Order-of-magnitude-tighter bound
on the electron electric dipole moment

A Dissertation
Presented to the Faculty of the Graduate School
of
Yale University
in Candidacy for the Degree of
Doctor of Philosophy

by
Zack Lasner

Dissertation Director: David DeMille

May 2019

Abstract

Order-of-magnitude-tighter bound on the electron electric dipole moment

Zack Lasner

2019

The electron's electric dipole moment (eEDM) is a time-reversal- (T-) violating interaction that is generically predicted to have a magnitude near or above the bounds of current experimental sensitivity in extensions to the Standard Model. We have completed an improved measurement of the electron's electric dipole moment with an order-of-magnitude greater sensitivity than the previous best measurement. The result is consistent with no interaction, $|d_e| < 1.1 \times 10^{-29}\, e \cdot \mathrm{cm}$. This upper bound is a factor of 8.6 smaller than the previous bound and correspondingly probes for new particles with masses at $\sim 3 - 30$ TeV, which is ≈ 3 times higher than previously explored in eEDM experiments. In this work, we describe the second-generation ACME experiment, models for and suppression of systematic errors, sources of phase noise, and preliminary work toward a third generation of the ACME apparatus.

Contents

List of Figures

List of Tables

For C.c.

Acknowledgments

First and foremost, I wish to thank my advisor, David DeMille, for his support in tackling every manner of obstacle. The attentive reader will find the ruins of more than one brick wall in these pages; Dave's scientific insight and conscientious guidance are largely responsible for them.

In addition, I owe a great deal to the other ACME PI's, who have made this project possible in the first place and fruitful in the end. John Doyle kindly made Harvard my university away from university and was always willing to provide his guidance or, failing that, coffee. Gerald Gabrielse's invaluable expertise is no less evident in this work by way of his students' many achievements, which have earned almost incessant mention throughout the following pages.

The other students, postdocs, and collaborators on the ACME II experiment—Vitaly Andreev, Daniel Ang, Jonathan Haefner, Nicholas Hutzler, Cole Meisenhelder, Brendon O'Leary, Cristian Panda, and Xing Wu—have, of course, been absolutely essential to this project. I have particular appreciation for Adam West and Elizabeth West for graciously taking on the job of training me when I was at my most useless to anybody.

This work was supported by the National Science Foundation. I am grateful to them and the U.S. tax payers who have, to my amazement, given up their own money so that foundational research like this can happen.

Outside of the ACME collaboration, Toshihiko Shimasaki made my first scientific experiences in graduate school both fun and productive. Toshi—you had to teach me a lot of things, huh?

On a less scientific note, I would like to thank several of my roommates during the past six years, who have made my time in graduate school inestimably more rewarding. Thank you to Benjamin Brubaker for your enthusiasm to pierce the heart of any matter. To William Sweeney for the lingering early-morning discussions and for keeping our blackboard well populated. To Arolyn Conwill for our cooking "together" and your persistent humor. And to my partner, Cheng Cheng, for your curiosity and grounded values that have been a continual inspiration to me—thank you.

Chapter 1

Introduction

Many foundational elements of the Standard Model theory of particle physics, including the seminal Higgs mechanism papers of 1964 [1–3], the development of electroweak gauge symmetry in the 1960's [4, 5], and the independent discoveries of asymptotic freedom in 1973 [6, 7], were developed over forty years ago. By the end of that period, the first collider using two circulating hadron beams was only a few years old and could achieve beam energies of no more than 15 GeV. Yet this already represented a major advance over previous accelerators, which used fixed targets and therefore had far lower center-of-mass collision energies [8]. With further advances in high-energy physics techniques made throughout the subsequent years, the Standard Model came under an onslaught of rigorous tests and was validated in every case. Major experimental milestones include the discovery of the charm quark in 1974 [9, 10], the tau lepton in 1975 [11], the bottom quark in 1977 [12], gluons in 1979 [13], the W and Z bosons in 1983 [14, 15], the top quark in 1995 [16, 17], the tau neutrino in 2001 [18], and the Higgs boson in 2012 [19]. To date, no laboratory measurement has stood in definitive contradiction with the Standard Model, despite four decades of technical developments including probes of energy scales three orders of magnitude higher than were possible when the Standard Model was being completed.

However, not all observations are laboratory observations, and cosmological measure-

ments have conclusively shown that the Standard Model, together with general relativity, provides an incomplete description of the natural world [20]. Many anomalies—appearing, for example, in galactic structure and the cosmic microwave background–can be accounted for by positing the existence of "dark matter" whose microscopic nature remains largely unspecified.

Since the Standard Model describes "low-energy" particles (which can be readily produced) and their "strong" non-gravitational couplings, new physics could appear in the form of massive particles ($m \gtrsim 100 \, \mathrm{GeV}$), weakly coupled particles (e.g., with coupling constants at inverse energy scales $g \ll G_F^{1/2} \sim 10 \, \mathrm{GeV}^{-1}$), or modifications to general relativity. As merely representative examples of each case, one may consider supersymmetric extensions to the Standard Model in which new particle masses are expected up to several TeV [21], light QCD axions (masses $m \sim \mathrm{meV}$, corresponding to coupling constants $g_{a\gamma} \sim 10^{-13} \, \mathrm{GeV}^{-1}$) [22], and modified Newtonian dynamics in which gravitational forces are reduced in the low-acceleration limit [23].

In the present work, we are primarily interested in the possibility of high-energy particles that interact, perhaps indirectly, with the electron and thereby modify its properties. Electrons are, of course, utterly abundant. Furthermore, laser technology allows us to manipulate electronic states in atoms and molecules with extreme precision. As we will see, these two features–large available numbers and precise control–make the electron a powerful window into the physics of other, higher-energy particles.

In particular, we have made a measurement of the electron's electric dipole moment (eEDM), which probes physics at the ~ 10 TeV energy scale. This chapter will give an introduction to EDM interactions and their connection with theories beyond the Standard Model (BSM), followed by a historical overview of EDM measurements and the measurement technique used by the Advanced Cold Molecule Electron EDM (ACME) collaboration.

1.1 Electric dipole moments of fundamental particles

An electric dipole moment d is a property of a particle defined by the interaction Hamiltonian $H_d = -\vec{d} \cdot \vec{\mathcal{E}}$, where $\vec{\mathcal{E}}$ is an electric field. In classical electromagnetic theory, such

an interaction arises between a spatial charge distribution and an external electric field, such that the dipole moment $\vec{d}$ lies along a spatial axis of charge separation. For example, a charge distribution consisting of two point charges, $+q$ and $-q$, is associated with an electric dipole moment that points from the negative to the positive charge. By contrast, in a fundamental particle, there is no special polar vector (i.e., a vector that is odd under parity) to define the axis of a dipole moment $\vec{d}$. Therefore, a fundamental dipole moment must lie along or against the particle's spin (axial vector) axis, $\vec{S}$. No fundamental particle EDM has been observed to date.

Much like the electron's magnetic moment, an eEDM would be an intrinsic property, and should ultimately not be thought of in terms of a classical separated charge distribution. To belabor this point, note that the magnetic moment interaction is $H_\mu = -\vec{\mu} \cdot \vec{\mathcal{B}}$ with $\vec{\mu} \equiv -\mu_{\mathrm{B}} g_s \vec{s}$, where μ_{B} is the Bohr magneton, $g_s \approx 2$ is the electron g-factor, and $\vec{s}$ is the spin of the electron. (Throughout this thesis, I will use natural units, $\hbar = 1$.) In the language of quantum field theory (QFT), the electron is an irreducible representation of the Lorentz group and therefore has very particular properties under rotations, which are fully defined by the spin vector $\vec{s}$ [24, Sec. 3.1]. If it were to have some independent vector intrinsically associated with it ($\vec{\mu} \nparallel \vec{s}$), then it would cease to have the correct transformations, and this is simply not allowed. By the same reasoning, we must also find $\vec{d_e} \propto \vec{s}$, where $\vec{d_e}$ is the electric dipole moment vector of the electron. We let $H_{d_e} = -\vec{d_e} \cdot \vec{\mathcal{E}} \equiv -2 d_e \vec{s} \cdot \vec{\mathcal{E}}$ so that $H_{d_e} = -d_e |\mathcal{E}|$ when $\vec{s} \parallel \vec{\mathcal{E}}$ for a spin-1/2 particle.

The relationship $\vec{d_e} \propto \vec{s}$ is properly interpreted as an operator equation, but it can also be instructive to demonstrate this relationship directly in the context of expectation values. Recall the Wigner-Eckhart theorem, $\langle j', m_{j'} | T_q^k | j, m_j \rangle \propto \langle j', m_{j'} | k, q; j, m_j \rangle$, where T_q^k is the q-component of a rank-k tensor in the spherical basis with $q \in \{k, k-1, \cdots, -k\}$, while $|j, m_j\rangle$ is a state with total angular momentum j and projection along the z-axis of m_j (and likewise for $|j', m_{j'}\rangle$), and $\langle j', m_{j'} | k, q; j, m_j \rangle$ is a Clebsch-Gordan coefficient [25]. For an electron with, for example, $m_s = \frac{1}{2}$, the expectation value of the energy shift is

$$
\begin{aligned}
\langle H_{d_e} \rangle &= -\vec{\mathcal{E}} \cdot \langle s = \tfrac{1}{2}, m_s = \tfrac{1}{2} | \vec{d_e} | s = \tfrac{1}{2}, m_s = \tfrac{1}{2} \rangle \\
&= -\mathcal{E}_z d_{e,z},
\end{aligned}
\tag{1.1}
$$

since the x and y components of any vector operator have only off-diagonal matrix elements. Such an electron therefore experiences no energy shift associated with $\mathcal{E}_x$ or $\mathcal{E}_y$; its dipole moment lies along z, just like its spin.

1.1.1 P- and T-violation in the Standard Model and beyond

The fact that the dipole moment vector is along the electron spin has profound consequences for the transformation properties of H_{d_e} under parity transformation, P, and time reversal, T. Since $\vec{s}$ is an angular momentum, it is odd under time reversal but not parity transformation. On the other hand, an electric field is odd under parity transformation but not time reversal. Therefore, H_{d_e} transforms as follows:

$$
\begin{aligned}
T : H_{d_e} &\rightarrow -2d_e(-\vec{s}) \cdot \vec{\mathcal{E}} = -H_{d_e} \\
P : H_{d_e} &\rightarrow -2d_e \vec{s} \cdot (-\vec{\mathcal{E}}) = -H_{d_e}.
\end{aligned}
\tag{1.2}
$$

This is, perhaps, the most crucial difference between the EDM of a fundamental particle and the EDM of a classical charge distribution (in which $\vec{d}$ is even under T and odd under P). It also provides an "explanation" for the non-observation of EDM's: the Standard Model does not contain enough T-violating parameters to generate electric dipole moments at observable levels.[1] On the other hand, EDMs are a generic consequence of theories containing both T and P violation [26–28].

To see the significance of EDM's in the Standard Model and its possible extensions, I will examine possible sources of P and T violation. In what follows, I will equivocate between T and CP (symmetry or violation), where C represents charge conjugation. This is justified by the CPT theorem, which states that CPT is an exact symmetry under extremely general conditions. Therefore, if T is violated (conserved), then CP must be violated (conserved) in equal measure to preserve CPT. A heuristic treatment of the CPT theorem can be found in [29, Sec. 2.1]; technical treatments are in [30–32].

Parity violation is ubiquitous in the weak interaction of the Standard Model because

---

1. This gets the logic backwards from the history: in the case of the neutron EDM, d_n, stringent limits have constrained the relevant T-violating phase in the Standard Model to $|\theta_{QCD}| \lesssim 10^{-10}$ [26, Sec. 2.3]. The apparent unlikelihood of such a small dimensionless number is known as the Strong CP Problem .

W bosons interact only with left-handed matter fields, while the degree of chirality in interactions with Z bosons is parametrized by the Weinberg angle θ_W, where $\sin^2 \theta_W \approx 0.23$ [33, Ch. 9]. As a result, P violation is not a "limiting ingredient" to generating EDM interactions in the SM.

T violation also occurs in the weak interaction Lagrangian, characterized by a phase in the Cabbibo-Kobayashi-Maskawa (CKM) matrix that appears in the charged current interaction of the quarks [34]. However, any effect of this phase on the electron's properties must be mediated by quarks and W bosons, and the lowest-order non-zero contribution to the eEDM is a three-loop Feynman diagram. Khriplovich and Pospelov showed that all three-loop diagrams exactly cancel [35]; following their proof, they comment, "We cannot get rid of the feeling that this simple result... should have a simple transparent explanation. Unfortunately, we have not been able to find it."[2] The most recent estimate for the value of d_e, arising from four-loop diagrams involving the CP-violating phase in the CKM matrix, is $|d_e| \sim 10^{-44}\, e \cdot \text{cm}$ [38], about fifteen orders of magnitude below current experimental sensitivity.[3]

The final CP-violating phase in the SM, θ_{QCD}, is constrained by ^{199}Hg and neutron EDM experiments to be less than 10^{-10} [26, 41, 42]. Although the size of θ_{QCD} is not strictly in conflict with the SM, its smallness plausibly calls for an explanation from physics outside the SM. Explaining the anomalous size of θ_{QCD} is known as the Strong CP Problem and is regarded as one of the major problems in theoretical physics [43, Sec. 5]. The origin of the Strong CP Problem in an EDM measurement demonstrates the suitability of EDM

2. Other authors give a heuristic argument for the cancellation in terms of the GIM mechanism [36], but do not provide a formal proof [37]. The cancellation has its mathematical root in the antisymmetry of the reduced Jarlskog invariant [38], but if this isn't transparent enough for Khriplovich and Pospelov then it certainly isn't for me. For an alternative presentation of and commentary on Khriplovich and Pospelov's result, independently discovered in 1993, also see [39].

3. A couple of brief asides may clear up some possible confusion. Most earlier references give estimates of order $|d_e| \sim 10^{-38}\, e\cdot$cm, an estimate that can be traced to [40], which disclaims that "it is difficult however to make a halfway accurate estimate." In particular, most of the discrepancy with [38] can be accounted for by the inclusion of a highly uncertain numerical prefactor in the earlier work: "The fact that the factor...is quite large partly reflects...the **possible** presence of large logarithms like $\ln^3(m_t^2/m_u^2) \approx 7\times10^3$" (emphasis mine). Furthermore, taking the SM value of $|d_e| \sim 10^{-44}\, e \cdot$cm as given, we will see that the dominant signature of T violating physics in the ACME experiment arises in the SM from a scalar-pseudoscalar electron-nucleon coupling, rather than an eEDM. In any case, the SM background for the ACME measurement is many orders of magnitude below current sensitivity.
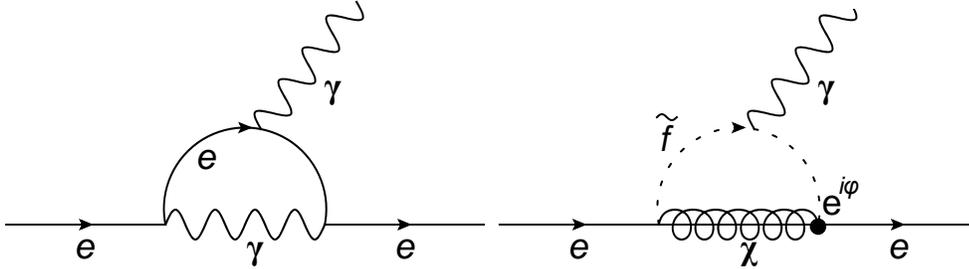
Figure 1.1: (Left) Leading contribution to the anomalous magnetic moment of the electron. (Right) An example one-loop contribution to the eEDM arising in supersymmetric models [26]. Here e is an electron, γ is a photon, χ is a chargino, and $\tilde{f}$ is a sfermion. The CP-violating phase ϕ is introduced at one of the interaction vertices involving $\tilde{f}$.

experiments to probe surprising features of fundamental physics. Nevertheless, the neutron EDM is far more sensitive to θ_{QCD} than the eEDM, which receives a contribution of only $d_e \sim 10^{-28}\, e \cdot \text{cm} \times \theta_{\text{QCD}} < 10^{-38}\, e \cdot \text{cm}$, far below current sensitivity [44]. It is interesting to note, though, that experimental constraints on θ_{QCD} are consistent with a contribution to d_e on the same order as the contribution from the CKM phase (or much larger, depending on the estimate of the CKM contribution).

Additional contributions to the eEDM appear in the minimal extensions to the classic SM that are necessary to account for neutrino masses. The particular nature of the neutrino is unknown, including whether it is a Majorana or Dirac fermion. If the neutrino is a Dirac particle, then there is an analogue to the CKM matrix in the lepton sector known as the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix, which contains additional T-violating phases.[4] However, contributions from these phases are suppressed by the small masses of the neutrinos, and the resulting contributions to d_e are at the $10^{-107}\, e \cdot \text{cm}$ level [45, Sec. 4.1]. On the other hand, simple models of Majorana neutrinos give more complicated results [45], with generic estimates of $d_e \sim 10^{-43}\, e \cdot \text{cm}$ near the CKM contribution and fine-tuned estimates of $d_e \sim 10^{-33}\, e \cdot \text{cm}$ only a few orders of magnitude below current experimental sensitivity [46].

One might get the impression from this summary that it is difficult to construct enough T violation to generate an electron EDM at the $\sim 10^{-30}\, e \cdot \text{cm}$ level. In fact, the SM is rather special with regard to its dearth of T violation. In extensions to the SM, it

---

4. I always use "SM" to refer to this model rather than the classic SM with massless neutrinos.

6

is straightforward to generate a relatively large eEDM. As a simple example, consider a theory with a new particle X, CP-violating phase ϕ, and coupling constant to the electron f. A one-loop Feynman diagram, analogous to the dominant contribution to the anomalous magnetic dipole moment, $(\mu_e - \mu_\mathrm{B})/2 = (\alpha/2\pi)\mu_\mathrm{B}$, could contribute to the dipole moment d_e at the order of [26, Sec. 6]

$$d_e \sim \left[\left(\frac{f}{e} \right)^2 \sin\phi \left(\frac{m_e}{m_X} \right)^2 \right] \left(\frac{\alpha}{2\pi} \right) \mu_\mathrm{B} \sim 10^{-3} \left(\frac{m_e}{m_X} \right)^2 \mu_\mathrm{B}, \tag{1.3}$$

where the square braces provide the scaling relative to the anomalous magnetic dipole moment and the right-most side applies in the naive case $f \sim e$ and $\phi \sim 1$ (see Fig. 1.1). This relationship is written in natural units. If we assume new particle masses comparable to those probed by the Large Hadron Collider (LHC), $m_X \sim 20$ TeV, then we can compute $d_e \sim 10^{-29}\, e$ cm, at the level probed by the recently completed ACME measurement, "ACME II." In-depth calculations have allowed eEDM searches to significantly constrain the parameter spaces of more complicated theories [28]. Increased measurement precision will constrain theories further and, in the possible event of a non-null result, could even provide the first definitive observation of physics beyond the Standard Model.

While Eq. 1.3 gives a "generic" relationship between a new physics mass scale and $|d_e|$, it is also of interest to note the *largest* mass scale that can be probed by EDM searches. The previous estimate suffers from the large suppression factor $(m_e/M_X)^2$; however, in models with large flavor violation, one factor of m_e may be replaced by m_τ, where $m_\tau \approx 3 \times 10^3\, m_e$ is the tau lepton mass [47].[5] Applying this enhancement directly to Eq. 1.3, we would anticipate sensitivity to $M_X \sim 1$ PeV. In fact, models of interest with large flavor violation tend to predict mass scales of a few hundred TeV for an eEDM near current sensitivity [47, 48].

Atomic and molecular electric dipole moment measurements are generically sensitive to two additional sources of T violation: the scalar-pseudoscalar electron-proton coupling $C_{S,p}$ and electron-neutron coupling $C_{S,n}$. For simplicity of discussion, we define $C_S(N, Z) =$

---

5. Note that flavor violation is not especially exotic, as neutrino oscillations are already present in any version of the SM consistent with observations (i.e., in which neutrinos have mass).

$(ZC_{S,p} + NC_{S,n})/(Z + N)$, a weighted average of $C_{S,p}$ and $C_{S,n}$ with the weights given by the number of protons Z and neutrons N, respectively, in an atomic or molecular system of interest. In the SM, $C_S \sim 10^{-18}$ can be generated by the CKM phase, corresponding to $d_e^{\text{equiv}}(C_S) \sim 10^{-38}\, e \cdot \text{cm}$ [38]. Here, $d_e^{\text{equiv}}(C_S)$ is the size of an electron EDM needed to mimic the effects of C_S in an atomic or molecular T violation measurement such as ACME. It has also been estimated in [49] that a contribution $C_S \sim 0.06 \times \theta_{\text{QCD}}$ could dominate the CKM contribution. Assuming the bound $\theta_{\text{QCD}} \sim 10^{-10}$, we would infer $d_e^{\text{equiv}}(C_S) \sim 10^{-31}\, e \cdot \text{cm}$, remarkably near present-day sensitivity. (However, as neutron and diamagnetic species EDM experiments also advance, it is unlikely that eEDM experiments could lead in sensitivity to θ_{QCD} for the foreseeable future.) Likewise, it is important to realize that in some BSM theories, $d_e^{\text{equiv}}(C_S) \gg d_e$ [28, Sec. 4.2].

Finally, it is useful to point out that in some BSM theories, two-loop contributions to the eEDM dominate. In these cases, current eEDM measurements can probe energy scales up to several TeV rather than several tens of TeV [50, 51].

1.1.2 T violation is generic

Careful treatment of particular BSM theories is far outside the scope of this work; however, it may be instructive to treat T violation in generic BSM models heuristically. Toward this end, I will consider an effective field theory (EFT), in which high-energy phenomena are modeled with new interactions among SM particles. Fermi famously used such an approach in 1933 to model beta decay via a four-fermion contact interaction (i.e., an incoming neutron becomes a proton, electron, and antielectron neutrino at a single vertex) [52, 53]. Although beta decay is properly explained by W boson exchange in the weak interaction [33], such mediating particles need not appear explicitly in order to account for low-energy observations.

If the SM is an effective field theory of a more complete high-energy theory, then we can model low-energy phenomena using only SM particles, just as Fermi modeled beta decay without explicit reference to the W boson. In this case, it is instructive to categorize BSM effects by the "mass dimension" of an EFT interaction. An interaction among particles in quantum field theory can be expressed by a term in the Lagrangian density $\mathcal{L}$, such that the
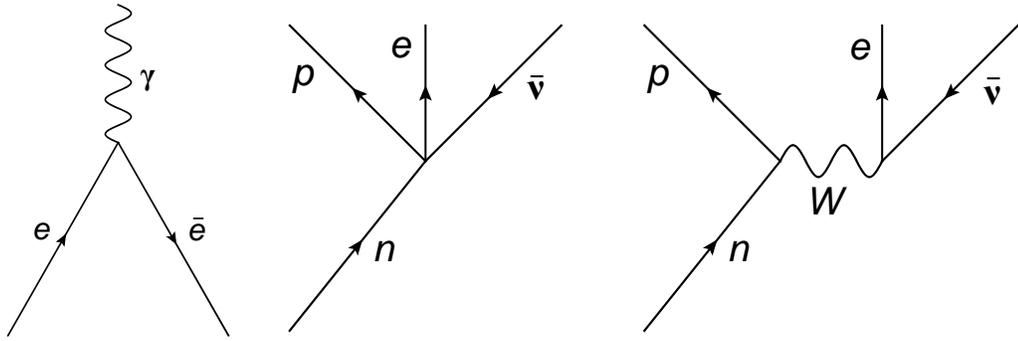
Figure 1.2: (Left) Feynman diagram representing electron-photon interaction in QED. (Center) Beta decay in Fermi's model. (Right) Beta decay in the Standard Model.

action is defined by $S = \int \mathcal{L} d^3x\, dt$. In natural units, $[S] = [\hbar] = [1]$ and $[x] = [t] = [m]^{-1}$, so the Lagrangian density has units of $[\mathcal{L}] = [m]^4$. Interaction terms contain products of field operators corresponding to the kinds of particles that can interact at tree level.

For example, the QED interaction term between a fermion and photon is $eA_\mu \bar{\Psi} \gamma^\mu \Psi$, where e is the elementary charge, A_μ is the electromagnetic four-potential operator, $\bar{\Psi}$ is an antifermion spinor, γ^μ is a (unitless) Dirac matrix, and Ψ is a fermion spinor. This is graphically represented in a Feynman diagram, where a fermion, antifermion, and photon meet at a vertex (see Fig. 1.2 left).

Each type of field operator (scalar, fermion, gauge boson, etc.) has characteristic units, and in natural units these are expressed as powers of mass. Correspondingly, the units, or "mass dimension," of any product of field operators can be straightforwardly calculated. In the example of the QED vertex, we find $[A\bar{\Psi}\Psi] = [m]^4$ as required for a term in the Lagrangian density.[6]

In an EFT, the effective low-energy interaction is composed of products of operators with mass dimension greater than four, and therefore requires a dimensionful "coupling constant" g that characterizes the interaction strength. In the example of a four-fermion interaction such as $[\bar{\psi}\bar{\psi}\psi\psi] = [m]^6$, the coupling constant must have units of $[m]^{-2}$ so that $[g\bar{\psi}\bar{\psi}\psi\psi] = [m]^4$ has the appropriate units for a Lagrangian density (see Fig. 1.2 center). These mass dimensions are generated, from the perspective of the high-energy theory, by

---

6. Fermions carry $[m]^{3/2}$ and the four-potential carries $[m]^1$.

suppressing the effective interaction with products of high-energy particle masses. The coupling constant in Fermi's four-fermion interaction is $G_F \propto 1/M_W^2$, where M_W is the mass of the W boson. At energy scales significantly below $M_W \sim 100$ GeV, we can describe the strength of beta decay in terms of only G_F, without any reference to the W boson (Fig. 1.2 right).

EDMs can be generated from EFT interactions such as $\psi\bar{\psi}hF$, where h is a higgs field and F is the electromagnetic field tensor, among other mechanisms [54]. This is a dimension-six operator and must therefore be suppressed by two powers of high-energy particle masses. It turns out that there are no T-violating dimension-five operators, but (counting distinctions among flavors) there are 2499 dimension-six EFT operators using SM particles [55]. Among these, 1149 exhibit T violation. While the observational effects of any such operator must be evaluated within the context of a particular model, this simple counting argument serves to illustrate the ubiquity of T violation in BSM theories.

1.1.3 Empirical support for T violation

In addition to the foregoing theoretical arguments, searches for T-violating BSM physics are strongly motivated by cosmological observations. In particular, it is obvious that our immediate surroundings in the solar system contain only trace amounts of antimatter, by virtue of the fact that we are not instantly annihilated. Careful astronomical observations reveal the same paucity of antimatter in the rest of the observable universe. Sakharov has identified several conditions necessary to account for the excess of matter over antimatter, assuming the universe initially contained equal quantities of each [56]:

1. There must exist processes that violate baryon number B (i.e., the difference between the number of baryons and the number of antibaryons); this is obvious if $B(t=0)=0$ but $B(t>0) \neq 0$.

2. There must exist C violation. If there exists a process that converts ψ to $\bar{\psi}$ at rate $R(\psi \to \bar{\psi})$, then the inverse process must occur at a greater rate, $R(\bar{\psi} \to \psi) > R(\psi \to \bar{\psi})$.

3. There must exist CP violation (which is equivalent to T violation provided CPT is

an exact symmetry). This closes a loophole in the previous condition: even if, e.g., $R(\bar{\psi}_L \to \psi_L) > R(\psi_L \to \bar{\psi}_L)$, the overall production of matter and antimatter would be balanced by the CP-inverted relationship $R(\psi_R \to \bar{\psi}_R) > R(\bar{\psi}_R \to \psi_R)$ *unless CP* is also violated.

4. The evolution of the early universe must not occur in thermal equilibrium. In thermal equilibrium, the relative numbers of matter and antimatter particles are governed by the Boltzmann factors, $\exp[-m_\psi c^2/kT]$ and $\exp[-m_{\bar{\psi}} c^2/kT]$. Since matter and antimatter particles have the same mass, they would be thermally produced in equal proportion.

Naturally, the formal proof of these conditions is more complicated than the remarks above.

We are particularly interested in understanding how condition (3) might be satisfied; see [57] for a general overview of the approaches discussed here. The T-violating effects in the SM are believed to be insufficient to account for the degree of matter-antimatter asymmetry in the universe, necessitating BSM theories. Alternatively, one might look for loopholes in Sakharov's conditions. The most obvious candidate is that perhaps the universe had a matter anti-matter imbalance as a matter of initial conditions, rather than dynamically generating the imbalance. While this could potentially be consistent with observations, it is strongly disfavored by inflationary cosmological models, which, though subject to some degree of controversy [58], currently represent the dominant cosmological paradigm (see [59] for a favorable review of the literature). Another attempt to evade Sakharov's conditions is to postulate that matter and antimatter are equally prevalent over sufficiently large spatial volumes, but that they are to be found in separate domains of the universe. Analysis of the cosmic microwave background suggests that any such domains, if they exist, would have to be larger than the observable universe (a hypothesis which could presumably never be contradicted by observations). Therefore, the most parsimonious explanation is that the baryon asymmetry was generated via some BSM CP-violating process.[7]

---

7. I do not wish to overstate this case: if we continue to fail to observe signatures of BSM physics, or a compelling alternative to inflation is discovered, or some other theoretical loophole is found, then at some point the weight of evidence may turn against the arguments presented here. So while there *could* be no eEDM above the SM value, for the time being the motivations for EDM searches appear quite robust.

1.1.4 Allowed electromagnetic interactions

I would like to close this section by discussing a motivation for eEDM searches that is entirely independent of T violation: namely, that the EDM is the only allowed electromagnetic interaction that has not yet been observed for a fundamental spin-1/2 particle.

The well-known electromagnetic moments of the electron are the electric monopole (charge) and magnetic dipole moments. In general, an electromagnetic source (charge and current) distribution can be characterized, far from all sources, as an infinite series of electric and magnetic multipole moments. In the near field, an infinite series of toroidal multipole moments must also be included. For example, in the static limit, the toroidal dipole moment is equivalent to an anapole moment and couples directly to the external current density, $H_a = -\vec{a} \cdot \vec{J}$ [60].

However, a particle's spin places constraints on the non-zero electromagnetic moments it can have. Just as a spin-1/2 particle cannot support independent magnetic and electric dipole moments (but rather they must be aligned or anti-aligned with respect to each other), it cannot support quadrupole or higher moments. In particular, only the electric charge and three dipole moments may be non-zero [61].

More formally, the Lagrangian density of the classical electromagnetic field in the presence of matter is $\mathcal{L} = -(1/4)F_{\mu\nu}F^{\mu\nu} + j_\mu A^\mu$, where $F_{\mu\nu}$ is the electromagnetic field tensor, j_μ is the external current density, and A^μ is the four-potential. In quantum electrodynamics (QED), F and A are promoted to operators, and $j^\mu_{\text{QED}} = e\bar{\Psi}\gamma^\mu\Psi$ is the current density operator. The QED current density describes, at tree level, the coupling of an elementary charge and magnetic dipole moment equal to the Bohr magneton. Higher-order loop corrections generate the anomalous magnetic moment. However, additional terms in the Lagrangian of a theory, $\mathcal{L} = \mathcal{L}_{\text{QED}} + \mathcal{L}_{\text{add.}}$, can modify the effective coupling of a particle to the electromagnetic field so that $j^\mu_{\text{eff}} \neq j^\mu_{\text{QED}}$. If we require the coupling of a Dirac particle to be Lorentz covariant and invariant under electromagnetic gauge transformations, then only four kinds of couplings are possible, corresponding to the charge, magnetic dipole moment, electric dipole moment, and anapole moment [61].

The value of the electron's electric charge is protected from corrections by the Ward

identity, which is a consequence of gauge invariance [24, Sec. 7.5]. Experimentally, any surviving correction could likely be inferred by comparing the proton and electron charges, which have been established to differ by no more than 1 part in 10^{20} [62]. The anomalous magnetic moment has been measured at better than the 10^{-9} level [63] and agrees with eighth-order perturbation theory calculations using QED [64].

In the case of the electron, the numerical value of the anapole moment is not invariant under electroweak gauge transformations and therefore can't be regarded as a physically meaningful quantity [61][8]. Therefore, the only unmeasured purely electromagnetic moment of the electron allowed by electroweak gauge invariance and Lorentz invariance is the electric dipole moment.

1.2 Brief history of EDM measurements

1.2.1 Separated oscillatory field measurement

The EDM Hamiltonian $H_{d_e} = -(2d_e)\vec{s}\cdot\vec{\mathcal{E}}$ has the same form as the magnetic dipole moment Hamiltonian $H_\mu = +(g_S\mu_\mathrm{B})\,\vec{s}\cdot\vec{\mathcal{B}}$, up to the numerical prefactors [9]. In 1949, Norman Ramsey developed a technique known as "separated oscillatory fields" to measure nuclear magnetic moments [69, 70]. Most eEDM experiments have used some variant of this technique since. The method is simplest to understand in the context of a classical magnetic moment or, equivalently, a two-level quantum system for which the orientation on the Bloch sphere replaces the orientation of the classical angular momentum vector. There are five steps: (1) A beam of particles is prepared in an eigenstate (e.g., by a magnetic state selector that deflects undesired spin projections) along the z-axis of the Bloch sphere; (2) a $\pi/2$ pulse is

---

8. Confusingly, a different gauge-invariant quantity can be defined and is sometimes called the anapole moment [65]. However, this latter quantity can't be viewed strictly as an electromagnetic coupling. Nevertheless, nuclear anapole moments are of interest, in part because the relative contribution of electromagnetic and weak interactions to the anapole moment in a particular gauge have different scalings with the atomic weight of a nucleus. To date, the nuclear anapole moment has been measured in only one species [66].

9. The convention for H_d is that $\langle m_s = s|H_d|m_s = s\rangle = -d_e\mathcal{E}$, independent of the magnitude of s, while for a magnetic moment (consistent with the classical interpretation of a spinning charge), the convention is that $\langle m_s = s|H_\mu|m_s = s\rangle \propto s$. Further, because the electron is negatively charged, $\hat{\mu} = -\hat{s}$, with a convention that $g_S > 0$. Confusingly, we always use a molecular Zeeman interaction $H_\mu = -g\mu_\mathrm{B}M\mathcal{B}$ for a z-aligned magnetic field: note the sign difference in the interaction for a positive g-factor (e.g., see [67]). I am, of course, not the first to notice that this discrepancy is unfortunate; a rather thorough critique of the electron g-factor sign convention can be found in [68].

applied by an electromagnetic field (e.g., from radio-frequency coils) oscillating at frequency ω, populating a coherent superposition of the two eigenstates that lies in the xy-plane of the Bloch sphere; (3) the molecules traverse a region with constant field along the z-axis that causes the state to precess in the xy-plane of the Bloch sphere at Larmor frequency $\omega_{\rm L}$; (4) another $\pi/2$ pulse is applied at frequency ω, mapping the Bloch sphere azimuthal angle (precession phase) onto the polar angle (state population) according to the frequency difference $\omega_{\rm L} - \omega$; (5) the population in each state is recorded (e.g., by another magnetic state selector). This allows for a precise determination of the precession frequency $\omega_{\rm L}$.

The experimental protocol used in the ACME experiment actually differs from the classic Ramsey scheme in important respects, and will be described in detail in Sec. 2.2. The state preparation and readout (roughly analogous to steps 1-2 and 4-5, respectively, in a Ramsey measurement) are achieved in ACME via laser-induced electronic transitions. The constant-field, or "interaction" region (analogous to step 3 above) contains both static electric and magnetic fields. The signature of an eEDM is a change in precession frequency upon reversal of the relative direction between the electron's spin and electric field experienced by the electron.

Application of the energy-time uncertainty principle gives the correct scaling for the measurement precision [29, Sec. 3.1] as follows. The observed precession frequency ω has an uncertainty $\Delta\omega$ for a single-particle measurement given by $\Delta\omega\,\tau \sim 1$, where τ is the precession time. Since the EDM-induced precession is given by $\omega = -\vec{d_e} \cdot \vec{\mathcal{E}}$, we see that $\Delta d_e \sim \hbar/(\mathcal{E}\,\tau)$. Ideally, the noise is dominated by quantum projection, so the uncertainty in the eEDM for a measurement of N particles is at least roughly $\Delta d_e \sim \hbar/(\mathcal{E}\,\tau\,\sqrt{N})$. In practice, only a small fraction ϵ of molecules is typically detected, so the dominant noise source is shot noise, e.g., from fluorescence photons. In this case, the precision scales as $1/\sqrt{N\epsilon} \equiv 1/\sqrt{N_{\rm det}}$. For this reason, we simply consider the number of detected molecules directly rather than the number of molecules involved in the measurement, N. Further, $N = \dot{n}T$ is constrained by a practical integration time T (typically a few weeks) for a given achievable preparation rate $\dot{n}$. An ideal system maximizes the product $\mathcal{E}\,\tau\,\sqrt{\dot{n}}$ while remaining robust against systematic errors.

Atom	Z	S	Ref.
Rb	37	28	[74]
Cs	55	133	[74]
Tl	81	-585	[75]
^{129}Xe	54	-8×10^{-4}	[76]
^{199}Hg	80	1×10^{-2}	[76]

Table 1.1: Relativistic enhancement factors, S, for important atoms.

1.2.2 Systems for eEDM measurements

Searches for the eEDM cannot be performed precisely with free electrons since the applied electric field responsible for the EDM energy shift will accelerate the electrons out of the interaction region, thus dramatically limiting the precession time per electron. In order to obtain useful observation times, we must therefore use bound electrons. A naive computation of the expected interaction energy in an electrostatically bound system would give $\langle H_{d_e} \rangle = -\langle \vec{d_e} \cdot \vec{\mathcal{E}} \rangle$, where $\langle \rangle$ indicates the quantum-mechanical average over an electron orbit. Since the electron spin is typically fixed over an orbit, one might expect $\langle H_{d_e} \rangle = -\vec{d_e} \cdot \langle \vec{\mathcal{E}} \rangle = 0$, where $\langle \vec{\mathcal{E}} \rangle = 0$ is a necessary condition of electrostatic binding. This result, in the context of a more sophisticated calculation using the Dirac equation, is known as Schiff's theorem [71].

Relativistic effects modify the argument above because the dipole moment vector d_e is contracted along the direction of the electron's momentum and is thus dependent on the electron's position in the Coulomb potential from the nucleus.[10] As a result, $\langle \vec{d_e} \cdot \vec{\mathcal{E}} \rangle \neq \langle \vec{d_e} \rangle \cdot \langle \vec{\mathcal{E}} \rangle$ and an EDM interaction may produce a non-zero first-order energy shift in an electrostatically bound system [72, 73]. The effective electric field is defined by $d_e^{\text{rest}} \mathcal{E}_{\text{eff}} = \langle \vec{d_e} \cdot \vec{\mathcal{E}} \rangle$. In many cases, the magnitude of the effective electric field can be significantly enhanced beyond an applied laboratory field.

Experiments have been performed and proposed to measure the eEDM in a variety of systems, categorized below.

---

10. Obviously, in both the relativistic and non-relativistic case, the electric field, $\vec{\mathcal{E}}$, acting on the electron also depends on the electron's position in the nuclear Coulomb potential. However, it is only necessary for either $\vec{d_e}$ or $\vec{\mathcal{E}}$, not both, to be position-independent in order to factor $\langle \vec{d_e} \cdot \vec{\mathcal{E}} \rangle$ in the non-relativistic case.

1.2.3 Paramagnetic atoms

It wasn't suggested until 1950 that elementary particles might possess electric dipole moments [69]. Limits on the eEDM were initially extracted from Lamb shift ($2 \times 10^{-13}\, e$ cm) [77], electron g-factor (3×10^{-15}–$4 \times 10^{-16}\, e$ cm) [78, 79], and atomic scattering experiments (10^{-15}–$2 \times 10^{-16}\, e$ cm) [80, 81].

The eEDM interaction in a paramagnetic atom couples the unpaired electron spin to the atom's effective internal electric field, whose magnitude depends on the polarization of the atom. To maximize the relativistic enhancement, heavy atoms with highly-charged nuclei are preferred. In free space, the Hamiltonian of the atomic system is even under parity inversion so that $\vec{\mathcal{E}}_{\text{eff}} = 0$ even with the relativistic enhancement. However, an applied electric field can mix opposite-parity states such that the eigenstates of the system are not eigenstates of parity, and thus $\vec{\mathcal{E}}_{\text{eff}}$, a vector quantity that is odd under parity, may be non-zero. Since opposite-parity atomic energy levels in heavy atoms are typically separated on the order of 1 eV, the atomic polarization remains in the linear perturbative regime even for the largest achievable laboratory fields, $\vec{\mathcal{E}}_{\text{applied}}$. Thus, in atoms, $\vec{\mathcal{E}}_{\text{eff}} = S\vec{\mathcal{E}}_{\text{applied}}$, where S is known as the enhancement factor. See Table 1.1 for the enhancement factors of several atomic species. The enhancement factors, S, of paramagnetic species are of order $\sim 10 \times Z^3 \alpha^2$. The proportionality to $Z^3 \alpha^2$ can be found with elementary arguments for a hydrogen-like atom due to a combination of smaller orbital radius, larger electric field, and larger orbital velocity (i.e., relativistic effects) as Z increases (see [29, Sec. 6.2]).

An early eEDM limit was set by an experiment using optical pumping of Rb atoms in a vapor cell ($3 \times 10^{-20}\, e$ cm using S_{Rb} from [74]) [82]. Atomic beam experiments using Cs quickly surpassed this technique due to the ability to apply larger electric fields (2×10^{-21}–$2 \times 10^{-23}\, e$ cm) [74, 83–87][11], although ultimately the best limit set with Cs used a vapor cell ($2 \times 10^{-25}\, e$ cm) [88]. More modern approaches plan to use laser-cooled and trapped Cs atoms to achieve longer coherence times [89–91].

