combination of these influences, which are interesting from a trapping perspective, but probably do not aid us in a quest for more optical depth. However, it remains possible that other experiments requiring ringlike geometries would benefit by harnessing a trap of this kind.

12.4 Review of Achieved Cold Atom Traps

Here we review what has been experimentally achieved in other laboratories with regard to cold atom trap parameters, primarily the measured atom number and density. These together with the atomic cross section imply the optical depth (with assumptions about cloud shape) as described above.

Typical traps have used trapping beam optical powers of order 10 mW and beam diameters of several millimeters. These standard MOTs are typically in the multiple-scattering or two-component regimes (see above or [149]). These traps have typical atom numbers in the low 10^7 range and densities in the range of 10^{11} cm^{-3} [161, 152, 156, 149].

Still considering the standard MOT configuration, but with an increased trapping beam size and power one can create a large MOT with many more atoms (below we see that the atom number scales as the dimension to the fourth power). These "monster MOTs" typically have beam sizes of greater than a cm in radius and overall trapping power of greater than 100 mW, producing a MOT in the optically thick regime. The Chu group created a very large cesium MOT with 4 cm diameter beams and huge 25 mW/cm^2 intensities, giving an atom number of 4×10^{10} and density of 4×10^{10} cm^{-3} (uniform across the sample) [153]. This implies an optical depth of near 40. The Ketterle group created a Sodium MOT (3 cm diameter beams, 10 mW/cm^2) with numbers around 10^{10} and a density of 10^{10} cm^{-3} [162]. Our standard cesium MOT has a beam diameter of around 2.5 cm and an intensity of around 6 mW/cm^2, giving atom numbers of around 3×10^9 and densities in the 10^{10} cm^{-3} range consistent with the other experiments above. Our peak optical depths are as high as 50 (beam smaller than cloud), but including all the atoms the effective optical depth is smaller by a factor of two or so. In conclusion, with larger more powerful MOTs, the number increases via brute force, but the density tends to decrease with size.

Dark-SPOT traps have been shown to increase the density by an order of magnitude or more [162, 163]. This is consistent with my measurements, although the number of trapped atoms does not typically increase dramatically, and the optical depth increases less than the density because of a typical drop in cloud size. (This is in contrast to our previous work [27] where unbelievably large atom numbers for our dark-SPOT parameters 10^{11} were reported, with a density of 2×10^{12} cm^{-3}.) Typical temporal Rubidium dark-SPOTs have numbers of 10^8 and densities around 10^{11} cm^{-3} before being evaporated to 10^6, 10^{14} cm^{-3} [164]. Mesoscopic traps can get to similar densities but with less number with 10^4 atom at 3×10^{14} cm^{-3} [165].

Several groups have reported partial success with compression techniques for increasing the density. In [166], a cesium MOT with number 4×10^8 and density 10^{12} cm^{-3} is created via transient field compression. In [167], a Rubidium MOT with 1×10^7 and density 5×10^{11} cm^{-3} is prepared similarly.

Finally, by using sideband cooling, one can prepare around 3×10^8 cesium atoms at a density of 10^{12} cm^{-3} as reported in [168].

12.5   MOT Physics

In this section, we detail various diagnostics for measuring the parameters associated with the atom trapping and cooling process. Some of these are quite standard while others are less so, but we include them all for the sake of completeness with adequate referencing and with no claims of originality. Good reviews include the references [149, 151, 152, 153, 154]. There are many great atomic physics books available for describing the below physics, including [65, 1, 38, 66].

12.5.1   Cooling Benchmarks

Following [1] we can label several physical benchmarks for cooling alkali atoms. When $\Gamma_D = \gamma$ we have:

$$T^* = \gamma^2 \frac{M_{Cs}c^2}{2k_B\omega^2} = 158 \text{ mK}. \tag{12.4}$$

The Doppler limit is obtained when match atomic cooling rates with heating rates in simple model

$$T_D = \frac{\hbar\gamma}{2k_B} = 125 \ \mu\text{K} \tag{12.5}$$

and the recoil limit is obtained when one considers the energy emitted to an atom when one last photon is emitted in a random direction

$$T_{rec} = \frac{\hbar^2\omega^2}{2k_B M_{Cs} c^2} = 99 \ \text{nK}. \tag{12.6}$$

The velocity corresponding to this temperature is $v_{rec} = 3.5$ mm/s. Note $T_D^2 = T^* T_{rec}$.

The Doppler limit was surprisingly beaten by order of magnitude via polarization gradient cooling (PGC) in early attempts at laser cooling. Describing this process requires a model with more than a two level atom. This PGC process typically results in atomic temperatures in the $1 - 10 \ \mu\text{K}$ [152]. It turns out that the recoil limit can also be broken, as mentioned in [1].

12.5.2 MOT Loading

All of the atoms loaded into the trap come from the background gas of cesium at room temperature $T = 300$ K, pressure $P_{Cs} = n_{Cs} k_B T \approx 10^{-9}$ Torr, and density $n_{Cs} \approx 3.2 \times 10^7$ cm^{-3}. For comparison, the spatial density if there were one atom per wavelength cubed would be $n = 1.6 \times 10^{12}$ cm^{-3} or $50,000$ times greater. From Maxwell-Boltzmann statistics, we have the following speeds [169]:

$$\text{Root mean square: } v_{rms} = \sqrt{3k_B T/M_{Cs}} = 237 \ \text{m/s} \tag{12.7}$$

$$\text{Mean: } v_{mean} = \sqrt{8k_B T/\pi M_{Cs}} = 219 \ \text{m/s} \tag{12.8}$$

$$\text{Most probable: } v_{prob} = \sqrt{2k_B T/M_{Cs}} = 194 \ \text{m/s}. \tag{12.9}$$

The largest force that a beam of any intensity can impart upon an atom (saturated excited level) is

$$F_{max} = \hbar k \gamma / 2. \tag{12.10}$$

Now define the capture velocity as that of an atom that loses all of its kinetic energy after

having passed through a beam imparting this maximal force

$$v_c = \sqrt{\frac{1}{2}}\sqrt{\frac{1}{2}}\sqrt{\frac{2F_{max}d}{M_{Cs}}} \tag{12.11}$$

where one factor of $\sqrt{1/2}$ is from halving the distance (on average) and the other is from halving the force. Factors of order unity are necessarily treated loosely in this analysis (and most others). The value of the capture velocity for cesium is $v_c = 17\sqrt{d}$ m/s/$\sqrt{cm}$, so for typical cm size beams the capture velocity is tens of m/s.

The number of atoms per second entering the trap volume V with low enough velocities $v < v_c$ to be captured [152] is calculated to be

$$R = \frac{1}{2}nV^{2/3}v_c^4\left(\frac{M_{Cs}}{2k_BT}\right)^{3/2}. \tag{12.12}$$

Because $v_c \propto \sqrt{d}$ and $V \propto d^3$, we have $R \propto d^4$ so the best way to increase the number of atoms in a basic MOT is to increase the size of the beams (keeping the intensity near saturation level). Alternatively, one can load with a directed, cooled beam from another atom source [170]. In an experiment by the Chu group, a very large MOT was prepared with large 5.5 cm beam diameters, intensities of 25 mW/cm^2 (approximately 2 W of power) resulting in 3.6×10^{10} atoms [153].

The collision rate with other cesium atoms is given by [171, 142]

$$\frac{1}{\tau_{Cs}} = n_{Cs}\sigma_{Cs}v_{rms}. \tag{12.13}$$

There are also collisions with Helium governed by $1/\tau_{He}$ where

$$P_{He} = \frac{k_BT}{\tau_{He}\beta_{He}}. \tag{12.14}$$

The loading dynamics of the MOT can then be described by

$$dN/dt = R - N(1/\tau_{Cs} + 1/\tau_{He}). \tag{12.15}$$

The solution of the loading equation is

$$N(t) = N_{ss}(1 - \exp[-t/\tau]) \tag{12.16}$$

with $\tau = (1/\tau_{Cs} + 1/\tau_{He})^{-1}$ and $N_{ss} = R\tau$. Note that if $\tau_{Cs} \ll \tau_{He}$ (i.e., the cesium-cesium collisions dominate with high background cesium pressure), then N_{ss} is independent of the background cesium density n because both R and $1/\tau$ are proportional to n. In practice, it is desirable to have a cesium pressure greater than the Helium pressure (or others), but not so great that the MOT beams are significantly absorbed or the collisional lifetime is made too small.

In [154], another source of loss due to two-body collisions in the trap is proposed as $dN/dt = -\beta N^2/V$. With the reported value of $\beta = 10^{-11}$ cm^3/s, this will only be a concern if $\beta n \approx 1/\tau$ [171, 142] . Because $\tau \approx 1$ s and $n \approx 10^{11}$ cm^{-3} in our experiment, the collisional terms are just starting to be an issue.

12.5.3 Pressure Analysis

Suppose we do not trust our pressure gauge and, in addition, we want to determine the background cesium and helium pressures. We can do this with the MOT by placing the background cesium density at two points (n_1 and n_2) by, for example, varying the temperature of the cesium reservoir (or closing the valve). Then, at each point we measure the steady-state loaded atom number and the loading time, giving us the points (N_1, τ_1) and (N_2, τ_2). One can then show from the above relations in the MOT Loading section that we have

$$\tau_{He} = \frac{\tau_2(N_1/N_2) - \tau_1}{(N_1/N_2) - 1} \tag{12.17}$$

$$\tau_{Cs,i} = \frac{\tau_{He}\tau_i}{\tau_{He} - \tau_i}. \tag{12.18}$$

These equations are useful for determining the background pressures independently.

12.5.4 Temperature Analysis

We start by defining the joint phase-space distribution for the atom cloud, which we assume to be Maxwell-Boltzman in the velocities (in other words, there is a well-defined temperature) and a Gaussian spatial distribution. That gives us,

$$\Pi(x, p, t = 0) = \frac{1}{2\pi\sqrt{r_0^2 M_{Cs} k_b T}} \exp\left[-\left(\frac{x^2}{2r_0^2} + \frac{p^2}{2M_{Cs} k_b T}\right)\right] \tag{12.19}$$

where r_0 is the initial 1-σ radius of the atom cloud, p is the transverse momentum coordinate and x is the transverse position coordinate (page 265 of [169]). I will choose the transverse direction to be perpendicular to the direction of free fall (vertical on the camera) such that there is no momentum change due to gravity. We want propagate this phase-space density in time under the assumption that the cloud expands freely such that the temperature remains constant,

$$\Pi(x, p, t) = \frac{1}{2\pi \sqrt{r_0^2 m k_b T}} \exp\left[-\left(\frac{(x + p\,t/M_{Cs})^2}{2r_0^2} + \frac{p^2}{2M_{Cs}k_b T} \right) \right]. \tag{12.20}$$

We want to compute the RMS velocity because that is the proper quantity related to the temperature. So we start by computing the expectation value of the square of the transverse cloud radius using our time-dependent phase-space distribution,

$$\langle r^2(t) \rangle = \int\int x^2 \Pi(x, p, t)\,dx\,dp = r_0^2 + \frac{k_b T}{M_{Cs}} t^2. \tag{12.21}$$

The RMS cloud *radius* as a function of time is given by,

$$\bar{r}_{\text{rms}} \equiv \sqrt{\langle r^2(t) \rangle} = r_0 \sqrt{1 + \frac{t^2}{\tau^2}} \tag{12.22}$$

where I have defined the *expansion time constant*, τ,

$$\tau = \sqrt{\frac{M_{Cs} r_0^2}{k_b T}}. \tag{12.23}$$

To compute the cloud temperature from measurements of the cloud radius as a function of time, we use a numerical fit to these equations. Although not as robust a technique in practice, the temperature can also also extracted from just the $t = 0$ derivative using

$$\bar{v}^2\big|_{t=0} \equiv \left(\frac{dr_{\text{rms}}}{dt} \right)^2 \bigg|_{t=0} = \frac{r_0^2}{\tau^2} = \frac{k_B T}{M_{Cs}} \tag{12.24}$$

and thus

$$T = \frac{M_{Cs}}{k_b} \left(\frac{dr_{\text{rms}}}{dt} \right)^2. \tag{12.25}$$

One can also measure the temperature by the decay time of an absorption/probe measurement signal [162]. For example, one could perform a measurement where the spins are

optically pumped and rotated vertically to give a polarization rotation signal. If the detuning is made large (or intensity small) enough, so that the spontaneous emission and tensor Hamiltonian decay terms are relatively negligible, the decay of the signal will be purely from the atoms leaving the probe beam laterally. Furthermore, if there are no mean lateral forces then the centroid of the distribution will always stay within the probe and one can easily extract the temperature from the decay curve. In reality, the atoms will either hit the bottom of the cell (from below this happens at 250 ms), or move out of the probe beam due to the momentum impulse of the pump beam (from below we see that this timescale will typically be 30 ms) but these timescales will be somewhat longer than the temperature diffusion timescale. We give a rough estimate of this method as follows. Assume the optical depth is simply

$$\text{OD} = \frac{N\sigma}{\pi r(t)^2} \tag{12.26}$$

$$= \frac{N\sigma}{\pi r_0^2(1 + t^2/\tau^2)} \tag{12.27}$$

$$= \frac{N\sigma}{\pi r_0^2(1 + t^2 k_B T/M_{Cs} r_0^2)} \tag{12.28}$$

where the above equation for the radius has been used. Factors of order unity are being ignored here. If this decay curve reaches half of its initial value at time τ, then the temperature will be roughly given by

$$T = \frac{M_{Cs} r_0^2}{\tau^2 k_B}. \tag{12.29}$$

For the timescale of 30 ms (from the pumping impulse) and assuming a MOT of $r_0 = 1$ mm, this gives a temperature limit of this technique of 18 μK.

12.5.5   Stern-Gerlach Analysis

This section will describe the position-momentum distribution of an atomic ensemble evolving under the forces of gravity and a magnetic field gradient. In the end, we predict the expected signal in a Stern-Gerlach time of flight experiment. Much of this analysis can be found in [105].

The simple model we will consider is that of a particular sublevel ($m_f = -4, ..., 4$) with

an initial Gaussian one-dimensional position-momentum distribution ($\Pi(x_0, p_0)$, ignoring the lateral dimensions and assuming gravity acts along x). The atoms in the sublevel m_f will experience both the force of gravity (g) and the force from the field gradient (dB/dz):

$$F = M_{Cs}g - m_f g_F \mu_B \frac{dB}{dz} = M_{Cs}a \tag{12.30}$$

where a is the acceleration. The evolution of the position and momentum will be

$$x = x_0 + p_0 t / M_{Cs} + at^2/2 \tag{12.31}$$

$$p = p_0 + M_{Cs}at. \tag{12.32}$$

Inverting, we write x_0 and p_0 as functions of x and p to give

$$x_0 = x - p_0 t / M_{Cs} - at^2/2 = x - pt/M_{Cs} + at^2/2 \tag{12.33}$$

$$p_0 = p - M_{Cs}at \tag{12.34}$$

$$dx_0 = dx \tag{12.35}$$

$$dp_0 = dp. \tag{12.36}$$

Now we can change the variables of an initially Gaussian distribution (with variances σ_x and σ_p) to give the updated distribution:

$$
\begin{aligned}
\Pi(x_0, p_0) dx_0 dp_0 &= C \exp\left[-\frac{x_0^2}{2\sigma_x^2} - \frac{p_0^2}{2\sigma_p^2} \right] dx_0 dp_0 & (12.37) \\
&= C \exp\left[-\frac{(x - pt/M_{Cs} + at^2/2)^2}{2\sigma_x^2} - \frac{(p - M_{Cs}at)^2}{2\sigma_p^2} \right] dx dp & (12.38) \\
&= \Pi(x, p, t) dx dp & (12.39)
\end{aligned}
$$

where C is for normalization. We now wish to calculate the quantity $\Pi(x, t) = \int_{-\infty}^{+\infty} dp\, \Pi(x, p, t)$, which describes the number of atoms falling through a particular point in space as a function of time. This value will be proportional to the experimental fluorescence signal from a thin "sheet" of laser light located a distance below the falling cloud. First we rewrite the distribution as a Gaussian in p:

$$\Pi(x, p, t) dx dp = C \exp\left[+a_2 p^2 + a_1 p + a_0 \right] dx_0 dp_0 \tag{12.40}$$

with the parametrization

$$a_0 \;=\; -\frac{1}{2}\left(\frac{(x+at^2/2)^2}{\sigma_x^2} + \frac{(M_{Cs}at)^2}{\sigma_p^2}\right) \tag{12.41}$$

$$a_1 \;=\; \frac{t(x+at^2/2)}{M_{Cs}\sigma_x^2} + \frac{M_{Cs}at}{\sigma_p^2} \tag{12.42}$$

$$a_2 \;=\; -\frac{1}{2}\left(\frac{1}{\sigma_p^2} + \frac{t^2}{M_{Cs}^2\sigma_x}\right). \tag{12.43}$$

Performing a simple "complete-the-square" integral over p gives the desired distribution:

$$\Pi(x,t)dx \;=\; \left(\int_{-\infty}^{+\infty} dp\,\Pi(x,p,t)\right)dx \tag{12.44}$$

$$=\; C\sqrt{-\pi/a_2}\,\exp\left(a_0 - \frac{a_1^2}{4a_2}\right)dx. \tag{12.45}$$

For the the repelled sublevels m_f there exists a value of the gradient for which gravity is overcome by the magnetic force. For the $m_f = 4$ sublevel of the $f = 4$ ground state of cesium, the value at which the total force is zero is

$$\frac{dB}{dz} \;=\; \frac{M_{Cs}g}{m_f g_f \mu_B} = 23.32 \text{ G/cm.} \tag{12.46}$$

Of course, for gradients larger than this in the full three-dimensional model the atoms will be initially pushed up, but then spill over the edges and eventually fall. Other tomography techniques are discussed in section 6.8.

12.5.6 Fluorescence Analysis

The total power scattered from a steady-state normal MOT [84] will be:

$$P_{scatt} \;=\; N\hbar\omega\Gamma\frac{(I_t/2I_s)}{1 + 4(\Delta/\Gamma)^2 + (I_t/I_s)}. \tag{12.47}$$

The intensity I_t used is the total intensity from the six beams of the MOT and the saturation intensity used is the $I_s = 27$ W/m^2 for isotropic polarized light [84]. Now the total power collected from a detector of collection area A_d at a distance of r from the MOT center is

$$P_{coll} \;=\; \frac{A_d}{4\pi r^2}P_{scatt}. \tag{12.48}$$

By using the above P_{coll} can be used to infer N with all other parameters known.

The optical depth can be inferred from the total measured number combined with a knowledge of the cloud size obtained from a CCD camera. The following two methods allow one to determine the optical depth more directly.

12.5.7 Measuring OD with Absorption Spectroscopy

Here we derive the absorption versus frequency spectrum so that we can determine the optical depth of a given sample. For example, if we would like to determine the scattering from the atoms populating the $f = 3$ ground state of cesium, we can calculate the absorption profile as follows. The power being radiated from an atom in a beam of intensity I with detuning close to resonance with the f line is given by $R_f = \sum_{f'} R_{f,f'}$ where

$$R_{f,f'} = \hbar\omega\Gamma P_{f'} \tag{12.49}$$

and

$$P_{f'} = \frac{(I/2I_{s,f,f'})}{1 + 4(\Delta_{f'}/\Gamma)^2 + (I/I_{s,f,f'})} \tag{12.50}$$

is the population of the excited state f'. For *linearly* polarized input light, the saturation intensity from one level is $I_{s,f,f'} = I_{s,lin}/S_{f,f'}$ where for the cesium D2 line [84]

$$I_{s,lin} = 16.5 \text{ W/m}^2 \tag{12.51}$$

and

$$S_{3,2} = 5/14, S_{3,3} = 3/8, S_{3,4} = 15/56 \tag{12.52}$$

$$S_{4,3} = 7/72, S_{4,4} = 7/24, S_{4,5} = 11/18. \tag{12.53}$$

Note that these transition strength factors satisfy $\sum_{f'} S_{f,f'} = 1$ to give the appropriate far-detuned saturation intensity.

Now the intensity of the probe beam as it crosses an element of length dx follows the

equation:

$$\frac{dI}{dx} = -nR_f. \tag{12.54}$$

If $I \ll I_{s,f,f'}$ for each f' then we can ignore the I dependence in the denominator of $P_{f'}$ and express the loss rate as $R_f = \sigma_f I$ where

$$\sigma_f = \sum_{f'} \frac{\Gamma\hbar\omega}{2I_{s,f,f'}} \frac{1}{1 + 4(\Delta_{f'}/\Gamma)^2} \tag{12.55}$$

is interpreted as the cross section in units of area. Under this assumption, the above differential equation can then be solved to give the intensity leaving the uniform sample as $I(x) = I_0 \exp[-\sigma_f(\Delta)nx] = I_0 \exp[-\mathrm{OD}(\Delta)]$, where $\mathrm{OD}(\Delta) = \sigma_f(\Delta)nx$.

Note that the on-resonance cycling transition ($f = 4$, $m_f = 4$ to $f = 5$, $m_f = 5$) optical depth is given by $\mathrm{OD}_0 = \sigma_0 nx$ where $\sigma_0 = 3\lambda^2/2\pi = \Gamma\hbar\omega/2I_{s,circ}$ and $I_{s,circ} = 11.0 \ \mathrm{W/m}^2$ for the cesium D2 line [84]. Thus, to convert any of the above hyperfine-specific on-resonance cross sections (or optical depths) we use the conversion

$$\sigma_{0,lin,f,f'} \equiv \frac{\Gamma\hbar\omega S_{f,f'}}{2I_{s,lin}} = \sigma_0 \frac{I_{s,circ} S_{f,f'}}{I_{s,lin}}. \tag{12.56}$$

If not otherwise noted, any detuning independent optical depth stated in this thesis will be the on-resonance optical depth, $\mathrm{OD} = \mathrm{OD}_0$, by default.

Now we introduce temperature dependence into the cross sections, primarily for the purpose of deriving the distribution of the room temperature background gas. Hence we allow for a Gaussian distribution (with temperature implied by the parameter α_D) of frequencies about the nominal detuning and convolve as follows.

$$\sigma_{lin,f,f'}(\Delta, \alpha_D) = \sigma_{0,lin,f,f'} v_{f,f'}(\Delta, \alpha_D). \tag{12.57}$$

The $v_{f,f'}$ is the so-called Voigt lineshape and is given by

$$v_{f,f'} = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \frac{\exp[-t^2]}{1 + 4\left(\frac{\Delta_{f'} - \alpha_D t}{\Gamma}\right)^2} dt. \tag{12.58}$$

Notice that for $\alpha_D = 0$ we return to the above case $v_{f,f'} = 1/(1 + 4(\Delta_{f'}/\Gamma)^2)$. In the

opposite high temperature limit where $\alpha_D \gg \Gamma$ we have

$$v_{f,f'} \approx \frac{\Gamma}{\alpha_D} \frac{\sqrt{\pi}}{2} \exp[-\Delta_{f'}^2/\alpha_D^2]. \tag{12.59}$$

The parameter α_D is related (ignoring factors of unity) to the mean velocity via the Doppler shift by $\alpha_D = 2\pi\Delta f = 2\pi(c/\lambda)(v/c)$ so $v = \alpha_D\lambda/2\pi$. In practice, for room temperature speeds of approximately 100 m/s the Doppler shift will be a few hundred MHz.

Similar calculations for sodium exactly reproduce the absorption spectrum of Figure 2 in the Ketterle group dark-SPOT paper [162]. The two atomic spectra look remarkably different for similar OD_0 because the cesium hyperfine levels are much farther apart in frequency than the sodium levels.

12.5.8 Measuring OD with Faraday Rotation

The optical depth can also be measured by observing the Faraday rotation of the probe beam (described at length in previous sections and experimentally in coming chapters) and using the OD as the only free parameter. Of course, this requires assumptions about both the model and the pumping efficiency. With regard to the model, there is sufficient faith in the atomic physics modeling that the only uncertainty should come from order-unity estimations such as in the probe beam area and intensity. The assumptions about the pumping can be justified if the degree of pumping is independently confirmed. Several methods of estimating the pumping are described in the tomography section of this chapter. One such method uses the same probe signal but in different configurations where the nonlinearity of the tensor effects can be used to estimate the efficiency with the ratio of two measurements, such that the pumping efficiency is the only fit parameter. Using probe measurements to fit the OD is completely valid, but one should be careful in comparing measured ODs from different stages of the experimental procedure (before or after pumping, etc.) because systematic effects could change the number of atoms within the beam.

12.5.9 Optical Pumping Simulations

A key stage in all of the experiments considered in this work is state preparation via optical pumping. Successful optical pumping provides an oriented (separable) collective spin-state, which can then be used for measuring rotations or preparing more complicated entangled

states. The concept of optical pumping was developed many decades ago, and the logic behind it is (in retrospect) quite intuitive given a basic understanding of transition strengths and selection rules. Nevertheless, it is a somewhat complicated task to numerically simulate the optical pumping of an alkali atom with full consideration of all transition strengths and all hyperfine levels. In this section, we briefly describe two simulations of optical pumping that both give information regarding expected pumping times, efficiencies, etc.

The first model simply uses the single-atom adiabatically eliminated master equation (including spontaneous emission) from equation (5.44). The assumptions of that derivation assumed that the saturation parameter was small, limiting the simulation to situations the $\Delta \gg \gamma$, which may not always be the case in a realistic experiment. However, the advantage of this model is that the excited states are not included and the detuning does not appear directly as a numerator term in the Hamiltonian, thus the numerical evolution can use relatively large timesteps and be less computationally intensive.

The second model is the full unconditional master equation of equation (5.15) prior to adiabatic elimination. Because the excited states are used and the detuning appears as a direct term in the Hamiltonian (so we must have $\Delta \approx \gamma$), the numerical evolution is necessarily more computationally intensive than the adiabatically eliminated model. However, this model allows pump and repump parameters that saturate the transition and faster pumping scenarios can be effectively simulated.

The first state prepared is the usual coherent spin-state of $|f = 4, m_f = 4\rangle$ by applying circular polarized light at the 4–4' frequency along with repump light to move atoms from $f = 3$ to $f = 4$. Because there is no $|f = 4, m_f = 5\rangle$ sublevel, the target state is a dark state.

The second state is the $m = 0$ state of $|f = 4, m_f = 0\rangle$ by applying linearly polarized (π) light on the 4–4' transition and repump light. This state is a dark state because of the zero transition strength between $|f = 4, m_f = 0\rangle$ and $|f' = 4, m_f = 0\rangle$.

The third and final state is the incoherent mixture of the two oppositely oriented extended states:

$$
\begin{aligned}
\rho_{split} &= (|f = 4, m_f = 4\rangle\langle f = 4, m_f = 4| + \\
&\quad |f = 4, m_f = -4\rangle\langle f = 4, m_f = -4|)/2 \qquad (12.60) \\
&= (|\uparrow\rangle\langle\uparrow| + |\downarrow\rangle\langle\downarrow|)/2. \qquad\qquad\qquad (12.61)
\end{aligned}
$$

This "split" state can be prepared with linear polarized π light on the 4–3' transition and suitable repump light, because there are no $|f' = 3, m_f = \pm 4\rangle$ sublevels. This state is potentially useful for experiments that aim to create an "effective" two-level system so as to remove tensor effects [112]. It could also be used to simulate Polzik's "double-cell" experiments, but with both atomic ensembles cold and located at the same position. Such an experiment would not create spatially separated entanglement, unless the populations were physically separated in a state-dependent fashion after pumping (and probing).

The simulations are useful because they allow one to calculate the speed and efficiency of a particular pumping scheme. Also, potential problems can be avoided and realistic scenarios can be tested. For example, the necessity of a holding field along a quantization axis can be determined given an assumed background field uncertainty or noise. The importance of the repump polarization, intensity, and detuning can be investigated. Finally, the effect of including all of the hyperfine excited states can be evaluated. The code can also be easily adapted to the D1 transition, which may allow for more efficient pumping schemes in some cases.

Once a pumped state has been prepared one needs to be able to measure the degree to which the pumping has been achieved. Relevant tomography techniques are discussed in section 6.8, including most notably the use of the tensor Hamiltonian to determine the pumping efficiency. Corresponding measurements of the pumping efficiency are presented in section 13.6.

12.6   Motional Effects

In this section, we consider both how much atomic motion (due to temperature, gravity, and pump/probe forces) will affect the subsequent probing.

12.6.1   Motion Out of the Probe Beam

Now we consider, in rough fashion, how much temperature and gravity will move the atoms out of the probe beam. In the experiment, the atoms are trapped, then released, and then probed as they diffuse and fall out of the probe beam (without further trapping). So both temperature and gravity will place hard constraints on how long we are able to measure the atoms.

First, considering temperature, the average speed (in one direction) of the atoms will be $v = \sqrt{k_B T / M_{Cs}}$ due to equipartition. For reference sake, this gives $v = 8, 25, 79$ mm/s for $T = 1, 10, 100$ μK respectively and $v = 137$ m/s at room temperature $T = 300$ K. (Note that the recoil velocity limit for cesium is $v_{rec} = 3.5$ mm/s, which corresponds to sub-μK temperatures.) Thus, for typical temperatures of 10 μK the time to move across one wavelength of 852 nm is 34 μs and the time to move across a trap of width 1 mm is 40 ms. This latter time is acceptable because most measurement times are limited to order 1 ms, but if measurements are made with a smaller beam waist, the lateral temperature diffusion will become a significant issue. Note however that the dark-SPOT temperature is 1 mK [162], thus the above time will be reduced from 40 ms to 4 ms. So cooling is absolutely necessary.

Gravity places similar constraints on the experiment time. In the experiment, the probe beam is oriented vertically, thus gravity alone will not in principle remove atoms from the probe beam. However, we are also concerned with atoms falling out of the horizontal pump beam during pumping. The time it takes the atoms to fall a distance d due to gravity is simply $t = \sqrt{2d/g}$. For the typical trap size of 1 mm this gives a time of 14 ms and the time to fall to the bottom of the vacuum chamber (12 in.) is 250 ms.

12.6.2   Pump Beam Forces

One concern is how much momentum a single pumping beam will impart onto the atoms. If it is too much, then the atoms may move out of the probe beam. Assume that N_p photons are absorbed on the way to the dark state of $|f = 4, m_f = 4\rangle$, these will impart a momentum of $\Delta p = N_p \hbar k = M_{Cs} \Delta v$ onto an individual atom. The single photon recoil velocity is $v_r = \hbar k / M_{Cs} = 3.5$ mm/s. So if we assume an $N_p = 8$ (required to move from $m_f = -4$) we get $\Delta v = 2.8$ cm/s. For a typical measurement time of $\tau = 1$ ms, the distance moved during the measurement time is $\Delta d = \Delta v \tau = 28$ μm. The typical probe beam size is a few mm, so this can mostly be ignored, but for not too much smaller traps this would be a major source of misalignment.

12.6.3   Probe Beam Forces

There is possible interest in moving the atoms to increase the optical depth. The previously derived equations, mostly the scalar term, give the energy of the atoms as a function of space

(when the intensity of the probe beam is given as a function of space). The propagating probe beam will impart two types of forces: dipole and spontaneous emission [38]. The dipole force for a one-way free-space beam will act lateral to the beam, the spontaneous emission (absorption force) will act along the direction of the beam.

First consider the situation for a two-level atom, which is valid at large detunings. The energy shift for the ground state is given by

$$U = -\frac{\hbar\Omega^2}{4\Delta} \tag{12.62}$$

where $\Omega = E_0 d/\hbar$ and the detuning is in radians/sec [38]. As before we use $E = \sqrt{2P/\epsilon_0 cA}$. The lateral force is then given by the derivative of this quantity in the directions perpendicular to the beam. If the radius of the beam is r, the force will then be of order $F = U/r$. If we assume this to be a harmonic force, with $F = -kr$ we get $k = U/r^2$ and the frequency of oscillation $\omega = \sqrt{k/M_{Cs}}$. The time to reach the center of the well from the edge is then a quarter of the period $\tau = (1/4)2\pi/\omega = 2\pi\sqrt{M_{Cs}r^2 4\Delta/\hbar\Omega^2} \propto r^2 \propto \sqrt{1/P}$ at fixed power.

Note that the dipole force is inversely proportional to the distance scale, thus if we had counter-propagating beams creating a one-dimensional lattice, the intensity would vary over λ instead of r and the force would be about $r/\lambda \approx 1$ mm/1 μm $= 1000$ times greater. The absorptive force on the other hand is

$$\begin{aligned} F_{sp} &= \hbar k\gamma\rho_{ee} \tag{12.63} \\ &= \frac{\hbar k s_0 \gamma/2}{1 + s_0 + (2\Delta/\gamma)^2} \tag{12.64} \end{aligned}$$

where $s_0 = I/I_{sat}$ [38]. Let us convert this into the time it takes to move an atom across one radius, $r = (F_{sp}/M_{Cs})\tau^2/2$. This leads to $\tau = \sqrt{2rM_{Cs}/F_{sp}}$. For cesium we consider these quantities at $P = 500$ mW and $r = 1$ mm. From this simulation we see, that even at these high powers the timescales are typically 100 ms for detunings of 10s of GHz, 10 ms for GHz and, around 1 ms for 100 MHz. At small detunings, F_{sp} is larger and the beam pushes the atoms more than it contracts them laterally, at larger detunings the contractive dipole force F_d is larger. The crossover detuning depends only on the width of the beam and, for the cesium D2 line, is $\Delta/2\pi \approx 57r$ GHz/mm.

Chapter 13

Semiclassical Data

In this chapter, we begin to present experimental results from the apparatus described in chapter 11. We describe data relevant to the optimization of the experiment and the confirmation of the *semiclassical* measurement theory introduced in chapter 6. Hence we consider only the evolution of the quantum atomic state due to the classical moments of the input beam or vice versa. Collective noise effects are neglected entirely as we are only interested in the evolution of the mean moments. A single atom picture is completely adequate and the effect of measuring many atoms at once is only stronger forward scattering and a high single-shot signal-to-noise ratio, which allows us to get meaningful results without averaging over many trials. In contrast, the next chapter describes attempts to leverage this high signal-to-noise ratio to observe collective quantum projection noise as described in chapter 8.

We begin by describing the experimental procedure for optimizing the signal-to-noise ratio. Then several experimental results are summarized that confirm the tensor Hamiltonian theory previously developed. The experimental results presented in section 13.4 are from the paper [29], while every other result is previously unpublished. Included are demonstrations of the magic angle in Larmor precession and the use of the tensor Hamiltonian for practical tasks such as measuring the optical pumping efficiency. We close by discussing the magnetic-resonance scenario, which is relevant to the stroboscopic modulation scheme of the next chapter.

13.1 Basic Experimental Procedures

13.1.1 Aligning the Experiment

In this section, we describe the procedure for aligning our apparatus in order to trap and observe the spin-state of a cold atomic cloud. Our specific magneto-optical trap (MOT) apparatus has been described in chapter 11, and the physical considerations related to the preparation of a MOT are described in chapter 12 and the references therein. Prior to the alignment of the MOT, the DC magnetic field and first-order gradients (due largely to the ion pump) are cancelled with the large Helmholtz coils detailed in appendix A using a fluxgate magnetometer. The MOT trapping light is derived from an injection-locked diode laser and is passed through an acousto-optic modulator, which acts as a switch (along with a slower mechanical shutter). This beam is then telescopically expanded to fill standard one-inch diameter optics. When being reflected or transmitted at nonnormal incidence, larger two-inch diameter mirrors and optics are used. At the output of the expansion an iris is placed such that the size of the subsequent trapping beams can be adjusted. The beam is then split into three trapping beams, one for each axis, with waveplates and large polarizing beam-splitters, before each is appropriately circularly polarized, sent through the cell, and retro-reflected back through the cell with a mirror and another quarter-waveplate. Irises are also placed along each of the three input beams to aid with alignment.

The optical access ports in the magnetic shield are registered to the inner magnetic field coils including the MOT coils, thus the total shield structure provides a convenient reference for coarsely aligning the MOT beams. The structure is placed about the quartz cell and the beams are then aligned to it. For each beam, one input mirror is used to align the beam to the input shield port, while another is used to align it to the output shield port. The retro-reflecting mirror is then adjusted such that the returning beam is aligned to the closed-down input iris viewed with an IR camera. Aligning in this fashion makes a MOT easily visible using typical MOT parameters. In our case, we use 90 mW total of trapping light, a detuning of 15 to 20 MHz and several mW of repump power along one axis, and a magnetic field gradient of approximately 10 G/cm. Further alignment of the beams is achieved by shrinking the iris common to all trapping beams and aligning to achieve a small MOT centered at the magnetic field zero. Fluorescence measurements (see section 12.5.6) can then be used to optimize the atom number, but given the previous alignment

221

procedure, improvements greater than order unity will not be possible.

The parameter of interest in the experiment is the optical depth that the probe beam "sees" as it passes through the atomic ensemble. To align the probe beam to the MOT, it is first sent vertically through the cell and aligned to the input and output shield ports. The trapping beams are then reduced in size via the iris to produce a small MOT that is well localized at the magnetic center. To align the probe beam to the MOT center, the probe detuning is tuned near the 4–5' resonance, placed at full power, and used to destroy the MOT as observed by the CCD camera. The probe beam can also be reduced in size with an iris to further improve the alignment. The alignment can also be optimized by turning off the trap and maximizing the fluorescence observed from the probe beam as the atoms fall. The optical pumping beam is similarly aligned, but along the horizontal x-axis, as opposed to the z-axis of the probe.

We are also concerned with the polarization alignment of the optical beam because certain elements can rotate the polarization in unintended ways as discussed in section 3.2.5. First, we would like to prepare an optical beam polarized along the x-direction at the location of the atoms. Because the input window (as well as input mirrors) can impart nontrivial rotations (if the incidence angle is slightly nonideal) we place two compensating waveplates after the polarizing beam cube and before the mirrors that direct the probe beam into the cell. These waveplates are coarsely adjusted such that the input polarization is roughly along x and then the waveplates are further fine tuned using the atoms as will be discussed in section 14.5.2. Second, the mirrors on the output of the cell can arbitrarily rotate the optical polarization. Thus we must include a set of compensating waveplates at the output of the cell and prior to the polarimeter. As opposed to the input compensating waveplates however, we must arrange the output compensating waveplates to compensate the mirror's rotation for *any* input state. In practice this can roughly be done with three waveplates (e.g., two quarter-waveplates and one half-waveplate). The procedure involves using the polarimeter in multiple modes (e.g., measuring $\hat{S}_x$ and $\hat{S}_z$) to adjust the compensating waveplates for multiple known input states (e.g., linear polarization of arbitrary angle should give zero $\hat{S}_z$).

For the data presented in this chapter, we typically use one of the two subtracting detectors, which were described in section 11.6. For example, with absorptive measurements where it is desirable that the probe has minimal effect on the atoms, we use the low-power

detector (which is optical shotnoise limited above 1 μW). For measurements where we are investigating the effect of the probe itself on the atomic state, e.g., spontaneous emission decay, we require larger intensities and hence we use a high-power detector (which is optical shotnoise limited above 1 mW). The beam profile in all of these measurements varies, but unless otherwise noted is generally non-Gaussian (near flat-top profile) with around 1.5 mm radius overall but closer to 1 mm when operationally measured assuming a Gaussian profile (see `http://qwiki.caltech.edu/wiki/How_To_Measure_A_Beam_Waist`).

13.1.2 Absorptive Measurements: Optical Depth Optimization

The on-resonance optical depth (OD) is the key parameter for optimizing the experimental signal-to-noise ratio, as discussed in chapter 8. In principle, it is possible to measure the OD with either the absorption or dispersion of the probe beam. In the next section, we describe the dispersion case where the sample is cooled, optically pumped, and measured via Faraday rotation. In this section, we describe absorption measurements that have two distinct advantages over dispersion measurements when measuring the trap density characteristics. First, especially for the dark-SPOT configuration, the OD can be measured *continuously* with the absorption technique such that it can be quickly and conveniently optimized in real time. This is as opposed to the dispersion case where the sample must be prepared, released, and pumped, which takes several seconds per trial. In addition, with absorptive techniques one can neglect state preparation uncertainties (e.g., pumping efficiency) that come into any dispersion analysis. The preparation of a known initial state is often necessary in dispersion measurements because the distribution of spin population in a trap is often nearly isotropic such that no mean phase-shift or polarization rotation is expected. In practice, one does see such a rotation due to misalignments, but it is difficult to make sense of the complicated signal.

To perform the absorption measurement, the probe beam is coarsely aligned to the MOT (or dark-SPOT) as described previously and the power detected on one channel of the polarimeter (with the polarizing beamsplitter removed). The frequency of the beam swept over the cesium hyperfine manifold via the PZT input on the Vortex diode laser that seeds the slave probe laser. Of course, one must ensure that the slave laser remains injection locked over the entire sweep range prior to making the measurement. The probe power is set to be small and the sweep time is set to be sufficiently fast (around 0.1 ms) that the time
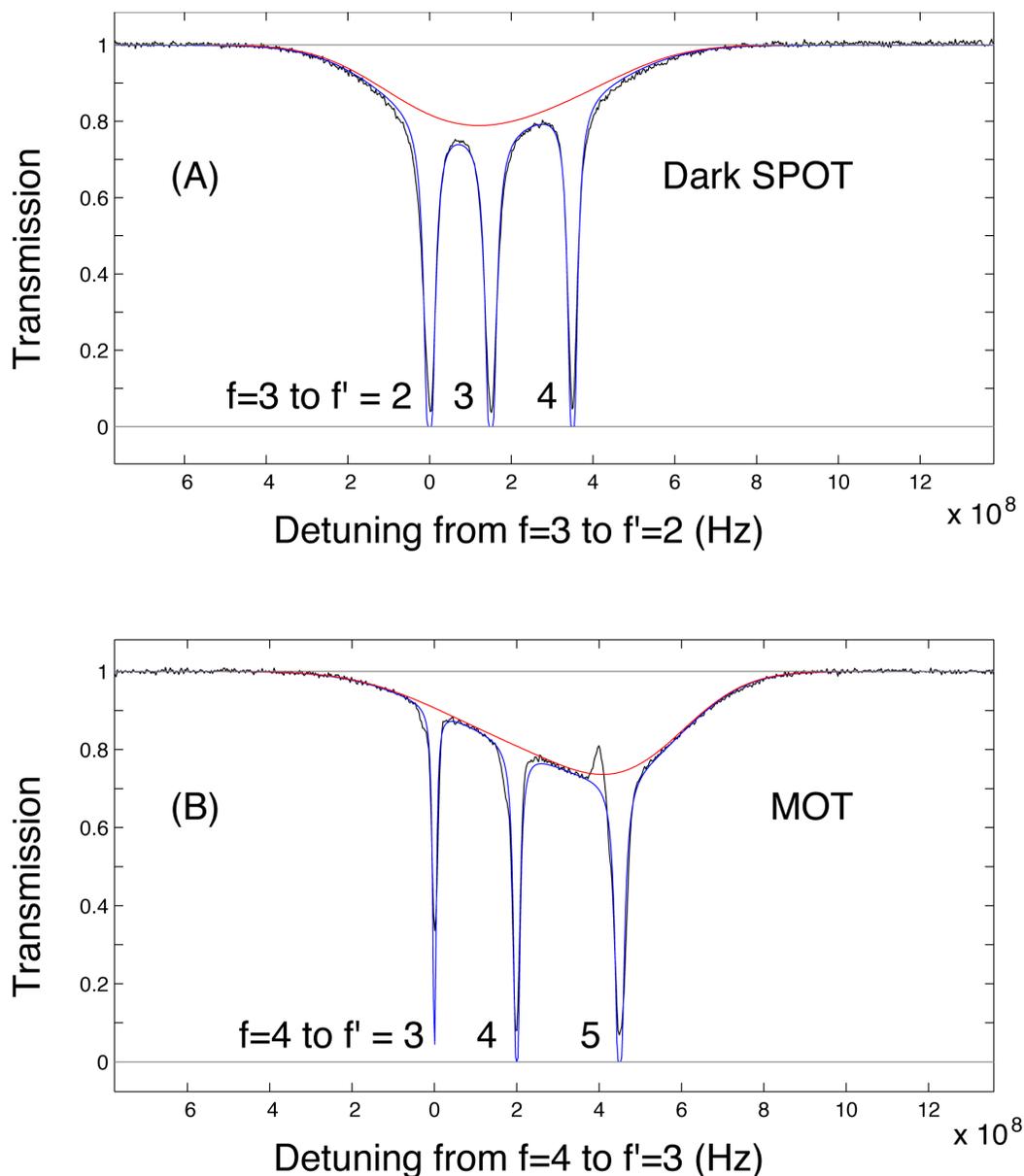
Figure 13.1: Absorption spectrum measurements of the cold atom cloud optical depth. (A) Absorption spectrum for a dark-SPOT with light tuned to $f = 3$ lines ($OD_0 = 117$). (B) Absorption spectrum for a MOT with probe tuned to $f = 4$ lines ($OD_0 = 49$). The Doppler broadened background gas contribution is independently fit without cold atoms, leading to the upper red curve. The black curve is the noisy absorption measurement and the blue curve the fit. For both measurements, the probe beam was set at a weak power of 2 μW and a beam radius of approximately 100 μm and the low-power detector was used. The time to take each trace is ≈ 40 μs.

spent near resonance is small and the effect of the probe on the trap dynamics is negligible. If a smaller beam is used one will get larger measurements of the OD because only the maximal density region of the cloud is measured.

Measuring the absorption spectrum is a much more robust technique than measuring the absorption at only a single probe frequency for several obvious reasons. First, with large OD samples if one tries to measure the absorption near resonance, the technical noise will prohibit any measurement of the very small amount of remaining light that gets through the atomic cloud, thus any useful information will only come from sufficiently far-off resonant absorption. Second, the probe beam has nonnegligible but slow power and frequency fluctuations that will make any single-point measurement unnecessarily uncertain. By measuring the spectrum, the frequency and power level can also be fit (knowing the hyperfine splitting of cesium), and the OD will be determined by the quality of the fit in the off-resonant regions of the spectrum. The theory for fitting the absorption spectrum to extract the on-resonance OD was described in section 12.5.7.