The best limit to date using paramagnetic atoms was set using an atomic beam of Tl,

---

11. Be warned that multiple papers from this group report limits *less* stringent than those reported by the same group in earlier years. After the systematic errors in atomic EDM experiments became better understood, the interpretation of reported limits becomes much more straightforward. My reading of the literature is that the results before 1969 should be regarded with a skeptical eye.

which has a favorable relativistic enhancement factor, $S = -585$ (4×10^{-24}–$2 \times 10^{-27}\,e$ cm) [92–96]. The most recent Tl experiment was limited by three kinds of systematic errors that can be suppressed in carefully-chosen molecules. First, a beam of molecules with velocity $\vec{v}$ in an electric field $\vec{\mathcal{E}}$ experiences a "motional" magnetic field $\vec{\mathcal{B}}_{\mathrm{mot}} \propto \vec{v} \times \vec{\mathcal{E}}$. This field can couple to a magnetic dipole moment and cause $\vec{\mathcal{E}}$-correlated spin precession. Second, laboratory magnetic field gradients can couple to the motional magnetic field to produce $\vec{\mathcal{E}}$-correlated geometric phases. Third, the leakage current between the electric field plates produces a magnetic field correlated with the applied voltage. In the Tl experiment, $|\mathcal{E}_{\mathrm{applied}}| \approx 120$ kV/cm, making leakage currents difficult to eliminate. These three effects cause spin precession (via the Zeeman effect) that is correlated with the applied electric field, mimicking an eEDM signal.

1.2.4 Diamagnetic atoms

Since diamagnetic atoms don't have an unpaired electron spin, they are less sensitive to the eEDM interaction than paramagnetic atoms. However, second-order effects can lead to an eEDM-induced atomic dipole moment. For example, the hyperfine interaction $H_{\mathrm{HFS}} \propto \vec{\sigma} \cdot \vec{I}$, where $\vec{\sigma}$ is an electron spin Pauli operator and $\vec{I}$ is the nuclear spin, can induce singlet-triplet mixing that generates a small component of the state with net spin $\vec{S} \equiv \sum_i \vec{\sigma}_i$ that lies along $\vec{I}$. This spin can couple to the effective electric field via the eEDM interaction in a polarized atom. In fact, the dominant contribution to an eEDM signature in diamagnetic atoms is a third-order effect due to a relativistic enhancement of the interaction between the dipole moment and the magnetic field of the nucleus [97]. Detailed calculations can be found elsewhere [98]. The resulting "enhancement factors" are no more than $S \sim 10^{-2}$ (see Table 1.1).

Despite this disadvantage, eEDM limits from experiments using diamagnetic atoms can be competitive due to the long achievable coherence times. Their small magnetic moment suppresses the impact of systematic errors like the ones that limited the Tl experiment. Further, inter-atomic collisions occur too quickly for spin-exchange interactions to occur between nuclei [99]. Another common source of decoherence, spin relaxation upon collisions with the walls of a vapor cell, are believed to occur due to interactions with local

paramagnetic sites, although the theory of wall collisions is not well-understood. The coupling of these paramagnetic sites to the nucleus is much smaller than for an unpaired valence electron.

In addition, P- and T-violating effects in the nucleus of a diamagnetic atom can induce an atomic dipole moment [100] characterized by the Schiff moment. Permanent EDM searches in diamagnetic atoms thus remain powerful probes of T-violating physics even with the enhanced sensitivity of molecular experiments to the eEDM, discussed above.

An early xenon-based eEDM search used a metastable paramagnetic state with enhancement factor $S = 130$ to find $|d_e| < 2 \times 10^{-24}\,e$ cm, comparable to the thallium result achieved in the same year [101]. Subsequent vapor-cell experiments with ^{129}Xe in the diamagnetic ground state were unable to achieve the same level of sensitivity to the eEDM $(1 \times 10^{-23}$–$5 \times 10^{-24}\,e$ cm$)$, but placed new limits on the Schiff moment [102, 103].

Later vapor-cell experiments with ^{199}Hg were able to set tight constraints on the Schiff moment and ultimately placed a limit on the eEDM comparable to the thallium result $(2\times10^{-24}$–$7\times10^{-28}\,e$ cm$)$ [42, 104–108]. Certain classes of diamagnetic atoms with octupole-deformed, or "pear-shaped," nuclei have significantly enhanced sensitivity to T-violating phenomena [109, 110]. To exploit this enhancement, initial measurements of the dipole moment of ^{225}Ra in an optical trap have already been made [111, 112], and an experiment with ^{223}Rn atoms is now under development [113].

1.2.5 Molecules

Sandars pointed out as early as 1967 that molecules could be an advantageous class of systems for EDM searches due to the high polarizability associated with nearby opposite-parity rotational levels. TlF was suggested due to its suitability for molecular beam experiments and the large relativistic enhancement from the heavy Tl nucleus [114]. However, TlF is diamagnetic and is thus best suited to constrain T-violating interactions in the nucleus rather than the eEDM. Both the groups of Edward Hinds and Norman Ramsey performed a series of measurements using this molecule, which were typically an order of magnitude less restrictive of the eEDM than concurrent experiments with paramagnetic atoms [115–118]. A new TlF experiment using modern molecular physics techniques, to probe the same

nuclear T-violating phenomena, is now underway [119, 120].

The Hinds group subsequently performed an experiment with paramagnetic YbF in the ground state, which had become feasible with advances in laser and molecular beam production technologies. They initially produced YbF via effusion of Yb and AlF_3 out of an oven source, but later developed a higher-flux source using laser-ablated molecules cooled by supersonic expansion in a carrier gas [121]. Although their flux was still small compared to atomic experiments, they were able to achieve a limit $|d_e| < 1 \times 10^{-27}\,e$ cm with longer integration time. Further, the high molecular polarizability allowed them to achieve an effective electric field greater by a factor of 220 than the Tl experiment. Perhaps most importantly, they evaded previously-limiting systematic errors by operating at a lower applied electric field than is feasible with atoms and by exploiting the large molecular tensor Stark shift, which naturally suppresses the effect of motional magnetic fields and geometric phases [122]. Major upgrades to their apparatus are currently being developed to measure the eEDM in a molecular fountain, which would allow for a greater precession time [123].

A later experiment using PbO sought to improve on the YbF result in two major respects [124, 125]. First, a vapor cell was used instead of a beam so that high densities could be obtained. Second, the measurement state has Ω-doublet structure, which leads to nearly-degenerate states of opposite parity. These states are split by the interaction between the electronic angular momentum and the molecular rotation. Such splittings are tens to hundreds of kHz, in contrast to rotational levels, which are typically split by tens of GHz. This allows molecules in Ω-doublet states to be fully polarized with electric fields on the order of 10 V/cm.

The quantum number $\mathcal{N}$ characterizes whether an Ω-doublet molecule is aligned or anti-aligned with the polarizing field. Opposite-$\mathcal{N}$ states can be resolved spectroscopically and respond identically to magnetic fields to a good approximation. Energy shifts common to the two $\mathcal{N}$ states can thus be used to infer magnetic fields, while energy shifts opposite between the two $\mathcal{N}$ states are a signature of an eEDM. For this reason, molecules with Ω-doublet structure are said to have an "internal comagnetometer."

Ultimately, the PbO experiment was unable to suppress stray electromagnetic fields and field gradients at a sufficient level to set a new limit ($|d_e| < 2 \times 10^{-26}\,e$ cm), in part

due to the complex construction of the vapor cell apparatus. However, this experiment demonstrated the power of Ω-doublets to probe and suppress systematic errors.

The ACME collaboration chose to perform an experiment with ThO based on the following criteria[126]:

1. A valence electron is in a σ orbital (analogous to an atomic s orbital, i.e., with no orbital angular momentum) so it can orbit near a heavy atomic nucleus and experience a relativistic enhancement

2. The molecule has Ω-doublet structure for systematic error rejection

3. The electronic orbital angular momentum is twice the electronic spin in magnitude and opposite in direction for systematic error suppression, as discussed below

4. Long coherence times are feasible

5. It is diatomic for spectroscopic simplicity

6. It can be efficiently produced in a beam

7. Its relevant energy transitions are accessible by affordable lasers.

These conditions narrowly constrain the list of appealing molecules.

Condition (3) can be written $\Lambda = -2\Sigma$, where Λ is the electronic orbital angular momentum and Σ is the total electronic spin (projected along the internuclear axis; see Sec. 2.1). This ensures that the state is very magnetically insensitive since the total magnetic g-factor is $g_{\text{tot}} = g_l\Lambda + g_s\Sigma \approx \Lambda + 2\Sigma = 0$. The reduced sensitivity to magnetic fields suppresses classes of systematic errors commonly found in other experiments. In practice, $g_{\text{tot}} \neq 0$ due to mixing with other states and corrections to the electronic g-factor, g_s. In the H state of ThO, $|\mu| < 10^{-2}\mu_{\text{B}}$ [127].

The upper limit for "long" coherence times in condition (4) is set by feasible length scales for a molecular beam experiment. For a relatively slow beam with speed $v \sim 200$ m/s and interaction length $L \sim 20$ cm, useful lifetimes are on the order of 1 ms and above. The lifetime of the metastable H state in ThO is approximately 2 ms[126].

The apparatus used in the ACME experiment is discussed in Section 2.3. The previous result set an eEDM limit at $|d_e| < 9 \times 10^{-29}\,e$ cm [128, 129].

Many similar considerations have motivated another eEDM search using the ground $^3\Delta_1$ state of WC [130, 131]. An eEDM experiment in the ground state of BaF has also been recently proposed to exploit some of the favorable attributes of molecules [132].

Building on rapid advances in molecular physics techniques, there has also been a proposal to perform an eEDM experiment in a polyatomic molecule (i.e., with at least three atoms) [133]. Such molecules, in contrast to atoms or diatomic molecules, have angular momentum associated with vibrational degrees of freedom, allowing for technically simple polarization and rejection of systematic errors in a manner analogous to the Ω-doublet structure of some diatomic molecules. However, these advantagous features are not intrinsically linked to the electronic state (unlike in an Ω-doublet), allowing for independent optimization of electronic properties for eEDM sensitivity and control, most notably through laser cooling. An experiment pursuing polyatomic Yb-based molecules is now in development.

Finally, some groups have proposed performing an eEDM using trapped molecular ions with $^3\Delta_1$ structure, which have many of the advantages of neutral molecules discussed above [134, 135]. Only a few ions can be trapped simultaneously before significant systematic errors are produced due to strong ion-ion interactions, but coherence times as long as 1 second are achievable. Since the eEDM sensitivity scales as the coherence time but only as the square root of molecule number, these systems could be competitive with beams of neutral molecules. The most noteworthy experiment of this type, based at JILA, uses HfF$^+$ ions, which recently reported a measurement at comparable sensitivity to the first-generation ACME result [136].

1.2.6   Solid-state systems

The effect of an eEDM can be enhanced in a solid-state system due to the possibility of a large number of unpaired electron spins. Two different approaches have been used for solid-state eEDM searches. First, an electric field can be applied across a solid-state sample. The eEDM will cause the electron spins to become aligned with the external field, and the sample will become magnetized due to the spins' magnetic moments. The

magnetization can then be detected with a superconducting quantum interference device (SQUID). The first solid-state eEDM experiment was of this type and found $|d_e| < 1 \times 10^{-22}\,e$ cm [137]. Subsequent technological improvements have allowed this technique to achieve sensitivities greater by several orders of magnitude using gadolinium gallium garnet (GGG) and $\mathrm{Eu_{0.5}Ba_{0.5}TiO_3}$ (8–6 $\times 10^{-25}\,e$ cm) [138, 139]. Avenues for improving these measurements include lower sample temperatures to reduce leakage currents, higher applied voltage on the field plates and reduced sample thickness to increase the electric field, and improved coupling efficiency to the SQUIDs. Significant advances in these areas could yield statistical sensitivities comparable to molecular experiments, although it is difficult to predict the role of systematic errors in this regime.

The second approach to solid-state eEDM measurements uses the reverse process: a magnetic field is applied to align unpaired electron spins. If the electrons have an EDM, then they will generate a voltage difference across the sample. This technique was demonstrated in 2005 using gadolinium-iron garnet (GdIG) and achieved $|d_e| < 5 \times 10^{-24}\,e$ cm [140].

A rather different idea has recently been proposed to leverage the enormous numbers (and correspondingly favorable shot noise limits) in solid-state-based systems: it may be possible to perform an eEDM measurement by embedding polar molecules (e.g., YbF) in a solid matrix of chemically inert species (Ne, Ar, etc.) [141, 142].

1.2.7 Summary of progress in EDM searches

The history of experimental EDM searches is less a march than a six-decade-long charge of progress. The most important results to date from paramagnetic atoms and molecules are shown in Fig 1.3. The y-axis gives d_e sensitivity in terms of the energy scale probed in BSM models, according to a representative scaling found in [21]. Fig. 1.3 is known as a Livingston plot, after the accelerator physicist Milton Stanley Livingston, who observed that the energies probed by particle accelerators grows exponentially in time. Unfortunately, that trend is not projected to hold into the future.[12] For comparison, I have shown the

---

12. As [143] notes, "yet the figure of merit of energy reach in the constituent center-of-momentum frame is far from the entire story of enabling capabilities." Considerations such as collider luminosity and constituent species play an important role in the physics probed. In a similar spirit, it should be clear that EDM searches do not necessarily probe the same parameter space as colliders.

Energy reach of eEDM measurements and colliders
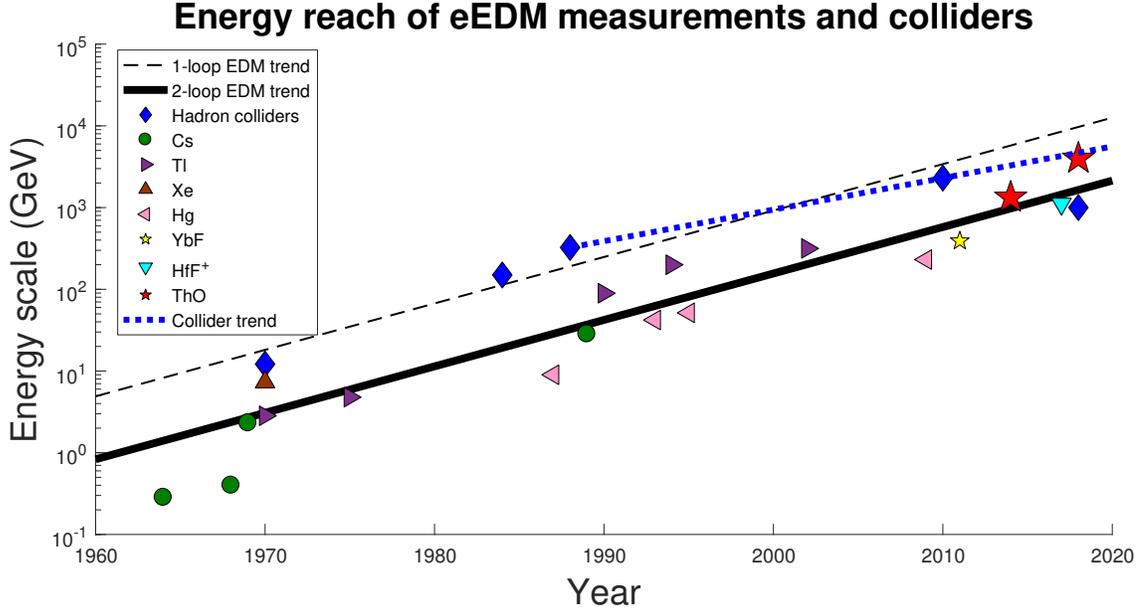
Figure 1.3: Livingston plot showing energy scale probed by hadron colliders (blue diamonds) and atomic and molecular eEDM measurements. Collider data from [143] represent the "reach" of experiments, generally smaller than the center-of-mass collision energy. Trendline shows fit to eEDM experiments since 1969, with an order of magnitude improvement in energy reach every ~ 18 years. Up to now, collider energies have advanced at a similar exponential rate, but future planned colliders fall well below the historical trend line. All eEDM results are scaled to give corresponding sensitivities to the stop mass as computed from two-loop effects in [50] (see Fig. 9 therein). The dashed line represents the sensitivity to the (even higher) energy scale of a selectron contributing to one-loop effects with typical parameter values as shown in Fig. 8 of the review [21]. While alternate theoretical models may shift the data points (and corresponding trendlines) up or down, the conclusion of exponential growth in energy reach with time is model-independent. For reference, trend for recent collider results is shown (dotted blue line). The datum for hadron colliders from 2018 represents an up-to-date bound from the LHC on the stop mass, for direct comparison with two-loop eEDM constraints on the stop mass. Note that while EDM measurements have been sensitive to one-loop effects occuring at energy scales comparable to or exceeding colliders for decades, they are now probing even two-loop effects at energy scales exceeding those accessible to the LHC.

energy scale of historical colliders together with experimental eEDM results. Remarkably, the exponential growth of energy sensitivity found in EDM experiments has been maintained for at least five decades and shows no sign of slowing.

Chapter 2

Experimental overview

The methodology and apparatus of the ACME experiment are described in detail in prior theses [144–151] and publications [67, 126–129, 152–155]. The most comprehensive overview of the first ACME measurement, ACME I, may be found in [129]. The recently completed measurement, ACME II, is similar in most essential respects; see the recent result paper, [156]. In this chapter, I will describe the basic measurement approach and experimental apparatus, with particular attention to the "rotational cooling" used to increase the population of ThO molecules in a single quantum state prior to the EDM measurement

2.1 Molecular structure

It is useful to review molecular structure before describing in detail how to perform an EDM measurement using $^{232}\text{Th}^{16}\text{O}$ molecules. Comprehensive references on diatomic molecules include [157, 158]. Previous ACME theses contain more concise summaries of the aspects of molecular structure that are relevant to the ACME experiment; for a reasonably comprehensive overview, I recommend Nick Hutzler's dissertation [146]. The goal of this section is to give a conceptual grounding in the most essential features of diatomic molecules rather than to make any formal derivations.

2.1.1 Vibrational motion

In order for a diatomic molecule to exist, there must of course be an attractive potential between the two atomic nuclei, which dominates when the atoms are far apart, and a compensating repulsive barrier so the nuclei do not collide at short range. The equilibrium distance between nuclei occurs at the minimum of the overall potential energy curve (as a function of internuclear distance), and the potential is approximately harmonic in the vicinity of this equilibrium position ($r_e = 1.8$ Å in ThO [159]). As a result, the motion of the nuclei with respect to each other is characterized by a ladder of quantum energy levels, $E_v = \omega_v(v + \frac{1}{2})$, where E_v is the "vibrational energy," v denotes the vibrational quantum number, and $\omega_v = 896 \, \text{cm}^{-1}$ is the energy constant of a vibrational excitation in ThO [159]. Especially at larger vibrational amplitudes, $v > 1$, the potential may have significant anharmonic contributions. However, the essential picture of a ladder of quantized energy levels, describing the oscillation of nuclei like masses connected by a spring, remains.

2.1.2 Rotational motion

For simplicity, now consider a diatomic molecule with its nuclei "frozen" at their equilibrium separation. The geometry of such a molecule is that of a dumbell, and it has an associated moment of inertia I for rotations about an axis passing through its center of mass and perpendicular to the internuclear axis. Classically, the energy of such a rotation with angular momentum N is $H_N = \frac{1}{2I}N^2$. For a molecule, the angular momentum $N = n\hbar$ is an integer multiple of the reduced planck constant and the energy is the eigenvalue of the operator H_N. For a molecular state with a well-defined value of N, we write $H_N|N\rangle = \frac{1}{2I}N^2|N\rangle$ and thus interpret the rotational energy as $E_N = \frac{1}{2I}N(N+1)$. We conventionally write $\frac{1}{2I}$ as the "rotational constant" B, approximately 0.333 cm^{-1} in the ground electronic state of ThO [160].

In general, the rotational constant is actually a weak function of the vibrational state v because the vibration modifies the moment of inertia I. Likewise, a high degree of molecular rotation induces a large centrifugal force on the nuclei and stretches out the equilibrium separation, thus modifying the moment of inertia. Fortunately, these higher-order effects

will not be critical to the ACME measurement.

Note that the rotational energy level spacing grows quickly with rotational quanta, but is characteristically much smaller than the vibrational spacing.

2.1.3 Electronic structure

So far we have entirely ignored the properties of the electrons in the molecule (except implicitly insofar as they are essential to the bonding of the molecule). Depending on the relative strengths among various energy scales (spin-orbit coupling, rotation, and coupling of the electron orbit to the internuclear axis by the internal molecular electric field), different quantum bases may be more or less convenient to describe an electronic configuration. The limiting situations are known as Hund's cases.

In Hund's case (a), which will usually be most convenient for discussing properties of ThO in the ACME experiment, the electronic orbital angular momentum $\vec{L}$ is strongly coupled to the internuclear axis by electrostatic binding. Therefore, the projection of $\vec{L}$ along the internuclear axis is well-defined: $\vec{L} \cdot \hat{n}|\psi\rangle = \Lambda|\psi\rangle$, where $|\psi\rangle$ is the molecular state and $\hat{n}$ is the internuclear axis defined by our convention to point from the negative to positive nucleus in a polar diatomic molecule (and thus from O to Th). Colloquially, we say the orbital angular momentum is pinned to the internuclear axis.

A further defining feature of Hund's case (a) molecules is that the spin-orbit coupling is strong enough to "pin" the electron spin axis along the electron orbital angular momentum, and thus also along the internuclear axis: $\vec{S} \cdot \hat{n}|\psi\rangle = \Sigma|\psi\rangle$, where $\vec{S}$ is the electron spin operator.

Furthermore, in Hund's case (a), the rotational energy level spacing is small compared to the spin-orbit coupling. As a result, we interpret states with different spin-orbit components (e.g., opposite relative orientations of $\vec{L}$ and $\vec{S}$) as completely distinct electronic manifolds, each of which have their own ladders of vibrational and rotational energy levels. Although the opposite case will not be relevant to the present work, it is important to be aware that many molecules are well-described by Hund's case (b)), in which the spin-orbit coupling is small compared to the rotational spacing and states with different spin-orbit character are regarded as comprising additional sub-structure of a single electronic state rather than

distinct "electronic manifolds."

Because Λ and Σ are independently good quantum numbers in Hund's case (a), it follows that their sum, $\Omega \equiv \Lambda + \Sigma$, is also a good quantum number. Furthermore, the total electron spin S is well-defined. The total angular momentum J is always a good quantum number due to rotational invariance (in the lab frame) of the molecule's Hamiltonian. We typically write the state of a molecule well-described by Hund's case (a) using a molecular term symbol, $^{2S+1}|\Lambda|_{|\Omega|}(v, J, m)$, where the $2S + 1$ superscript denotes the total electron spin degeneracy, $|\Lambda|$ is the magnitude of orbital angular momentum projected along the internuclear axis, and $|\Omega|$ is the magnitude of total spin angular momentum projected along the internuclear axis. The quantities (v, J) may or may not be written explicitly, but I have included them as a reminder that the molecular vibration and rotation also need to be considered. The value of $|\Lambda|$ is conventionally indicated by capital Greek letters $\Sigma, \Pi, \Delta, \cdots$ for $|\Lambda| = 0, 1, 2, \cdots$ (which is particularly unfortunate since Σ is also the projection of the electron spin along the internuclear axis).

For example, the ground state of ThO is $X^1\Sigma_0$, denoting $S = 0$, $\Lambda = 0$, and $\Omega = 0$. The "science state" in ThO, used in the eEDM measurement, is $H^3\Delta_1$, indicating $S = 1$, $|\Lambda| = 2$, and $|\Omega| = |\Lambda + \Sigma| = 1$. It follows that $\Sigma = \pm 1 = \mp\Omega$ with $\Lambda = \mp 2\Sigma = \pm 2\Omega$. It is critical that there are two possible configurations of the H state, with the total electronic angular momentum either along or against the internuclear axis. These states, with $\Omega = \pm 1$, are very nearly degenerate. This (near) double-degeneracy is known as Ω-doubling and will be described in Sec. 2.1.4.

In most states of ThO, the spin-orbit coupling is so strong that only the total electronic orbital angular momentum $\vec{J_e} \equiv \vec{L} + \vec{S}$ can be considered. As a result, Λ and Σ are no longer good quantum numbers, but $\Omega = \vec{J_e} \cdot \hat{n}$ remains a good quantum number. This is described by Hund's case (c), where the only good quantum numbers are Ω and J. It is typically convenient to decompose a Hund's case (c) state as a sum of states with Hund's case (a) character. For example, calculations show that we can decompose the C state as $C \approx 0.77\,^1\Pi_1 + 0.20\,^3\Pi_1 + 0.02\,^3\Delta_1$ [159]. Of course, even a relatively good Hund's case (a) state, like X or H, will typically have some small admixture of additional molecular terms.

2.1.4   Ω-doubling

Here we will explore important consequences of the double-degeneracy of $H^3\Delta_1$ state. To do this, we will return to the rotational Hamiltonian and rewrite it in terms of other quantum numbers: $H_{rot} = BR^2 = B(\vec{J} - \vec{J_e})^2 = BJ^2 + BJ_e^2 - 2B\vec{J}\cdot\vec{J_e}$, where $\vec{R}$ is the rigid-body rotational angular momentum of the molecule. Within a given electronic manifold, BJ_e^2 is a constant, which we will therefore not consider further. The first term, BJ^2, leads to a ladder of states with energies $BJ(J+1)$ as we'd expect for a purely rotational energy contribution. We therefore conventionally refer to states with distinct J as "rotational" levels, even though $\vec{J}$ is actually defined to include all angular momenta of the molecule (not just rotation). The last term, $-2B\vec{J}\cdot\vec{J_e}$, generates Ω-doubling. Although a careful calculation is somewhat involved (e.g., see [145, Sec. 3.0.2]), we will motivate its essential consequences.

Working in the molecular frame with $\hat{z} = \hat{n}$, this term contributes the "Coriolis" interaction $H_{\mathrm{Cor}} = -2B\vec{J}\cdot\vec{J_e} = -2B(\vec{J}\cdot\hat{n}\,\vec{J_e}\cdot\hat{n} - J^-J_e^+ - J^+J_e^-)$, where $J^{\pm}$ are raising and lowering operators. Since the rotational angular momentum $\vec{R}$ has no component along the internuclear axis, the first term in parentheses reduces to Ω^2, a constant of the electronic manifold. The remaining terms raise and lower the component of the electronic angular momentum along the internuclear axis, i.e. Ω. Since electronic states have well-defined Ω, these terms cannot create a first-order energy shift, $\langle J_e^{\pm}\rangle = 0$. However, states of opposite Ω, i.e. $\Omega = \pm 1$ in the H state of ThO, can be coupled to each other at second order in this Coriolis interaction via an intermediate state with $\Omega = 0$. The expected energy shift at second order in perturbation theory is therefore of the order $\sim \frac{B^2}{A}$, where $A \sim 400\,\mathrm{cm}^{-1}$ is the spin-orbit constant in ThO that characterizes the separation between an $|\Omega| = 1$ and $\Omega = 0$ state (such as the $A^3\Pi_0$ and $B^3\Pi_1$ states). This second-order interaction causes a splitting of the nominal double-degeneracy in $H^3\Delta_1(J = 1)$ by 360 kHz [146, Sec. D.3].

We now turn to the question of the parity of these states. Writing a state in a Hund's

case (a) basis, the typical parity transformation P behaves as [158, Sec. 6.9.3]

$$
\begin{aligned}
P|S, \Sigma\rangle &= (-1)^{S-\Sigma}|S, -\Sigma\rangle \\
P|J, \Omega, M\rangle &= (-1)^{J-\Omega}|J, -\Omega, M\rangle \\
P|\Lambda\rangle &= (-1)^{\Lambda}|-\Lambda\rangle,
\end{aligned}
\tag{2.1}
$$

where I have separated the wavefunction into parts depending on the electron spin and its projection, total angular momentum and its projections, and the projection of electron orbital angular momentum. The quantum number M is the projection of total angular momentum $\vec{J}$ along the laboratory z-axis. Using these transformation rules, one can see that parity eigenstates are

$$
|J, M, \pm\rangle = \frac{1}{\sqrt{2}}(|J, \Omega, M\rangle \pm (-1)^{J-S}|J, -\Omega, M\rangle),
\tag{2.2}
$$

where I have dropped kets involving implicitly constrained quantum numbers (e.g., $\Sigma = -\Omega$ in H). Therefore, in the absence of an external electric field to break the parity symmetry, the eigenstates of H are even and odd superpositions of states with opposite Ω, and these nearly-degenerate states form a "parity doublet."

2.1.5 Molecule orientation

In the absence of external fields, the Hamiltonian of a molecule is rotationally symmetric and therefore $\langle \vec{\mathcal{E}}_{\text{eff}} \rangle = 0$. Therefore, we must orient ThO molecules with respect to our laboratory frame in order to probe the EDM interaction, $H = -2d_e \vec{s} \cdot \vec{\mathcal{E}}_{\text{eff}}$. An oriented molecule is *not* a parity eigenstate: let us denote the state of a molecule oriented along the lab z-axis by $|\Uparrow\rangle$ and one oriented against the lab z-axis by $|\Downarrow\rangle$. Then the parity eigenstates are superpositions of oriented molecules, $|P = \pm\rangle = |\Uparrow\rangle \pm |\Downarrow\rangle$. Conversely, the oriented states $|\Uparrow\rangle$ and $|\Downarrow\rangle$ are superpositions of parity eigenstates $|P = \pm\rangle$.

As a result, opposite-parity states must be mixed, typically by an external electric field, in order to obtain molecules that are oriented in the lab frame. The eigenstates of the molecule approach $\{|\Uparrow\rangle, |\Downarrow\rangle\}$ when the perturbing (parity-mixing) energy greatly exceeds the energy splitting Δ_P between opposite-parity states, $|D\mathcal{E}_{\text{lab}}| \gg \Delta_P$, where D is the
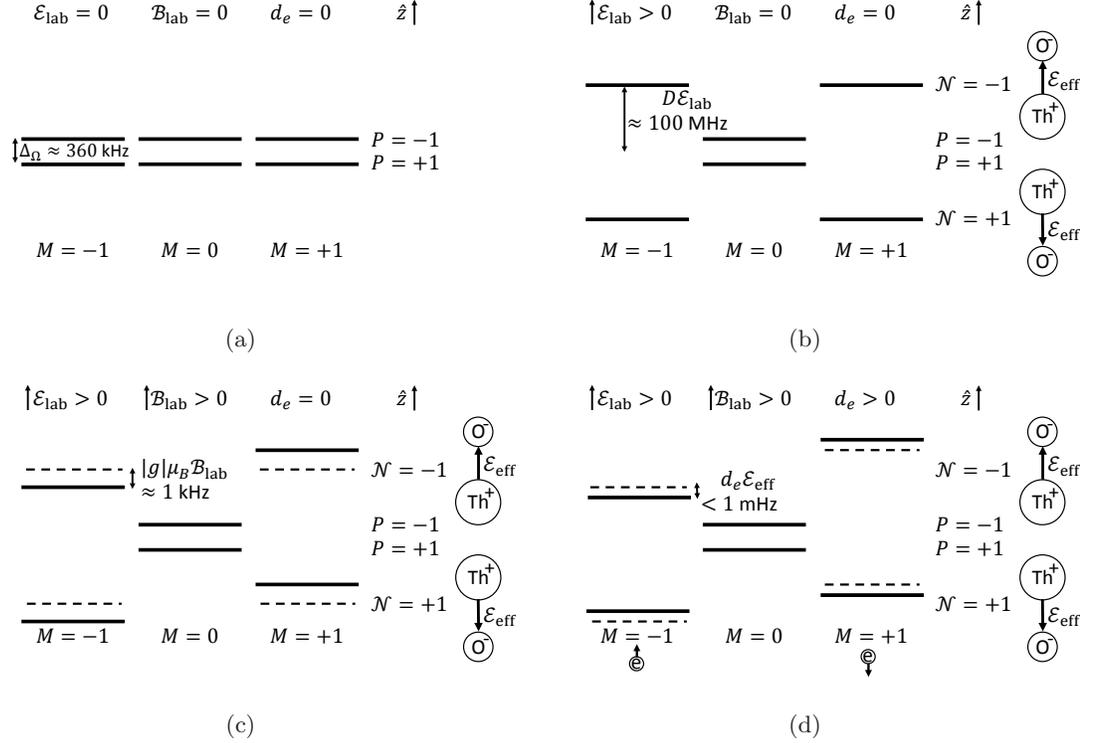
Figure 2.1: The level structure of $H^3\Delta_1(J=1)$ in externally applied electric and magnetic fields along the laboratory z-axis, $\mathcal{E}_{\text{lab}}$ and $\mathcal{B}_{\text{lab}}$, as well as the EDM shift $H_d = -d_e\mathcal{E}_{\text{eff}} M \times \tilde{\mathcal{N}}\tilde{\mathcal{E}}$. (a) Energy levels in the absence of external fields. Opposite-parity states are split by $\Delta_\Omega \approx 360$ kHz. (b) Laboratory electric field $\mathcal{E}_{\text{lab}} \sim 100$ V/cm along the z-axis Stark shifts $M = \pm 1$ by ~ 100 MHz. The aligned and anti-aligned molecule states correspond to $\tilde{\mathcal{N}} = \pm 1$. The $M = 0$ states are unperturbed to a good approximation. (c) Laboratory magnetic field $\mathcal{B}_{\text{lab}} \sim 10$ mG along the z-axis Zeeman shifts $M = \pm 1$ by ~ 1 kHz. Note that $g_H < 0$, so $M = 1(-1)$ is shifted up (down). (d) Electric dipole moment interaction shifts states in opposite directions when either M or $\tilde{\mathcal{N}}$ is reversed, up to $\sim 100\,\mu$Hz. Note that the electron spin is *down* for $M = +1$ and *up* for $M = -1$ since M is the projection of total, rather than electron, angular momentum on the laboratory z-axis (since the total angular momentum is predominantly set by the electronic orbital angular momentum, which points against the electron spin in a $^3\Delta_1$ state; see Sec. 2.1.6).

molecular electric dipole moment in the molecule-fixed frame and $\mathcal{E}_{\text{lab}}$ is the applied electric field in the laboratory frame. In ThO, the relevant pair of opposite-parity states are those that constitute the Ω-doublet in the H electronic state: $\Delta_P = \Delta_\Omega = 2\pi \times 360$ kHz.

The most significant consequence of the closely-spaced opposite-parity states in the H electronic state of ThO is that the molecule can be "fully polarized" in modest laboratory fields (provided $M \neq 0$). The "molecule-fixed" dipole moment has magnitude $|D_H| = 1.67(4)\,ea_0$, where e is the electron charge and a_0 is the Bohr radius. The dipole moment in a given rotational level is $D_J = \frac{D_H}{J(J+1)}$, which for $J = 1$ is $D_1 = h \times 2.13(2)$ MHz/(V/cm) in laboratory units [127]. Therefore, with only $\mathcal{E}_{\text{lab}} = 10$ V/cm, the ratio $|D_1 \mathcal{E}_{\text{lab}}|/\Delta_\Omega \approx 60 \gg 1$. This represents a significant technical advantage over atoms or molecules without Ω-doublet structure, where full polarization is not achieved even with tens of kV/cm.

The H state level structure is shown in Fig. 2.1. We now introduce the quantum number $\tilde{\mathcal{N}}$, which has value $+1$ when the molecule is aligned with the electric field (i.e., its energy is reduced due to $\mathcal{E}_{\text{lab}}$) and -1 when the molecule is anti-aligned (its energy is increased)[1]. Since the molecule's dipole moment, like the internuclear axis, points from the negatively to the positively charged nucleus by definition, the Stark shift has the sign of $\text{sign}(H_{\text{Stark}}) = -\hat{D} \cdot \hat{\mathcal{E}} = -\hat{n} \cdot \hat{\mathcal{E}} \equiv -\tilde{\mathcal{N}}$ (i.e., when the Stark shift is negative, $\tilde{\mathcal{N}}$ is positive). Introducing the notation $\tilde{\mathcal{E}} \equiv \hat{\mathcal{E}}_{\text{lab}} \cdot \hat{z} = \pm 1$ to describe the state of the experiment at a given time, we can therefore see that $\hat{n} = +\tilde{\mathcal{N}}\tilde{\mathcal{E}}\,\hat{z}$ when an H-state molecule is fully polarized[2]. These conclusions can also be obtained from formal derivations of the Stark shift (e.g., [145, 146]). For reference, I reproduce the Stark shift formula here [146, Eq. 2.26]:

$$\langle J, M, \Omega; \Lambda, S, \Sigma | H_{\text{Stark}} | J, M, \Omega; \Lambda, S, \Sigma \rangle = -\mathcal{E}_{\text{lab}} D \frac{M\Omega}{J(J+1)}. \tag{2.3}$$

Notice from this expression that there is no first-order energy shift to $M = 0$ states due to the Stark interaction.

---

1. Throughout this thesis, a tilde denotes a quantity that can take values ± 1, though sometimes the tilde will be omitted when convenient.

2. Be wary of these, and other, sign conventions. For instance, Ch. 2 of Nick Hutzler's thesis (generally the best reference for detailed treatment of these issues) records the opposite sign from that written here, but Fig. 5.1 therein is consistent with my expression [146].

As an alternative to the approach used above, it is sometimes useful to express $\hat{n}$ in terms of Ω and M rather than $\tilde{\mathcal{N}}$. This can be done by formally deriving eigenstates in the presence of an electric field (see [146, Ch. 2]), but it is also useful to obtain this relationship by thinking through the geometry. Consider the two cases $\hat{n} = \pm\hat{z}$ separately, and take $J = 1$, $M = \pm 1$ for simplicity. (Since there is no Stark shift when $M = 0$, the molecule cannot be oriented in that case.) For $\hat{n} = +\hat{z}$, the total angular momentum projection M along the lab z-axis is exactly the same quantity as Ω, the angular momentum projection along the internuclear axis. Therefore, $\Omega = M$ in this case. When $\hat{n} = -\hat{z}$, the same reasoning shows $\Omega = -M$. Therefore, in either case $\Omega = M\,\hat{n}\cdot\hat{z} = M\tilde{\mathcal{N}}\tilde{\mathcal{E}}$, and $\hat{n} = M\Omega\hat{z}$.

2.1.6 EDM interaction in the H state

Using the foregoing knowledge of molecular structure, we can now show how the EDM interaction manifests in the $H^3\Delta_1$ state of ThO. First, notice that there are two valence electrons in this state, $S = 1$. One of these electrons is in a σ orbital and experiences the relativistic enhancement of the EDM interaction near the Th nucleus. The other electron is in a δ orbital (analogous to an atomic d-orbital) and contributes both to producing the Ω doubling and to causing the near-cancellation of the g-factor[3]. When we colloquially speak of "the electron spin" in the EDM interaction, it is the spin of the electron in the σ orbital that matters. However, since both spins are aligned with each other we need not make this distinction carefully in practice. Just as for a free electron spin, we define $H_d = -\vec{d}_e \cdot \vec{\mathcal{E}}_{\text{eff}}$ such that the average energy shift is $H_d = -d_e\mathcal{E}_{\text{eff}}$ for a spin aligned with the effective electric field in the H state of ThO. Note, however, that this Hamiltonian acts on the molecular state (which includes both valence electrons) rather than a free electron.

To see the effect of the EDM interaction in terms of molecular and experimental quan-

---

3. As discussed in Sec. 1.2.5, this cancellation occurs because $g \approx g_S\Sigma + g_L\Lambda \approx 2 \times 1 - 1 \times 2 = 0$ for a $^3\Delta_1$ state, where $g_{S(L)}$ is the g-factor of the electron spin (electron orbit).

tities $\tilde{\mathcal{N}}$, $\tilde{\mathcal{E}}$, J, and M, recall that $\vec{\mathcal{E}}_{\text{eff}} = -\mathcal{E}_{\text{eff}}\hat{n}$, so

$$
\begin{aligned}
H_d &= \quad -\vec{d_e} \cdot \vec{\mathcal{E}}_{\text{eff}} \\
&= \quad +d_e\mathcal{E}_{\text{eff}}\hat{S} \cdot \hat{n} \\
&= \quad +d_e\mathcal{E}_{\text{eff}}\Sigma \\
&= \quad -d_e\mathcal{E}_{\text{eff}}\Omega.
\end{aligned}
\tag{2.4}
$$

We saw previously that $\Omega = M\tilde{\mathcal{N}}\tilde{\mathcal{E}}$, so

$$
H_d = -d_e\mathcal{E}_{\text{eff}}M \times \tilde{\mathcal{N}}\tilde{\mathcal{E}}
\tag{2.5}
$$

in the $H(J = 1)$ state.

There are two critical features of this Hamiltonian. (1) even though M is the projection of the total molecular angular momentum J along the lab z-axis, H_d has almost the same form as for a free electron where M would be the projection of the electron spin along the laboratory z-axis. (2) By preparing the molecules in a state with a positive Stark shift ($\tilde{\mathcal{N}} = -1$) vs. negative Stark shift ($\tilde{\mathcal{N}} = +1$), **or** by reversing the direction of the laboratory electric field, $\tilde{\mathcal{E}} \to -\tilde{\mathcal{E}}$, the sign of the interaction term can be reversed.

2.1.7 Coupling the molecules to laser light

All of our state preparation and readout depends on state-specific couplings to laser light. We select electronic, vibrational, rotational, and Ω-doublet states by the frequency of laser light, while we select the composition of M states using angular momentum selection rules. We consider here the transition matrix elements involving the H state that are relevant to the state preparation and readout steps of the ACME experiment. Although these are standard molecular matrix element calculations, because they differ somewhat from atomic calculations, I will walk through the logic in detail.

Of particular interest are states in the $H(J = 1)$ manifold of the form

$$
|\psi(\phi), \tilde{\mathcal{N}}\rangle = \frac{1}{\sqrt{2}}(e^{-i\phi}|M = +1, \tilde{\mathcal{N}}\rangle + e^{+i\phi}|M = -1, \tilde{\mathcal{N}}\rangle).
\tag{2.6}
$$

This state has the molecular dipole moment either aligned ($\tilde{\mathcal{N}} = +1$) or anti-aligned ($\tilde{\mathcal{N}} = -1$) with the laboratory electric field, and has angular momentum aligned in the xy-plane at an angle ϕ with respect to $\hat{x}$. Under normal conditions (with $\mathcal{E}_{\text{lab}} \sim 100$ V/cm), opposite-$\tilde{\mathcal{N}}$ states are separated by ~ 100 MHz and hence fully resolved by optical transitions, which have typical linewidths of a few MHz. In the ACME experimental protocol, we only couple $|\psi(\phi), \tilde{\mathcal{N}}\rangle$ to states of the form $|J = 1, M = 0, \tilde{\mathcal{P}}\rangle \equiv |\tilde{\mathcal{P}}\rangle$, where $\tilde{\mathcal{P}} = \pm 1$ is the parity of the coupled state. (In particular, we address such states in both the I and C electronic states' rotational manifolds; see Fig. 2.2.) The E1 transition amplitude is proportional to $\langle \tilde{\mathcal{P}} | \vec{\epsilon} \cdot \vec{r} | \psi(\phi), \tilde{\mathcal{N}} \rangle$. We can compute this amplitude by decomposing the Hund's case (c) excited state into its rotational state kets, $|J, M, \tilde{\mathcal{P}} = \pm\rangle = \frac{1}{\sqrt{2}}(|J, \Omega, M\rangle \pm (-1)^{J-\Omega}|J, -\Omega, M\rangle)$, where $J = 1$, $\Omega = \pm 1$, and $M = 0$ for the states of interest: up to an arbitrary overall phase,

$$|\tilde{\mathcal{P}}\rangle = \frac{1}{\sqrt{2}}(|\Omega = +1\rangle + \tilde{\mathcal{P}}|\Omega = -1\rangle)|J = 1, M = 0\rangle. \tag{2.7}$$

We now decompose $|\psi(\phi), \tilde{\mathcal{N}}\rangle$ in terms of the Ω quantum number and consider the matrix element

$$
\begin{aligned}
\langle \tilde{\mathcal{P}} | \vec{\epsilon} \cdot \vec{r} | \psi(\phi), \tilde{\mathcal{N}} \rangle \quad = \quad & \tfrac{1}{2}(\langle \Omega = +1| + \tilde{\mathcal{P}}\langle \Omega = -1|)\langle J = 1, M = 0|)\vec{\epsilon} \cdot \vec{r} \\
& \times (e^{-i\phi}|M = +1, \Omega = +\tilde{\mathcal{N}}\tilde{\mathcal{E}}\rangle + e^{+i\phi}|M = -1, \Omega = -\tilde{\mathcal{N}}\tilde{\mathcal{E}}\rangle)|J = 1\rangle.
\end{aligned}
\tag{2.8}
$$

This matrix element involves one lab-frame vector, $\vec{\epsilon}$, and one molecule-frame vector, $\vec{r}$. This can be handled using the approach outlined in Brown and Carrington [158], but we instead refer to the convenient formulas compiled in Nick Hutzler's thesis [146, Ch. 2]. In particular,

$$
\begin{aligned}
\langle J M \Omega; \Lambda S \Sigma | \vec{\epsilon} \cdot \vec{r} | J' M' \Omega'; \Lambda' S' \Sigma' \rangle = {} & \delta_{SS'}\delta_{\Sigma\Sigma'}(-1)^{M'-\Omega}\langle \Lambda | T^1_{\Omega-\Omega'}(\vec{r}) | \Lambda' \rangle T^1_{M'-M}(\vec{\epsilon}) \\
\times \sqrt{(2J+1)(2J'+1)} & \begin{pmatrix} J & 1 & J' \\ -\Omega & (\Omega - \Omega') & \Omega' \end{pmatrix} \begin{pmatrix} J & 1 & J' \\ -M & (M - M') & M' \end{pmatrix},
\end{aligned}
\tag{2.9}
$$

35

where δ_{ij} is a Kronecker delta, $T^1_{M'-M}(\vec{\epsilon})$ is a rank-1 vector operator depending on $\vec{\epsilon}$, and analogously for $T^1_{\Omega-\Omega'}(\vec{r})$. The factors in parentheses are three-J symbols. This matrix element entails a selection rule $\Delta\Omega = 0, \pm1$. In our case of interest, the ground state H has $\Omega' = \pm1$ and the excited state (C or I) has $\Omega = \pm1$, and the cross-terms between $\Omega' = \pm1 \to \Omega = \mp1$ must vanish. One then finds that

$$\langle\tilde{\mathcal{P}}|\vec{\epsilon}\cdot\vec{r}|\psi(\phi),\tilde{\mathcal{N}}\rangle \propto e^{-i\phi}T^1_+(\vec{\epsilon}) + \tilde{\mathcal{P}}e^{-i\phi}T^1_-(\vec{\epsilon}) \tag{2.10}$$

for any $\vec{\epsilon}$.[4] For linearly polarized light in the xy-plane with polarization angle θ with respect to the x-axis, $\vec{\epsilon} = \cos\theta\hat{x} + \sin\theta\hat{y}$, $T^1_\pm(\vec{\epsilon}) \propto e^{\pm i\theta}$ and thus the transition probability is proportional to $\cos^2(\phi - \theta)$ when $\tilde{\mathcal{P}} = +1$ and $\sin^2(\phi - \theta)$ when $\tilde{\mathcal{P}} = -1$. We therefore write in the general case that the transition probability is $\cos^2(\phi - \theta + \frac{\pi}{4}(\tilde{\mathcal{P}} - 1))$.

Consider the case of $\tilde{\mathcal{P}} = +1$ first. We refer to $|\psi(\phi = \theta), \tilde{\mathcal{N}}\rangle$ as the "bright" state, where the transition amplitude is maximized, and $|\psi(\phi = \theta + \frac{\pi}{2}), \tilde{\mathcal{N}}\rangle$ as the "dark" state, where the transition amplitude vanishes. When driving a transition to $\tilde{\mathcal{P}} = -1$, the role of the dark and bright states reverses. Thus, changing $\tilde{\mathcal{P}}$ has an equivalent effect on the transition amplitudes as rotating the laser light polarization by $\frac{\pi}{2}$.

2.2 Spin precession measurement

In Sec. 1.2.1, we described an archetypal EDM measurement in terms of Ramsey's method of separated oscillatory fields. The approach used in ACME actually deviates from this method in key ways, so it will be treated explicitly here.

We want to probe the interaction Hamiltonian in the $H(J = 1)$ state of ThO, $H_d = -d_e\mathcal{E}_{\text{eff}}M \times \tilde{\mathcal{N}}\tilde{\mathcal{E}}$, where[5] $\vec{\mathcal{E}}_{\text{eff}} = -(78\,\text{GV/cm})\hat{n}$ in the relevant state of $^{232}\text{Th}^{16}\text{O}$. For the time being, we'll take $\tilde{\mathcal{N}}\tilde{\mathcal{E}} = +1$ so that H_d takes a simpler form, $H_d \to -d_e\mathcal{E}_{\text{eff}}M$.

---

4. See Appendix A for the definition of, and handy identities for, the spherical basis and computations with $T^1_\pm$.

5. $\vec{\mathcal{E}}_{\text{eff}}$ is quoted here as the average of the latest calculations [161, 162]. Note that this is slightly smaller than the value quoted in the first result paper [128], which used the result of [163]. The current theoretical uncertainty in $\mathcal{E}_{\text{eff}}$ is estimated to be a few percent, which will only significantly inhibit the precision of an EDM measurement once the mean value of d_e can be probed to two significant digits.

2.2.1 A brief analogy to $J = \frac{1}{2}$

We can understand the procedure used to measure the energy shift H_d in the ACME experiment by briefly considering the analogous Zeeman shift, $H_\mu = -\mu \vec{s} \cdot \vec{\mathcal{B}}$[6], for a spin-1/2 system. It is a fact unique to spin-1/2 objects that for any unit vector $\hat{u} = \hat{u}(\theta, \phi)$ parametrized by polar angle θ and azimuthal angle ϕ, there is a state $|\psi(\theta, \phi)\rangle$ that is an eigenstate of the spin-component operator $S_{\hat{u}}$. (For larger spins, an arbitrary state $|\psi\rangle$ is described by more parameters than are needed to specify a unique spatial direction, and therefore an arbitrary state $|\psi\rangle$ is not necessarily an eigenstate of any operator $S_{\hat{u}}$.) In particular, up to an overall phase,

$$|\psi(\theta, \phi)\rangle = \cos\frac{\theta}{2} e^{-i\phi/2}|+\rangle + \sin\frac{\theta}{2} e^{i\phi/2}|-\rangle, \tag{2.11}$$

where I am using the shorthand $|\pm\rangle \equiv |M = \pm\frac{1}{2}\rangle$. Consider a state with $\theta = \frac{\pi}{2}$, corresponding to a spin oriented in the xy-plane, or on the equator of the Bloch sphere:

$$|\psi(\phi)\rangle = \frac{1}{\sqrt{2}}(e^{-i\phi/2}|+\rangle + e^{i\phi/2}|-\rangle). \tag{2.12}$$

Under the influence of H_μ for a duration τ, the spin orientation will precess around the polar axis: $\phi \rightarrow \phi + \mu\mathcal{B}\tau$. With this concept in mind, let's examine the case of $J = 1$.

2.2.2 Phase measurement in $J = 1$

As we saw before, we want to observe the effect of the Hamiltonian $H_d = -d_e\mathcal{E}_{\text{eff}}M$ acting on the $H(J = 1)$ state of ThO. By analogy to the $J = \frac{1}{2}$ case, define $|\pm\rangle \equiv |M = \pm 1\rangle$ and consider the state

$$|\psi(\phi)\rangle = \frac{1}{\sqrt{2}}(e^{-i\phi}|+\rangle + e^{i\phi}|-\rangle), \tag{2.13}$$

so that evolution of the state under H_d produces an accumulation of the phase $\phi \rightarrow \phi + d_e\mathcal{E}_{\text{eff}}\tau$ over a duration τ. With the $J = \frac{1}{2}$ case in mind from Sec. 2.2.1, it is tempting to interpret this behavior in terms of the $\vec{J}$ vector precessing around the z-axis. However, that

---

6. For an electron, $\mu = -g_S\mu_B$ so that, e.g., $\langle M|H_\mu|M\rangle = +g_S\mu_B M\mathcal{B}$ for a z-aligned $\mathcal{B}$ field.

is not what is going on.

As we noted previously, for a $J > \frac{1}{2}$ system, not every state is an eigenstate of the angular momentum operator along any particular axis. In particular, states of the form Eq. 2.13 have *no* angular momentum orientation at all: $\langle\psi(\phi)|J_{\hat{x}}|\psi(\phi)\rangle = \langle\psi(\phi)|J_{\hat{y}}|\psi(\phi)\rangle = \langle\psi(\phi)|J_{\hat{z}}|\psi(\phi)\rangle = 0$. Therefore, we should be extremely careful about importing too much intuition from the simple case of a two-level system.

While $|\psi(\phi)\rangle$ does not have any angular momentum orientation $\langle J_{\hat{u}}\rangle$, it *must* have components of angular momentum *alignment*, $\langle J_{\hat{u}}^2\rangle$, because $\langle J_{\hat{x}}^2 + J_{\hat{y}}^2 + J_{\hat{z}}^2\rangle = J(J+1) = 2$. In particular, for the state in Eq. 2.13, we see that $\langle J_{\hat{z}}^2\rangle = 1$, $\langle J_{\hat{x}}^2\rangle = \cos^2\phi$, and $\langle J_{\hat{y}}^2\rangle = \sin^2\phi$. In other words, $|\psi(0)\rangle$ is aligned in the xz-plane and $|\psi(\frac{\pi}{2})\rangle$ is aligned in the yz-plane. When speaking of the angular momentum alignment for these states, we typically ignore the component along the z-axis and simply say that $|\psi(0)\rangle \equiv |X\rangle$ is "aligned along the x-axis," and $|\psi(\frac{\pi}{2})\rangle \equiv |Y\rangle$ is "aligned along the y-axis."

It turns out that for a system restricted to a subspace with $J = 1$, we can simultaneously measure the alignment along both $\hat{x}$ and $\hat{y}$, $\langle J_{\hat{x}}^2\rangle = \cos^2\phi$ and $\langle J_{\hat{y}}^2\rangle = \sin^2\phi$, which each independently provide information about the angle ϕ of the spin alignment in the xy-plane[7]. With the geometric picture of a vector aligned in the xy-plane in mind, we can do this simply by projecting an arbitrary state along either $|X\rangle$ or $|Y\rangle$. Indeed, $|\langle X|\psi(\phi)\rangle|^2 = \cos^2\phi$ and $|\langle Y|\psi(\phi)\rangle|^2 = \sin^2\phi$. In particular, we project the molecular population onto the $|X\rangle$ state with probability $\cos^2\phi$, and necessarily also project onto the orthogonal $|Y\rangle$ state with probability $\sin^2\phi$, by exciting molecules with laser light such that those in $|X\rangle$ emit a fluorescence photon and those in $|Y\rangle$ do not. If the number of total molecules were known and every fluorescence photon were detected, this would enable us to determine $\cos^2\phi$, and therefore ϕ, with minimum uncertainty. However, given experimental limitations such as finite photon detection efficiency, it is possible to improve the estimate of ϕ by subsequently exciting the molecules in $|Y\rangle$ with a laser that causes them to emit fluorescence photons.

---

7. This result holds because of two facts. First, the operators $J_{\hat{x}}$ and $J_{\hat{y}}$ cannot change an angular momentum quantum number J when acting on an angular momentum eigenstate $|JM\rangle$. Therefore, it is useful to define $J_{\hat{x}}^{2(1)}$ and $J_{\hat{y}}^{2(1)}$ to be the matrix representations of $J_{\hat{x}}^2$ and $J_{\hat{y}}^2$ acting on the restricted subspace of angular momentum eigenstates with $J = 1$. Second, it turns out that these matrices satisfy $[J_{\hat{x}}^{2(1)}, J_{\hat{y}}^{2(1)}] = 0$. It follows that $\langle J_{\hat{x}}^2\rangle$ and $\langle J_{\hat{y}}^2\rangle$ can be simultaneously measured for a system restricted to $J = 1$. The same reasoning holds, in fact, for $J = 0$ and $J = 1/2$, but not for any $J > 1$.

Molecules that were originally projected onto $|X\rangle$ do nothing, while those that had been projected onto $|Y\rangle$ emit a fluorescence photon.

Following the approach just described, suppose that we project an ensemble of molecules in state $|\psi(\phi)\rangle$ onto either $|X\rangle$ or $|Y\rangle$ and measure the average projected populations, $S_X \propto |\langle X|\psi(\phi)\rangle|^2$ and $S_Y \propto |\langle Y|\psi(\phi)\rangle|^2$. We can then construct the asymmetry

$$\mathcal{A} = \frac{S_X - S_Y}{S_X + S_Y} = \cos(2\phi), \tag{2.14}$$

where the RHS is the expected value of $\mathcal{A}$. The denominator, $S_X + S_Y$, normalizes for the total number of molecules and detection efficiency (which typically fluctuate from measurement to measurement). In the phase regime $\phi = \pm\frac{\pi}{4} + \delta\phi$ where the sensitivity to small changes in the phase, $\delta\phi$, is maximized, we find $\mathcal{A} = \mp 2\delta\phi$. In practice, experimental imperfections lead to a slightly reduced sensitivity to small changes in the phase, which we parametrize by the contrast $\mathcal{C}$. We therefore write

$$\mathcal{A} = \mp 2|\mathcal{C}|\delta\phi. \tag{2.15}$$

We want to measure $\delta\phi = d_e \mathcal{E}_{\text{eff}} \tau$, given the asymmetry $\mathcal{A}$ and some calibration of $|\mathcal{C}|$. Therefore, $\delta\phi = \mp\frac{\mathcal{A}}{2|\mathcal{C}|}$.

Returning briefly to the $J = \frac{1}{2}$ analogy, note that $\langle\psi(\frac{\pi}{2})|\psi(0)\rangle \neq 0$ in that system, so that the method of alternately projecting the spin orientation along the x-axis and y-axis would not work (or at minimum would need to be modified). Indeed, it is necessary in that case to alternately project along, e.g., $|\psi(0)\rangle$ and $|\psi(\pi)\rangle$, but then the geometric interpretation of measuring the alignment along orthogonal *spatial* (rather than Hilbert space) axes does not apply. Finally, except in special cases where the total angular momentum is restricted to a given value $J \leq 1$, it is not generally possible to measure the spin alignment along the x- and y-axes simultaneously (see discussion in Footnote 7). In this sense, our measurement procedure can only be applied directly to systems restricted to a subspace with $J = 1$, and one should be careful about leaning too much on analogies with simpler (or for that matter,
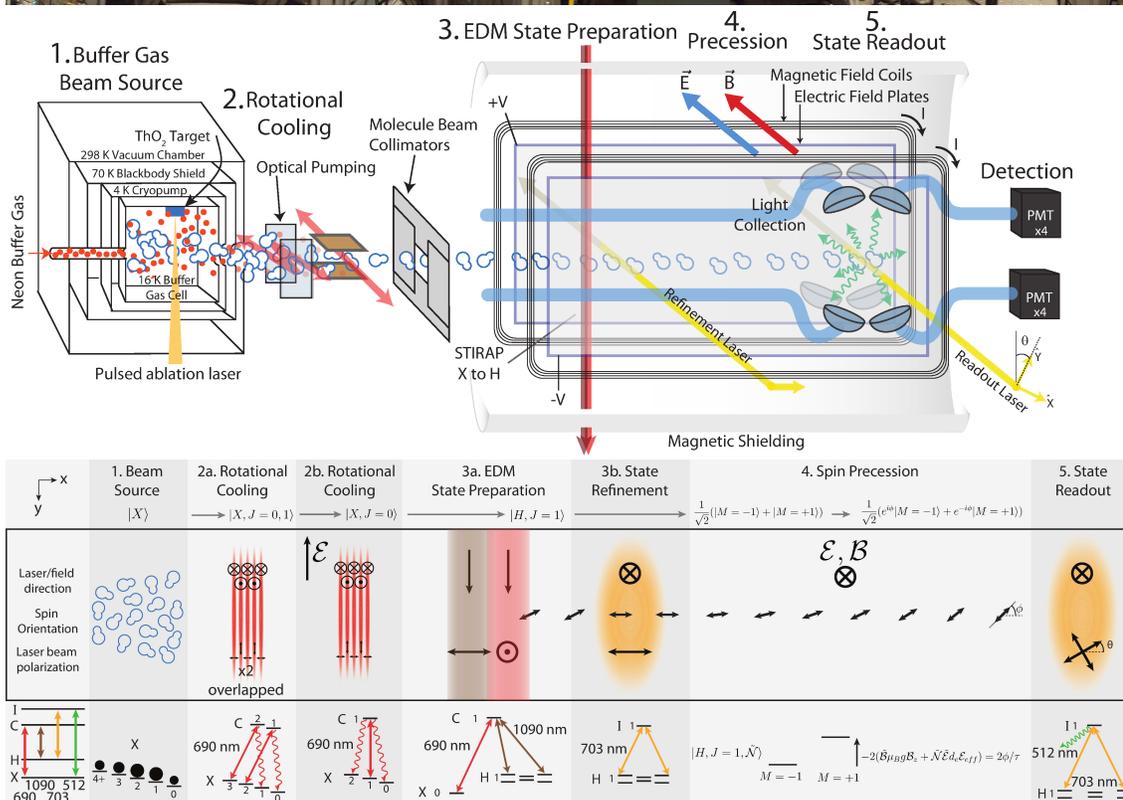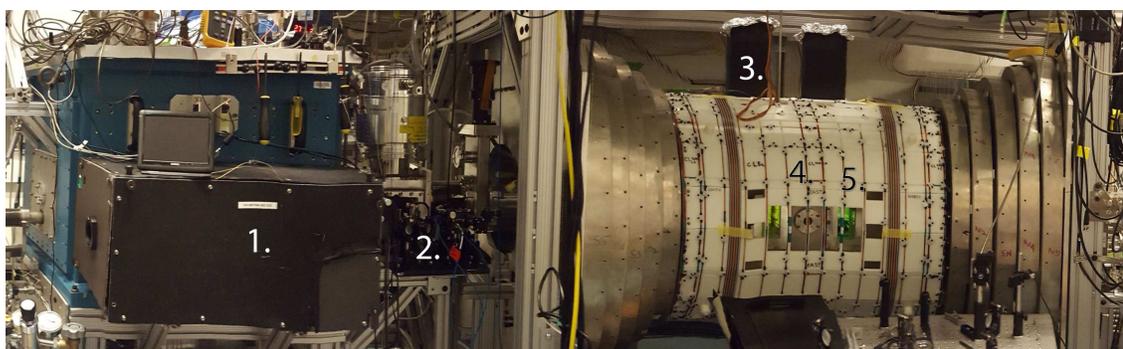
Figure 2.2: Three representations of the ACME apparatus, in increasing abstraction from top (photograph) to bottom. Components are approximately aligned in each subfigure. Center shows cartoon version of the physical apparatus (minimally modified from a figure by Brendon O'Leary); bottom shows molecular state, laser and field configurations, and energy level diagrams for each stage (significantly modified from version by Adam West). Detailed description in caption continued on next page (pages best viewed side-by-side in print).

Figure 2.2: Caption continued from previous page.

[Top panel] Photograph of the apparatus. (1.) A pulsed Nd:YAG ablation laser enters the black enclosure to the left and is directed into the cryogenic "beam box" (blue vacuum chamber), where it is incident on a ThO_2 "target," producing ThO gas. Neon gas entrains the ThO molecules in a beam, travelling along $\hat{x}$ from left to right. (2.) Up to three (typically two) optical pumping lasers consolidate population to $J = 0$. (3.) The molecules enter the interaction region about 1.1 m from the ablation cell. In the photograph, the applied magnetic field coils are visible (white frame), as well as the five sets of μ-metal "endcaps." STIRAP laser beams travel vertically through the left-hand black enclosure above the interaction region. The refinement laser passes through the horizontal vacuum window on the left (green appearance due to anti-reflection coating) to reproject the molecular state (4.) The ≈ 20 cm region between vacuum windows where molecules freely precess in electric and magnetic fields. (5.) The detection region. Lasers pass through the vacuum window on the right, exciting molecules that subsequently decay. The fluorescence photons are collected and detected on photomultiplier tubes (PMTs). Molecules pass through end of the interaction region and are pumped away to the right of the photograph.

[Center panel] Simplified schematic of the apparatus, not to scale. (1.) All radiation shielding and cryopumping layers of the beam box are shown. (2.) The first two rotational cooling lasers optically pump molecules in a field-free environment. The last rotational cooling laser operates in an $\mathcal{E}$-field applied along $\hat{y}$ (brown plates). (3.) As the molecules enter the interaction region, they are collimated by a square arrangment of razor blades. Vertical STIRAP lasers transfer population to H and a horizontal refinement laser, passing through the field plates along $\hat{z}$, reprojects the state. (4.) The molecules precess for ≈ 20 cm. (5.) Lasers with orthogonal polarizations stimulate molecules to an excited state and fluorescence photons (green wavy arrows) are detected via an array of lens doublets, each of which focuses light into bent light pipes. The photons are subsequently detected in PMTs.

[Bottom panel] Top section shows laboratory coordinate system and the molecular state(s) resulting from each stage of the experiment. Center section shows laser, external field, and spin alignment configurations. Bottom section shows electronic energy levels and lasers used in the experiment (left), and energy levels and transitions occurring in each stage. (1.) In the beam source, molecules are produced in a variety of rotational (J) states, with black circles representing the relative thermal populations. (2a.) Two overlapped lasers, stretched along $\hat{y}$, pump $|J = 2, 3\rangle$ to $|J = 0, 1\rangle$. (2b.) A single laser partially transfers $|J = 1\rangle$ to $|J = 0\rangle$ in the presence of an electric field. All rotational cooling lasers travel along $\pm\hat{z}$ with alternating polarization and direction on each pass. (3.) Two partially overlapped lasers, travelling along $\hat{y}$, transfer population from $X \rightarrow H$ via STIRAP. Any imperfection in the spin alignment (double-headed arrow) is attenuated by the refinement laser, travelling along $\hat{z}$, which optically pumps away the unwanted spin orientation via the $H \rightarrow I$ transition. (4.) The molecules precess in electric and magnetic fields, accumulating a phase ϕ. (5.) The phase is read out by detecting fluorescence photons via $H \rightarrow I \rightsquigarrow X$ with alternating laser polarizations in the $H \rightarrow I$ excitation lasers, travelling horizontally.

more complex) systems.[8]

One last analogy can also be useful at times. Note that photons carry one unit of angular momentum, analogous to a molecule with $J = 1$, but longitudinal polarizations do not exist. Because of this restriction, the angular momentum of a molecular system in a linear combination of states $|J = 1, M = \pm 1\rangle$ has a one-to-one correspondence with the angular momentum of a photon, which is described by a linear combination of polarization states $|\epsilon = \pm 1\rangle$. Here, ϵ denotes the photon's angular momentum projection along its propagation axis.

2.3 Experimental apparatus

With the essential concepts of the ACME spin precession measurement in hand, we can describe the components of the apparatus and the measurement procedure in more detail. See Fig. 2.2 for a photograph of the apparatus, cartoon schematic, and diagram of energy levels and field configurations in each stage. In each representation, we have labeled the apparatus as consisting of five parts: (1) Buffer gas beam source; (2) Rotational cooling; (3) EDM state preparation; (4) Spin precession; (5) State readout. I will describe each of these stages in turn.

2.3.1 Beam source

For details of the buffer gas beam source, see [146, 152]; no significant modifications of the beam source were made between ACME I and ACME II[9]. Briefly, we cool a copper cell to ~ 16 K using a pulse tube refrigerator, and ablate a cylindrical "target" of ceramic ThO_2 via a high-power pulsed laser ($\gtrsim 40$ mJ/pulse, ~ 15 ns pulse duration, 50 Hz repetition

---

8. Of course, a physical interpretation of what the molecular state is doing is entirely unnecessary, as long as we can design a protocol that gives us a measurement of the EDM energy – which is certainly possible for any J.

9. While the buffer gas source is generally quite robust, it does require regular maintenance, especially replacing cells and ablation "targets" that generate our beam of ThO molecules – work made much less convenient by the radioactivity of thorium. Furthermore, the high-power pulsed ablation laser requires regular cleaning, flash lamp replacements, and repairs much more frequent than we would wish. Although this was one of my particular roles on the experiment, there is little of novel scientific interest to describe in this dissertation.

rate). This produces a cloud of gaseous ThO (aalong with, almost certainly, several other neutral and ionic species, which are not addressed by any of our lasers and therefore have no significant effect on our experiment). During this time, neon is continuously flowed through a fill line at the back of the cell and thermalizes with the 16 K copper walls. The neon, which is at a much higher density than the ThO produced during ablation, thermalizes with the ThO molecules and entrains them in a flow out of the ablation cell, creating a beam of both neon and ThO. An "ion sweeper" downstream, consisting of two metal plates held at a potential difference of 500 V, deflects any ions produced during ablation out of the molecular beam.

The beam diverges with a half-angle on the order of $\sim 45°$ full-width-half-maximum (FWHM). The central region passes through a "skimming" aperture, 6 mm in diameter, while the rest of the beam sticks ("is cryopumped") to the 4 K surfaces surrounding the 16 K cell. This is by far the lossiest step in our experiment: ultimately, only one in several thousand molecules will reach the detection region due to molecular beam divergence.

2.3.2 Rotational cooling

The ThO molecules emerge from the buffer gas beam source overwhelmingly in their ground vibrational and electronic state. However, their rotational tepmerature is ~ 4 K, while the spacing between energy levels is set by $B = 0.333\,\mathrm{cm}^{-1} \sim 0.5$ K. As a result, the first three rotationally excited states have significant fractions of the overall population ($> 10\%$). Therefore, after the ThO molecules ballistically exit the buffer gas beam source, we optically pump the first three excited rotational levels to the rotational ground state, enhancing the signal in a single quantum state by a factor of $\approx 2 - 2.5^{10}$. I will describe this optical pumping, which we refer to as "rotational cooling," in much more detail in subsequent sections. Briefly, we drive $X(J = 2) \to C(J = 1) \rightsquigarrow X(J = 0)$ and $X(J = 3) \to C(J = 2) \rightsquigarrow X(J = 1)$, where $\to$ denotes a change of state by laser light and $\rightsquigarrow$ denotes a change of state by spontaneous emission. The $X(J = 1)$ population is subsequently driven

---

10. In practice, we typically optically pump only the first two excited states to avoid the technical inconvenience of keeping an additional laser locked for the modest increase in signal due to the $X(J = 3)$ population.

(partially) to the ground state by $X(J = 1) \rightarrow C(J = 1) \rightsquigarrow X(J = 0)$.

2.3.3 State preparation

After the molecules are rotationally cooled, the molecular beam passes into the magnetically shielded region of the apparatus that we refer to as the "interaction region" since this is where we observe the interactions between the molecules and external electric and magnetic fields. Here, five nested layers of mu-metal in a cylindrical geometry reduce the magnetic field to at most a few hundred μG. Within the mu-metal shields, we have a set of magnetic field coils that allows us to apply any average magnetic field, $\mathcal{B}_x$, $\mathcal{B}_y$, $\mathcal{B}_z$, or independent gradient $\partial_x \mathcal{B}_x$, $\partial_y \mathcal{B}_y$, $\partial_x \mathcal{B}_y$, $\partial_x \mathcal{B}_z$, $\partial_y \mathcal{B}_z$, the other gradients being related to these by Maxwell's equations. The molecular beam is collimated by fixed razor blades (square geometry with 24 mm side length) and passes between electric field plates made from Corning 7980 OA glass coated with a 20 nm layer of indium tin oxide (ITO). The plates are transparent and separated by 4.5 cm, allowing for horizontal laser beams to pass through them as well for vertical laser beams to pass between them.

Before proceeding, it is useful to introduce a rigorously defined coordinate system. Let $\hat{z} \propto \langle \vec{\mathcal{E}}_{\text{lab}} \rangle$ lie along the average electric field experienced by the molecules between state preparation and state readout (discussed shortly), while $\hat{x} \propto \vec{v} - v_z \hat{z}$ lies approximately along the average molecular velocity but is constrained to be perpendicular to $\hat{z}$. This uniquely defines $\hat{y} = \hat{z} \times \hat{x}$, approximately aligned with gravity. Under normal conditions, the applied electric and magnetic fields are always aligned along or against $\hat{z}$. As seen already, we specify the direction of the electric field by $\hat{\mathcal{E}}_{\text{lab}} = (\hat{\mathcal{E}}_{\text{lab}} \cdot \hat{z})\hat{z} \equiv \tilde{\mathcal{E}}\hat{z}$. Analogously, we specify the direction of the magnetic field by $\hat{\mathcal{B}}_{\text{lab}} = (\hat{\mathcal{B}}_{\text{lab}} \cdot \hat{z})\hat{z} \equiv \tilde{\mathcal{B}}\hat{z}$.

STIRAP

After passing between the field plates, the molecules are coherently transferred from $X(J = 0)$ to $H(J = 1, M = \pm 1)$ via the intermediate state $C(J = 0, P = -1)$ using the STImulated Raman Adiabatic Passage (STIRAP) technique. The implementation of STIRAP was led by Cris Panda; see [155] for full technical details, and Brendon's thesis [149] for some nice discussion, as well as Cris's thesis (in preparation as of this writing). The two STIRAP

laser beams travel vertically, along $\hat{y}$, and are focused along the x-direction to a waist size of $w_x \approx 150\,\mu$m to achieve high intensities with the laser powers available, but extended along the z-direction to address the entire molecular beam cross-section (≈ 25 mm). The laser beams are partially overlapped along $\hat{x}$ such that molecules first pass through the $C \leftrightarrow H$ laser (1090 nm, ~ 10 W) and then through the $X \leftrightarrow C$ laser (690 nm, ~ 100 mW).

The molecules therefore experience a time-dependent intensity profile that looks like: the $C \leftrightarrow H$ laser turns on, the $C \leftrightarrow H$ laser begins to turn off as the $X \leftrightarrow C$ laser turns on, and finally the $X \leftrightarrow C$ laser turns off. The foolproof way to see how STIRAP works is, as always, to write down the Hamiltonian and compute the time-evolution. In particular, consider the three-level system $|g_1\rangle = |X\rangle$, $|e\rangle = |C\rangle$, and $|g_2\rangle = |H\rangle$, with time-dependent couplings (Rabi frequencies) $\Omega_{g_1e}(t)$ and $\Omega_{g_2e}(t)$ via laser light at frequencies ω_{g_1e} and ω_{g_2e}. Then the Hamilton in the $\{|g_1\rangle, |g_2\rangle, |e\rangle\}$ basis is

$$
H = \begin{bmatrix} E_{g_1} & 0 & \Omega_{g_1e}(t)\cos(\omega_{g_1e}t) \\ 0 & E_{g_2} & \Omega_{g_2e}(t)\cos(\omega_{g_2e}t) \\ \Omega^*_{g_1e}(t)\cos(\omega_{g_1e}t) & \Omega^*_{g_2e}(t)\cos(\omega_{g_2e}t) & E_e \end{bmatrix}. \tag{2.16}
$$

For a pedagogic discussion of STIRAP, and in particular how to use Eq. 2.16 to transfer population from $|g_1\rangle$ to $|g_2\rangle$ without populating $|e\rangle$, see [164].

To help build some intuition for how STIRAP works in general, I will consider an extremely simple system, not quite describing the state preparation in ACME, where the "ground states" $|x\rangle = |\psi(0)\rangle$ and $|y\rangle = |\psi(\frac{\pi}{2})\rangle$ are degenerate substates of a $J = 1$ manifold, of the form Eq. 2.13, and they are coupled to $|e\rangle = |J = 0\rangle$ via laser light with orthogonal linear polarizations, e.g., $\hat{x}$ and $\hat{y}$. In this case, the transition matrix elements are proportional to $\langle e|\vec{\epsilon}\cdot\vec{r}|g\rangle$. If $\vec{\epsilon} = \cos\theta\hat{x} + \sin\theta\hat{y}$ is an arbitrary linear polarization vector in the xy-plane, then the transition rate is proportional to $|\langle e|\vec{\epsilon}\cdot\vec{r}|g\rangle|^2 \propto \cos^2(\theta - \phi)$: in particular, the state with angular momentum aligned with the polarization vector is "bright," while the state with angular momentum alignment perpendicular to the polarization vector is "dark" (i.e., the transition rate vanishes).

Suppose we begin with a state $|\psi(\phi = 0)\rangle = |X\rangle$ and turn on laser light addressing the dark state, $\theta = \frac{\pi}{2}$. Then the transition rate vanishes, $\cos^2(\theta - \phi) = 0$, and the population

remains in the dark state. If we now rotate the laser polarization sufficiently slowly, $\theta \to \theta + \dot{\theta}t$, where $\dot{\theta}$ is the rate of polarization rotation, then the population remains in the instantaneous dark state according to the adiabatic theorem [165, Ch. 10][11]. When the polarization has rotated by a full 90 degrees, so that $\theta = \pi$, the dark state is parametrized by $\phi = \frac{\pi}{2}$. In other words, the population is now in $|\psi(\frac{\pi}{2})\rangle = |Y\rangle$ and complete population transfer has been performed from $|X\rangle$ to $|Y\rangle$, without ever populating an excited state.

Although we just described the STIRAP mechanism in terms of rotating the polarization vector, $\theta \to \theta + \dot{\theta}t$, one could equivalently consider the powers in the $\hat{x}$-polarized and $\hat{y}$-polarized components over time: $\hat{\epsilon}_x = \cos(\theta + \dot{\theta}t)$ and $\hat{\epsilon}_y = \sin(\theta + \dot{\theta}t)$. Starting at $\theta = 0$, this looks like we slowly ramp down the power of the $\hat{x}$-polarized light as we ramp up the power of the $\hat{y}$-polarized light.

In the ACME experiment, the "ground states" $|g_1\rangle = |X\rangle$ and $|g_2\rangle = |H\rangle$ are coupled to $|e\rangle = |C\rangle$ via laser light with different wavelengths (rather than different polarizations), but otherwise the essential picture is unchanged. If we first address the $|H\rangle \leftrightarrow |C\rangle$ transition, and then slowly ramp down the $|H\rangle \leftrightarrow |C\rangle$ as we ramp up the $|X\rangle \leftrightarrow |C\rangle$ power, the population will be adiabatically transferred from $|X\rangle$ to $|H\rangle$ without ever populating $|C\rangle$. We have demonstrated transfer of $\approx 75\%$ of the molecules into the H state using this method.

Because the STIRAP lasers travel along $\hat{y}$, they can have polarization in the xz-plane. We choose to polarize the 690 nm $X \leftrightarrow C(\tilde{\mathcal{P}} = -1)$ laser along $\hat{z}$ and the 1090 nm $C(\tilde{\mathcal{P}} = -1) \leftrightarrow H$ laser along $\hat{x}$, which fixes the initial molecular population to be in the bright state $|\psi(\phi = \frac{\pi}{2})\rangle$ under ideal conditions, where $\tilde{\mathcal{P}}$ denotes the parity of the $C(J = 1, M = 0)$ Ω-doublet that is addressed by the lasers. It would be possible to also implement STIRAP via the intermediate $C(\tilde{\mathcal{P}} = +1)$ state, but doing so would require the use of additional acousto-optical modulators (AOMs), which entails power loss. Because the STIRAP lasers (especially the 1090 nm $C \leftrightarrow H$ laser) are not fully saturated, reducing the power further would reduce the preparation efficiency and make the measurement scheme more susceptible

---

11. A careful application of the adiabatic theorem requires an energy splitting, Δ, between the bright and dark states, satisfying $\Delta \gg \dot{\theta}$. In the case considered here, this energy splitting would arise from the AC Stark shift of the bright ground state (or Autler-Townes shift in the case of exactly resonant excitation).

to systematic errors that depend on the STIRAP two-photon lineshape.

State refinement

In practice, AC Stark shifts, coupled to other imperfections in our apparatus, can lead to an unwanted component of the phase prepared during the STIRAP state transfer, $|\psi(\phi \neq \frac{\pi}{2})\rangle$ (see Sec. 4.6 for details). As a result, we "refine" the molecular state, $|\psi(\phi)\rangle \to |\psi(\frac{\pi}{2})\rangle$[12] in the manner described below. By controlling the frequency of the $C \leftrightarrow H$ STIRAP laser, we can populate either $\tilde{\mathcal{N}} = \pm 1$ manifold (separated in energy by ~ 100 MHz due to the Stark shift).

Immediately downstream of the STIRAP lasers, we reproject the spin orientation with a 703 nm laser travelling horizontally through the field plates with linear polarization at angle θ_{prep}, which drives $H \leftrightarrow I(\tilde{\mathcal{P}} = +1)$, optically pumping away the component of the STIRAP-prepared state along $|\psi(\phi = \theta_{\mathrm{prep}})\rangle$. Thus the dark state of the refinement laser, $|\psi(\phi = \theta_{\mathrm{prep}} + \frac{\pi}{2})\rangle$, is reliably prepared. In order to maximize the surviving population, we align the polarization approximately along $\hat{x}$, $\theta_{\mathrm{prep}} \approx 0$. We tune the frequency of the refinement lasers to address whichever $\tilde{\mathcal{N}}$ manifold was prepared by STIRAP. We regard the state prepared by the refinement (also referred to as "cleanup" or "preparation") laser as the initial state of the spin precession measurement. Just as with the STIRAP transition, it would be possible to refine the molecular state via the $H \leftrightarrow I(\tilde{\mathcal{P}} = -1)$ transition by using additional AOMs, at the expense of a reduction in the refinement laser power. However, in order to maximize the power available for state refinement, we have chosen not to implement a refinement laser parity switch.

2.3.4 Spin precession

After the molecules have been prepared in $|\psi(\phi = \theta_{\mathrm{prep}} + \frac{\pi}{2})\rangle$, they travel ballistically for ≈ 20 cm through the interaction region, where there is an approximately uniform electric field $\vec{\mathcal{E}}$ and magnetic field $\vec{\mathcal{B}}$. The molecular angular momentum alignment precesses with

---

12. As we will see, we need not prepare $\phi = \frac{\pi}{2}$ exactly: what actually matters is simply that we prepare a *consistent* phase ϕ prior to the spin precession stage of the measurement.

frequency

$$
\begin{aligned}
\omega &\approx \omega_\mu + \omega_{d_e} \\
&= -(g\mu_B \mathcal{B}\tilde{\mathcal{B}} + d_e \mathcal{E}_{\text{eff}} \tilde{\mathcal{N}} \tilde{\mathcal{E}}),
\end{aligned}
\tag{2.17}
$$

where $g = -0.00440(5)$ in $H(J = 1)$[67], $\mathcal{B} \equiv |\vec{\mathcal{B}}|$ is the magnitude of the applied magnetic field, and $\tilde{\mathcal{B}} \equiv \text{sign}(\hat{\mathcal{B}} \cdot \hat{z}) = \pm 1$ denotes the direction of the magnetic field, in analogy to $\tilde{\mathcal{E}}$. This induces the prepared state to evolve after a time τ,[13]

$$
|\psi(\tau), \tilde{\mathcal{N}}\rangle = \frac{1}{\sqrt{2}}(e^{-i(\theta_{\text{prep}} + \frac{\pi}{2} + \omega\tau)}|+, \tilde{\mathcal{N}}\rangle + e^{i(\theta_{\text{prep}} + \frac{\pi}{2} + \omega\tau)}|-, \tilde{\mathcal{N}}\rangle).
\tag{2.18}
$$

The molecular velocity is typically ~ 200 m/s; hence, the precession time is ~ 1 ms. However, due to variations in ablation conditions between molecular pulses, and velocity dispersion within each pulse, τ may vary by several hundred microseconds.