Typical measurements of the OD for both a standard MOT and a dark-SPOT are shown in figure 13.1. The optical depths shown here are larger (by order unity) than the optical depths used in subsequent dispersion measurements because we use a small beam to measure the optical depth at the peak of the cloud distribution and a large beam to measure all of the atoms dispersively. Notice that we have accounted for the presence of a significant amount of Doppler broadened background gas and included its density and temperature as fit parameters. Thus this is a useful means of measuring the background cesium pressure whose result should be consistent with the MOT loading time technique described in section 12.5.3. Also note that the quality of the fit near the on-resonance points is not very good for technical reasons and the actual information is contained in the high quality fit of the off resonant frequencies. Near resonance one can expect some dispersive deflection of the beams that also affect the measurement. Because of mode-matching (i.e., beam size versus cloud size) considerations, the measurement of OD should be taken as valid to within only order unity because we have not accounted for, or imaged, the spatial dependence of the OD. However, given a fixed beam profile, relative measurements of the OD can be much more precise. Notice that in figure 13.1B, when the probe frequency matches the the MOT trap frequency (red of the 4–5' transition) the probe is actually amplified, consistent with previous studies of nonlinear MOT spectroscopy (see page 72 of [172]).

One can observe this absorption spectrum continuously in time by sweeping the probe frequency at greater than kHz frequencies. This is an essential step for optimizing the OD of the dark-SPOT trap in real time by observing the degree of absorption while tweaking trap parameters (beam alignment, depump power, repump power, etc.). This technique can also be used to measure the OD as a function of time by turning off the trap and then measuring several spectra in a row before the atoms have fallen and diffused away. Because the atoms typically take several milliseconds to disappear and each sweep can be performed in around 0.1 ms (limited by the diode PZT sweep rate) one can get sufficient time resolution to observe interesting atomic compression and/or expansion dynamics.

13.1.3   Dispersive Transient Measurements: Faraday Rotation

To perform the Faraday rotation measurement we use the following procedure orchestrated by the computer system described in chapter 11. The atoms are trapped and cooled with either a MOT or a dark-SPOT trap as described previously. The trap is then turned off and the anti-Helmholtz magnetic field decays in approximately 1 ms, the rate being enhanced by the FET switch discussed in section A.5.4.1. We optionally use polarization gradient cooling at this stage. While the MOT field is decaying, the atoms are optically pumped for a few milliseconds using circularly polarized, 100 μW beam with a diameter of approximately a cm. During pumping a magnetic field (approximately 100 mG) is applied along the x-axis to define a quantization axis. The probe beam is then switched on at a preset power and detuning and the measured Stokes component recorded in time. (The probe beam parameters are such that the saturation parameter is small and the adiabatic elimination analysis of chapter 5 is valid.) Concurrent with the probe beam being turned on, magnetic fields are switched on and modulated according to the particular experiment being performed. Our magnetic field apparatus allows us to apply Gauss level fields over relatively fast timescales (μs) in an arbitrary direction.

First, we demonstrate the continuous observation of Larmor precession due to a perpendicular magnetic field via Faraday rotation of the probe beam. Figure 13.2A displays the relative orientation of the spin-state, magnetic field, and probe polarization after pumping. The spin-state is initially aligned along the x-direction with an optical pumping beam and magnetic field, then all fields and beams are turned off. Next, a field is applied along the y-direction, the probe beam is turned on, and the linear polarization rotation of the probe
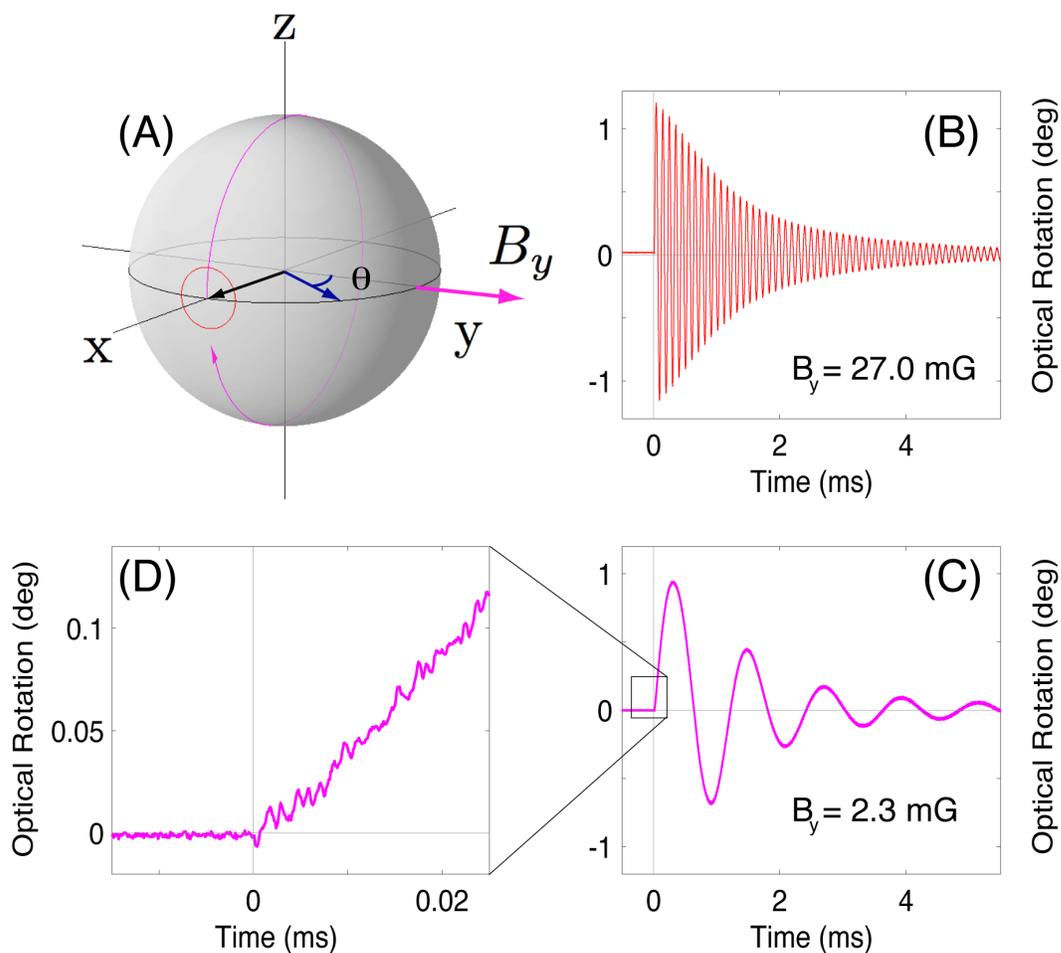
Figure 13.2: Real time, single-shot observation of Larmor precession via Faraday rotation. The probe beam was relatively weak at power 50 μW, radius 1.1 mm, and detuning of 500 MHz, and the low-power polarimeter detector was used. The atomic sample optical depth is approximately 20. (A) Schematic showing the path of the spin-state on the collective Bloch sphere and the orientation of the probe polarization (in this case $\theta = 90$ degrees). (B) The perpendicular DC field is 9.5 kHz/γ_{B4} where $\gamma_{B4} = 350$ kHz/G for the f=4 ground state. Fields giving frequencies of several hundred kHz can be obtained with Gauss level fields. (C) A weaker DC field of 0.8 kHz/γ_{B4} is applied. (D) A zoom in of the lower frequency plot. At small times, the estimation of the field amounts to finding the slope of the curve masked by optical shotnoise.
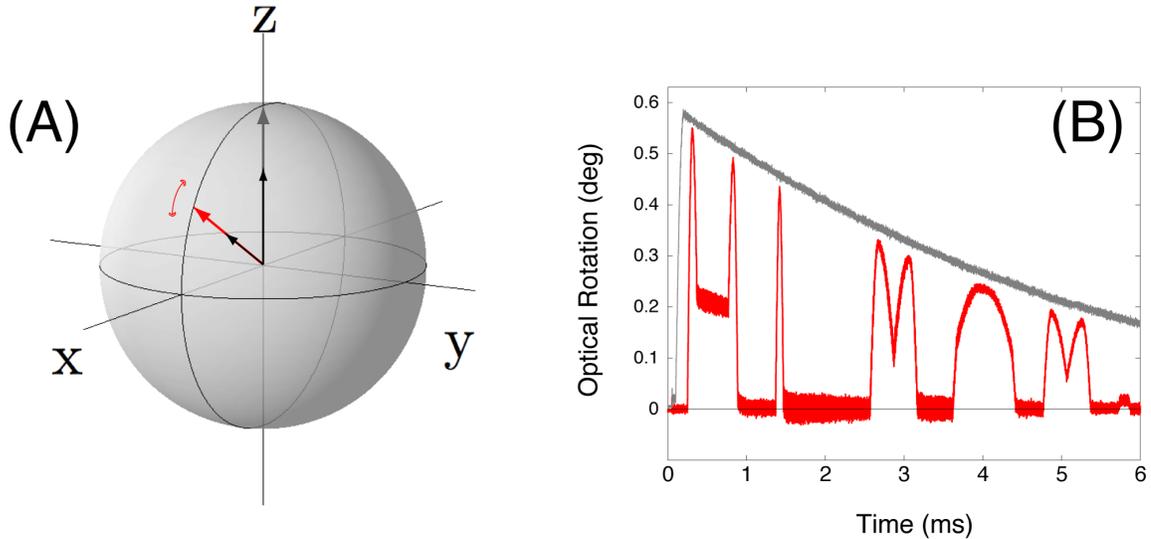
Figure 13.3: After optically pumping along x, the holding field is left on, parallel to the spin-state. The field is then slowly redirected in the xz-plane at a rate slower than the Larmor frequency of the field (110 kHz/γ_{B4}) and the spin-state adiabatically follows. The probe measures the z-component of the spin-state as the field is *arbitrarily* rotated in time by the experimentalist. The probe beam for this data was set at power 50 μW, radius 1.1 mm, and detuning 1 GHz and the low-power polarimeter detector used.

beam is measured with the polarimeter. As described in chapter 6 we observe the Larmor precession of the moment $\hat{F}_z$ via Faraday rotation of the probe beam and a measurement of $\hat{S}_y$. Figure 13.2B shows the resulting single-shot measurement and figure 13.2C shows a similar measurement with a smaller field. These plots and a zoom-in of figure 13.2C displayed in figure 13.2D show the high single-shot signal-to-noise ratio. This last figure displays the electronic and photon shotnoise along with the early evolution of the spin-state. The resulting signal-to-noise ratio is as predicted in section 8.2.2. Magnetometry with weak fields amounts to estimating the slope of this line in the presence of masking optical shotnoise as discussed in chapter 10.

The only dynamics of the spin-state here appear to be the Larmor precession and decay. If the measurement were truly nondemolition this would be the case, but we see in the next section that one can observe the effect of the tensor Hamiltonian on the evolution of the oscillating spin-state via the lifetime of the decay. After this discussion we discuss the measurement and sources of the decay more fully.

Another simple experiment involves keeping the magnetic field applied along the spin polarization direction (initially x) and then rotating the field (while keeping the field at a

constant magnitude) such that the spin-state adiabatically follows the field direction. As long as the rotation rate (typically kHz scale) is much less than the Larmor precession rate about the field (typically hundreds of kHz) then the adiabatic approximation will remain valid. One example of the arbitrary way in which the spin-state can be dragged around is shown in figure 13.3. Notice that at certain nonsmooth points of the trajectory, the spin-state "falls off" the field direction nonadiabatically, leading to small observed Larmor precession. This holding field technique is used to isolate the tensor Hamiltonian effect of the spins on the light (while reducing the tensor Hamiltonian effect of the light on the spins) in section 13.4. (Note that when the background cesium pressure is sufficiently high, the background gas alone will give a Faraday signal that changes as the field changes, which must be distinguished from the contribution of the cold atoms.)

When tweaking up the experiment, the above Larmor precession (or the adiabatic dragging signals) are used to optimize the experiment over various experimental parameters, including the pumping beam parameters, the trap parameters, and the cooling procedure. Further, the background fields can be zeroed and the applied fields calibrated in strength by measuring the frequency of the induced Larmor oscillations.

13.2 Faraday Rotation and the Magic Angle

As discussed in section 6.5, and shown in figure 5.1, the tensor Hamiltonian leads to interesting nontrivial dynamics when the spin-state is rotating due to a magnetic field. Here, following the results of [25], we present experimental results confirming the expected behavior. In figure 13.2A, we show the angle of the probe linear polarization relative to the applied magnetic field, which is perpendicular to the spin-state oscillating in the xz-plane. At zero degrees, for example, the tensor Hamiltonian effect on the spin-state will cause the observed spin decay to occur faster than the decay expected due to spontaneous emission or other sources. This is a single-atom dephasing effect that is purely Hamiltonian and reversible in principle. If given sufficient time to act and ideal experimental circumstances, this dephasing should eventually lead to rephasing and hence an echo in the observed decay signal. Without spontaneous emission or inhomogeneities, this revival would be complete. However, we do not usually observe the echo cleanly primarily due to probe intensity inhomogeneity over the atomic sample, which causes different subsamples to dephase at different
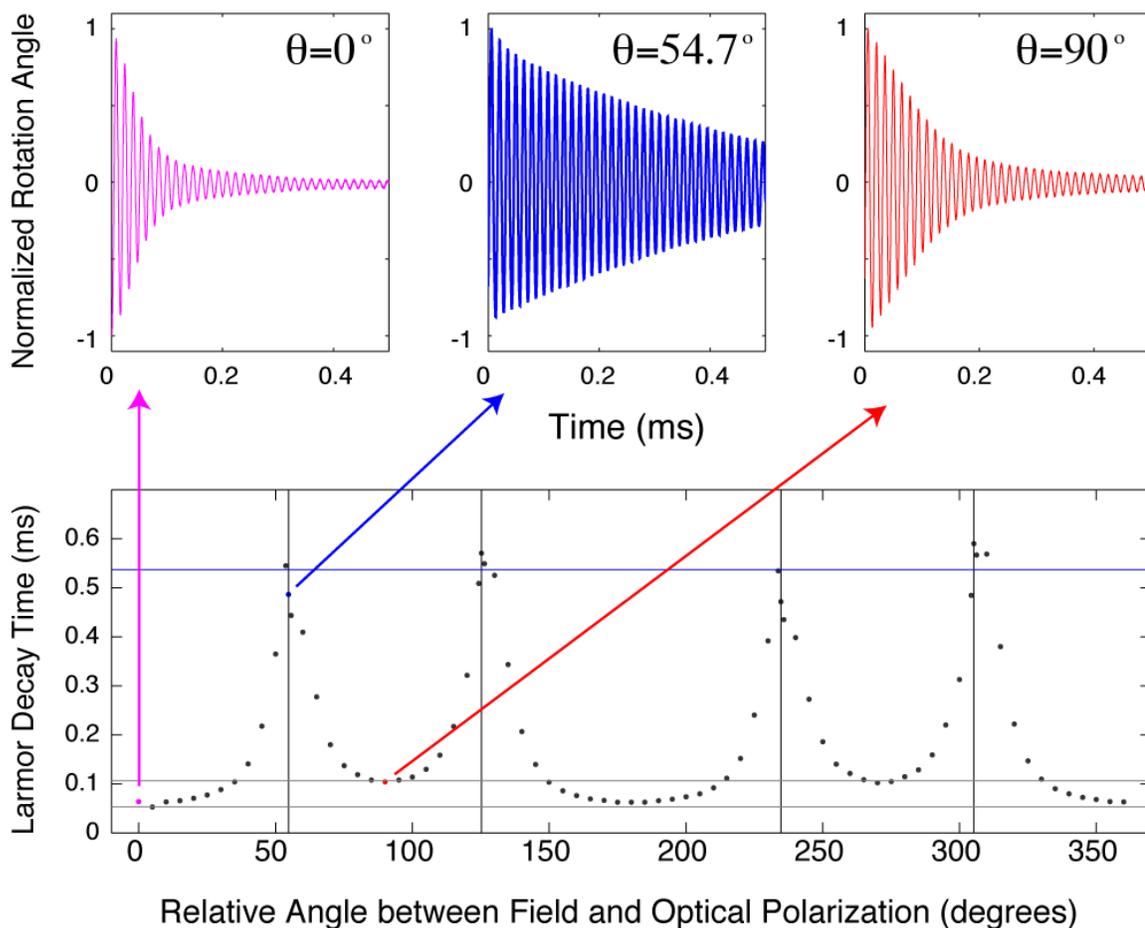
Figure 13.4: Measurements of the Larmor precession decay time as a function of the relative angle between field and optical polarization shown in figure 13.2A. At the magic angle of 54.7 degrees, the tensor dephasing vanishes and the decay is due to only the spontaneous emission. Here we used a probe beam of power 5 mW, radius 1.3 mm, and detuning 815 MHz with a high power shotnoise limited detector. The DC magnetic field was 70 kHz/γ_{B4} and 0.2 ms of each trace was used to obtain the fit giving the decay time constant.

rates. To get rid of these technical nonidealities, one could imagine either making the probe more homogenous or performing a reversal along the lines of usual spin-echo techniques, although we have not spent much time on either front.

In any case, we can observe the dependence of the initial decay time on the relative angle of the field and optical polarization. We expect from section 6.5 that the decay time should be smaller than that due to spontaneous emission alone for all angles other than the so-called "magic angle" of $\theta = 90 - \arctan(1/\sqrt{2}) = 54.7$. Furthermore, it is seen that the Hamiltonian (hence the decay strength) in equation (6.92) should be a factor of two different for relative angles of 0 and 90 degrees. This behavior was also mentioned earlier in respect to figure 5.1.

Figure 13.4 shows the fit decay time as a function of the relative angle between the field and the optical polarization. In practice we fix the input optical polarization along x, adiabatically rotate the spin-state to be aligned along z, and apply the field in the appropriate direction in the xy-plane while turning on the probe light. The factor of two difference in decay times is clearly seen for angles of 0 and 90 degrees. To demonstrate the full symmetry, the field angle is varied completely from 0 to 360 degrees, displaying all four magic angle points. Under different probe imaging configurations one can observe echo like dynamics in the non-magic-angle inset time traces, whereas here we only see a leveling off of the Larmor precession indicating a inhomogenously mixed version of the echo. While we have observed echo dynamics, the results of [25] pay more attention to imaging and intensity homogeneity, thus more cleanly observe the shape of the echos shown in figure 5.1. We have also observed evidence of the nonrotating frame DC echos displayed in figure 5.2A.

13.3 Decay Data

In this section, we detail attempts at characterizing the decay of the atomic state due to the measurement process itself (spontaneous emission) and more technical sources. We present two methods for measuring the decay, both of which avoid the tensor Hamiltonian induced dephasing mentioned in the previous section. In the first method, we simply measure the decay of the Larmor precession with the probe polarization placed at the magic angle (as a function of probe beam parameters). In the second, we investigate the decay of the

state while it is held with a parallel magnetic field, which gives a Hamiltonian that is large compared to the tensor probe Hamiltonian.

In section 5.2.2, we calculated the expected decay of a spin-state due to spontaneous emission alone. There we showed that, as expected, the decay rate depends upon the relative orientation of the probe polarization to the spin-state alignment. This results from the difference between Clebsch-Gordan coefficients that determine the decay rate under different orientations. In the subsequent data interpretation, we reference two representative rates of decay, one for a polarized spin-state parallel to the optical polarization, and one for the spin-state perpendicular to the optical polarization.

Figure 13.5 shows the measurement of the decay of the Larmor precession at the magic angle as a function of both probe power and probe detuning (as seen in the middle inset of figure 13.4). For a high power (or small detuning), the spontaneous emission loss is dominant and the decay rate follows the expected scaling in detuning (or power). During Larmor precession, the spin-state is alternately perpendicular and parallel to the probe polarization hence we expect the overall decay rate to be intermediate between the expected rates for perpendicular and parallel polarizations alone, which is indeed where the data fall. The value at which the rate levels off is likely given by many sources, among them the diffusion of the atoms out of the beam due to temperature and gravity, magnetic field inhomogeneities, and residual tensor effects due to slight geometrical misalignment.

The second way in which the spontaneous emission scaling has been investigated is by observing the decay of the spin-state while it is held stationary with a parallel magnetic holding field. This situation was previously simulated in figure 5.2. Figure 13.6A shows experimental results with three different scenarios. In the middle (blue) curve, the coherent spin-state is first prepared via pumping along x, then it is adiabatically rotated to the z-direction, then a probe beam with a linear polarization along x is applied and used to measure the z-component of the spin. Thus this curve represents "perpendicular" decay and results from one single measurement. In contrast, the lower (red) curve represents a situation in which the coherent spin-state is first prepared via pumping along x, the spin-state held along x, the probe beam is turned on and off for the time shown, the spin is adiabatically rotated to the z-direction, and the probe beam is turned on again to measure the resulting decay from the previous pulse. In this curve, which represents "parallel" decay, each point represents a distinct trial. Finally, the upper (black) curve is obtained from the
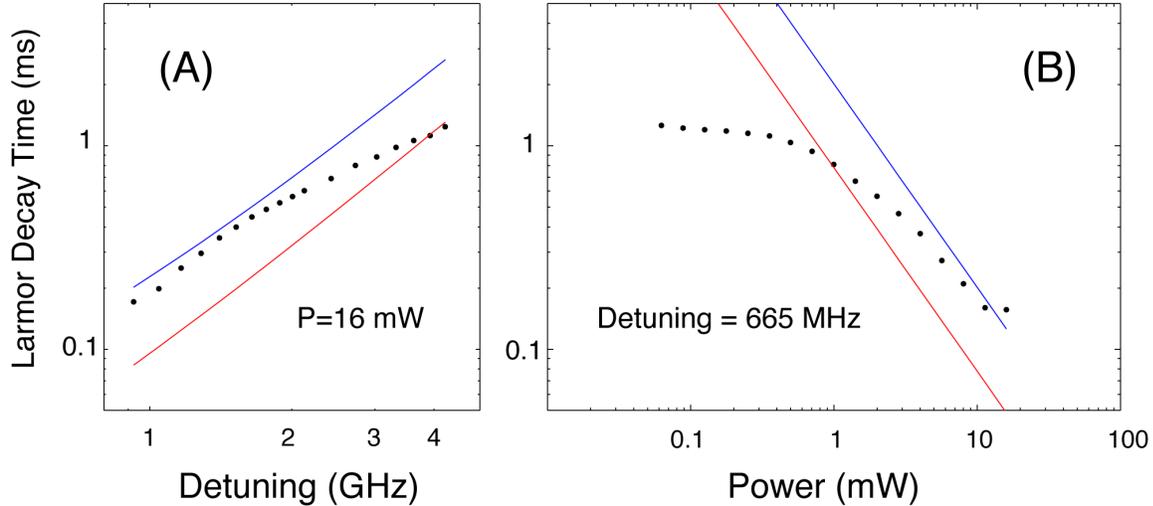
Figure 13.5: The decay rate of Larmor precession (for a perpendicular field of 30 kHz/γ_{B4}) with probe polarization at the magic angle with respect to the field. (A) At a fixed probe power of 16 mW (and radius 1.3 mm) for various detunings. At small detuning the spontaneous emission dominates and the decay time scales as Δ^2. (B) At a fixed probe detuning of 665 MHz, for various powers. At large powers the decay time scales as $1/P$. In both plots, we show the expected decay rates for parallel (lower, red) and perpendicular (upper, blue) spins.

previous procedure, but without the first pulse of probe light. Thus this curve represents the nonspontaneous emission decay and each point represents a distinct trial. (Actually, to avoid the complication of intertrial atom number fluctuations, each procedure also included a first rotation up to z and a very short measurement of the initial vector by which the subsequent curve was normalized.) We see that the upper curve shows us the timescale at which non-probe-induced decay takes place, the middle curve includes additional spontaneous emission decay, and the lower curve represents even more spontaneous emission decay (as predicted in section 5.2.2).

Figure 13.6B shows the exponential timescale of (perpendicular orientation) decay for a fixed probe power at different detunings (each point from a fit of the middle curve in figure 13.6A). First we notice that for smaller detunings ($\Delta < 2$ GHz) the scaling of the decay timescale is as expected from the predicted blue line, even showing the elbow at $\Delta \approx 1$ GHz resulting from a consideration of all hyperfine levels. Notably, however, there is a difference of around a factor of three in the absolute value of the prediction (from an independent measurement of the probe beam intensity) and the experimental result. We attribute this
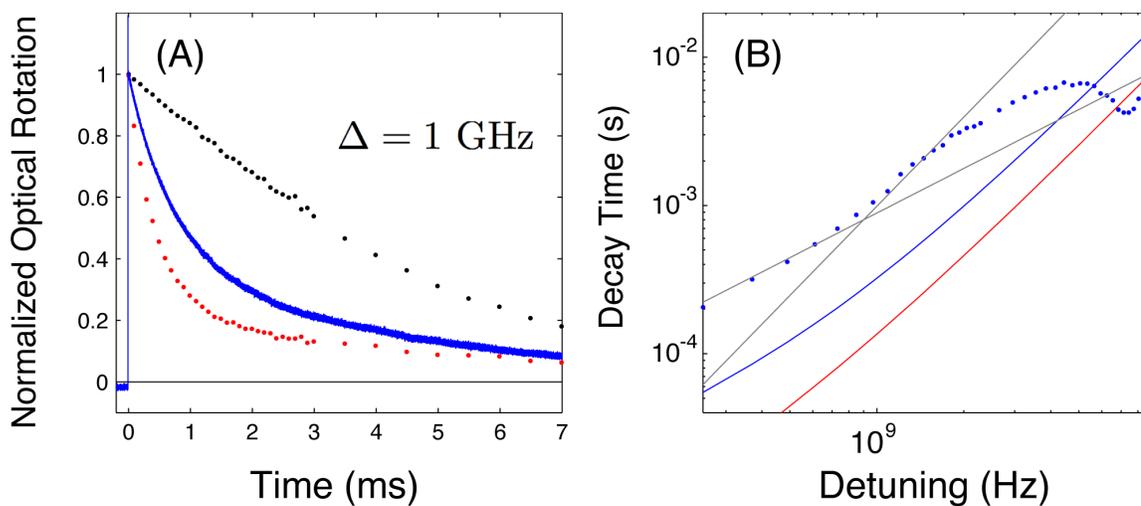
Figure 13.6: The measurement of spontaneous emission decay with a magnetic holding field present. The measurements were performed as discussed in the text, with a holding field of 110 kHz/γ_{B4}, a probe beam power of 15 mW and radius 1.5 mm, and the high-power polarimeter detector. For all plots a 1.5 ND filter is used to avoid saturation of the high power detector. (A) At a fixed detuning of 1 GHz, the decay of the spin-state is measured: without probe induced spontaneous emission (top, black), with perpendicular polarization decay (middle, blue), and with parallel polarization decay (bottom, red). (B) The decay of the perpendicular decay example measured as a function of detuning. The light gray curves represent the scalings of Δ and Δ^2 for reference, while the predicted values for perpendicular and parallel decay are below.
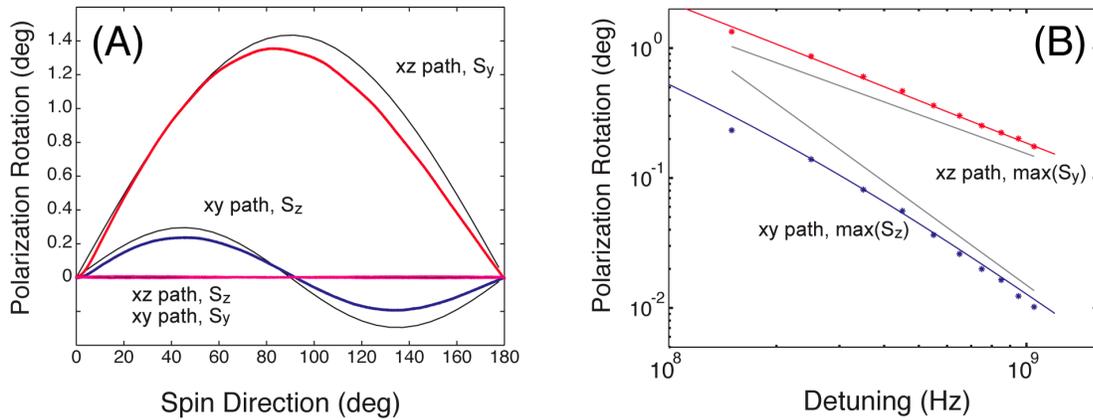
Figure 13.7: Measurements of the semiclassical effect of the spin direction (adiabatically held with a 110 kHz/γ_{B4} holding field) on the polarization of the probe beam. In all measurements the probe beam has power 10 μW and radius 1.5 mm. Each trajectory is averaged 10 times. (A) For an input probe beam with x-polarization and a detuning of 150 MHz, $\hat{S}_y$ and $\hat{S}_z$ were measured for both the xz and xy trajectories (described in the text) resulting in the solid curves. All trajectory times are $\tau = 2$ ms, during which we observe some atomic decoherence that causes the prediction (dotted curves) to stray from the data. (B) As a function of probe detuning, we plot the peak of the $\hat{S}_y$ measurement (for the xz trajectory), which depends only on rank-1 terms, and the peak of the $\hat{S}_z$ measurement (for the xy trajectory), which depends only on rank-2 terms. The predicted behavior (solid curves) shows good agreement with the data out to large detunings where the curves asymptote to the $1/\Delta$ and $1/\Delta^2$ lines provided to guide the eye.

difference to probe inhomogeneity effects. At detunings in the range of 2 GHz $< \Delta <$ 6 GHz the decay time increases but at a slower rate and eventually saturates near 8 ms. For larger detunings closer to the 9 GHz, we begin to measure the atoms that have been depumped into the $f = 3$ ground state. While the fits are poor in this region, the feature reminds us that this complication exists and constrains where the probe detuning should be placed in further experiments. Given that the power (and intensity) setting is near maximal for these plots and that we want to be operating in a spontaneous emission dominated regime, the optimal detuning is near a few GHz. A smaller detuning would increase the relative size of the tensor contribution and complicate the QND description of any quantum limited experiment.

13.4 Dragging Data

In this section we summarize the basic result of [29] where the tensor Hamiltonian dynamics were investigated by adiabatically dragging the spin coherent state to a particular position with a magnetic holding field and observing the semiclassical effect on the polarization of the probe light. The theory of this procedure was previously discussed in section 6.4. In figure 13.7A we show the measurement record for two different paths and two different measured Stokes variables. In the first, the spin-state is dragged in the xz-plane and in the second it is dragged in the xy-plane perpendicular to the probe propagation direction. The input light for all cases is polarized along x. As expected, the measurement of $\hat{S}_y$ for the xz-path gave the largest optical rotation due mainly to the vector term, while the measurement of $\hat{S}_z$ for the xy-path gave a smaller but measurable oscillation at twice the frequency due to the tensor term. The other terms ($\hat{S}_z$ for xz-path and $\hat{S}_y$ for xy-path) are quadratic in the overall rotation vector (which is small) and their effect can barely be measured.

Figure 13.7B shows the value of the peaks of figure 13.7A as a function of detuning. As expected from chapter 6, the vector term falls off as $1/\Delta$ and the tensor term falls off as $1/\Delta^2$ for detunings larger than the hyperfine splitting. More complete spectra for the vector and tensor terms are presented in section 13.5. The technique used here can also be used to measure the pumping efficiency of the coherent spin-state as we show in section 13.6.

13.5 Dispersion Spectrum

In section 6.3, we displayed and interpreted the spectra of the prefactors of the tensor Hamiltonian. The zeros of the coefficients at detunings comparable to the hyperfine splittings were also calculated. The plot of figure 13.7B displays a limited version of this spectrum and now we generalize to measure the spectrum over the entire hyperfine frequency range. The spectra are measured with a single shot of an experimental run with the following procedure. After optical pumping along x, the coherent spin-state is adiabatically rotated with a magnetic field to a particular position and held at that position. A weak probe beam is then turned on and the frequency swept over the hyperfine manifold. This sweep is performed quickly compared to the overall probe-induced decay time.

For measuring the vector spectrum, the input light is polarized along x, the spins are
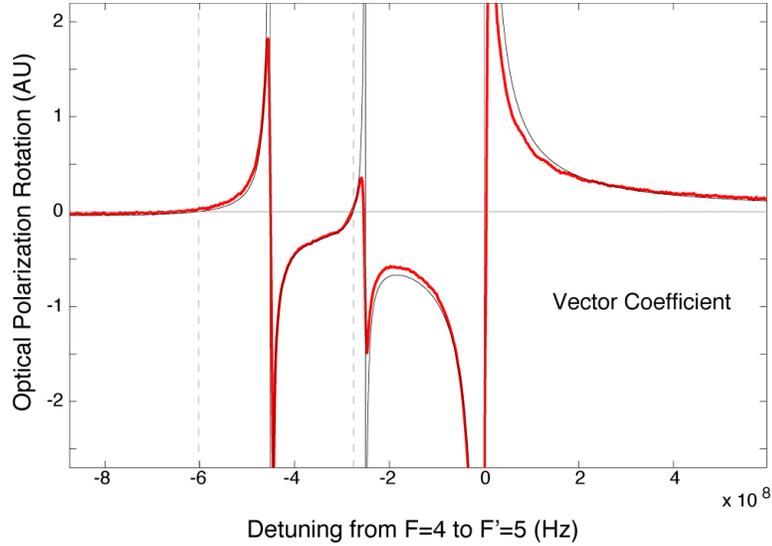
Figure 13.8: Vector Hamiltonian coefficient measured as described in the text in real time. The frequency was swept over 0.5 ms to give the trace shown at fixed spin and optical polarization positions. The probe beam had power of 15 μW, and radius 1.1 mm and a low-power polarimeter detector was used and detected with the high-power detector. The holding field was 70 kHz/γ_{B4}. The optical depth of the sample was approximately 10.

aligned along the z-axis, and the quantity $\hat{S}_y$ is measured while the frequency is swept. This corresponds to the peak of the larger curve in figure 13.7A. The full vector spectrum is shown in figure 13.8. For measuring the tensor spectrum, the input light is polarized along x, the spins are aligned along the $(\mathbf{e}_x + \mathbf{e}_y)/\sqrt{2}$ direction, and the quantity $\hat{S}_z$ is measured while the frequency is swept. This corresponds to the peak of the smaller curve in figure 13.7A. The full tensor spectrum is shown in figure 13.9.

The only free parameters in the fit of the theory to the experiment were the amplitude (constant prefactor) and the frequency offset and slope (versus time) of the theory curve. The fit is surprisingly good given the hasty nature of the measurement procedure. The fit strays from the experiment near resonance primarily because of the limited bandwidth of the detector (approximately 1 MHz). Note that the measurement also reproduces all of the appropriate signs and zero-crossings of both the vector and tensor terms.
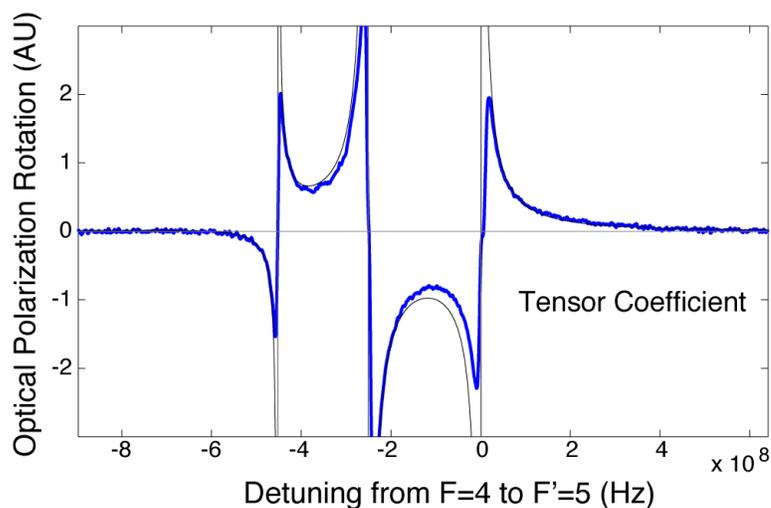
Figure 13.9: Tensor Hamiltonian coefficient spectrum measured as described in the text in real time. The frequency was swept over 0.5 ms to give the trace shown at fixed spin and optical polarization positions. The probe beam had power of 5 μW, and radius 1.1 mm and a low-power polarimeter detector was used and detected with the high-power detector. The holding field was 112 kHz/γ_{B4}. The optical depth of the sample was approximately 10.
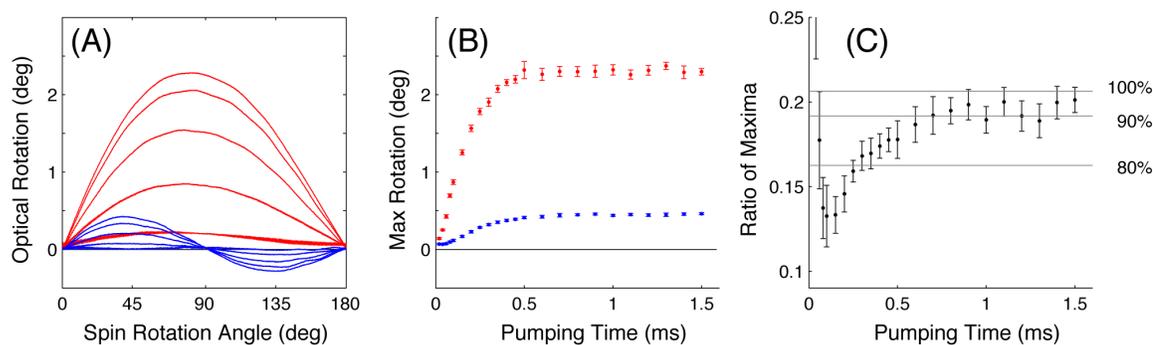
Figure 13.10: A measurement of pumping efficiency versus pumping time as described in the text. The probe beam had power 5 μW, detuning 150 MHz, and radius 1.1 mm and the high-power detector was used. The pump beam had a power of 45 μW and a radius of several mm. The DC holding field was set at 112 kHz/γ_{B4}, the optical depth was around 10 and each point was measured 10 times. (A) The curves of figure 13.7A for increasing pumping time. (B) The peak values of those curves. (C) The ratio, which can be mapped to the pumping efficiency via theory.

13.6   Pumping Measurement

Despite the fact that the tensor Hamiltonian complicates the simple QND model of our experiment, it is also useful for many practical tasks. Here we demonstrate the use of it for estimating the efficiency of the optical pumping (towards a coherent spin-state) preparation process used in our experiment. In the next chapter, we will also see that the the tensor term allows us to fine-tune the geometric alignment of the experiment (as can also be seen from the result of section 13.2).

In section 6.8 we discussed how to potentially measure the pumping efficiency of a coherent spin-state by measuring the ratio of the two curves in figure 13.7A. When the probe parameters are known and the tensor coefficients are used the ratio gives the pumping efficiency in a way that is independent of many experimental parameters, including the beam size, mode matching issues or atom number. Figure 13.10A displays the same measurement as figure 13.7A, but for different pumping times. Figure 13.10B then shows the peak value of these curves as a function of pumping time. (A model for how these values increase given the pump beam parameters can be found in section 12.5.9.) The ratio of these two curves is then taken and plotted as a function of pumping time in figure 13.10C and converted to a pumping efficiency using the thermal model discussed in section 6.8 (although at the pumping efficiencies we observe, this mapping is largely model independent). We see that, as expected, the ratio monotonically increases with pumping time saturating at a level that corresponds (via the theory) to a value between 90% and 100%.

To extract the pumping efficiency using the technique described above one needs to do two separate measurements. However, it is a trivial matter to make a pumping efficiency measurement in a *single shot* by dragging the spins in a combination of the paths in the xy- and xz-planes while measuring an optical polarization variable such as $\hat{S}_z + \hat{S}_y$ with the polarimeter and taking a ratio of two points in one trajectory. This still requires that the probe beam is sufficiently weak (to avoid too much decay) and the rotation is sufficiently small (such that the measurement is linear in the angle). One could presumably also do an analysis that includes decay and optical shotnoise to determine the single-shot measurement fidelity one could possibly attain. However, here we were only interested in ensuring our pumping efficiency at the level of 10%, which is clearly possible.

13.7 Magnetic Resonance Data

In the next chapter, we discuss an alternate means of observing spin-squeezing with a modulation scheme. This corresponds to a magnetic-resonance scenario, which is as follows. As before, the atomic cloud is optically pumped along the x-direction with a magnetic holding field. The pump beam is then turned off but the magnetic holding field along x is kept on. The probe beam is then turned on (to measure $\hat{F}_z$) and at the same time an oscillating field along y is applied. The frequency of the field is chosen to be near or at resonance with the Larmor frequency given by the holding field. The spin dynamics given this scenario are analyzed in full detail in appendix D.

A typical magnetic resonance measurement is presented in figure 13.11A with parameters given in the caption. As expected the measurement signal is a modulated sinusoid of the Larmor frequency, with the beat frequency given by the magnitude of the small, resonant perpendicular field. In practice the cancellation is not complete because of field inhomogeneities that cause the signal to be not completely resonant with every atom. Figure 13.11B shows the measured signal for a perpendicular field sinusoid of a larger amplitude (but equal resonant frequency) than the case of figure 13.11C.

In the short-time limit of figure 13.11C, shown in figure 13.11D, notice that the signal is simply a sinusoid of linearly increasing amplitude masked by optical shotnoise. Thus when performing magnetometry at small times, one needs to derive the field from the slope of the sinusoidal amplitude rather than just the slope as in the DC case of figure 13.2D. To perform quantum limited magnetometry, discussed in chapter 10 for the DC case, one should use stroboscopic measurement as discussed in the next chapter.

In this experimental scenario, without a perpendicular magnetic field applied, one can vary the holding field to observe the ambient field fluctuations at particular frequencies (using an FFT of the signal to identify frequencies in the neighborhood of the DC field frequency). The sources of these high-frequency fields are further discussed in appendix A.
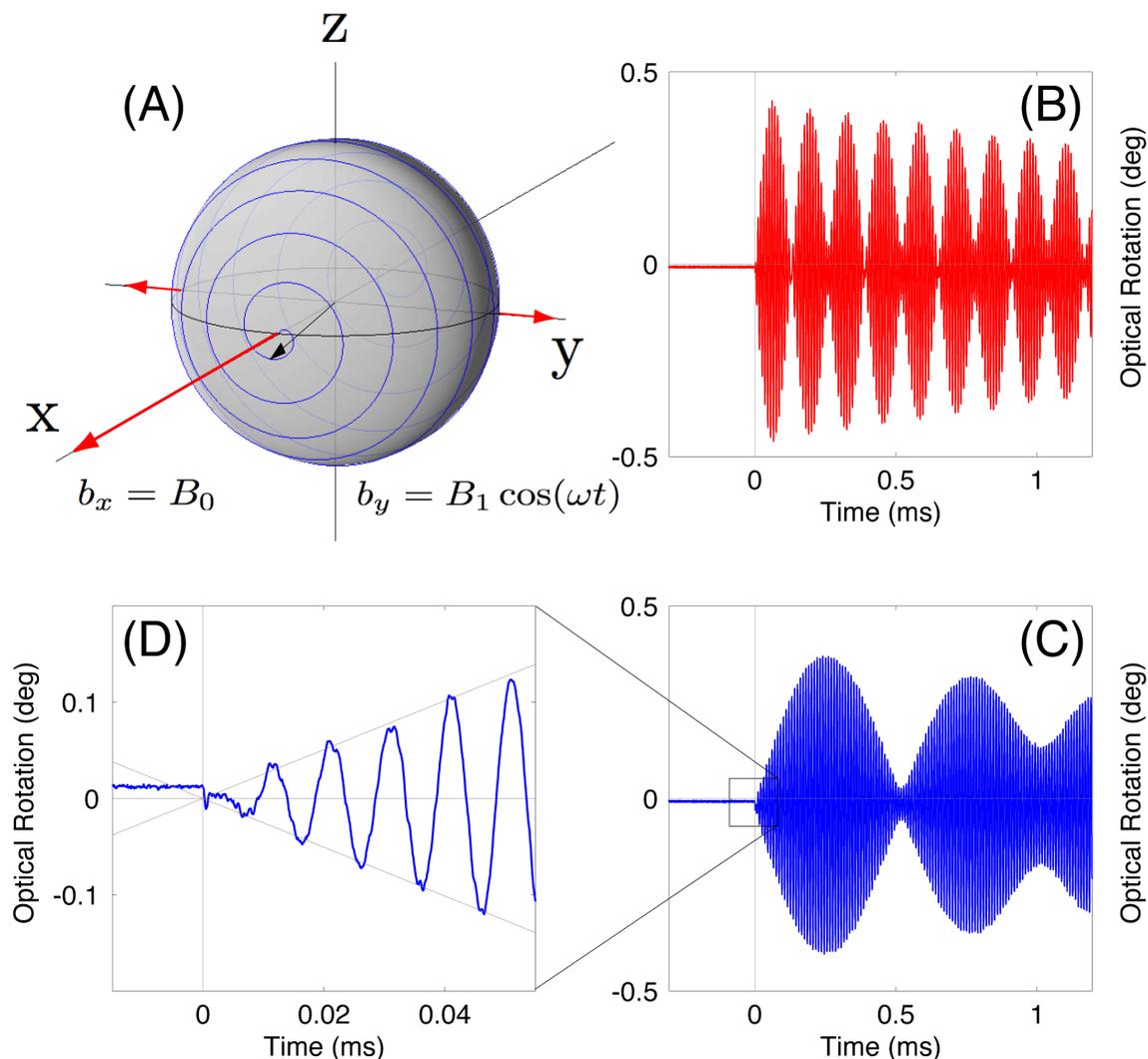
Figure 13.11: A continuously observed magnetic resonance signal. (A) The geometry of the measurement. A large field (100 kHz/γ_{B4}) is applied along x, where the spin-state begins. A small field, resonant with the large field, and of varying amplitude is applied along y causing the spiraling motion of the spin-state. (B) The small resonant field has an amplitude leading to a 4 kHz beatnote frequency. (C) The small resonant field has a smaller amplitude leading to a 1 kHz beatnote frequency. (D) A zoom in of the previous figure, showing the sine wave with linearly increasing amplitude. For all of these measurements, the probe beam was of power 50 μW, radius 1.1 mm, and detuning 1 GHz and the low-power detector was used. The signal was amplified and filtered with a low-pass at 1 MHz. The optical depth was around 20.