2.3.5 State readout

As we described schematically in Sec. 2.2.2, we "read out" the state by alternately projecting along orthogonal states. In particular, we use two lasers in the state readout region, ≈ 20 cm downstream of the refinement laser, addressing the $H \leftrightarrow I(J = 1, M = 0, \tilde{\mathcal{P}})$ transition where $\tilde{\mathcal{P}} = \pm 1$ may be chosen freely (unlike STIRAP and state refinement, where we only use $\tilde{\mathcal{P}} = -1$ and $\tilde{\mathcal{P}} = +1$, respectively). Let us first excite molecules in $|\psi(\tau), \tilde{\mathcal{N}}\rangle$ via laser light, travelling horizontally through the field plates, which has polarization $\theta_X \equiv \theta_{\text{prep}} + \delta\theta$. We refer to this as the X laser. Note that when the refinement laser is polarized along $\hat{x}$ and $\delta\theta = 0$, the X laser is polarized along $\hat{x}$ as well; however, in general the X laser may have any linear polarization in the xy-plane. (For simplicity, we are neglecting the possibility of laser beam misalignment, which would allow a component of the polarization along $\hat{z}$, and polarization ellipticity.) This laser projects the molecules onto the bright state/dark state basis. The molecules projected onto the bright state are excited to the short-lived I state, which subsequently decays and emits a fluorescence photon for each bright-state molecule.

---

13. Most internal references in ACME handle the factors of $e^{\pm i\pi/2}$ by introducing a negative sign between the two terms and absorbing the overall factor of i as an arbitrary phase. During this discussion, I have chosen instead to always write the molecular state in a form that makes the angular momentum alignment explicit in order to facilitate a physical interpretation.

In particular, the projected population, and therefore fluorescence photon "signal," is on average proportional to

$$
\begin{aligned}
\langle S_X \rangle \;\; &\propto \;\; \cos^2[(\theta_{\mathrm{prep}} + \tfrac{\pi}{2} + \omega\tau) - (\theta_{\mathrm{prep}} + \delta\theta) + \tfrac{\pi}{4}(\tilde{\mathcal{P}} - 1)] \\
&= \;\; \cos^2[\omega\tau - \delta\theta + \tfrac{\pi}{4}(\tilde{\mathcal{P}} + 1)].
\end{aligned}
\tag{2.19}
$$

We then excite molecules in the orthogonal state by exciting with laser light that has polarization at angle $\theta_Y \equiv \theta_{\mathrm{prep}} + \delta\theta + \tfrac{\pi}{2}$, and the projected population is proportional to

$$
\begin{aligned}
\langle S_Y \rangle \;\; &\propto \;\; \cos^2[(\theta_{\mathrm{prep}} + \tfrac{\pi}{2} + \omega\tau) - (\theta_{\mathrm{prep}} + \delta\theta + \tfrac{\pi}{2}) + \tfrac{\pi}{4}(\tilde{\mathcal{P}} - 1)] \\
&= \;\; \sin^2[\omega\tau - \delta\theta + \tfrac{\pi}{4}(\tilde{\mathcal{P}} + 1)].
\end{aligned}
\tag{2.20}
$$

Note the averaging notation $\langle S_{X(Y)} \rangle$, which reminds us that in any given run of the experiment the signals measured from the X and Y fluorescence might deviate from their average values. We can point out a few important features already: the "global" polarization angle θ_{prep} drops out entirely (except insofar as it couples to imperfections in the experiment not considered here, or modulates the overal signal size $S_X + S_Y$). Only the precession phase $\omega\tau$, relative to the angle between the preparation and readout bases $\delta\theta$, and the parity of the excited state $\tilde{\mathcal{P}}$ can affect the relative strength of signals S_X and S_Y. In particular, changing $\tilde{\mathcal{P}}$ from $+1$ to -1 or vice versa is equivalent to rotating the readout basis by $\tfrac{\pi}{2}$. This fact will be important in suppressing certain classes of systematic errors.

In a bit more detail, we spatially overlap the two lasers at polarization angles θ_X and θ_Y and alternately turn them on for $1.9\,\mu$s at a time, switching between them at a rate of 200 kHz. We integrate the photoelectron signal (from detected fluorescence) over approximately the duration that each laser is on, giving signals S_X and S_Y. There is a $0.6\,\mu$s gap when neither laser is on, allowing the molecular population to decay away so that the signal nominally from molecules excited by the Y laser has negligible contamination from molecules that were previously excited by the X laser. The dead time between laser excitation also prevents any coherent transfer between the bright states of each laser, which would complicate the data analysis and possibly lead to systematic errors.

As discussed in Sec. 2.2.2, we construct the asymmetry $\mathcal{A} = \frac{S_X - S_Y}{S_X + S_Y}$, which normalizes

against fluctuations in the total number of molecules, and removes any need to calibrate the overall detection efficiency.. Then $\langle\mathcal{A}\rangle = \cos[2(\omega\tau - \delta\theta + \frac{\pi}{4}(\tilde{\mathcal{P}} + 1))]$. Let us define $\omega\tau - \delta\theta \equiv n\pi \pm \frac{\pi}{4} + \delta\phi$, choosing the sign of $\pm$ and value of integer n such that $\delta\phi$ is minimized (and in particular, $\delta\phi < \frac{\pi}{4}$ strictly). Then $\langle\mathcal{A}\rangle = \pm\tilde{\mathcal{P}}\sin(2\delta\phi)$. We can always set $\delta\theta$ in the experiment such that $\delta\phi \ll 1$, and then find

$$\langle\mathcal{A}\rangle \approx (\pm\tilde{\mathcal{P}})2\delta\phi. \tag{2.21}$$

For example, when $\omega\tau \approx +\frac{\pi}{4}$ and we address $\tilde{\mathcal{P}} = +1$, $\delta\phi < 0$ corresponds to a greater expected signal S_Y (corresponding to molecules with angular momentum alignment closer to that prepared by STIRAP) and thus a *negative* asymmetry.

In any particular run of the experiment, this relationship allows us to *define* the measured phase, $\delta\phi_m \equiv \pm\tilde{\mathcal{P}}\frac{\mathcal{A}}{2}$, so that the average value of the measured phase (defined in this way) equals the true value, $\langle\delta\phi_m\rangle = \delta\phi$. Note that due to statistical noise, the "measured" phase in any particular measurement will not precisely equal the actual phase of the molecules.

In practice, we operate in one of two regimes: (1) at "low" magnetic field values, $\omega\tau \ll 1$ and $\delta\theta \approx \frac{\pi}{4}$; and (2) at "high" magnetic field values, $\omega\tau \approx \frac{\pi}{4}$ and $\delta\theta \approx 0$. Thus in the low magnetic field regime, we choose the $-$ sign in Eq. 2.21, while in the high magnetic field regime, we choose the $+$ sign, reversing the slope of asymmetry vs. phase deviation, $\frac{d\langle\mathcal{A}\rangle}{d\delta\phi}$.[14] Further, this slope is proportional to $\tilde{\mathcal{P}}$, the parity of the excited state addressed by the readout lasers. Finally, an additional way to reverse the sign of this slope is to rotate the entire readout polarization basis by $\pm\frac{\pi}{2}$, effectively swapping the roles of the X and Y lasers. In our formal treatment, this operation reverses the sign of the slope $\frac{d\langle\mathcal{A}\rangle}{d\delta\phi}$ because it interchanges $\pm \leftrightarrow \mp$ simultaneously with changing the value of n by one.

Thus the asymmetry contains information about small deviations in phase from a reference condition. We extract quantities of interest, such as d_e, by operating the experiment in

---

14. We will see later that it is useful to absorb the sign of this slope into a definition of the asymmetry averaged over two experiment "states," or settings of the experimental switches discussed in Sec. 2.5 such as the direction of the applied electric and magnetic fields. Historically, an asymmetry with this sign absorbed was referred to as the "$\mathcal{B}$-corrected asymmetry" precisely because the magnitude of the magnetic field, $\mathcal{B}$, changes the sign of this slope.

different conditions and observing a modulation in the extracted values of $\delta\phi$. For example, using the simplified expression for the precession frequency ω in Eq. 2.17, the EDM interaction can be obtained by measuring the asymmetry with both $\tilde{\mathcal{N}}\tilde{\mathcal{E}} = +1$ and $\tilde{\mathcal{N}}\tilde{\mathcal{E}} = -1$ and computing $\langle\mathcal{A}(\tilde{\mathcal{N}}\tilde{\mathcal{E}} = +1)\rangle - \langle\mathcal{A}(\tilde{\mathcal{N}}\tilde{\mathcal{E}} = -1)\rangle \approx -4d_e\mathcal{E}_{\text{eff}}\tau$. We will examine our exact procedure for extracting physical quantities in more detail after considering all experimental switches.

2.4 Experiment time scales

In this section, I will describe the time structure of our data. We record voltages from 8 PMT's that encode a time-dependent molecular fluorescence. The signal has the following structure, from the pulse timescale to longer timescales:

- An independent molecular **pulse** occurs every 20 ms. During each pulse, we save 10 ms of data from each PMT separately. The data from these PMT's constitute a single **shot**. 25 shots are averaged together to form a **trace**. Therefore, a trace takes 0.5 s to record.

- Between traces, we change experimental parameters ($\tilde{\mathcal{N}}, \tilde{\mathcal{E}}$, etc., described in Sec. 2.3). A **block** consists of 64 traces with a degeneracy of 4 in each $\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\theta}\tilde{\mathcal{B}}$ state, and takes approximately one minute to measure. The delay between traces varies depending on which switches are performed.

- A **superblock** consists of 16 blocks, with a degeneracy of 2 in each $\tilde{\mathcal{P}}\tilde{\mathcal{L}}\tilde{\mathcal{R}}$ state. A superblock takes about 15 minutes to measure.

- Many superblocks (order tens) are measured within a **run**.

Each pulse has sub-structure as well.

- The trace has two main regions (see Fig. 2.3(b)):

  - For the first ~ 3 ms of a trace, no molecules are present in the interaction region. We use this region of the signal to obtain a background measurement.
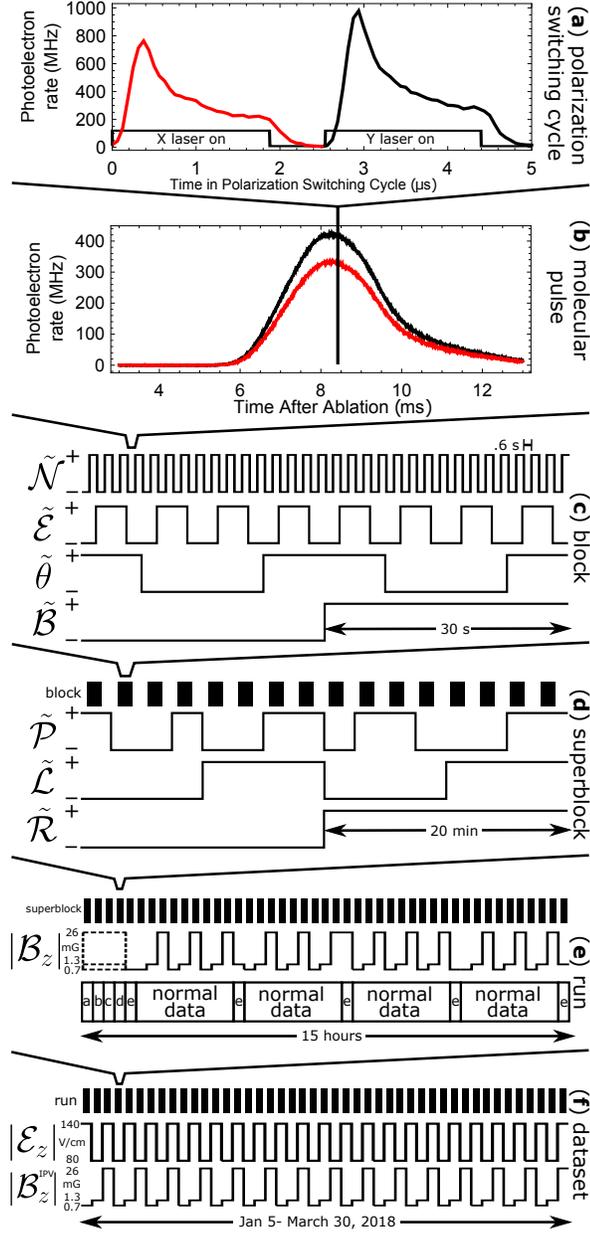
Figure 2.3: Timescales of the ACME II measurement. Figure modified by Cris Panda from version by Brendon O'Leary, and published in the ACME II result paper [156]. (a) Fluorescence signals S_X and S_Y over timescale of one polarization switching cycle. (b) The envelope of a molecular pulse. (c) Block switches, required to extract a precession frequency ω. (d) Superblock switches, used to search for and suppress systematic errors. (d) Changes to the magnitude of the magnetic field, $|\mathcal{B}|$, and "intentional parameter variations" (IPVs, see 4.1). (f) Day-to-day changes in experimental conditions, the magnitude of the electric field, $|\mathcal{E}|$, and the magnetic field at which IPVs are measured, $|\mathcal{B}^{\mathrm{IPV}}|$, throughout the ACME II EDM dataset.

– After 3 ms of background in the trace, the molecules begin to arrive. The pulse envelope rises relatively sharply and has a long falling edge. The region of usable molecules is ~ 3 ms.

- Every 2.5 μs, the linear readout laser polarization is switched from X to Y or vice versa, modulating the molecule signal. Each region with a fixed laser polarization is a polarization **bin**, while a complete two-bin cycle constitutes a **bin pair**. Each bin has three main features (see Fig. 2.3(a)):

  – The beginning of each bin has a brief period (0.6 μs) in which no lasers are on, so the signal is low. The laser then turns on, giving a sharp rise in signal.

  – After a brief time, when the fluorescence rate is dominated by molecules that are already Rabi flopping rather than entering the readout beam from upstream, the signal decays exponentially with rate $\sim \gamma/2$ because each molecule spends half its time in the excited state.

  – The readout laser is turned off again for 0.6 μs at the end of the bin, and the signal decays at rate γ. The "dead time" is set so that the total leftover fluorescence measured in the following bin will be only $\sim 0.1\%$ as large as the total signal in that bin due to newly-excited molecules.

- The signal voltage from the PMT's is acquired through a field-programmable gate array (FPGA) data acquisition (DAQ) system, which records 80 **samples** per polarization bin pair.

Additionally, there are two important timescales that depend on our choices in the data analysis routine:

- Tens of bin pairs are combined to form **groups**. The main advantage of this is to form an empirical uncertainty of a grouped quantity by checking the scatter in individual quantities used to form the group. In addition, grouping data allows us to carry around less data at once, speeding up the analysis. We vary the group size from ~ 10–50 as a systematic check of the analysis routine.

- A **sub-bin** consists of a well-defined region within a polarization bin that we consider to have "good" fluorescence signal. We vary the size and position of the sub-bin as a systematic check of the analysis routine.

2.5 Experimental switches

As we have seen, in order to extract physically meaningful quantities from the raw experiment signals S_X and S_Y, it is necessary to operate the experiment under multiple conditions and compare the resulting changes in phase $\delta\phi$ from the "baseline" of $\delta\phi \approx 0$. In particular, we perform four "block" switches, which are necessary in our data analysis routine to extract a precession frequency from measured asymmetries, and three "superblock" switches on slower time scales, which suppress known and possible systematic errors. See Fig. 2.3 for the timescales of all regular switches.

The $\tilde{\mathcal{N}}$ switch

Every 25 molecular pulses (every ≈ 0.5 seconds), we change which $\tilde{\mathcal{N}}$ state is prepared by STIRAP, refined in the preparation region, and addressed by the two readout lasers, by tuning the frequencies of all lasers coupled to the H state manifold. This is the first of our two "EDM switches," which reverse the sign of the precession frequency associated with a possible EDM.

The $\tilde{\mathcal{E}}$ switch

Approximately once for every four $\tilde{\mathcal{N}}$ state reversals (i.e., 100 molecular pulses, ≈ 2 seconds), we reverse $\tilde{\mathcal{E}}$, the direction of the electric field applied by the field plates in the interaction region, by reversing the voltage on each plate. This is the second "EDM switch." Note that the sign of the EDM contribution to the precession frequency reverses when *either* $\tilde{\mathcal{N}}$ or $\tilde{\mathcal{E}}$ is changed, but not when both are changed simultaneously.

The $\tilde{\theta}$ switch

Approximately once for every four $\tilde{\mathcal{E}}$ state reversals (i.e., 400 molecular pulses, ≈ 10 seconds), we reverse $\tilde{\theta}$, which modulates the relative angle between the readout and polarization bases, $\delta\theta = \delta\theta_0 + \Delta\theta\tilde{\theta}$. Here, $\delta\theta_0 \approx 0$ or $\frac{\pi}{4}$, depending on the strength of the magnetic field (and thus whether $\omega\tau \approx 0$ or $\omega\tau \approx \frac{\pi}{4}$ due to Zeeman precession). We typically set $\Delta\theta = 6°$ by adjusting the angle of a half-wave plate that both readout lasers pass through before entering the interaction region[15]. See [147, Sec. 6.3.2] for an analysis of the optimal step size.

The purpose of this switch is to enable us to measure the sensitivity of the asymmetry to small changes in phase. Recall that we defined $\delta\phi$ by $\omega\tau - \delta\theta \equiv n\pi \pm \frac{\pi}{4} + \delta\phi$, so that $\langle\mathcal{A}\rangle \approx \pm\tilde{\mathcal{P}}\,2\delta\phi$. However, this equation only holds for the ideal case previously considered. Even with perfectly controlled precession frequencies ω, our molecular beam emerges with a spread in velocities $\Delta\tau$, which results in a spread in precession phases $\Delta\phi = \omega\Delta\tau$. This spread in precession phases leads to a reduced sensitivity to the *average* accumulated phase $\omega\bar{\tau}$, where $\bar{\tau}$ is the average molecular precession time probed [144, Sec. 2.2.2].

When we take into account the applied polarization dither $\Delta\theta$, we instead define $\omega\bar{\tau} - \delta\theta_0 \equiv n\pi \pm \frac{\pi}{4} + \delta\phi$, in which case $\langle\mathcal{A}(\tilde{\theta})\rangle = \mp 2\mathcal{C}(\delta\phi - \Delta\theta\tilde{\theta})$, where $\mathcal{C}$ is the measurement "contrast," generically with $|\mathcal{C}| < 1$. With precise control of $\Delta\theta$ and measurements of both $\tilde{\theta}$ states, we can infer the contrast $\mathcal{C}$ by $\langle\mathcal{A}(\tilde{\theta} = +1) - \mathcal{A}(\tilde{\theta} = -1)\rangle = \pm 4\Delta\theta \times \mathcal{C}$. This procedure will be elaborated upon in Sec. 3.1.

The $\tilde{\mathcal{B}}$ switch

Approximately once for every two $\tilde{\theta}$ reversals (i.e., 800 molecular pulses, ≈ 30 seconds), we reverse $\tilde{\mathcal{B}}$, the direction of the applied magnetic field field with respect to $\hat{z}$. The primary purpose of this switch is to reverse the Zeeman interaction, $\omega_\mu = -g\mu_B\mathcal{B}\tilde{\mathcal{B}} = \frac{\phi_\mu}{\tau}$, allowing us to determine the precession time τ and therefore compute precession frequencies ω from measured phases ϕ. The exact means of extracting all quantities of interest will be discussed

---

15. You are warned that factors of two can be confusing here: the polarization angle changes by twice the waveplate rotation, and of course the total range over which the polarization changes is twice the "step size" $\Delta\theta$.

in Sec. 3.1. Note that the $\tilde{\mathcal{N}}$, $\tilde{\mathcal{E}}$, $\tilde{\theta}$, $\tilde{\mathcal{B}}$ switches occur in a "block," which is the minimal set of data required to compute a precession frequency given our switch sequence.

The $\tilde{\mathcal{P}}$ switch

Approximately once for every two blocks (i.e., 3200 molecular pulses, ≈ 2 minutes), we reverse $\tilde{\mathcal{P}}$, the parity of the excited state addressed by the readout lasers. Recall that this is equivalent to rotating the relative angle between the readout and preparation laser polarization bases by $\frac{\pi}{2}$. This operation therefore removes "asymmetry effects," described in more detail in Sec. 4.12, which appear in our measurement as an asymmetry due to imbalances between the X and Y readout beams. For example, suppose we were to systematically acquire more light from the X laser due to optical scatter into our photodetectors. Then the asymmetry, $\mathcal{A} = \frac{S_X - S_Y}{S_X + S_Y}$, would have a positive contribution even when $\phi = 0$ due to $S_X > 0$. However, if we interchange the roles of the X and Y lasers, then $S_X \leftrightarrow S_Y$ and the asymmetry will have a *negative* contribution due to the spurious signal. (In fact, the "roles" of X and Y are distinguished implicitly by the sign of the contrast; this will be explained in more detail in Sec. 4.12.) Therefore, these "asymmetry" effects, unrelated to the actual molecular phase, can be identified and removed by comparing data in the $\tilde{\mathcal{P}} = \pm 1$ states.

The $\tilde{\mathcal{L}}$ switch

Every four blocks (i.e., 6400 molecular pulses, ≈ 2 minutes), we reverse $\tilde{\mathcal{L}}$, which denotes the correspondence between which of the two voltage supplies is connected to which of the two field plates. In particular, let us denote the supplies by 1 and 2, and the field plates by E and W (for east and west). In $\tilde{\mathcal{L}} = +1$, the connections are 1→E and 2→W, while in $\tilde{\mathcal{L}} = -1$, the connections are 1→W and 2→E. Further, when changing the $\tilde{\mathcal{L}}$ state, we reverse the nominal voltage on each power supply to keep the applied electric field in the apparatus unchanged (i.e., if supply 1 is at positive voltage, then we switch it to negative voltage). The switch is implemented via mercury-wetted relays controlled by TTL pulses.

Notice that the $\tilde{\mathcal{L}}$ switch should ideally have no effect on the apparatus. However, if there are non-reversing voltages associated with one or the other power supply, then this will create an electric field contribution in the apparatus that does not switch with $\tilde{\mathcal{E}}$ (which

is implemented by reversing the applied voltage on the supplies). For example, if supply 1 is at voltage 100.1 V in the nominal $\tilde{\mathcal{E}} = +1$ state but at -99.9 V in the nominal $\tilde{\mathcal{E}} = -1$ state, then the non-reversing component of the applied voltage is 0.1 V. However, when we switch the correspondence between the power supplies and the field plates, i.e. in $\tilde{\mathcal{L}} = -1$, the non-reversing component of the applied voltage becomes -0.1 V.

Because non-reversing voltages correspond to non-reversing electric fields, which could contribute to systematic errors in the EDM measurement, we need to be able to remove any contributions to the measured precession frequencies arising from power supply voltage offsets. By comparing data in the $\tilde{\mathcal{L}} = \pm 1$ states, we can identify these contributions and remove them in the data analysis. Note that under normal conditions in ACME II, the $\tilde{\mathcal{L}}$ switch has no statistically significant effect on our EDM measurement, as we would expect in the ideal case (i.e., we could have run without it and our results would still be valid).

The $\tilde{\mathcal{R}}$ switch

The last "superblock" switch in ACME II is the $\tilde{\mathcal{R}}$ switch, which rotates the readout polarization basis by $\frac{\pi}{2}$ via mechanically rotating a waveplate that both X and Y laser beams pass through. This provides an additional method, independent of the $\tilde{\mathcal{P}}$ switch, of interchanging the roles of the X and Y beams to remove "asymmetry effects."

Omission of the $\tilde{\mathcal{G}}$ switch

I would like to briefly comment on the absence of one superblock switch that was used in ACME I, namely the $\tilde{\mathcal{G}}$ switch. In ACME I, this switch rotated both the preparation and readout polarization bases by $\frac{\pi}{2}$. As discussed in 2.3.5, this should – and did – have no effect on measured quantities.

In ACME II, this switch cannot be straightforwardly implemented because STIRAP can only prepare an angular momentum alignment approximately along $\hat{y}$. Therefore, rotating the preparation (refinement) laser polarization basis by $\frac{\pi}{2}$ would completely eliminate our experimental signal. One alternative method of implementing an analogue of the $\tilde{\mathcal{G}}$ switch is to drive STIRAP through $\tilde{\mathcal{P}} = +1$ instead of $\tilde{\mathcal{P}} = -1$ in order to prepare the angular momentum state aligned with $\hat{x}$, and simultaneously rotate the refinement laser polarization

by $\frac{\pi}{2}$. We ultimately rejected this approach because it would significantly increase the experimental complexity and require laser powers beyond available levels (since the ability to address the opposite $\tilde{\mathcal{P}}$ state in the preparation would require additional AOMs, and corresponding laser power losses). Because we have no model in which the $\tilde{\mathcal{G}}$ switch would have suppressed systematic errors, and it was observed to be extraneous in ACME I, we are confident that omitting this switch does not pose any challenge to the validity of our results.

Systematic error checks

On slower timescales, during both "normal" EDM data and during our campaign to search for systematic errors (prior to the published EDM dataset), we modify any one or several of a variety of experimental parameters, including the magnitude of the electric or magnetic fields, laser detunings, laser powers, transverse magnetic fields $\mathcal{B}_{x,y}$, magnetic field gradients $\partial_i \mathcal{B}_j$, etc. In keeping with the nomenclature for faster switches, we sometimes call these "uberblock" switches. The full list, and effects, of parameter adjustments will be considered in detail in Ch. 4.

2.6 Rotational cooling in detail

This section will consider the theory and implementation of rotational cooling in much greater detail than Sec. 2.3.2. Preliminary considerations of the rotational cooling scheme for ACME II were explored by Brendon O'Leary [166].

As mentioned previously, we use optical pumping to consolidate molecules in a single quantum level, $X(J = 0)$, before transferring them to the $H^3\Delta_1$ science state. It is not necessary to vibrationally cool, so all references to X-state molecules refer to $X(v = 0)$. Rotational cooling[16] is achieved by optically pumping through the electronic C state. This section describes our considerations when selecting an optical pumping scheme and provides

---

16. The prepared states do not follow a thermal distribution, so there is not a well-defined rotational "temperature" achieved by rotational cooling. However, there is a reduction in the entropy associated with the internal state distribution, which is physically allowed due to the dissipative effect of spontaneous emission.

the methods of calculation used to estimate the ideal efficiency of transfer to the rotational ground state. The observed efficiency is in reasonably good agreement with the calculations, but slightly lower as generically expected.

This section is structured as follows: first, I compute the distribution of molecules among J states within the $X(v = 0)$ manifold assuming thermal equilibrium. Then, I estimate the branching ratio for decays $C(v = 0) \rightsquigarrow X(v = 0)$ based on the Franck-Condon factor and estimated electronic dipole matrix elements. I compute branching ratios for different transitions $J \rightsquigarrow J'$. (In this section, primes always denote ground states. Unprimed states typically refer to excited states, but may also refer to ground states when the electronic state is unambiguous). Two rotational cooling schemes are then described and I estimate the efficiency of the cooling, defined as the total fraction of $X(v = 0)$ molecules ending up in $J = 0$.

Rotational cooling to higher rotational levels (e.g., $J = 1$) would generally be less efficient because we can only extract population from a single magnetic sublevel using STIRAP $X \rightarrow C \rightarrow H$. Due to geometric constraints, we cannot apply a well-defined quantization axis (i.e., external electric or magnetic field) throughout the region between rotational cooling and the interaction region, so we expect the magnetic sublevels to remix while the molecules travel from the rotational cooling region into the interaction region. As a result, the rotational cooling efficiency cannot possibly exceed $1/(2J + 1)$ for cooling to rotational level J. Since the cooling to $J = 0$ is estimated to result in efficiencies greater than $1/3 \geq 1/(2J + 1)$ for $J \geq 1$, it is not worthwile considering cooling to any other rotational levels.

2.6.1 Initial population distribution

The initial rotational population is assumed to be thermally distributed at temperature T. The rotational energy of a state with angular momentum J is given by $J(J + 1)B_R$, where $B_R = 0.33264 \, \text{cm}^{-1}$ in the X state of ThO [160]. The degeneracy of the J manifold is $(2J + 1)$. The normalized initial population in level J is therefore

J	$P_J^0(\%)$
0	12
1	27
2	28
3	19
4+	14

Table 2.1: Initial rotational distribution for $T = 4$ K

$$P_J^0 = \frac{1}{Z}(2J+1)\exp\left[-J(J+1)\frac{B_R}{k_B T}\right], \tag{2.22}$$

where the normalization $\sum_J P_J^0 = 1$ is enforced by the partition function $Z = \sum_J (2J + 1)\exp\left[-J(J+1)\frac{B_R}{k_B T}\right]$. An example population distribution is shown in Table 2.1 at $T = 4$ K. Internal measurements [167, 168] and published data [152] are consistent with rotational temperatures in the range of 3-6 K, with somewhat more evidence toward the lower end of the range (3-4 K). In this document, I will take $T = 3 - 4.5$ K to be a reasonable working range.

2.6.2 Branching ratios

The following approach to computing branching ratios is described clearly in [169, Sec. 6.5]. In the Born-Oppenheimer approximation ([170], English translation [171]), we approximate a total energy eigenstate of a molecule by the product of electronic, vibrational, and rotational eigenstates, $|\Psi\rangle = |\Psi\rangle_E |\Psi\rangle_V |\Psi\rangle_R$. The transition strength between two states $|\Psi\rangle$ and $|\phi\rangle$ is, to leading order, proportional to the square of the dipole matrix element $\langle\phi|\mu|\Psi\rangle$. Assuming the transition under consideration is between distinct electronic energy levels, the nuclear dipole moment matrix element vanishes and $\langle\phi|\mu|\Psi\rangle = \langle\phi|\mu_e|\Psi\rangle$, where μ_e is the electronic dipole moment $q_e r_e$. Working in this special case, we simply write μ in place of μ_e for everything that follows.

The unnormalized transition strength can then be factored into

$$|_E\langle\phi|\mu|\Psi\rangle_E|^2 \times |_V\langle\phi|\Psi\rangle_V|^2 \times S, \tag{2.23}$$

where $|_\text{V}\langle\phi|\Psi\rangle_\text{V}|^2$ is a Franck-Condon factor and S is a "rotational line strength" or "Hönl-London factor."

Electronic branching

To determine the efficiency of rotational cooling, we need to know the branching ratio from $C(v = 0) \rightsquigarrow X(v = 0)$. We can first crudely estimate the branching ratio for $C(v = 0) \rightsquigarrow X$ and then find the proportion that land in the $v = 0$ ground-state manifold using Franck-Condon factors.

The C state has a higher energy than the X, H, Q, A, and B states. We first expand these states in terms of a series of molecular term symbols, corresponding to a Hund's case (a) basis [159].

$$
\begin{aligned}
C &= 0.77\,{}^1\Pi_1 &+& 0.20\,{}^3\Pi_1 &+& 0.02\,{}^3\Delta_1 \\
X &= {}^1\Sigma_0^+ \\
H &= 0.98\,{}^3\Delta_1 &+& 0.01\,{}^3\Pi_1 &+& 0.01\,{}^1\Pi_1 \\
Q &= 0.94\,{}^3\Delta_2 &+& 0.04\,{}^1\Delta_2 &+& 0.02\,{}^3\Pi_2 \\
A &= 0.95\,{}^3\Pi_0 &+& 0.05\,{}^1\Sigma_0^+ \\
B &= 0.77\,{}^3\Pi_1 &+& 0.18\,{}^1\Pi_1 &+& 0.05\,{}^3\Sigma_1
\end{aligned}
$$

We must remember to take the square root of these coefficients when expanding molecule states, e.g., $|C\rangle = \sqrt{0.77}|^1\Pi_1\rangle + \sqrt{0.20}|^3\Pi_1\rangle + \sqrt{0.02}|^3\Delta_1\rangle$, to ensure that the states are properly normalized.[17] Relative phases between terms are assumed to be unimportant. We are not able to precisely calculate the electric dipole matrix elements between the C state and final electronic states f, so we assume that all allowed transitions have comparable matrix elements. Transitions are forbidden when they would (1) change the total spin (e.g., $S = 0 \leftrightarrow S = 1$); (2) change the spin projection (e.g., $\Sigma = 0 \leftrightarrow \Sigma = 1$); (3) or violate angular momentum conservation (e.g., when $\Lambda = 0^+ \leftrightarrow \Lambda = 0^-$ or $|\Delta\Lambda| > 1$, and likewise for Σ, Ω). A table of forbidden transitions from the components of the C state is given in Table 2.2.

---

17. Beware that this point is neglected in [166], leading to a discrepancy with the results below.

	$^1\Pi_1$	$^3\Pi_1$	$^3\Delta_1$
$^1\Sigma_0^+$	✓	1	12
$^3\Delta_1$	12	2	✓
$^3\Pi_1$	1	✓	2
$^1\Pi_1$	✓	1	12
$^3\Delta_2$	1	✓	2
$^1\Delta_2$	✓	1	12
$^3\Pi_2$	12	2	3
$^3\Pi_0$	12	2	✓
$^3\Sigma_1$	12	2	3

Table 2.2: Forbidden transitions are marked by the selection rule(s) they violate. Allowed transitions are indicated by ✓.

We then estimate the electronic overlap using the composition of each state and the allowed transitions. As an example, $C \rightsquigarrow Q$ has allowed contributions

$$\sqrt{(0.77)(0.04)} \left[^1\Pi_1 \rightsquigarrow {}^1\Delta_2 \right] + \sqrt{(0.20)(0.94)} \left[^3\Pi_1 \rightsquigarrow {}^3\Delta_2 \right] = 0.61. \qquad (2.24)$$

Repeating this calculation for each possible decay from C, we can obtain the overlap $|\langle f|C\rangle|^2$ for each state f. The decay rates are proportional to these overlap factors (up to slight differences in the transition dipole moments, which we don't know) and to ω^3, where ω is the angular frequency of the transition (e.g., see [172, Sec. 3.3]). The relative probability of a branching $C \rightsquigarrow f$ is then approximately $\omega_f^3 |\langle f|C\rangle|^2 / \sum_k \left[\omega_k^3 |\langle k|C\rangle|^2 \right]$, with the denominator for normalization. These branching ratios are computed in Table 2.3 and are consistent with the values computed in [148]. Transition frequencies can be found in [160].

A crude lower-bound estimate of the branching ratio to a given state is computed assuming that the dipole operator for that transition is 25% smaller than in the "default" case (where all allowed moments are assumed equal) and that all other dipole operators are 25% larger. Likewise, an upper-bound estimate is computed assuming that the dipole operator for that transition is 25% larger than in the default case while other dipole operators are 25% smaller. These estimates clearly do not constitute hard bounds. In fact, as we'll see, the efficiency of rotational cooling is consistent with the lowest end of this branching ratio,

$\lvert f \rangle$	$\lvert \langle f \vert C \rangle \rvert^2$	ω (cm^{-1})	Computed branching (%)	Range (%)
X	0.77	14490	88.0	73-95
H	0.09	9174	2.7	1-7
Q	0.37	8362	8.2	3-20
A	0.13	3890	0.3	0.1-0.8
B	0.58	3361	0.8	0.3-2.3

Table 2.3: Electronic branching ratios from C

and we can infer that larger uncertainties may be warranted.[18]

Vibrational branching

The vibrational contribution to branching ratios is given by Franck-Condon factors, which express the degree to which the nuclear wave functions overlap between two states. Calculations of Franck-Condon factors can be performed by treating nuclear wave functions in the harmonic oscillator approximation provided rotational and vibrational constants are known. However, at present we only need to know the vibrational branching $C(v = 0) \rightsquigarrow X(v = 0)$ among all $C(v = 0) \rightsquigarrow X$ decays, which can be calculated from well-known molecular constants to be 84%, based on a simple Morse potential model [173]. We therefore expect the total branching ratio $C(v = 0) \rightsquigarrow X(v = 0)$ to be in the range 61-80%, with a best guess at 74%.

Note that the Steimle group has measured the branching ratio for $I(v = 0) \rightsquigarrow X(v = 0)$ to be 91%, which is better than the best possible branching ratio for $C(v = 0) \rightsquigarrow X(v = 0)$ [174]. Since the I state also has $\lvert \Omega \rvert = 1$, its rotational structure is identical to that of C. Therefore, the scheme discussed here could be applied to the I state by changing the laser wavelengths. As discussed in Sec. 2.6.5, this is a compelling option for future experiments, but we are unlikely to pursue it due to the increased effort required and merely marginal improvement over rotational cooling through C.

18. The full story of these branching ratios is, unfortunately, confusing. Earlier optical pumping measurements by Emil Kirilov and Cristian Panda appeared to be consistent with an electronic branching ratio closer to $\sim 85\%$, while very careful direct measurements of the branching ratio by Daniel Ang are around 65% (assuming the previously measured Franck Condon factor of 84%, discussed below). My take-away is that the direct branching ratio measurements, consistent with our ACME II rotational cooling data, probably give the most reliable value in the range of $\sim 65 - 75\%$.

Zero-field rotational branching ratios

Like atoms, molecules primarily decay through E1 transitions for which the effective interaction Hamiltonian is $\vec{\mathcal{E}} \cdot \vec{\mu}$, where $\vec{\mathcal{E}}$ is an electric field and $\vec{\mu}$ is a transition dipole moment. In spontaneous decay, the effective electric field $\vec{\mathcal{E}}$ arises due to vacuum fluctuations, and $\vec{\mu}$ is a function of the electronic state but doesn't depend on the molecular rotation.

We have already considered such matrix elements in Sec. 2.1.7. The rotational line strength S_M for decay from a particular excited sublevel $|J, \Omega, M\rangle$ is given by

$$
\begin{aligned}
S_M \quad &\propto \quad \sum_{M'} |\langle J, M, \Omega; \Lambda, S, \Sigma | \vec{\mathcal{E}} \cdot \vec{\mu} | J', M', \Omega'; \Lambda', S', \Sigma' \rangle|^2 \\
&\propto \quad (2J+1)(2J'+1) \begin{pmatrix} J & 1 & J' \\ -\Omega & (\Omega - \Omega') & \Omega' \end{pmatrix}^2 \\
&\qquad \times \sum_{M'} \begin{pmatrix} J & 1 & J' \\ -M & (M - M') & M' \end{pmatrix}^2,
\end{aligned}
\tag{2.25}
$$

where I retained only factors that depend on rotational quantities. The remaining sum reduces to $1/(2J+1)$[19]. This leads to

$$
S_M \equiv (2J'+1) \begin{pmatrix} J & 1 & J' \\ -\Omega & (\Omega - \Omega') & \Omega' \end{pmatrix}^2,
\tag{2.26}
$$

where I am now defining S_M such that the constant of proportionality in Eq. 2.25 is unity. This result doesn't depend on M, consistent with the rotational symmetry of spontaneous decay. More conventionally, rotational branching ratios are given by Hönl-London factors, defined as $S \equiv \sum_M S_M = (2J+1)S_M$. I give Hönl-London factors normalized over all decays in Table 2.4.

---

19. An internal ACME document, [175], contains the detailed calculation.

Final	$J = 1$	$J = 2$
Lost	$1 - \ell$	$1 - \ell$
$J' = 0$	$(2/3)\eta_0 \ell$	
$J' = 1$	$(1 - \eta_0)\ell$	$(3/5)\ell$
$J' = 2$	$(1/3)\eta_0 \ell$	
$J' = 3$		$(2/5)\ell$

Table 2.4: Normalized Hönl-London factors used in rotational cooling. Primed rotational levels are in the ground state. The value of η_0 depends on the parity addressed in the excited state and on the electric field (see Sec. 2.6.2). The parameter ℓ describes loss to other vibrational or electronic states (see Sec. 2.6.3). It is assumed that decays from $J = 2$ occur from the even-parity state in the absence of an electric field.

Rotational branching in an applied electric field

An applied electric field mixes parity states, which modifies the Hönl-London factors. However, in the absence of any external fields, the Hamiltonian of the molecule commutes with the parity operator, and so the eigenstates of the Hamiltonian have fixed parity. We will consider the effect of an external electric field shortly, but first consider a case where the initial manifold involves superpositions of states with definite parity,

$$|J\,\Omega\,M\rangle = a_{M,+}|J,\,\Omega,\,M,\,+\rangle + a_{M,-}|J,\,\Omega,\,M,\,-\rangle, \tag{2.27}$$

but the final state has definite parity p'. The Hönl-London factor is effectively modified by a factor of $|a_{M,-p'}|^2$ since $E1$ transitions connect only opposite-parity states. The effective Hönl-London factor from a sublevel M is then

$$S_M = (2J' + 1) \begin{pmatrix} J & 1 & J' \\ -\Omega & (\Omega - \Omega') & \Omega' \end{pmatrix}^2 |a_{M,-p'}|^2. \tag{2.28}$$

The $X^1\Sigma_0$ state has opposite-parity states separated in energy only by the rigid-body molecular rotation (~ 20 GHz), so that there is negligible parity mixing. However, the C state, predominantly $^1\Pi_1$, has Ω-doublet states split by 50.4 MHz, which are fully mixed in applied electric fields of $\mathcal{E}_{\mathrm{lab}} \sim 100$ V/cm [145]. Therefore, consider the situation where an electric field may significantly mix parity states in the C manifold but not in the X manifold. The effect of an electric field on parity mixing is described in [145, Sec. 3.1].

65

Given a zero-field energy splitting Δ_Ω, dipole moment matrix element D, electric field $\mathcal{E}$, and parity eigenstates $|e\rangle$ and $|f\rangle$, the perturbed eigenstates can be written as

$$|\tilde{e}\rangle = \cos\tfrac{\theta}{2}\ |e\rangle + \sin\tfrac{\theta}{2}\ |f\rangle$$

$$(2.29)$$

$$|\tilde{f}\rangle = -\sin\tfrac{\theta}{2}\ |e\rangle + \cos\tfrac{\theta}{2}\ |f\rangle$$

where

$$\theta = \arctan(\beta M),$$

$$\beta = -\frac{2D\mathcal{E}|\Omega|}{\Delta_\Omega J(J+1)}.$$

$$(2.30)$$

It follows that the parity components in an electric field have magnitudes

$$|\langle s'|\tilde{s}\rangle|^2 = \frac{1}{2}\left(1 \pm \frac{1}{\sqrt{1+\beta^2 M^2}}\right),$$

$$(2.31)$$

where $s = e, f$ and the positive sign applies with the positive sign applies for the component present when $\mathcal{E} = 0$ (e.g., $s' = e$ when $\tilde{s} = \tilde{e}$) and the negative sign applies otherwise. Inserting this result into the expression for the mixed-parity branching ratio, we see that

$$S_M = \frac{1}{2}(2J'+1)\begin{pmatrix} J & 1 & J' \\ -\Omega & (\Omega - \Omega') & \Omega' \end{pmatrix}^2 \left(1 \pm \frac{1}{\sqrt{1+\beta^2 M^2}}\right).$$

$$(2.32)$$

I define $\phi(\beta, M) \equiv 1/\sqrt{1+\beta^2 M^2}$ and note that $\phi \in [0,1]$. Therefore, the 0-field, parity-allowed Hönl-London factors are modified by a multiplicative factor of $\eta_\pm = (1/2)(1 \pm \phi) \in [0,1]$, where η_- corresponds to decays disallowed in 0 field and η_+ corresponds to decays allowed in 0 field. Note that $\eta_\pm = 1 - \eta_\mp$.

Magnetic branching ratios

To compute the branching ratios from a given excited sublevel to the sublevels in a particular rotational manifold, Clebsch-Gordan coefficients (or equivalently, 3-j symbols) may be used as usual. If we're only interested in the distribution among ground state sublevels that a given excited sublevel M decays to, then we can ignore any effect of parity mixing.
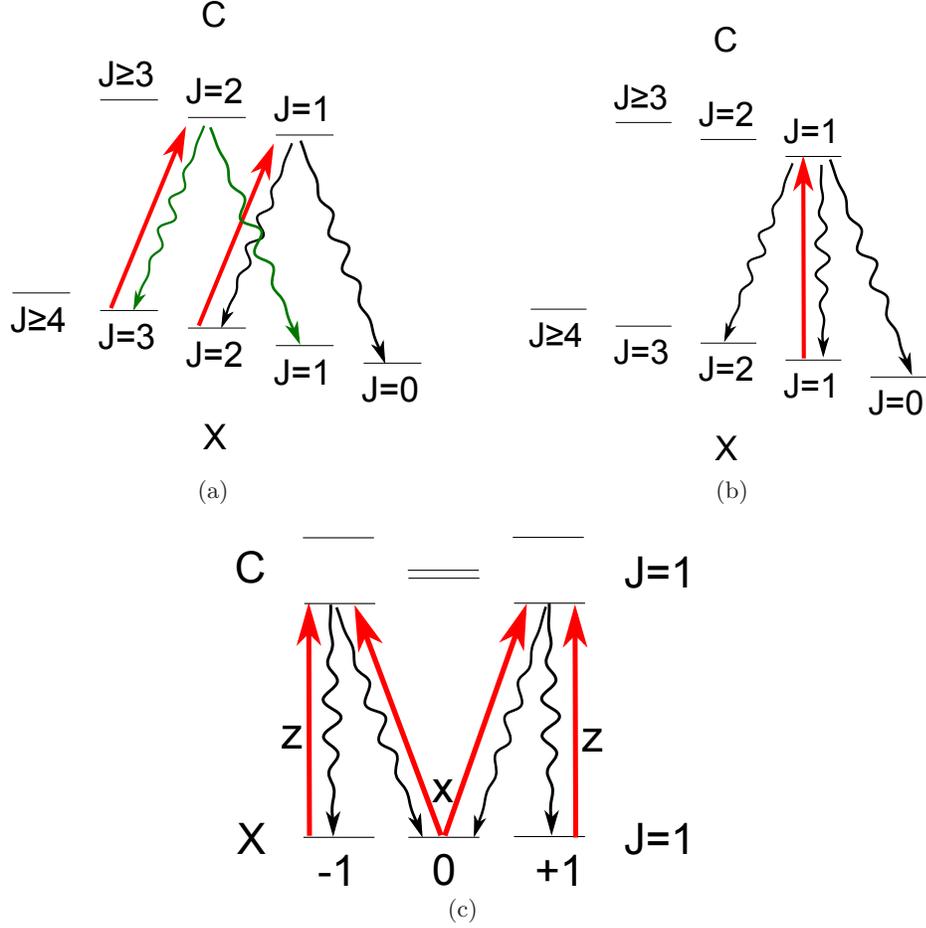
Figure 2.4: Rotational cooling scheme used in ACME II. (a) First step of rotational cooling, with $\mathcal{E} = 0$. Transfer $J' = 2, 3$ to $J' = 0, 1$. (b) Second step of rotational cooling, with $\mathcal{E} > 0$. Transfer $J' = 1$ to $J' = 0$. Decays to $J' = 2$ also occur. (c) Magnetic sublevels in second step of rotational cooling, with $M = \pm 1$ and $M = 0$ addressed by distinct laser polarizations ($\hat{z}$ and $\hat{x}$, respectively).

2.6.3 Rotational cooling scheme

Step 1: Pumping to $J' = 0, 1$

A diagram of both rotational cooling steps is shown in Fig. 2.4. In the first step of the rotational cooling scheme, shown in panel (a), we pump $X \to C$ using $J' = 3 \to J = 2$ and $J' = 2 \to J = 1$ with no applied electric field. The parity of the X state is $p' = (-1)^{J'}$, while there exists a parity doublet in the C manifold. By parity selection rules, the excited states addressed by the lasers have parity $p = (-1)^J$. Therefore, decays $J = 2 \rightsquigarrow J' = 2$ and $J = 1 \rightsquigarrow J' = 1$ are forbidden. As a result, the only allowed decays to the X state are

$J = 2 \rightsquigarrow J' = 1, 3$ and $J = 1 \rightsquigarrow J' = 0, 2$. Molecules that land back in $J' = 2, 3$ are excited again. Let the fraction decaying from the excited state back to the $X(v = 0)$ manifold be ℓ. We expect $\ell \approx 0.75$ as described in Sec. 2.6.2. The branching ratios are then given in Table 2.4 with $\eta_0 = 1$.

Pumping with a given polarization will result in dark states. For example, any superposition of magnetic sublevels $|M' = \pm J'\rangle$ is dark when pumping $J' \rightarrow J = J' - 1$ with $\hat{z}$-polarized light. However, these states are not dark when pumping with $\hat{x}$-polarized light.[20] To eliminate all dark states, we switch between $\hat{x}$- and $\hat{y}$-polarized pumping light in the first step (note that the lasers travel along $\hat{z}$).

Any given molecule that is not lost during the pumping process is only transferred among three rotational states (namely, $J' = 1, 3$ and $J = 2$, or $J' = 0, 2$ and $J = 1$). Let the probability of decay to the initial state be p_i and the decay to the target state be p_t. Note that these sum to the branching fraction $\ell < 1$ for $C(v = 0) \rightarrow X(v = 0)$. Then the probability of transfer to the target state is

$$P(i \rightarrow t) = p_t + p_i \cdot p_t + p_i^2 \cdot p_t + \cdots, \tag{2.33}$$

where the n-th term represents an event where the molecule is transferred after the n-th excitation from i. This can be computed as a geometric series,

$$P(i \rightarrow t) = p_t \sum_{n=0}^{\infty} p_i^n = \frac{p_t}{1 - p_i}. \tag{2.34}$$

Referring to Table 2.4 for branching ratios p_i and p_t, we can compute the transfer efficiencies for $J' = 3 \rightarrow J' = 1$ and $J' = 2 \rightarrow J' = 0$. Recall that the initial population in rotational level J' is denoted by $P_{J'}^0$; in general, we will label the population after step n by $P_{J'}^n$. Then the populations in $J' = 0, 1$ after step 1 are given by

---

20. One can always compute the matrix elements directly, but an alternative way to see this is to note that an arbitrary dark state of $\hat{z}$-polarized light is $(a|+J\rangle + b|-J\rangle)$. However, $J_x(a|+J\rangle + b|-J\rangle)$ is some superposition of $|J - 1\rangle$ and $|-J + 1\rangle$, and the subspace of dark states of $\hat{z}$-polarized light is not closed under J_x. By rotational symmetry, the dark states of $\hat{x}$-polarized light should form a closed subspace under J_x just as the dark states of $\hat{z}$-polarized light form a closed subspace under J_z. Therefore, the dark states of $\hat{z}$-polarized light can't also be dark states of $\hat{x}$-polarized light. Likewise, $\hat{x}$- and $\hat{y}$-polarized light can't give rise to the same dark states.

$$P^1_{J'=0} = P^0_{J'=0} + \frac{(2/3)\ell}{1-(1/3)\ell} \; P^0_{J'=2}$$

$$\text{(2.35)}$$

$$P^1_{J'=1} = P^0_{J'=1} + \frac{(3/5)\ell}{1-(2/5)\ell} \; P^0_{J'=3}$$

Step 2: Empty $J' = 1$

At this point, we would ideally like to transfer all molecules from $J' = 1$ to $J' = 0$. Since C is an $|\Omega| = 1$ state, there is no $C(J = 0)$ level. Therefore, we must pump through $J = 1$. Since $J' = 1, 0$ have opposite parities, we must apply an electric field to mix the parity of the $J = 1$ excited state in any two-photon transfer process.

Further, assume that the $M = 0$ sublevel is spectroscopically resolved from the $M = \pm 1$ levels due to Stark splitting. We address the $M = \pm 1$ states[21]. Due to the spectroscopic resolution, the ground $M' = \pm 1$ states are dark to both $\hat{x}$- and $\hat{y}$-polarized light (which only couple to $M = 0$, ~ 100 MHz off resonance), so we switch between $\hat{x}$- and $\hat{z}$-polarized light. These have dark states $M' = \pm 1$ and $M' = 0$, respectively. This step is shown schematically in Fig. 2.4(b-c).

Recall that in the presence of an electric field, the $M' = \pm 1$ excited state parity doublets mix and Stark shift in opposite directions. For now, we're free to leave the addressed excited state manifold, with either a positive or negative Stark shift, unspecified. We let η_0 denote whichever of $\eta_\pm(J = 1, J' = 0, \beta)$ is appropriate for the chosen excited state that is spectroscopically addressed by the $J' = 1 \leftrightarrow J = 1$ laser. The normalized branching ratios for decay from $J = 1$ are then given in Table 2.4.

From this point, we can use the same geometric series trick as in Sec. 2.6.3 to determine the final populations. The result is

$$P^2_{J'=0} = P^1_{J'=0} + \frac{(2/3)\eta_0\ell}{1-(1-\eta_0)\ell} \; P^1_{J'=1}$$

$$\text{(2.36)}$$

$$P^2_{J'=2} = \frac{(1/3)\eta_0\ell}{1-(1-\eta_0)\ell} \; P^1_{J'=1}$$

---

21. For convenience, here only we will take the quantization axis to be along the applied electric field in the rotational cooling region (approximately vertical) even though it does not coincide with the laboratory axis convention used elsewhere.
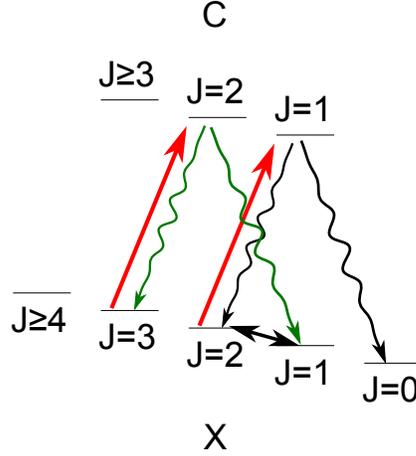
Figure 2.5: Alternative rotational cooling scheme, in a single field-free region, using microwaves.

We consider two figures of merit for the rotational cooling scheme. The efficiency E of the process is just the relative population in $J = 0$ at the end of step 2, and represents how much more efficiently we could hypothetically perform rotational cooling. The gain G, which is of more direct experimental interest, is the ratio of this population to the initial population in $J = 0$:

$$E = P^2_{J'=0}, \tag{2.37}$$

$$G = P^2_{J'=0}/P^0_{J'=0}. \tag{2.38}$$

To compute η with a given electric field, we find $\beta(\mathcal{E}) = -\frac{2D\mathcal{E}|\Omega|}{\Delta_\Omega J(J+1)}$. In the C state, $\Delta_\Omega = 2\pi \times 51$ MHz [154] and $D = 1.00\, ea_0$ [148, Sec. 6.5]. Thus in $J = 1$, $D_{C,1} = 0.63\, \text{MHz}/(\text{V/cm})$ and $\mathcal{E}_C = 51\, \text{MHz}/(0.63\, \text{MHz}/(\text{V/cm})) = 81\, \text{V/cm}$, where $\mathcal{E}_C$ is defined such that $\beta(\mathcal{E}_C) = 1$. In order to "fully mix" the parity in the C state, we need to apply an electric field on the order of $2\mathcal{E}_C \sim 150$ V/cm.

Rejected alternative: optical pumping with microwaves

As an alternative optical pumping scheme to populate $J' = 0$, we considered mixing the populations in $J' = 1, 2$ by using microwaves. If the $J' = 2 \to J = 1$ laser depopulates

$J' = 2$ in the ground state, then $J' = 1$ must be depopulated as well. See Fig. 2.5.

In computing the efficiency of this process, we may consider the $J' = 3 \to J = 2 \rightsquigarrow J' = 1$ process to occur before the microwave remixing. Then the population in $J' = 1$ before remixing is simply $P^1_{J'=1}$ as computed before. The optical pumping $J' = 2 \to J = 1 \rightsquigarrow J' = 0$ proceeds as usual with the substitution $P^0_{J'=2} \to P^0_{J'=2} + P^1_{J'=1}$. Making these substitutions,

$$E = P^0_{J'=0} + \frac{(2/3)\ell}{1 - (1/3)\ell}(P^0_{J'=2} + P^1_{J'=1}) \tag{2.39}$$

$$G = E/P^0_{J'=0} \tag{2.40}$$

This is more efficient than the first scheme considered because the branching ratio for $J = 1 \rightsquigarrow J = 0$ is more favorable in the absence of an electric field and because any molecules that land in $J' = 2$ during "Step 2" of the previous scheme are simply lost, while in the alternative scheme using microwaves they can be excited by the laser again. A minor inconvenience is that this alternative requires twice as much power on the $J' = 2 \to J = 1$ laser because each molecule is in the addressed ground state for only half the time (in the case that the microwaves saturate the $J' = 2 \leftrightarrow J' = 1$ transition).

More importantly, there is an off-resonant excitation of the $H(J = 1) \leftrightarrow H(J = 2)$ transition since the rotational constants in each electronic level are simliar: $B_H = 0.326\,\mathrm{cm}^{-1}$, compared to $B_X = 0.333\,\mathrm{cm}^{-1}$ [160]. This transition is thus off-resonant by $\Delta \sim 2\pi \times 1$ GHz in the H state if X is driven on resonance. This will not drive any significant population transfer, but it could lead to AC Stark shifts in the $H(J = 1)$ state, and thus systematic errors.

In more detail: in a two-level system driven with detuning Δ and resonant Rabi frequency Ω, the energy shift in the ground state is $\delta\omega = \frac{1}{2}(\Delta - \sqrt{\Delta^2 + \Omega^2})$, as can be found by diagonalizing the Hamiltonian in the rotating frame with the rotating wave approximation (see [149, Sec. 4.3] for a careful treatment of AC Stark shifts in the ACME experiment). If, e.g., $\Omega = 2\pi \times 10$ MHz, then the AC Stark shift is of order $\delta\omega \sim 2\pi \times 50$ kHz. To leading order, the AC Stark shift is quadratic in the Rabi frequency. In certain configurations, AC

Quantity	Meaning	Likely range	Best guess	Observed
T	Initial rotational temperature	3-4.5 [K]	3.75 K	-
ℓ	Branching ratio for $C(v=0) \rightsquigarrow X(v=0)$	0.6-0.8	0.74	-
η_0	Multiplicative factor for Hönl-London factor from parity selection rules. We set this with the electric field $\mathcal{E}$.	0-1	0.5	-
E	Efficiency (proportion of all molecules ending up in $J' = 0$)	34-56%	47%	-
G	Gain (ratio of number of $J' = 0$ molecules after cooling to number before cooling)	2.7-4.6	3.8	2.7

Table 2.5: Definition of quantities and the figures of merit expected based on the model considered here. The gain expected from the microwave scheme is $\approx 20\%$ higher than in the ACME II scheme, while using the I excited state is anticipated to improve $J' = 0$ yield by $\approx 40\%$.

Stark shifts on the order of a few kHz could be problematic, so the effect of these shifts must be better understood if we use this rotational cooling scheme in the future. We expect that using the rotational cooling scheme with microwaves would improve the $J = 0$ population yield by approximately 20%, which at present is not enough of an improvement to overcome concerns about systematic errors.

2.6.4 Anticipated and measured rotational cooling gain

Based on the models in previous subsections, I compute the efficiency of the laser-based and microwave-based schemes separately. I also compare these to the case in which we use the I excited state, where $\ell = 0.91$ compared to the C state where $\ell \approx 0.75$. Because T and ℓ are not known precisely, I indicate a range for E and G based on the reasonable parameter intervals $T \in (3, 4.5\,\mathrm{K})$, $\ell \in (0.6, 0.8)$, and $\eta_0 \in (0.1, 0.9)$. A summary of various quantities is given in Table 2.5 for reference. For detailed plots of all parameter space, and with gains from each optical pumping laser given separately, see [175] (internal ACME document).

Comparison to measurements

In the following discussion, I refer to the $J' = 3 \to J = 2$ laser, the $J' = 2 \to J = 1$ laser, and the $J' = 1 \to J = 1$ laser as $J3$, $J2$, and $J1$, respectively. We have measured individual increases to the $J = 0$ population to be, conservatively, $\sim 70\%$ from $J1$, $\sim 80\%$ from $J2$, and $\sim 20\%$ from $J3$ (with the $J1$ laser also on), for an overall gain of ~ 2.7. I have quoted gains for each laser relative to the no-cooling case. If we instead take the configuration with the $J1$ and $J2$ lasers as our baseline, then the $J3$ laser contributes only an additional $\sim 8\%$ to the signal. For this reason, we typically leave the $J3$ laser off during data collection, in order to make the system more robust overall (as there is one fewer laser to keep locked).

These results are consistent with the most pessimistic scenario in which $\ell \approx 0.6$ and $T \approx 3$ K. Given that there are some imperfections not included in the model (e.g., less than perfect saturation or alignment), it is not too surprising that the overall gain is toward the lower range of our predictions.

There is little effect of the applied electric field on the efficiency of the $J1$ pumping, over the range $50 - 200$ V/cm. It is more efficient to address the positively Stark-shifted excited state than the negatively Stark-shifted excited state due to a more advantageous parity admixture.

We also observe less than 10% difference in $J2$ gain between configurations where the laser passes are quite close together and widely spaced. The closely-spaced configuration is slightly better, consistent with the fact that the tight spacing allows more laser passes and therefore higher total power seen by the molecules. This shows that adiabatic following of dark states with closely-spaced laser passes is not a major problem in our operating regime.

The polarization switching is implemented by double-passing the laser light through a quarter-wave plate, which rotates the linear polarization by $90°$ between each subsequent pass of the laser through the vacuum chamber. However, if the fast axis of this wave plate is aligned with the initial linear polarization, then no polarization switching should occur. Therefore, we can check the effect of polarization switching on the rotational cooling gain by rotating the wave plate axis relative to the initial laser polarization. Contrary to our model, we do not see any effect of adjusting the quarter-wave plate angle on the $J1$

transition; in other words, direct polarization switching has no effect on the optical pumping efficiency. We believe that fringing electric fields between the field plates (whose design was not optimized at all for homogeneous fields) create a rotating quantization axis, which has a similar effect as an alternating polarization between each laser pass. The effect of the quarter-wave plate angle was as expected for the $J2$ transition, where any residual electric fields are negligible.

All three lasers appear reasonably well saturated with ~ 10 mW, consistent with expectations.

2.6.5   Future considerations

The general approach to rotational cooling is likely to be unchanged in any future ACME measurements. For a time, we planned to use an electrostatic lens, which "focuses" the molecules as they traverse the beam line via a radially confining harmonic Stark potential. In the original plan, this would require population in the $X(J = 2, M = 2)$ state. We have demonstrated efficient $X \to C \to X$ STIRAP previously in the ACME experiment (first by Emil Kirilov and subsequently by Cris Panda). It was therefore determined that the optimal state preparation for an electrostatic lens in the $X(J = 2, M = 2)$ state would involve rotational cooling to $J = 0$, just as we do in the ACME II measurement, followed by STIRAP to the focusing state, rather than direct optical pumping into $J = 2$.

We now expect that, if an electrostatic or magnetostatic lens is used, any future lensing state will be in the $Q^3\Delta_2(J = 2, M = 2)$ state, which is even less favorable for optical pumping state preparation than $X(J = 2, M = 2)$. Therefore, optical pumping into $J = 0$ (using the same methods as in ACME II) followed by STIRAP into the focusing state is likely to be the only viable state preparation procedure.

As already mentioned, the I state has a more favorable branching ratio for optical pumping ($\approx 91\%$ vs. $\approx 75\%$ branching to $X(v = 0)$), but a less convenient wavelength for home-made lasers (512 nm vs. 690 nm). Daniel Ang has set up a 512 nm laser from Toptica, demonstrating that rotational cooling via the I state should be feasible. Preliminary measurements by Daniel and Cristian Panda validate the expectation that optical pumping through I leads to slightly more efficient preparation of $X(J = 0)$. However, Daniel and

Cole Meisenhelder have already set up more stable and robust 690 nm lasers (Toptica DL Pro vs. the homemade lasers used in ACME II) for future rotational cooling, making the purchase of several more Toptica DL Pro lasers for the $X \to I$ transition less appealing, considering the relatively modest ($\approx 30 - 40\%$) improvement in expected signal.

Finally, for more technical details of the rotational cooling power requirements and alignment, see Appendix B.

Chapter 3

Data analysis

It is said the number is the number of birds

that can nest in an ancient tibrol tree,

less three grams of honest work,

but Vivec in his later years found a better one

and so gave this secret to his people.

The Thirty-Six Lessons of Vivec

3.1 Schematic view of the analysis routine

We are ultimately interested in measuring precession frequencies, and in particular the contribution to precession frequencies arising from the EDM interaction, but these are encoded somewhat indirectly in the raw data. I will go over our general analysis approach quickly in this section, relying on the conventions used in [129]. Note that certain choices in the analysis might vary in other versions.[1] In Sec. 3.2, I will revisit the data analysis procedure with an emphasis on the statistical assumptions involved.

---

1. In ACME II, we had two "independent" analyses–mine and Cristian Panda's–and two "satellite analysis checks"–Jonathan Haefner's and Daniel Ang's. The latter two were developed with substantial cross-checking with the former two, which introduces, in principle, the possibility of correlated errors. Cris's and my codes were written with as little communication about analysis choices as possible, to ensure that the likelihood of replicating any mistake is minimized.

Average degenerate traces

In the first step of the analysis, I typically gather "degenerate" traces, which are all of the signals acquired with the same experimental switch settings within a block, and average them together, reducing the effective number of traces per block from 64 to 16, each of which involves data taken with a different configuration of block-level switches ($\tilde{\mathcal{N}}$, $\tilde{\mathcal{E}}$, $\tilde{\theta}$, $\tilde{\mathcal{B}}$).

We found that a small proportion of single traces (less than one in 10^4) was triggered incorrectly such that nominal precession phases for one trace in a block can represent large outliers from the other three, nominally degenerate, traces in the same experimental state within a block. My analysis routine checks for such an event and, when a single trace appears to be a significant outlier, removes such traces and averages the remaining traces together.

Bin-level fluorescence and background subtraction

Scattered light and electronic offsets contribute to the PMT signal, so we perform a background subtraction measurement before computing any quantities nominally related to the ThO molecules. In particular, we compute an integrated fluorescence $F_{X,Y}$ for each of the ~ 4000 polarization bins in a trace. Each $F_{X,Y}$ is integrated only over the designated sub-bin, which is defined as an analysis parameter (for an example, see Fig. 4.15). We define the background of X and Y bins separately, so that $B_{X,Y} \equiv \langle F_{X,Y} \rangle_{\mathrm{bg}}$, where the brackets indicate an average over a designated background region that can be varied as an analysis parameter (the first 1-2 ms of the trace). The "signal" part of the fluorescence is then defined to be $S_{X,Y} \equiv F_{X,Y} - B_{X,Y}$.

Individual asymmetries

In order to determine precession frequencies, we first compute precession phases $\phi = \omega\tau$. These are given by the difference in integrated signal between X and Y bins. In particular, the asymmetry is, as defined previously,

$$\mathcal{A} \equiv \frac{S_X - S_Y}{S_X + S_Y} \propto \cos[2(\omega\tau - \delta\theta + \frac{\pi}{4}(\tilde{\mathcal{P}} + 1))], \tag{3.1}$$

where as before, $\delta\theta$ designates the angle between the linear polarization of the X beam and the preparation laser polarization, and $\tilde{\mathcal{P}}$ is the parity of the excited state addressed by the readout lasers. Recall that $\delta\theta = \delta\theta_0 + \Delta\theta\tilde{\theta}$ depends implicitly on the $\tilde{\theta}$ state. The asymmetry is now defined for all ~ 2000 polarization bin pairs in each trace (which, recall, contains data from only a single configuration of the experiment).

Grouped asymmetries

We now "group" asymmetries in each of the 16 states together. If the group size is 20, for example, then we partition the ~ 2000 asymmetry values into ~ 100 sets of 20. In each set, we compute a linear regression, giving an estimate of the mean asymmetry value within the group and the uncertainty in the mean. We now have ~ 100 asymmetry values for each of 16 states within a block. In everything that follows, I only combine data with the same group index (but possibly from different states) unless stated otherwise. As a result, we can examine the data as a function of time after the beginning of a trace ($\approx$time after ablation, up to a constant offset, since the same time after ablation will always correspond to the same group index within a block).

In ACME I, it was only necessary to compute the sample mean and corresponding uncertainty in the mean for data within a group. With our improved signal-to-noise ratio in ACME II, we can resolve the linear slope, arising from velocity dispersion in the molecular beam, among asymmetry values within a group. Therefore, we must measure the mean asymmetry in the group using linear regression to appropriately estimate the uncertainty in the mean. Fitting the data to a line, rather than the mean, inflates the uncertainty in the mean by a negligible factor of $\sqrt{(N_{\mathrm{grp}} - 1)/(N_{\mathrm{grp}} - 2)}$, where $N_{\mathrm{grp}} \sim 20$ is the number of asymmetry values in the fit for each group. It is not practical to instead use smaller group sizes (e.g., $N_{\mathrm{grp}} \sim 2$), within which the local slope of asymmetry values cannot be resolved, because the uncertainty in the mean of all groups would be inflated by the factor of approximately $\sqrt{N_{\mathrm{grp}}/(N_{\mathrm{grp}} - 1)}$. Here, the numerator arises because there are more groups in a given molecular pulse when smaller group sizes are used, while the second factor is proportional to the uncertainty in the mean for a single group.

Asymmetry and contrast

We combine asymmetries in $\tilde{\theta} = \pm 1$ states to produce a $\tilde{\theta}$-averaged asymmetry and contrast:

$$\mathcal{A}(\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}) = \tfrac{s}{2}[\mathcal{A}(\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}, \tilde{\theta} = +1) + \mathcal{A}(\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}, \tilde{\theta} = -1)]$$

$$(3.2)$$

$$\mathcal{C}(\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}) = \tfrac{t}{2}[\mathcal{A}(\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}, \tilde{\theta} = +1) - \mathcal{A}(\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}, \tilde{\theta} = -1)]$$

where $s = \text{sign}(\mathcal{C}) = \tilde{\mathcal{B}}\tilde{\mathcal{P}}\tilde{\mathcal{R}}$ at large values of $\mathcal{B}$ and $s = \tilde{\mathcal{P}}\tilde{\mathcal{R}}$ otherwise; further, $t = -180°/(2\pi\Delta\theta)$, with $\Delta\theta \sim 6°$ giving the full readout beam half-waveplate dither range in degrees (recall Sec. 2.5).[2]

Older references refer to $\mathcal{A}$ defined in this way as the "$\mathcal{B}$-corrected asymmetry" for historical reasons. I refer to it instead by the more general name, "$\mathcal{C}$-corrected asymmetry." The factor of $\text{sign}(\mathcal{C})$ in the definition of $\mathcal{A}(\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}})$ prevents us from accidentally revealing the EDM blind by examining the asymmetry in different superblock states, where $\text{sign}(\mathcal{C})$ can reverse. In particular, building in an explicit factor of $\text{sign}(\mathcal{C})$ guarantees that asymmetry offsets introduced for blinding during analysis are indistinguishable from real precession phases, whose signature in the uncorrected asymmetry (i.e., computed without the factor of $\text{sign}(\mathcal{C})$) would otherwise reverse when the sign of $\mathcal{C}$ reverses. This will become clearer when we discuss the blinding procedure.

I never compute the "uncorrected" asymmetry for any reason, and therefore generally drop the explicit "$\mathcal{C}$-corrected" label outside this section.

Phase

The phase (in each $\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}$ state and group) is

$$\phi = \frac{\mathcal{A}}{2|\mathcal{C}|}. \tag{3.3}$$

This is dominated by Zeeman precession in each state, and therefore reverses sign with $\tilde{\mathcal{B}}$ to leading order. Velocity dispersion in the molecule beam causes a non-trivial time-

---

2. Reminder: The full waveplate dither range is *half* the full polarization dither range, which is in turn the polarization step size $\Delta\theta$.

dependent structure of $\phi(t)$ within the molecule pulse: slower molecules precess longer and have larger phase accumulation.

Note that most references in the group use the definition $\phi = \frac{\mathcal{A}}{2\mathcal{C}}$, which is appropriate for the $\mathcal{C}$-uncorrected asymmetry. Since, as stated previously, I only use the $\mathcal{C}$-corrected asymmetry, the absolute magnitude of contrast in the denominator effectively cancels out the "$\mathcal{C}$-correcting" factor of $\text{sign}(\mathcal{C})$ implicit in the definition of $\mathcal{A}$.

Parity sums

We will want to relate measured phases to energy shifts, $\langle H \rangle = \alpha M$, corresponding to $\phi = \alpha\tau$ where α is some prefactor that may depend on experimental switch states, τ is the precession time, and M is the projection of total angular momentum on the laboratory z-axis. Since the precession time is extracted from the $\tilde{\mathcal{B}}$-odd Zeeman precession phase, we have to change from the state basis $\{\phi(\tilde{\mathcal{N}} = \pm, \tilde{\mathcal{E}} = \pm, \tilde{\mathcal{B}} = \pm)\}$ to the parity basis $\{\phi^{(\tilde{\mathcal{N}})(\tilde{\mathcal{E}})(\tilde{\mathcal{B}})}\}$, where in the parity basis each switch factor in parentheses might or might not appear explicitly. (The exact notation will be explained shortly.) The function of the parity basis is most clearly understood from the definition,

$$\phi(\tilde{\mathcal{N}}, \tilde{\mathcal{E}}, \tilde{\mathcal{B}}) = \phi^{nr} + \phi^{\tilde{\mathcal{N}}}\tilde{\mathcal{N}} + \phi^{\tilde{\mathcal{E}}}\tilde{\mathcal{E}} + \phi^{\tilde{\mathcal{B}}}\tilde{\mathcal{B}} + \phi^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}\tilde{\mathcal{N}}\tilde{\mathcal{E}} + \phi^{\tilde{\mathcal{N}}\tilde{\mathcal{B}}}\tilde{\mathcal{N}}\tilde{\mathcal{B}} + \phi^{\tilde{\mathcal{E}}\tilde{\mathcal{B}}}\tilde{\mathcal{E}}\tilde{\mathcal{B}} + \phi^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}}\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}. \quad (3.4)$$

Therefore, each term $x^{\cdots}$ (referred to as a "parity sum") is the part of quantity x that reverses sign when any of the superscripted switches is changed. Each parity component is isolated by inverting this equation (more precisely, this set of 8 equations). The parity basis notation is unfortunately overloaded: a superscript $\tilde{S}$ indicates that a quantity reverses sign when the factor $\tilde{S}$, representing a state of a switch, changes. Further, we should remember that a particular component such as $\phi^{\tilde{\mathcal{N}}}$ *must not* reverse sign under switches not explicitly listed, so we have to know which unlisted switches are implicitly included in the definition of a particular parity component. Despite these caveats, the parity basis notation is extremely useful when used carefully.

In many cases, it is easiest to think of the measurement values $\phi(\tilde{\mathcal{N}} = \pm, \tilde{\mathcal{E}} = \pm, \tilde{\mathcal{B}} = \pm)$

as components of a vector, $\vec{\phi} = \phi(+++)\hat{e}_{+++} + \cdots + \phi(---)\hat{e}_{---}$, where $\hat{e}_{+++}, \cdots, \hat{e}_{---}$ form an orthonormal basis. We are free to rewrite this vector in any convenient basis via an invertible change-of-basis matrix P. If we want to preserve orthonormality in the new basis, we additionally require that $P^{-1} = P^T$. Once we construct the matrix P, we can almost effortlessly change any quantity (phase, asymmetry, contrast, etc.) back and forth between bases.

After constructing the change-of-basis matrix $P = P_{\text{state}\to\text{parity}}$, I compute $\vec{\phi}_{\text{parity}} = P\vec{\phi}_{\text{state}}$. This allows us to infer the energy shifts contributing to measured phases,

$$\{\langle H \rangle^{(\tilde{\mathcal{N}})(\tilde{\mathcal{E}})(\tilde{\mathcal{B}})}\} = \{\phi^{(\tilde{\mathcal{N}})(\tilde{\mathcal{E}})(\tilde{\mathcal{B}})} M/\tau\}. \tag{3.5}$$

An alternative perspective on the "parity sum," as a special case of ordinary least squares regression, is given in Appendix C.

Precession time

The applied magnetic field causes Zeeman precession due to the Hamiltonian term $\omega^{\tilde{\mathcal{B}}} = -g\mu_B|\mathcal{B}|$, leading to a $\tilde{\mathcal{B}}$-odd phase $\phi^{\tilde{\mathcal{B}}} = -g\mu_B|\mathcal{B}|\tau$. The g-factor in the $J = 1$ manifold of the H state is $g = -0.00440(5)$ [129]. We have calibrated our magnetic field coils (see Sec. 5.2.4) and found that the precession region experiences 1.337(3) mG per mA of applied current. The total precession time is then

$$\tau = \frac{\phi^{\tilde{\mathcal{B}}}}{-g\mu_B|\mathcal{B}|} = \frac{\phi^{\tilde{\mathcal{B}}}}{+|g|\mu_B\frac{d\mathcal{B}}{dI}I}, \tag{3.6}$$

where $\frac{d\mathcal{B}}{dI} = 1.337$ mG/mA is a calibration factor, I is the applied current, and $|g| = 0.00440$.

This formula gives the total Zeeman precession time, but recall that we only measure phases up to multiples of $\pi/4$ in order to maximize the sensitivity to small changes in phase. In particular, we set the preparation and readout bases to have a relative angle of $\delta\theta \approx n\frac{\pi}{4}$ where n is an integer, such that $|\phi_{\text{measured}}| < \frac{\pi}{4}$ in all states. Then

$$\tau = \frac{\phi^{\tilde{\mathcal{B}}}_{\text{measured}} + n\frac{\pi}{4}}{|g|\mu_B\frac{d\mathcal{B}}{dI}I}. \tag{3.7}$$

In the data analysis, we estimate n from the expected Zeeman precession,

$$
\begin{aligned}
\phi^{\tilde{\mathcal{B}}} &= |g|\mu_B \tfrac{d\mathcal{B}}{dI} I\tau \\
&\approx n\tfrac{\pi}{4} \\
\to \quad n &= \left[\frac{|g|\mu_B \frac{d\mathcal{B}}{dI} I\tau_0}{\pi/4}\right],
\end{aligned}
\tag{3.8}
$$

where the brackets denote that we round n to the nearest integer and $\tau_0 = 1$ ms is an estimate of the precession time that is always accurate to much better than the 50% required for this procedure to correctly assign the value of n.

Precession frequencies

We compute precession frequencies in the state basis by

$$
\omega(\tilde{\mathcal{N}}, \tilde{\mathcal{E}}, \tilde{\mathcal{B}}) = \frac{\phi(\tilde{\mathcal{N}}, \tilde{\mathcal{E}}, \tilde{\mathcal{B}})}{\tau}.
\tag{3.9}
$$

Since $\vec{\omega}_{\text{parity}} = P_{\text{state}\to\text{parity}}(\vec{\phi}_{\text{state}}/\tau) = \phi_{\text{parity}}/\tau$, I compute parity components of frequency directly from the parity components of phase. In principle, we should exercise care in computing the uncertainty of frequency parity components, since $\phi(\tilde{\mathcal{N}}, \tilde{\mathcal{E}}, \tilde{\mathcal{B}})$ and τ are non-trivially correlated, but it turns out handling this carefully doesn't give results significantly different from the naive method of propagating uncertainty (i.e., ignoring covariances).

Blind the EDM

It is plausible that, if we expected the EDM to have a certain value (e.g., zero), then we would subconsciously "massage" the system until we managed to measure that particular value because, alas, we are humans. In order to avoid this potential source of bias, we add a "blind" to the EDM and related channels, $\mathcal{A}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$, $\phi^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$, and $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$. Since the EDM "lives" in the frequency ω, we apply the blind there first and propagate backward. It is easiest to work this out thoroughly using vector notation. Suppose we have a parity-basis frequency vector $\vec{\omega}_p$ and want to apply a parity-basis blind $\vec{\Omega}_p$. The frequency blind in the state basis is then properly $\vec{\Omega}_s = P^{-1}\vec{\Omega}_p = P^T\vec{\Omega}_p$, where P is the orthogonal transformation matrix from the state basis to the parity basis. In our case, the only non-zero component of $\vec{\Omega}_p$ is

$\Omega_p^{\check{N}\tilde{\mathcal{E}}}$.

$$\begin{aligned}
\vec{\omega}_p &\rightarrow & \vec{\omega}_p + \vec{\Omega}_p \\
\vec{\phi}_p &\rightarrow & \vec{\phi}_p + \vec{\Omega}_p \tau \\
\vec{\phi}_s &\rightarrow & \vec{\phi}_s + P^T \vec{\Omega}_p \tau.
\end{aligned} \tag{3.10}$$

To propagate to the asymmetry, we'll switch to index notation:

$$\begin{aligned}
\mathcal{A}_s^i &\rightarrow & \mathcal{A}_s^i + 2|\mathcal{C}|_s^i \sum_j (P^T)^{ij} \Omega_p^j \tau \\
&= & \mathcal{A}_s^i + 2|\mathcal{C}|_s^i \sum_j P^{ji} \Omega_p^j \tau \\
&= & \mathcal{A}_s^i + 2|\mathcal{C}|_s^i P^{\check{N}\tilde{\mathcal{E}},i} \Omega_p^{\check{N}\tilde{\mathcal{E}}} \tau \\
\mathcal{A}_p^i &\rightarrow & \mathcal{A}_p^i + 2\Omega_p^{\check{N}\tilde{\mathcal{E}}} \tau \sum_j P^{ij} |\mathcal{C}|_s^j P^{\check{N}\tilde{\mathcal{E}},j}.
\end{aligned} \tag{3.11}$$

Let's pause and consider the ideal case when $|\mathcal{C}|_s^j = |\mathcal{C}|_s$ for all j. Then orthogonality of P implies that $\mathcal{A}_p^{i \neq \check{N}\tilde{\mathcal{E}}} \rightarrow \mathcal{A}_p^{i \neq \check{N}\tilde{\mathcal{E}}}$ is unchanged by the blind, and the blinded EDM-correlated asymmetry is $\mathcal{A}_p^{\check{N}\tilde{\mathcal{E}}} \rightarrow \mathcal{A}_p^{\check{N}\tilde{\mathcal{E}}} + 2\Omega_p^{\check{N}\tilde{\mathcal{E}}} \tau |\mathcal{C}|$ as we might expect.