Chapter 14

Quantum Data

In the previous chapter, we presented several experimental measurements that confirmed the single-atom, semiclassical theory of the interaction Hamiltonian between the cold atoms and the optical probe beam. Now we turn the focus to the fully quantum, collective description of the experiment, which accounts for the fact that all atoms simultaneously interact with the same probe beam. As discussed previously, the full description predicts that, for a sufficiently optimized experiment, the collective atomic state should impart projection noise onto the probe beam that, when detected, implies the conditional preparation of collective spin-squeezed states. With the observation of this projection noise being the ultimate goal, this chapter discusses the technical experimental difficulties in this pursuit and proposes various techniques to overcome them.

We begin by reviewing the expected behavior of the projection noise and how it relates to conditional spin-squeezing. Next, to place our work in the context of the larger community, we briefly summarize the particulars of other experiments in the field that have aimed to produce spin-squeezing. The basic technical noise sources and limitations are then introduced including detector and magnetic field fluctuations. Two regimes of the experiment are then considered in contrast. First, we consider the most basic low-frequency ("DC") version of the experiment. Our initial experiments were of this type, and we discuss why those experiments were both flawed and technically limited. Second, we consider a high-frequency modulated ("AC") version of this experiment using a magnetic resonance configuration that in principle removes the primary obstacle of low-frequency detector noise. This technique draws inspiration from "stroboscopic QND" proposals originally introduced in the context of measuring quantum harmonic oscillators below the standard quantum limit.

While, in the end, we fail to prepare well characterized spin-squeezed states, the tech-
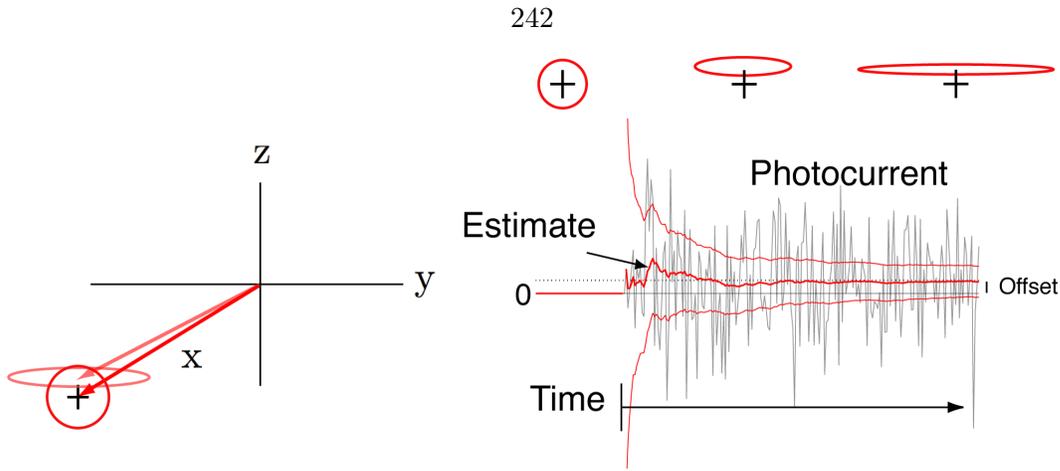
Figure 14.1: Schematic depicting conditional preparation of spin-squeezed states due to filtering of noisy measurement record (DC Scheme).

niques developed here will hopefully play a role in future work that does prepare correlated atomic states via measurement for use in metrology applications.

14.1 Ideal Shot Noise and Averaging (DC)

Here we summarize the ideal scenario for producing spin-squeezed states via the dispersive interaction with and QND measurement of a polarized probe beam. The basic theory of this concept was analyzed in detail in chapter 9 where the (experimentally difficult) long-time limit was also considered. And in chapter 8 we related this idealized theory to our physical Faraday rotation experiment with multilevel cesium atoms, deriving the relevant signal-to-noise ratio that predicts the expected degree of squeezing.

We again refer to the standard configuration of figure 14.1, where the spin-state is initially aligned along x. The probe beam propagates along the z-axis, linearly polarized along x. The quantity $\hat{S}_y$ is then measured with the polarimeter. As seen in figure 14.1 we expect the statistics to be such that white optical shotnoise masks a constant offset (atomic projection noise) on every trial. This offset is nominally constant within a trial (neglecting decay), but between trials the offsets are random and the standard deviation of this projection noise distribution is proportional to the number of atoms measured. Conditional squeezing amounts to no more than averaging away the optical shotnoise to reveal the underlying constant with increasing certainty. Thus, the observation of the projection noise for a single sample under this idealized scenario is equivalent to the preparation of a

conditional spin-squeezed state.

From chapter 8, the degree of squeezing expected is

$$W \equiv \frac{\langle \Delta \hat{F}_z^2 \rangle_t}{\langle \Delta \hat{F}_z^2 \rangle_0} \tag{14.1}$$

$$= \frac{1}{1 + \text{SNR}^2}. \tag{14.2}$$

For a coherent state of N spin $f = 4$ atoms, the signal-to-noise ratio is given as

$$\text{SNR}^2 = \eta \text{OD} \frac{\tau_{sc}}{\tau_{ss}} \tag{14.3}$$

where the optical depth is $\text{OD} = N\sigma_0/A = \rho L \sigma_0$, A is the cross sectional area and L the length of the cloud. The decay timescale τ_{sc} is given by equation (5.46) and the spin-squeezing time is given by equation (8.38) and the ratio discussed in section 8.2.4.

As we have discussed, maximizing the optical depth is critical for achieving a measurable degree of spin-squeezing in free-space. However, the experimental game is not exactly that simple as there are many other technical points to consider. The probe interaction must be strong enough that the spontaneous emission timescale τ_{sc} is dominant relative to the other sources of decay (e.g., atom diffusion). Furthermore, we must ensure that we work in a regime where the tensor interaction (and possibly related propagation effects) does not destroy the simplified description above. Finally, there are sources of optical and magnetic field noise that could potentially obfuscate the observation of projection noise entirely.

Once all of these problems are eliminated and the projection noise identified with certainty, several additional experiments trivially follow, e.g., quantum limited magnetometry (chapter 10) and quantum feedback for deterministic state preparation (chapter 9 and [15]). However, the purpose of this chapter is to not get ahead of ourselves and to focus on the unambiguous identification of the projection noise.

14.2 Status of Experimental Atomic Squeezing

Here we briefly review the status of experimental progress creating large scale atomic entanglement as was touched upon in a larger context in section 1.3.

Certainly the most developed means of entangling a handful of particles has been in ion

trapping groups around the world, e.g., [40]. However, in parallel to these developments, there has been much work in the last decade in entangling a large collection of atoms at once. The idea behind the experiment of this thesis is perhaps one of the simplest for achieving this kind of mass entanglement. Much of the current literature cites as the first demonstration of spin-squeezing the work of [37], which is similar to our effort in some respects but crucially different in many others. That experiment did not prepare a single sample of atoms in one squeezed state. Rather it claims to have measured the (antisqueezing) enhanced projection noise from subsamples of a room-temperature cell full of atoms, which is concurrently pumped and probed. Furthermore, the experiment used a magnetic field to observe this projection noise at high-frequencies but without the stroboscopic considerations discussed below. Thus, the end result of their measurement is a thermally mixed collection of atomic-subsamples, each of which has been measured but in a non-QND manner. In contrast, we aim to prepare a single sample of atoms all of which are collectively squeezed into (ideally) a pure squeezed state as the conditional result of a single-shot measurement pulse.

Also within the past decade, there has been much highly publicized work by the Polzik group where two ensembles are entangled into two-mode squeezed states via a common measurement pulse [36]. This work supposedly avoids the above thermal subsample problem by pumping and probing the samples sequentially in time and measuring all of the atoms with a combination of a large beam and appropriate timescales. More importantly, this work uses a trick that allows them to perform a truly joint variable QND measurement at a high frequency by applying independent and opposite magnetic fields to each cell. (The Polzik group has also reported the preparation of a spin-squeezed state in a cold atomic cloud via the absorption of a squeezed optical beam [39].)

Our work is different from both of the above in that we work with cold atoms and aim to prepare a single squeezed state. Several other groups are similarly interested in measuring the state of a cold atom cloud continuously in time, including the Jessen group [24, 25, 46] and the Takahashi group [43, 44, 173]. However, while both of these groups have continuously observed spin dynamics, as we have presented in chapter 13, they have not yet observed quantum projection noise.

Along with [37] another work that is commonly cited as an experimental demonstration of spin-squeezing is our own work [27] and the following magnetometry paper [28]. Below, in particular section 14.4, we explain why this work is flawed and discuss subsequent

improvements.

Several groups are continuing this work currently. The Polzik and Jessen groups are extending this scheme to the clock transition [45, 46], while the Kasevich group is developing the means to perform these measurements in a cavity [47] where much more squeezing is possible.

14.3   Technical Issues (DC and AC)

The idealized quantum measurement scenario presented above is complicated by several technical issues and constraints. Here we go over the various problems and potential solutions before presenting experimental attempts to observe atomic projection noise.

14.3.1   Polarimeter Noise

In the simplified measurement scenario above, the quantum projection noise on a given trial is represented by a random DC offset. Clearly, if there is low frequency technical noise on the balanced polarimeter detector of a nonatomic origin, this noise will mimic and hence obscure any underlying quantum fluctuation. Indeed such low-frequency noise (e.g., $1/f$ noise) is prevalent in all of experimental science, hence the prevalent use of high-frequency modulation schemes, lock-in amplifiers, etc. Unfortunately, in some quantum limited measurement scenarios, broadband detection and control is necessary, for example in the "adaptive phase measurement" experiment of [12]. In this section we identify the sources of this detection noise in the polarimeter detectors described in section 11.6 and possible ways to avoid this technical problem.

As mentioned previously, we have two sets of balanced polarimeter detectors in use, one of which is capable of operating at relatively high optical powers. The currents from each of the two photodiodes in a detector are subtracted prior to a transimpedance amplifier such that in unbalanced configuration the detector can only take so much power before the amplifier rails, whereas in nearly balanced configuration (with nearly equal powers on each detector) the nominal signal is zero and the detector can take much more power. As the powers are increased in balanced configuration the response of the detector eventually becomes nonlinear due to a variety of effects including heating and diode current supply issues. Also technical noise fluctuations on the polarimeter become so large that they

imbalance and rail the detector. Our detectors are optically shotnoise limited over their broad bandwidth between two powers. On the low end, the not-quite-white electronic noise floor sets a bound, and at a particular noise equivalent power (NEP) the optical shotnoise equals this native noise. On the high end, the detector nonlinearities set in and the noise floor is no longer due to primarily the optical shotnoise. In between these powers, the white noise floor from detection of the probe beam scales as the square root of the power as previously discussed. For our "low-power" detector, the NEP= 1 μW and the detector becomes nonlinear around a mW (bandwidth is DC to 1 MHz). For our "high-power" detector, with significantly larger photo-diodes, the NEP= 1 mW and it becomes nonlinear around 30 mW (the bandwidth is 0 to 5 MHz).

To balance the detector properly, one should do more than zero the DC output of the detector because amplifiers and other effects can add a DC offset. Therefore the best way to zero the signal is to sinusoidally modulate the power of the probe beam (at relatively low frequencies) and align optical elements until the amplitude is minimal. The balancing is typically performed by simply tweaking the waveplate prior to the analyzer beam-splitter. One could also balance by changing the effective efficiency of one of the detectors (e.g., by adjusting a circuit element within the detector), but this should not be done by adjusting the alignment of a beam on one of the photodiodes as this will possibly exacerbate pointing noise fluctuations. It should be noted that we are interested in relatively low-frequency balancing (less than the MHz bandwidth of our detectors), thus we do not necessarily need phase-matching elements between the photodiodes within the detector (as are used, for example, in the high frequency heterodyne detectors of the Kimble lab optical squeezing experiments).

Even in an aligned and balanced configuration, it becomes clear from the amplified (nominally zero) output that there is a small, but significant amount of low-frequency noise on the signal, which, as mentioned above, can obscure the desired quantum fluctuation signal. For typical atom numbers in the experiment, this amounts to resolving the polarization rotation to less 10^{-5} degrees and the technical noise is generally much larger. The amount of experimentally observed low-frequency noise is indicated in figure 14.3 and corresponds to a fluctuation of order 10^{-2} degrees of optical rotation. The details of this figure are discussed further later where this noise is compared to the superior high-frequency noise levels.

The low-frequency noise offset noise is potentially caused by various effects. For smaller photodiode detectors, one must ensure that the diode is detecting all of the light and slight pointing fluctuations will not affect the balancing significantly. Even for larger diode detectors, there may be native polarization fluctuations or pointing noise that gets translated into polarization fluctuations. This is not entirely surprising as the beam passes through many elements (reflective and transmissive) that can potentially cause polarization and/or pointing noise over the long timescales observed (greater than 0.1 s) via mechanical and thermal effects. Slow mechanical fluctuations of the vacuum apparatus (with non-AR coated interfaces) and optical elements are one possible source, which could be improved with better mechanical stabilization (passive or active). Air-current density fluctuations are also a possible source of noise, which could be improved by surrounding the apparatus in a controlled gas environment. The expected order-of-magnitude of this effect has not been calculated by us, but other labs have reason to believe it is important for sensitive detection. Pointing noise can also be caused by the AOM, which serves as a shutter for the light beam (with the probe beam coming from the first-order beam). With all of these suspicions, I attempted to reduce the measured fluctuations with a test setup where the beam was sent through an optical fiber after the AOM and immediately into the polarimeter with minimal intervening optics. Although this effort was not exhaustive, improvements in the offset noise were marginal at best.

Signal noise can also come from frequency noise on the probe light translated into amplitude and/or polarization noise on the post cell beam by way of interacting (by absorption or phase shift) with the room temperature thermal distribution of atoms and slightly imperfect optical balancing. This effect is much worse at the maximally sloped region of the absorption profile at several hundred MHz detuning. In practice, the New Focus Vortex probe laser has frequency fluctuation of a few MHz in depth at near kHz modulation rate and the problem obviously becomes worse closer to resonance. One could minimize this sort of problem (and speed up the MOT loading time) with nonthermal, beam loading of the MOT (e.g., with a separate two-dimensional MOT), but this technical improvement has been put off for subsequent generations of the experiment.

Other possible sources of noise include imbalanced detection of background light (although this is not likely given attempts at shrouding the detectors) and also an unidentified source of low frequency noise from the detector electronics present only when the photodi-

odes are generating current.

We point out that, depending on our goal, we do not need to completely remove technical offset noise because it is possible to observe projection noise on top of the technical noise. In this case however, the smaller the relative difference between the size of the noise levels, the more data are needed to resolve them. And if the technical noise is larger than the expected quantum projection noise, this means that we may be able to observe the projection noise but *not* prepare conditional spin-squeezed states due to the ignorance imposed by the technical offset.

Now we turn to possible solutions for eliminating or avoiding the technical low-frequency detection noise. First, because of the fact that some technical noise becomes worse with increasing power, one might be inclined to work at lower powers. However, the experiment demands that the overall measurement time should be approximately the spontaneous emission decay time and to keep this less than technical atomic decay timescales, the probe power needs to be sufficiently large. Alternately, one could decrease the power and decrease the detuning, but then the tensor Hamiltonian complications come into play, as discussed below. As mentioned in section 8.1.4, it is possible to do some detection tricks whereby the overall power detected is decreased (and polarization rotation "magnified") at no cost to the signal-to-noise ratio. However, when using this technique, the electronic noise floor becomes effectively larger and eventually the detection is no longer shotnoise limited.

Another technique involves subtracting a reference offset signal on every trial, i.e., the atomic and technical offset are measured, the atoms quickly removed, then just the technical offset measured. The atomic signal is then extracted on a per-trial basis from the subtracted signals. As long as the technical noise fluctuations are slow enough, this will be an effective method. However, the reference signal needs to measured and averaged for much longer than the atomic signal such that the subtraction does not add shotnoise related noise to the result. When the atomic measurement time needs to be practically of order a millisecond and the reference time much longer, in practice it becomes difficult to maintain the slow fluctuation assumption this technique requires.

Finally, we discuss the potential for modulation schemes to avoid low-frequency noise sources. As with all modulation schemes one needs to modulate the probe such that only the system of interest responds at that same frequency. However we must also be mindful of whether the desired quantum measurement dynamics are retained under the new

measurement scheme.

As the first modulation scheme, we consider frequency modulation, with inspiration from successful techniques in cavity based spectroscopy (e.g., NICE-OHMS). Suppose in our case that the probe frequency is modulated (e.g., by a depth of several MHz about an overall detuning of a several hundred MHz) at a rate of a few hundred kHz. In reference to the signal prefactor theoretical plot figure 6.2B or its experimental confirmation in figure 13.8, the frequency dither would map to a modulation of the vector coefficient prefactor along some point of the dispersive spectrum. There are a few problems with this technique. First, if we dither about an arbitrary detuning center point where the signal is nonzero, and keep only the sinusoidal dither component, we will be throwing away the DC signal, which contains information about the state, thus significantly reducing the effective quantum detection efficiency of the measurement scheme. In contrast, if we were to dither about the 4–2' point in figure 6.2B where the vector coefficient crosses zero, then the loss of effective efficiency would not be a problem, with the additional benefit that the tensor coefficient is also zero near this detuning (figure 6.2C). Second, the cold atoms may not be the only element along the beam path that responds to the frequency dither. The background gas absorption (along with imperfect balancing), frequency dependent power fluctuations on the probe laser, and etalon effects may all create a noisy background signal at the desired dither frequency.

As the second modulation scheme, we consider the stroboscopic Larmor scheme, discussed in detail below in section 14.5. As opposed to the above scheme, this technique has the advantage that the frequency signal is created with a DC field. As a result, neither the unpumped background gas (nor any other element) will create a response signal at the expected Larmor frequency of the cold atoms. However, to retain the QND nature of the measurement the probe light has to be strobed at twice the Larmor frequency.

Indeed, there be a more clever or practical means of avoiding low frequency noise sources than is discussed in this section. However, we hope to have given the reader some sense of the challenges in detecting quantum fluctuations in from a technically noisy dispersive measurement signal.

14.3.2 Magnetic Field Noise

From figure 13.2D in the constant field case (and figure 13.11D in the high frequency field case) it is clear that the application of an unknown, random magnetic field will also degrade the amount of spin-squeezing or make it altogether impossible depending on the estimation procedure used.

In Table 10.2 we quantify the degree to which spin-squeezing is obscured when performing simultaneous spin and (constant) field estimation. While the field and spin estimation noise decouples after some timescale (from the other's prior), it is generally prohibitive to have too much field noise present when trying to create a conditionally spin-squeezed state, especially when that preparation is marginal to begin with.

In appendix A we go into more detail on the field cancellation required and techniques for managing the magnetic field. Here we summarize only a few points pertinent to spin-squeezing. The larger number of atoms present, the smaller the subtended uncertainty angle is for the variance at the end of a coherent state, thus the smaller the field fluctuations must be to observe squeezing. Strategies that we have employed to reduce field fluctuations include both passive shielding, cursory attempts at active shielding, and line-locking the experiment to the 60 Hz frequency of the lab power supply. One further approach is to increase the measurement rate to be high enough that field does not have time to rotate the state significantly over the course of the measurement. However, this requires more optical power (and detection capabilities) than what we currently have.

An advantage of the high-frequency modulation scheme described below, in addition to the negation of low frequency detector noise, is that the spins are only sensitive to fields with a frequency in the vicinity of the Larmor precession frequency given by the holding field. Thus, not only can one move away from the noisy low-frequency region, but the particular frequency band can be chosen to be one that happens to be relatively quiet. Below we discuss how the high-frequency spectrum is measured with the spins and how a sufficiently quiet working frequency is chosen.

14.3.3 Tensor Dynamics

In section 7.2, we calculate the dynamical effects of the tensor Hamiltonian on the first and second-order moments describing the collective spin-state. From that analysis, it is

clear that close to resonance the tensor dynamics are capable of destroying the simple QND description described earlier, no matter what the relative polarization orientation is. Although the single atom evolution is not influenced by tensor terms for a perfectly aligned system (with probe polarization parallel to the spins, see section 6.7), the collective tensor terms appear for any orientation. Furthermore, slight misalignments of even the single-atom master equation will lead to non-QND effects near resonance. Coupled with technical atom number fluctuations, the tensor Hamiltonian can confuse the interpretation of the data significantly.

There are several strategies available for resolving this issue. With all of them the range of applicability, e.g., for metrology applications, is modified, but the primary aim is producing squeezed states in some basis. First, one could simply work with a truly two-level system such that the tensor dynamics do not exist in the first place, such as the clock transition [45, 46]. Second, the initial state of the system can be manipulated such that the dynamics are effectively two level, as in [112]. Third, the parameters of the system (e.g., spin direction, or probe frequency) could possibly be adjusted in time such that the tensor dynamics (at the quantum level) cancel and only the vector dynamics remain, in analogy to the magic angle technique demonstrated in section 13.2. Such a technique would ideally allow one to take advantage of the better scattering timescale to measurement timescale close to resonance (see section 8.2.4) without worrying about the tensor dynamics.

The final strategy is to simply work at a large enough detuning that the ratio between tensor and vector Hamiltonian components is sufficiently small (see figure 6.4). The primary drawback here is that it takes a significant amount of optical power to keep the measurement time small compared to other decay timescales.

Even for the high-frequency scheme considered below, at relatively large (few GHz) de-tunings, the tensor Hamiltonian introduces dynamics that, combined with atom number fluctuations, leads to noise that masks the desired quantum projection noise signal. Measurements demonstrating the effect (and how it is used to align the system) are presented in section 14.5.2.

14.3.4   Atomic Calibration and Optimization

When attempting to observe the atomic projection noise, it builds confidence in the final results to have as many independent calibrations of the experiment parameters as possible.

In our case, it is critical to know the probed atom number to know where to expect the projection noise level should be. Ideally, the measured projection noise would not only be near the independently calibrated theoretical expectation, but also scale in the appropriate way as the number is increased. Unfortunately, in free-space, when small amounts of squeezing are expected, the dynamic range in the number over which the scaling can be tested will not be great.

In principle the atom number can be calculated from the dispersive measurement of the optical depth (through the height of the Larmor precession trajectory) and knowledge of the beam and atom cloud size. To double check these numbers one can use the same probe beam to measure the absorption induced by the atoms, either at a single frequency or with a frequency sweep as discussed in section 13.1.2 for atoms in the steady-state trap or as they diffusively fall. At the same time, we can measure the resulting fluorescence with a detector off the axis of the initial beam, which should match the in-line absorption measurement by energy conservation. The nice thing about these measurements is that they measure the atoms that the probe beam actually sees. As a final check, one can measure the total number of cold atoms (some of which the probe beam detects, some of which it does not) with fluorescence detection from the MOT beams themselves. This, along with spatial imaging of the MOT cloud, gives an indication of the size and density of the cloud, which combined with knowledge of the probe beam size can be used to get a final estimate for the probed atom number. When comparing all of these techniques, it is important to distinguish measurements that are performed under different conditions (continuously trapped versus free falling, optically pumped or not, etc.).

Given an experimental scenario with constraints (such as a limited optical trapping power budget), the goal is then to maximize the observed number and, more importantly, the optical depth. In chapter 12 and chapter 13, we allude to many of the techniques we have tried. The available options include adjusting the geometry of the trap (e.g., to create a large optical depth cigar shaped cloud), the steady-state trapping technique (e.g., MOT versus dark-SPOT), and post-trapping transient techniques (e.g., cooling and compression via magnetic field gradients).

14.3.5   Trade-offs and Choosing Probe Parameters

In summary, we can use the following procedure to pick the probe parameters for our spin-squeezing experiment, minding various technical trade-offs and constraints. We first choose the measurement timescale we wish to work with. Because of the finite bandwidth of the detectors and the acquisition rate of our computer card (and the bandwidth of potential feedback actuators), we desire the measurement time to be significantly greater than 1 μs. However, because of the timescale of decay is limited to several milliseconds (see figure 13.6) by technical factors such as atomic diffusion we want to keep the measurement time smaller than this so that the measurement is spontaneous emission limited. We also want to keep the measurement time short to avoid the long-term influence of undesired background fields.

Now three parameters go into the measurement timescale (spontaneous emission given by equation (5.46)), the probe detuning, power, and cross sectional area. If our strategy is to avoid the tensor effects, then we desire the detuning to be large ($\Delta \gg 300$ MHz) such that the ratio of tensor to vector coefficients is small (given in figure 6.4). However, we cannot make the detuning very large because for one, we have a limited power budget to keep the timescale sufficiently small, and two, we eventually run into measuring the $f = 3$ state close to 9 GHz if the probe is blue detuned (see figure 3.1).

As for the area of the probe beam, we can consider two options. In the first, we can choose to measure all of the atoms in the cloud at once, making the area of the probe beam just greater than the cross sectional area of the atom cloud (typically of scale mm to cm). This is aesthetically and functionally pleasing as including all atoms ignores the issue of subensemble mixing (by diffusion) and thus the sample stays squeezed for longer after the state preparation. In contrast, one could make the area of the beam much smaller than the cross sectional area of the atom cloud, but still large compared to the wavelength to avoid extreme diffraction issues. This is possibly appealing for the reason that given a fixed power budget we can make the measurement time short by reducing the area to increase the intensity. Furthermore, this decreases the number of atoms used and thus increases the relative size of the atomic projection noise. However, one must be mindful of the atoms diffusing between the measured and nonmeasured regions as discussed in section 12.6. In practice, we typically keep the beam large for this reason.

After setting the desired timescale, detuning, and beam size, the probe power is then determined and is often in the range of many mW. First, one needs to ensure that the resulting intensity and detuning do not violate the weak saturation parameter condition (equation (5.25)) necessary for the adiabatic elimination procedure and dispersive interaction to hold up as a valid approximation. The goal is then to build an optical shotnoise limited detector for this power, with an electronic noise floor (NEP) sufficiently below the desired power optical shotnoise over a bandwidth acceptable for the experiment (e.g., DC to several MHz). Of course, one must also build a laser that can supply this power cleanly, as we have done with the master slave setup described in section 11.6. For power in the range of several mW, large mm scale photodiodes are necessary to ensure a linear response. For larger powers, one could make the diodes larger, but then the capacitance becomes larger and the detector bandwidth begins to suffer.

14.4   Low-Frequency Measurement (DC)

Here we consider the pursuit of the basic nonmodulated spin-squeezing scheme. In the end, the technical complications of our experiment have prohibited the possibility of observing DC spin-squeezing. We explain this disappointing result in the historical context of our previous papers where we temporarily believed otherwise [27, 28]. In the next section, we consider a modification of the experiment to high frequencies to avoid low-frequency sources of noise.

Because we want to avoid non-QND tensor Hamiltonian complications (seen, e.g., in figure 5.2C-D) it makes the most sense to measure the atoms with a probe detuning much larger than several hundred MHz, which requires powers larger than a mW (for mm size beams) to keep the spontaneous emission the dominant source of decay at millisecond timescales. At these large powers, the offset noise fluctuations are simply dominant and much larger than the expected atomic projection noise. Magnetic field fluctuations and residual tensor Hamiltonian induced dynamics are also a large concern, but the primary limiting factor is the low frequency polarimeter offset fluctuations. Using all of the techniques discussed above, I could not reduce this technical noise below the expected atomic projection noise level even at small μW level powers. This, in retrospect, is the simple motive for moving to the modulation scheme of the next section, although this fact has been significantly

obscured by premature conclusions about the capabilities of our experiment.

Historically, our understanding of the experimental limitations and results has been convoluted to say the least. Our initial Larmor measurement results were consistent with the semiclassical results of the last chapter, although we had then misunderstood many effects. At the time, although crude two-level approximations existed, we had not derived the expected multilevel atomic signal and noise levels, and neither did we appreciate the non-QND role of the tensor Hamiltonian at the near-resonance probe detunings we were using. However, we proceeded with the naïve QND model working at an arbitrary set of probe parameters.

As a result of measurements taken by JM Geremia near the end of 2003, we then quickly published the work of [27], which appeared to show projection noise and hence conditional spin-squeezing. Although the claims were somewhat tempered and the work was not entirely calibrated, the paper gave the unmistakable impression that spin-squeezing enhanced by real time quantum measurement control had been achieved. In addition, the work claimed the observation of the antisqueezing side effect of the measurement and magnetic field feedback was used to make the process more deterministic. Following on these results, a paper implementing the scheme to perform magnetic field estimation in the presence of the supposed projection noise was eventually published [28]. Unfortunately, it is now apparent that there are many reasons to doubt the quantum-level claims of these papers, a few of which we mention here.

In the course of the following two years, I resumed control of the experiment and began to work out the detailed theory of the atom-light interaction resulting in much of this thesis. Inspired by the Jessen group observations [24], I eventually confirmed experimentally the predicted tensor dynamics as seen in both [29] and, more completely, the results of the chapter 13.

After attempting to reproduce the squeezing results, it became clear that our previous work [27, 28] can be disputed on many levels. First, regarding the experiment, it appears now that the technical polarimeter noise (discussed above), among other problems, prohibits spin-squeezing at experimentally feasible atomic and probe parameters. Second, ignoring technical noise and comparing the results of [27] to the updated atomic theory alone, it is clear that there are several inconsistencies. Regarding the supplemental information section of [27] there is an error in which the signal-to-noise ratio is taken with respect to

the full scale signal ($\langle \hat{F}_z \rangle \propto N$) as opposed to the coherent state noise ($\sqrt{\langle \Delta \hat{F}_z^2 \rangle} \propto \sqrt{N}$), although the calculated numbers given there seem to reflect the latter adjustment. (The correct derivation of the SNR is given in section 8.2.3.) Despite this error, for the fixed set of probe parameters, the crude two-level theory given in the supplement roughly gets the same order-of-magnitude signal and noise sizes (and decay times) as the updated multilevel theory. (However, the updated theory should obviously be used to get precise results for general probe parameters.) The primary problem with [27] is that the atom number cited therein (10^{11}) is nearly two orders of magnitude larger than what can currently be achieved (around 3×10^9) with our experimental parameters. As discussed in chapter 12, the smaller number is also more in line with theoretical expectations for the trap parameters. If one could possibly achieve this larger atom number then (ignoring technical noise) one could expect an SNR^2 of order unity for averaging times near 10 μs and much greater for the larger spontaneous emission limited timescale (of order milliseconds). Even if this atomic number (and the absence of technical noise) were correct, there is still a large discrepancy between the apparent observed squeezing (several dB) and the smaller amount expected from theory for the short measurement times used. Also, the averaging procedure of [27], only calculates statistical correlations using 10 μs windows even though the probe beam is on prior to these windows for 40 μs. In other words, not all of the measurement record available is used in the statistical analysis. Thus one expects much more antisqueezing than was observed and the squeezed/antisqueezed uncertainty product should have been larger.

The anomalously large atom number of 10^{11} cited in [27] implies an on-resonance optical depth of 8700, which is much larger than the parameters achieved in more well-equipped experiments (see section 12.4) and would cause a full-scale Larmor precession signal that completely rotated the optical rotation by thousands of degrees at typical probe detunings. I have never observed either absorption or dispersion calibrations indicating this atom number and the typical results I have seen are two orders of magnitude smaller as shown in chapter 13. Even if the trap did contain this large number of atoms, independently observed magnetic field noise would randomly rotate the spin by much greater than the coherent state projection noise angle during the measurement time, obscuring the quantum noise severely.

Finally, these papers do not address at all the possibility of unconditional evolution of the

uncertainty distributions. In figure 5.2 we see that the single-atom uncertainties can increase due to the tensor dynamics (perpendicular polarizations) and/or the spontaneous emission (parallel polarizations). Furthermore, the results of section 7.2 indicate that the many-atom uncertainties can evolve in entirely nontrivial ways (for both parallel and perpendicular polarizations) due to the non-QND tensor Hamiltonian close to resonance. It is a telling sign of the confusion in the previous work that the dependence of the spin-squeezing on the measurement time and the effect of the decay and tensor effects on the uncertainty distributions are not investigated at all.

If the source of the statistical fluctuations were not true projection noise, then the question remains, "What caused the state-dependent statistics?" I did not take the data, however, it could have been related to a semiclassically driven nonzero signal (from tensor dynamics) coupled with classical noise fluctuations (in the polarimeter offset and/or atom number).

In summary, the work of [27, 28] is flawed because an overestimate of the likely atom number was used, there exist unknown reasons for the measurement statistics matching a naïve theory with that atom number, non-QND tensor Hamiltonian and spontaneous emission effects were not adequately considered, and the expected effects of independently observed technical noise were not accounted for. The claims of [27, 28] were over stated and these papers should not have been published as they are currently written. Due to the litany of problems and confusion surrounding this early work, much of my subsequent time in the lab was spent pursuing the modulation scheme presented in the next section.

14.5 High-Frequency Measurement and Oscillator Dynamics (AC)

To avoid the pervasive technical noise in the spin-squeezing scenario above, we now introduce a scheme to create spin-squeezed states using resonant high-frequency modulation of the spin-state and probe light. In section 13.7 and appendix D, we have already discussed and demonstrated the dynamics of the semiclassical spin-state undergoing evolution in the presence of a large magnetic field, initially parallel to the coherent spin-state direction, as a small resonant time-varying magnetic fields are applied along the perpendicular direction. Now we consider the same basic scheme but without time-varying fields and with a focus
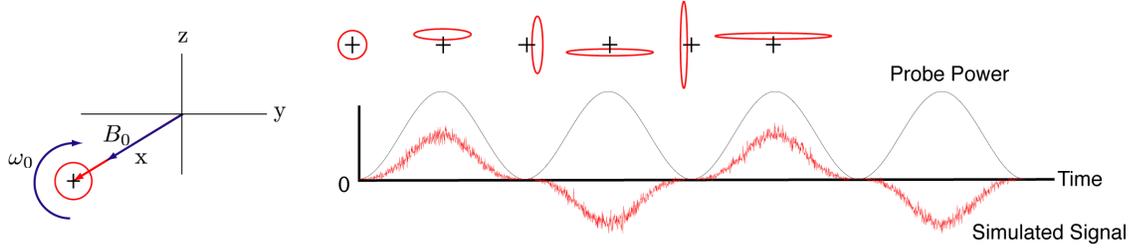
Figure 14.2: The stroboscopic measurement scheme. With the spin-state aligned along x, a parallel constant magnetic field causes the spin uncertainty distribution to oscillate at a frequency ω_0. The probe intensity is modulated sinusoidally at a frequency of $2\omega_0$, such that the spin-state is measured with max intensity when the squeezed axis lines up with the measured axis and with minimum intensity when the antisqueezed axis is along the measured axis. With the probe on continuously the measurement is not QND, but by strobing it the measurement can be made more effectively QND.

on the quantum dynamics of the collective spin-moments in the yz-plane, as shown in figure 14.2.

This configuration approximately "simulates" a harmonic oscillator as follows. The collective spin-state is initially pumped into a coherent spin-state along the x-axis, with possibly some small deviation from the "origin" of $\langle \hat{F}_\text{y} \rangle = 0$, $\langle \hat{F}_\text{z} \rangle = 0$. When a magnetic field B_0 is applied along the x-axis, the small y- and z-components will oscillate into each other at a frequency $\omega_0 = |\gamma_f B_0|$. If we somewhat artificially define the quantum operators for "position" and "momentum" as

$$\hat{X} = \frac{\hat{F}_\text{z}}{\sqrt{\langle \hat{F}_\text{x} \rangle}} \tag{14.4}$$

$$\hat{P} = \frac{\hat{F}_\text{y}}{\sqrt{\langle \hat{F}_\text{x} \rangle}} \tag{14.5}$$

where $\hat{F}_\text{x} \approx F$ is approximated as a classical number, we see that the dynamics and commutator properties of the system are equivalent to that of a harmonic oscillator with natural frequency ω. As long as the measurement and magnetic field dynamics ensure that the collective spin-state remains in the local region around $\langle \hat{F}_\text{x} \rangle = F$, $\langle \hat{F}_\text{y} \rangle = 0$, $\langle \hat{F}_\text{z} \rangle = 0$ then this analogy will hold. Essentially, the states are restricted into a local phase space that does not extend far with respect to the curvature of the sphere, and hence appears locally flat. This Gaussian approximation is discussed in more detail in section 9.7. Now if we

include the measurement of $\hat{X}$ (or $\hat{F}_z$), then we see that the position distribution is initially squeezed some, but then the harmonic oscillator Hamiltonian rotates the antisqueezed momentum distribution into the measurement direction. In other words, the measurement of position is no longer QND because the measured variable no longer commutes with the full Hamiltonian.

The solution is to use *stroboscopic measurement* to squeeze the position over timescales longer than the period of the oscillator. The idea is to only turn on the probe of the position (stroboscopically) at times $t_n = n\pi/\omega_0$ (with n integer), when the squeezed distribution is aligned with the measurement axis. Using a strobed measurement to measure oscillator below the standard quantum limit is an idea that dates back to papers from the 1970s and 1980s in relation with LIGO, where the goal is to measure gravity waves via the displacement of the position of an interferometer mirror [78, 41]. More recently, the idea has been presented in the context of measuring nanomechanical oscillators with continuous quantum measurement [174]. The basic conclusion in these latter papers is that when one expects a larger degree of total squeezing, the duty-cycle of the measurement should be made smaller, i.e., the measurement pulse time should be made smaller relative to the oscillator period. For small amounts of squeezing, a sinusoidal probe power is expected to be adequate.

In the case of the spin-simulation, oscillating perpendicular magnetic fields play the role of harmonic oscillator forces. In appendix D, this is analogy holds true for short times, as demonstrated in figure 13.11D. The analysis of noise in the stroboscopic scheme is also presented theoretically in appendix D.

In the following sections, we consider two distinct modes of measurement. For the first, we consider performing the measurement in an explicitly non-QND way, with a holding field but with a constant probe intensity. In this case, the measurement record is still analyzed at the appropriate high frequency (given by the field), but the probe beam is not strobed. Our strategy is to first use this scheme to eliminate technical sources of noise and reach a regime where projection noise is expected. Given this success, our next goal is to then move on to strobing the probe light and observing projection noise in a more QND manner. It remains an open question what the best way is to observe the projection noise in the constant probe, non-QND scene. As opposed to the QND scenario where the spin effectively assumes a constant offset masked by shotnoise, in this case the spin projection will diffusively wander with competing dynamical timescales given by the measurement rate

and the Larmor frequency.

The rest of this chapter proceeds as follows. In the next section, the advantages of the high-frequency scheme are considered, and experimental noise cancellation demonstrated. We then detail the tensor Hamiltonian induced complications that occur if the system is not very precisely aligned. In section 14.5.3, we demonstrate the high-frequency (but nonstroboscopic) measurement of the collective atomic state and show how close the experiment is to being atomic projection noise limited. Finally, in section 14.5.4, we summarize our progress with stroboscopic measurement.

14.5.1 High-Frequency Advantages and Disadvantages

In this section we consider the experimental details of the stroboscopic detection scheme and discuss its practical advantages and disadvantages. Many of the specifics for the data processing of the stroboscopic signals are left to appendix D.

In practice, the strobing of the probe beam is controlled by modulating an AOM with a beatnote generated by mixing two high quality signals of slightly different frequencies (see section 11.6). The signal generators are synchronized with the rest of the experiment and the probe is modulated at the beatnote frequency. The holding field Larmor frequency is tuned to be half of the beatnote frequency by observing large angle Larmor precession about the holding field and adjusting the magnetic field size to maintain this resonance condition for the lifetime of the atomic signal. We then extract the sine and cosine quadratures (at the Larmor frequency) from each signal trajectory and average the in-sync quadrature to obtain the estimate of the spin amplitude as detailed in appendix D. Noise variance statistics are then calculated and compared to the theoretically expected noise levels from independent calibrations.

First and foremost among the advantages is that the scheme, as designed, renders low frequency fluctuations of the polarimeter output unimportant. This can be seen in the experimental data of figure 14.3, where the high-frequency noise statistics of the probe beam are analyzed without the influence of the atomic contribution. The data was taken with the probe beam on continuously, and the strobing case is considered later. Each measurement in the plot consists of two parts, the continuous running estimate is a spiky line, and the dots at the local minima of those curves are the estimates for times corresponding to an even number of cycles. The even cycle points correspond to minima because the offset noise from
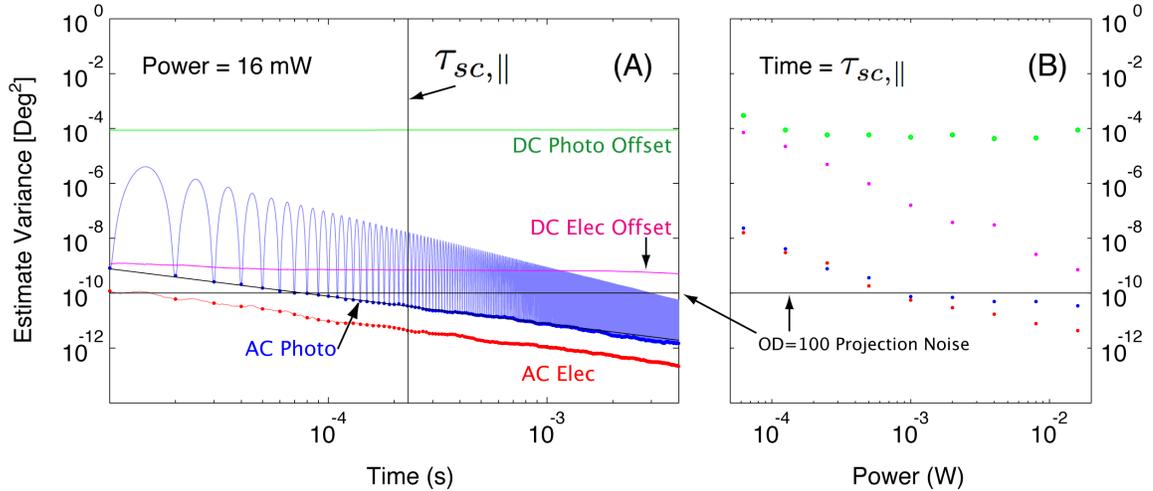
Figure 14.3: Optical noise data without atoms and without strobing. (A) The estimate variance as a function of time for various estimate procedures. The upper (green) flat curve is the DC offset variance from the optical signal (0.01 degree fluctuations). The lower (magenta) flat curve is the DC offset variance for the electronic signal of the detector alone. The lowest (black) flat line represents the expected level of DC projection noise at a (large) optical depth of OD = 100, a detuning $\Delta = 2$ GHz, and beam radius $r = 1.1$ mm. The vertical black line represents the expected decay time at these parameters τ_{sc}. The blue (optical noise) and red (electronic noise) sloped curves represent the estimate at a frequency of 100 kHz. The difference indicates the degree that the detectors are optical shotnoise limited. (B) The value of the previous components for different powers at the decay time τ_{sc}. The photodetector becomes optically shotnoise limited above 1 mW and, without technical noise, it appears that spin-squeezing should be possible at higher powers and an optical depth of around 100 (as the blue dips below the black line).

each pulse needs to cancel their subtracted partners. The spiked peaks of the incomplete estimate are indicative of the high level of low frequency noise. Thus the success of the procedure is indicated by the fact that the minima lie on the expected shotnoise curves. The technique will not be as effective for more complicated, higher frequency background noise spectra, but for typical set of probe parameters the plot in figure 14.3 shows that it works well enough.

Of course there are still many technical hurdles to overcome even with the modulation scheme. Quite generally, if the mean spin-state direction becomes misaligned with the large holding field for any reason, this will produce a mean signal at the Larmor frequency with an amplitude proportional to the degree of misalignment. If there are classical fluctuations in the number of atoms in the probed volume (as there typically are at the level of 10%) this will create amplitude noise on the misaligned signal, which will then mask the desired

quantum fluctuations.

There are several imaginable ways by which the spins could become misaligned with the field. The first is through residual tensor dynamics, as discussed in section 7.2 and below in section 14.5.2. Even if the spins start aligned with the field, a slightly misaligned polarization of a strong enough probe beam can drive the spins off of the field direction, after which a signal at the Larmor frequency is apparent. Second, the spin-state may start slightly misaligned with the field in a random manner. Assuming the optical pumping beam direction to be fixed, the holding field could nominally drift about this due to low frequency field fluctuations, which could possibly lead to an initially misaligned state and field direction. Fortunately it appears that the spin-state is more aligned to the field over the pump beam direction after the pumping process such that this is not a problem. Third, magnetic field fluctuations resonant with the Larmor frequency of the holding field can lead to misalignments due to the magnetic resonance interaction demonstrated in section 13.7. Fortunately, only frequencies in the neighborhood of the Larmor frequency matter, and this can be tuned to a value (several hundred kHz) where ambient noise is minimal. Thus the spins can be used as a field spectrometer to locate regions quiet enough to perform spin-squeezing measurements. In practice, the noise generally decreases at higher frequencies, although there are spikes at particular frequencies due to equipment sources.

A final complication of the stroboscopic scheme is that the measurement time is dilated by a factor proportional to the duty cycle of the strobing. This is an issue because while the needed time to create squeezing increases, other limits on the time of the measurement (e.g., atom diffusion) do not, thus there is less flexibility in adjusting the measurement time relative to other timescales. For large degrees of targeted squeezing, this is an extreme problem because the duty cycle needs to be such that the probe is on for a total time much less than the total oscillation period [174]. However, for smaller amounts of targeted squeezing, the dilation factor is of order unity and is not a critical problem.

14.5.2   Tensor Oscillation

As mentioned in and section 6.7 and section 7.2, the tensor Hamiltonian causes nontrivial evolution of the mean spin moments for near resonant probe detunings. From equation (6.97) and equation (6.98), it is clear that if the probe light does not maximize $\langle \hat{S}_x \rangle$ (i.e., $\langle \hat{S}_y \rangle$ or $\langle \hat{S}_z \rangle$ are nonzero), then the spin-state initially aligned along x will rotate in either
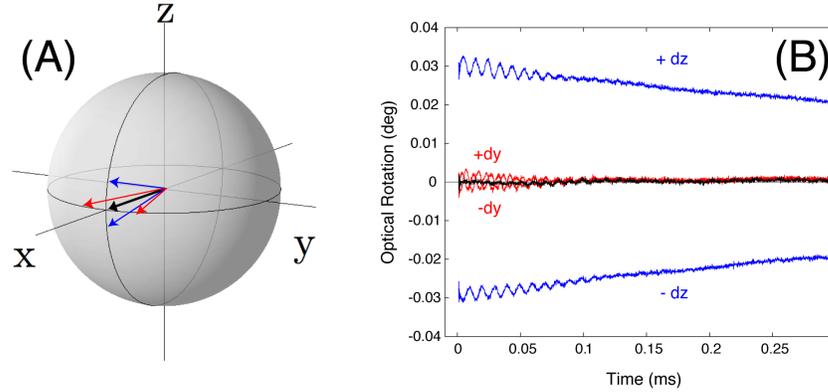
Figure 14.4: The spin-state is optically pumped along x, adiabatically rotated slightly, then an optical probe beam polarized along x and propagating along z is applied. A holding field, corresponding to a Larmor frequency of 110 kHz, is on for the duration of the measurement and is used to adiabatically rotate the spins. The optical polarization rotation is measured (nominally proportional to $\langle \hat{F}_z \rangle$). The blue curves (top and bottom) represent the state rotated ± 10 degrees toward the z-axis, the red curves represent the state rotated ± 10 degrees towards the y-axis, and the black (zero curve) represents the state held along x. When misaligned from the x-axis, the tensor Hamiltonian drives the state off of the field direction, after which it rotates at the Larmor frequency of the field. For the blue z-rotation curves, the z moment is observed as the decaying offset via the usual vector measurement Hamiltonian. The probe parameters used were: power 20 mW, detuning 2 GHz, beam radius 1.1 mm and the optical depth was approximately 20.

the y- or z-direction, leading to a misalignment between the spins and the holding field with subsequent oscillations due to the field. Again, atom number fluctuations on top of this signal can lead to noise that masks the desired quantum projection noise fluctuation. Thus, it is our goal to align the linear optical polarization axis to the spin and field direction, or vice-versa, depending on what is chosen to define the x-axis. In practice, this is achieved through minimizing the observed oscillation signal (possibly closer to resonance than usual to amplify the signal) by adjusting the input waveplates that determine the polarization of the probe beam.

In figure 14.4, we demonstrate representative trajectories for a fixed probe polarization and a rotated spin-state. For trajectories where an oscillation and offset are observed the spin-state has a small initial nonzero $\hat{F}_z$-component and for trajectories with only an oscillation the state has a small initial nonzero $\hat{F}_y$-component. It is clear for many reasons that these oscillations are induced by the tensor Hamiltonian and not, for example, simply the spins nonadiabatically slipping off of the field direction during the rotation that is not

shown. First, the shape and symmetry of these curves are well reproduced by simple single-atom master equation simulations (from chapter 5), which show the feature that the tensor induced oscillations decohere faster than the spontaneous emission damping offset. Second, the peak-to-peak size of the oscillations increases relative to the vector offset as the detuning is decreased. Third, the phase of the oscillation is fixed to the (variable) time that the probe is turned on.

These effects as well as the data in section 13.2 emphasize the utility of the tensor Hamiltonian for aligning the system in a way that would not be possible with the QND vector Hamiltonian alone. The tensor effect has also been shown to be useful in pumped state tomography in section 6.8 and section 13.6. Despite these practically useful features of the tensor Hamiltonian, the tensor Hamiltonian potentially ruins the idealized QND description of the interaction. Even if the system is aligned as above to eliminate the initial evolution of the mean spin moments, the results of section 7.2 indicate that the tensor Hamiltonian also affects the spin uncertainty distributions in this configuration (parallel spins and probe polarization). The deterministic evolution of the uncertainty distribution induced by the tensor Hamiltonian can be both single-atom and collective in origin. Near resonance, these terms are of the same order as the uncertainty evolution via conditional QND detection and will influence the optimal description of the spin-squeezing process. For the stroboscopic scheme of this section, the nontrivial tensor dynamics will be different from the constant probe scheme (as seen in figure 5.2), but still important when the detuning is close to resonance.

Despite these complications, the measurements discussed below aim to first get to the level where projection noise can be expected. After this level of sensitivity has been achieved, we can begin to consider distinguishing between the QND projection noise and the tensor induced uncertainty evolution.

14.5.3 High-Frequency Nonstroboscopic Measurements

Here we present the results of spin variance measurements using the high-frequency measurement scheme, but *without* strobing the probe beam. In the next section, we consider the same measurements but with strobing, which is the scenario under which one ideally expects to observe the projection noise most cleanly. (Without strobing under otherwise ideal circumstances, the spin state would diffuse in a way that our typical averaging fil-

ter would not optimally observe.) However, unfortunately, it turns out that the technical problems presented in this section also plague the stroboscopic case and prevent the clean observation of projection noise and, hence, squeezing in any case.

Before the experiment is run, we optimize by maximizing the observed optical depth, aligning the polarizations to avoid tensor fluctuations, and locating the holding field Larmor frequency at which the ambient field noise is minimal. Then we reproduce the procedure of figure 14.3 (where only the optical shotnoise and technical noise was considered), but with optically pumped, cold atoms present.

As discussed previously, there will be a few sources of technical noise that potentially obscure our observation of projection noise, even with modulation. One is the tensor oscillations due to misalignment as shown in figure 14.4. The second is the presence of magnetic fields on-resonance with our bias field Larmor frequency. Before showing the high-frequency atomic variance results along the lines of figure 14.3, we briefly discuss estimating an applied on-resonance field.

In chapter 10 we discussed the theory of estimating a small DC field at the quantum limit. Then, in section 13.7 and appendix D, we discussed the semiclassical magnetic resonance method for observing a particular high frequency field in a tunable fashion. An analogous version of a tunable AC magnetometer, except in a high temperature vapor cell, is presented in [175]. Now we investigate the level at which we can observe AC magnetic fields with our cold atom apparatus. In analogy with the DC scheme of chapter 10, the optimal procedure would be to fit the early time signal to a sinusoid with linearly increasing amplitude. The quantum projection noise offset in this case is represented by the initial amplitude and the field value by the slope of the amplitude in time.

The result of estimating AC fields in a suboptimal way is shown in figure 14.5. The signal for each field is multiplied by a sine of the appropriate phase and frequency (we assume both as known). At early times, the sign of the multiplied signal is then all positive, and to estimate the field we simply perform a running average and normalize appropriately in time (assuming we know the atom number) as given in equation (D.22). This is equivalent to suboptimal procedure for the DC case presented in [28] where, instead of independently fitting the slope and offset of the short-time signal (discussed in chapter 10), we estimate the slope by measuring the average of the total curve. This procedure is suboptimal because it unnecessarily convolves the spin-uncertainty into the field-uncertainty. However, for the
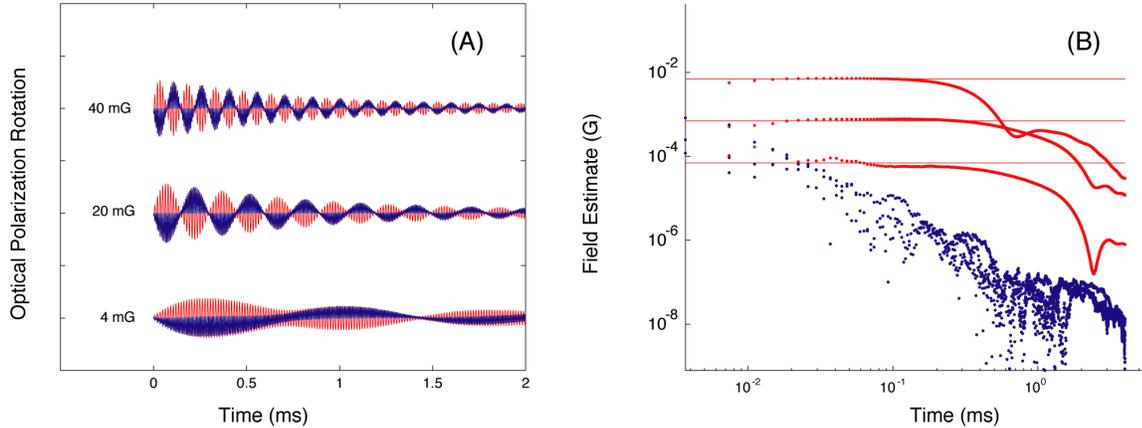
Figure 14.5: (A) Red curves are the photocurrent trajectories for the NMR scheme with different amplitudes for the perpendicular oscillating fields. The blue curves are the photocurrents multiplied by a sine wave of the same frequency (all positive for the first half cycle). (B) Estimates of the perpendicular fields (for smaller field values) over short times using equation (D.22). The parameters for this measurement were OD $= 12$, $\Delta = 2$ GHz, $P = 3$ mW, $f = 270$ kHz.

current situation, we are not yet atomic projection noise limited and the two procedures will give a similar result.

In [28], for the DC case, we investigated the difference in procedures that accounted for the initial offset (fitting the slope) and those that did not. The expected difference was apparent, however, in retrospect we believe that those experiments were not projection noise limited and the offset noise was technical in nature.

Now we measure the noise-floors explicitly, as in figure 14.3 but with atoms. The results are shown in figure 14.6A and described further in the caption. Unfortunately, there are two primary problems. One is that the optical depth is not large enough to observe the expected leveling off of the measured signal before the decay time sets in. The second is that there are still sources of technical noise that are proportional to the atom number and optical depth. Although they are not apparent in this plot, the tensor fluctuations cannot be completely zeroed and the atomic technical noise floor is typically greater than the photon shotnoise floor. (A more typical result can be seen in the data of figure 14.7.) Thus, to solve the first problem, we need to increase the optical depth further with more creative, less technically limited methods. However, as the optical depth is increased the second problem becomes worse, and we must simultaneously reduce the tensor oscillation and other technical sources of noise.
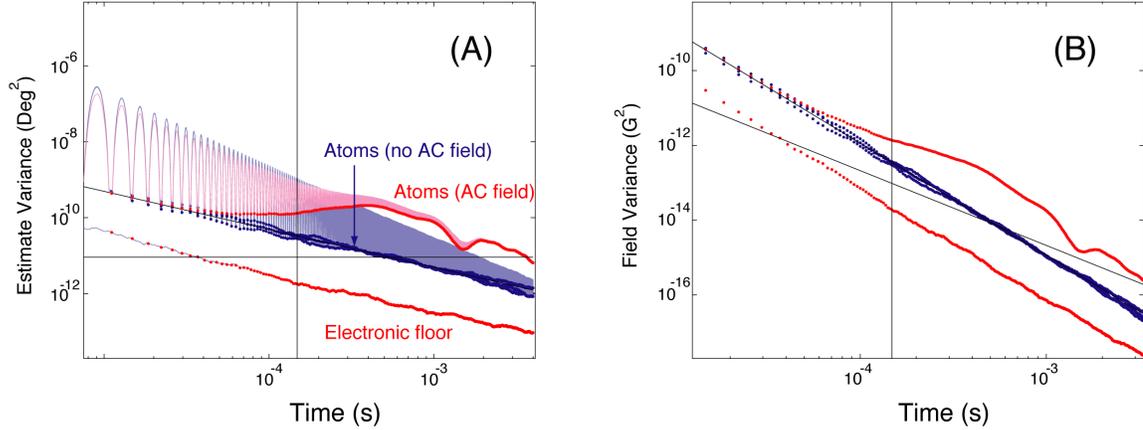
Figure 14.6: High-frequency, constant probe noise statistics. (A) Spin-noise of high frequency signal (270 kHz). The vertical line represents the expected decay time. The horizontal line represents the expected level of the atomic projection noise. The bottom data gives the electronic noise floor. The black sloped curve is the expected optical shotnoise floor. The top data points show the atomic noise with a small AC field present. The blue data points show the atomic noise floor (at zero field) and lie roughly on top of the optical shotnoise floor. The fact that the projection noise level is below the optical shotnoise level at the decay time indicates that we do not expect to see much squeezing. (B) The corresponding field sensitivity using an averaging estimator as described in the text. The parameters for this measurement were OD = 18, $\Delta = 2$ GHz, $P = 25$ mW, $f = 270$ kHz.

In figure 14.6B, the spin noise data of figure 14.6A are converted into the corresponding estimate of the fluctuating field. For larger fields, atom number uncertainty multiplies each estimate leading to a noise above the photon shotnoise floor of size BdN/N, where B is the size of the field and dN/N is the relative fluctuation of the atom number (typically around 10%). This uncertainty does not play a role for near-zero fields or if feedback is used to zero the field. For near-zero fields, the optical depth is small so that the measurement is limited by the photon shotnoise and the atomic projection noise does not play a role in the final sensitivity. Regardless, figure 14.6B shows that the measurement has a sub-μG sensitivity over millisecond timescales.

14.5.4 High-Frequency Stroboscopic Measurements

Here we present the results of spin variance measurements using the high-frequency measurement scheme, but now using stroboscopic modulation of the probe beam. If the experiment were capable of observing quantum projection noise and hence preparing conditionally spin-squeezed states, this would be apparent in these statistics as a leveling off of the variance
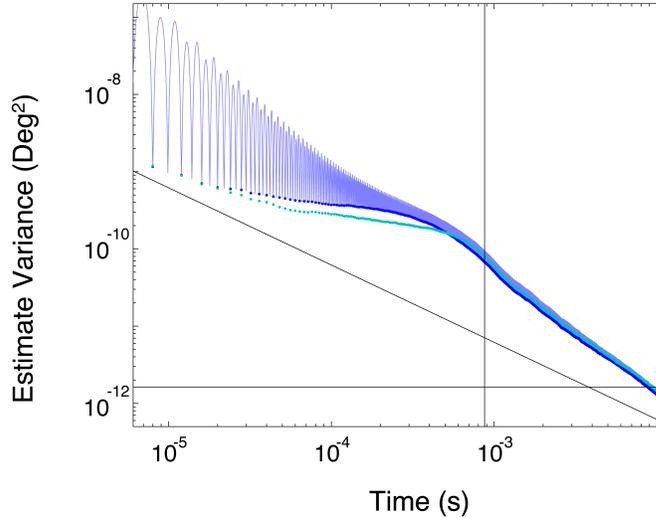
Figure 14.7: High-frequency, constant probe noise statistics, more typical than figure 14.6. The presence of slight tensor oscillations shown in figure 14.4, combined with atom number and alignment fluctuations lead to a noise level greater than the expected projection noise level. The lower data represent an attempt at normalizing the data by a later measurement of the total Larmor height, resulting in only marginal improvement. The parameters for this measurement were OD = 25, $\Delta = -5$ GHz, $P = 20$ mW, $f = 250$ kHz.

of the estimate in time at the expected projection noise limit. However, in the end, the technical problems demonstrated in the last section also plague this scenario.

First we consider only the noise statistics of the strobed optical beam to confirm that no additional technical noise is introduced via the sinusoidal pulsing method schematically shown in figure 14.2. Similar to the plot of figure 14.3, the noise statistics for a typical optical strobing parameter set are displayed figure 14.8 and described in the caption. The sine quadrature is by definition in phase with the pulsing and the cosine quadrature out of phase. As detailed in section D.2, the sine quadrature variance is 3/8 smaller than the DC case and the cosine quadrature is 1/8 smaller. However, each of the data curves are bumped up by the nonnegligible electronic noise floor. The frequency of the strobing is actually limited at higher frequencies where the AOM will put out a pulse fringe with nonoptimal visibility. The phase of the pulsed light (fixed by phase-locking the function generator triggering the experiment and those controlling the probe modulation) is determined by measuring the output of the polarimeter in slightly unbalanced configuration.

In the nonstrobe case, the quadratures were equivalent and the conditioning frequency could be chosen post-measurement to match the experimental field. In the strobe case,
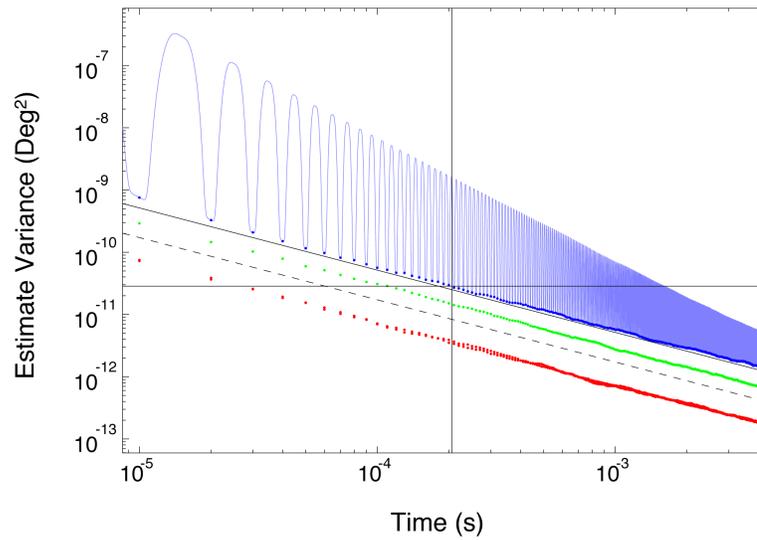
Figure 14.8: All optical variance data (without atoms). The probe light of power $P_{\max} = 18$ mW is strobed at $f_{\mathrm{strobe}} = 200$ kHz and both quadratures extracted from the measurement record. The upper blue data display the in-phase sine quadrature (again with the points at an even number of cycles), and the green data the out-of-phase cosine quadrature. Each is bumped slightly above the expected line (solid black for sine and dashed for cosine) because of the presence of electronic noise shown below in red. The vertical decay time line $\tau_{sc,\parallel}$ and the horizontal projection line are shown assuming a detuning of $\Delta = 2$ GHz and a large possible optical depth OD $= 100$. The number of trials used was 500.
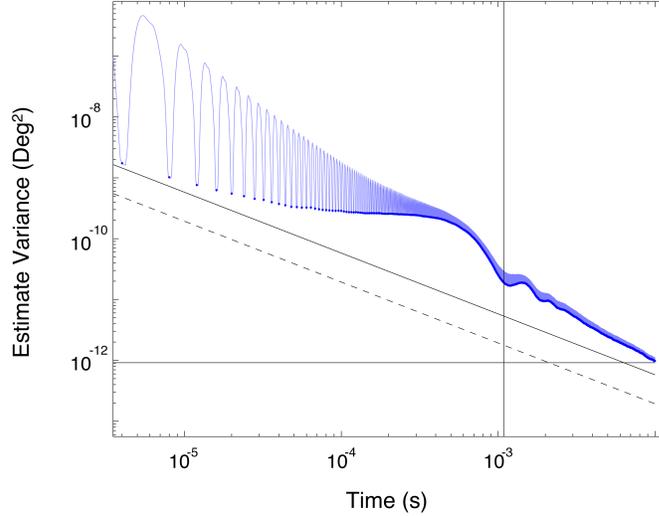
Figure 14.9: Atomic variance data. The probe light of power $P_{\mathrm{max}} = 16$ mW is strobed at $f_{\mathrm{strobe}} = 500$ kHz and both quadratures extracted from the measurement record. The upper blue data display the in-phase sine quadrature (again with the points at an even number of cycles), which is above the black optical shotnoise level. The vertical decay time line $\tau_{sc,\parallel}$ and the horizontal projection line correspond to the detuning of $\Delta = -5$ GHz and the independently measured optical depth OD $= 25$.

however, one needs to ensure that the strobing frequency matches the field frequency before running the experiment. This calibration is performed by aligning the spins perpendicular to the holding field and allowing them to Larmor precess at full scale about it. The strobe frequency is then tweaked between trials until the strobe and the Larmor precession remain in phase for the duration of the measurement time. In practice, this is easier to achieve if one waits a few extra ms after optical pumping to ensure that the decay of the MOT field is no longer causing the holding field to change in time.

A representative plot of the atomic noise with strobing is shown in figure 14.9, with specifications given in the caption. As with the nonstrobing case, there are again two problems. First the expected projection noise level is below the point where we expect to see anything. Second, even for this relatively low optical depth, the atom dependent technical noise obscures even the optical noise floor. The observed atomic noise is technical in origin and the size is expected from a measurement of the nonzeroed, deterministic signal resulting from misalignment, combined with expected atom number fluctuations of around 10%. Again, the deterministic nonzero signal is due to both tensor misalignment effects and possible stray fields. The tensor effect would get better with larger detunings, but then the

decay times at available power becomes unacceptably long. The field noise generally gets better at larger frequencies (larger holding fields), but then we are limited by the size of the holding field, the speed that we can strobe at, and the sample rate at which we can record data.

In summary, we have yet to unambiguously observe the atomic projection noise, or infer the preparation of spin-squeezed states, but technical noise levels are now well characterized and approaching the quantum projection noise limit. There are several experimental refinements that can be expected to lower the technical noise sources, and once the projection noise is observed there are several directions for future experiments. Suggestions for both are mentioned in the next chapter.

Chapter 15

Future Directions

By looking at the long list of references in this thesis, it should be quite clear that the work presented here stands on an intimidating amount of research stretching back more than a century. But, despite its length, this thesis is incomplete in many ways and represents a potential touchstone for future investigations. As I leave, I more than anybody know the holes that exist here, and I also have my own tastes for how I would like to see this research develop. In this final section, I outline a list of topics, both theoretical and experimental, which I would be very interested in working on if I were to somehow start graduate school again with what I know now.

15.1 Theory

There are a host of interesting research directions that build on the more theoretically inspired work of this thesis, including the entanglement results presented in chapter 4 and the control of stochastic differential equations discussed in chapter 9. However, due to my increasingly applied interests, all of the theoretical questions I discuss here are more directly motivated by particular quantum metrology experiments.

15.1.1 Polarimetry Conditional Master Equation

In chapter 8, we reasoned by analogy why the stochastic master equation for our experiment should be similar to the one derived in the context of homodyne measurement. Although we have derived the unconditional version of the master equation in chapter 7, we have yet to derive the conditional master equation from a physical model of the polarimetry experiment. The derivation of the SME will be different than the usual homodyne case where the local

oscillator is taken to infinity to simplify the expressions. Instead, the input beam will likely have to be divided into a semiclassical mean and additional quantum fluctuations. I expect the resulting SME to be as we have expected via physical considerations, but this does not take away from the importance of knowing this is true from a proper derivation.

(Update: Led by Luc Bouten, our group has recently finished a paper in which the polarimetry stochastic master equation is derived and shown to be as assumed in this thesis [111]. The paper uses a formal quantum stochastic differential equation (QSDE) formalism and also shows that the scheme is equivalent to performing homodyne detection on the y channel when the probe beam is initially polarized along x. This technique is potentially useful for limiting the power on the detectors under a situation where the local oscillator power is much less than the total probe power but much greater than the probe light scattered into the y channel.)

15.1.2 Three-Dimensional Scattering Physics

In this thesis, we have mostly chosen to work with an extremely simplified one-dimensional physical model to focus clearly on such things as the single atom interaction Hamiltonian, quantum filtering, and magnetometry. However, in truth we are dealing with a very complicated physics problem, involving a three-dimensional probe beam propagating through a spatially extended cloud composed of more than a billion diffusing atoms. Classically, one can imagine this cloud of atoms as a random array of dipole antennas that preferentially scatter the input beam in the forward direction. Under simplified circumstances, this would indeed give us the one-dimensional approximation we desire. However, under different atom-probe geometries, the validity of this approximation will vary, as the beam may be focused and distorted significantly.

One of the fundamental ideas of this work, as discussed in the introduction, is the fact that truly symmetric measurement, with physical indistinguishability of particles, can enable the preparation of entangled states. So the question becomes, given a realistic spatial model of the atom-light interaction, "Does the measurement of the scattered light allow us, in principle, to distinguish the spins or not?" Of course this question is very complicated and model dependent, and the trick is to simplify it, but not so much that the core idea is lost. Several papers have considered three-dimensional diffraction from the an optically probed ensemble in the context of quantum measurement [118, 115, 176], but the full story is still

incomplete. The issue of decay via spontaneous emission is often modeled naïvely in spin-squeezing applications, but it is also intimately tied up with the scattering physics discussed here [96, 177]. Ideally, one would like to think of a preferred atomic mode and a preferred optical mode that interact to give our simple model, but there exist nonpreferred atomic and optical modes that cause additional uncertainty. The leaking of information from the atomic mode into unobserved optical modes is what we call spontaneous emission and the leaking of information from nonpreferred atomic modes into the observed optical mode is another source of uncertainty [118]. From a more abstract perspective, the work of [178] considers the case of inhomogenous coupling to the atoms and calculates the general consequences of mode-mismatching between entangled state preparation and read-out stages.

Although our experiment involves many atoms, it is physically insightful to consider the example of two radiating atoms. If the atoms are close enough, some modes of radiating light will be symmetric, while detecting in other modes will allow one to distinguish the particles. Many of the techniques needed to understand this formally in free-space can be found in the work of Van Enk [179]. Although not discussed in the measurement terms used here, this type of example has been demonstrated in an experiment that observes interference fringes in the radiation from two trapped ions [180].

Even in one-dimension the existence of a non-QND Hamiltonian like the tensor Hamiltonian discussed here can lead to very interesting propagation effects [26]. Furthermore, one can show that simply due to angular-momentum conservation, the probe beam may be deflected as it propagates through the ensemble [32]. Clearly, when probing atoms with spatially extended probe beams there are many geometrical effects that may need to be considered. Using an optical cavity could simplify the problem to a degree because the modes then become more natural to describe, but there will still be similar mode-matching issues. Hopefully, simplifying approximations will enable efficient modeling of these systems, but those approximations should always be justified from the full physical model.

15.1.3   Nonsymmetric States for Metrology

For most of this thesis we have considered only symmetric states for metrology for several good reasons. First it is often experimentally convenient to manipulate all of the atoms at once, thus symmetric states are often easier to produce in the lab. Second, there are several symmetric states, namely the spin-squeezed state and the GHZ state that have been shown

to be useful in rotation measurement applications. However, the symmetric subspace is only a tiny fraction of the total Hilbert space available for a spin ensemble, and one is tempted to believe that by accessing more of that space higher precision measurements may indeed be possible. Of course, one needs to address the issue that with less symmetry comes the demand for more experimental resources to manipulate the parts of the entire system, but in many situations the trade-off may work out in favor of the nonsymmetric states.

Our group is currently working on investigating these ideas in an experimentally relevant context. Much of this work is inspired by, for example, [7] where it is shown that particular nonsymmetric spin-states are optimal for transmitting reference frame direction information. Again, it will be important to consider the mode-matching and symmetry properties of the experiment to determine entanglement production capabilities of the experiment [178].

15.1.4 Non-QND Collective Physics

In chapter 7, we derived the unconditional master equation describing the collective spin-state of the atomic ensemble with each atom evolving due to the full tensor Hamiltonian. Because of the assumption of full symmetry (indistinguishability), this master equation appears able to produce spin-squeezing unconditionally. However, in a physical experiment with a spatially extended cloud this approximation would not be completely valid. The non-QND tensor terms would lead to nontrivial semiclassical propagation effects as described in [26]. The QND approximation is often desired because of its logical simplicity, but it would be interesting to know what the expected quantum effects of the non-QND terms are, both because they occur for real alkali atoms and because the results of chapter 7 show them to be potentially useful.

On the other hand, one could try to avoid the tensor effects altogether to recover the QND approximation. The easiest way to do this is to move the probe detuning far away from resonance, but this is not always technically feasible. Other strategies involve defining new level schemes such that the tensor effects are avoided [112]. However, then one needs to address the issue of what metrology applications are suited to the newly defined level scheme. In any case, the experimental significance of the tensor terms needs to be acknowledged, especially for experiments operating somewhat close to resonance.

15.1.5   Stroboscopic Modeling

In section 14.5, we gave a very limited introduction to the idea of stroboscopic measurement of an oscillating spin state. In the case that the technical problems of this thesis are sufficiently resolved, it will be a priority to more seriously consider a stochastic master equation model including stroboscopic measurement. Again, these ideas were brought up in the context of LIGO [78, 41] and more recently considered in the context of nanomechanical oscillators [174]. However, we will be interested in extending the model to account for the particulars of our experiment beyond the intuitive, but potentially naïve, perspectives used earlier.

15.2   Experiment

15.2.1   Free-Space Continued

There still remains much work to be done on the core "free-space" experiment described in this thesis. First, one could attempt to reduce the low frequency optical noise on the polarimeter output that has made prohibited DC squeezing thus far. As discussed in chapter 14, this noise likely comes from many sources, and it may not be possible to achieve sufficiently low levels for realistic experimental parameters. Therefore it may be necessary to continue with the modulation scheme discussed earlier, which successfully bypasses low-frequency detector and magnetic field noise.

The optical depth can continue to be optimized via a combination of dark-SPOT and compression techniques. Even if a significantly larger OD is obtained however, there may also be larger amounts of technical noise that also increase with the optical depth. Thus, methods of identifying and reducing these sources of technical noise will be critical.

As mentioned in section 8.2.4 and chapter 14, we could also try to implement an approach that allows us to adjust the probe detuning closer to resonance, where the measurement rate is expected to increase relative to the spontaneous emission rate, thus allowing more squeezing. Unfortunately, this also means that the tensor terms become more important. However, if we work in a particular rotating frame it is conceivable that the tensor terms will time average away while the spin-squeezing dynamics remain. It is definitely worth pursuing such a scheme if the needed separation of timescales proves experimentally feasible.

However, it may still turn out to be necessary to work at even larger detunings, where the problems associated with a high power probe beam must be confronted.

After unambiguously observing the projection noise and spin-squeezing, one could proceed with doing truly quantum limited metrology. Thus we could demonstrate quantum-limited magnetometry measurements and also inertial sensing experiments, at low or high frequencies depending on the experimental limitations.

15.2.2   Free-Space Alternatives

Besides the spin-squeezing and magnetometry experiments discussed here, there are several other ideas that might be implemented in the current free-space configuration. Following a proposal presented in [86], one can imagine producing polarization squeezed light by passing the probe beam through the cold atomic ensemble multiple times. A high-frequency modulation scheme could possibly be designed to take advantage of the double-pass scenario of that proposal to successfully avoid ubiquitous low-frequency noise.

Another possibility is to somehow divide a single atomic cloud into two independently accessible, oppositely polarized subensembles that can be jointly measured with a common probe beam. In this manner, one could imagine reproducing the experiments of the Polzik group [36], where two spatially separated vapor cells were entangled (into a two-mode spin-squeezed state) with a common probe beam. This experiment, which could also be reproduced with two well-separated MOT clouds, has its own set of difficulties, but takes advantage of the feature that the relevant information is at a high (Larmor) frequency, thus avoiding low-frequency noise.

This idea may benefit from using a "split" state (an incoherent mixture of coherent states pointing opposite directions). This state and others can be prepared by modified optical pumping schemes as discussed in section 12.5.9. These alternate states may also prove useful for verifying quantum noise levels. For example, as detailed in section 3.3.4, the transverse variance of the $m = 0$ state should differ from that of the coherent spin-state by an expected factor of five.

Finally, one could do more sophisticated AC magnetometry than has been discussed here, e.g., detecting the phase of an oscillating magnetic field. This measurement could even be done adaptively, in analogy with the optical phase measurements from the first experiment I was involved with [12], thus bringing our research full circle.

15.2.3 Next Generation

Much of the motivation for creating spin-squeezed states comes from using those states in devices such as atomic clocks or interferometers, where the clock transition defines the pseudospin of interest. Accordingly, there has been much interest in using the techniques described in this thesis to create spin-squeezed states on the clock transition. Several groups have made great progress towards these ends by observing the dynamics of the clock transition via continuous measurement [45, 46, 47]. However, as described earlier, in free-space the amount of available spin-squeezing is limited by the optical depth of the ensemble, and to achieve significant degrees of squeezing the use of an optical cavity is necessary. Along these lines, the Kasevich group has made significant progress in achieving a robust cavity system, with many atoms contained in a relatively large cavity mode [47]. As a result of their efforts, they have achieved record cooperativity parameters, the key metric for attaining squeezing in a cavity mode. To increase measurement times there are also efforts to contain the atoms in an optical lattice within the cavity. While the use of optical cavities and lattices present a host of new technical problems for the experimentalist to combat, they represent the future of spin-squeezing research, and the research of our group as well as many others is moving quickly in this direction.

Appendix A

Magnetic Field Management

A.1   Introduction

The art of controlling magnetic fields is of extreme importance in many atomic physics experiments. Whether we intend to cool and trap atoms or measure their collective spin-state, we must cancel stray fields and apply driving fields with high levels of precision. This is especially true in experiments with broadband measurement and control, where low frequency sources cannot be avoided. Here we discuss some of the experimental details related to our use of magnetic fields in the experiment described in this thesis.

First, to motivate one to pay sufficient attention to field control, we state the stringent field cancellation requirements that our experiment demands. Next, we detail the sensors we have available to measure magnetic fields. Then we briefly characterize all of the field sources within a typical lab setting as measured by these sensors. In the next section, we discuss the specifics of applying fields with current coils. Then we discuss the electronics needed to drive these coils. Finally, we detail techniques for either canceling or avoiding stray fields.

A.2   Field Requirements

To be able to resolve the projection noise of measured atomic spin ensembles, the rotation of the spin vector due to stray fields during the measurement must be small compared to the expected angular projection of the measurement. For an ensemble with N spins polarized

perpendicular to the probe direction, we expect the projection angle to be

$$\theta_p \quad \approx \quad \frac{\sqrt{\langle \Delta \hat{F}_z^2 \rangle}}{\langle \hat{F}_x \rangle} = \frac{\sqrt{fN/2}}{fN} = \frac{1}{\sqrt{2fN}}. \tag{A.1}$$

A transverse field of size b will cause the vector to rotate via the angle $\theta_b \approx \gamma bt$. Setting these equal, shows that we need to reduce the field size to

$$b \quad \approx \quad \frac{1}{2\pi \gamma_{B4} t \sqrt{2fN}}. \tag{A.2}$$

For $f = 4$, $\gamma_{B4} = 350$ kHz/G, and typical parameters of $N = 10^9$ and $t = 1$ ms, this gives $b \approx 5$ nG. This is a small field when compared to typical ambient laboratory DC fields (around 1 G) as well as AC fields (approximately 1 mG at 60 Hz and 0.1 mG at sub-Hz frequencies).

In addition to nulling the average field over the sample, we must also take care to cancel any field gradients over the extended sample region. Any gradient over the atomic cloud would cause unnecessary dephasing of the atomic coherence, because different atoms would see different fields. If we demand that the field difference across the sample is approximately the needed field resolution from above $\Delta b = 5$ nG with a sample diameter of $\Delta x = 2$ mm this gives a needed gradient of $\Delta b / \Delta x \approx 25$ nG/cm. The typical gradients found in the lab (near sources such as ion pumps) are of order 10 mG/cm, thus we need to cancel this gradient by a considerable amount.

A.3 Magnetic Field Sensors

Now we discuss a wide array of commercially available magnetic field sensors. Under optimal working conditions, the highest resolution sensor in the lab may in fact be the atoms themselves, but it is of critical importance to have a collection of other sensors (with a variety of trade-offs) available for the following reasons: (1) to measure the fields quickly, continuously, and conveniently, (2) to characterize and cancel the ambient fields (DC and AC) sufficiently well that a signal from the atoms is even possible, and (3) to calibrate applied fields.

There are a number of excellent reviews in the literature for understanding the inner-

workings and relative performances of existing field sensors [181, 182]. Typically, sensors are divided into two classes: those that measure fields larger than 10 G or so (termed "Gaussmeters" or "Teslameters") and those that measure smaller fields well (called "magnetometers"). A further division can be made of sensors into vector magnetometers (which can measure orthogonal field components independently) and scalar magnetometers (which only measure the absolute field value).

A.3.1 Gaussmeters

Gaussmeters are useful for coarse zeroing of fields within atomic physics labs, but are typically used to measure very large fields in, for example, condensed matter labs. Gaussmeters include the following sensors, many of which can measure with resolutions of mG: mechanical devices (reed switches, relays, compass needles, wire deflection), Hall effect devices, magneto-diodes, magneto-transistors, and magneto-resistors.

A.3.1.1 Hall Effect

The F. W. Bell 9950 Gaussmeter was the first field sensor purchased by the Mabuchi lab. It is mostly good for measuring large fields and zeroing fields coarsely, but it has reliability issues and frequently needs to be reset and zeroed. This is a Hall effect device operating in the ranges of 3 G–300 kG, with resolutions 1 mG–10 G, and bandwidth 0–400 Hz in DC mode and up to 50 kHz in AC mode.

A.3.2 Magnetometers

Magnetometers do not typically have the range of Gaussmeters (in open-loop configuration) but have much better resolution Different types of magnetometers include: magneto-resistors, induction coils, fluxgates, fiber-optic sensors, proton precession magnetometers (scalar), optically pumped magnetometers (scalar), and SQUIDs.

A.3.2.1 Magnetoresistive

These are the cheap, midrange sensors that we bought initially intending to surround the atomic sample symmetrically to get an average measure of the field at the atom cloud position, as done in [183]. Their noise floor is about a factor of 1000 worse than the best

fluxgates, but these are reasonably fast (1 kHz), an order of magnitude less expensive, and smaller (around 1 inch). We purchased a set of 3-axis (Honeywell HMC2003) sensors for about \$200 a piece. The chips have a range of -2 to 2 G, a noise floor of 30 μG/$\sqrt{\text{Hz}}$ at 10 Hz, a sensitivity of 1 V/G, and a bandwidth of 0–1 kHz (which can be made slightly larger by removing capacitors). A reset pulse is needed to realign the sensor after significant field (e.g., the field from a MOT coil) is applied to the sensor.

A.3.2.2 Induction (Search) Coils

It is quite easy to create a homemade sensor from an induction coil, which serves as an antenna for measuring magnetic fields along its axis. By attaching a coil to an amplifier, the current running through the coil can be converted to a voltage. Because a changing field leads to a changing flux that leads to a current, this sensor can be good at sensing alternating fields. Without further loop shaping, the output voltage would be the derivative of the field (thus containing a factor of frequency in the spectrum). This is the easiest and most reliable method for locating and measuring higher frequency field sources (greater than 10 kHz).

A.3.2.3 Fluxgates

Fluxgates are essentially induction coils but with an internal high permeability material to amplify the field and the signal. One could use an unaided ferromagnetic material inside, but the permeability would only be high in the unstable region where the material is not net magnetized. To account for the inherent instability of ferromagnetic materials, the fluxgate uses a driving coil to alternately polarize the material one way and then the other (at some fast 10s of kHz repetition rate). In between being polarized, the permeability spikes to a high value (as opposed to when the spins are polarized/saturated, giving a permeability that is no better than vacuum). By demodulating the signal at the repetition rate, one can get a high resolution measurement of *both* DC and AC fields, which are slower than that rate. Although many different designs exist, most fluxgates consist of a regular cylindrical pickup coil with a toroidal ferromagnet inside which is wrapped with a driving coil. Because the driving coil is wrapped on itself, the field it produces does not contribute to the pickup coil signal. See [184, 181, 182] for more on the design principles of fluxgate magnetometers.

There are several kinds of fluxgate probes available. The high-end fluxgate designs are

the best available sensors for use in the lab, short of using a SQUID. Vendors for fluxgate devices include MEDA, Walker, Applied Physics Systems, and Bartington (vendor GMW). We borrowed an APS unit from the Kirschvink lab at Caltech temporarily before opting to purchase a Bartington ($4000 sensor $6000 signal conditioning unit) because the noise floor was almost an order of magnitude lower and the bandwidth was an order of magnitude greater than the APS. The fluxgate unit is a Bartington Mag03MS100 (with Low Noise Option) with a range of 1 G max (1 mG min) 0.04–0.06 μG/$\sqrt{\text{Hz}}$ at 1 Hz and bandwidth 0–3 kHz.

A.3.2.4 SQUIDs

Although SQUIDS (and optical magnetometers) are among the best available sensors in the world, we do not consider them for obvious reasons: they are too big, clunky, labor intensive, and expensive. Also, their bandwidths are typically slow compared to other devices (of order Hz). SQUIDs are cited to have resolutions around 1 nG [182]. Typical noises are 10 pG/$\sqrt{\text{Hz}}$ (1 fT/$\sqrt{\text{Hz}}$) but noise floors as low as 0.1 pG/$\sqrt{\text{Hz}}$ have been reported. Although the optical magnetometers appear similar there are a variety of other criteria that make SQUIDs more desirable for microscopy applications. For example, the inability to work at high fields is a serious flaw of optical magnetometers compared to SQUIDs [185].

A.3.2.5 Optically Pumped

The GEM Potassium magnetometer has a specified resolution of 1 nG. For commercial optical magnetometry information see `http://www.gemsys.ca`. In the lab, noises as small as 5 pG/$\sqrt{\text{Hz}}$ (0.5 fT/$\sqrt{\text{Hz}}$) have been reported in a different optical setup [34] where the theoretical limit of the device is expected to be around 0.1 pG/$\sqrt{\text{Hz}}$. These magnetometers operate on much of the same physics as our experiment, but in a hot vapor cell, and in a configuration that is concurrently pumped and probed.

A.3.3 Current Measuring Devices

There exist a number of devices that use the principles of magnetic field measurement for measuring the current passing through a wire (ammeters, current probes). Of course, these are useful in the lab for a variety of reasons, including field management. A typical device

(such as the Tektronix A622) will clamp around a wire, using an induction coil measure the AC field and a Hall probe measure the DC field.

A.4 Stray Fields

Now that we have described the limitations of our measurement capabilities, we shall describe the ambient field properties in our "typical" lab environment. Although it is often difficult to distinguish, these can be divided into two rough categories, fields due to people and fields due to natural processes (which, for example, could be measured at similar levels in the middle of a desert). Understanding all of the field noise present is a necessary prerequisite for designing a field canceling system.

A.4.1 Measured Total Stray Fields

Here I summarize what field spectrum our sensors see when placed at the center of our optical table. Before identifying a field source, one should make sure that the noise seen is not purely electrical, especially at low frequencies. One can confirm the signal is truly magnetic with the use of shielding, or looking for coincidence with multiple sensors.

At any point in the lab we typically observe a DC magnetic field of in the neighborhood of 0.1–1 G, with the largest component pointing down (due to the earth, the table, the ion pump). The AC field on the other hand is dominated by the power line noise (including electrical fan motors) and we observe mG scale fields fluctuating at the multiples of 60 Hz. We also observe fluctuations of a few hundred μG at sub-Hz frequencies. The people at LIGO have measured a similar noise floor to the one we see with the same sensor: `http://www.ligo-la.caltech.edu/PEM_Ref/magref/magref.htm`.

We are also concerned with the spatial variation of the AC and DC noise fields. For the DC fields, we are only able to cancel the field and gradient well, while the curvature is more difficult to deal with. For the AC fields, the spatial variation has important consequences for how we decide to cancel AC fluctuations (e.g., getting a lockable line signal from a nearby sensor). The ion pump alone produces DC gradients around 50 mG/cm and DC curvatures of 1 mG/cm^2. For the AC fields, the phase of the near-field line noise changes over a relatively small length scale. Moving the sensor across the table by hand (near scopes, function generators, etc.) shows that the phase of the 1 mG signal changes phase by 180

degrees in about a meter.

A.4.2   Low-Frequency Sources

Here is a list of low frequency (less than 1 kHz) field sources and expected contributions to the noise spectrum. For those objects that are stated as giving only DC fields, note that those sources may move and create slow time-varying fields.

A.4.2.1   Natural

Lets first consider signals you would possibly see in a hypothetical location with no electronics around except for the sensor itself.

The earth below: There is a large body of literature that discusses the earth's magnetic field, for example [186]. The earth's field is typically around 0.5 G but it is difficult to isolate given other Gauss-like sources in the lab. The website `http://williams.best.vwh.net/` `magvar.htm` will calculate the expected field components of the earth for your location. For our zip code, it gives $B_x = 0.23$ G (north), $B_y = 0.06$ G (east), and $B_z = 0.44$ G (down). The DC measured field in the center of the table with the trim coils off is $B_x = 0.04$ G (north), $B_y = 0.14$ G (east), and $B_z = 0.42$ G (down) (whereas near door the field is [0.35, 0.16 0.27]), and the field from the pump is probably canceling most of the earth's x-field.

The atmosphere above: Certain atmospheric processes (e.g., solar winds driving the magnetosphere) are capable of contributing around 0.1–1 mG at sub-Hz frequencies [186]. The scale of noise depends on environmental conditions and can be much larger during magnetic storms.

A.4.2.2   Experimental System

Now we move from the desert to consider things closest to our atomic system. These sources, residing on and above the optical table, are necessary for the experiment itself and need to be where they are or else we would move them far away.

Optical table: Our magnetic optical table is a significant source of DC and AC fields. The DC field perpendicular to the table ranges from 1 G near the surface to 100 mG a few

feet above. The DC field parallel to the table due to the table is much less. The AC field perpendicular to the table is around 1.5 mG p-p near the surface and 0.5 mG p-p a few feet above, whereas the parallel AC field is not affected much by the distance from the table and is around 0.5 mG p-p. If one wants to avoid these fields, note that it is possible to buy a table made from nonmagnetic steel (e.g., from TMC).

Ion pump: We use a VacIon Plus 40 ion pump in our vacuum system. The manual contains approximate plots of the fields that are confirmed by measurement. The maximum field from the pump is 1–3 G within 20 cm of flange and less than 0.5 G beyond. The maximum gradients are of order 50 mG/cm. The maximum curvatures are around 1 mG/cm^2. There are no known significant AC fields from the pump because it operates with DC electric and magnetic fields. Currently the pump is shielded with two layers of mu-metal, so the fields are much less.

Vacuum chamber: There exist both magnetic and nonmagnetic steel parts for vacuum chambers. Although our primary system is within a quartz cell and we have tried to keep most materials in the near vicinity nonmagnetic, the adjoining steel parts give some amount of field, but most of it is nonmagnetic. At our cell distance (10s of cm), magnetic steel gives approximately 100 μG fields, while nonmagnetic steel gives less than 1μG fields.

Magnetic mounts: Standard magnetic mounts produce many mG at a meter and 100s of μG across the room.

Electronics: All electronics boxes (function generator, scopes, etc.) put out order 1 mG at 60 Hz, largely due to the magnetic motors of cooling fans. For some reason, the SRS function generator puts out 20, 40, 150 Hz into air.