However, in the more general case that state-by-state contrasts are not identical (which is indeed the case for any particular dataset), $|\mathcal{C}|_s^j \neq |\mathcal{C}|_s$ for all j, self-consistency would require some contribution from the blind to all parity channels of the asymmetry, $\mathcal{A}_p^{i \neq \check{N}\tilde{\mathcal{E}}}$. We normally blind by modifying only the EDM parity channel. This means that computing phases or frequencies from blinded asymmetry data could result in an inconsistency–and even worse, it could reveal the blind. As long as we are careful about this possibility, however, we are free to blind only the EDM-correlated asymmetry channel for the sake of simplicity. In short: we can compute unblinded phases and frequencies from unblinded asymmetries, but once the blind is added to the EDM-channel asymmetry, we can never again use asymmetries to compute phases or frequencies.

Because each parity component of a quantity has equal-magnitude contributions from every state component (e.g., the non-reversing component is an equal-weight sum of all state components), with distinct parity components distinguished only by the relative signs of each state contribution, it follows that $P^{ij} = \pm\frac{1}{8}$. We can simplify the expression for the EDM-channel asymmetry by exploiting the identity $(P^{ij})^2 = (\pm\frac{1}{8})^2$, in which case

$$\mathcal{A}_p^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} \to \mathcal{A}_p^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} + 2\Omega_p^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}\tau\langle|\mathcal{C}|_s\rangle, \tag{3.12}$$

where $\langle|\mathcal{C}|_s\rangle$ is the average magnitude of the contrast across all states.

The blind is simply stored as a number in a binary file, and interpreted as a precession frequency drawn from a normal distribution with $1\sigma = 13$ mrad/sec (approximately the 90% confidence bound on $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ from ACME I). We have taken some caution to write our analysis code such that it is difficult to unblind by accident. We only checked the blind after the entire collaboration had agreed that the data analysis and systematic error control were finalized. After unblinding, we did not change the results of our analysis.[3]

Difference between g-factors: η

For each block, after computing precession frequencies, we compute the molecular quantity $\eta \equiv -\frac{\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{B}}}}{\mu_B|\mathcal{E}\mathcal{B}|} \approx -0.8$ nm/V, which describes the electric-field-dependence of the H state g-factor (see Sec. 4.8). This should be normally distributed with uncertainty set by the precession frequency. This quantity is used as a sanity check and diagnostic, but isn't necessary to compute the EDM.

Superblock switches

We compute superblock parity components for $\mathcal{A}$, $\mathcal{C}$, ω, τ, η, and auxiliary parameters used to look for systematics (beam box temperature, vacuum pressure, etc.). Data taken in different superblock states are uncorrelated, so simple error propagation is sufficient. For a given superblock state within a superblock, we can compute the weighted average of every quantity for each group separately, or average together quantities (e.g., $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$) in a group-independent way.

The non-reversing superblock parity component is equivalent to an unweighted average over all superblock states. If different amounts of data are taken in each superblock state, then the uncertainty in each parity component has the largest contribution from the state

---

3. It would be possible to construct the blind such that it is actually impossible for us to unblind before this. However, we are trying to guard against subconscious bias, not outright malice.

with the least data. However, a weighted average could bias the result in the presence of a systematic error, so we use only unweighted error propagation when combining data taken in different states and accept the slight loss in sensitivity.

Uberblock switches

Anything slower than a superblock switch is an "uberblock" switch, and is not regarded as necessary for an EDM measurement. In particular, we will regard the run-level quantities (i.e., the average value of $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ for a given run) as being given by *weighted* averages over all superblocks (irrespective of the uberblock states, since uberblock switches are not necessary to obtain valid EDM values). However, we use the uberblock switches to confirm control of systematic effects.

To do this, we compute regressions (typically linear but sometimes quadratic) against all uberblock switches (e.g., the power of the refinement laser, P_{prep}). For example, we may compute $\frac{d\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{dP_{\mathrm{prep}}}$. We often run with different magnitudes of the applied magnetic field, $\mathcal{B}$, and I typically compute regressions (e.g., $\frac{d\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{dP_{\mathrm{prep}}}$) for each value of $\mathcal{B}$ separately, as well as for all $\mathcal{B}$ combined.

In some cases, we will intentionally check for a systematic with multiple uberblock switches (e.g., detuning of the STIRAP lasers Δ_{STIRAP} and magnetic field gradient $\partial\mathcal{B}/\partial z$), in which case we compute regressions involving both parameters (e.g., $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = \omega_0^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} + \omega_1^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} s_1 + \omega_2^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} s_2 + \omega_{12}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} s_1 s_2$ for parameters s_1 and s_2). We carefully examine all resulting regressions and confirm that we can understand their behavior.

3.2 Statistical distributions in an ideal measurement

By default, we rely on Gaussian statistics, but our underlying data is not exactly Gaussian to begin with. Worse still, we perform operations on the data that can introduce correlations and map even normal variables to non-normal variables. Although we check the statistical behavior of our results as we go via calculations of reduced chi-squared statistics, comparison to simulations, and by varying analysis parameters, it can be useful to have at least a broad picture of the statistical assumptions that we are otherwise sweeping under the rug.

Here, I will proceed with a detailed statistical analysis for the case of an ideal measurement, where there is no technical noise source such as electronics or experiment timing. We will see in Sec. 4.15 that in fact, we have significant noise beyond the level expected for an ideal measurement. The treatment of the ideal case here has been useful in ruling out statistical artifacts as the source of our excess experimental noise. In future iterations of the ACME experiment it will be necessary to suppress the additional technical noise (since otherwise an improved measurement cannot be made), and the statistics of an ideal experiment will hopefully be recovered.

3.2.1 PMT signal: compound Poisson process

Implicitly, we assume certain statistical properties of the electronic signal generated by the photomultiplier tubes used to detect fluorescence signals in the ACME experiment. These assumptions inform the manner in which we analyze data (e.g., we assume that the photon shot-noise limit on the phase sensitivity can be reached), even though we do not directly "pre-process" the PMT signals. Therefore, it is instructive to consider how exactly these electronic signals arise, and what limits their statistical noise. In subsequent sections, we will analyze the statistical properties of "processed" data, keeping the underlying statistics of the PMT signals in mind.

In the simplest case, we model our data as being generated from two Poisson processes, corresponding to signal photoelectrons and background photoelectrons with characteristic numbers $N_{S(B)}$ for each sample within a trace. Note that we are already making a huge simplifying assumption, namely that the gain of photoelectrons is uniform.

The gain of the PMT's used to detect photons would be better (though still approximately) described by a Poisson distribution [176]. The measured fluorescence in a particular sample is $F = \sum_{i=1}^{n} f_i$, where f_i is the photoelectron gain of the ith detected photon, and is drawn from some gain distribution, while $n \sim \text{Poisson}(N_{\text{pe}})$ and N_{pe} is the expected number of photoelectrons incident on the PMT's. The notation $x \sim D(p_1, \cdots, p_n)$ denotes that a random variable x is drawn from the distribution D parametrized by $p_1, \cdots, p_n$. Here, D represents a distribution's name rather than a proper function of the parameters.

If we model the gain of each dynode stage with a Poisson distribution, then the varia-

tion in the gain of the first stage dominates the overall variance of the PMT gain (since the total gain of the subsequent stages is averaged over the electrons generated at the output of the first dynode). Therefore, we can write $F = \sum_{i=1}^{n} \alpha G_i$, where $G_i \sim \text{Poisson}(N_1)$, N_1 is the expected gain of the first stage of the PMT's, and $\alpha = N_{\text{PMT}}/N_1$ is the expected product of gains from the remaining PMT stages. (Note that I'm still making the simplifying assumption that all 8 PMT's have the same gain distribution.) That is, the number of photoelectrons is the sum of Poisson-distributed variables, where the number of terms in the sum is drawn from another Poisson distribution. This is a particular case of the "compound Poisson distribution," which has no simple form in general. In the limit that the expected number of photoelectrons is large, $N_{\text{pe}} \to \infty$, the central limit theorem gives $F \sim \text{Normal}(N_{\text{pe}} m_1, N_{\text{pe}} m_2)$ where $m_i = E[f^i]$ is the expectation value (i.e., mean) of f^i and $\text{Normal}(\mu, \sigma^2)$ denotes a normal distribution with mean μ and variance σ^2 (e.g., see the lecture notes [177]). In the specific case $f_i/\alpha \sim \text{Poisson}(N_1)$, the number of detected signal photons is given by

$$F \sim \text{Normal}(\alpha N_1 N_{pe},\ \alpha^2 (N_1^2 + N_1) N_{pe}). \tag{3.13}$$

Thus when we have a large number of photoelectrons, the fluorescence signal is well-approximated by a normal distribution.

When, in addition, the gain of the first PMT stage is reasonably large such that $N_1^2 \gg N_1$, the fluorescence signal is drawn from a distribution approximated by a normal distribution with mean $\mu = \langle G \rangle N_{pe}$ and standard deviation $\sigma = \langle G \rangle \sqrt{N_{pe}}$, exactly as we would naively hope. Here, $\langle G \rangle$ is the expected gain from the PMT's. The exact standard deviation in this model is $\sigma = \langle G \rangle \sqrt{N_{pe}} (1 + 1/\sqrt{N_1})$. This result is a specific consequence of the more general fact that the spectral noise density, S, of a compound Poisson process is given by $S = 2 q_e I \times F_{\text{e.n.}}$, where $F_{\text{e.n.}} = 1 + \text{var}(G)/\langle G \rangle^2$ is the "excess noise factor," I is the photoelectron current, and q_e is the magnitude of the electron charge.

In ACME II, Cristian Panda measured the gain distribution of the PMT's by integrating the current response of single-photon detection events, for many photons. By taking the ratio of the mean and variance of these distributions, we find that the PMT's have excess noise factors in the range $1.2 - 1.3$. I typically take a representative value of $F_{\text{e.n.}} = 1.25$.

Since we empirically determine the variance of asymmetry values, this excess noise factor is automatically accounted for in the nominal error bars of all quantities of interest that are extracted from the data. However, it may be useful to realize that we typically neglect it (and in particular, I neglect it in the folllowing sections) when considering the statistical properties of the EDM data. Since the excess noise factor is "reasonably close" to 1, we expect this approximation not to invalidate the main conclusions.

3.2.2 Background subtraction

We model the background and signal as being produced by separate Poisson processes with characteristic photoelectron numbers N_{BG} and N_S. The background measurement, made over several hundred bins, allows us to determine the characteristic background rate to a precision far greater than the fluctuations of the background in any particular bin, so we assume that N_{BG} is known exactly. There are ~ 25 background photons per bin after summing over 25 shots/trace and 4 degenerate traces/state.

We measure only total counts $c = s + b$, where s is the true signal count and b is the true background count. We subtract a measured average background level N_{BG} to obtain an inferred signal $\tilde{s} = c - N_{BG} \neq c - b$, where the inequality reminds us that we do not actually know the exact number of background counts in a particular bin. However, $E[c - N_{BG}] = E[c - b]$, so our procedure gives an unbiased estimator of the true signal counts. Here, $E[x]$ denotes the expected value of a random variable x over many iterations of an experiment.

The total counts are composed of the sum of two Poisson-distributed variables, and are thus also Poisson distributed, $c \sim \text{Poisson}(N_S + N_{BG})$. Even the background counts are large enough (~ 25 per bin) to justify replacing the Poisson distribution with a normal distribution, $\text{Poisson}(N) \to \text{Normal}(N, N)$[4], so we approximate $c \sim \text{Normal}(N_S + N_{BG}, N_S + N_{BG})$. Therefore $\tilde{s} = s - N_{BG} \sim \text{Normal}(N_S, N_S + N_{BG})$.

In conclusion, background-subtraction is valid when the background and signal count rates within a bin are both significantly larger than one, at the unavoidable expense of a

---

4. As a reminder, the first argument of the normal distribution is its mean, while the second argument is its variance.

slight increase in the variance of the inferred signal relative to the true (background-free) signal variance, which can never be independently measured. That said, a word of caution is necessary: while our background-subtraction works in expectation, it can give unphysical, negative inferred signal counts when $N_S \approx N_{BG}$ or smaller. Avoiding this is one (though not the only) advantage of using a fluorescence threshold cut during data analysis, in which any bin with fewer than ~ 500 detected photons is discarded from the data set.

More complicated background-subtraction procedures exist, which give Bayesian estimations of signal counts, given a known typical background rate and observed total count [178–180]. These methods appear to have been developed mostly for astrophysical observations, where the signal-to-noise ratio can be very low in observing faint sources. However, the efficacy of these methods depends sensitively on the signal and background models, and they tend to be much more computationally intensive. For these reasons, I chose to use the simplest version of background subtraction as above.

3.2.3 Asymmetry

In everything that follows, we consider only background-subtracted signals in the X and Y bins, which we denote S_X and S_Y, respectively. From these, we compute the asymmetry

$$\mathcal{A} = \frac{S_X - S_Y}{S_X + S_Y}. \tag{3.14}$$

In the regions of interest, $S_{X,Y}$ can be modelled as Gaussian-distributed variables (with $\mu_{X,Y} \approx \sigma_{X,Y}^2 \approx N_{X,Y}$). The sum or difference of normally distributed variables is itself normally distributed, $S_{\pm} \equiv S_X \pm S_Y \sim \text{Normal}(\mu_X \pm \mu_Y, \sigma_X^2 + \sigma_Y^2)$. Further, note that $\text{cov}(S_+, S_-) = \sigma_X^2 - \sigma_Y^2 \neq 0$ because the numerator and denominator, $S_X - S_Y$ and $S_X + S_Y$, are both constructed from the same independent variables. This might seem surprising at first, but in general, linear combinations of independent variables are not independent. The correlation coefficient is then $\rho \equiv \text{corr}(S_+, S_-) \equiv \text{cov}(S_+, S_-)/(\sigma_+ \sigma_-) = (\sigma_X/\sigma_Y) - (\sigma_Y/\sigma_X)$. We can approximate this simply as follows: for signals $S_{X,Y} = S_0(1 \pm \mathcal{A}/2)$, we expect $\sigma_{X,Y} \approx \sigma_0\sqrt{1 \pm \mathcal{A}/2}$ so that $\rho \approx \mathcal{A} + \mathcal{O}(\mathcal{A}^3) \leq 0.2$ typically.

In general, the distribution of the ratio between two correlated normally-distributed

variables is extremely complicated. However, it's clear that in the limit that the denominator approaches a constant, the ratio should approach an ordinary normal distribution. This idea is made more rigorous in terms of the "coefficient of variation," σ/μ, and explored via Monte Carlo simulations in [181]. They conclude that for $\rho < 0.5$, $|\sigma/\mu|_{\text{num.}} > 0.19$, and $|\sigma/\mu|_{\text{den.}} < 0.09$ (where the "num." and "den." subscripts specify the coefficient of variation of either the normally distributed numerator or denominator only), the ratio distribution is reasonably well-described by a normal distribution with mean[5]

$$\mu_{\mathcal{A}} \approx \frac{\mu_X - \mu_Y}{\mu_X + \mu_Y} + (\sigma_X^2 + \sigma_Y^2)\frac{\mu_X - \mu_Y}{(\mu_X + \mu_Y)^3} - \frac{\sigma_X^2 - \sigma_Y^2}{(\mu_X + \mu_Y)^2}. \tag{3.15}$$

The first term is the ideal asymmetry. Using $\sigma_{X,Y}^2 \to \mu_{X,Y}$, the correction terms cancel exactly. Thus the higher-order terms are deeply suppressed.

The variance of the ratio distribution in this regime leads to

$$\sigma_{\mathcal{A}}^2 \approx \frac{(\sigma_X^2 + \sigma_Y^2)(\mu_X - \mu_Y)^2}{(\mu_X + \mu_Y)^4} + \frac{\sigma_X^2 + \sigma_Y^2}{(\mu_X + \mu_Y)^2} - \frac{2(\sigma_X^2 - \sigma_Y^2)(\mu_X - \mu_Y)}{(\mu_X + \mu_Y)^3}. \tag{3.16}$$

Using the same approximation $\sigma_{X,Y}^2 \to \mu_{X,Y} \to N_{X,Y}$ and writing $\frac{\mu_X - \mu_Y}{\mu_X + \mu_Y} \to \mathcal{A}$, we can express this as

$$\sigma_{\mathcal{A}}^2 \approx \frac{\mathcal{A}^2 + 1 - 2\mathcal{A}^2}{N_X + N_Y} = \frac{1 - \mathcal{A}^2}{N_X + N_Y}. \tag{3.17}$$

The $\mathcal{A}^2$ term is typically much smaller than the previous leading term since $\mathcal{A} \ll 1$ in the usual regime. If we neglect this correction term, then we recover the shot-noise limit, $\sigma_{\mathcal{A}} \approx 1/\sqrt{N_{\text{tot}}}$, where N_{tot} is the total number of molecules measured. Thus we can use $\mathcal{A}_{\text{measured}} \sim \text{Normal}(\mathcal{A}, 1/(N_X + N_Y))$ as usually assumed.

Before moving on, we must check the conditions $|\sigma/\mu_{\text{num.}}| > 0.19$, and $|\sigma/\mu|_{\text{den.}} < 0.09$. Note that $|\sigma/\mu|_{\text{num.}}$ is minimal for maximal values of $|\mathcal{A}| < 0.2$. Letting $\sigma_{\text{num.}} \to \sqrt{N}$, we find that the first condition is satisfied for $\sqrt{N} > 0.04$, which is trivially true. Likewise, $|\sigma/\mu|_{\text{den.}} \approx 1/\sqrt{N} < 0.09$ implies $N > 125$. My signal threshold is typically set so that

---

5. Following the conventional notation, I generally denote the mean of a normal distribution by μ.

$N > 500$ photons per bin pair are measured, providing a comfortable margin.[6]

I'll note, somewhat out of turn, that the asymmetry computed from combined datasets (e.g., summed degenerate traces) is equivalent to the weighted average of asymmetries from individual datasets, where the weights are given by the shot noise limit: $\mathcal{A}_{\text{tot}} \equiv \frac{\sum S_{X,i} - \sum S_{Y,i}}{\sum S_{X,i} + \sum S_{Y,i}} = \sum[\mathcal{A}_i N_i]/N_{\text{tot}} \equiv \bar{\mathcal{A}}$. This justifies summing degenerate traces at the beginning of analysis.

3.2.4 Grouped asymmetry

We compute a linear regression of asymmetry values within a local region of each trace ($\sim 10 - 50$ bin pairs). Assuming the linear model of asymmetry values vs. bin pair index is accurate, the inferred mean value within the group is $\bar{\mathcal{A}} \sim \text{Normal}(\mathcal{A}_0, \sigma_{\mathcal{A}}^2/n)$, where n is the number of points in the linear regression and $\mathcal{A}_0$ is the population mean at the center of the the group.

We also assign the variance of the asymmetry values by computing the sample variance of the residuals (using a denominator of $n-2$ instead of the usual $n-1$ to take account of the two degrees of freedom in the linear regression). Although this gives an unbiased estimator for the population variance, the assigned variance S^2 is itself a random variable and is distributed according to $(n-2)S^2/\sigma_{\mathcal{A}}^2 \sim \chi_{\nu=n-2}^2$, where χ_ν^2 is a chi-squared distribution with ν degrees of freedom. Nick Hutzler discovered this in the context of an ordinary mean (rather than linear regression) during ACME I and discusses some interesting consequences in his thesis [146, Sec. 4.4.4]. We can adapt much of his insight, with the caveat that for a linear regression, $(\bar{\mathcal{A}} - \mathcal{A}_0)/S \sim t_{n-2}$ instead of t_{n-1}, where t_ν is a Student's t-distribution with ν degrees of freedom (since two degrees of freedom are used to calculate the offset and slope of a linear regression, rather than one used to calculate the sample mean as considered in [146]). As a consequence, $(\bar{\mathcal{A}} - \mathcal{A}_0)^2/S^2 \sim F(1, n-2)$, where $F(d_1, d_2)$ is the F-distribution with (integer) parameters d_1 and d_2, which are related to the distribution's mean and variance in a somewhat complicated way (again, see [146, Sec. 4.4.4] for a detailed

---

6. Interestingly, the condition $N > 125$ would have been impossible to satisfy in ACME I due to the lower signal, so the variance of individual asymmetry points would be expected to have a more complicated functional dependence on signal size and asymmetry than in the current generation.

outline of the basic approach used here). When we compute χ^2 from a collection of group-averaged asymmetries using sample variances, we should not expect the result to follow an ideal χ^2-distribution.

Looking forward a bit, it's important to realize that when we compute parity sums, we propagate uncertainties by taking linear combinations of state-basis sample variances. The distribution of a linear combination of χ^2 variables is expressed as an infinite gamma series.[7] Needless to say, we don't want to delve into this if we don't have to.

However, we can gain some insight about the behavior of assigned variances for parity sums by noting that the sample variance in each state is comparable, and so the assigned variance in a parity component is approximately given by an *unweighted* sum of χ^2-distributed variables. In general, if these have degrees of freedom $\nu_1, \ldots, \nu_N$, then the sum is distributed as $\chi^2_{\nu_1 + \cdots + \nu_N}$; e.g., see the lecture notes [183]. The χ^2-distribution approaches a normal distribution as the number of degrees of freedom increases, so we can expect the χ^2_{red} value computed from parity components of asymmetries to be much better-behaved than the discussion in [146] would imply for state-basis asymmetries. Specifically, we can replace $n - 1 \to 16(n - 2)$ in [146, Eq. 4.67] to see that the expected correction factors for parity-basis asymmetries are very close to unity. Explicitly, the reduced chi-squared statistic of parity-basis asymmetry values is

$$\left[\chi^2_{\mathrm{red}}\right]^{(\mathrm{parity})} = \frac{16n - 33}{16n - 35} \pm \sqrt{\frac{2}{N - 1}} \times \frac{16n - 33}{16n - 35} \sqrt{\frac{16n - 34}{16n - 37}}, \qquad (3.18)$$

where N is the total number of groups used to compute χ^2_{red}, and (as a reminder) n is the number of asymmetry values included in each group. The RHS of the expression above is written as $(\mathrm{mean}) \pm (\mathrm{standard\,deviation})$. We see that the mean of the reduced chi-squared statistic for values in the parity basis appears quickly approaches 1 for $n \gtrsim 3$, while the standard deviation approaches $\sqrt{2/(N - 1)}$, as would be the case for gaussian data. This conclusion holds even though the mean asymmetry values within a group, computed in the state basis, are not normally distributed.

---

7. This result "is obtained in a straightforward manner," as the abstract of [182] helpfully informs us.

3.2.5 State-by-state asymmetry and contrast

For each $\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}$ state, recall that we collect data in $\tilde{\theta} = \pm 1$ states. We construct "state-by-state" asymmetry and contrast values by taking linear combinations of asymmetries values, $\mathcal{A}_\pm$, measured with $\tilde{\theta} = \pm 1$:

$$
\begin{aligned}
\mathcal{A} &= \tfrac{s}{2}(\mathcal{A}_+ + \mathcal{A}_-) \\
\\
\mathcal{C} &= \tfrac{t}{2}(\mathcal{A}_+ - \mathcal{A}_-)
\end{aligned}
\tag{3.19}
$$

where $s = \mathrm{sign}(\mathcal{C})$ and $t = -180°/(2\pi\Delta\theta)$, with $\Delta\theta \sim 6°$ giving the full waveplate dither range in degrees. Here, I leave $\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}$ state designations implicit. The noise in the $\tilde{\theta} = \pm 1$ states is uncorrelated, but as mentioned previously, the noise in linear combinations of uncorrelated variables is generally correlated. In particular, $\sigma_{\mathcal{A}}^2 = s^2(\sigma_+^2 + \sigma_-^2)/4$, $\sigma_{\mathcal{C}}^2 = t^2(\sigma_+^2 + \sigma_-^2)/4$, and $\mathrm{cov}(\mathcal{A},\mathcal{C}) = st(\sigma_+^2 - \sigma_-^2)/4$. We can interpret the covariance as follows: if $\mathcal{A}_+$ tends to have larger swings than $\mathcal{A}_-$, then an excursion from the mean in $\mathcal{A}$ is likely to be associated with an excursion in the mean of $\mathcal{C}$ in the same direction (when $st > 0$). On the other hand, if $\mathcal{A}_-$ tends to have larger swings than $\mathcal{A}_+$, then their excursions will be anti-correlated. In the case that the variances in the original states are the same, these two effects cancel out and $\mathrm{cov}(\mathcal{A},\mathcal{C}) \to 0$.

On average, the fluorescence and therefore $\sigma_{\mathcal{A}}^2$ will be unchanged between $\tilde{\theta}$ states, so $E[\mathrm{cov}(\mathcal{A},\mathcal{C})] = 0$. However, any particular state might have slightly difference fluorescence amplitudes due to typical fluctuations in the molecular beam properties, so we ought to propagate this covariance correctly. I've confirmed that the assigned uncertainty in ϕ ends up being $5 - 10\%$ percent higher when the covariance is included, and the associated reduced-chi-squared values (measured for "well-behaved" channels like $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ and computed by comparing groups within a block), χ^2_{red}, indicate a better uncertainty assignment when $\mathrm{cov}(\mathcal{A},\mathcal{C}) \neq 0$ is used. Here, $\chi^2_{\mathrm{red}} = \frac{1}{N-1}\sum_{i=1}^{N} \frac{(x_i - \bar{x})^2}{\sigma_i^2}$ characterizes the size of actual uncertainties in a data set $\{x_i\}$ relative to nominal ("assigned") uncertainties σ_i, and $\bar{x}$ is the sample mean of the data. For normally distributed data with variances σ_i^2, the expected value of χ^2_{red} is 1; a value larger than one indicates that uncertainties have been

underestimated.

Linear combinations of normal variables are distributed normally, so $\mathcal{A}$ and $\mathcal{C}$ should be described approximately by correlated normal distributions.

3.2.6 Phase

In order to understand in detail how we extract the phase from the state-dependent asymmetry and contrast, we should dwell on the physical origin of the contrast $\mathcal{C}$. Amar Vutha shows that if the phase is distributed randomly with mean α and variance σ_α^2, then the expected asymmetry is attenuated [144]:

$$\langle \mathcal{A} \rangle = |\mathcal{C}| \cos[2\alpha], \tag{3.20}$$

defining the unsigned contrast $|\mathcal{C}|$. In the particular case considered by Amar, $|\mathcal{C}| = \exp(-2\sigma_\alpha^2)$, but the existence of some factor $|\mathcal{C}| < 1$ does not depend on a specific model of the phase distribution. Comparing to Sec. 2.3.5, we see that $\alpha \equiv \omega\tau - (\delta\theta_0 + \Delta\theta\tilde{\theta}) + \frac{\pi}{4}(\tilde{\mathcal{P}}+1)$ for our measurement in a state with $\tilde{\theta} = \pm 1$ and $\tilde{\mathcal{P}} = \pm 1$. Once again, we define $\omega\tau - \delta\theta_0 \equiv n\pi \pm \frac{\pi}{4} + \delta\phi$ so that $\langle \mathcal{A}(\tilde{\theta}) \rangle = \pm\tilde{\mathcal{P}}|\mathcal{C}| \sin[2(\delta\phi - \Delta\theta\tilde{\theta})] \equiv \mathcal{C} \sin[2(\delta\phi - \Delta\theta\tilde{\theta})]$, where we've chosen a convention for the contrast sign. Here, I have written the asymmetry explicitly as a function of the polarization dither state $\tilde{\theta}$ to distinguish from the $\tilde{\theta}$-averaged and $\mathcal{C}$-corrected asymmetry of Eq. 3.19.

Let us first consider the $\delta\phi \to 0$, $\Delta\theta \to 0$ limit, where $\langle \mathcal{A}(\tilde{\theta}) \rangle \approx 2\mathcal{C}(\delta\phi - \Delta\theta\tilde{\theta})$. This motivates a definition of the measured contrast,

$$\mathcal{C}_m = -\frac{\mathcal{A}(\tilde{\theta} = +1) - \mathcal{A}(\tilde{\theta} = -1)}{4\Delta\theta}, \tag{3.21}$$

so that $\langle \mathcal{C}_m \rangle = \mathcal{C}$ in this limit. Then taking the asymmetry $\mathcal{A}$ averaged over $\tilde{\theta}$ states (and multiplied by the sign of the contrast) as defined in Eq. 3.19, the equation $\langle \mathcal{A} \rangle = 2|\langle \mathcal{C}_m \rangle|\delta\phi$ is valid.

We will now extend this to the next-leading order in small phases $\delta\phi$ and $\Delta\theta$. Here, $\delta\phi \ll 1$ refers to the average phase deviation. Then the asymmetry averaged over $\tilde{\theta}$ switches has expectation value

$$\langle \mathcal{A} \rangle \approx 2|\mathcal{C}|\, \delta\phi (1 - 2\Delta\theta^2 - \frac{2}{3}\delta\phi^2). \tag{3.22}$$

We can similarly evaluate the expected value of $\mathcal{C}_m$ to third order:

$$\frac{\langle \mathcal{C}_m \rangle}{\mathcal{C}} \approx 1 - 2\delta\phi^2 - \frac{2}{3}\Delta\theta^2. \tag{3.23}$$

Then we can find $\frac{\langle \mathcal{A} \rangle}{2\langle |\mathcal{C}_m| \rangle} = \delta\phi(1 - \frac{4}{3}\Delta\theta^2 + \frac{4}{3}\delta\phi^2)^8$. In other words, the ratio of the measured asymmetry to twice the measured contrast gives the phase deviation $\delta\phi$ at leading order, but also gives correction terms that are cubic in small phases, $\delta\phi^3$ and $\delta\phi \times \Delta\theta^2$. Thus at leading order, we can define the measured phase to be

$$\delta\phi_m^{(1)} = \frac{\mathcal{A}}{2|\mathcal{C}_m|}, \tag{3.24}$$

with the property that $\langle \delta\phi_m^{(1)} \rangle = \delta\phi + \mathcal{O}(\delta\phi, \Delta\theta)^3$: the measured phase defined in this way is unbiased up through second order in small quantities.

To take the higher-order terms into account, we try the solution

$$\delta\phi_m^{(1-3)} = \frac{\mathcal{A}}{2|\mathcal{C}_m|}\left(1 + \frac{4}{3}\Delta\theta^2 - \frac{4}{3}\left(\frac{\mathcal{A}}{2|\mathcal{C}_m|}\right)^2\right). \tag{3.25}$$

The tricky part arises in the third term, proportional to $\left(\frac{\mathcal{A}}{2|C_m|}\right)^3$. When we take the expectation value of this term, we do *not* get precisely the cube of the expectation value of $\frac{\mathcal{A}}{2|C_m|}$. In particular, taking $\frac{\mathcal{A}}{|C_m|}$ to be approximately normally distributed, we find that $\langle \frac{\mathcal{A}}{|C_m|}^3 \rangle = \langle \frac{\mathcal{A}}{|C_m|} \rangle^3 + 3\langle \frac{\mathcal{A}}{|C_m|} \rangle \sigma^2_{\mathcal{A}/|C_m|}$. However, the correction term, $\sim \langle \frac{\mathcal{A}}{|C_m|} \rangle \sigma^2_{\mathcal{A}/|C_m|}$, is of comparable order as the terms we neglect in using the approximation $\langle \frac{\mathcal{A}}{|C_m|} \rangle \approx \frac{\langle \mathcal{A} \rangle}{|\langle C_m \rangle|}$ and in particular is quite small for the generic values of $N \gg 1$ in the ACME II measurement. Thus, Eq. 3.25 accurately gives $\langle \delta\phi_m^{(1-3)} \rangle = \delta\phi + \mathcal{O}(\delta\phi, \Delta\theta)^5$.

The term $\frac{4}{3}\Delta\theta^2$ is a small ($\sim 1\%$) scaling correction to every phase, which is not very dangerous. The last term, however, is potentially more interesting because it characterizes the leading-order nonlinear relationshionship between asymmetry and the measured phase.

---

8. We take the shortcut of computing the ratio of expectation values, rather than the expectation value of the ratio $\mathcal{A}/\mathcal{C}_m$, because the fractional uncertainty in $\mathcal{C}_m$ is small (see [181]).

The nonlinear term guarantees that parity sums of $\frac{\mathcal{A}}{2|\mathcal{C}_m|}$ will not isolate components with distinct physical sources. We normally ignore the nonlinear term during analysis, instead using $\delta\phi_m^{(1)} \equiv \frac{\mathcal{A}}{2|\mathcal{C}_m|}$. Using this approximation is equivalent to *adding* a cubic term to the actual phase (i.e., $\langle \delta\phi_m^{(1)} \rangle = \delta\phi + \mathcal{O}(\delta\phi^3)$), which allows for nonlinear couplings between phase components. For example, with nonzero $\phi^{\tilde{\mathcal{N}}\tilde{\mathcal{B}}}$, $\phi^{\tilde{\mathcal{B}}}$, and $\phi^{\tilde{\mathcal{E}}}$, there will be a corresponding non-zero contribution to $\phi_m^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$. Nevertheless, we have not seen evidence that any such terms are problematic at the precision of ACME II, and we have no models by which they should be. As a result, all analysis is performed by computing $\delta\phi_m^{(1)}$ only.

Since both $\mathcal{A}$ and $|\mathcal{C}|$ are described by correlated normal distributions, ϕ is described by a ratio distribution. (Taking the absolute magnitude of $\mathcal{C}$ introduces no difficulties since fluctuations never change the sign of $\mathcal{C}$ in the region of signal used; however, we must remember that $\mathrm{cov}(\mathcal{A}, |\mathcal{C}|) = s \times \mathrm{cov}(\mathcal{A}, \mathcal{C})$.) The correlation coefficient between them, produced simply by random fluctuations in signal between $\tilde{\theta} = \pm 1$ measurements, is certainly less than 0.5. As described before, the numerator, $\mathcal{A}$, has a coefficient of variation that is trivially within the required range. The fractional uncertainty of contrast is nearly shot-noise limited, so the denominator will have a sufficiently small coefficient of variation provided $N > 125$ in each bin pair, just as before. The phase is therefore normally distributed with mean

$$\mu_\phi \approx \frac{1}{2}\left(\frac{\mathcal{A}}{|\mathcal{C}|} + \frac{\sigma_\mathcal{C}^2}{\mathcal{C}^2}\frac{\mathcal{A}}{|\mathcal{C}|} - \frac{t(\sigma_+^2 - \sigma_-^2)}{\mathcal{C}^2} \right). \tag{3.26}$$

The second term is a very small correction, of order $\frac{t^2}{4N} \sim 1/N$, to the first term. The last term should average to 0 since we don't expect a systematic bias toward more signal in either $\tilde{\theta}$ state, but it could dominate the first two terms in any particular measurement within the region of the molecule pulse where $\mathcal{A} \approx 0$ (typically close to the fluorescence peak). Numerically, assuming 5% number fluctuations between $\tilde{\theta}$ states, the last term in the parentheses is of order $0.1/N$, which can be neglected without danger.

The variance of the distribution of ϕ measurements is

$$\sigma_\phi^2 \approx \frac{\sigma_\mathcal{C}^2}{\mathcal{C}^2}\phi^2 + \frac{\sigma_\mathcal{A}^2}{\mathcal{A}^2}\phi^2 - \frac{2t(\sigma_+^2 - \sigma_-^2)}{\mathcal{C}^2}\phi. \tag{3.27}$$

To compare the first two terms, note that

$$\frac{\sigma_\mathcal{A}^2}{\mathcal{A}^2} > \frac{s^2(\sigma_+^2+\sigma_-^2)/4}{0.2^2}$$
$$\approx 6.3(\sigma_+^2 + \sigma_-^2)$$

$$\frac{\sigma_\mathcal{C}^2}{\mathcal{C}^2} \approx \frac{t^2(\sigma_+^2+\sigma_-^2)/4}{0.95^2}$$
$$\approx 6.3(\sigma_+^2 + \sigma_-^2). \tag{3.28}$$

Therefore, the second term usually dominates (in the high-signal regions of the molecular beam pulse, where $\mathcal{A} \ll 0.2$), but can be comparable to the first term in the large-asymmetry regime near the edges of the molecule pulse where $\mathcal{A}$ is larger than normal. Indeed, we choose $\Delta\theta$ so that the uncertainty in the contrast is negligible over most of the region of interest, but we do not make it large enough to stray too far from the center of the Ramsey fringe where sensitivity to $\delta\phi$ is maximized.

In ACME I, a state-averaged contrast was used instead of the group-averaged contrast in order to suppress $\sigma_\mathcal{C}^2$ and reduce the uncertainty of the phase in each group. I have opted not to do this for several reasons: (i) the contrast term mostly inflates the error bars where the asymmetry is large, which occurs in our apparatus due to velocity dispersion only near the tails of the molecule pulse. In this region, the signal-to-noise ratio (SNR) is already relatively low, and changing the analysis routine to improve the SNR there will not significantly improve the final result. (ii) By using a state-averaged contrast in the computation of ϕ, we would correlate nominal phase values in different groups within the molecular pulse. Intuitively, using a state-averaged contrast does not add information and therefore shouldn't actually reduce the final statistical error bar. (iii) The ability to treat distinct groups as statistically independent is a significant advantage in conceptual and practical simplicity.

The covariance term in the formula for σ_ϕ^2 is not negligible because it is only linear in ϕ rather than quadratic. Empirically, I see that including it increases the assigned uncertainty in ϕ by several percent. This marginally improves the χ_red^2 values of parity-basis phase measurements, calculated among groups for data within a block.

3.2.7 Change to parity basis

We change from the state basis to the parity basis by taking a linear combination of state-basis quantities. Since we have determined above that all measured quantities ($\mathcal{A}$, $\mathcal{C}$, ϕ) are approximately normally distributed for an ideal measurement, and using the fact that a linear combination of normally-distributed variables is itself normally-distributed, we obtain the expected parity-basis quantities. Since asymmetries, contrasts, and phases have uncorrelated noise across $\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}$ states, the variance assigned to a parity-basis quantity is the average of variances in the state basis. Parity-basis quantities are correlated because they are computed from the same data as each other (this is analogous to the correlation between $\mathcal{A}$ and $\mathcal{C}$ in a particular $\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}$ state). In particular, if the covariance matrix in the state basis is $\Sigma^{\mathrm{state}} = \mathrm{diag}(\vec{\sigma}_{\mathrm{state}})$, then the covariance matrix in the parity basis is $\Sigma^{\mathrm{parity}} = P\Sigma^{\mathrm{state}}P^T$, which is non-diagonal. This is important for computing precession frequencies later.

3.2.8 Precession time

Recall that

$$\tau = \frac{\phi^{\tilde{\mathcal{B}}}_{\mathrm{measured}} + n\frac{\pi}{4}}{|g|\mu_B\frac{d\mathcal{B}}{dI}I}, \tag{3.29}$$

where n accounts for the possible rotation of the readout laser polarization basis, relative to the preparation laser polarization basis, to ensure that $\phi \ll 1$ in all experimental configurations.

Here we rely on parameters that are not measured directly via the fluorescence signal. There is a systematic uncertainty on $|g|$, $\frac{d\mathcal{B}}{dI}$, and I, with the fractional uncertainty in $|g|$ being the largest ($\approx 1\%$). We assume that n is known exactly, although the uncertainty with which the angle can be set between the preparation and readout bases should also introduce a systematic uncertainty to n. In practice, the systematic uncertainty in τ due to $|g|$ is larger than the statistical uncertainty in any high-signal group, even for a single block. (This was not true in ACME I, where the statistical uncertainty was much higher.) However, it is cumbersome to keep systematic and statistical uncertainties separate during normal data

analysis, since the assigned statistical uncertainty should integrate down with more data but the assigned systematic uncertainty should not. Further, the uncertainty in τ is propagated to the precession frequencies ω, but we don't usually care about the absolute uncertainty in the value of τ. Instead, we will primarily be concerned with whether a value shifts within its *statistical* uncertainty as we adjust parameters in the system (e.g., whether $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ depends on laser detunings). Therefore, including the systematic uncertainty in τ could mask shifts larger than those expected from statistical fluctuations alone. Whenever we compute a final value (e.g., of $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$), in principle we ought to add the systematic uncertainty in τ back in. Similar arguments apply for other parameters like η, which is computed from quantities with systematic uncertainty. Note, however, that a 1% systematic uncertainty in τ increases the ultimate uncertainty in a precession frequency by 1%. For the EDM, this means that a $\sim 100\sigma$ effect would have to be measured before the systematic error in τ could introduce a 1σ shift in the mean.

The uncertainty in τ is therefore computed in a relatively straightforward manner by propagating the error in $\phi^{\tilde{\mathcal{B}}}$. Since $\phi^{\tilde{\mathcal{B}}}$ is normally distributed under the assumptions previously discussed, the value of τ within any particular group should be normally distributed with an offset and scaling determined by n and $\mathcal{B}$. An especially large caveat is warranted here: this analysis presumes that the molecular beam velocity passing through the detection region, at a particular time delay from ablation, is perfectly stable over the $\sim$minute time scales of a block. We will see in Sec. 4.15.2 that small fluctuations in the molecular velocity can cause phase noise beyond the level of shot noise.

3.2.9   Precession frequencies

Error propagation for precession frequencies is non-trivial because errors among phases in the parity basis are correlated, and τ is constructed from $\phi^{\tilde{\mathcal{B}}}$. The full procedure I use is as follows:

- Compute the (non-diagonal) covariance matrix for phases in the parity basis.

- Do a "coordinate transformation" from the $\{\phi\}$ "basis" to the $\{\phi, \tau\}$ "basis"[9] and compute the covariance matrix in the new coordinates using the standard methods of error propagation. Due to the dependence of τ on $\phi^{\tilde{\mathcal{B}}}$, there will be non-zero covariances between τ and each phase component (just as there will be covariances between τ and each frequency component).

- Do another coordinate transformation to the ω variables and compute the new covariance matrix using the standard methods.

- The diagonal entries of the new covariance matrix are the variances of each frequency component in the parity basis.