Speaker shutters: The flagged speaker shutters are obviously DC magnetic and when switched give a changing field. Close to the switch (around 1 foot) the on/off change in field can be as high as 0.5 mG. At the distance of the MOT from the switches, the on/off change in field is more like 10 μG. Because of the solenoid geometry, most of the field change is in the vertical direction.

Transformers: Many devices require transformers to be powered. These transformers are huge antennaes of 60 Hz noise. For example, the Thorlabs detector transformer gives significant noise around 1 mG p-p at the 60 Hz multiples. We would move them farther away but there is so much 60 Hz noise that it is hardly worth it.

A.4.2.3 Peripheral

Most common items in the room produce a field. The following produce fields of a few hundred μG: lab doors (inner and outer), screen partition, chairs, stools, metal floor planks. At the larger range of mG are things like metal carts, file cabinet drawers, and screwdrivers. The ladders in lab are actually mostly nonmagnetic but still produce fields of about 10 μG.

CRT monitors: The adaptive computer monitor does not appear to give any noise (at least less than 1 μG) in operation, but gives a 10 μG kick when turned on. However, the black CRT monitor on the desk causes modulation of the 60 Hz noise at a few Hz. It probably contains a transformer from 50 Hz that creates the beatnote. This should definitely be turned off during experiments as it increases the line-locked variance by a factor of 10 or so.

Battery chargers: There is one emergency light in every room with a battery that is continuously being charged. We started to suspect these as sources of noise after seeing the noise from Striker's battery, but we have not measured any significant noise from them. If necessary, they can be removed from the wall.

Surge protectors: The APC surge protectors will give 0.5 mG at 2 Hz signal when plugged in but not connected to anything.

Printer: The outside printer gives a characteristic 50–100 μG oscillating signal as it is moving the paper.

VWR oven: After much searching, I found one of the most annoying sources of field noise in the lab. It turns out that the VWR oven/incubator in the bio-lab will draw spikes of heater current every 2 seconds (as can be seen from the front LED) when it is near its temperature equilibrium. I initially noticed a 50 μG signal of this character in the ensemble

lab and tracked the noise source up to the conduits on the ceiling. The noisiest conduit was followed back the the bio-lab where the oven source was eventually discovered. This noise is intermittent because apparently the oven occasionally loses lock and the peaks are of random orientation for some unknown reason.

A.4.2.4   Neighbors

Machine Shop: In the next door machine shop, there were toy magnetic pendula that gave around 20 μG p-p noise at 0.5 Hz, which have since been removed. There is also a backup "trickle-charge" battery in this room that can be observed via its field noise.

Low-Temperature Labs: In the Mabuchi Lab (subbasement of Sloan), there is a cryostat with a superconducting solenoid in the Eisenstein lab directly east of the bio-lab. I performed some tests with Xerxes (a grad student in the Eisenstein group) on 3/19/04 to measure the fields produced on each of our three optical tables from their magnet. Here are the results with x pointing up, y north, and z west, and numbers stated in mG produced in our lab versus Tesla inside their coil. In the bio-lab: $B_x = 2500$ μG/T, $B_y = 700$ μG/T, and $B_z = 1750$ μG/T. In the QED (center) lab: $B_x = 350$ μG/T, $B_y = 250$ μG/T, and $B_z = 350$ μG/T. In the adaptive/ensemble lab: $B_x = 300$ μG/T, $B_y = 100$ μG/T, and $B_z < 50$ μG/T. Their coils are capable of going as high as 9 T and typical ramp rates for them are 1 T/min although they can be faster. When they are doing experiments the magnet is ramping 90% of the time.

Miscellaneous: It is expected that the elevator, cars, and subway lines can contribute mG signals but none of these have been investigated.

A.4.3   High-Frequency Field Sources

For the DC magnetometry experiments, we must be very careful in creating a low frequency magnetic environment and negate the sources above. As mentioned earlier, an AC scheme is also possible where fields at a particular high frequency (50–500 kHz) can be measured via magnetic resonance (see appendix D). To observe a particular frequency with the atoms, a holding field is placed along the spins and motion of the spin-state off of the holding field is only possible given a near-resonant, perpendicular AC field.

Although the noise spectra in general gets quieter at higher frequencies there are certain frequencies with a lot of noise, mainly due to power supply switching. If a quantum limited experiment is to be performed, it should be at a frequency largely absent of this noise.

Here is a brief summary of some high frequency sources in the lab, as observed with either the atoms or an inductive search coil: Fluxgate (16 kHz multiples, from internal modulation), CRT Toshiba (very large 15.75 kHz multiples), Mini-Vac ion pump high voltage source (17.75 kHz multiples, especially 53.25 kHz), Vortex diode laser supply (27.6 kHz), Gaussmeter (13.6 kHz (large) 4 kHz spaced chain), Acopian power supplies (small 32 kHz), Fluorescent lights (7.5 kHz, huge 46 kHz, mound at 60 kHz, 90 kHz), HP supplies (many peaks, spaced by 1kHz), Oscilloscope (23.5 kHz, 38 kHz, 47 kHz, 53 kHz), Spectrum Analyzer (47 kHz), Laptop (60 kHz).

Above 200 kHz or so the "forest" of peaks becomes less dense and it is easier to find a frequency at which to perform a quiet measurement.

A.4.4 Fundamental Limits of Field Noise

Ultimately, the field noise floor is limited by the thermal noise. To get a sense of the scale of this noise consider the set of large shim coils, which have a conversion ratio of approximately 1 G/A and a total resistance of $R = 8\ \Omega$. The Johnson voltage noise floor (in $\mathrm{V}/\sqrt{\mathrm{Hz}}$) is $v_R = \sqrt{4kTR}$ and the current noise floor is $i_R = \sqrt{4kT/R}$. Using $k = 1.38e - 23$ J/K, $T = 300$ K, and the conversion factor below, the latter gives $b_R = 5 \times 10^{-11}\ \mathrm{G}/\sqrt{\mathrm{Hz}}$. See page 351 of [1] and [34], for a similar analysis that predicts a comparable noise floor near a thermal conductor.

The current shotnoise is given by $i_S = \sqrt{2qI_b}$ where q is the electron charge and I_b is the bias current. Using $I_b = 1$ A and the above conversion factor this gives $b_S = 6 \times 10^{-10}\ \mathrm{G}/\sqrt{\mathrm{Hz}}$.

A.5 Applied Field Sources: Coils

In this section, we discuss how to apply magnetic fields with current carrying coils of wire. First we discuss basic magnetostatic physics and how different fields can be produced with certain coil geometries. We summarize field formulas for basic coil configurations and provide references for more refined geometries. Subsequently, we discuss more practical con-

siderations for how the coils are constructed. We then include specifications for particular coils in the experiment. Finally, we describe the precision current drivers needed for each set of coils.

We also consider only moderate size fields and do not discuss Tesla size fields produced by superconducting magnets (sold, for example, by Oxford for condensed matter applications) or small-scale fields as produced on atom chips. The highest fields we might possibly need are for strong field gradients for a Stern-Gerlach analysis of the atoms as discussed in [187]. At some point we may want to include a discussion of RF field sources (horns), especially if we ever use the clock transition, but currently we do not.

A.5.1 Coil Physics and Modeling

A.5.1.1 Biot-Savart

All one needs to do most magnetostatic analysis of current carrying wires is the Biot-Savart law, which is discussed in any elementary electromagnetics book [188]. To get an order-of-magnitude sense of current to field conversion, consider the simplest possible wire geometry: an infinitely long straight wire. Recall that the transverse field a distance R from the wire with current I running through it is

$$B = \mu_0 I / 2\pi R. \tag{A.3}$$

For a distance of $R = 1$ m from a wire carrying $I = 1$ A, there is a field of $B = 200$ nT $= 2$ mG. Given the simple Biot-Savart law, one is equipped to design a wide array of static field sources. See `http://www.netdenizen.com/emagnet/index.htm` for fields produced by common wire geometries.

A.5.1.2 Coil Fields

A typical concern is to produce a uniform field of constant magnitude in an extended space. Of course, the best way to do this is to make a large solenoid around the sample with many turns. This is not often done because, among other problems, it limits optical access to the sample. The most common solution is to use two coils in a Helmholtz configuration, which is an approximation of the solenoid, where two coils are oriented in the same direction and produce a large field along the midpoint of the axis. Reference [137] discusses higher-

order approximations with more than two coils that allows for field uniformity over greater volumes.

For square coils of length L (per side), the coils should each be $d = 0.27L$ from the center (such that their spacing is $2d$) for optimal uniformity. In this configuration, the central field is $B = 18/L$ mG/A with L in meters.

Another concern, for example when creating a magneto-optic trap (MOT) is to produce a point in space with zero field, but a large field gradient. The simplest solution to this problem is to use two oppositely oriented coils in anti-Helmholtz configuration. The field on the axis of a single circular current carrying coil is (p. 211 [188])

$$B \quad = \quad \frac{\mu_0 I}{2} \frac{R^2}{(R^2 + x^2)^{3/2}} \tag{A.4}$$

for coil of radius R at a point x from the center. Placing two oppositely oriented coils of radius R a distance d from the center position (such that their total separation is $2d$) and measuring a distance $z \ll R, d$ from that point along the axis, we get a central axial field gradient of

$$\frac{\delta B}{\delta z} \quad = \quad 3\mu_0 I \frac{d}{R^3} \frac{1}{(1 + (d/R)^2)^{5/2}}. \tag{A.5}$$

Setting $R = 2d$ maximizes the gradient at the value $0.86\mu_0 I/R^2$.

Finally, one can imagine trying to design a coil that produces higher-order derivatives of the field with more and more funky shapes. This is a common goal in MRI where the goal is to cancel residual fields to very high derivative order.

A.5.1.3   Inductance

Given a coil geometry another quantity of interest to calculate is the self-inductance. The inductance fundamentally limits the bandwidth that the coil can be driven to the frequency $\omega = R/L$ where R is the appropriate resistance from the driving circuit. Inductances are a little nasty to calculate, but the equations for most common geometries are available in standard references (see [189] and the following website that uses those results: `http://emcsun.ece.umr.edu/new-induct/`).

Here we state the inductance formula for the two most standard coil shapes: a square

and a circle. The inductance of a square coil is given by:

$$L \approx N^2 \frac{2\mu_0 L}{\pi} [\ln(L/a) - 0.774] \tag{A.6}$$

where N is the number of turns, L is the length of a side, and a is the radius of the wire. The inductance of a circular coil is given by:

$$L \approx N^2 R \mu_0 [\ln(8R/a) - 2] \tag{A.7}$$

where N is the number of turns, R is the radius, and a is the radius of the wire.

A.5.1.4 Modeling

For standard coil geometries, the above approximate analytic treatment is completely adequate. For more complicated wire geometries, like microtraps, simulation software is very useful. In the past, we have used a magnetic field calculation program called Vector Fields (OPERA). Currently, we are using FEMLab (for Finite Element Method), a generalized Matlab PDE solver, which can do electromagnetic problems as well as thermal, stress, etc.

A.5.2 Coil Construction

The first thing to consider when constructing a coil is what magnet wire to use. Considerations here include the gauge (diameter), cross section (square/circular), insulation, temperature range, etc. The company we typically order wire from is MWS. If using small currents, a convenient option might be to use ribbon cable for the coils.

When constructing the mounts there are a few things to consider. One should definitely not use a magnetic material, unless the design calls for amplification of some kind. Aluminum is typically a good nonmagnetic metal, but if one cares about eddy currents, metals are a problem. When going to a nonmetallic material one must worry about machinability and heat conduction. Teflon and Kevlar are among the nonmetallic materials with adequate structural and thermal properties.

Although we have not had to worry about this, when applying a significant current one must worry about heating (particularly runaway heating). There are a few cooling options available, one of which is to use hollow wires with water running through as a coolant.

Alternatively one could just cool the mount itself with separate cooling tubes.

For winding magnets, a setup using a lathe as the winder is usually most efficient. Although one might be attempted to use square cross section wire for neatness, it is somewhat difficult to keep the rows aligned when winding, and we do not think it is worth the frustration.

A.5.3   Lab Coil Specifications

Here we outline the specifications for the various magnetic coils used in the lab.

A.5.3.1   MOT Coils: Anti-Helmholtz

A MOT requires an anti-Helmholtz coil set up with zero field and gradient of approximately 10 G/cm at the center. See schematics for exact geometry of our coils, which have approximately $R = 2.75$ inches and $d = 1.75$ inches (the $2d$ separation is only slightly nonoptimal). With the number of turns $N = 175 - 180$, this gives approximately 3.7 G/cm/A. The wire used has 1.5 mm diameter, which, for the above single coil inductance gives $L \approx 12$ mH.

A.5.3.2   Bias Coils: Small Helmholtz

For our square coils, $L = 10.5$ inches (d is close to the optimal 2.86 inches), and the number of turns is $N = 4$, which gives $B = 270$ mG/A. The wire used is 1 mm diameter, which, for the above single coil inductance gives $L \approx 19$ μH. Via the field controller, the applied voltage gets transferred to the current with a sense resistor R, which we have as being 10 Ω for coarse fields and 1 kΩ for fine fields. This leads to applied fields of size 27 μG/mV and 0.27 μG/mV, which is almost exactly what is measured when calibrating with a fluxgate. In the latter case, with a measured noise voltage of 0.5 mV this gets mapped into an applied field noise of 135 nG, but this is probably a worst case estimate.

A.5.3.3   Trim Coils: Large Helmholtz

Around the entire experiment we have a frame made of 80/20 about which square Helmholtz coils are placed to zero the field and gradient at the sample location. The frame is 36" across, 30" tall, and 30" deep. For a single-turn Helmholtz coil, $L \approx 33$ inches gives $B = 22$ mG/A and (with 1 mm diameter wire) $L = 4.5$ μH.

- Top/Bottom Pair (and Close/Far Pair): Number of turns is 28 per coil. With $N = 28$ the field becomes $B = 616$ mG/A and the inductance $L = 3.5$ mH. If operated in anti-Helmholtz configuration, the gradient generated is 5 mG/cm/A.

- Left/Right: Number of turns is 50 per coil. With $N = 50$ the field becomes $B = 1.1$ G/A and the inductance $L = 11$ mH. If operated in anti-Helmholtz configuration, the gradient generated is 9 mG/cm/A.

With the power supplies of current resolution 0.01 A, the corresponding field resolution is 6 mG for one set of coils and 11 mG for the other. This is too large to be useful hence we typically have another set of single turn coils with a homemade current supply for fine zeroing.

A.5.4 Current Drivers

A.5.4.1 MOT Coils: Anti-Helmholtz

The MOT coils are supplied with a few amps of current via a Bipolar Operational Power Supply (Kepco BOP 20-10M). During our experiment we need to turn off the large magnetic fields quickly because the atoms begin to fall/diffuse and there is not more than a few milliseconds that we can waste. If we simply turned off the supply voltage the system would ramp down at the $\tau = L/R$ timescale, which for $R = 1.3$ Ω and $L = 2 \times 12$ mH is $\tau = 18.5$ ms. This is too long. We reduce the damping time by more than an order of magnitude by using a circuit design from [190] using a high-voltage MOSFET (International Rectifier, IRFP140N), a high power resistor (actually four high power resistors in parallel, 100 Ω, 50 W, giving an overall resistance of 25 Ω), and a diode (MUR415 9423).

In operation, the Kepco (in voltage mode) supplies the box and the box supplies the coils. A TTL signal from the computer turns off the MOSFET switch after the atoms have loaded. To keep the current flowing the other side of the MOT coils go to a high voltage (which the MOSFET can withstand) and the current flows through the parallel high power resistors and diode. Because the high-power resistor is larger than the natural resistance of the MOT, the damping time is considerably decreased.

Alternatively, one can decrease the ramp-down time of the MOT fields and the eddy current fields from nearby metals, by using techniques from [191].

A.5.4.2   Bias Coils: Small Helmholtz

The key magnetic components in our experiment are the field drivers used to apply fast and precise fields along all three spatial directions in the lab with the "cube-coils" mentioned previously. These coils can apply up to 1 G in less than 1 μs, which is a task needed for performing the optical pumping, the DC field experiments, and the AC field (magnetic resonance) experiments described in this thesis.

The ramping time of a magnetic coil is limited by the inductance of the coil. Even though the Helmholtz coils are made relatively small to keep the inductance low, the L/R time for the coils described would not be as fast as 1 μs without further effort. The idea is to achieve faster switching times with a high speed, high voltage power supply. If one requires a fast step function of the field, instead of just applying the end-voltage to start, one applies a much larger transient voltage to "overpower" the inductor into achieving a faster performance. In practice we use a High voltage MOSFET Power Op Amp (APEX Microtechnology PA05) with a magnetic deflection design taken from the Apex Application Note 5.

For the power supply, we have used two Lambda 150048-RA supplies, which provide ±48 V to the op-amps (and up to 32 A or 1500 W). Unfortunately, as with most power supplies, these supplies put out a significant amount of high frequency switching noise (in particular at 60 kHz), which is an experimentally relevant noise source especially for our AC measurements. To alleviate this problem, Anthony Miller has built a charger circuit such that we can supply the Apex chips with a quiet battery supply during the experiment and charge the batteries with the Lambda supplies when the experiment is offline. More on the charger can be found at `http://qwiki.caltech.edu/index.php/QwikiCharger`.

A.5.4.3   Trim Coils: Large Helmholtz

There are actual two large (1 m) Helmholtz pairs of coils for each axis meant to zero the DC field and gradient. The first set of coils has many turns and is supplied with HP E3630A voltage supplies (six total). The second set of coils is only a single turn and is supplied by a homemade current supply box provides six bi-lateral, manual/programmable supplies of ±0.8 A. The design can be found in my electronics notebook, but is based on a design from page 388 of Horowitz and Hill [192]. The design uses a voltage divided from two

precision references (± 7 V, National Semiconductor LM399) as the manual signal, which is amplified with a power operational amplifier (National Semiconductor LM675). The same basic circuit design can be found on page 7 of the LM675 data sheet.

A.5.5 Field Calibration

Once the experimental system is setup the Larmor precession signal from the atoms can be used to both calibrate and align (or orthogonalize) the fields. The absolute calibration comes from a knowledge of the gyromagnetic ratio. The more subtle orthogonalization procedure uses effects of the tensor Hamiltonian.

A.6 Nulling Stray Fields

Given the stark contrast between the stringent field cancellation requirements stated above and the native fields present in the lab, it is clear that we need to do an intelligent job of zeroing the fields. One way around this nasty problem is to use a spin system that is field insensitive (e.g., the clock transition, with RF induced transitions). Although low frequency fields are hard to cancel, by the same token, they are easy to produce, and we accept this trade-off.

Here we outline strategies for reducing the effects of stray fields. The first is the most obvious: avoid magnetic sources. By understanding what devices produce significant fields and how to hunt down sources of unique signals, much frustration can be saved. The next strategy is the standard approach of using magnetic shielding of some kind. However, because our experiment requires a significant amount of three-dimensional optical access, the shielding cannot be a complete solution. Subsequently, we discuss using cancellation coils to both passively and actively negate ambient fields. Short of full-scale feedback, given stable AC noise, one can use precision timing (line-locking) to significantly reduce the effects of that noise. Finally, we discuss in detail how to use our best sensor (the atoms) to periodically zero the field.

A.6.1   Passive

A.6.1.1   Nonmagnetic Mounts

Clearly, one wants to keep the volume close to the atomic sample free from magnetic material and metals. Nonmagnetic metals (aluminum, copper, nonmagnetic steel) are better than magnetic materials, but when ramping large fields, these may cause annoying eddy currents. If possible use nonmagnetic and nonmetallic materials (e.g., teflon, nylon, kevlar, plexiglass) to avoid this. See [193] for more on materials and nonmagnetic steel.

We avoid a lot of these problems by using a quartz cell (ordered from either Techniglass, for our old cell, and Starna, for new cell) rather than a metal chamber (like those from Kimball Physics). One tradeoff here is that one cannot reliably AR coat the inside of the cell. As a compromise we ordered a bunch of AR coated cover slips that we put on the outer faces of the cell with an index matched fluid in between.

A.6.1.2   Shielding

Given noise fields that are unavoidable, a common technique to cancel fields is to surround the sample or source with shielding material to either keep fields out or contain them. The basic idea of all shielding is that if an ambient field is applied to a high-permeability shielding material, that material will produce its own field that cancels the original. In general, the shielding performance is geometry dependent. For example, when shielding a source, one wants to "wrap" the field lines onto themselves in an intelligent way. For most applications though, one usually just makes as smooth of an enclosed cage as possible around the volume of interest.

The best magnetic shields in existence are superconducting bottles. For example, the superconducting shield in the Kirschvink lab gives a factor of 10^{12} suppression (120 dB). Unfortunately, we cannot use these shields because they do not allow the optical access we need (in addition to being expensive to maintain and bulky). Other shielded rooms have shown shielding factors of around 10^5 at 60 Hz and 500 at DC [194].

We opt for the simpler approach of using high-permeability materials such as mu-metal (from Magnetic Shield Corp.). These come in a variety of thicknesses, but even thin foils give approximately a factor of 3 suppression per sheet. The highest suppression obtained with multilayer mu-metal cans is a few hundred (as in the Kirschvink lab at Caltech). A

single layer mu-metal shield around the fluxgate sensor gives suppression factor of 5 for DC (50 mG to 10 mG) and 8 for AC (400 μG to 50 μG). There exist more refined techniques for simultaneously shielding both DC and AC fluctuations at once by using a multilayer shield with alternating mu-metal (for DC) and aluminum, which damps AC fluctuations because of its conductivity. When creating shields with angular geometries, it is useful to use thin foils (either mu-metal or cobalt alloy foil), especially for corners. Mu-metal, like any other iron alloy, may become saturated if applied to fields that are too large. If this happens, the shield needs to be degaussed. One technique for degaussing is to apply a large fluctuating field powered by a Variac.

If it is possible to have a sample in a room where a minimal number of sources are present, a good strategy is to shield the entire room, keeping as many sources of noise outside as possible. Ritek is one of many companies that specializes in creating shielded rooms.

A.6.1.3   Shields in Lab

In the lab, we have shielded both the ion pump (2 layers) and the glass vacuum cell (4 layers), such that the driving coils are inside the shield. See chapter 11 for more on these shields, including pictures.

A.6.1.4   Passive Coils

Given the large trim coils previously described we can null both the absolute field and the gradient by supplying current to each of the two coils independently. Note that this only provides cancelation for the gradients $\delta B_i/\delta x_j$ where $i = j$, *not* for $i \neq j$.

Because of the NMR/MRI industry, there is a lot of knowledge out there on how to passively cancel gradients and higher-order derivatives of the field using funky shaped coils. Actually it is not so complicated; the shape of the coils are related to each term in a Taylor-like functional decomposition of the field. For more on this, Russ Jacobs at Caltech is a useful resource.

A.6.2 Active

Passive cancellation of the field is of little use if the field is fluctuating with significant amplitude at experimentally important timescales, as we know it is from the above discussion. One possibility is to try to cancel the fluctuations continuously in real time with feedback.

Originally we had intended to do this, along the lines of [183], by placing the magnetoresistive sensors symmetrically around the cell to give a reading of the field at the center. This sensed field was fed back to the small drive coils in real time. As in [183], we succeeded in canceling the 60 Hz fluctuations from about 1 mG p-p to 100 μG, but the signal-to-noise of the sensors did not allow better performance. With any feedback process, there is going to be a bandwidth/precision trade-off, which is inherently related to the signal-to-noise ratio of the sensors.

An additional complication arises if one wants to cancel fluctuations below some frequency, but also apply fast fields that do not get cancelled themselves. I developed a feed-forward style controller, which satisfied this goal for all three channels. Ultimately, we decided to skip this feedback for the sake of simplicity opting to use only passive shielding and line-locking techniques.

A.6.2.1 Line Lock

Assume that there is field noise within the lab, but that it is stable and repeatable with respect to some reference. The 60 Hz noise in our lab is of this kind because, if we trigger on the power line, the phase of the field noise with respect to the line is fixed. If our experiment takes place over very short timescales τ compared to the period of this noise $T = 16.6$ ms, such that $\tau \ll T$ we can take advantage of this fact. Assuming the noise is sinusoidal (with amplitude B_n, if we start every experiment on the side of the fringe then the total change in field during the experiment is the linear $\delta B = B_n(\tau/T)$. Even better, if we start at the top of the fringe (and apply DC field $-B_n$) then the total change is the quadratic $\delta B = B_n(\tau/T)^2/2$. For $B_n = 1$ mG, $T = 16.6$ ms, and $\tau = 100$ μs, this gives $\delta B = 6$ μG for the linear case and $\delta B = 18$ nG for the quadratic case. Clearly, if our timing and triggering are good we can do quite well.

There are a few qualifications to this analysis. The first is that the field in different directions will certainly have different phases, thus if we arrange the triggered phase (or

delay) such that one direction gets the quadratic benefit, the other directions may do as poorly as the linear case. Second, the zeroing here is only as good as the timing of the trigger, thus if the trigger signal has noise on it, this will lead to corresponding variance in the field, which may be worse than the drift calculated. In practice, we use a high signal-to-noise ratio line signal, so we think this is not much of an issue. Finally, it is obvious that some slow sources are not phase locked to the line. If there is a slow sub-Hz noise source that is faster than our DC rezeroing rate, then it can render the line-locking less effective.

Appendix B

Conditional Dynamics of an Atom in a Two-Sided Cavity

This appendix regards the conditional dynamics of a single three-level atom interacting with a Fabry-Perot cavity mode, which leaks out *both* sides of the cavity. There were multiple motivations for investigating this problem. First, I wanted to understand the realistic case of a two-sided cavity, as opposed to a typically idealized one-sided cavity, and confirm the fact that ignoring one output of a symmetric cavity is equivalent to cutting the effective detection efficiency in half. Second, the Kimble lab was at one point considering an experiment that investigated measuring both sides of the cavity and conditioning the internal atomic state accordingly. Third, and most important, these notes provide a nice tutorial example for using the quantum measurement formalism introduced in chapter 2. In particular, the example nicely illustrates the notion of unravellings by considering detecting the dual outputs of the cavity in multiple, physically distinct, ways.

B.1 Abstract

We discuss in tutorial fashion the formalism for describing the conditional evolution of one of the simplest open quantum systems: a driven cavity containing an atom with all output fields potentially monitored in real time with photo-detectors. We emphasize the flexibility of the formalism for modifying the form of the measurement: mixing output fields with themselves or local oscillators, imperfect quantum efficiencies, etc. Examples (with accompanying simulations) are included that emphasize the role of the stochastic master equation as a filter that updates the observer's state of knowledge according to the

measurement record.

B.2 Introduction

A single atom strongly coupled to the electromagnetic field within a high finesse Fabry-Perot cavity is one of the simplest open systems that can be used to study conditional quantum dynamics. Experimental progress is being made in the control of such systems [195, 196] and it is becoming ever more relevant to consider developing the rigorous theoretical tools that will help the experimentalist optimally map the measurement record to the proper state description of the joint intracavity quantum system.

This appendix is meant to serve as a tutorial demonstration of the quantum conditioning process, using a generalized jump operator formalism, with the atom-cavity system as the guiding example. With an experimental perspective in mind, we aim to derive the stochastic master equation that the observer will use to update his state of knowledge optimally as measurement results arrive. Much of foundational work summarized here can be found in several references [52, 69], notably the thesis of Wiseman [57].

With any quantum system, even the optimal state of knowledge will be incomplete in that predictive uncertainty about some potential measurements will remain. Learning how to manage this uncertainty is the primary objective of quantum filtering and control. Although the estimation equations discussed here will always apply, they are perhaps most useful when the observer wants to control the system in real time with a system parameter that is made a function of the optimal estimate. However, we do not discuss feedback directly here, as we are mostly concerned with the experimental details of observation and conditioning dynamics.

This appendix is organized as follows. In section B.3, we outline the components of the model to be subsequently analyzed, including the relevant atom and cavity parameters. Section B.4 provides the derivation of the stochastic master equation that serves as the filter from measurement results to optimal state description using jump operator notation. This includes imperfect photo-detection and reformatting the output beams to give different unravellings of the unconditional dynamics. We begin the analysis of the conditioning equation by discussing the special unconditional case in B.5. Finally, in section B.6 we examine special cases of the conditional dynamics. This includes the conditioning equations
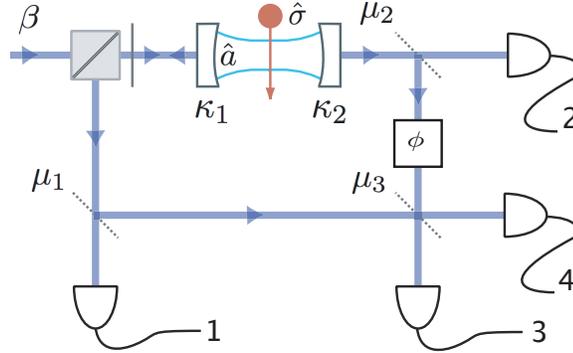
Figure B.1: An experimental schematic very similar to the scheme of [197]. The parameter β represents the strength of the input coherent state, which is coupled into the cavity with a polarizing beamsplitter and a rotating waveplate to allow for the back-reflection to be directed to detector 1. The μ_i are the reflectivities of the beam-splitters and the κ_i are the optical decay rates out of either mirror. The cavity mode is described with operator $\hat{a}$ and the atom with operator $\hat{\sigma}$. In the text, we consider two simulations, one with $\mu_1 = \mu_2 = 0$ and one with $\mu_1 = \mu_2 = 1$, $\mu_3 = 0.5$. In the latter case, we allow the freedom to adjust the phase ϕ of the cavity output light.

for when one output channel is ignored and also for when the two output channels are subsequently mixed on a beam-splitter, changing the effect of conditioning substantially and creating a "which way" (or "welcher weg") type of interferometer [197].

Each simulation used in this appendix relies on the quantum optics toolbox and its stochastic integration routines [99].

B.3   System

In this section, we describe the components of our system of interest in preparation for the conditional equations derived in the next section. Each component is displayed in the schematic of figure B.1, which is similar to that from [197].

B.3.1   Cavity

The high finesse Fabry-Perot cavity at the center of the schematic contains the field that will be one component of our quantum system, the other being an atom's internal state. The two mirrors of the cavity are coupled to the environment modes by the rates κ_1 and κ_2, which we do not necessarily specialize to either the symmetric or one-sided configuration. If the mirrors are somewhat transmissive, the cavity modes can be both driven and measured.

We will not model absorptive loss from the mirrors for simplicity, although this can be included in a quantum efficiency term. The cavity is of length L so the spacing between modes, or free spectral range, is $\mathtt{FSR} = c/2L$. The inclusion of multiple modes would allow for the possibility of pulse propagation within the cavity. For simplicity we assume that all relevant rates, detunings, and bandwidths are much smaller than the $\mathtt{FSR}$ so that only the mode with frequency ω_C nearest the atomic frequency ω_A and the driving laser frequency ω_L will ever be populated. This is a good approximation for physically realizable cavities.

The annihilation operator for the cavity mode will be labeled $\hat{a}$ and the free Hamiltonian of the cavity is simply $\hat{H}_C = \omega_C \hat{a}^\dagger \hat{a}$, which, in the rotating frame oscillating at ω_L, becomes

$$\hat{H}_C \;=\; (\omega_C - \omega_L)\hat{a}^\dagger \hat{a} \equiv \Delta_C \hat{a}^\dagger \hat{a}. \tag{B.1}$$

The ω_L term comes from the time derivative of the density matrix $\hat{\rho}$ when the entire master equation is transformed into the rotating frame as follows.

B.3.1.1 Rotating Frame

In a frame rotating at the laser frequency ω_L we transform a general operator $\hat{X}$ to $\hat{X}_R$ by

$$\hat{X}_R \;=\; \exp[i\omega_L t(\hat{a}^\dagger \hat{a} + \hat{\sigma}^\dagger \hat{\sigma})]\hat{X}\exp[-i\omega_L t(\hat{a}^\dagger \hat{a} + \hat{\sigma}^\dagger \hat{\sigma})]. \tag{B.2}$$

For the cavity mode and atomic lowering operators this gives

$$\hat{a}_R \;=\; \exp[i\omega_L t(\hat{a}^\dagger \hat{a} + \hat{\sigma}^\dagger \hat{\sigma})]\hat{a}\exp[-i\omega_L t(\hat{a}^\dagger \hat{a} + \hat{\sigma}^\dagger \hat{\sigma})] \tag{B.3}$$

$$\;=\; \exp[i\omega_L t\hat{a}^\dagger \hat{a}]\hat{a}\exp[-i\omega_L t\hat{a}^\dagger \hat{a}] \tag{B.4}$$

$$\;=\; \exp[-i\omega_L t]\hat{a} \tag{B.5}$$

$$\hat{\sigma}_R \;=\; \exp[i\omega_L t(\hat{a}^\dagger \hat{a} + \hat{\sigma}^\dagger \hat{\sigma})]\hat{\sigma}\exp[-i\omega_L t(\hat{a}^\dagger \hat{a} + \hat{\sigma}^\dagger \hat{\sigma})] \tag{B.6}$$

$$\;=\; \exp[i\omega_L t\hat{\sigma}^\dagger \hat{\sigma}]\hat{\sigma}\exp[-i\omega_L t\hat{\sigma}^\dagger \hat{\sigma}] \tag{B.7}$$

$$\;=\; \exp[-i\omega_L t]\hat{\sigma}. \tag{B.8}$$

Where the commutation relations for either can be used to show the final identity [68]. We also have $\hat{a}_R^\dagger = \exp[i\omega_L t]\hat{a}^\dagger$ and $\hat{\sigma}_R^\dagger = \exp[i\omega_L t]\hat{\sigma}^\dagger$, $\hat{a}_R^\dagger \hat{a}_R = \hat{a}^\dagger \hat{a}$, and $\hat{\sigma}_R^\dagger \hat{\sigma}_R = \hat{\sigma}^\dagger \hat{\sigma}$. Throughout this appendix, we implicitly work in the rotated frame and remove the R

subscript.

B.3.2 Atom

During the course of our measurements, we allow a single atom to interact with the cavity mode. For illustrative purposes, we consider a three level atom with an excited state $|e\rangle$, a long-lived ground state $|g_c\rangle$, which is coupled to the excited state by the cavity light, and another stable ground state $|g_u\rangle$, which is uncoupled to any field.

The annihilation operator for this atom will be labeled $\hat{\sigma}$ (where $|g_c\rangle = \hat{\sigma}|e\rangle$) and the free Hamiltonian of the atom is simply $\hat{H}_A = \omega_A \hat{\sigma}^\dagger \hat{\sigma}$, which, in the rotating frame oscillating at ω_L, becomes

$$\hat{H}_A = (\omega_A - \omega_L)\hat{\sigma}^\dagger \hat{\sigma} \equiv \Delta_A \hat{\sigma}^\dagger \hat{\sigma}. \tag{B.9}$$

Again the extra term comes from the transformation of the time derivative of $\hat{\rho}$ in the full master equation.

Under the rotating wave approximation (RWA) we allow the atom to interact with the field mode via the Hamiltonian (in the rotating frame)

$$\hat{H}_J = ig(\hat{\sigma}^\dagger \hat{a} - \hat{\sigma}\hat{a}^\dagger) \tag{B.10}$$

where g is mode volume dependent coupling constant. Now we define the system Hamiltonian (ignoring bath couplings) as

$$\hat{H}_0 = \hat{H}_C + \hat{H}_A + \hat{H}_J \tag{B.11}$$

$$= \Delta_C \hat{a}^\dagger \hat{a} + \Delta_A \hat{\sigma}^\dagger \hat{\sigma} + ig(\hat{\sigma}^\dagger \hat{a} - \hat{\sigma}\hat{a}^\dagger). \tag{B.12}$$

In practice there are two simple situations to consider with regard to the relative positions of the atom and field mode. In the first, we simply pin down the position of the atom such that g is a constant, which is becoming more realistic for long experimental timescales [195]. In the second we allow the atom to drift through the field mode in some way (as in [196]). In this case one could imagine either estimating the position of the atom concurrently with its internal state (see [198] for more on quantum position tracking) or removing the position from the estimation by approximating it as a perfectly known classical variable,

which is currently unrealistic. In later sections, we consider the atom to drift through the mode in a known way such that g is a Gaussian in time. In practice, without trapping the atom, the incoming trajectory is random but approximately Gaussian. We also ignore light atom forces and axial g variations to focus on the conditioning physics.

The atom will not couple solely to the cavity mode because it can spontaneously emit radiation out the side of the cavity. The rate at which this emission happens is labeled γ. Just as the mirrors leak information at the rates κ_i, this provides another output channel and its effect on the dynamics will be described shortly.

We assume all of the output channels to be destructively measured (whether or not the results are ignored), and our description of the system that we wish to update according to these measurements includes both the field mode and the atom, $\hat{\rho} = \hat{\rho}_{AC}$. We also assume the initial field and atomic states to be uncorrelated such that

$$\hat{\rho}(0) = \hat{\rho}_A(0) \otimes \hat{\rho}_C(0). \tag{B.13}$$

The coupling Hamiltonian $\hat{H}_J$ will produce entanglement between the two such that $\hat{\rho}(t)$ cannot be written this way for all times.

In principle, the field mode has a basis of infinite number, but in practice the driving field can only populate so many number states so that in simulation we can get away with using only N_A Fock states and the total density matrix will have size $(3 \times N_A) \times (3 \times N_A)$

B.3.3 Environment Coupling

No system is completely closed. If one were, we would not be able to talk about it. Here we describe the interaction of the field modes with the environment field modes. The interaction of the atom with the noncavity modes it spontaneously emits into is similar and the result will simply be written down later. The input state to the mirrors will always be coherent states, with *known* complex amplitude, which are by definition unentangled with any other system. Mostly, we will only consider driving the cavity from side 1 with a coherent state $\beta_1(t)$, but for generality we initially consider also driving the other side with $\beta_2(t)$, which can be set to zero subsequently. Also we assume the environment to be roughly at zero temperature, which is a good approximation at optical frequencies.

Thus in timestep dt, on either side we let the coherent states of the the traveling envi-

ronment modes interact with the cavity mode via the evolution operator

$$\hat{U}(dt) \;=\; \exp[-i\hat{H}_0 dt + \sqrt{dt}(\sqrt{\kappa_1}[\hat{b}_1^\dagger \hat{a} - \hat{a}^\dagger \hat{b}_1] + \sqrt{\kappa_2}[\hat{b}_2^\dagger \hat{a} - \hat{a}^\dagger \hat{b}_2])]. \qquad \text{(B.14)}$$

In the Heisenberg picture, application of $\hat{U}(dt)$ leads to the evolution of the input field operators $\hat{b}_{i,in}$. Expanding to first-order in dt and using the commutation relation for $\hat{b}$, the input evolves into the output field operator as

$$\hat{b}_{i,out} \;=\; \hat{U}(dt)^\dagger \hat{b}_{i,in} \hat{U}(dt) \qquad \text{(B.15)}$$

$$\;=\; \sqrt{\kappa_i}\hat{a} + \hat{b}_{i,in}. \qquad \text{(B.16)}$$

After bouncing off the cavity mirror, the output traveling mode is no longer independent of the intracavity state because it depends on $\hat{a}$.

B.3.4   Detection Scheme

Even though we can determine the field operators emanating from each cavity mirror in terms of the cavity mode operator and the input field operators, we are not forced to measure these fields directly. There are many options on how to perform the measurement of either field. One could perform any combination of mixing the fields with each other, mixing with local oscillators coherent fields, and phase shifting before doing direct photon detection at the end of the optical setup. Just as the driving fields are coherent states, we also imagine all input states to beam splitters to be coherent states (possibly vacuum) uncorrelated to anything else.

Of course, the unconditional master equation must be invariant to any of this "reformatting" of the output fields because the traveling modes interact with the cavity in a Markovian manner. How this happens in detail will be discussed in the next section. The important point is that some set of transformations of the measurement operators, which leave the unconditional evolution invariant, may correspond to a realistic physical way of analyzing the output fields. The physical unravelling that is chosen in the lab will obviously have important consequences for the conditional dynamics.

B.4 Constructing the Conditional Quantum Filter

Because the unconditional dynamics are a special case of the conditional description, we begin by formulating the procedure for updating $\hat{\rho}(t)$ based on continuous projective measurement of the formatted output fields. In the next section, we detail example situations of varying levels of conditioning, from complete ignorance of the results to perfect efficiency.

B.4.1 Detectors 1 and 2

Imagine the following procedure for deriving the conditional dynamics for the atom cavity system. Each timestep of length dt we let a different spatially localized traveling mode interact with the system from either side of the cavity through $\hat{U}(dt)$. After dt has elapsed we make a projective measurement of each output channel with detectors 1 and 2 without any optical adjustment of the beams. We make dt small enough such that there are three possibilities, zero photons in each output channel, or one photon in either (but not both). (We take for granted the highly nontrivial association of a photo-detection event with projective number measurement of the traveling mode.) The conditional state of the atom and cavity mode before dt is labeled by $\hat{\rho}(t)$, and we let the traveling modes on either side be in the coherent states, $|\beta_1(t)\sqrt{dt}\rangle_1$ and $|\beta_2(t)\sqrt{dt}\rangle_2$. These states are unentangled with either the system state or any other state. The final total state is then described by

$$\hat{\rho}_T(t + dt) \;=\; \hat{U}(dt)\hat{\rho}_T(t)\hat{U}^\dagger(dt) \tag{B.17}$$

$$\;=\; \hat{U}(dt)\left[\hat{\rho}(t) \otimes \hat{\rho}_E(t)\right]\hat{U}^\dagger(dt) \tag{B.18}$$

where the environment state is

$$\hat{\rho}_E(t) \;\equiv\; |\beta_1(t)\sqrt{dt}\rangle_1|\beta_2(t)\sqrt{dt}\rangle_2\langle\beta_1(t)\sqrt{dt}|_1\langle\beta_2(t)\sqrt{dt}|_2. \tag{B.19}$$

If we label the results of the subsequent measurement by 0 for no clicks on either output, 1 for a click on detector 1 but not 2, and 2 for a click on detector 2 but not 1. Then the state of the atom and cavity prior to the next dt is given by

$$\hat{\rho}(t + dt) = \frac{\hat{\Omega}_i(dt)\hat{\rho}(t)\hat{\Omega}_i^\dagger(dt)}{\text{Tr}[\hat{\Omega}_i(dt)\hat{\rho}(t)\hat{\Omega}_i^\dagger(dt)]} \tag{B.20}$$

309

where from the projective measurement rule, we have

$$\hat{\Omega}_0(dt) = \langle 0|_1 \langle 0|_2 \hat{U}(dt)|\beta_1(t)\sqrt{dt}\rangle_1|\beta_2(t)\sqrt{dt}\rangle_2 \tag{B.21}$$

$$\hat{\Omega}_1(dt) = \langle 1|_1 \langle 0|_2 \hat{U}(dt)|\beta_1(t)\sqrt{dt}\rangle_1|\beta_2(t)\sqrt{dt}\rangle_2 \tag{B.22}$$

$$\hat{\Omega}_2(dt) = \langle 0|_1 \langle 1|_2 \hat{U}(dt)|\beta_1(t)\sqrt{dt}\rangle_1|\beta_2(t)\sqrt{dt}\rangle_2. \tag{B.23}$$

To get the actual form of these jump operators, we expand $\hat{U}(dt)$ and the coherent state to first-order in dt

$$|\beta_i(t)\sqrt{dt}\rangle_i \approx (1 - |\beta_i|^2 dt/2)|0\rangle_i + \beta_i\sqrt{dt}|1\rangle_i \tag{B.24}$$

to get

$$\hat{\Omega}_0(dt) = 1 - iH_0 dt - ((\kappa_1 + \kappa_2)\hat{a}^\dagger \hat{a} dt/2 \tag{B.25}$$

$$-(\sqrt{\kappa_1}\beta_1 + \sqrt{\kappa_2}\beta_2)\hat{a}^\dagger dt - (|\beta_1|^2 + |\beta_2|^2)dt/2) \tag{B.26}$$

$$\hat{\Omega}_1(dt) = (\sqrt{\kappa_1}\hat{a} + \beta_1)\sqrt{dt} \tag{B.27}$$

$$\hat{\Omega}_2(dt) = (\sqrt{\kappa_2}\hat{a} + \beta_2)\sqrt{dt}. \tag{B.28}$$

The probability for getting the result i is simply

$$P_i = \text{Tr}[\hat{\Omega}_i(dt)\hat{\rho}(t)\hat{\Omega}_i^\dagger(dt)] \tag{B.29}$$

and by conservation of probability we have

$$\Sigma_{i=0}^2 \hat{\Omega}_i^\dagger(dt)\hat{\Omega}_i(dt) = I. \tag{B.30}$$

Notice that, in the case where no photons are detected, we still update the state in a non-trivial manner because any result corresponds to a gain of information. The unconditional evolution is the incoherent sum of all possibilities

$$\hat{\rho}(t + dt) = \Sigma_{i=0}^2 \hat{\Omega}_i(dt)\hat{\rho}(t)\hat{\Omega}_i(dt)^\dagger. \tag{B.31}$$

B.4.2 Formatting the Measurement: Output Channel Division and Mixing

The collapse operators are the same as the output field operators (with the $\hat{b}$ operator replaced by the number β). So by manipulating the output fields we simultaneously manipulate the collapse operators and, if the fields are eventually detected, we can change the information content of detection event. We might imagine "formatting" the output fields in several standard ways: dividing with beam-splitters, adding phase shifters, mixing with coherent local oscillators for homodyne/heterodyne rather than direct photo-detection, etc. We assume all added beams to be in known coherent states, uncorrelated with anything and in the vacuum state unless otherwise stated. For example, we could split up the fields with beam-splitters with reflectivities μ_i as in the figure B.1, giving us four measurement channels. Assuming now that $\beta_1 = \beta \neq 0$ and $\beta_2 = 0$, the corresponding jump operators then become

$$
\begin{aligned}
\hat{\Omega}_0(dt) &= 1 - iH_0 dt - (\kappa_1 + \kappa_2)\hat{a}^\dagger \hat{a} dt/2 \\
&\quad - \sqrt{\kappa_1}\beta \hat{a}^\dagger dt - |\beta|^2 dt/2) & \text{(B.32)} \\
\hat{\Omega}_1(dt) &= \sqrt{1-\mu_1}(\sqrt{\kappa_1}\hat{a} + \beta)\sqrt{dt} & \text{(B.33)} \\
\hat{\Omega}_2(dt) &= \sqrt{1-\mu_2}\sqrt{\kappa_2}\hat{a}\sqrt{dt} & \text{(B.34)} \\
\hat{\Omega}_3(dt) &= [\sqrt{\mu_1}\sqrt{\mu_3}(\sqrt{\kappa_1}\hat{a} + \beta) \\
&\quad + \sqrt{1-\mu_3}\sqrt{\mu_2}\sqrt{\kappa_2}\exp[i\phi]\hat{a}]\sqrt{dt} & \text{(B.35)} \\
\hat{\Omega}_4(dt) &= [\sqrt{\mu_2}\sqrt{\mu_3}\sqrt{\kappa_2}\exp[i\phi]\hat{a} \\
&\quad - \sqrt{1-\mu_3}\sqrt{\mu_1}(\sqrt{\kappa_1}\hat{a} + \beta)]\sqrt{dt} & \text{(B.36)}
\end{aligned}
$$

where we have used the general beam-splitter relations for mapping input fields to output fields.

B.4.3 Formatting the Measurement: Adding a Local Oscillator

Besides dividing and mixing the fields, there exist other (unitary) transformations of the collapse operators under which the unconditional evolution is invariant. Some of these transformations correspond to physical ways of analyzing the fields emanating from the cavity. For example, we could add a coherent field γ_i $(i \neq 0)$ to each output channel by

reflecting the beam off of a mirror with infinitesimal transmission with a strong driving beam on the other input port. Measuring the photon number of the resulting beam corresponds to either a homodyne or heterodyne measurement depending on how the phase of γ_i is adjusted in time. We can write this transformation as

$$
\begin{aligned}
\hat{\tilde{\Omega}}_i(dt) &= \hat{\Omega}_i(dt) + \gamma_i \sqrt{dt} & \text{(B.37)} \\
\hat{\tilde{\Omega}}_0(dt) &= \hat{\Omega}_0(dt) - \Sigma_{i=1}^N (\gamma_i^* \hat{\Omega}_i(dt)\sqrt{dt} + |\gamma_i|^2 dt/2) & \text{(B.38)}
\end{aligned}
$$

where γ_i is a complex number representing the known coherent state of the local oscillator. The operator $\hat{\tilde{\Omega}}_0(dt)$ can be derived by demanding the unconditional master equation is independent of γ_i, i.e.,

$$
\Sigma_{i=0}^2 \hat{\Omega}_i(dt)\hat{\rho}(t)\hat{\Omega}_i(dt)^\dagger = \Sigma_{i=0}^2 \hat{\tilde{\Omega}}_i(dt)\hat{\rho}(t)\hat{\tilde{\Omega}}_i(dt)^\dagger. \tag{B.39}
$$

When the collapse operators are written out, there is a superficial similarity between the role of the driving β and the local oscillator γ, but the fact that one appears in $\hat{\tilde{\Omega}}_0(dt)$ with an annihilation operator and the other appears with a creation operator ensures that only the driving appears in the unconditional evolution. Because the addition of the local oscillator is downstream it must have no physical effect on the unconditional evolution.

In actuality, we would perform a balanced homodyne or heterodyne measurement with a 50/50 beam-splitter, thus increasing the number of output channels and adding a local oscillator.

B.4.4 Imperfect Efficiencies

Now consider what the appropriate conditioning procedure should be when each output channel is detected with an efficiency η_i, which is not necessarily unity. Proceeding every dt again, if we get a click on channel $i \neq 0$ then the appropriate procedure is to collapse the state with the appropriate operator as before

$$
\hat{\rho}(t + dt) = \frac{\hat{\Omega}_i(dt)\hat{\rho}(t)\hat{\Omega}_i^\dagger(dt)}{\text{Tr}[\hat{\Omega}_i(dt)\hat{\rho}(t)\hat{\Omega}_i^\dagger(dt)]} \tag{B.40}
$$

but now the probability of this event is

$$P_i \;\;=\;\; \eta_i \mathrm{Tr}[\hat{\Omega}_i(dt)\hat{\rho}(t)\hat{\Omega}_i^\dagger(dt)]. \tag{B.41}$$

When no click is measured in dt, we should consider both the possibility that no photon should have been detected $\hat{\Omega}_0(dt)$ and the possibility that a photon should have been detected but was not due to the poor detector. So if no click is registered we apply

$$\hat{\rho}(t+dt) \;\;=\;\; \frac{\hat{\Omega}_0(dt)\hat{\rho}(t)\hat{\Omega}_0^\dagger(dt) + \Sigma_{i=1}^{N}(1-\eta_i)\hat{\Omega}_i(dt)\hat{\rho}(t)\hat{\Omega}_i^\dagger(dt)}{\mathrm{Tr}[\hat{\Omega}_0(dt)\hat{\rho}(t)\hat{\Omega}_0^\dagger(dt) + \Sigma_{i=1}^{N}(1-\eta_i)\hat{\Omega}_i(dt)\hat{\rho}(t)\hat{\Omega}_i^\dagger(dt)]}. \tag{B.42}$$

Of course the average behavior is still described by the unconditional master equation.

B.4.5 Stochastic Master Equation

We have completely described the basic rules for updating the conditional density matrix stepwise in time. Now we collect these rules into one equation, henceforth referred to as the stochastic master equation:

$$
\begin{aligned}
\hat{\rho}(t+dt) \;\;=\;\; & \hat{\rho}(t) + d\hat{\rho}(t) \\
=\;\; & \hat{\Omega}_0(dt)\hat{\rho}(t)\hat{\Omega}_0^\dagger(dt) \\
& +\Sigma_{i=1}^{N}(1-\eta_i)\hat{\Omega}_i(dt)\hat{\rho}(t)\hat{\Omega}_i^\dagger(dt) \\
& +\Sigma_{i=1}^{N}dN_i\left[\frac{\hat{\Omega}_i(dt)\hat{\rho}(t)\hat{\Omega}_i^\dagger(dt)}{\mathrm{Tr}[\hat{\Omega}_i(dt)\hat{\rho}(t)\hat{\Omega}_i^\dagger(dt)]} - \hat{\rho}(t)\right] \\
& +\Sigma_{i=1}^{N}E(dN_i)\hat{\rho}(t)
\end{aligned}
\tag{B.43}
$$

where dN_i is either zero or one and becomes one in an infinitesimal time increment dt with probability

$$E(dN_i) = \eta_i \mathrm{Tr}[\hat{\Omega}_i(dt)\hat{\rho}(t)\hat{\Omega}_i^\dagger(dt)] \propto dt. \tag{B.44}$$

In equation (B.43), the $-\hat{\rho}(t)$ term is to cancel the $\hat{\rho}(t)$, which comes from the $\hat{\Omega}_0(dt)$ term when a jump is made. In turn, the final term is to cancel that term when the average evolution is taken such that the unconditional master equation is obtained. Note that taking all η_i to zero also returns the deterministic unconditional master equation. Thus there are two ways to get an unconditional solution, by averaging stochastic trajectories

with some $\eta_i \neq 0$ or simply evolving the above equation with all $\eta_i = 0$.

It is important to emphasize that the dN_i represent the physical measurement record. If our model was wrong, then the statistics of each dN_i would be different and, if we were smart enough to recognize this, we would alter our model (not the state) accordingly. Thus in simulation, one could imagine producing dN_i with the optimal filter and feeding these results to a less perfect observer with a nonoptimal filter. Here, however, we always assume complete knowledge of all system parameters and our only sources of ignorance are possibly the initial state (pure versus mixed) and the inherent randomness of the measurement.

In the case where the outputs are measured by homodyne/heterodyne detection one could use the $\tilde{\hat{\Omega}}_i(dt)$ jump operators described previously and take the large γ_i limit such that the dN_i noise is replaced with a white noise dW_i term [57]. Also notice that the transformation into a rotating frame only adds the energy terms to the SME because the collapse operators always appear in pairs so exponential terms cancel.

B.5 Unconditional Dynamics

At this point, we have derived the conditional master equation (B.43) that will be the focus of the rest of this appendix. We interpret it as an optimal filtering equation, always imagining an experimentalist updating his description of the system $\hat{\rho}(t)$ as the measurement results arrive. Before considering conditional situations, we first consider the case where this observer's detectors are all broken. Under these dire circumstances, the best the observer can do is set all $\eta_i = 0$ in the filtering equation and update $\hat{\rho}(t)$ with the resulting unconditional master equation.

B.5.1 Steady-States

Setting $\eta_i = 0$ for all output channels we get the deterministic unconditional master equation. We calculate the steady-states of the unconditional evolution by solving $d\hat{\rho}_{ss} = 0$. From $\hat{\rho}_{ss}$ we can calculate the average steady-state expectation value of any operator related quantity. For example, we can look at the average expected photon counting rate for the output channels in steady-state as a function of the detuning, Δ_C. Figure B.2 displays

this spectrum, where the reflected and transmitted powers are respectively

$$P_{R,ss} = \langle \hat{\Omega}_1^\dagger \hat{\Omega}_1 \rangle_{ss} = \langle (\sqrt{\kappa_1}\hat{a} + \beta_1)^\dagger (\sqrt{\kappa_1}\hat{a} + \beta_1) \rangle_{ss} \qquad \text{(B.45)}$$

$$P_{T,ss} = \langle \hat{\Omega}_2^\dagger \hat{\Omega}_2 \rangle_{ss} = \kappa_2 \langle \hat{a}^\dagger \hat{a} \rangle_{ss}. \qquad \text{(B.46)}$$

The steady-state transmission is a Lorentzian function of the detuning, attaining unity on resonance.

Now we can ask what happens when an atom is pinned within the cavity. Because the uncoupled ground state does not interact with the system, we must first state what fraction of population is in this state before calculating the steady-state. If the atom is completely in the uncoupled ground state, then the spectrum is unchanged from the central peaks of figure B.2. If the cavity contains an atom with initially all population in the coupled state, the steady-state of the entire system splits the transmission peak by the amount g. To conserve energy, we must also consider the power emitted from the atom via spontaneous emission

$$P_{SE,ss}(t) = \langle \hat{\Omega}_{SE}^\dagger \hat{\Omega}_{SE} \rangle_{ss} = \gamma \langle \hat{\sigma}^\dagger \hat{\sigma} \rangle_{ss}. \qquad \text{(B.47)}$$

If the atom is in either an initial coherent or incoherent superposition of ground states, then the transmission/reflection spectrum is the obvious weighted superposition of the curves in figure B.2.

B.5.2   Falling Atoms

In this section we consider the following unconditional, non-steady-state scenario. The cavity is initially empty (in the vacuum state) and the atom is prepared in a superposition of the coupled and uncoupled ground states far outside the cavity (hence $g(0) \approx 0$). At time zero, the input light is turned on (steplike) and after some time ($t \gg 1/\kappa$) the cavity reaches steady-state. Some time later we let the atom fall through the cavity mode with a Gaussian profile, which we model by simply letting g vary in time with a Gaussian profile. Again, we assume the position of the atom is known by the observer in possession of the SME.
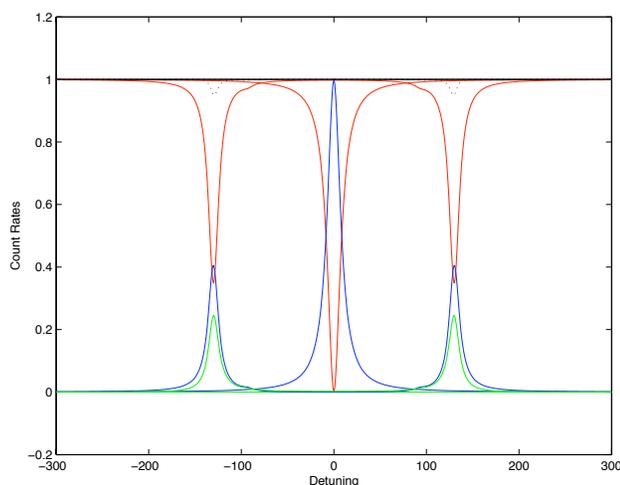
Figure B.2: Transmission and reflection spectra, with and without atom. When the cavity is empty, the transmission (blue) at zero detuning is one and the reflectivity (red) is zero. When the atom is added these peaks split and some power is lost to spontaneous emission (green).
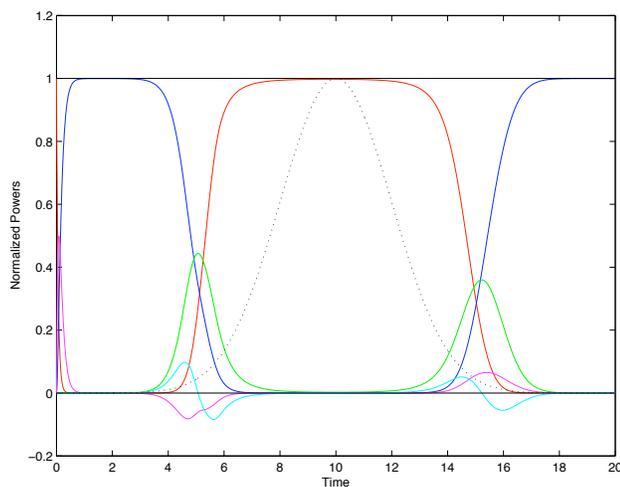
Figure B.3: Power levels as an atom traverses the cavity mode with initial atomic state $|\Psi\rangle(0) = |2\rangle$ (coupled ground state). After a small time the empty cavity achieves equilibrium. Subsequently, the atomic g increases as the atom falls through the cavity mode (dashed line). When the atom is in the cavity, the system switches from fully transmitting to fully reflecting. Spontaneous emission (green) only occurs when g is in the weak coupling regime; when the atom becomes strongly coupled all light is rejected from the cavity and no excited state population is possible. Notice the time rate of change of the internal energy is needed to conserve energy (cyan is atomic energy change $dE_A(t)/dt$ and magenta is light energy change $dE_C(t)/dt$).
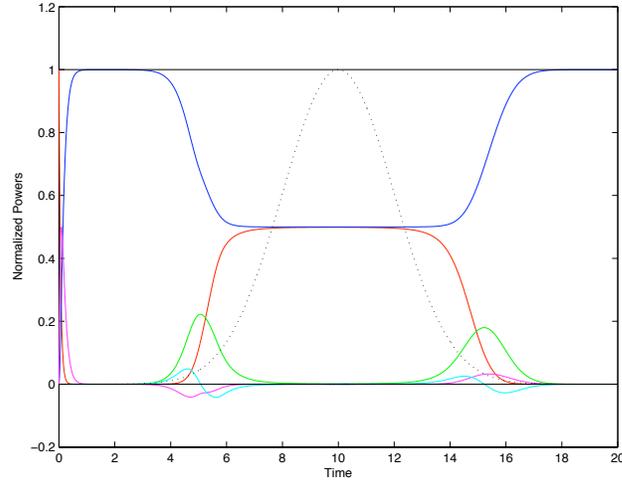
Figure B.4: With initial atomic state $|\Psi\rangle(0) = (|2\rangle + |3\rangle)/\sqrt{2}$ or $\hat{\rho}(0) = (|2\rangle\langle2| + |3\rangle\langle3|)/2$. The final atomic state will be $\hat{\rho}_F = (|2\rangle\langle2| + |3\rangle\langle3|)/2$.

As the atom enters and leaves, the average energy must be conserved, so that

$$P_{in}(t) = P_R(t) + P_T(t) + P_{SE}(t) + dE_A(t)/dt + dE_C(t)/dt \tag{B.48}$$

where

$$P_{in}(t) = |\beta_1(t)|^2 \tag{B.49}$$

$$P_R(t) = \langle\hat{\Omega}_1^\dagger\hat{\Omega}_1\rangle(t) = \langle(\sqrt{\kappa_1}\hat{a} + \beta_1)^\dagger(\sqrt{\kappa_1}\hat{a} + \beta_1)\rangle(t) \tag{B.50}$$

$$P_T(t) = \langle\hat{\Omega}_2^\dagger\hat{\Omega}_2\rangle(t) = \kappa_2\langle a^\dagger\hat{a}\rangle(t) \tag{B.51}$$

$$P_{SE}(t) = \langle\hat{\Omega}_{SE}^\dagger\hat{\Omega}_{SE}\rangle(t) = \gamma\langle\hat{\sigma}^\dagger\hat{\sigma}\rangle(t) \tag{B.52}$$

$$E_A(t) = \langle\hat{\sigma}^\dagger\hat{\sigma}\rangle(t) \tag{B.53}$$

$$E_C(t) = \langle\hat{a}^\dagger\hat{a}\rangle(t). \tag{B.54}$$

The latter two terms were needed to account for all energy sources and sinks during the transient times. In steady-state, of course, the above condition holds with $dE_A/dt = dE_C/dt = 0$. Figure B.3 displays the evolution of these energy rates as the atom falls through the cavity initially in the coupled state. After the atom has left, it is again in the pure coupled ground state.

If the atom is initially in a coherent superposition of the coupled and uncoupled ground

states

$$|\Psi\rangle(0) \;=\; (|2\rangle + |3\rangle)/\sqrt{2} \qquad\qquad (B.55)$$

or in the incoherent completely mixed state

$$\hat{\rho}(0) \;=\; (|2\rangle\langle 2| + |3\rangle\langle 3|)/2 \qquad\qquad (B.56)$$

then we get figure B.4 and the final state is the completely mixed state

$$\hat{\rho}_F \;=\; (|2\rangle\langle 2| + |3\rangle\langle 3|)/2. \qquad\qquad (B.57)$$

B.6   Conditional Dynamics

In the unconditional case, our ignorance during the measurement leads to a loss of information. The experimentalist's job is to fight this condition and gather as much information as possible emanating from the system of interest. In the ideal case, all quantum efficiencies are one, and a pure quantum state description is eventually attained. Before considering the dynamics in this case, we notice a few simple properties of the conditional master equation when certain output channels are ignored but others are not.

B.6.1   Partial Ignorance

Let us restrict ourselves to the case where direct photon detection is performed with detectors 1 and 2 with only one side driven ($\beta_1 \neq 0$ and $\beta_2 = 0$). Whether or not atoms are inside the cavity is inconsequential in the following argument. The stochastic master

equation for this situation above can be rewritten in the form

$$
\begin{aligned}
d\hat{\rho}(t) \;=\; & -i[\hat{H}, \hat{\rho}(t)]dt \\
& -(1-\eta_1)i\sqrt{\kappa_1}[i\beta_1^* a - i\beta_1 \hat{a}^\dagger, \hat{\rho}(t)]dt \\
& +(1-\eta_1)\mathcal{D}[\sqrt{\kappa_1}\hat{a}]\hat{\rho}(t)dt \\
& +(1-\eta_2)\mathcal{D}[\sqrt{\kappa_2}\hat{a}]\hat{\rho}(t)dt \\
& +\eta_1\mathcal{H}[\kappa_1 \hat{a}^\dagger \hat{a}/2 + \sqrt{\kappa_1}a\beta_1]\hat{\rho}(t)dt \\
& +\eta_2\mathcal{H}[\kappa_2 \hat{a}^\dagger \hat{a}/2]\hat{\rho}(t)dt \\
& +\mathcal{G}[\sqrt{\eta_1}(\sqrt{\kappa_1}\hat{a} + \beta_1)]\hat{\rho}(t)dN_1 \\
& +\mathcal{G}[\sqrt{\eta_2}(\sqrt{\kappa_2}\hat{a})]\hat{\rho}(t)dN_2
\end{aligned}
\tag{B.58}
$$

where dN_i are as before and we have used the definitions

$$
\mathcal{D}[\hat{c}]\hat{\rho} \;\equiv\; \hat{c}\hat{\rho}\hat{c}^\dagger - (\hat{c}^\dagger \hat{c}\hat{\rho} + \hat{\rho}\hat{c}^\dagger \hat{c})/2
\tag{B.59}
$$

$$
\mathcal{H}[\hat{c}]\hat{\rho} \;\equiv\; \hat{c}\hat{\rho} + \hat{\rho}\hat{c}^\dagger - \langle \hat{c} + \hat{c}^\dagger \rangle \hat{\rho}
\tag{B.60}
$$

$$
\mathcal{G}[\hat{c}]\hat{\rho} \;\equiv\; \frac{\hat{c}\hat{\rho}\hat{c}^\dagger}{\mathrm{Tr}[\hat{c}\hat{\rho}\hat{c}^\dagger]} - \hat{\rho}
\tag{B.61}
$$

$$
\langle \hat{c} \rangle \;\equiv\; \mathrm{Tr}[\hat{c}\hat{\rho}].
\tag{B.62}
$$

Notice that, when written in this form, the β_1 term in the $\hat{\Omega}_0$ operator turns into a Hamiltonian driving term. Now if we ignore channel 1 and set $\eta_1 = 0$, we can rewrite this equation as

$$
\begin{aligned}
d\hat{\rho}(t) \;=\; & -i[\hat{H}, \hat{\rho}(t)]dt \\
& -i\sqrt{\kappa_1}[i\beta_1^* \hat{a} - i\beta_1 \hat{a}^\dagger, \hat{\rho}(t)]dt \\
& +(1-\eta_1')\mathcal{D}[\sqrt{\kappa_2'}\hat{a}]\hat{\rho}(t)dt \\
& +\mathcal{G}[\sqrt{\eta_2'}(\sqrt{\kappa_2'}\hat{a})]\hat{\rho}(t)dN_2' \\
& +\eta_2'\mathcal{H}[\kappa_2' \hat{a}^\dagger \hat{a}/2]\hat{\rho}(t)dt
\end{aligned}
\tag{B.63}
$$

where we have used $\eta_2' \equiv \eta_2/(1 + \kappa_1/\kappa_2)$ and $\kappa_2' \equiv \kappa_2(1 + \kappa_1/\kappa_2)$ and $dN_2' = dN_2$.

Notice that we would arrive at exactly the same equation if we had instead started with

the full master equation and taken the limits

$$\eta_1 \rightarrow 0 \tag{B.64}$$

$$\kappa_1 \rightarrow 0 \tag{B.65}$$

$$\beta_1 \rightarrow \beta_1/\sqrt{\kappa_1} \tag{B.66}$$

$$\eta_2 \rightarrow \eta_2' \tag{B.67}$$

$$\kappa_2 \rightarrow \kappa_2' \tag{B.68}$$

where, to keep the driving term the same, we had to increase the input power as $\kappa_1 \rightarrow 0$. Thus we have two equivalent situations, if we choose to ignore the one side in the original configuration, it is as if we were monitoring one side of a completely asymmetric cavity with a single mirror loss rate $\kappa_1 + \kappa_2$ and reduced efficiency $\eta_2' < \eta_2$. This type of equivalence is clearly invariant to the way in which we examine the output channel 2 (e.g., with homodyne or heterodyne detection).

B.6.2   Welcher Weg!

Now we examine a particular application of our stochastic master equation. Consider the "which way" configuration of figure B.1, which is discussed in [197] where the atom-cavity system is proposed as a "quantum gate" for light. As in that paper, we consider a mode of operation where the cavity is off resonance (reflective) and the atom brings it into resonance (transmissive). If the corner mirrors are perfectly transmissive, then we detect only with detectors 1 and 2. When the atom is initially in a coherent superposition of the coupled and uncoupled ground states

$$|\Psi\rangle(0) = (|2\rangle + |3\rangle)/\sqrt{2} \tag{B.69}$$

a burst of transmissive light on detector 2 will tend to project the atom into the coupled state (see figure B.5), while the absence of light or too little light will tend to project the atom into the uncoupled state (see figure B.6). This is intuitive as light in either channel is directly correlated to one or the other ground state.

Now consider the corner mirrors as being perfectly reflective, such that we detect only with detectors 3 and 4 (at the output of the beamsplitter). If we set the phase shift on
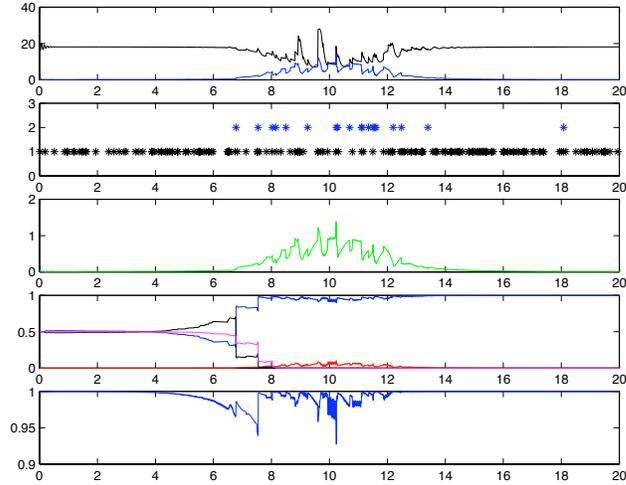
Figure B.5: The conditional dynamics of a superposition state that gets projected into the coupled state. The first plot shows expected counts reflected (black) and transmitted (blue). The second plot shows actual detection events reflected and transmitted. The third plot shows the expected photon number inside the cavity, which is here large (since light is being let through by the coupling atom). The fourth plot shows the populations in the coupled ground state (blue), uncoupled (black) and excited state (red), as well as the coherence between the two ground states (magenta). The final plot shows the entropy of the atom subsystem (a low number represents entanglement between the cavity mode and atomic state).

one arm appropriately ($\phi = 0$), one output port of the final beam-splitter can be made independent of the intracavity state (hence that mode only contains β not $\hat{a}$). For certain parameter regimes, the information gained from the remaining channel will not distinguish between the different ground states and the final state of the atom will not be projected into one or the other, but a superposition of the two (see figure B.7). Still, from the unconditional dynamics we know that the average of all final states must be the completely mixed state, so the final atomic state per trajectory must be projected randomly in some basis. In this configuration, we have lost information on "which way" the light went, but of course this is a meaningless question under the rules of our conditioning.
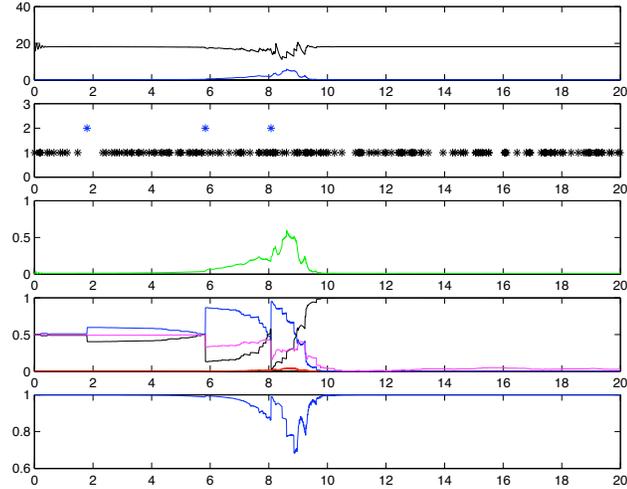
Figure B.6: See figure B.5 for an explanation of the vertical axes. Here not enough clicks were detected to project the atom into the coupled state so it gets projected into the uncoupled state. The estimator "makes this decision" around time 9.
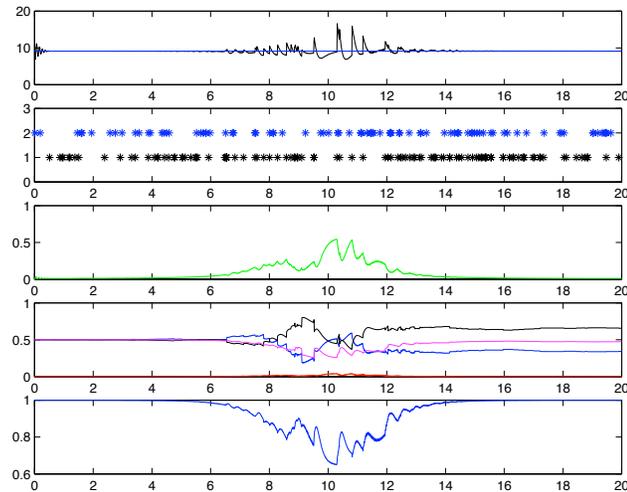
Figure B.7: See figure B.5 for an explanation of the vertical axes. Here the corner mirrors have been made reflective and the clicks on detectors 3 and 4 are shown. Notice that the blue curve is flat in the first plot because that channel is independent of the state. The second plot shows that each detector is getting an approximately equal number of clicks. The fourth plot shows that the system is not projected into either one of the ground states after the atom has passed through.

Appendix C

Conditional Dynamics of an Optical Parametric Oscillator

After the experiment of [12], we briefly considered performing a similar experiment but using squeezed states of light as opposed to the more trivial coherent states. In this case, the squeezed output of an Optical Parametric Oscillator (OPO) would be adaptively measured in order to estimate a parameter describing the squeezed state. The following theoretical notes were written in this context, expanding on the work of [199].

C.1   Introduction

Squeezed states of light are one of the foundational pillars of quantum optics research. Experimental applications of squeezing have either operated in the pulsed or continuous wave (cw) regime. In the cw case, the second-order moments of individual trajectories are constant. When a decaying squeezed state is continuously measured though, the squeezed state will evolve to the vacuum in a nontrivial manner. Much of the groundwork for this situation with a homodyne measurement scheme has already been established in a slightly different context [199]. Here we summarize and extend that work to consider situations where the trajectory formalism is most appropriate. The compact and intuitive trajectory description is shown to be efficient at calculating nonstationary correlations. We also discuss the use of this formalism to derive optimal adaptive measurement schemes in single-shot experiments.

C.2   Conditional Squeezed State Dynamics

The unconditional evolution of light leaking from a cavity is given by the master equation

$$\dot{\rho}(t) = -i[\hat{H}, \hat{\rho}(t)] + \hat{a}\hat{\rho}(t)\hat{a}^\dagger - \frac{1}{2}(\hat{a}^\dagger\hat{a}\hat{\rho}(t) + \hat{\rho}(t)\hat{a}^\dagger\hat{a}) \tag{C.1}$$

where we have assumed the decay rate to be unity. In terms of a quantum trajectory, or jump operator, formalism, the unconditional evolution can be written as

$$\hat{\rho}(t + dt) = \sum_r \hat{\Omega}_r(dt)\hat{\rho}(t)\hat{\Omega}_r^\dagger(dt). \tag{C.2}$$

One choice of measurement operators consistent with equation (C.1) is

$$\hat{\Omega}_1(dt) = \hat{a}\sqrt{dt} \tag{C.3}$$

$$\hat{\Omega}_0(dt) = 1 - (i\hat{H} + \frac{1}{2}\hat{a}^\dagger\hat{a})dt \tag{C.4}$$

where the form of $\hat{\Omega}_0(dt)$ comes from the normalization condition. This set of operators can be physically derived if we consider projecting the output field of the cavity into the number basis via photon counting.

A unitary rearrangement, or unravelling, of the unitary operators does not change the unconditional evolution. One possible transformation physically corresponds to homodyne detection where a large classical local oscillator $\gamma = |\gamma|\exp[i\Phi]$ is added to the output light of the cavity. This transformation is

$$\hat{a} \rightarrow \hat{a} + \gamma \tag{C.5}$$

$$\hat{H} \rightarrow \hat{H} - i\frac{1}{2}(\gamma^*\hat{a} - \gamma\hat{a}^\dagger). \tag{C.6}$$

In the limit of large $|\gamma|$ we have a Poisson distributed detection process.

To generate squeezed light, we impose the OPO Hamiltonian $\hat{H} = \frac{i}{4}\chi(\hat{a}^2 - (\hat{a}^\dagger)^2)$ where $0 < \chi < 1$ below threshold. Working in a rotating frame, the conditional stochastic master

equation becomes:

$$d\hat{\rho} = \left(\frac{1}{4}\chi[\hat{a}^2 - (\hat{a}^\dagger)^2, \hat{\rho}]\right) dt$$

$$+ \left(\hat{a}\hat{\rho}\hat{a}^\dagger - \hat{a}^\dagger\hat{a}\hat{\rho}/2 - \hat{\rho}\hat{a}^\dagger\hat{a}/2\right) dt$$

$$+ \sqrt{\eta}\left(\exp[-i\Phi]\hat{a}\hat{\rho} + \hat{\rho}\hat{a}^\dagger\exp[i\Phi] - \langle\hat{a}\exp[-i\Phi] + \hat{a}^\dagger\exp[i\Phi]\rangle\hat{\rho}\right) dW. \quad \text{(C.7)}$$

Define the field quadratures $\hat{x} = \hat{a} + \hat{a}^\dagger$ and $y = -i\left(\hat{a} - \hat{a}^\dagger\right)$. If the initial density matrix has a bivariate Gaussian form in these variables, then it can be shown that $\hat{\rho}$ retains this form under the evolution of equation (C.7). The evolution of this Gaussian state can be written [199] in terms of the differential equations of the distribution parameters

$$d\langle\hat{x}\rangle = -\frac{1}{2}(1+\chi)\langle\hat{x}\rangle dt + \sqrt{\eta}\left(\cos\Phi(V_x - 1) + \sin\Phi V_{xy}\right) dW \quad \text{(C.8)}$$

$$d\langle\hat{y}\rangle = -\frac{1}{2}(1-\chi)\langle\hat{y}\rangle dt + \sqrt{\eta}\left(\cos\Phi V_{xy} + \sin\Phi(V_y - 1)\right) dW \quad \text{(C.9)}$$

$$\frac{dV_x}{dt} = 1 - (1+\chi)V_x - \eta\left(\cos\Phi(V_x - 1) + \sin\Phi V_{xy}\right)^2 \quad \text{(C.10)}$$

$$\frac{dV_{xy}}{dt} = -V_{xy} - \eta\left(\cos\Phi(V_x - 1) + \sin\Phi V_{xy}\right)\left(\cos\Phi V_{xy} + \sin\Phi(V_y - 1)\right) \quad \text{(C.11)}$$

$$\frac{dV_y}{dt} = 1 - (1-\chi)V_y - \eta\left(\cos\Phi V_{xy} + \sin\Phi(V_y - 1)\right)^2 \quad \text{(C.12)}$$

where the variances are defined as

$$V_x = \langle(\hat{x} - \langle\hat{x}\rangle)^2\rangle \quad \text{(C.13)}$$

$$V_y = \langle(\hat{y} - \langle\hat{y}\rangle)^2\rangle \quad \text{(C.14)}$$

$$V_{xy} = \frac{1}{2}\langle(\hat{x} - \langle\hat{x}\rangle)(\hat{y} - \langle\hat{y}\rangle) + (\hat{y} - \langle\hat{y}\rangle)(\hat{x} - \langle\hat{x}\rangle)\rangle. \quad \text{(C.15)}$$

Notice that the variances evolve deterministically.

Let us assume $\eta = 1$, perfect measurement efficiency, for the remainder of the discussion. The unconditional stationary solution to equation (C.1) (or the average of solutions to

equation (C.7)) is given by

$$\langle \hat{x} \rangle = 0 \tag{C.16}$$

$$\langle \hat{y} \rangle = 0 \tag{C.17}$$

$$V_x = \frac{1}{1+\chi} \tag{C.18}$$

$$V_y = \frac{1}{1-\chi} \tag{C.19}$$

$$V_{xy} = 0. \tag{C.20}$$

This is a mixed state. Notice that $V_x > \frac{1}{2}$, so that the unconditional state cannot become infinitely squeezed.

The situation is qualitatively different for the conditional states. With $\Phi = 0$, the conditional solution to equation (C.7) is given by

$$\langle \hat{x} \rangle \neq 0 \tag{C.21}$$

$$\langle \hat{y} \rangle = 0 \tag{C.22}$$

$$V_x = 1 - \chi \tag{C.23}$$

$$V_y = \frac{1}{1-\chi} \tag{C.24}$$

$$V_{xy} = 0. \tag{C.25}$$

Here a particular pure state solution is an ideal squeezed state ($V_x = 1/V_y$). The displacement $\langle \hat{x} \rangle$ undergoes bound diffusion (projection jitter), which supplies a degree of noise such that *on average* these pure states reproduce the unconditional solution. This is a particularly intuitive way of thinking about the squeezing spectrum. For short times differences the conditional variance can be infinitely squeezed ($V_x = 1 - \chi$), while for larger time differences the correlations wash out because of the projection jitter ($\langle \hat{x} \rangle \neq 0$).

With $\Phi = \pi/2$, the conditional solution to equation (C.7) is given by

$$\langle \hat{x} \rangle = 0 \tag{C.26}$$

$$\langle \hat{y} \rangle \neq 0 \tag{C.27}$$

$$V_x = 1 + \chi \tag{C.28}$$

$$V_y = \frac{1}{1 + \chi} \tag{C.29}$$

$$V_{xy} = 0. \tag{C.30}$$

The discussion here is the same as above, but with less squeezing and a different axis of projection noise.

Because we know that the conditional pure state is an ideal squeezed state, we will now simplify our description. The definition of an ideal squeezed state [68] is

$$|\alpha_0^S, z_0^S\rangle = \hat{D}(\alpha_0^S)\hat{S}(z_0^S)|0\rangle \tag{C.31}$$

where $\hat{S}$ squeezes the vacuum with the complex number z_0^S and $\hat{D}$ displaces the resulting state in the direction of α_0^S. The 0 subscript indicates the starting values of the state, and the S superscript labels a state variable. Another way to represent the squeezing parameter is as $B_0^S = \exp[i(\phi_0^S + \pi)]\tanh(r_0^S)$ where $z_0^S = r_0^S \exp[i\phi_0^S]$. Notice that the limits of a coherent state and infinite squeezing are represented by the $|B_0^S|$ (r_0^S) magnitudes of 0 (0) and 1 (∞) respectively. For the situation where we pump up the cavity with $\Phi = 0$ the conditional amount of squeezing is $|B_0^S| = \tanh(-\ln(1-\chi)/2)$. Finally, we can relate all of these descriptions to the usual specification of squeezing as

$$S_{dB} = 10\log\left(\frac{\langle \Delta x^2 \rangle}{1}\right) \tag{C.32}$$

$$= 10\log(e^{\pm 2r_t^S}) \tag{C.33}$$

$$= \pm 20\log(e)r_t^S \tag{C.34}$$

$$= \pm 8.7 r_t^S \tag{C.35}$$

$$= \pm 8.7\,\mathrm{atanh}(|B_t^S|). \tag{C.36}$$

In terms of this new set of parameters, consider what happens when we stop pumping the OPO ($\chi = 0$) and let the state decay to the vacuum. Initially, we can imagine determining

the initial state by using homodyne tomography for a sufficient length of time [200] before turning off the pump. Again, assume the state is being measured with a standard homodyne detection scheme with (at time t) difference photocurrent $I(t)$ and local oscillator phase $\Phi(t)$. In the limit where the local oscillator magnitude is large, the photocurrent evolution is given by

$$I(t)dt = 2\mathrm{Re}(\alpha_t^S e^{-i\Phi(t)})dt + dW(t) \tag{C.37}$$

where α_t^S is the conditional displacement at time t and $dW(t)$ is a Wiener increment, which encapsulates the intrinsic randomness of the measurement [11].

Remarkably, the conditional state is a function of only the initial state parameters and two other complex measurement parameters:

$$
\begin{align}
R_t &= \int_0^t I(s)e^{-s/2}e^{i\Phi(s)}ds \tag{C.38} \\
S_t &= -\int_0^t e^{-s}e^{2i\Phi(s)}ds. \tag{C.39}
\end{align}
$$

In terms of these numbers, the conditional state at time t is given by

$$|\alpha_t^S, z_t^S\rangle = N \exp[-\frac{1}{2}\hat{a}^\dagger \hat{a}t] \exp[\frac{1}{2}(S_t^* \hat{a}^2 + R_t^* \hat{a})]|\alpha_0^S, z_0^S\rangle \tag{C.40}$$

where N is a normalization factor [113]. In other words, with the perfect detection efficiency we have assumed, a pure squeezed state remains a pure squeezed state under the conditional evolution. Although we will use this set of variables to describe the evolution, we can easily remove the decay envelope of the state with a change of the time variable $v = 1 - e^{-t}$. This leads to redefined measurement results

$$
\begin{align}
R_t \to A_v &= \int_0^v I(u)e^{i\Phi(u)}du \tag{C.41} \\
S_t \to B_v &= -\int_0^v e^{2i\Phi(u)}du \tag{C.42}
\end{align}
$$

with which one can simulate the photocurrent evolution as

$$I(v)dv = 2\mathrm{Re}(\alpha_v^S e^{-i\Phi(v)})dv + dW(v). \tag{C.43}$$

Using Equations (2.12) and (A8) from [201], we can manipulate equation (C.40) to find

the new state parameters (α_t^S, B_t^S) as a function of the initial values (α_0^S, B_0^S) and the measurement parameters (R_t, S_t). After rearranging terms, these turn out to be

$$
\begin{aligned}
\alpha_t^S &= e^{-\frac{t}{2}} \left(\alpha_0^S [1 + B_0^S(e^{-t} - 1) - B_0^{S*}S_t + |B_0^S|^2(S_t - e^{-t})] + [B_0^S R_t^* + |B_0^S|^2(e^{-t}R_t - S_t R_t^*)] \right) \\
&\quad \times \left(1 - 2\mathrm{Re}(B_0^S S_t^*) + |B_0^S|^2(|S_t|^2 - e^{-2t}) \right)^{-1} \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad \text{(C.44)} \\
B_t^S &= e^{-t} \left(\frac{1}{B_0^S} - S_t^* \right)^{-1}. \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad \text{(C.45)}
\end{aligned}
$$

There are a few important points to make about these equations. First, the magnitude of B_t^S decreases monotonically. If the state is not initially squeezed ($B_0^S = 0$), then the measurement cannot induce squeezing. However, if the squeezing ellipse is not lined up with the measurement axis (B_t^S complex), the direction of squeezing will rotate (B_t^S becomes more real). Also note that α_t^S only depends on the measurement results if there is a nonzero degree of squeezing. In other words, the state must be quantum mechanical to experience back-action. Otherwise, it decreases monotonically because of the cavity decay.

These points can also be seen by ignoring the measurement parameters and writing out the stochastic differential equations (SDEs) for the state parameters.

$$
\begin{aligned}
d\alpha_t^S &= -\frac{1}{2}\alpha_t^S dt + \frac{B_t^S dW}{1 - |B_t^S|^2}[(B_t^S)^* e^{i\Phi(t)} + e^{-i\Phi(t)}] \quad\quad\quad\quad\quad\quad \text{(C.46)} \\
dB_t^S &= -B_t^S(1 + e^{-2i\Phi(t)}B_t^S)dt. \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad \text{(C.47)}
\end{aligned}
$$

Again the $d\alpha_t^S$ equation is stochastic only if there is nonzero squeezing. Notice that the dB_t^S equation is not stochastic and can be integrated easily to give equation (C.45). However, the B_t^S equation can become stochastic indirectly if we allow an adaptive measurement. In this case, dB_t^S may depend on the random dW if we allow $\Phi(t)$ to be a function of $I(t)$.

To get a feel for the above equations, imagine performing the following experimental procedure as depicted in figure C.1 and figure C.2. Set the local oscillator phase to zero and measure the photocurrent for each of two squeezed vacuum states: amplitude squeezed ($B_0^S = -0.9$, north-south) and phase squeezed ($B_0^S = 0.9$, east-west). In both cases α_t^S stays real due to equation (C.46), and the squeezing ellipse retains its direction due to equation (C.47). For the phase squeezed state, the B_t^S decay looks like an exponential. For the amplitude squeezed state, the B_t^S decay does not look like an exponential, but stays larger for longer.