Once again, even when we don't include the systematic uncertainty in the computation of τ, it is important in principle to properly propagate covariances between τ and party-sum components of the phase ϕ. However, unlike what we found for $\mathcal{A}$ and ϕ, using this "proper" procedure produces essentially the same results as the naive procedure when $n = 1$. This is because the relative fluctuations in τ are suppressed by the noiseless $n\frac{\pi}{4}$ term, which dominates $\phi^{\tilde{\mathcal{B}}}$ in the typical signal regime.

The precession frequency $\omega = \phi/\tau$ should once again be described by a ratio distribution. It is straightforward to check that the conditions for approximate normality are satisfied. We should therefore expect the mean value of ω, μ_ω, to be given by the usual expression for the ratio of normally distributed variables [181]:

$$\mu_\omega = \frac{\phi}{\tau} + \frac{\sigma_\tau^2}{\tau^2}\frac{\phi}{\tau} - \frac{\mathrm{cov}(\phi, \tau)}{\tau^2}. \tag{3.30}$$

The second term is a very small ($\sim 10^{-3}$) overall correction factor to the first term. The last term is also negligible: in the worst case, we would have $\mathrm{cov}(\phi, \tau) = \sigma_\phi \sigma_\tau$ so that the last term would be equivalent to $\frac{\sigma_\phi}{\phi}\frac{\sigma_\tau}{\tau}\frac{\phi}{\tau}$, whose first two factors are each much smaller than unity. Furthermore, the covariance between any component of ϕ and τ could in principle

---

9. Of course, this is not a true basis in the sense of linear algebra because the components are linearly dependent. What I mean here is that there exists a set of nine quantities, $f_i = \phi^i$ for $i \in [1, \ldots, 8]$ and $f_9 = \tau$, that can be related to $\vec{\phi}$ by $\vec{f} = A\vec{\phi}$ for some (non-square) matrix A. The covariance matrix for $\vec{f}$ can be computed from the covariance matrix for $\vec{\phi}$ and the transformation matrix A using standard techniques.

be non-negligible in one block, but it shouldn't be strongly biased in one direction over another[10]. Therefore, we don't have any reason to believe that this last term causes any difficulties.

The variance is

$$\sigma_\omega^2 = \frac{\sigma_\tau^2}{\tau^2}\omega^2 + \frac{\sigma_\phi^2}{\phi^2}\omega^2 - \frac{2\text{cov}(\phi,\tau)}{\tau^2}\omega. \tag{3.31}$$

3.2.10 Summary of statistical analysis

It may be useful to recapitulate some key points of this section:

- The excess noise factors of the PMTs inflate the photon shot-noise limit by a factor of ≈ 1.25, but is expected to have no significant effect on the statistical behavior of the data otherwise.

- A data cut on low-fluorescence regions of the molecular beam pulse (below ~ 125 photons per bin pair) is required to ensure good guassian behavior of asymmetry data. This cut also ensures that our method of background-subtraction is valid (i.e., we never record unphysical negative photon counts in a polarization bin used to compute the EDM).

- Although the averaged asymmetry values recorded for groups in the state basis are not normally distributed (in particular, they are t-distributed), all values in the parity basis are normally distributed to a good approximation (because taking linear combinations of state-basis quantities "washes out" the non-normal behavior).

- We have carefully considered effects of higher-order (e.g., cubic) corrections to the calculated phase, based on the sinusoidal profile of the Ramsey fringe and deviations of the phase from the region of maximum sensitivity, and found that they are negligible.

---

10. Recall that $\text{cov}(\phi,\tau)$ expresses a covariance associated with statistical fluctuations, not to be confused with an actual physical relationship between the average values of ϕ and τ. For example, $\langle\phi^{\tilde{\mathcal{N}}\tilde{\mathcal{B}}}\rangle \propto \langle\tau\rangle$ since $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{B}}} \neq 0$, owing to a non-zero difference between the magnetic g-factor in $\tilde{\mathcal{N}} = \pm 1$ states. However, this fact has no implications for the covariance in the _measured_ values of $\phi^{\tilde{\mathcal{N}}\tilde{\mathcal{B}}}$ and τ under constant physical conditions (i.e., constant unknown _true_ values of $\phi^{\tilde{\mathcal{N}}\tilde{\mathcal{B}}}$ and τ, which are unavoidably measured with statistical uncertainty).

- The polarization basis "dither," $\Delta\theta$, is set so that the contrast measurement contributes only a small fraction of the uncertainty in the measurement of the phase. The threshold to achieve this condition depends on the largest magnitude of the asymmetry, $|\mathcal{A}|$, that is used to compute the phase.

- I have argued for the advantages of using a "group-averaged" contrast measurement, as opposed to the "state-averaged" contrast that was used in ACME I.

- The possibility of covariances among quantities computed from the same data, for example $\mathcal{A}$ and $\mathcal{C}$ when computing $\phi = \frac{\mathcal{A}}{2|\mathcal{C}|}$, should be considered for the most accurate assignment of uncertainties. In practice, however, these effects are insignificant in the ACME II data set compared to the effects from technical noise sources that inflate the uncertainty beyond the level expected for an ideal measurement. Note that for the ACME II data set, my analysis code fully handles covariances among quantities throughout.

3.3 Cuts

We apply many data cuts, each of which can be adjusted by analysis parameters. These ensure that the data used during analysis has the expected properties and is reasonably "well-behaved." These are discussed in detail below. All concrete numbers are merely illustrative and are adjusted as a check on the robustness of the analysis code.

When a block- or higher-level cut is detected, a warning message is displayed and saved to a text file. A "cut" block is actually completely analyzed, but is not combined with data from "good" (uncut) blocks when computing superblock- or run-level values (e.g., it is not used to compute the EDM from the EDM data set). Similarly, "cut" groups are simply excluded from averaging together with good groups when computing block-, superblock-, or run-level values. This allows us to check the properties of "cut" groups or blocks as desired. Any cut can be turned off in the analysis. We have found that reasonable ranges of all analysis parameters produce resulting EDM values that are consistent within the statistical uncertainty.

Block-level setpoints

In addition to the raw fluorescence signal for each PMT during each trace, we save a text file that contains all additional information relevant to each block. This file is known as the "header" and contains a separate section for each of the 64 traces. Each section is subdivided into the following categories: (0) Trace #, Start time, End time; (1) Switch times; (2) DAQ properties; (3) Switch states; (4) Instrument setpoints; (5) Logging measurements. See Appendix G for more details on the information recorded in each trace.

Most of the "DAQ properties," including (although not actually a DAQ property) the positioning of the ablation laser on the ThO_2 target, should not change throughout the entire block. If a property that should not be changed in fact does change in the middle of a block, then I throw out that block.

State-level setpoints

The four degenerate traces for each state should have identical instrument setpoints (e.g., applied magnetic fields), as given by the header. If they do not all agree in any state, then I throw out the block.

PMT signal

For each of the 64 traces in a block, I compute the total number of background-subtracted photons detected for each of the 8 PMT's, giving a 64×8 table of total photon numbers. If the ratio between the maximum and minimum entries of this matrix exceeds a certain threshold (typically 10), then the block is discarded. This could happen because a PMT is broken (which has never occurred) or because there was no signal in any one of the 64 traces.

Fourier components

I compute the Fourier transform of the fluorescence signal in each state and flag anomalous peaks. In particular, I use the MATLAB findpeaks function, which returns the frequency, amplitude, width, and prominence of local maxima in the spectrum. The prominence is

the height of a peak above the nearest local minimum. I throw out any peaks whose prominence exceeds the median prominence by less than a certain factor (typically 10). I then throw out any peaks that correspond to slow-timescale dynamics (less than 10 kHz), for which we don't have a clean expected model of the spectrum. Finally, I ignore any peaks detected within the frequency ranges $n \times f_{\text{ps}} \pm \Delta f_{\text{ps}}$, where $f_{\text{ps}} = 200$ kHz is the polarization switching frequency (at which we obviously expect structure in the fluorescence signal) and $\Delta f_{\text{ps}} = 100$ Hz.

This cut was implemented because we found that noise at 78.9 kHz was contaminating the PMT's in early runs due to broadcast electronic noise from unshielded cables; see Sec. 4.15.1. The default spectrum cut parameters easily flag this problematic frequency component. If an anomalous frequency is detected, the block is discarded.

Absolute fluorescence threshold

The most important cut in Gen. I was the fluorescence threshold. For each state, any group with less than ~ 500 photons per bin pair is excluded from the analysis. More precisely, all computations are performed for every group (e.g., $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ is computed even in the low-signal region), but I only combine data from groups that are "good" in every state when computing averages over group index or across different blocks.

Fractional pulse height threshold

We have seen strong indications that the statistics of asymmetry values is poorly behaved near the wings of the molecule pulse. We believe this is due to molecule velocities that are inconsistent on the trace-to-trace timescale even when considering a fixed time after ablation. Therefore, I implement a cut on the region of the molecule pulse whose amplitude falls below some fractional threshold, typically $\sim 15 - 25\%$. This accounts for a $\sim 5 - 10\%$ reduction in total signal. In practice, this cut is stronger than all others; i.e., the groups that are omitted from the analysis are typically exactly those that fail this cut.

Asymmetry threshold

For each state, any group with unphysically large asymmetries, e.g.,$|\mathcal{A}| > 0.9$, is discarded. This should only occur due to noise in low-signal regions.

Contrast threshold

For each state, groups with low contrast, e.g., $|\mathcal{C}| < 0.8$, are discarded.

Normality check

For each state, we also perform a statistical test of asymmetry data within each group (i.e., the individual asymmetry values used to compute the group-by-group asymmetry) in order to flag highly non-normal data. In ACME I, Brendon O'Leary and Nick Hutzler used a Pearson χ^2 test, while Ben Spaun did not include a normality check. Neither Brendon nor Nick found that this cut significantly affected the result, but I have chosen to implement an analogous cut based on the Shapiro-Wilk test (see [184]), which I will describe below.

All the discussion below is fairly informal, and is intended only to give some conceptual clarity to what questions these statistical tests are asking. It's a good idea to get a feeling for statistical tests by applying them to known distributions. I have confirmed that the p-values are correctly calibrated for normally distributed data (with any mean and standard devation), and that the Shapiro-Wilk test does a reasonably good job of rejecting several types of non-normal data.

With any statistical test, we must set a threshold α that determines how much good data we throw away. Larger values of α allow more aggressive rejection of bad data at the cost of losing more good data. In order for a group to survive in a block, it must independently pass the test for all 16 states. Therefore, if we want to cut a proportion α_{block} of data, then we set a significance threshold of $\alpha = 1 - (1 - \alpha_{\text{block}})^{16} \approx \alpha_{\text{block}}^{16}$ in each state. Since we know that asymmetry values can have a significant slope within a group, I apply the normality test to the residuals of the linear regression of asymmetry values within each group.

Comment on the Pearson χ^2 goodness-of-fit test

The Pearson χ^2 goodness of fit test essentially constructs a histogram of observations with N bins and computes $\chi^2 = \sum_{i=1}^{N} (O_i - E_i)^2 / E_i$, where O_i is the actual number of observations in bin i and E_i is the expected number in bin i, given some model. The result should be drawn approximately from a χ^2 distribution. A p-value, giving the probability that data described by the model would generate a χ^2 value exceeding the actual value, is compared to a threshold α. This means that for a threshold $\alpha = 0.05$ (for example), 5% of data that is actually described by the model will be rejected.

There are a couple of drawbacks to the Pearson test: (i) Most importantly, it assumes reasonably large sample sizes. The rules of thumb are often that one should have $N \geq 5$ bins and at least ~ 5 points per bin. This already means we probably should not use it on data within a group, where there are typically only ~ 20 points. (ii) The result depends somewhat on the way the data is binned. (iii) The threshold α allows one to reliably reject a known quantity of data that is actually described by the preferred model, but it does *not* necessarily reject data that is *not* described by that model. This is ultimately a limitation for all tests of normality, but there exist other tests that much more reliably reject most non-normal data for the same p-value threshold. (iv) The construction of the p-value only allows one to reject data with χ^2 values exceeding some limit, but does not throw out data with extraordinarily low values of χ^2.

Shapiro-Wilk test

We would like to determine whether N values are drawn from a normal distribution. In the Shapiro-Wilk test, the values are ordered from smallest to largest, $y_1, \ldots, y_N$. We then look up the expected ordered values in a sample of N points drawn from a standard normal distribution, $x_1, \cdots, x_N$. (In practice, the "expected" values are obtained via Monte Carlo simulations.) Now imagine creating a scatter plot with points (x_i, y_i). If the values of y_i are drawn from a standard normal distribution, then this scatter plot will show approximately a straight line. If the values of y_i are drawn from a normal distribution with non-zero mean but unit variance, then the line on the scatter plot will simply shift up or down. Finally, if

the values of y_i are drawn from a normal distribution with non-unit variance, the slope of the line will change accordingly.

The Shapiro-Wilk test essentially computes the correlation coefficient between the ordered values $\{y_i\}$ and expected ordered values for a standard normal distribution, $\{x_i\}$. Since the correlation coefficient is unaffected by changes to the offset or slope of a scatter line, it should be large when the data $\{y_i\}$ are drawn from any normal distribution. Ordered data drawn from other distributions will not generally be well-correlated with $\{x_i\}$.

The closely related Shapiro-Francia test uses simple least squares rather than generalized least squares to estimate the slope of the regression line between $\{y_i\}$ and $\{x_i\}$, ignoring the covariances among ordered values [185]. In some cases (depending on which non-normal models should be rejected most powerfully), this can be preferable to the Shapiro-Wilk test.

Molecular pulse width

If the total duration of "good" groups in a block is less than 0.5 ms, then the entire block is cut. This could happen, for example, if a laser becomes unlocked.

3.4   Correlations

We typically log ~ 40 parameters (such as currents through the magnetic field coils, vacuum pressures, room temperature, etc.) in the header during each trace, and we log many other parameters on slower time scales. Further, we compute 64 parity components (including superblock switches) of $\mathcal{A}$, $\mathcal{C}$, ϕ, and ω and we can optionally compute auxiliary quantities like parity components of fluorescecence signals. We therefore have an overwhelming number of pairs of parameters that we could check for correlations (e.g., "is the room temperature correlated with $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$?"). Furthermore, we can examine the autocorrelation of parameters as a function of group index (i.e., time after ablation), block number, or superblock number, as well as correlations between quantities and analysis parameters like fluorescence sub-bin region or group size.

I especially consider correlations between any superblock component of

$$\{\omega, \mathcal{C}\}^{\{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}, \tilde{\mathcal{E}}\tilde{\mathcal{B}}, \tilde{\mathcal{N}}\tilde{\mathcal{E}}, \tilde{\mathcal{E}}, \tilde{\mathcal{N}}\}} \tag{3.32}$$

with any block and superblock component of ω, $\mathcal{C}$, or logging parameters (which have 128 parity components, including the $\tilde{\theta}$ switch). The explicitly listed block components of ω and $\mathcal{C}$ are "well behaved" in the sense that they should not change with respect to any aspect of the experimental configuration. As a contrasting example, $\omega^{\tilde{\mathcal{B}}}$ is "badly" behaved and is expected to change with parameters such as (obviously) the applied magnetic field magnitude. For "diagonal" entries, like the correlation of $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ with itself, I compute the single-step autocorrelation function over superblocks instead of the trivial self-correlation. With ~ 40 logging parameters, this gives us $\sim 80 \times 5,000 \sim 400,000$ correlation coefficients. We're obviously not going to inspect each of these individually, so we need a method to automatically flag anomalously large correlations.

To deal with this, I compute the Pearson correlation coefficient,

$$\rho_{X,Y} = \frac{\text{cov}(X,Y)}{\sigma_X \sigma_Y}. \tag{3.33}$$

Note that the autocorrelation coefficients $\rho_{X,X} \equiv \rho_{X(1,\cdots,N-1),X(2,\cdots,N)}$ must be computed with fewer observations than the correlation coefficients $\rho_{X,Y} \equiv \rho_{X(1,\cdots N),Y(1,\cdots,N)}$. These will be treated in more detail in the next subsection since they are distributed differently than the ordinary correlation coefficients.

The more aggressively we flag potentially significant correlations, the more flags we will see for data that is not truly correlated. As in ACME I, I typically set a significance threshold such that we expect ~ 1 false positive per run analysis. In particular, for a given pair of parameters X and Y whose sample correlation coefficient is $\rho_{X,Y} = r$, we can compute the probability p of generating a sample correlation whose magnitude is at least as large as $|r|$ under the hypothesis that X and Y are drawn from independent normal distributions (equivalently, the pairs (X,Y) are assumed to be drawn from a bivariate normal distribution with vanishing correlation). If the p-value exceeds a pre-determined

108

threshold α, then the pair of variables is flagged as having a possible correlation. The computed p-values will not be accurate for binary or highly discrete variables (electronics lead configurations, STIRAP translation stage position, etc.), so I do not include those in my correlation searches. We use $\alpha \sim 1/N_{\text{correlation coefficients}}$ to maintain a reasonable ability to flag correlations without having to wade through too many false positives.

The p-value is computed assuming the null hypothesis of vanishing correlation, in which case the quantity $t \equiv r\sqrt{\frac{N-2}{1-r^2}} \sim t_{N-2}$ follows a Student t-distribution with $N-2$ degrees of freedom [186, Sec. 16.28]. Therefore, the probability that a sample correlation coefficient with magnitude larger than $|r|$ is observed by chance is $p(|t[r]|) = 2 \times (1 - \text{TCDF}(|t[r]|, N-2))$, where TCDF is the cumulative distribution function for the t-distribution.

The statistical properties of the sample correlation coefficient r are not completely straightforward in general. The statistic r is the maximum-likelihood estimator (MLE) of the true correlation ρ for a bivariate normal distribution, but this is biased toward smaller magnitudes [187].[11] An expression for the minimum-variance unbiased estimator of ρ can be written in terms of a hypergeometric function if needed. However, because we can easily compute correctly calibrated p-values for the null hypothesis (both variables tested are normally distributed and uncorrelated), we can maintain the standard interpretation of our significance testing even with a biased estimator of ρ. The primary consequence of the biased estimator is to reduce the power of the test somewhat (that is, the number of false positives is unaffected but the number of true positives for discovering correlated variables is slightly reduced).

If we add too many logging parameters or parity sums, then we must set $\alpha \approx 0$ (in order to not have to examine an unwieldy number of false positives), impairing our ability to resolve true and potentially important correlations, for example between the EDM channel and some other important channel like $\omega^{\tilde{\mathcal{B}}}$. To deal with this, I set different significance thresholds for different "classes" of correlations. For each of the combinations of parameters in Table 3.1, we set $\alpha \sim 1/N_\rho$.

---

11. The MLE of a random variable is often biased. For example, the MLE estimator of the variance is computed as $\sigma^2_{\text{MLE}} = \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2$ using the same variance with the intuitive normalization factor $1/N$, but the unbiased estimator uses normalization $1/(N-1)$.

X	Y	N_X	N_Y	$N_\rho = N_X N_Y$	$1/N_\rho$
$\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$	$\{\omega,\mathcal{C}\}^{\text{block}}$	1	16	16	6×10^{-2}
	$\{\omega,\mathcal{C}\}^{\text{superblock}}$	1	128	128	8×10^{-3}
	$\{\text{logging}\}^{\text{block}}$	1	~ 500	~ 500	2×10^{-3}
	$\{\text{logging}\}^{\text{superblock}}$	1	~ 4000	$\sim 4 \times 10^3$	3×10^{-4}
$\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\{\text{SB}\}}$	$\{\omega,\mathcal{C}\}^{\text{block}}$	8	16	128	8×10^{-3}
	$\{\omega,\mathcal{C}\}^{\text{superblock}}$	8	128	$\sim 1 \times 10^3$	1×10^{-3}
	$\{\text{logging}\}^{\text{block}}$	8	~ 500	$\sim 4 \times 10^3$	3×10^{-4}
	$\{\text{logging}\}^{\text{superblock}}$	8	~ 4000	$\sim 3 \times 10^4$	3×10^{-5}
$\{\omega,\mathcal{C}\}^{\text{"good"}}$	$\{\omega,\mathcal{C}\}^{\text{block}}$	10	16	160	6×10^{-3}
	$\{\omega,\mathcal{C}\}^{\text{superblock}}$	10	128	$\sim 1 \times 10^3$	1×10^{-3}
	$\{\text{logging}\}^{\text{block}}$	10	~ 500	$\sim 5 \times 10^3$	2×10^{-4}
	$\{\text{logging}\}^{\text{superblock}}$	10	~ 4000	$\sim 4 \times 10^4$	3×10^{-5}
$\{\omega,\mathcal{C}\}^{\text{"good"}\{\text{SB}\}}$	$\{\omega,\mathcal{C}\}^{\text{block}}$	80	16	$\sim 1 \times 10^3$	1×10^{-3}
	$\{\omega,\mathcal{C}\}^{\text{superblock}}$	80	128	$\sim 1 \times 10^4$	1×10^{-4}
	$\{\text{logging}\}^{\text{block}}$	80	~ 500	$\sim 4 \times 10^4$	3×10^{-5}
	$\{\text{logging}\}^{\text{superblock}}$	80	~ 4000	$\sim 3 \times 10^5$	3×10^{-6}

Table 3.1: Data sets for which we search for correlations. There are assumed to be ~ 30 logging parameters decomposed into $\tilde{\theta}$-even and $\tilde{\theta}$-odd parts. The X column is organized from most important to least important parameters to find correlation with, and the Y column consists of larger data sets that we can try to cross-correlate with the quantities of interest. Here, "SB" denotes any superblock parity component, and "good" denotes the most well-behaved block components, $\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}$, $\tilde{\mathcal{E}}\tilde{\mathcal{B}}$, $\tilde{\mathcal{N}}\tilde{\mathcal{E}}$, $\tilde{\mathcal{E}}$, $\tilde{\mathcal{N}}$. Some of the parameters in the Y column will not be normally distributed under some circumstances (e.g., $\omega^{\tilde{\mathcal{B}}}$ is bimodal when a $|\mathcal{B}|$ switch is implemented), in which case the observed sample correlation coefficients will tend to be smaller in magnitude than in the null hypothesis assumes. The column $1/N_\rho$ gives the threshold α required to see one false positive for each type of correlation on average.

A consistent correlation that emerges from this sort of analysis, which we will discuss at length in Sec. 4.9, is between $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ and $|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$, as well as between $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ and $|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}}$.

Note that the p-values are calibrated for normally distributed data (as with almost any standard statistical test). For real data, these calibrations are not going to be particularly accurate for extremely small p-values (e.g., when searching for correlations between any component of a frequency or contrast and any component of a logging parameter, which requires calibrated p-values on the order of 10^{-6}). Therefore, although this machinery is set up to search for extreme correlations among all kinds of parameters, it should be expected to be most useful when we are searching in categories where relatively modest correlations are of interest, with $\alpha \sim 0.01 - 0.1$.

Sample autocorrelation coefficients

As mentioned in the previous section, we consider autocorrelation coefficients for quantities of interest throughout the EDM data set (and any particular experimental run). The single-lag sample autocorrelation coefficient is defined as

$$r_{X,X} = \frac{\sum_{i=1}^{N-1}(x_i - \bar{x})(x_{i+1} - \bar{x})}{\sum_{i=1}^{N}(x_i - \bar{x})^2}, \tag{3.34}$$

where $\bar{x}$ is the sample mean over the entire data set. Unlike for the ordinary correlation coefficient, the sample autocorrelation coefficient will tend to be non-zero even when the underlying data is randomly distributed. Define the deviation from the mean, $d_i \equiv x_i - \bar{x}$, so that the average deviation among other measurements must be nonzero, $E[x_{j\neq i} - \bar{x}] = -\frac{d_i}{N}$, following from $E[x - \bar{x}] = 0$. This suggests that data in a random sample have a negative autocorrelation on average. In fact, more careful analysis shows that $E[r_{X,X}] = -\frac{1}{N}$ for a random sample, while $\text{Var}[r_{X,X}] = \frac{(N-2)^2}{N^2(N-1)}$. It turns out that the distribution of the sample autocorrelation is extremely complicated and–rather surprisingly–there is no consensus about how to test the null hypothesis of vanishing correlation in the underlying distribution. An excellent starting point is [188], which discusses many surprising results that one ought to be aware of. In no particular order: $r_{X,X}$ has nontrivial minimum and maximum possible values that depend on the sample size N (but never exceeding unity in

magnitude); the first four central moments $E[(r_{X,X})^i]$ are given as function of N and the lag (always equal to one in this document); and the efficacy of various approximate hypothesis tests is explored.

Traditionally, an asymptotic normal approximation $\text{Normal}(\mu = 0, \sigma^2 = \frac{1}{N})$ is often used to fit the probability density function $P(r_{X,X})$, but because $P(r_{X,X})$ has a sharp cutoff somewhere within $-1 < r_{X,X} < 1$, this tends not to work very well near the tails where accuracy is most important for hypothesis testing. The results can be improved by using a normal approximation with the exact mean and variance, $\text{Normal}(\mu = -\frac{1}{N}, \sigma^2 = \frac{(N-2)^2}{N^2(N-1)})$, but this suffers from a similar problem in the tails. I've confirmed via simulations that these models don't perform very well for small p-values (i.e., below ~ 0.01).

A preferable model is to use the four-parameter Pearson distribution, which is obtained as the solution to a particular differential equation. Depending on the relationships among parameters, the Pearson distribution is traditionally broken into seven types, corresponding to a normal distribution, generalized beta distribution, gamma distribution, Student t-distribution, etc., including non-standard distributions. We parametrize the Pearson distribution in terms of its mean, variance, skewness, and kurtosis, all of which we can compute for $P(r_{X,X})$ by using the known central moments in [188]. Note that the Pearson distribution is merely a very flexible four-parameter distribution that can be readily computed numerically, and which can be made to match "reasonably" arbitrary statistical distributions, but it is not chosen for any other special properties.[12]

For the sample autocorrelation coefficient with $N \geq 4$, we obtain a Pearson Type I distribution (with $N = 4$ actually a Type II, which is a special case of Type I). This distribution has finite support; i.e., it vanishes beyond a certain range. Thus fitting to this Pearson distribution largely avoids the problems of a normal distribution in the tails, where we want to calibrate very small p-values.[13]

12. John von Neumann supposedly said, "with four parameters I can fit an elephant." One can easily imagine that he had the Pearson distribution in mind.

13. In the special case that the support ranges from $r = 0$ to $r = 1$, the Pearson Type I distribution becomes a β distribution; for this reason, the Pearson Type I distribution is also called a "generalized β distribution." For $N = 3$, the resulting parameters give a Pearson Type IV distribution, which is not related to any standard distribution but is, in any case, supported over the entire real line. My p-value calibrations using a Pearson distribution for $N = 3$ are very poor, as we might expect from this fact.

Using the four-parameter Pearson distribution, we get acceptable p-value calibration for $N \geq 5$ and excellent calibration for $N \geq 10$. This calibration is fairly insensitive to the chosen rejection threshold α. By contrast, the built-in MATLAB autocorrelation test, based on the Ljung-Box Q-test, has a significantly larger miscalibration that depends fairly sensitively on the chosen threshold, as would be expected from the discussion in [188]. The Ljung-Box Q-test is also much slower, probably since the test statistic requires the computation of the autocorrelation function at all possible lags.

Effect of uncertainty

It isn't straightforward to adapt the standard computation of the correlation coefficient for the case in which the quantities being correlated have time-dependent uncertainty, or "heteroscedasticity." This is generically the case at some level because superblocks don't always have the same signal level. Intuitively, we might expect uncertainties not to affect testing of the null hypothesis: if x and y are uncorrelated, then the observed values $X = x + \epsilon_x$ and $Y = y + \epsilon_y$ will also be uncorrelated, where $\epsilon_{x,y}$ represent some uncorrelated measurement errors on any point (x, y). Therefore, a test of the hypothesis that X and Y are uncorrelated is a valid test of the hypothesis that x and y (the true values) are uncorrelated. In the case that $\epsilon_{x,y} \sim \text{Normal}(0, \sigma_{x,y}^2)$ with $\sigma_{x,y}^2$ constant over all measurements, this reasoning turns out to be correct. More discussion along these lines can be found in [189, Sec. 1.3]. In other words, hypothesis testing can be performed just as if there were no measurement uncertainties. The primary effect of these uncertainties is to reduce the power of the test so that non-zero correlations are less likely to be identified as measurement uncertainties grow larger.

This intuitive reasoning breaks down when uncertainties vary from measurement to measurement; see [190, Sec. 9.1.2]. Unfortunately, this is always the case for us due to signal fluctuations. That said, the problems with naive null hypothesis testing appear to be most severe with large sample sizes, which we often don't have in a dataset of interest (e.g., a systematic check).

Nevertheless, I compute naive correlation and autocorrelation coefficients, with the understanding that neglecting to take measurement uncertainties into account will reduce our
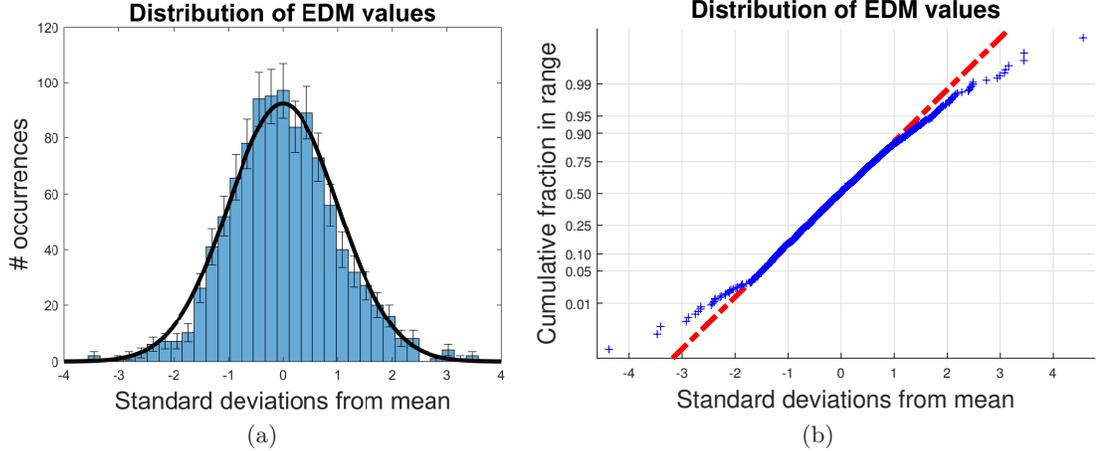
Figure 3.1: Statistics of the EDM data set (1050 superblocks). Data are combined from four magnetic field magnitudes, $\mathcal{B} = 0.7$, 1.3, 2.6, and 26 mG. Because one of the mechanisms for excess noise (beyond shot noise) depends on $\mathcal{B}$, we compute $\chi^2_{\mathrm{red}}(\mathcal{B})$ separately for each value of $\mathcal{B}$, and show distributions of $\left[\omega_i^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} - \langle \omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} \rangle\right]/[\sigma_i \chi^2_{\mathrm{red}}(\mathcal{B}_i)]$, where $\omega_i^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ is the i-th value of the EDM precession frequency, σ_i is its shot-noise uncertainty, and $\mathcal{B}_i$ is the magnitude of the magnetic field for the i-th data point. (a) Histogram of EDM-channel frequencies in units of their expected uncertainty (shot noise rescaled by χ^2_{red}), relative to the weighted mean over the entire data set. Errorbars show the expected variation in the height of each bin of the histogram. A black standard normal distribution is overlayed. The fit is best in the central region, but begins to stray from the observed distribution beyond ± 1.5 standard deviations. (b) Normal probability plot, comparing the cumulative probability distribution of EDM data to that expected for a standard normal distribution (red line). The deviation is largest beyond ± 2 standard deviations.

power to resolve non-zero correlations and, to some extent, cause our p-values to be miscalibrated. We accept all of these drawbacks to avoid excessive complication in the analysis. These considerations are, at any rate, swamped by the non-gaussian distribution of measured values throughout the EDM dataset caused by experimental imperfections (see Sec. 4.15).

3.5 Measurement with non-ideal statistics

The statistical analysis in Secs. 3.2-3.4, for the most part, assumes ideal normally distributed data. We have shown that for an ideal measurement, data are (to a good approximation) normally distributed, and this was seen in ACME I [129]. Unfortunately, in ACME II there are multiple noise sources that can contribute to additional noise, not arising from photon shot noise, to the distribution of $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$. These noise sources will be examined in Sec.

4.15. This noise has two effects: (1) the width of the distribution of $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$, and therefore the uncertainty in the mean value, $\sigma_{\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}$, is increased relative to the shot-noise level; and (2) the values of $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ are not normally distributed. See Fig. 3.1 for the empirical distribution of the EDM data set.

If the noise only caused the uncertainty of $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ to increase, but did not change the distribution of values, then it would be fairly trivial to handle (though obviously still undesirable) using standard methods such as increasing the errorbar by a factor of $\sqrt{\chi^2_{\mathrm{red}}}$ (a measure of how much the errorbars are underestimated). For data that is not normally distributed, more advanced techniques should be used. In particular, for noise sources whose distribution is not well understood, it is appropriate to use non-parametric statistics; i.e., we do not rely on the data being described by any particular distribution at all, though certain reasonableness properties, such as finite variance, are still assumed to hold. We have chosen to use the M-estimator approach, where M stands for "maximum-likelihood." To justify this choice, let's review standard statistical parameter estimation.

3.5.1 Robust parameter estimation

There are two parts to an ordinary statistical problem: (1) How do the data depend on some parameters on average, and (2) how are the residuals of the data about that average distributed? For example, in the simplest linear regression, we suppose (1) the i-th measurement has value $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$, where β_0 and β_1 are parameters (offset and slope), x_i is the value of a variable in the i-th measurement, and ϵ_i are residuals. We then further suppose that (2) $\epsilon_i \sim N(0, \sigma^2)$ is normally distributed with mean 0 and variance σ^2, identically with and independently of the residuals for all other values of i. The goal is to compute, from data $\{y_i\}$, some estimates $\hat{\beta}_0$ and $\hat{\beta}_1$ of the true parameter values β_0 and β_1. The estimates will have uncertainty; i.e., they will not exactly equal β_0 and β_1 for any particular set of results $\{y_i\}$.

In the approach of maximum-likelihood estimation, we ask: what values of $\hat{\beta}_0$ and $\hat{\beta}_1$ maximize the probability of observing the set of results $\{y_i\}$? Suppose we assume that $\beta_{0(1)} = \tilde{\beta}_{0(1)}$. Then the likelihood of observing $\{y_i\}$ is $\mathcal{L} = \Pi_i f(\frac{y_i - \tilde{\beta}_0 - \tilde{\beta}_1 x_i}{\sigma})$, where $f(\epsilon) = \frac{1}{\sqrt{2\pi}} \exp\left[-\epsilon^2/2\right]$ is a standard normal distribution. Note that for this particular distribution,

$\mathcal{L} = \Pi_i f(\frac{y_i - \tilde{\beta}_0 - \tilde{\beta}_1 x_i}{\sigma}) = f(\sum_i \frac{y_i - \tilde{\beta}_0 - \tilde{\beta}_1 x_i}{\sigma})$. Then $2\sigma^2 \ln(\sqrt{2\pi}\mathcal{L}) = -\sum_i (y_i - \tilde{\beta}_0 - \tilde{\beta}_1 x_i)^2$. If the values of $\tilde{\beta}_{0(1)}$ maximize the likelihood $\mathcal{L}$, then they minimize the sum of squared residuals, $\sum_i (y_i - \tilde{\beta}_0 - \tilde{\beta}_1 x_i)^2$. This is a straightforward way to see why least-squares regression is used to obtain estimates $\hat{\beta}_{0(1)}$ of the parameters describing linear data.

The key point for us is that the "sum of squares" rule comes directly from an assumption about the distribution of residuals; and hence, when the distribution of residuals is non-normal, our statistical procedures must be modified. Suppose, for the moment, that the probability density of residuals is $f(\epsilon_i) \propto \exp[-g(\epsilon_i)]$. Then by same reasoning we used for normally distributed residuals, the maximum likelihood estimator minimizes $\sum_i g(\epsilon_i)$. We *must* assume some form for $g(\epsilon_i)$ or we simply cannot use statistics; there is absolutely no way out of this.

This is rather unfortunate since we stated before that we would like to use a "non-parametric" model; i.e., we don't know what form $g(\epsilon_i)$ takes in reality. It turns out that there are certain models for the residuals (i.e., certain functions $g(\epsilon_i)$) are "robust" in the sense that resulting parameter estimates are accurate (on average) and reasonably precise for a wide variety of actual, underlying residual distributions. In other words, our model for the residuals need not be correct as long as the resulting estimates are consistent with the true values. The study of such model distributions belongs to the field of "robust parameter estimation." See [191] for an extremely readable report on the subject; [192] also gives a broad overview, and more technical details are available in [193]. While these statistical techniques are used less widely than the simpler tools designed for gaussian-distributed data, they are widely accepted–see, for example, their application to high-energy physics analysis in [194].

Generally speaking, parameter estimation with robust M-estimators is equivalent to assuming a long-tailed distribution of residuals. As a result, any extreme outlier (e.g., an event at $+10\sigma$ from the central value) is taken as less strong evidence of a shifted mean. More formally, the choice of assumed residuals distribution $f(\epsilon_i)$ is equivalent to a choice of *weights* used in the computation of a weighted mean, $\bar{x} = \sum_i x_i w_i / \sum_j w_j$. In the case of normally distributed residuals with (possibly unequal) variances σ_i^2, the weights are simply $w_i = \sigma_i^{-2}$. When a generic distribution of residuals $f(\frac{\epsilon_i}{\sigma_i s})$ is assumed, where s is some scale

factor of the distribution (e.g., $s = 1$ for a normal distribution), the weights are a function of the residual values.

The particular situation relevant to the ACME II data is not typically covered in detail in standard texts, so let's go through it carefully. We have data points $\{x_i\}$ and associate each value with an estimated uncertainty σ_i. However, due to additional noise in the system, we allow for the possibility of an overall "scale factor" s so that the actual uncertainty in each point is $\sigma_i \to s\sigma_i$. This way, the relative uncertainty among data is constant, but a calculation of s will correct for any underestimation of the overall scale of uncertainty in the data.

We assume a (for now, generic) distribution of residuals $f(\frac{x_i - \bar{x}}{\sigma_i s})$, where $\bar{x}$ is the mean of the data set, which we want to estimate. For shorthand, we define the residual $\epsilon_i \equiv x_i - \bar{x}$. Let us write $f(\frac{\epsilon_i}{\sigma_i s}) = \exp(-g(\frac{\epsilon_i}{\sigma_i s}))$. Then we want to choose an estimate of the mean, $\bar{x}$, such that the likelihood of obtaining the observed data is maximized. As we saw before, this is equivalent to minimizing $\sum g(\frac{\epsilon_i}{\sigma_i s})$ with respect to $\bar{x}$. Then $\frac{\partial}{\partial \bar{x}} \sum g(\frac{\epsilon_i}{\sigma_i s}) = -\sum g'(\frac{\epsilon_i}{\sigma_i s})/(\sigma_i s) = 0$. Let us define $w_i \equiv g'(\frac{\epsilon_i}{\sigma_i s})/(\epsilon_i \sigma_i)$ so that $\sum w_i \epsilon_i = 0$ and, finally, $\bar{x} = \sum w_i x_i / \sum w_i$. Note that for normally distributed residuals where $g(\frac{\epsilon_i}{\sigma_i s}) = \frac{\epsilon_i^2}{2\sigma_i^2 s^2}$, the weights are $w_i = \frac{1}{\sigma_i^2}$ (up to constant factors of s that cancel in the estimation of $\bar{x}$).

Now let us specify s based on the data. For normally distributed data with correctly assigned uncertainties σ_i, the model will match the data for $s = 1$. We therefore want to calculate s such that it corresponds to a robust assessment of the characteristic standard deviation of residuals (normalized to the assumed standard deviations). The conventional way to characterize s is by computing the "median absolute deviation," $\mathrm{median}[|\frac{x_i - \bar{x}}{\sigma_i}|]$, where $\bar{x}$ is itself an estimate of the mean of data points. In particular, we define $s = 1.4826 \times \mathrm{median}[|\frac{x_i - \bar{x}}{\sigma_i}|]$ so that normally distributed residuals, with correctly assigned uncertainties σ_i, have an expected value of $s = 1$. The median absolute deviation is a more robust estimate of the scale of a distribution (i.e., more insensitive to disturbance by extreme outliers) than the sample standard deviation.

There is a slight problem with this prescription: the estimate of s requires an assumed

value of $\bar{x}$, and the estimate of $\bar{x}$ requires an assumed value of s[14]. For this reason, we compute an M-estimator iteratively: use ordinary gaussian statistics to obtain an estimate of $\bar{x}$, then compute an estimate of s. Use this to revise the estimate of $\bar{x}$, which is used in turn to revise the estimate of s. We repeat until the value of $\bar{x}$ changes by an amount less than some tolerance (e.g., a fractional change of 10^{-4} for consecutive iterations). This typically takes only a few iterations.

To compute the ACME II EDM result, we used the Huber weights,

$$
w_i = \begin{cases}
\frac{1}{\sigma_i^2} \times 1 & \left| \frac{\epsilon_i}{\sigma_i s} \right| < a \\
\frac{1}{\sigma_i^2} \frac{a}{\left| \frac{\epsilon_i}{\sigma_i s} \right|} & \left| \frac{\epsilon_i}{\sigma_i s} \right| \geq a.
\end{cases}
\tag{3.35}
$$

Here, we take a typical value of $a = 1.345$. This means that we *begin* to weight observations less strongly when residuals lie outside $1.345\sigma_i s$, where $\sigma_i s$ is an empirically rescaled uncertainty for the i-th obervation. For this particular distribution, no data are strictly "cut"; i.e., $w \neq 0$ for all values of the residuals $x_i - \bar{x}$.