Qualitatively, the phase squeezed state is very quickly 'projected' into its wings, becom-
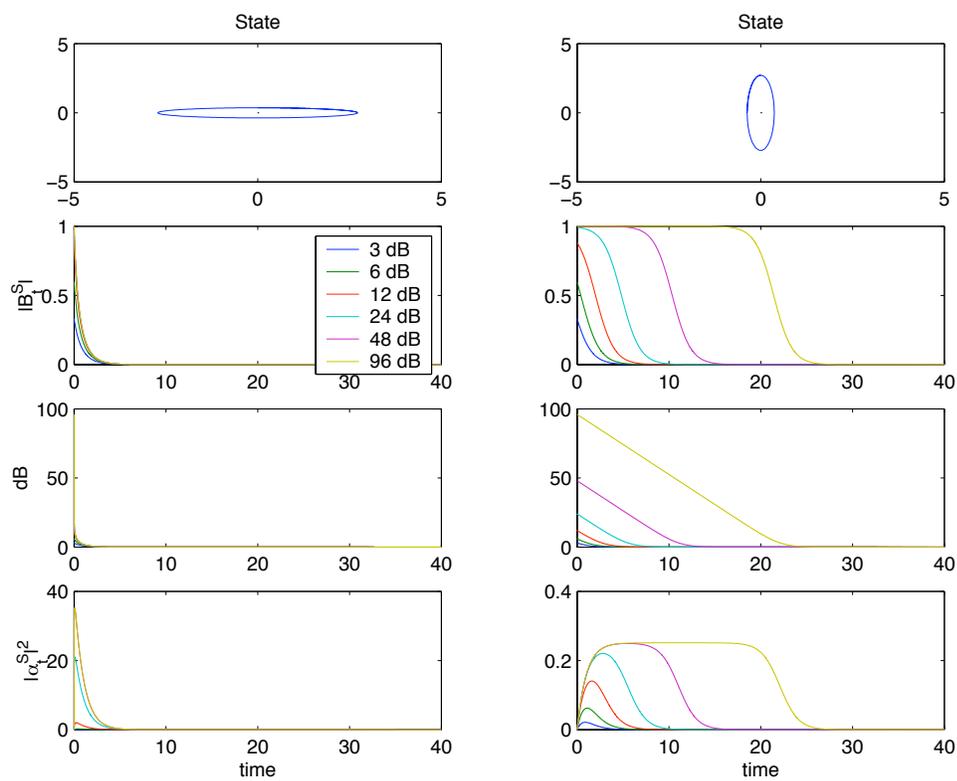
Figure C.1: Collapse of coherence depending on measurement quadrature. Notice that the state collapses much more slowly when measuring the squeezed quadrature (right) as opposed to the antisqueezed quadrature (left).
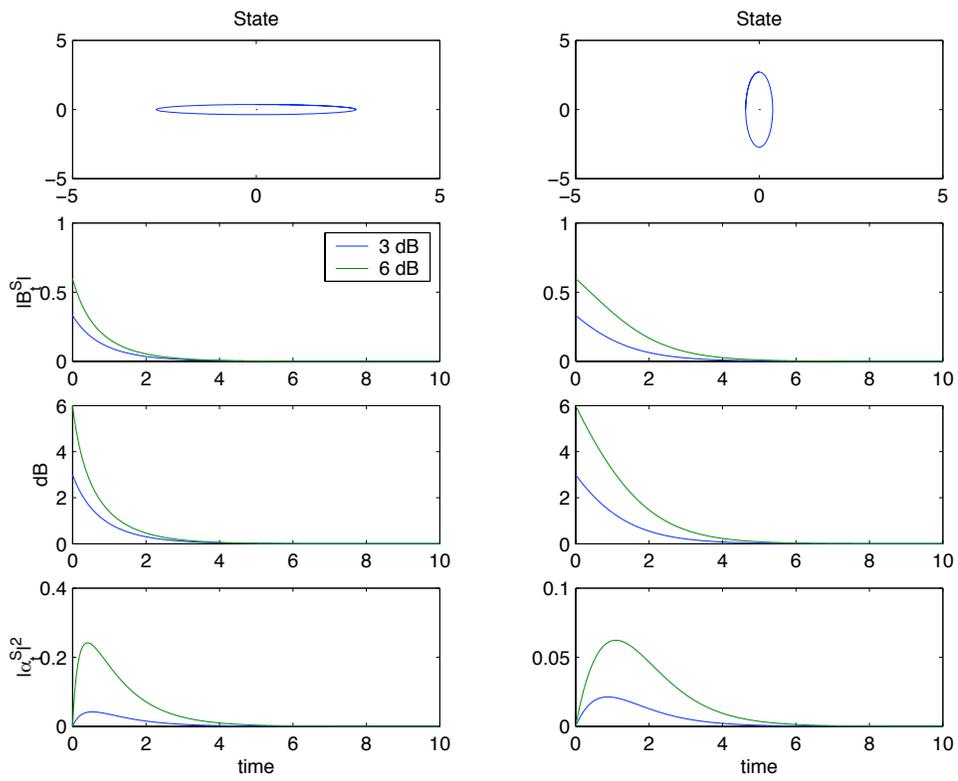
Figure C.2: A realistic subsample of figure C.1.

ing a coherent state with a large displacement limited only by the amount of squeezing. This can be seen in the $\langle |\alpha_t^S|^2 \rangle$ plots of figure C.1. The amplitude squeezed state is able to stay squeezed for longer because of its minimal projection onto the real axis. For large squeezing, its mean displacement saturates to $\langle |\alpha_t^S|^2 \rangle = 0.25$ within the (unity) timescale of decay. Once the deterministic squeezing ellipse decays to near vacuum, this value also decays back to zero. The time that the amplitude squeezed state stays in this $\langle |\alpha_t^S|^2 \rangle$ equilibrium is limited only by the amount of initial squeezing.

Of course, we could describe the two situations as the same state, but with different constant local oscillator phases of 0 and $\pi/2$. Despite the large qualitative difference between the two trajectory pictures, the unconditional density matrix $\widetilde{\hat{\rho}}(t)$ must be equal to the average projector onto the conditional pure states of either case described above:

$$\hat{\rho}(t) = \overline{|\psi_i(t)\rangle\langle\psi_i(t)|}\big|_{\Phi=0} = \overline{|\psi_i(t)\rangle\langle\psi_i(t)|}\big|_{\Phi=\frac{\pi}{2}}. \tag{C.48}$$

Again, this requirement stems from the fact that changing the local oscillator phase is a unitary reformatting of the measurement operators. Physically, this just means that the unconditional description is independent of how we look at the system. One must change the Hamiltonian of the system to alter $\hat{\rho}(t)$.

To reconcile equation (C.48) with the trajectory picture we note that at intermediate times $\hat{\rho}(t)$ is a mixed state with an associated variance matrix. We assume it starts pure and know it ends pure as the vacuum. At the point in time where one set of trajectories is a coherent state and the other is still a squeezed state, the averaged variances of both are still the same because there are two sources of noise: the uncertainty and the "projection jitter." The coherent states have unity uncertainty and no projection jitter, while the squeezed states have small uncertainty but nonzero projection jitter (see equation (C.46)).

C.3 Potential Experiments

Our goal now is to design an experiment that demonstrates the trajectory equations above are an essential description in single-shot measurements. First, we will outline experiments that display characteristics of the above dynamics, given that we know the initial state of the system, i.e., system verification. Subsequently, we will assume we do not know the initial

state and discuss possible measurement and feedback strategies for performing squeezed state tomography.

C.3.1 System Verification

First assume we know initial state of the system. As mentioned above, this can be achieved in practice by pumping the OPO to equilibrium, measuring the output c-w state via tomography techniques, then instantaneously turning off the pump to start the decay experiment.

C.3.1.1 Nonstationary Correlations

In squeezed state measurements one considers measuring correlations of the photocurrent. Consider the quantity $\langle \Delta I(t) \Delta I(t') \rangle$ with the definitions

$$\Delta I(t) \quad = \quad I(t) - 2\mathrm{Re}(\alpha_t^C e^{-i\Phi(t)}) \tag{C.49}$$

$$\alpha_t^C \quad = \quad \alpha_0^S e^{-\frac{t}{2}}. \tag{C.50}$$

Here α_t^C is how the state would behave if there were no initial squeezing. After some manipulation of SDE integrals, we arrive at:

$$\langle \Delta I(t) \Delta I(t') \rangle = \delta(t-t') + 2e^{t-t'}\mathrm{Re}[e^{-i\Phi(t)}f(t')] + e^{-\frac{t+t'}{2}} \int_0^{t'} e^{-s}|f(s)e^{-i\Phi(t)} + f^*(s)e^{i\Phi(t')}|^2 ds \tag{C.51}$$

with $t > t'$ and $f(t) = \frac{B_t^S}{1-|B_t^S|^2}[(B_t^S)^* e^{i\Phi(t)} + e^{-i\Phi(t)}]$. Unlike the case in cw squeezed beams, the correlation function here is a nonstationary quantity. Without the use of the trajectory formulation, its calculation would be considerably more difficult. One could measure $\langle \Delta I(t) \Delta I(t') \rangle$ to verify the dynamics described here, however its interpretation is slightly subtle. For example, one might want to associate the power spectrum $\langle \Delta I(t)^2 \rangle$ directly with the uncertainty parameter B_t^S, which is invalid because of the noise from "projection jitter" mentioned above.

The nonstationary correlation function as a function of time is plotted in figure C.3. Clearly, the most interesting correlation function comes from the scenario where the squeezed quadrature is measured and the squeezing persists for a larger amount of time as seen in figure C.1.
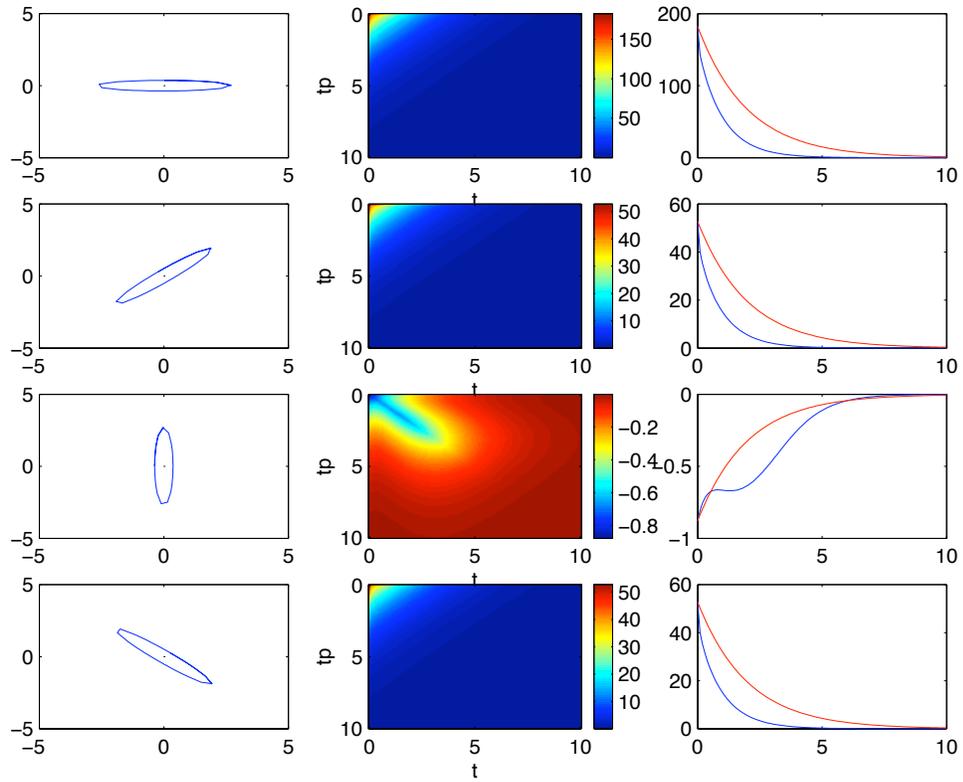
Figure C.3: Nonstationary correlations of a decaying squeezed mode. The first column represents the squeezing direction with respect to the measurement direction, the second column is $\langle \Delta I(t) \Delta I(t') \rangle$, the third column is a cut of the two-dimensional correlation function along the side axis (red) and along the diagonal (blue). Unity shotnoise has been removed from the diagonal for clarity. The most interesting nontrivial feature is in the shape of the diagonal correlation function when the squeezed quadrature is measured.

C.3.1.2 Rotate

Other games are available to verify the above dynamics. One possibility is to always keep the squeezed state diagonal with respect to the local oscillator axis such that the squeezing ellipse in the nonrotating frame is continuously rotating. This is achieved with a local oscillator value of $\Phi(t) = \frac{\arg(B_t^S)}{2} - \frac{\pi}{4}$.

C.3.1.3 Maximize Current

Feedback schemes are particularly adept at displaying the trajectory dynamics. Let us still assume we know the initial state (squeezed vacuum), but are given the task of maximizing the current. Looking at figure C.1, we should line up the squeezing ellipse with the local oscillator such that a large displacement is created. Then, since the displacement is now necessarily along only this axis, we only have to toggle Φ between 0 and π to keep the current large and positive. In the case that there is some initial displacement, the algorithm grows more complex and the change in Φ would also depend on our knowledge of α_t^S and the magnitude of B_t^S.

C.3.1.4 Minimize Current

Minimizing the current with feedback is a little tricky in realistic situations. Imagine we started with the ellipse lined up with the local oscillator, then the state would project into a coherent state with large displacement, and we would have to set $\Phi = \arg(\alpha_t^S) + \frac{\pi}{2}$. However, with a delay in the feedback loop this may result in a temporary undesirable peak in the photocurrent. A better (trivial) approach would be to keep the squeezing ellipse perpendicular to the local oscillator axis such that it stays squeezed for longer. While there is some small displacement during this time, it is acceptable compared to the large jump that would occur if the local oscillator was rotated.

C.3.2 Squeezed State Tomography

Let us now assume we have limited knowledge of the initial state of the system. There are 4 parameters to monitor (2 complex numbers) and, depending on the game we want to play, we may allow ourselves knowledge of some with the goal of measuring the others. Each measurement is assumed to consist of only a single shot, not an ensemble where standard

tomography techniques could be used to reproduce the state in full.

C.3.2.1 Displacement Phase Measurement

We first discuss the adaptive measurement of the phase of a single light pulse. Algorithms to do this have been studied extensively [11], but mainly with reference to coherent states where no back-action occurs. Similar approaches have been described with squeezed states where the direction of the squeezing is chosen along axis of the displacement to facilitate a low-noise measurement of the displacement phase [199]. In the case that the squeezed beam is continuous, there is no rotation of the squeezing ellipse thus it does not have to be accounted for in the feedback algorithm. However, when measuring the phase of a decaying squeezed state cavity, the direction of squeezing may rotate, complicating the feedback scheme. But if the algorithm is able to quickly set $\Phi \approx \arg(\alpha_t^S) + \frac{\pi}{2}$ then the squeezing ellipse will not significantly rotate and the algorithm can proceed as before.

C.3.2.2 Squeezing Phase Measurement

Now we discuss measuring the squeezing parameters, particularly the phase, in a single-shot measurement. Because these parameters manifest themselves as fluctuations the task is inherently more complicated. Still, the intuition of the above description is useful in designing heuristic algorithms.

First assume that the initial state is squeezed vacuum. Unless the state is very close to the amplitude-squeezed position (north-south), the squeezed state will quickly project into a coherent state with finite displacement along a direction depending on the squeezing parameter (B_t^S). One procedure that demonstrates the usefulness of feedback during single-shot measurements is to give the state some time to project into a coherent state, which is subsequently measured using standard procedures [11]. This phase will then be related to the desired squeezing parameter through equation (C.46). Information about both the magnitude and phase of B_t^S can be extracted from this relation. Also, if the state is initially near amplitude squeezed, a heterodyne type measurement can be made early on to get the state to project into a coherent state.

C.4   Measurement Formalisms

It is extremely important to realize that we do not always have to use the trajectory formalism to calculate experimental results, e.g., the correlation function. Experimentally convincing data consist of ensemble averages, not single traces that may be stochastically anomalous. If we continue to describe the bath quantum mechanically, the ensemble average of the measured current correlation functions can be expressed as the expectation value of an operator acting on both the bath and the system (i.e., the Heisenberg picture).

For the case of open-loop measurements, there may not be much difference between the two methods. However, if we perform an experiment to measure the correlation function of the system with feedback, the trajectory picture becomes a practical necessity.

Let us describe a very simplified feedback experiment for demonstration purposes. Suppose we prepare the cavity in a coherent state, which subsequently decays exponentially. Since there is no squeezing, the state will experience no nontrivial back-action. Further assume we are interested in the measurement of the ensemble averaged current, where we have linear "Markovian" feedback from the current to the homodyne phase. In practice there will be some delay in the feedback loop, but we assume that the delay is close to zero.

C.4.1   Trajectory Formalism

The conditional current for a single trajectory is given by

$$I_C(t) = 2\text{Re}(\alpha_0^S e^{-t} e^{-i\phi(t)}) + \xi(t) \tag{C.52}$$

where $\xi(t) = dW(t)/dt$ is thought of as the shotnoise from the Poisson detection process of detecting the large local oscillator.

Let us suppose the feedback demands $\phi(t) = gI_C(t)$ with $g \ll 1$. If we replace the phase with this expression, we have an intractable equation for $I_C(t)$. However, if we assume g to be small we can expand the exponentials and rearrange [57] to get

$$I_C(t) = (1 + 2ge^{-t}\text{Im}(\alpha_0^S))(2e^{-t}\text{Re}(\alpha_0^S) + \xi(t)) \tag{C.53}$$

$$\langle I(t) \rangle = 2e^{-t}\text{Re}(\alpha_0^S) + 4ge^{-2t}\text{Im}(\alpha_0^S)\text{Re}(\alpha_0^S). \tag{C.54}$$

C.4.2 Operator Formalism

Here we include the bath in our description and represent the current as an operator equal to the quadrature of the output field

$$\hat{I} = (\hat{a} + \hat{\nu})e^{-i\hat{\phi}} + (\hat{a}^\dagger + \hat{\nu}^\dagger)e^{i\hat{\phi}} \qquad (C.55)$$

where $\hat{\nu}$ is the annihilation operator on the input field of the cavity (considered to be vacuum). Most importantly, we have changed the local oscillator phase from a parameter into an operator.

Now we make the feedback into an operator equation: $\hat{\phi} = g\hat{I}$. Again, this leaves us with an intractable equation for $\hat{I}$, which we can solve if we let g be small. Here we must also symmetrize the resulting expression, as is normal when converting parameters into operators. The final form of the feedback current operator is then

$$\hat{I} = \hat{a} + \hat{\nu} + \hat{a}^\dagger + \hat{\nu}^\dagger + ig((\hat{a}^\dagger)^2 - \hat{a}^2 + (\hat{\nu}^\dagger)^2 - \hat{\nu}^2). \qquad (C.56)$$

Now for coherent state we find the average current to be

$$\langle I(t) \rangle = \mathrm{Tr}_{S,E}(|\alpha_0^S e^{-t}\rangle\langle\alpha_0^S e^{-t}|_S \otimes |0\rangle\langle 0|_E \hat{I}) \qquad (C.57)$$

$$= 2e^{-t}\mathrm{Re}(\alpha_0^S) + 4ge^{-2t}\mathrm{Im}(\alpha_0^S)\mathrm{Re}(\alpha_0^S) \qquad (C.58)$$

as we also found for the measurement scenario. In general, one would perform this trace with respect to the unconditional density matrix, but in this special case the conditional and unconditional state are the same. By the form of the above equation, it can be seen that using feedback rearranges the information content of the correlation functions, placing higher-order moments in functions that would normally have only lower-order moments of the field operators.

C.4.3 Discussion

Under these special circumstances the two approaches were similarly easy to work with. Instead of making ϕ a function of I let us now assume that ϕ is a function of the state variables at any point in time. In this case, we no longer have an intractable equation for I

and we do not have to make any small parameter approximations. This state-variable feedback can be implemented if we have perfect detection of the conditional pure state. Using the Heisenberg form of the operators in the no-measurement picture, we could now include delays in the feedback loop. Thus we could write down the analytic form of any correlation function as a trace of a complicated function of field operators over the unconditional density matrix.

To analytically simplify the resulting equation may be quite difficult. There may be multiple time arguments in each of a large number of operator averages. Of course, the same problem exists in the trajectory calculation, only with stochastically integrated scalars replacing the operators and white noise replacing the input field annihilation operators. For most calculations, it is useful to have both approaches available.

Practically, the trajectory picture is indispensable in adaptive scenarios. Trajectories are the classical way of thinking about the controlled object, thus it is only natural to extend the notion to quantum mechanical systems. Although feedback experiments need not rely on measurement, they often will, making a physically derived trajectory model the realistic description. Given the task of any of the control games mentioned above, it is generally easier to design a controller with the pure state conditional dynamics in mind than to think of the trace of a complicated operator expression over the unconditional density matrix.

Appendix D

Spin Magnetic Resonance Analysis

Much of this thesis has described the evolution of a collective spin-state due to a simple perpendicular constant field, resulting in Larmor precession. In our experiment, we possess the capability of modulating Gauss-sized fields at close to MHz frequencies, enabling an investigation of the spin evolution due to time varying fields. We have several motivations for investigating such dynamics including the possibility of eliminating technical noise with a frequency modulated spin-squeezing scheme and performing magnetometry at high frequencies. One such scheme is discussed in chapter 14 in the context of our experiment.

In this appendix, we go into more theoretical detail on the measurement of a collective spin-state evolving under time varying fields. This work is motivated primarily by the stroboscopic scheme presented in section 14.5. First, we describe the classical spin dynamics for the typical "magnetic resonance" configuration, with a large magnetic field in one direction and a small, time varying field along a perpendicular direction. Then we describe the expected noise characteristics for the sine and cosine quadratures of the measurement record under stroboscopic and constant measurement conditions. All of the intuition associated with the quantum dynamics and parameter estimation remains essentially the same as in earlier chapters, but many of the technical details are evaluated here.

D.1   Spin Magnetic Resonance

Now we calculate the evolution of an individual spin vector under the usual conditions of magnetic resonance, with a large holding field in one direction and a small oscillating field in a perpendicular direction. This problem is standard in many introductory quantum mechanics textbooks. Here we are mostly concerned with only the spin moments, e.g., $\langle \hat{f}_{\mathrm{z}} \rangle$,
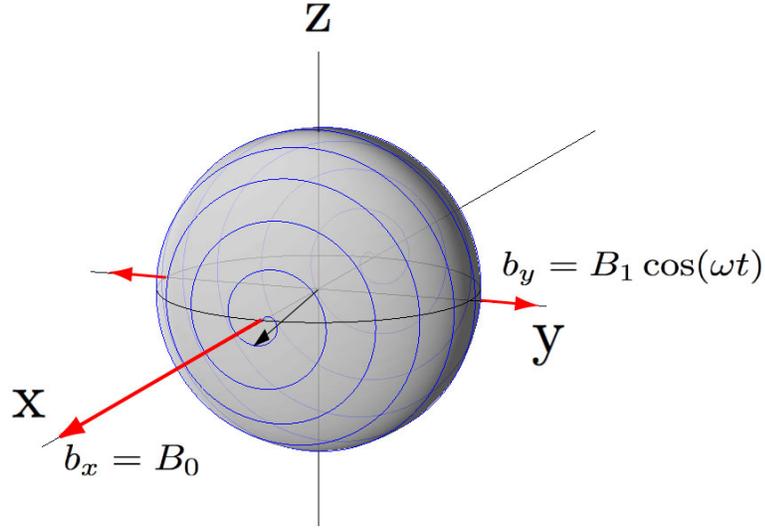
Figure D.1: Dynamics of the spin-vector initially aligned in the x-direction due to a large field B_0 in the x-direction and an oscillating field $B_1 \cos(\omega t)$. Here we plot the resonant solution with $\omega = \omega_0 = -\gamma_f B_0$.

thus the problem amounts to nothing more than solving the classical equation

$$\frac{d\mathbf{f}}{dt} = \gamma_f \mathbf{f} \times \mathbf{b} \tag{D.1}$$

for a particular time dependent field. (The $\gamma_f = -g_f \mu_B / \hbar$ used in this appendix is in units of rad/s/G and differs by a factor of 2π from the $\gamma_{Bf} = g_f \mu_B / h$ used in earlier chapters which is in units of Hz/G.) The results of the solution are presented here in the context of measuring the amplitude of the unknown perpendicular field, i.e., extending the quantum magnetometry scenario into the regime of measuring alternating (AC), rather than constant (DC), fields.

In solving the problem, we follow the treatment of reference [66], section 1.2.3, although we change coordinates and quantize along x (i.e., $z \to x$, $x \to -z$). Although the problem is specifically for a spin-1/2 particle, the spin moment solutions, $\langle \hat{f}_i \rangle$, are valid for any spin value, as the solution to the above equation is unique. The magnetic resonance Hamiltonian of interest is given by

$$\hat{H} = g_f \mu_B (B_0 \hat{f}_x + B_1 \cos(\omega t) \hat{f}_y) \tag{D.2}$$

where $B_0 \gg B_1$. The evolution from this Hamiltonian is graphically depicted in figure D.1 and we now go about solving the equations of motion. From the Hamiltonian, we define the following frequencies whose meaning will become apparent

$$\gamma_f = -g_f \mu_B / \hbar \tag{D.3}$$

$$\omega_0 = -\gamma_f B_0 \tag{D.4}$$

$$\omega_1 = -\gamma_f B_1 / 2 \tag{D.5}$$

$$\Omega_0 = \omega - \omega_0 \tag{D.6}$$

$$\Omega = \left(\Omega_0^2 + \omega_1^2\right)^{1/2}. \tag{D.7}$$

If we denote the quantum state as

$$|\Psi\rangle = c_1 |1\rangle + c_2 |2\rangle \tag{D.8}$$

(in the x-basis) and assume initial conditions $c_1 = 0$ and $c_2 = 1$, then one can go into the rotating frame of ω_0 about x and make the RWA (rotating wave approximation) ignoring $\omega + \omega_1$ frequencies to get the solution [66]

$$c_1 = -i\frac{\omega_1}{\Omega} \sin\left(\frac{\Omega t}{2}\right) \exp(-i\Omega_0 t/2) \tag{D.9}$$

$$c_2 = \left(\cos\left(\frac{\Omega t}{2}\right) - i\frac{\Omega_0}{\Omega} \sin\left(\frac{\Omega t}{2}\right)\right) \exp(+i\Omega_0 t/2). \tag{D.10}$$

This time dependent state has the following values for the spin components in the rotating frame:

$$\langle \hat{f}_x \rangle = \frac{1}{2}(|c_1|^2 - |c_2|^2)$$
$$= \frac{1}{2}\left(\frac{2\omega_1^2}{\Omega^2} \sin^2\left(\frac{\Omega t}{2}\right) - 1\right) \tag{D.11}$$

$$\langle \hat{f}_y \rangle = \frac{1}{2}(ic_1^* c_2 - ic_2^* c_1)$$
$$= \frac{1}{2}\frac{-2\omega_1}{\Omega}\left(\frac{\Omega_0}{\Omega} \sin^2\left(\frac{\Omega t}{2}\right) \sin\left(\Omega_0 t\right) + \sin\left(\frac{\Omega t}{2}\right) \cos\left(\frac{\Omega t}{2}\right) \cos\left(\Omega_0 t\right)\right) \tag{D.12}$$

$$\langle \hat{f}_z \rangle = -\frac{1}{2}(c_1^* c_2 + c_2^* c_1)$$
$$= -\frac{1}{2}\frac{2\omega_1}{\Omega}\left(\frac{\Omega_0}{\Omega} \sin^2\left(\frac{\Omega t}{2}\right) \cos\left(\Omega_0 t\right) - \sin\left(\frac{\Omega t}{2}\right) \cos\left(\frac{\Omega t}{2}\right) \sin\left(\Omega_0 t\right)\right). \tag{D.13}$$

Again, the 1/2 prefactor can be replaced with the full spin value $f = 4$ to get the solution for the larger spin. For the special on-resonance case of $\omega = \omega_0$ we have

$$\langle \hat{f}_\mathrm{x} \rangle = -\frac{1}{2} \cos(\omega_1 t) \tag{D.14}$$

$$\langle \hat{f}_\mathrm{y} \rangle = -\frac{1}{2} \sin(\omega_1 t) \tag{D.15}$$

$$\langle \hat{f}_\mathrm{z} \rangle = 0. \tag{D.16}$$

Thus, substituting f for $1/2$ to generalize, we have the behavior of the perpendicular spin $\langle \hat{f}_\mathrm{z} \rangle$ in the nonrotating frame as

$$\langle \hat{f}_\mathrm{z} \rangle = -f \sin(\omega_1 t) \sin(\omega_0 t)$$

$$\approx -f \omega_1 t \sin(\omega_0 t)$$

$$= \frac{f \gamma_f B_1}{2} t \sin(\omega_0 t) \tag{D.17}$$

where we have used the fact that $\omega_1 \ll \omega_0$ and approximated the solution for small times. The photocurrent (measuring $\hat{F}_\mathrm{z}$ as usual) will then be

$$y = \sqrt{S}\langle \hat{F}_\mathrm{z} \rangle + \sqrt{\zeta}dW$$

$$= \sqrt{S}\frac{\langle \hat{F}_\mathrm{x} \rangle \gamma_f b}{2} t \sin(\omega_1 t) + \sqrt{\zeta}dW. \tag{D.18}$$

The evolution from the above equations leading to this photocurrent is graphically depicted in figure D.1.

Now consider *estimating* the perpendicular field value from this photocurrent. A naïve estimate of the field will involve the average of the appropriate quadrature of the photocurrent

$$\bar{y}_s(t) = \frac{\int_0^t y(t') \sin(\omega_0 t')dt'}{t} \tag{D.19}$$

$$= \sqrt{S}\frac{\langle \hat{F}_\mathrm{x} \rangle \gamma_f b}{2t} \int_0^t \sin^2(\omega_0 t')t'dt'. \tag{D.20}$$

Using the integral

$$\int_0^t \sin^2(\omega_0 t')t'dt' = \frac{t^2}{4} - \frac{t \sin(2\omega_0 t)}{4\omega_0} - \frac{\cos(2\omega_0 t)}{8\omega_0^2} + \frac{1}{8\omega_0^2} \tag{D.21}$$

and solving for the field for times $t \gg 2\pi/\omega_0$, we get the naïve estimator as

$$\tilde{b} \approx \frac{8}{\sqrt{S}\langle \hat{F}_{\mathrm{x}} \rangle \gamma_f} \frac{\bar{y}_s(t)}{t}. \tag{D.22}$$

This is a simple and intuitive estimate of the perpendicular field magnitude that is nonoptimal for several reasons. First, even with this classical model of the dynamics, an optimal estimator would essentially "fit" the expected sinusoid with linearly increasing amplitude, which is not what we have done. Second, the quantum noise of the measurement has been completely neglected. The optimal estimator would surely account for the possibility of a nonzero, and random, initial amplitude (equivalent to the projection noise offset in the DC scenario). In the next section, we analyze the noise characteristics of the measurement with a simple model of the quantum noise, while neglecting the magnetometry considerations. This analysis includes the possibility of stroboscopic measurement, which convolves another sinusoid into the analysis.

D.2 AC Noise Analysis

The photocurrent of the measurement can be abstractly represented as

$$\begin{aligned} I(t)dt &= \sqrt{S}dFP(t)dt + \sqrt{N}\sqrt{P(t)}dW(t) \\ &= I_S dt + I_N dt \end{aligned} \tag{D.23}$$

where dF is the expected signal from the spins (random on every trial), I_S represents the signal part of the photocurrent, and I_N represents the shotnoise part of the photocurrent. Now define the quantities below with j being either S for signal component, or N for noise component:

$$m_j(t) = \frac{\int_0^t I_j(t')dt'}{t} \tag{D.24}$$

$$s_j(t) = \frac{\int_0^t I_j(t') \sin(\omega t')dt'}{t} \tag{D.25}$$

$$c_j(t) = \frac{\int_0^t I_j(t') \cos(\omega t')dt'}{t}. \tag{D.26}$$

Here $m_j(t)$ is the mean estimate of the photocurrent, $s_j(t)$ is the sine quadrature, and $c_j(t)$ is the cosine quadrature. Now we evaluate these quantities for the case of unmodulated and modulated (or sinusoidally strobed) probe power.

D.2.1 Unmodulated Power

For an unmodulated measurement with $P(t) = P_0$ we have

$$\langle m_S^2 \rangle_0 = SP_0^2 \langle dF^2 \rangle \tag{D.27}$$

$$\langle s_S^2 \rangle_0 = \langle m_S^2 \rangle_0 \left(\frac{1 - \cos(\omega t)}{\omega t} \right)^2 \tag{D.28}$$

$$\langle c_S^2 \rangle_0 = \langle m_S^2 \rangle_0 \left(\frac{\sin(\omega t)}{\omega t} \right)^2 \tag{D.29}$$

$$\langle m_N^2 \rangle_0 = \frac{NP_0}{t} \tag{D.30}$$

$$\langle s_N^2 \rangle_0 = \langle m_N^2 \rangle_0 \left(\frac{2\omega t - \sin(2\omega t)}{4\omega t} \right) \tag{D.31}$$

$$\langle c_N^2 \rangle_0 = \langle m_N^2 \rangle_0 \left(\frac{2\omega t + \sin(2\omega t)}{4\omega t} \right) \tag{D.32}$$

where the zero subscript indicates the absence of modulation. In calculating the noise expressions, Itô's rule has been used to simplify the expressions. There should be no correlation between the signal components and the noise components. Now we use these values for reference in evaluating the stroboscopic measurement.

D.2.2 Modulated Beam and Spins

Suppose the power of the probe beam is sinusoidally modulated according to

$$P(t) = P_0(1 - \cos(2\omega t))/2 \tag{D.33}$$

and, furthermore, suppose the spins are moving such that on any on trial

$$dF(t) = dF \sin(\omega t) \tag{D.34}$$

where $\langle dF \rangle = 0$ and $\langle dF^2 \rangle \neq 0$. This simple model of the quantum spin noise should be approximately valid given an x-field that causes the time evolution and a measurement that is sufficiently stroboscopic. Here we have assumed prior knowledge of the phase of the

spin-noise such that the sine-quadrature contains the spin-noise information. This is the case in the experiment because the modulated probe *initiates* the spin-projection, so the phase is always the same.

With these assumptions, we have the signal variances

$$\langle s_S^2 \rangle \;=\; \langle m_S^2 \rangle_0 \left(\frac{12\omega t - 8\sin(2\omega t) + \sin(4\omega t)}{32\omega t} \right)^2 \tag{D.35}$$

$$\rightarrow\; \langle m_S^2 \rangle_0 \left(\frac{3}{8} \right)^2 \tag{D.36}$$

$$\langle c_S^2 \rangle \;=\; \langle m_S^2 \rangle_0 \left(\frac{\cos(4\omega t) - 4\cos(2\omega t) + 3}{32\omega t} \right)^2 \tag{D.37}$$

$$\rightarrow\; 0 \tag{D.38}$$

and the uncorrelated shotnoise variances

$$\langle s_N^2 \rangle \;=\; \langle m_N^2 \rangle_0 \left(\frac{12\omega t - 8\sin(2\omega t) + \sin(4\omega t)}{32\omega t} \right) \tag{D.39}$$

$$\rightarrow\; \langle m_N^2 \rangle_0 \left(\frac{3}{8} \right) \tag{D.40}$$

$$\langle c_N^2 \rangle \;=\; \langle m_N^2 \rangle_0 \left(\frac{4\omega t - \sin(4\omega t)}{32\omega t} \right) \tag{D.41}$$

$$\rightarrow\; \langle m_N^2 \rangle_0 \left(\frac{1}{8} \right). \tag{D.42}$$

The right arrow indicates either long times or at the specific "canceling times" of $t_n = n\pi/\omega$ for n integer.

Now because the spin noise $\langle s_S^2 \rangle$ goes down by a factor of $(3/8)^2$, the optical noise floor $\langle s_N^2 \rangle$ goes down by a factor of $(3/8)$, and the spin decay time gets extended by a factor of 2 (average of power), the expected reduction in spin-squeezing (over the ideal DC case) will be $(3/8)^2/((3/8)(1/2)) = 3/4$. This is a naïve analysis but it indicates the expected result that a sacrifice is made in the ultimate level of achievable spin-squeezing when using the technically convenient modulation scheme.

D.2.3   Technical Offset Noise

This modulation technique is clearly useful for removing the effects of shot-to-shot DC fluctuations in the polarimeter. Assume that the polarimeter has DC noise

$$I(t)dt = dIdt \tag{D.43}$$

where $\langle dI \rangle = 0$, and $\langle dI^2 \rangle \neq 0$. Then without the canceling we have

$$\langle m^2 \rangle_0 = \langle dI^2 \rangle. \tag{D.44}$$

In contrast, for the sinusoid quadrature we have

$$\langle s^2 \rangle_0 = \langle dI^2 \rangle \left(\frac{1 - \cos(\omega t)}{\omega t} \right)^2 \tag{D.45}$$

and adding in the stroboscopic measurement we have

$$\langle s^2 \rangle = \langle dI^2 \rangle \left(\frac{\cos(3\omega t) - 9\cos(\omega t) + 8}{6\omega t} \right)^2. \tag{D.46}$$

Both of these latter quantities are clearly zero for the canceling times $t_n = n\pi/\omega$ with n integer. Thus, this technique completely ignores any truly constant offset that is random between trials. If the offset moves during the course of the measurement, however, the canceling is degraded, thus the measurement time is constrained by the spectrum of the polarimeter fluctuations.

Appendix E

Suboptimal Averaging Filter with Decay

In chapter 5, we derived the state-dependent expressions for the decay due to spontaneous emission in a single, many level alkali atom. However, our treatment of the effect of spontaneous emission on the generation of collective spin-squeezing has been crude. Typically, we have simply imposed a "cutoff time" beyond which our model without spontaneous emission is no longer used. Modeling the effect of the spontaneous emission at a collective level is a difficult task as we discuss more in chapter 15. In particular, the optimal filter used to update our knowledge of the spin-squeezed state will become nontrivial once many spontaneous emission events have occurred. In the following notes, we consider the simplified two-level atom model of [177], and analyze just how poorly the nonoptimal, but easy, averaging filter performs in the presence of decay. Although these results are not exactly appropriate for our experimental situation, the end result that the averaging filter can perform almost as well as the optimal filter is encouraging.

E.1 Introduction

Spontaneous emission inherently limits the degree of spin-squeezing that can be generated in an atomic ensemble via continuous measurement. In reference [177], Madsen and Mølmer quantify the degree at which the squeezing saturates for Faraday measurement of simple atoms with two ground state levels coupled to two excited state levels. Unlike the case without spontaneous emission, the resulting optimal filter mapping the photocurrent to the state estimate is *not* equivalent to a simpler average of the photocurrent. Here we expand

on the results of [177] to characterize how well the averaging filter performs compared to the optimal filter.

E.2   Averaging Filter

Here the same results and notation as [177] are adopted, with the additional associations

$$\langle \hat{p} \rangle(t) = \langle \hat{p}_{at} \rangle(t) \tag{E.1}$$

$$v(t) = \langle \Delta \hat{p}_{at}^2 \rangle(t) = \mathrm{Var}[\hat{p}_{at}] \tag{E.2}$$

$$v_0 = \mathrm{Var}[\hat{p}_{at,0}] = 1/2 \tag{E.3}$$

$$\kappa'(t) = \kappa \sqrt{2(1-\epsilon) \exp[-\eta t]} \tag{E.4}$$

$$M = \kappa^2(1-\epsilon)/N = \Phi \left(\frac{\chi}{\Delta} \right)^2 (1-\epsilon) \tag{E.5}$$

$$\mu = \frac{\eta}{2\kappa^2(1-\epsilon)}. \tag{E.6}$$

The photocurrent (left implicit in [177]) can be represented as

$$I(t)dt = \kappa'(t)\langle \hat{p} \rangle(t)dt + dW(t) \tag{E.7}$$

$$= \sqrt{4M} \frac{\langle \hat{J}_z \rangle(t)}{\hbar} dt + dW(t) \tag{E.8}$$

where $dW(t)$ is a normally distributed random variable with variance dt. It should be emphasized that the experimentalist only has access to $I(t)$. If, for example, the experimentalist is too lazy or ignorant to use the optimal filter and keep track of the optimal estimate $\langle \hat{p} \rangle(t)$, he cannot determine the innovation $dW(t)$ at every timestep.

We can summarize the optimal filter from [177] as

$$d\langle \hat{p} \rangle(t) = -\frac{\eta}{2}\langle \hat{p} \rangle(t)dt + \kappa'(t)v(t)dW(t) \tag{E.9}$$

$$= -\frac{\eta}{2}\langle \hat{p} \rangle(t)dt + v(t)\kappa'(t)(I(t) - \kappa'(t)\langle \hat{p} \rangle(t))dt \tag{E.10}$$

$$dv(t) = -\kappa'(t)^2 v(t)^2 dt - \eta v(t)dt + \eta \exp[-\eta t]dt. \tag{E.11}$$

The equation for the variance is deterministic and has solution

$$
\begin{aligned}
v(t) &= \frac{\beta}{2}\left(\frac{v_0 + \mu + \beta/2 + \exp[-2\beta\eta t/\mu](v_0 + \mu - \beta/2)}{v_0 + \mu + \beta/2 - \exp[-2\beta\eta t/\mu](v_0 + \mu - \beta/2)}\right)\exp[\eta t] \\
&\quad - \mu\exp[\eta t].
\end{aligned} \tag{E.12}
$$

What we actually care about though is not $v(t)$, but the squeezing parameter. When this optimal filter is used the squeezing parameter is

$$
\xi(t)^2 = \frac{N\langle\Delta\hat{J}_z^2\rangle(t)}{\langle\hat{J}_x\rangle(t)^2} \tag{E.13}
$$

$$
= \frac{\hbar N v(t)}{\langle\hat{J}_x\rangle(t)} \tag{E.14}
$$

with $\langle\hat{J}_x\rangle(t) = N\exp[-\eta t]/2$.

In principle, the above expression tells us what to expect of the squeezing when using the optimal filter. Now we calculate some other quantities in order to predict the performance of a suboptimal filter. Using the solution for a standard Ornstein-Uhlenbeck process (see page 115 of [52]), we can solve the stochastic differential equation (SDE) for the mean

$$
\langle\hat{p}\rangle(t) = \exp\left[-\int_0^t \frac{\eta}{2}dt\right]\langle\hat{p}\rangle(0) + \int_0^t \exp\left[-\int_{t'}^t \frac{\eta}{2}dt\right]\kappa'(t')v(t')dW(t') \tag{E.15}
$$

$$
= \kappa'(t)\int_0^t v(t')dW(t') \tag{E.16}
$$

where we have used the zero initial condition and the definition of $\kappa'(t)$. Notice that the solution of this SDE depends on the inherently random history of the measurement record. Using equation (E.16) and $\mathrm{E}[dW(t)dW(t')] = \delta(t-t')dt$ (or page 116 of [52]), it is easy to show that

$$
\mathrm{E}[\langle\hat{p}\rangle(t)^2] = \kappa'(t)^2\int_0^t v(t')^2 dt'. \tag{E.17}
$$

With some help from Mathematica, we can then plug in $v(t)$ and show

$$
\mathrm{E}[\langle\hat{p}^2\rangle(t)] = \mathrm{E}[\langle\hat{p}\rangle(t)^2] + v(t) \tag{E.18}
$$

$$
= \exp[-\eta t]v_0 + \sinh[\eta t]. \tag{E.19}
$$

Notice that this quantity remains constant for all time (and thus $\langle \hat{p}^2 \rangle(t)$ is a martingale) iff $\eta = 0$. This result turns out to be useful in evaluating integrals below.

We can also simply solve equation (E.10) using the standard integrating factor ODE technique to get the moment as an explicit function of the photocurrent

$$\langle \hat{p} \rangle(t) \quad = \quad \frac{\int_0^t \exp\left[\int_0^{t'} (\eta/2 + \kappa'(t'')^2 v(t'')) dt''\right] \kappa'(t') v(t') I(t') dt'}{\exp\left[\int_0^t (\eta/2 + \kappa'(t')^2 v(t')) dt'\right]}. \tag{E.20}$$

Notice that this filter is somewhat complicated as it depends on weighting the photocurrent with nontrivial time dependent functions.

In practice, one may be interested in obtaining the state estimate from the measurement record in real time. (This is especially relevant if feedback is being used, but here we do not consider additional control dynamics.) In principle one could use an FPGA [10] to do real time complicated state estimation (using, for example, equation (E.20)), however one may not have the resources to evolve the optimal filter in real time. For this and other reasons, it is of interest to ask how well a nonoptimal, but functionally simple filter works. Here we consider a filter that estimates the displacement of the optimal state via the average of the photocurrent. This particular filter can be implemented with analog circuitry with minimal delay and complexity. From the form of the photocurrent

$$I(t)dt \quad = \quad \sqrt{4M} \frac{\langle \hat{J}_z \rangle(t)}{\hbar} dt + dW(t) \tag{E.21}$$

we choose the intuitive estimate of the quantity $\langle \hat{J}_z \rangle(t)$ to be defined as

$$\tilde{z}(t) \quad = \quad \frac{\hbar}{\sqrt{4M}} \frac{\int_0^t I(s)ds}{t}. \tag{E.22}$$

Because the nonoptimal $\tilde{z}(t)$ estimate does not track the displacement $\langle \hat{J}_z \rangle(t)$ perfectly, the effective variance for the spin-squeezing parameter is greater than $\langle \Delta \hat{J}_z^2 \rangle(t)$ by the additive factor $\mathrm{E}[(\tilde{z}(t) - \langle \hat{J}_z \rangle(t))^2]$. In other words, the imperfect spin-squeezing parameter is defined as

$$\xi_a(t)^2 \quad = \quad \frac{N(\langle \Delta \hat{J}_z^2 \rangle(t) + \mathrm{E}[(\tilde{z}(t) - \langle \hat{J}_z \rangle(t))^2])}{\langle \hat{J}_x \rangle(t)^2}. \tag{E.23}$$

By using the definition of the averaging filter, $\mathrm{E}[dW(t)dW(t')] = \delta(t - t')dt$, and the result from equation (E.19), after some work we can show that

$$\xi_a(t)^2 \quad = \quad \frac{\exp[\eta t]}{\eta^2 t^2}\left(2 + 2\eta t + \exp[\eta t](\eta t(2\mu + \eta t) - 2)\right). \tag{E.24}$$

Notice that the N dependence comes through μ here.

The equations for the time at which $\xi_a(t)^2$ reaches a minimum cannot be solved algebraically. Nevertheless, the performance of this function can be seen as the upper (blue) curve in figure E.1. The middle (red) curve displays the performance of the optimal filter $\xi(t)^2$ (from [177]), which reaches a significantly flat minimum. The lower (green) curve is the expected squeezing without decay, $\eta = 0$. The first feature to notice is that at small times $t \ll 1/MN$, the squeezing performance of $\xi_a(t)^2$ is nonoptimal because it incorporates an effectively infinite prior for the variance. Define t_{min} as the time at which the averaging filter squeezing parameter achieves its minimal value. At times $1/MN < t < t_{min}$, the performance of the averaging filter is essentially optimal and $\xi_a(t)^2 \approx \xi(t)^2$. However, once $\xi(t)^2$ begins to flatten, the nonoptimal squeezing $\xi_a(t)^2$ begins to increase rapidly. Although the difference between minimal value of $\xi(t)^2$ and $\xi_a(t)^2$ is small, for the suboptimal filter the minimum time region is translated to somewhat lower times and is much less flat, therefore the squeezing from averaging is less robust to differences in stopping time. Therefore, if using the averaging filter, one should *not* stop at the minimum time for the optimal $\xi(t)^2$ or else the squeezing will be significantly degraded. From time t_{min} to time $1/M$ the Gaussian approximation remains valid, but the squeezing gets much worse.
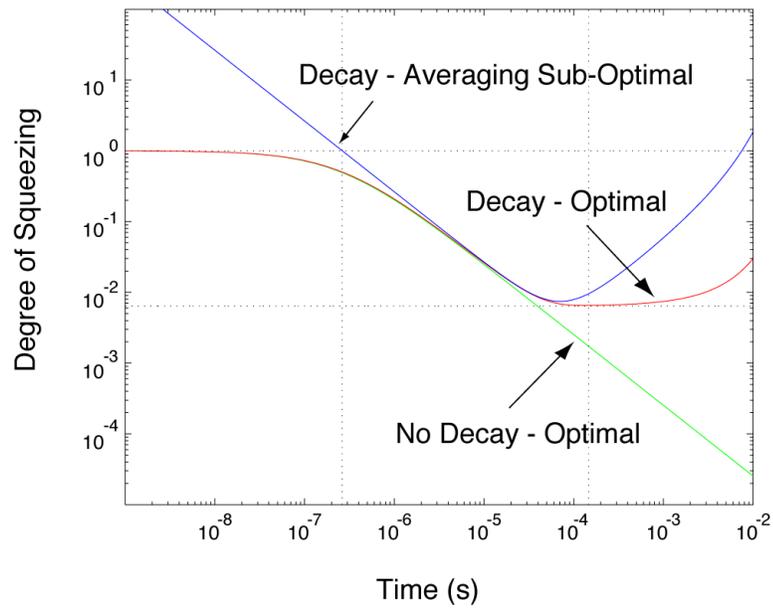
Figure E.1: Plot of squeezing parameters versus time. The top curve is the nonoptimal averaging filter spin-squeezing parameter. The middle curve is the optimal spin-squeezing parameter with decay. The bottom curve is the optimal spin-squeezing parameter without decay. The numerical parameters used are $N = 10^{11}$, $P = 100\mu\text{W}$, $\Delta = 1\text{GHz}$, $A = 2\text{mm}^2$.

Appendix F

On the Bound States of a Magnetic Quadrupole Waveguide

When I first arrived at Caltech, I began to work on a problem concerning a paper by Hinds and Eberlein on the propagation of spins in a magnetic quadrupole waveguide [202]. We, including most notably Clifford Hicks, found an error in this paper and wrote the following comment, which was not accepted. They ultimately wrote an erratum that addressed the issue and was accepted. During this research I learned how to simulate the motion of spins in magnetic traps. Tricky parts of the evolution include using the so-called "split operator" (between position and momentum) method for computational efficiency and accounting for spins that get antitrapped and lost out the side of the simulation volume. For more on magnetic trapping of atoms, see [142]. The following comment regarding [202] was originally written by me, Clifford Hicks, and Hideo Mabuchi.

F.1 Abstract

In a recent paper [202], Hinds and Eberlein discussed the quantized motion of neutral atoms in a two-dimensional quadrupole magnetic field. We argue that the bound states they propose for a spin-1/2 particle in this field cannot exist because the associated spinor wavefunction is not single-valued.

F.2 Comment

Laser cooling and high-resolution lithography have jointly enabled substantial progress towards the important goal of localizing atoms with position uncertainty much less than an

optical wavelength. Microfabricated wire patterns can produce magnetic fields with very high curvature, which may serve as extremely steep (but shallow) magnetostatic traps for neutral atoms [203, 204, 205].

Quadrupole traps are of particular interest in this context, as they are easy to produce with planar wire geometries. In practice, one will likely add a bias field to the quadrupole in order to suppress spin-flip transitions near the origin [206], but it remains a fundamental question whether bound states exist for spin-1/2 particles in a quadrupolar trap without bias. Here we will show that there are no bound states, in contradiction to Ref. [202] where a discontinuous, low angular momentum stable ground state is proposed.

The problem of a neutral spin-1/2 particle in a quadrupole magnetic field is defined by the Hamiltonian

$$H = \frac{\mathbf{p}^2}{2m} + g\mu_B \mathbf{S} \cdot \mathbf{B}(\mathbf{r}) = \begin{bmatrix} \frac{\mathbf{p}^2}{2m} & G\rho e^{i\phi} \\ G\rho e^{-i\phi} & \frac{\mathbf{p}^2}{2m} \end{bmatrix}, \tag{F.1}$$

where $\mathbf{B} = (\lambda x, -\lambda y, 0)$ in Cartesian coordinates, (ρ, ϕ, z) are corresponding cylindrical coordinates, and $G \equiv \frac{g\mu_B \hbar \lambda}{2}$. The matrix form of H is written in the eigenbasis of S_z. In what follows we will assume plane-wave dependence of the stationary states on z, and focus on spinor solutions in (ρ, ϕ).

Note that $[H, L_z - S_z] = 0$ follows from the quadrupole symmetry of the field. The general form for a simultaneous eigenstate of the two operators is then

$$\psi_q(\rho, \phi) = \frac{1}{\sqrt{2\pi}} e^{iq\phi} \begin{bmatrix} F_+(\rho) e^{\frac{i\phi}{2}} \\ F_-(\rho) e^{-\frac{i\phi}{2}} \end{bmatrix}, \tag{F.2}$$

where q is the eigenvalue of $L_z - S_z$, and the spinor is in the S_z basis. Hinds and Eberlein arrive at an equivalent expression for the general form of the solutions, although they label the discrete quantum number as l and refer to it as the angular momentum. In contrast, we label this quantum number as q to stress that it is neither an eigenvalue of L_z nor of $J_z \equiv L_z + S_z$. After arriving at equation (F.2), Hinds and Eberlein choose q to be an integer, despite the fact that *only half-integer values can ensure single-valuedness* over the entire coordinate space—any integer value of q yields $\psi_q(\rho, \phi)$ with a phase discontinuity as $\phi \to 0_+$ or $\phi \to 2\pi_-$. They proceed to solve the corresponding family of radial equations and find bound states for $q = 0$ and unbound states of finite lifetime for nonzero integer

q. In a similar analysis, we find that there are no bound states for any half-integral value of q, implying that the Hamiltonian of equation (F.1) has no single-valued bound states. So the choice of q as an integer or half-integer quantum number has significant physical consequences.

Historically, the single-valuedness requirement on wavefunctions has not always been obvious [207, 208], but multiple compelling arguments can be made without reference to experiment [209]. The problem with non-single-valuedness here is the singular nature (along the ray $\phi = 0$) of the angular derivative L_z, which appears implicitly in the Hamiltonian of equation (F.1). An equivalent irregularity with half-integral spherical harmonics is generally used as justification to exclude them from being admissible solutions in central-potential problems.

In the present case, it is important not to confuse the dependence on ϕ of an energy eigenspinor with the sign change obtained when the Bloch vector of a spin-1/2 particle is physically rotated through an angle of 2π. It is true that if a pointlike spin-1/2 particle adiabatically circles the origin of a quadrupole field, the spin must rotate by 2π in order to follow the local magnetic field lines. As a result, the spinor should obtain a well-known minus sign, i.e., a geometrical phase of π. One could presumably show this in a real or numerical experiment by sending a sufficiently localized (in ϕ) wavepacket on a circular trajectory around the quadrupole origin. But the stationary eigenstates of this system are completely delocalized in ϕ, and correlation between position and field orientation manifests itself in the fact that *half-integer* q in the spatial component of the spinor is required to obtain overall single-valuedness. Unusual quantum numbers are generically found in the quantization of systems that exhibit interesting geometrical phase properties semi-classically. For example, our situation recalls the dynamical Jahn-Teller effect in certain electron-lattice systems, where the quantum number must be half-integer because of the peculiar rotational symmetry of the problem [87].

Another particularly relevant example is the problem of a spin-1/2 particle in the circumferential field of a thin current filament [210, 211]. In this case, the total angular momentum J_z does commute with H due to rotational symmetry, thus its eigenvalues j_z are good quantum numbers. Yet the same logic applies and j_z is similarly quantized in half-integer units. In contrast to the quadrupole problem, this problem is exactly solvable and high angular momentum (high field seeking) bound states for spin-1/2 particles do exist.

Finally, we should point out possible spinor representations that make explicit the inherent coupling between spin and space variables, but also potentially obscure requisite boundary conditions. For example, we could choose to write equation (F.2) in a spatially dependent spinor basis that diagonalizes the potential terms. One option is

$$\psi_q(\rho, \phi) = \frac{1}{\sqrt{2\pi}} e^{iq\phi} (F_\uparrow(\rho)| \uparrow \rangle + F_\downarrow(\rho)| \downarrow \rangle) \tag{F.3}$$

where $| \uparrow \rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} e^{\frac{i\phi}{2}} \\ e^{-\frac{i\phi}{2}} \end{bmatrix}$ and $| \downarrow \rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} -e^{\frac{i\phi}{2}} \\ e^{-\frac{i\phi}{2}} \end{bmatrix}$ are respectively parallel and antiparallel to the field at every point in space. Another option is

$$\psi_q(\rho, \phi) = \frac{1}{\sqrt{2\pi}} e^{i(q-\frac{1}{2})\phi} (F_\uparrow(\rho)| \uparrow_{sv} \rangle + F_\downarrow(\rho)| \downarrow_{sv} \rangle) \tag{F.4}$$

where $| \uparrow_{sv} \rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} e^{i\phi} \\ 1 \end{bmatrix}$ and $| \downarrow_{sv} \rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} -e^{i\phi} \\ 1 \end{bmatrix}$ are also along the field, but are single-valued. The choice between the two options (which differ by a spatially varying phase factor) is a choice of the so-called spin-gauge freedom. Regardless of the representation that is used, each spinor component of the wavefunction must be single-valued in the global S_z basis, thus q is always half-integer. Unlike the $q = 0$ case, the radial equations always couple the stable $F_\uparrow(\rho)$ into the unstable $F_\downarrow(\rho)$ when q is half-integer, and all states are subject to decay.

Bibliography

[1] D. Budker, D. Kimball, and D. DeMille. *Atomic Physics*. Oxford University Press, 2004.

[2] M. Kitagawa and M. Ueda. Squeezed Spin States. *Phys. Rev. A*, 47:5138, 1993.

[3] M. Kitagawa and M. Ueda. Nonlinear-interferometric generation of number-phase-correlated fermion states. *Phys. Rev. Lett.*, 67:1852, 1991.

[4] D. J. Wineland, J. J. Bollinger, W. M. Itano, F. L. Moore, and D. J. Heinzen. Spin squeezing and reduced quantum noise in spectroscopy. *Phys. Rev. A*, 46:R6797, 1992.

[5] D. J. Wineland, J. J. Bollinger, W. M. Itano, and D. J. Heinzen. Squeezed atomic states and projection noise in spectroscopy. *Phys. Rev. A*, 50:67, 1994.

[6] A. Sørensen, L.-M. Duan, J. I. Cirac, and P. Zoller. Many-particle entanglement with Bose-Einstein condensates. *Nature*, 409:63, 2001.

[7] G. Chiribella, G. M. D'Ariano, P. Perinotti, and M. F. Sacchi. Efficient Use of Quantum Resources for the Transmission of a Reference Frame. *Phys. Rev. Lett.*, 93:180503, 2004.

[8] G. J. Milburn. Classical and quantum conditional statistical dynamics. *Quant. Semiclass. Opt.*, 8:269, 1996.

[9] W. A. Lin and L. E. Ballentine. Quantum tunneling and chaos in a driven anharmonic oscillator. *Phys. Rev. Lett.*, 65:2927, 1990.

[10] J. Stockton, M. Armen, and H. Mabuchi. Programmable logic devices in experimental quantum optics. *J. Opt. Soc. Am. B*, 19:3019, 2002.

[11] H. M. Wiseman and R. B. Killip. Adaptive single-shot phase measurements: The full quantum theory. *Phys. Rev. A*, 57:2169, 1998.

[12] M. A. Armen, J. K. Au, J. K. Stockton, A. C. Doherty, and H. Mabuchi. Adaptive Homodyne Measurement of Optical Phase. *Phys. Rev. Lett.*, 89:133602, 2002.

[13] M. A. Armen. Untitled. *PhD Thesis, California Institute of Technology*, 2007.

[14] K. M. Birnbaum. Cavity QED with Multi-Level Atoms. *PhD Thesis, California Institute of Technology*, 2005.

[15] L. K. Thomsen, S. Mancini, and H. M. Wiseman. Spin squeezing via quantum feedback. *Phys. Rev. A*, 65:061801–1–4(R), 2002.

[16] L. K. Thomsen, S. Mancini, and H. M. Wiseman. Continuous quantum nondemolition feedback and unconditional atomic spin squeezing. *J. Phys. B: At. Mol. Opt. Phys.*, 35:4937, 2002.

[17] J. K. Stockton, J. M. Geremia, A. C. Doherty, and H. Mabuchi. Characterizing the entanglement of symmetric many-particle spin-$\frac{1}{2}$ systems. *Phys. Rev. A*, 67:022112, 2003.

[18] J. M. Geremia, J. K. Stockton, A. C. Doherty, and H. Mabuchi. Quantum Kalman Filtering and the Heisenberg Limit in Atomic Magnetometry. *Phys. Rev. Lett.*, 91:250801, 2003.

[19] F. Verstraete, A. C. Doherty, and H. Mabuchi. Sensitivity optimization in quantum parameter estimation. *Phys. Rev. A*, 64:032111, 2001.

[20] J. K. Stockton, J. M. Geremia, A. C. Doherty, and H. Mabuchi. Robust quantum parameter estimation: Coherent magnetometry with feedback. *Phys. Rev. A*, 69:032109, 2004.

[21] J. K. Stockton, R. van Handel, and H. Mabuchi. Deterministic Dicke-state preparation with continuous measurement and control. *Phys. Rev. A*, 70:022106, 2004.

[22] R. van Handel, J. K. Stockton, and H. Mabuchi. Feedback control of quantum state reduction. *IEEE Trans. Automat. Control*, 50:768, 2005.

[23] R. van Handel, J. K. Stockton, and H. Mabuchi. Modelling and feedback control design for quantum state preparation. *J. Opt. B: Quantum Semiclass. Opt.*, 7:S179, 2005.

[24] G. A. Smith, S. Chaudhury, and P. S. Jessen. Faraday spectroscopy in an optical lattice: a continuous probe of atom dynamics. *J. Opt. B: Quant. Semiclass. Opt.*, 5:323, 2003.

[25] G. A. Smith, S. Chaudhury, A. Silberfarb, I. H. Deutsch, and P. S. Jessen. Continuous Weak Measurement and Nonlinear Dynamics in a Cold Spin Ensemble. *Phys. Rev. Lett.*, 93:163602, 2004.

[26] D. V. Kupriyanov, O. S. Mishina, I. M. Sokolov, B. Julsgaard, and E. S. Polzik. Multimode entanglement of light and atomic ensembles via off-resonant coherent forward scattering. *Phys. Rev. A*, 71:032348, 2005.

[27] J. M. Geremia, J. K. Stockton, and H. Mabuchi. Real-Time Quantum Feedback Control of Atomic Spin Squeezing. *Science*, 304:270, 2004.

[28] J. M. Geremia, J. K. Stockton, and H. Mabuchi. Suppression of Spin Projection Noise in Broadband Atomic Magnetometry. *Phys. Rev. Lett.*, 94:203002, 2005.

[29] J. M. Geremia, J. K. Stockton, and H. Mabuchi. Tensor polarizability and dispersive quantum measurement of multilevel atoms. *Phys. Rev. A*, 73:042112, 2006.

[30] D. Budker, V. Yashchuk, and M. Zolotorev. Resonant Magneto-Optical Rotation: New Twists in an Old Plot. 1999.

[31] W. Happer. Optical Pumping. *Rev. Mod. Phys*, 44:169, 1972.

[32] W. Happer and B. S. Mathur. Off-resonant light as a probe of optically pumped alkali vapors. *Phys. Rev. Lett.*, 18:577, 1967.

[33] D. Budker, W. Gawlik, D.F. Kimball, S.M. Rochester, V.V. Yashchuk, and A. Weiss. Resonant nonlinear magneto-optical effects in atoms. *Rev. Mod. Phys.*, 74:1153, 2002.

[34] I. K. Kominis, T. W. Kornack, J. C. Allred, and M. V. Romalis. A subfemtotesla multichannel atomic magnetometer. *Nature*, 422:596, 2003.

[35] P. Schwindt, S. Knappe, V. Shah, L. Hollberg, J. Kitching, L. Liew, and J. Moreland. Chip-scale atomic magnetometers. *Appl. Phys. Lett.*, 85:6409, 2001.

[36] B. Julsgaard, A. Kozhekin, and E. S. Polzik. Experimental long-lived entanglement of two macroscopic objects. *Nature*, 413:400, 2001.

[37] A. Kuzmich, L. Mandel, and N. P. Bigelow. Generation of Spin Squeezing via Continuous Quantum Nondemolition Measurement. *Phys. Rev. Lett.*, 85:1594, 2000.

[38] H. J. Metcalf and P. van der Straten. *Laser Cooling and Trapping.* Springer-Verlag, 1999.

[39] J. Hald, J. L. Sørensen, C. Schori, and E. S. Polzik. Spin Squeezed Atoms: A Macroscopic Entangled Ensemble Created by Light. *Phys. Rev. Lett*, 83:1319, 1999.

[40] V. Meyer, M. A. Rowe, D. Kielpinski, C. A. Sackett, W. M. Itano, C. Monroe, and D. J. Wineland. Experimental Demonstration of Entanglement-Enhanced Rotation Angle Estimation Using Trapped Ions. *Phys. Rev. Lett.*, 86:5870, 2001.

[41] V. B. Braginsky, F. Khalili, and K. S. Thorne (ed.). *Quantum Measurement.* Cambridge University Press, 1992.

[42] A. Kuzmich, N. P. Bigelow, and L. Mandel. Atomic quantum non-demolition measurements and squeezing. *Europhys. Lett.*, 42:481, 1998.

[43] Y. Takahashi, K. Honda, N. Tanaka, K. Toyoda, K. Ishikawa, and T. Yabuzaki. Quantum nondemolition measurement of the spin via the paramagnetic Faraday rotation. *Phys. Rev. A*, 60:4974, 1999.

[44] T. Isayama, Y. Takahashi, N. Tanaka, K. Toyoda, K. Ishikawa, and T. Yabuzaki. Observation of Larmor spin precession of laser-cooled Rb atoms via paramagnetic Faraday rotation. *Phys. Rev. A*, 59:4836, 1999.

[45] D. Oblak, J. K. Mikkelsen, W. Tittel, A. K. Vershovski, J. L. Sorensen, P. G. Petrov, C. L. Garrido Alzar, and E. S. Polzik. Quantum noise limited interferometric measurement of atomic noise: towards spin squeezing on the Cs clock transition. *Phys. Rev. A*, 71:043807, 2005.

[46] S. Chaudhury, G. A. Smith, K. Schulz, and P. S. Jessen. Continuous Nondemolition Measurement of the Cs Clock Transition Pseudospin. *Phys. Rev. Lett.*, 96:043001, 2006.

[47] A. K. Tuchman, R. Long, G. Vrijsen, J. Boudet, J. Lee, and M. A. Kasevich. Normal-mode splitting with large collective cooperativity. *Phys. Rev. A*, 74:053821, 2006.

[48] E. T. Jaynes and G. Larry Bretthorst. *Probability Theory: The Logic of Science.* Cambridge University Press, Cambridge, 2003.

[49] O. L. R. Jacobs. *Introduction to Control Theory.* Oxford University Press, New York, 2nd edition, 1996.

[50] J. Doyle, B. Francis, and A. Tannenbaum. *Feedback Control Theory.* Macmillan Publishing Co., 1990.

[51] K. Zhou and J. C. Doyle. *Essentials of Robust Control.* Prentice-Hall, New Jersey, 1st edition, 1997.

[52] C. W. Gardiner. *Handbook of Stochastic Methods.* Springer-Verlag, Berlin, 1985.

[53] B. Øksendal. *Stochastic Differential Equations.* Springer-Verlag, 5th edition, 1998.

[54] N. G. van Kampen. *Stochastic Processes in Physics and Chemistry.* Elsevier, Amsterdam, 1992.

[55] C. M. Caves, C. A. Fuchs, and R. Schack. Quantum probabilities as Bayesian probabilities. *Phys. Rev. A*, 65:022305, 2002.

[56] A. C. Doherty, S. Habib, K. Jacobs, H. Mabuchi, and S. M. Tan. Quantum feedback control and classical control theory. *Phys. Rev. A*, 62:012105, 2000.

[57] H. M. Wiseman. Quantum trajectories and feedback. *PhD Thesis, University of Queensland*, 1994.

[58] H. M. Wiseman and G. J. Milburn. *Unpublished Quantum Measurement Book.*

[59] T. A. Brun. A simple model of quantum trajectories. *Am. J. Phys.*, 70:719, 2002.

[60] M. B. Plenio and P. L. Knight. The quantum jump approach to dissipative dynamics in quantum optics. *Rev. Mod. Phys.*, 70:101, 1998.

[61] H. J. Carmichael. *Statistical Methods in Quantum Optics 1: Master Equations and Fokker-Planck Equations*. Springer-Verlag, 2003.

[62] H. P. Breuer and F. Petruccione. *The Theory of Open Quantum Systems*. Oxford University Press, Oxford, 2002.

[63] J. Preskill. *Lecture Notes on Quantum Computation*. Pasadena, 1998.

[64] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[65] C. Cohen-Tannoudji, J. Dupont-Roc, and G. Grynberg. *Atom Photon Interactions*. Wiley-Interscience, New York, 1992.

[66] J. Vanier and C. Audoin. *The Quantum Physics of Atomic Frequency Standards*, volume 1. Adam Hilger, Philadelphia, 1$^{\text{st}}$ edition, 1989.

[67] H. Mabuchi and A. C. Doherty. Cavity quantum electrodynamics: coherence in context. *Science*, 298:1372, 2002.

[68] L. Mandel and E. Wolf. *Optical Coherence and Quantum Optics*. Cambridge University Press, Cambridge, 1997.

[69] D. F. Walls and G. J. Milburn. *Quantum Optics*. Springer-Verlag, 1994.

[70] H. M. Wiseman and A. C. Doherty. Optimal Unravellings for Feedback Control in Linear Quantum Systems. *Phys. Rev. Lett.*, 94:070405, 2005.

[71] S. L. Adler, D. C. Brody, T. A. Brun, and L. P. Hughston. Martingale models for quantum state reduction. *J. Phys. A*, 34:8795, 2001.

[72] L. Bouten, M. Guta, and H. Maassen. Stochastic Schrodinger equations. *J. Phys. A: Mathematical and General*, 37:3189, 2004.

[73] V. P. Belavkin. Quantum Stochastic Calculus and Quantum Nonlinear Filtering. *J. Multivariate Anal.*, 42:171, 1992.

[74] V.P. Belavkin. *Proceedings, Bellman Continuum*, volume 121 of *Lecture notes in Control and Information Sciences*, chapter Nondemolition measurements, nonlinear filtering and dynamic programming of quantum stochastic processes. Springer-Verlag, Sophia-Antipolis, 1998.

[75] V. P. Belavkin. Measurement, Filtering and Control in Quantum Open Dynamical Systems. *Rep. on Math. Phys.*, 43:405, 1999.

[76] M. Brune, S. Haroche, V. Lefevre, J. M. Raimond, and N. Zagury. Quantum Nondemolition Measurement of Small Photon Numbers by Rydberg-Atom Phase-Sensitive Detection. *Phys. Rev. Lett.*, 65:976, 1990.

[77] G. Nogues, A. Rauschenbeutel, S. Osnaghi, M. Brune, J. M. Raimond, and S. Haroche. Seeing a single photon without destroying it. *Nature*, 400:239, 1999.

[78] C. M. Caves, K. S. Thorne, R. W. P. Drever, V. D. Sandberg, and M. Zimmermann. On the measurement of a weak classical force coupled to a quantum-mechanical oscillator. I. Issues of principle. *Rev. Mod. Phys.*, 52:341, 1980.

[79] M. Auzinsh, D. Budker, D. F. Kimball, S. M. Rochester, J. E. Stalnaker, A. O. Sushkov, and V. V. Yashchuk. Can a Quantum Nondemolition Measurement Improve the Sensitivity of an Atomic Magnetometer?. *Phys. Rev. Lett.*, 93:173002, 2004.

[80] K. Mølmer and L. B. Madsen. Estimation of a classical parameter with Gaussian probes: Magnetometry with collective atomic spins. *Phys. Rev. A*, 70:052102, 2004.

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

[82] M. R. James. Risk-sensitive optimal control of quantum systems. *Phys. Rev. A*, 69:032108, 2004.

[83] N. Yamamoto. Robust observer for uncertain linear quantum systems. *Phys. Rev. A*, 74:032107, 2006.

[84] D. A. Steck. Cesium D Line Data. 1998.

[85] E. Hecht. *Optics*. Addison-Wesley, San Francisco, 4$^\text{th}$ edition, 2002.

[86] J. F. Sherson and K. Mølmer. Polarization Squeezing by Optical Faraday Rotation. *Phys. Rev. Lett.*, 97:143602, 2006.

[87] J. J. Sakurai. *Modern Quantum Mechanics*. Addison-Wesley, 1994.

[88] R. H. Dicke. Coherence in Spontaneous Radiation Processes. *Phys. Rev.*, 93:99, 1954.

[89] T. Holstein and H. Primakoff. Field Dependence of the Intrinsic Domain Magnetization of a Ferromagnet. *Phys. Rev.*, 58:1098, 1940.

[90] C. Genes and P. R. Berman. Generating conditional atomic entanglement by measuring photon number in a single output channel. *Phys. Rev. A*, 73:013801, 2006.

[91] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri. Distinguishing Separable and Entangled States. *Phys. Rev. Lett.*, 88:187904, 2002.

[92] G. Vidal. Entanglement monotones. *J. Mod. Opt.*, 47:355, 2000.

[93] N. Gisin and H. Bechmann-Pasquinucci. Bell inequality, Bell states and maximally entangled states for n qubits. *Phys. Lett. A*, 246:1, 1998.

[94] M. A. Nielsen. Conditions for a Class of Entanglement Transformations. *Phys. Rev. Lett.*, 83:436, 1999.

[95] D. A. Varshalovich, A. N. Moskalev, and V. K. Khersonskii. *Quantum Theory of Angular Momentum*. World Scientific, 1988.

[96] P. Marte, R. Dum, R. Taieb, and P. Zoller. Resonance fluorescence from quantized one-dimensional molasses. *Phys. Rev. A*, 47:1378, 1993.

[97] N. Lutkenhaus, J. I. Cirac, and P. Zoller. Mimicking a squeezed-bath interaction: Quantum-reservoir engineering with atoms. *Phys. Rev. A*, 57:548, 1998.

[98] C. W. Gardiner and P. Zoller. *Quantum Noise*. Springer-Verlag, New York, 2nd edition, 2000.

[99] S. M. Tan. A computational toolbox for quantum and atomic optics. *J. Opt. B: Quant. Semiclass. Opt.*, 1:424, 1999.

[100] I. H. Deutsch and P. Jessen. Quantum state control in optical lattices. *Phys. Rev. A*, 57:1972, 1998.

[101] W. Happer and B. S. Mathur. Effective Operator Formalism in Optical Pumping. *Phys. Rev.*, 163:12, 1967.

[102] I. Carusotto and E. J. Mueller. Imaging of spinor gases. *J. Phys. B*, 37:S115, 2004.

[103] G. Baym. *Lectures on Quantum Mechanics*. W. A. Benjamin, 1969.

[104] A. Silberfarb and I. Deutsch. Continous measurement with travelling wave probes. *Phys. Rev. A*, 68:013817, 2003.

[105] S. E. Hamann. Quantum State Preparation in an Optical Lattice. *PhD Thesis, University of Arizona*, 1998.

[106] A. Silberfarb, P. S. Jessen, and I. H. Deutsch. Quantum State Reconstruction via Continuous Measurement. *Phys. Rev. Lett.*, 95:030402, 2005.

[107] G. A. Smith, A. Silberfarb, I. H. Deutsch, and P. S. Jessen. Efficient Quantum-State Estimation by Continuous Weak Measurement and Dynamical Control. *Phys. Rev. Lett.*, 97:180403, 2006.

[108] R. Boyd. *Radiometry and the Detection of Optical Radiation*. John Wiley and Sons, New York, 1983.

[109] R. J. Glauber. Coherent and Incoherrent States of the Radiation Field. *Phys. Rev. Lett.*, 131:2766, 1963.

[110] D. Chauvat, J. Guena, P. Jacquier, M. Lintz, M. A. Bouchiat, M. D. Plimmer, and C. W. Goodwin. Magnification of a tiny polarisation rotation by a dichroic plate in balanced mode polarimetry. *Optics Communications*, 138:249, 1997.

[111] L. Bouten, J. K. Stockton, G. Sarma, and H. Mabuchi. Scattering of polarized laser light by an atomic gas in free space: a QSDE approach. *unpublished*, 2007.

[112] S. R. de Echaniz, M. W. Mitchell, M. Kubasik, M. Koschorreck, H. Crepaz, J. Eschner, and E. S. Polzik. Conditions for spin squeezing in a cold 87Rb ensemble. *Quantum Semiclass. Opt.*, 7:S548, 2005.

[113] H. M. Wiseman. Quantum trajectories and quantum measurement theory. *Quant. Semiclass. Opt.*, 8:205, 1996.

[114] H. M. Wiseman. Quantum theory of continuous feedback. *Phys. Rev. A*, 49:2133, 1994.

[115] I. Bouchoule and K. Mølmer. Preparation of spin-squeezed atomic states by optical-phase-shift measurement. *Phys. Rev. A*, 66:043811, 2002.

[116] A. S. Sørensen and K. Mølmer. Entangling atoms in bad cavities. *Phys. Rev. A*, 66:022314, 2002.

[117] A. Andre and M. D. Lukin. Atom correlations and spin squeezing near the Heisenberg limit: Finite-size effect and decoherence. *Phys. Rev. A*, 65:053819, 2002.

[118] L. M. Duan, J. I. Cirac, and P. Zoller. Three-dimensional theory for interaction between atomic ensembles and free-space light. *Phys. Rev. A*, 66:023818, 2002.

[119] P. E. Kloeden, E. Platen, and H. Schurz. *Numerical Solution of SDE through Computer Experiments*. Springer-Verlag, New York, 1997.

[120] L. M. Duan and H. J. Kimble. Efficient Engineering of Multiatom Entanglement through Single-Photon Detections. *Phys. Rev. Lett.*, 90:253601, 2003.

[121] A. S. Sørensen and K. Mølmer. Measurement Induced Entanglement and Quantum Computation with Atoms in Optical Cavities. *Phys. Rev. Lett.*, 91:097905, 2003.

[122] R. G. Unanyan, M. Fleischhauer, N. V. Vitanov, and K. Bergmann. Entanglement generation by adiabatic navigation in the space of symmetric multiparticle states. *Phys. Rev. A*, 66:042101, 2002.

[123] A. Andre, A. Sørensen, and M. D. Lukin. Stability of atomic clocks based on entangled atoms. *Phys. Rev. Lett.*, 92:230801, 2004.

[124] C. Ahn, A. C. Doherty, and A. J. Landahl. Continuous quantum error correction via quantum feedback control. *Phys. Rev. A*, 65:042301, 2002.

[125] A. C. Doherty and K. Jacobs. Feedback control of quantum systems using continuous state estimation. *Phys. Rev. A*, 60:2700, 1999.

[126] S. D. Bartlett and H. M. Wiseman. Entanglement Constrained by Superselection Rules. *Phys. Rev. Lett*, 91:097903, 2003.

[127] A. Cabello. Bell's theorem with and without inequalities for the three-qubit Greenberger-Horne-Zeilinger and W states. *Phys. Rev. A*, 65:032108, 2002.

[128] D. C. Brody and L. P. Hughston. Efficient Simulation of Quantum State Reduction. *Journal of Mathematical Physics*, 43:5254, 2002.

[129] H. M. Wiseman, S. Mancini, and J. Wang. Bayesian feedback versus Markovian feedback in a two-level atom. *Phys. Rev. A*, 66:013807, 2002.

[130] W. P. Smith, J. E. Reiner, L. A. Orozco, S. Kuhr, and H. M. Wiseman. Capture and Release of a Conditional State of a Cavity QED System by Quantum Feedback. *Phys. Rev. Lett*, 89:133601, 2002.

[131] N. V. Morrow, S. K. Dutta, and G. Raithel. Feedback Control of Atomic Motion in an Optical Lattice. *Phys. Rev. Lett*, 88:093003, 2002.

[132] T. Fischer, P. Maunz, P. W. H. Pinkse, T. Puppe, and G. Rempe. Feedback on the Motion of a Single Atom in an Optical Cavity. *Phys. Rev. Lett*, 88:163002, 2002.

[133] J. Gambetta and H. M. Wiseman. State and dynamical parameter estimation for open quantum systems. *Phys. Rev. A*, 64:042105, 2001.

[134] H. Mabuchi. Dynamical identification of open quantum systems. *Quantum Semiclass. Opt.*, 8:1103, 1996.

[135] G. L. Bretthorst. *Bayesian Spectrum Analysis and Parameter Estimation.* Springer-Verlag, 1988.

[136] W. T. Reid. *Riccati Differential Equations.* Academic Press, New York, 1972.

[137] J. L. Kirschvink. Uniform Magnetic Fields and Double-Wrapped Coil Systems. *Bioelectromagnetics*, 13:401–411, 1992.

[138] J. R. Buck. Cavity QED in Microspheres and Fabry-Perot Cavities. *PhD Thesis, California Institute of Technology*, 2003.

[139] K. G. Libbrecht and J. L. Hall. A low-noise high-speed diode laser current controller. *Review of Scientific Instruments*, 64:2133, 1993.

[140] C. E. Wieman and L. Hollberg. Using diode lasers for atomic physics. *Review of Scientific Instruments*, 62:1, 1991.

[141] J. H. Shirley. Modulation transfer processes in optical heterodyne saturation spectroscopy. *Opt. Lett.*, 7:537, 1982.

[142] B. Lev. Magnetic Microtraps for Cavity QED, Bose-Einstein Condensates, and Atom Optics. *PhD Thesis, California Institute of Technology*, 2005.

[143] K. Singer, S. Jochim, M. Mudrich, A. Mosk, and M. Weidemuller. Low-cost mechanical shutter for light beams. *Review of Scientific Instruments*, 73:4402, 2002.

[144] P. C. D. Hobbs. *Building Electro-Optical Systems: Making It All Work.* John Wiley and Sons, New York, 2000.

[145] J. Graeme. *Photodiode Amplifiers.* McGraw-Hill, 1995.

[146] D. Budker, D. F. Kimball, V. V. Yashchuk, and M. Zolotorev. Nonlinear magneto-optical rotation with frequency-modulated light. *Phys. Rev. A*, 65:055403, 2002.

[147] J. Guena, D. Chauvat, Ph. Jacquier, M. Lintz, M. D. Plimmer, and M. A. Bouchiat. Differential-mode atomic polarimetry with pulsed lasers: high-precision zero adjustment. *J. Opt. Soc. Am. B*, 14:271, 1997.

[148] M. Bouchiat, D. Chauvat, J. Guena, P. Jacquier, M. Lintz, and M. D. Plimmer. High precision balanced mode polarimetry with a pulsed laser beam. *Optics Communications*, 119:403, 1995.

[149] C. G. Townsend, N. H. Edwards, C. J. Cooper, K. P. Zetie, C. J. Foot, A. M. Steane, P. Szriftgiser, H. Perrin, and J. Dalibard. Phase-space density in the magneto-optical trap. *Phys. Rev. A*, 52:1423, 1995.

[150] K. B. Davis, M. O. Mewes, M. R. Andrews, N. J. van Druten, D. S. Durfee, D. M. Kurn, and W. Ketterle. Bose-Einstein Condensation in a Gas of Sodium Atoms. *Phys. Rev. Lett.*, 75:3969, 1995.

[151] K. Lindquist, M. Stephens, and C. Wieman. Experimental and theoretical study of the vapor-cell Zeeman optical trap. *Phys. Rev. A*, 46:4082, 1992.

[152] C. Monroe, W. Swann, H. Robinson, and C. Wieman. Very cold trapped atoms in a vapor cell. *Phys. Rev. Lett.*, 65:1571, 1990.

[153] K. E. Gibble, S. Kasapi, and S. Chu. Improved magneto-optic trapping in a vapor cell. *Opt. Lett.*, 17:526, 1992.

[154] A. M. Steane, M. Chowdhury, and C. J. Foot. Radiation force in the magneto-optical trap. *J. Opt. Soc. Am. B*, 9:2142, 1992.

[155] T. Walker, D. Sesko, and C. Wieman. Collective behavior of optically trapped neutral atoms. *Phys. Rev. Lett.*, 64:408, 1990.

[156] D. W. Sesko, T. G. Walker, and C. E. Wieman. Behavior of neutral atoms in a spontaneous force trap. *J. Opt. Soc. Am. B*, 8:946, 1991.

[157] V. S. Bagnato, L. G. Marcassa, M. Oria, G. I. Surdutovich, R. Vitlina, and S. C. Zilio. Spatial distribution of atoms in a magneto-optical trap. *Phys. Rev. A*, 48:3771, 1993.

[158] I. Guedes, H. F. Silva Filho, and F. D. Nunes. Theoretical analysis of the spatial structures of atoms in magneto-optical traps. *Phys. Rev. A*, 55:561, 1997.

[159] D. Felinto, L. G. Marcassa, V. S. Bagnato, and S. S. Vianna. Influence of the number of atoms in a ring-shaped magneto-optical trap: Observation of bifurcation. *Phys. Rev. A*, 60:2591, 1999.

[160] D. Felinto, H. Regehr, J. W. R. Tabosa, and S. S. Vianna. Fluctuations in ball- and ring-shaped magneto-optical traps at low densities. *J. Opt. Soc. Am. B*, 18:1410, 2001.

[161] E. L. Raab, M. Prentiss, A. Cable, S. Chu, and D. E. Pritchard. Trapping of Neutral Sodium Atoms with Radiation Pressure. *Phys. Rev. Lett.*, 59:2631, 1987.

[162] W. Ketterle, K. B. Davis, M. A. Joffe, A. Martin, and D. E. Pritchard. High densities of cold atoms in a dark spontaneous-force optical trap. *Phys. Rev. Lett.*, 70:2253, 1993.

[163] C. G. Townsend, N. H. Edwards, K. P. Zetie, C. J. Cooper, J. Rink, and C. J. Foot. High-density trapping of cesium atoms in a dark magneto-optical trap. *Phys. Rev. A*, 53:1702, 1996.

[164] N. Lundblad. *Personal communication*, 2006.

[165] J. Sebby-Strabley, R. T. R. Newell, J. O. Day, E. Brekke, and T. G. Walker. High-density mesoscopic atom clouds in a holographic atom trap. *Phys. Rev. A*, 71:021401, 2005.

[166] M. T. DePue, S. L. Winoto, D. J. Han, and D. S. Weiss. Transient compression of a MOT and high intensity fluorescent imaging of optically thick clouds of atoms. *Optics Communications*, 180:73, 2000.

[167] W. Petrich, M. H. Anderson, J. R. Ensher, and E. A. Cornell. Behavior of atoms in a compressed magneto-optical trap. *J. Opt. Soc. Am. B*, 11:1332, 1994.

[168] A. J. Kerman, V. Vuletić, C. Chin, and S. Chu. Beyond Optical Molasses: 3D Raman Sideband Cooling of Atomic Cesium to High Phase-Space Density. *Phys. Rev. Lett.*, 84:439, 2000.

[169] F. Reif. *Fundamentals of Statistical and Thermal Physics*. McGraw-Hill, New York, 1st edition, 1965.

[170] J. Schoser, A. Batär, R. Löw, V. Schweikhard, A. Grabowski, Yu. B. Ovchinnikov, and T. Pfau. Intense source of cold Rb atoms from a pure two-dimensional magneto-optical trap. *Phys. Rev. A*, 66:023410, 2002.

[171] P. A. Willems and K. G. Libbrecht. Creating long-lived neutral-atom traps in a cryogenic environment. *Phys. Rev. A*, 51:1403, 1995.

[172] Z. Hu. Quantum optics with cold atoms – nonlinear spectroscopy and road toward single-atom trap. *PhD Thesis, California Institute of Technology*, 1995.

[173] M. Takeuchi, T. Takano, S. Ichihara, A. Yamaguchi, M. Kumakura, T. Yabuzaki, and Y. Takahashi. Fast polarimetry system for the application to spin quantum non-demolition measurement. *Appl. Phys. B*, 83:33, 2006.

[174] R. Ruskov, K. Schwab, and A. N. Korotkov. Squeezing of a nanomechanical resonator by quantum nondemolition measurement and feedback. *Phys. Rev. B*, 71:235407, 2005.

[175] I. M. Savukov, S. J. Seltzer, M. V. Romalis, and K. L. Sauer. Tunable Atomic Magnetometer for Detection of Radio-Frequency Magnetic Fields. *Phys. Rev. Lett.*, 95:063004, 2005.

[176] J. H. Müller, P. Petrov, D. Oblak, C. L. Garrido Alzar, S. R. de Echaniz, and E. S. Polzik. Diffraction effects on light˘atomic-ensemble quantum interface. *Phys. Rev. A*, 71:033803, 2005.

[177] L. B. Madsen and K. Mølmer. Spin squeezing and precision probing with light and samples of atoms in the Gaussian description. *Phys. Rev. A*, 70:052324, 2004.

[178] A. Kuzmich and T. A. B. Kennedy. Nonsymmetric Entanglement of Atomic Ensembles. *Phys. Rev. Lett.*, 92:030407, 2004.

[179] S. J. van Enk. Atoms, dipole waves, and strongly focused light beams. *Phys. Rev. A*, 69:043813, 2004.

[180] U. Eichmann, J. C. Bergquist, J. J. Bollinger, J. M. Gilligan, W. M. Itano, D. J. Wineland, and M. G. Raizen. Young's interference experiment with light scattered from two atoms. *Phys. Rev. Lett.*, 70:2359, 1993.

[181] J. E. Lenz. A Review of Magnetic Sensors. *Proceedings of the IEEE*, 78:973, 1990.

[182] S. A. Macintyre. Magnetic Field Measurement. 2000.

[183] J. Ringot, P. Szriftgiser, and J. C. Garreau. Subrecoil Raman spectroscopy of cold cesium atoms. *Phys. Rev. A*, 65:013403, 2001.

[184] F. Primdahl. The fluxgate magnetometer. *J. Phys. E.: Sci. Instrum*, 12:241, 1979.

[185] R. H. Koch, J. R. Rozen, P. Woltgens, T. Picunko, W. J. Goss, D. Gambrel, R. Wiegert, and D. Overway. High performance superconducting quantum interference device feedback electronics. *Rev. Sci. Instrum.*, 67:2968, 1996.

[186] R. Merrill, M. McElhinny, and P. McFadden. *The Magnetic Field of the Earth*. Academic Press, 1st edition, 1998.

[187] G. Klose. Density matrix reconstruction of a large angular momentum. *PhD Thesis, University of Arizona*, 2001.

[188] D. J. Griffiths. *Introduction to Electrodynamics*. Prentice-Hall, 2nd edition, 1989.

[189] F. W. Grover. *Inductance Calculations*. Courier Dover Publications, 2004.

[190] C. J. Dedman, K. G. H. Baldwin, and M. Colla. Fast switching of magnetic fields in a magneto-optic trap. *Review of Scientific Instruments*, 72:4055, 2001.

[191] C. L. Garrido Alzar, P. G. Petrov, D. Oblak, J. H. Muller, and E. S. Polzik. Compensation of eddy-current-induced magnetic field transients in a MOT. *physics/0701251*, 2007.

[192] P. Horowitz and W. Hill. *The Art of Electronics*. Cambridge University Press, 1989.

[193] J. H. Moore, C. C. Davis, and M. A. Coplan. *Building Scientific Apparatus*. Perseus Books, Cambridge, 3rd edition, 2003.

[194] Y. P. Ma and Jr. J. P. Wikswo. Magnetic shield for wide-bandwidth magnetic measurements for nondestructive testing and biomagnetism. *Review of Scientific Instruments*, 62:2654, 1991.

[195] J. McKeever, J. R. Buck, A. D. Boozer, A. Kuzmich, H.-C. Naegerl, D. M. Stamper-Kurn, and H. J. Kimble. State-Insensitive Cooling and Trapping of Single Atoms in an Optical Cavity. *Phys. Rev. Lett.*, 90:133602, 2003.

[196] C. J. Hood, T. W. Lynn, A. C. Doherty, A. S. Parkins, and H. J. Kimble. The Atom-Cavity Microscope: Single Atoms Bound in Orbit by Single Photons. *Science*, 287:1447, 2003.

[197] K. M. Gheri and H. Ritsch. Single-atom quantum gate for light. *Phys. Rev. A*, 56:3187, 1997.

[198] D. A. Steck, K. Jacobs, H. Mabuchi, T. Bhattacharya, and S. Habib. Quantum Feedback Control of Atomic Motion in an Optical Cavity. *Physical Review Letters*, 92:223004, 2004.

[199] D. W. Berry and H. M. Wiseman. Phase measurements at the theoretical limit. *Phys. Rev. A*, 63:013813, 2000.

[200] G. Breitenbach and S. Schiller. Homodyne tomography of classical and non-classical light. *Journal of Modern Optics*, 44:2207, 1997.

[201] M. J. Collett. Exact density-matrix calculations for simple open systems. *Phys. Rev. A*, 38:2233, 1988.

[202] E. A. Hinds and C. Eberlein. Quantum propagation of neutral atoms in a magnetic quadrupole guide. *Phys. Rev. A*, 61:033614, 2000.

[203] E. A. Hinds and I. G. Hughes. Magnetic atom optics: mirrors, guides, traps, and chips for atoms. *J. Phys. D*, 32:R119, 1999.

[204] J. Reichel, W. Hänsel, and T. W. Hänsch. Atomic Micromanipulation with Magnetic Surface Traps. *Phys. Rev. Lett.*, 83:3398, 1999.

[205] J. Denschlag, D. Cassettari, and J. Schmiedmayer. Guiding Neutral Atoms with a Wire. *Phys. Rev. Lett.*, 82:2014, 1999.

[206] C. V. Sukumar and D. M. Brink. Spin-flip transitions in a magnetic trap. *Phys. Rev. A*, 56:2451, 1997.

[207] E. Schrödinger. The ambiguity of the wave function. *Ann. Physik (5)*, 32:49, 1938.

[208] W. Pauli. *Helv. Phys. Acta*, 12:147, 1939.

[209] E. Merzbacher. Single valuedness of wave functions. *Am. J. Phys.*, 30:237, 1962.

[210] R. Blümel and K. Dietrich. Quantum states of neutrons bound in the magnetic field of a rectilinear current. *Phys. Rev. A*, 43:22, 1991.

[211] L. V. Hau, J. A. Golovchenko, and M. M. Burns. Supersymmetry and the Binding of a Magnetic Atom to a Filamentary Current. *Phys. Rev. Lett.*, 75:1426, 1995.