Another common choice of weights in robust parameter estimation is the Tukey bisquare:

$$
w_i = \begin{cases}
\frac{1}{\sigma_i^2} \left(1 - \left(\left| \frac{\epsilon_i}{\sigma_i s} \right| / a \right)^2 \right) & \left| \frac{\epsilon_i}{\sigma_i s} \right| < a \\
0 & \left| \frac{\epsilon_i}{\sigma_i s} \right| \geq a,
\end{cases}
\tag{3.36}
$$

where for this distribution the conventional parameter value is $a = 4.685$. Note that for this choice of weights, extreme outliers (residuals greater than $4.685\sigma_i s$) are not included in the computation of $\bar{x}$ at all.

In both of these cases, the value of a is chosen so that the resulting estimates of $\bar{x}$ are "95% efficient" in the case of perfectly gaussian data: in other words, the variance of estimated values of $\bar{x}$ using least squares (the optimal procedure for perfectly gaussian data) is 5% smaller than the variance of estimated values using the robust M-estimators. However, for a wide variety of non-gaussian data, the M-estimators are *more* efficient than least squares; i.e., estimates using the naive least squares have higher variance in such cases. For

---

14. Except in the model of normally distributed residuals, where the scale factor completely drops out of the weights w_i!

both the Huber and Tukey bisquare weights, the results of gaussian statistics are recovered when $a \to \infty$. We can generally interpret the parameter a as the characteristic scale (in units of standard deviations) where the associated distributions of residuals deviates significantly from a normal distribution.

We have found for the EDM data set that our results are consistent (within 10% of both the mean value and errorbar) when computed using weighted least squares (with uncertainty scaled by $\sqrt{\chi^2_{\mathrm{red}}}$), Huber weights, Tukey bisquare weights, or a trimmed mean[15].

3.5.2   Uncertainty of the M-estimator mean: bootstrapping

Once we've obtained the mean (e.g., of the EDM value) using a robust M-estimator, we must compute the associated uncertainty. One reliable way to do this for non-gaussian data is the bootstrap technique. See [196] for a broad overview, and [197] for the seminal paper.

To understand the bootstrap approach, let's review the interpretation of the "usual" uncertainty in the mean. If we draw N values $\{x_i\}$ from a normal distribution with mean $\bar{x}$ and standard deviation σ, then we will obtain some estimate of the mean $\tilde{x}$, where $\langle \tilde{x} \rangle = \bar{x}$ but typically $\tilde{x} \neq \bar{x}$ in any particular set of measurements. The uncertainty in the mean essentially tells us: if we draw another set of N samples from the same distribution, how far is the new sample mean, $\tilde{x}_{\mathrm{new}}$, likely to deviate from the observed sample mean $\tilde{x}_{\mathrm{obs.}}$? Specifically, what is the standard deviation of observed values of $\tilde{x}$ over many, many instances where we take N samples from the distribution?

In the simple case of a gaussian distribution with standard deviation σ, the uncertainty in the mean is $\sigma/\sqrt{N}$. However, we run into two problems with realistic data: (1) the true standard deviation of the distribution, σ, is not known, and (2) the form of the distribution is not typically known. Therefore, we may not be able to apply simple formulas to calculate the uncertainty in $\tilde{x}$–and of course we cannot typically check $\tilde{x}$ empirically by recollecting entirely independent data sets many times.

To deal with this, we model the underlying distribution by assuming that the data

---

15. This is another method of robust parameter estimation where some fraction of highest and lowest values are "thrown out" before estimating the mean. Note that the YbF EDM experiment used this approach[122]; see [195] for details.

are drawn from by the *observed* distribution of data. In particular, we use the "empirical distribution function," which is a step-wise continuous cumulative distribution function that increases by $\frac{1}{N}$ at the value of every observed data point. For the empirical distribution function of the EDM data set (however, without lines drawn to join data points), see Fig. 3.1(b).

Now assuming that the actual data are drawn from the empirical distribution, we can simulate many "independent" experiments. Specifically, we can draw a "bootstrap sample" of N values $\{x_i'\}$, taken (with replacement) from the empirical distribution, and then we compute the sample mean of these values, also called a bootstrap mean, $\tilde{x}'$. We repeat, drawing another N values $\{x_i''\}$ and compute $\tilde{x}''$. After doing this for many bootstrap samples (typically a few thousand), we can examine the "bootstrap distribution," i.e., the distribution of bootstrap means. We assume that the uncertainty in the observed mean is well-described by the width of the distribution of bootstrap means. In particular, we define the $\pm 1\sigma$ uncertainty interval of our EDM result to be the interval that contains 68.27% of bootstrap samples around the mean of the bootstrap distribution. For the EDM dataset, the bootstrap distribution is, to a good approximation, a normal distribution.

Note that our "measured" mean is still the M-estimator value of the mean from the empirical dataset, *not* the mean value of the bootstrap distribution.

As mentioned in Sec. 3.5.1, the uncertainty assigned in this way is reasonably insensitive to the method chosen to calculate the mean (i.e., least squares scaled by $\sqrt{\chi^2_{\text{red}}}$, Huber weights, Tukey bisquare weights, or trimmed mean).

Chapter 4

Systematic errors and excess noise

A piston that had been thrusting

left-right, left-right, for millennia

suddenly began shifting right-left.

Nothing broke, but everything changed.
2920, The Last Year of the First Era

In this chapter, I will review how we search for systematic errors, how they can generally arise given our experimental protocol, and describe systematic errors that were discovered in the ACME II measurement. All such effects were suppressed to a level below the statistical sensitivity. Further, I will consider mechanisms of noise that increased the statistical errorbar beyond the shot-noise level.

4.1 Measuring a systematic error

A systematic error is any effect that contributes an offset to $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ other than the electron EDM interaction. All systematic errors should arise from "normal" physics, i.e., they are not actually T-violating. They therefore cannot arise solely from the internal dynamics of the ThO molecules, but instead must couple to some "laboratory" parameters such as external fields or characteristics of our lasers. A list of parameters varied in our search for systematic errors is given in Table 4.1.

As a concrete example, suppose we wish to check the dependence of the nominal EDM

Category I parameters	Limit $< \sigma_{\text{stat.}}$
Magnetic fields	
$\mathcal{B}$-field gradients (nr and $\tilde{\mathcal{B}}$): $\frac{\partial \mathcal{B}_z}{\partial z}, \frac{\partial \mathcal{B}_z}{\partial y}, \frac{\partial \mathcal{B}_x}{\partial x}, \frac{\partial \mathcal{B}_y}{\partial y}, \frac{\partial \mathcal{B}_y}{\partial x}, \frac{\partial \mathcal{B}_z}{\partial x}$	✓
Non-reversing $\mathcal{B}$-fields: $\mathcal{B}_z^{\text{nr}}$	✓
Transverse $\mathcal{B}$-fields: $\mathcal{B}_x^{\text{nr}}, \mathcal{B}_y^{\text{nr}}$	✓
Transverse $\mathcal{B}$-fields: $\mathcal{B}_x^{\tilde{\mathcal{B}}}, \mathcal{B}_y^{\tilde{\mathcal{B}}}$	✗
$\tilde{\mathcal{E}}$-correlated $\mathcal{B}$-field: $\mathcal{B}_z^{\tilde{\mathcal{E}}}$	✓
Electric fields	
Non-reversing $\mathcal{E}$-field: $\mathcal{E}^{\text{nr}}$	✓
Field plate ground voltage offset	✓
Laser detunings	
Detuning of refinement and readout lasers: $\Delta_{\text{ref}}, \Delta_{\text{read}}$	✓
1-photon, 2-photon detunings of STIRAP lasers	✓
$\tilde{\mathcal{P}}$-correlated detuning: $\Delta^{\mathcal{P}}$	✓
$\tilde{\mathcal{N}}$-correlated detuning: $\Delta^{\mathcal{N}}$	✓
Detuning of rotational cooling lasers	✗
Laser powers	
$\tilde{\mathcal{N}}\tilde{\mathcal{E}}$-correlated power: $P^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$	✓
Power of refinement and readout lasers: $P_{\text{ref}}, P_{\text{read}}$	✗
$\tilde{\mathcal{N}}$-correlated power: $P^{\tilde{\mathcal{N}}}$	✗
$\tilde{\mathcal{P}}$-correlated power: $P^{\tilde{\mathcal{P}}}$	✗
Readout X- and Y-dependent laser power	✗
Laser pointings/position along $\hat{x}$	
Pointing change of the refinement and readout lasers	✗
Readout X- and Y-dependent laser pointing	✗
Position of refinement beam along $\hat{x}$	✓
Molecular beam clipping	
Clipping of the molecular beam along $\hat{y}$ and $\hat{z}$	✓

Category II parameters
Experiment timing
Readout X and Y polarization switching rate
Allowed settling time between block switches
Analysis
Signal size cuts, magnitude cuts, contrast cuts
Spatial dependence of fluorescence recorded by the 8 PMTs
Variation with time within the molecular pulse
Variation with time within the X and Y polarization cycle
Search for correlations with all ω and $\mathcal{C}$ parity components
Search for correlations with auxiliary monitored parameters
Four sets of analysis code by different people

Table 4.1: Parameters varied to search for systematic errors. Category I parameters are varied far from typical values. Category II parameters have no ideal value, but can be still be varied to search for unexpected effects. Second column: not all parameters have enough range (and/or integration time) to set a limit below the statistical sensitivity.
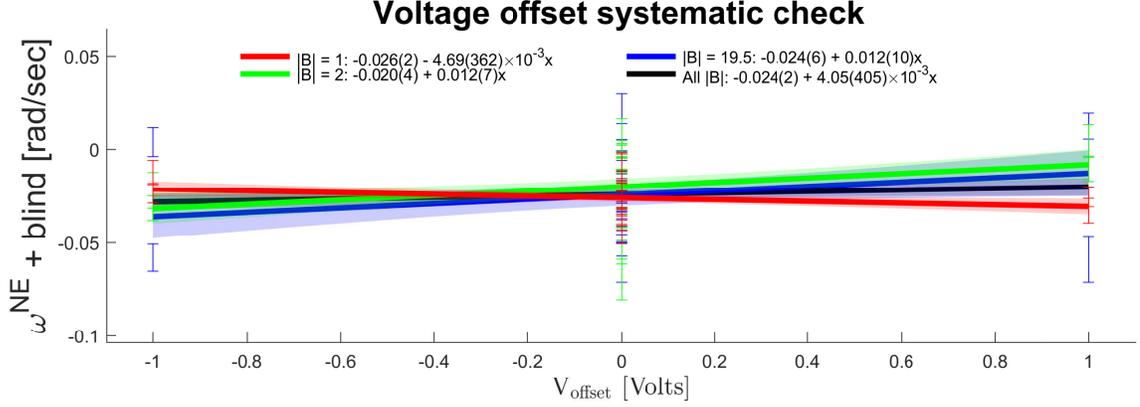
Figure 4.1: A benign systematic error check showing the dependence of $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ on a common voltage offset to both field plates, V_{offset}. We measure the slope separately for multiple values of the applied magnetic field $\mathcal{B}$ (here with 1, 2, and 19.5 mA applied to the field coils, corresponding to $\approx$1.3, 2.6, and 26 mG), and also compute an overall slope based on combining all data. The EDM channel is shown with the blind, so only the slopes contain physically meaningful information. Partially transparent bands show 1σ uncertainty on the fits. The slopes are consistent with 0.

value against an overall offset voltage of the electric field plates, common to both plates so the applied electric field is unchanged, V_{offset}. In this case, the ideal value is $V_{\text{offset}} = 0$ (even though we don't have a model for why a "non-ideal" value would be problematic in this case!). We measure the actual value, and variations in that value over time, using a digital multimeter connected to the field plate leads to obtain a typical deviation from the ideal value $|\Delta V_{\text{offset}}| < 20$ mV. We then deliberately apply an imperfection much greater than the typical value (when possible), e.g., $V_{\text{offset}} \to 1000$ mV and measure the EDM under these non-ideal conditions. By comparing to the EDM obtained under ideal conditions, we can infer a "systematic slope" $\frac{d\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{dV_{\text{offset}}}$. The EDM shift associated with this slope under ordinary conditions is $d\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}_{V_{\text{offset}}} = \frac{d\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{dV_{\text{offset}}}\Delta V_{\text{offset}}$. Our uncertainty in this systematic shift is computed by propagating the uncertainty in both the slope and measurement of ΔV_{offset}, using standard error propagation.

The general approach always follows this model: for a parameter p, we must measure the ordinary deviation from ideal conditions, Δp, as well as a systematic slope $\frac{d\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{dp}$ computed by exaggering p beyond its typical range under normal conditions. We then compute a systematic shift

$$d\omega_p^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = \frac{d\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{dp}\Delta p \qquad (4.1)$$

and compute an uncertainty in the shift by propagating uncertainties in the slope $\frac{d\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{dp}$ and offset Δp. The slope $\frac{d\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{dp}$ is calculated via linear regression with data under both ideal conditions and with a deliberately large value of p. The uncertainty in the slope is computed by propagating uncertainties using standard methods, and rescaling the resulting error bars by $\sqrt{\chi^2_{\text{red.}}}$ computed from the residuals of the linear fit (see Fig. 4.1).

The characteristic deviation from ideal conditions (which can be either an average value or a limit, depending on the parameter) is computed either from ordinary experimental data or from auxiliary measurements. For example, we will see that for $p = \mathcal{E}^{\text{nr}}$, the non-reversing component of the applied electric field, the typical deviations from ideal conditions $\Delta\mathcal{E}^{\text{nr}}$ are measured using microwave spectroscopy. On the other hand, when we consider a possible leakage of $\omega^{\tilde{\mathcal{E}}}$ into $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ due to a small difference in g-factors between the $\tilde{\mathcal{N}} = \pm 1$ states, we measure the average $\Delta\omega^{\tilde{\mathcal{E}}} = \langle\omega^{\tilde{\mathcal{E}}}\rangle$ from the EDM data set.

A list of varied parameters is shown in Table 4.1. What we would like is that (1) the systematic shift is consistent with zero, and (2) the uncertainty in the systematic shift is smaller than the statistical sensitivity of the ACME II EDM measurement, $\delta\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}_{\text{stat.}} = 373\,\mu\text{rad/sec}$.

Most of this chapter will be concerned with parameters for which desideratum (1) fails: sometimes, we unexpectedly found that a systematic shift is not consistent with zero. In this case, we must understand the mechanism for the non-zero systematic slope, $\frac{d\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{dp}$, and suppress it until we are confident that its corresponding systematic shift $d\omega_p^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ is significantly smaller than $\delta\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}_{\text{stat.}}$.

Furthermore, we cannot always achieve (2): that is, we do not always set a limit on a systematic shift that is smaller than the statistical uncertainty of the EDM measurement. To see why, suppose we deliberately apply an experimental imperfection with magnitude dp, while Δp is the degree of the imperfection under ordinary conditions. Further, suppose that we take data with a deliberately large value of p for duration T_p, while the EDM dataset under ideal conditions is collected over a duration T. Then the statistical uncertainty in

the systematic shift is approximately $\delta\omega_p^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} \sim \sqrt{\frac{T}{T_p}}\frac{\Delta p}{dp}\delta\omega_{\text{stat.}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$.

In reality, we spend several months collecting EDM data under ideal conditions in order to obtain $\delta\omega_{\text{stat.}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = 373\,\mu\text{rad/sec}$, and it is not practical to collect this much data for every systematic check parameter p. As a result, in cases where $|\frac{\Delta p}{dp}| \lesssim 10$, the statistical uncertainty in the systematic shift $d\omega_p^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ is typically comparable to or larger than the statistical uncertainty in the nominal EDM-channel value $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$. For example, we check that the nominal EDM measured does not depend on laser pointings, but it is impractical to use $10\times$ the "ordinary" laser pointing misalignment ($\approx 200\,\mu\text{rad}$). In particular, misaligning the laser beams to such a large degree would significantly change the addressed Doppler profile and the position of the preparation and readout of the molecular population. Any data obtained under these dramatically different experimental conditions would be of questionable relevance to the ordinary EDM measurement. See the second column of Table 4.1 for parameters where a limit on the systematic shift was observed to be smaller than the statistical sensitivity of the EDM measurement.

We also distinguish parameters according to two categories. Category I parameters have an ideal value in the experiment (zero in most cases). The common voltage offset to the electric field plates is a Category I parameter. On the other hand, Category II parameters have no specially ideal value. For example, the rate of polarization switching between X and Y readout lasers is essentially arbitrary (provided it is sufficiently fast that the excitation of every molecule is saturated for both lasers). We normally use 200 kHz but also check for an EDM shift when using a 100 kHz polarization switching rate. We furthermore regard all of our analysis parameters as being in Category II: for example, there is no *a priori* ideal value for the number of asymmetry values used to compute a grouped asymmetry, but we do check that our EDM result is robust against different reasonable choices for such a value.

For most systematic error mechanisms where a non-zero systematic slope $\frac{d\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{dp}$ was observed, we interleave "intentional parameter variations" (IPVs) with data taken under ideal conditions. During an IPV measurement, some parameter is deliberately applied with magnitude far beyond its ordinary conditions, $dp \gg \Delta p$. This allows us to measure the systematic effect under nearly identical conditions as used in the EDM measurement (e.g., we take data both with and without IPVs on the same day, with the same optics alignment,

etc.). The central EDM value, before subtracting any systematic error shifts, is computed only from data taken under ordinary conditions. As we will see, we use five distinct IPVs. In the notation of Fig. 2.3: $a = P_{\text{ref}}$, where the power of the refinement beam is tuned to zero (via a mechanical beam block); $b = \mathcal{E}^{\text{nr}}$, where a large non-reversing component of the electric field (150-300 mV/cm) is applied; $c = P^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/P^{\text{nr}}$, where a large correlation of the refinement laser power with the EDM switches $\tilde{\mathcal{N}}\tilde{\mathcal{E}}$ is applied ($P^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = 0.1 \times P^{\text{nr}}$); $d = \omega_{\text{ST}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$, where a large EDM-correlated precession frequency (~ 5 rad/sec) associated with the STIRAP state preparation is induced; and $e = \frac{\partial \mathcal{B}}{\partial z}$, where a magnetic field gradient of ~ 1 mG/cm is applied in the interaction region. We will see the reasons for each of these intentional parameter variations in the following sections.

4.2   A classification of systematic error mechanisms

In this section, we will step back and consider an abstract description of our experiment to see that there are two distinct classes of systematic errors that can arise.

Let the position-dependent detection efficiency be $\epsilon(\vec{x})$, where $\vec{x}$ is the position of a molecule when it is probed, the population density of the molecules that radiate photons in the probe region be $\rho(\vec{x})$, and the average accumulated phase for a molecule that is detected at position $\vec{x}$ be $\phi(\vec{x})$. The average measured phase in a particular state is then

$$\bar{\phi} = \frac{\int dV \, \epsilon(\vec{x})\rho(\vec{x})\phi(\vec{x})}{\int dV \, \epsilon(\vec{x})\rho(\vec{x})}, \tag{4.2}$$

where the integral runs over the entire detection volume V. In the ideal case that the phase is independent of position, $\phi \neq \phi(\vec{x})$, we recover the simple result $\bar{\phi} = \phi$. Alternatively, in the ideal case that the detected population density is constant in some finite occupied volume V and zero elsewhere, $\rho\epsilon \neq \rho(\vec{x})\epsilon(\vec{x})$, we also recover the intuitive result $\bar{\phi} = \int dV \, \phi(\vec{x})/V$. Any parity sums of state-by-state measured phases will thus behave well in these two cases.

For the nearly-general case, suppose only that $\epsilon(\vec{x}) = \epsilon^{\text{nr}}(\vec{x})$, so that the detection efficiency of a photon emitted at position $\vec{x}$ is independent of the experimental state[1]. For

---

1. This assumption breaks down badly when the $\tilde{P}$ switch is included because the spatial distribution of emitted photons is $\tilde{\mathcal{P}}$-dependent, but this caveat isn't essential to the main point of this discussion.

simplicity, I'll only explicitly include the non-reversing and $\tilde{\mathcal{N}}\tilde{\mathcal{E}}$-odd components of any quantities here, and I will keep position-dependence of ρ, ϵ, and ϕ implicit. The measured phase in some state of the experiment is then computed as

$$\bar{\phi} = \frac{\int dV \, \epsilon \left[\rho^{\mathrm{nr}} \phi^{\mathrm{nr}} + \tilde{\mathcal{N}}\tilde{\mathcal{E}} \rho^{\mathrm{nr}} \phi^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} + \tilde{\mathcal{N}}\tilde{\mathcal{E}} \rho^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} \phi^{\mathrm{nr}} + \rho^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} \phi^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} \right]}{\int dV \, \epsilon \left[\rho^{\mathrm{nr}} + \tilde{\mathcal{N}}\tilde{\mathcal{E}} \rho^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} \right]}. \tag{4.3}$$

We define $N^x \equiv \int dV \, \epsilon \rho^x$, where here $x = \mathrm{nr}$ or $\tilde{\mathcal{N}}\tilde{\mathcal{E}}$, and Taylor expand the denominator to second order in $n^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} \equiv \frac{N^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{N^{\mathrm{nr}}}$. Looking forward, I'll also define $\rho^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = n^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} \rho^{\mathrm{nr}} + \delta\rho^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ so that by definition $\int dV \epsilon \delta\rho^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = 0$. Further, let $\phi^x = \phi_0^x + \delta\phi^x$ such that $\int dV \epsilon \rho^{\mathrm{nr}} \delta\phi^x = 0$. We therefore interpret $\delta\rho^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ as the correlated *shape* of the population distribution, $n^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ as the state-correlated relative population distribution amplitude, and $\delta\phi^x$ as the position-dependent part of the (possibly state-dependent) phase. After some algebra, we find to second order in small quantities $\delta\phi^x$, $\delta\rho^x$, and $n^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ that

$$\bar{\phi}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = \frac{\int dV \, \epsilon (\rho^{\mathrm{nr}} \phi_0^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} + \delta\rho^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} \delta\phi^{\mathrm{nr}})}{N^{\mathrm{nr}}}. \tag{4.4}$$

The first term, $\int dV \, \epsilon \rho^{\mathrm{nr}} \phi_0^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} / N^{\mathrm{nr}}$, gives the overall $\tilde{\mathcal{N}}\tilde{\mathcal{E}}$-correlated phase averaged over the molecular population. This term contains both the EDM-dependent phase and a large class of systematic errors. I refer to systematic errors arising in this way as Class I systematic errors.

The second term, $\int dV \, \epsilon \delta\rho^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} \delta\phi^{\mathrm{nr}} / N^{\mathrm{nr}}$, describes a distinct class of systematics, in which a *position-dependent* correlated population distribution $\delta\rho^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ couples to a *position-dependent* uncorrelated phase $\delta\phi^{\mathrm{nr}}$. I refer to systematic errors arising in this way as Class II systematic errors.

We see that any possible correlation in the population distribution *amplitude*, $n^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$, has no effect on the measured phase due to our normalization scheme. Terms with only one of $\delta\rho^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ or $\delta\phi^x$ integrate to 0 by construction. Of course, we have defined our distributions so there is no $\delta\rho^{nr}$ to couple to $\delta\phi^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$.

4.2.1 Magnetic field gradients to probe Class II systematic errors

Applied magnetic fields that do not reverse with any switches can impart very large non-reversing phases ϕ^{nr}, making magnetic field gradients (e.g., $\frac{\partial \mathcal{B}_z}{\partial z}$) a powerful way to probe the second class of systematic errors: any spatial dependence of the population distribution that is correlated with the $\tilde{\mathcal{N}}\tilde{\mathcal{E}}$ switches will couple to the artificially large spatially-dependent phase $\delta\phi^{\mathrm{nr}}$ to produce a systematic shift $\bar{\phi}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$. Thus any systematic of the second class will be observable with a large enough applied magnetic field gradient. To make this more concrete, suppose we apply a constant magnetic field gradient such as $\mathcal{B} = \mathcal{B}_0 + \frac{\partial \mathcal{B}}{\partial z} z$. Then the accumulated phase, as a function of position in the readout beam, will be approximately $\phi(z) = -g\mu_B\tau(\mathcal{B}_0 + \frac{\partial \mathcal{B}}{\partial z} z) = \phi_0 + \delta\phi(z)$, with $\delta\phi(z) = \frac{\partial \phi}{\partial z} z$. Then the contribution from the Class II systematic error is $\bar{\phi}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = \frac{\partial \phi}{\partial z} \int dV \, \epsilon \delta\rho^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} z / N^{\mathrm{nr}}$. We can compute an EDM-correlated center of mass, $z_{\mathrm{CM}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = [\int dV \, \epsilon\rho z / \int dV \, \epsilon\rho]^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = \int dV \, \epsilon\delta\rho^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} z / N^{\mathrm{nr}}$, so we can express the systematic phase shift as $\bar{\phi}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = \frac{\partial \phi}{\partial z} z_{\mathrm{CM}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$. If instead a linear gradient is applied along the y-direction, then $\bar{\phi}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = \frac{\partial \phi}{\partial y} y_{\mathrm{CM}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$. Thus we can interpret a Class II systematic error, in the special case of a constant phase gradient, as the product of a non-reversing phase gradient and an $\tilde{\mathcal{N}}\tilde{\mathcal{E}}$-correlated movement of the center of mass of the detected molecular population. Note that in a full treatment, we could also have terms with other switch behavior, e.g., $\bar{\phi}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = (\frac{\partial \phi}{\partial z})^{\tilde{\mathcal{B}}} z_{\mathrm{CM}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}}$, etc., although we have seen no evidence of such contributions in ACME II.

We can monitor $\delta\rho^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ indirectly by analyzing the fluorescence collected in each of the 8 PMTs. For example, we can compute a total fluorescence signal F from PMTs positioned along $+z$ and compare to the total fluorescence signal from PMTs positioned along $-z$. This can be useful because the PMTs along $+\hat{z}$ preferentially detect photons from molecules along $+\hat{z}$, and likewise for $-\hat{z}$. We denote half the difference between these signals by $F^{\tilde{z}}$, consistent with our usual parity sum notation. Further, we can examine the dependence of $F^{\tilde{z}}$ on the experimental state to see whether the center of mass of the molecules shifts under different experimental conditions. In the parity sum notation, we consider whether or not $F^{\tilde{z}\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = 0$. Under conditions where Class II systematic errors are significant, we do see that $F^{\tilde{z}\tilde{\mathcal{N}}\tilde{\mathcal{E}}} \neq 0$, providing strong evidence of an $\tilde{\mathcal{N}}\tilde{\mathcal{E}}$-correlated shift in the molecular

center of mass that is probed.

No systematic errors from this second class, involving spatially-dependent correlated populations combined with spatially-dependent uncorrelated phases, were observed in ACME I.

4.3 A perturbative model of Class I systematic errors

We will spend a while longer considering systematic errors in the abstract before describing specific systematic errors observed in ACME II. In particular, we will consider a fairly general framework of time-dependent perturbation theory describing the evolution of the molecular state. I've found this perspective useful in checking systematic error models; for example, see Appendix D for order-of-magnitude estimates of various AC Stark and AC Zeeman effects using the framework developed in this section.

Time-dependent perturbation theory

First, let's review time-dependent perturbation theory. Let a time-dependent Hamiltonian be given by $H = H_0 + V(t)$, where the eigenstates and eigenvalues of H_0 are known. Denote the eigenstates by $|n\rangle$ and their eigenvalues by E_n. In the absence of a perturbation, $V(t) = 0$, we know that $|\psi(t)\rangle = \sum_n c_n e^{-iE_n t}|n\rangle$, where the values of c_n are specified by the initial condition. When $V(t) \neq 0$, the coefficients are perturbed, $c_n \rightarrow c_n(t)$. Time-dependent perturbation theory is typically formulated for the special case in which some $c_i(t = 0) = 1$ and $c_{n \neq i}(t = 0) = 0$. Unlike in standard time-independent perturbation theory, where the perturbed eigenstates and eigenvalues are computed (with no reference to initial conditions), in this section we will directly compute the time-evolution of a specific initial state given the perturbing Hamiltonian. To make this explicit, I write $|\psi_i(t)\rangle = \sum_n c_{ni}(t)e^{-iE_n t}|n\rangle$, where the i subscript refers to the initial state. Because the time-evolution operator is linear in state vectors, a state that is not initially in an eigenstate of the unperturbed Hamiltonian, $|\psi(t = 0)\rangle = \sum_i C_i |i\rangle$, can simply be written $|\psi(t)\rangle = \sum_i C_i \sum_n c_{ni}(t)e^{-iE_n t}|n\rangle$.

The standard results of time-dependent perturbation theory, up to second order in $V(t)$, are [25, Sec. 5.6]:

$$
\begin{aligned}
c_{ni}^{(0)}(t) &= \delta_{ni} \\
c_{ni}^{(1)}(t) &= -i \int_0^t dt'\, V_{ni}(t') e^{i\omega_{ni}t'} \\
c_{ni}^{(2)}(t) &= -\sum_{k\neq i} \int_0^t dt''\, V_{nk}(t'') e^{i\omega_{nk}t''} \int_0^{t''} dt'\, V_{ki}(t') e^{i\omega_{ki}t'},
\end{aligned}
\tag{4.5}
$$

where $\omega_{nm} = E_n - E_m$. These results implicitly assume that $V(t)$ is entirely off-diagonal. It is not particularly difficult to show that in the general case, the diagonal perturbations modify these results by everywhere substituting $E_n t \to E_n t + \int_0^t dt'\, V_{nn}(t')$. In the specific cases I have examined, any diagonal perturbations correct the standard results only at a higher order suppressed by $\sim \frac{|V|}{\Delta}$, where $|V|$ denotes a characteristic perturbation energy and Δ is a characteristic energy splitting. Therefore, I neglect any diagonal time-dependent perturbations. Of course, a time-*in*dependent diagonal perturbation can be trivially absorbed into H_0.

Note that in many previous treatments of systematic errors in ACME (e.g., throughout [144]), the first-order corrections to eigenstates are computed, followed by the associated energy shifts that appear at second order in the perturbing Hamiltonian. This is ultimately equivalent to computing coefficients $c_{nn}^{(2)}(t)$ via time-dependent perturbation theory, but omits the second-order interactions between $M = \pm 1$ levels in our measurement, $c_{M(-M)}(t)$. As we will see, this omission is more or less justified.

Evolved state with time-independent perturbation

We can gain some insight by using the formalism of time-dependent perturbation theory for a time-independent perturbation $V \neq V(t)$. As stated previously, any time-independent diagonal perturbation can be absorbed into H_0, so I assume that $V_{mm} = 0$ for all m. I further assume that all "single-step" detunings ω_{nm} are large compared to the characteristic measurement frequency $1/t$. Note, however, that "two-step" detunings $\omega_{nm} - \omega_{mi} = \omega_{ni}$ might be still small (e.g., the splitting between $M = \pm 1$ states). Direct calculation yields corrections to the initial state amplitude,

$$\begin{aligned}
c_{ni}^{(1)}(t) &= \frac{V_{ni}}{\omega_{ni}} \\
c_{ii}^{(2)}(t) &= \sum_m i\frac{|V_{mi}|^2}{\omega_{mi}}t \\
c_{ni}^{(2)}(t) &= \sum_m \frac{V_{nm}V_{mi}}{\omega_{mi}}t\left(\frac{\exp(i\omega_{ni}t)-1}{\omega_{ni}t} + \frac{1}{\omega_{nm}t}\right).
\end{aligned} \qquad (4.6)$$

The first-order perturbation, aside from very rapidly oscillating transients $\sim \exp(i\omega_{ni}t)$ that I have neglected here, saturates to have amplitude $\sim |V|/\Delta$. The second-order perturbation behaves rather differently depending on whether we consider the perturbed amplitude of the initial eigenstate, $c_{ii}^{(2)}$, or the mixture of some new eigenstate of the unperturbed Hamiltonian, $c_{ni}^{(2)}$. Note that $c_{ii}^{(2)}(t)$, to leading order, can be interpreted as the phase accumulated due to a perturbed energy, $c_{ii} \approx 1 + i\left(\sum_m \frac{|V_{mi}|^2}{\omega_{mi}}\right)t \approx \exp\left[i\left(\sum_m \frac{|V_{mi}|^2}{\omega_{mi}}\right)t\right] \equiv \exp\left[i\left(\sum_m E_{mi}^{(1)}\right)t\right]$. This is the second-order perturbation that has typically been considered in previous treatments of systematic errors (e.g., Amar Vutha's exploration of geometric phases [144, Sec. B] and Brendon O'Leary's estimates of AC Stark/Zeeman shifts [149, Sec. 4.4.2]).

However, terms of the form $c_{ni}^{(2)}(t)$ may be non-negligible in the case that the states $|i\rangle$ and $|n\rangle$ under consideration are near-degenerate, and in particular when $\omega_{ni} \sim t^{-1}$ or smaller. (The second term is of order $\frac{|V|^2}{\Delta^2}$ and will typically be negligible for realistic perturbations relevant to the ACME measurement.) In the $\omega_{ni} \ll t^{-1} \ll \omega_{nm}$ limit, we have $c_{ni}^{(2)} \to \sum_m i\frac{V_{nm}V_{mi}}{\omega_{mi}}t$, similar in form to $c_{ii}^{(2)}(t)$. In the regime $t^{-1} \ll \omega_{ni} \ll \omega_{nm}$, we have $c_{ni}^{(2)} \to -\left(\frac{1}{\omega_{ni}t}\right) \times \sum_m i\frac{V_{nm}V_{mi}}{\omega_{mi}}t$, so that the scale of the perturbation is suppressed by $\omega_{ni}t \gg 1$. The effect these perturbations have on our phase measurement is clarified in the next section.

Effect of perturbed states on the phase measurement

We will simplify the notation of our phase measurement somewhat from that used in Ch. 2. For an ideal measurement, let us nominally prepare the state

$$|\psi(t=0)\rangle = \frac{1}{\sqrt{2}}(|M=+1\rangle + |M=-1\rangle) \qquad (4.7)$$

in the $|H, J = 1\rangle$ manifold and measure the state after a time t such that the state is

$$|\psi(t)\rangle = \frac{1}{\sqrt{2}}(e^{-i(\pi/4+\phi)}|M = +1\rangle + e^{+i(\pi/4+\phi)}|M = -1\rangle), \tag{4.8}$$

where typically $\phi \ll 1$. We alternately project into states $|X\rangle(|Y\rangle) \equiv \frac{1}{\sqrt{2}}(|M = +1\rangle \pm |M = -1\rangle)$ and observe the projected numbers $S_X \equiv |\langle X|\psi(t)\rangle|^2$, and likewise for S_Y. We then compute the asymmetry $\mathcal{A} = \frac{S_X - S_Y}{S_X + S_Y}$ and normalize by the contrast to obtain $\frac{\mathcal{A}}{2\mathcal{C}} = \phi_{\text{meas.}}$. Note that we are neglecting the complications arising from the possibility of changing the relative preparation and readout basis, or projecting through either $\tilde{\mathcal{P}}$ state. However, none of our conclusions depend on these details.

Due to a perturbation over the evolution of the state, let

$$|M = \pm 1\rangle \rightarrow [(1 + \epsilon_\pm) + ib_\pm]|M = \pm 1\rangle + (c_\pm + id_\pm)|M = \mp 1\rangle. \tag{4.9}$$

This perturbation describes any possible mixing of the $M = \pm 1$ states, or shifts in energy to one or the other state individually. Further, suppose that we initially prepare a small population imbalance so that

$$|\psi(t = 0)\rangle = \frac{1}{\sqrt{2}}[(1 - \mu)|M = +1\rangle + (1 + \mu)|M = -1\rangle]. \tag{4.10}$$

This occurs, for example, if the preparation laser has non-zero ellipticity, a possibility we neglected for the sake of clarity in Ch. 2. We can then find the measured phase $\phi_{\text{meas.}}$ (asymmetry normalized by $2\mathcal{C}$) by some straightforward algebra. If $\epsilon_\pm, b_\pm, c_\pm$, and $d_\pm$ are second-order small, while ϕ and μ are first-order small, then at third order in small quantities,

$$\phi_{\text{meas.}} = \phi - \frac{b_+ - b_-}{2} - \frac{c_+ + c_-}{2} - (d_+ - d_-)\phi - \frac{2}{3}\phi^3 + (c_+ - c_-)\mu - 2\phi\mu^2. \tag{4.11}$$

Ideally, $\phi_{\text{meas.}} = \phi$, but in the presence of perturbations or imperfections, this will not necessary hold true. Because of our normalization procedure, any perturbation to the

real component of a state, $\epsilon_\pm$, drops out of the phase measurement. The terms involving $b_\pm$ give the differential phase picked up due to M-dependent energy shifts. The terms involving $c_\pm$ give the M-independent real mixing between M states; we will consider this term at some length shortly. Further terms involve imperfections in the state preparation. Those involving ϕ can be relatively easily distinguished by, for example, taking data with different applied magnetic field amplitudes, for which the induced phase precession ϕ will vary significantly. The term $(c_+ - c_-)\mu$ couples the M-dependent real mixing between M states to an initial population imbalance, which could arise (for example) from ellipticity in the preparation laser and would therefore be modulated experimentally by varying the preparation laser ellipticity.

Let's return to the third term, $(c_+ + c_-)/2$. This term involves second-order mixing coefficients between M states, which are contained in the state coefficient $c^{(2)}_{M(-M)}$. As a concrete example, consider a constant perturbation and suppose we can neglect the terms that are necessarily of order $\frac{|V|^2}{\Delta^2}$. Then $c_\pm$ will include contributions proportional to

$$\text{Im}[V_{Mk}V_{k(-M)}]\,\text{Im}[\exp(iM\mu_H\mathcal{B}t) - 1]/(M\mu_H\mathcal{B}t), \qquad (4.12)$$

and

$$\text{Re}[V_{Mk}V_{k(-M)}]\,\text{Re}[\exp(iM\mu_H\mathcal{B}t) - 1]/(M\mu_H\mathcal{B}t), \qquad (4.13)$$

where k labels some intermediate state and the splitting between $\pm M$ is known to be dominated by the Zeeman interaction. The script symbols Im and Re denote "the imaginary part of" and "the real part of."

The expansion of $[\exp(iM\mu_H\mathcal{B}t) - 1]/(M\mu_H\mathcal{B}t)$ has only imaginary M-even terms and real M-odd terms, so the real contribution to $c^{(2)}_{M(-M)}$ in this case arises only from the imaginary and M-even part of $V_{Mk}V_{k(-M)} = V_{Mk}V^*_{(-M)k}$, or the real and M-odd part of the same quantity. Assuming a general form $V_{Mk} = \alpha + i\beta + M\gamma + iM\delta$, it turns out that the real part of $V_{Mk}V^*_{(-M)k}$ is entirely M-even and the imaginary part is entirely M-odd. Therefore, the term $(c_+ + c_-)/2$ in the measured phase vanishes for a static perturbation (provided, again, that we can ignore the term of order $\frac{|V|^2}{\Delta^2}$). I suspect, but have not

133

proved in generality, that this reasoning applies to a generic time-dependent perturbation as well–i.e., the term $(c_+ + c_-)/2$ is at most of order $\frac{|V|^2}{\Delta^2}$. If we accept the conjecture that this is generally true (and not just in the specific cases I've examined), and that the terms proportional to ϕ and μ are relatively easily identified by their explicit dependences on lab parameters like the magnetic field magnitude or preparation laser ellipticity, the perturbations $c^{(2)}_{M(-M)}$ can indeed be ignored. Under these conditions, we can justify the approach seen in [144, 149] of only considering the phase accumulated due to perturbed energies in $M = \pm 1$ separately rather than also computing second-order couplings between $M = \pm 1$.

4.4 $\mathcal{E}^{\text{nr}}$ and correlations with $\tilde{\mathcal{N}}\tilde{\mathcal{E}}$

In order to generate a systematic error, some mechanism must correlate measured phases with the experimental switches $\tilde{\mathcal{N}}$ and $\tilde{\mathcal{E}}$. It is comparatively easy to find a model where $\omega^{\tilde{\mathcal{N}}} \neq 0$ or $\omega^{\tilde{\mathcal{E}}} \neq 0$, but the fact that a systematic must be odd under either switch independently makes our system relatively robust. However, there is one way in which $\tilde{\mathcal{N}}\tilde{\mathcal{E}}$-correlated effects can easily enter into our measurement procedure: a non-reversing component of the applied electric field, $\mathcal{E}^{\text{nr}}$. In particular, $D\mathcal{E}^{\text{nr}} = \Delta^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$, where Δ is the detuning of the preparation or probe laser and D is the dipole moment in the $H(J = 1)$ state. Following our usual parity sum notation, $\mathcal{E}^{\text{nr}}$ is defined here by the applied electric field in any given experimental state, $\vec{\mathcal{E}} \cdot \hat{z}(\tilde{\mathcal{E}}) = \mathcal{E}\tilde{\mathcal{E}} + \mathcal{E}^{\text{nr}}$, where I am ignoring switches other than $\tilde{\mathcal{E}}$ and for the sake of simpler notation we always write $\mathcal{E}$ instead of $\mathcal{E}^{\tilde{\mathcal{E}}}$.

To see this, consider Fig. 4.2 and note that changing either $\tilde{\mathcal{N}}$ or $\tilde{\mathcal{E}}$ reverses the sign of the laser detuning from the $H \to I$ transition used for both state preparation and readout. Thus the laser detuning is correlated with the EDM switches, $\Delta = \Delta^{\text{nr}} + \tilde{\mathcal{N}}\tilde{\mathcal{E}}\,\Delta^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} + \cdots$, where in general there are additional (but typically less dangerous) non-zero terms such as $\Delta^{\tilde{\mathcal{N}}}$ and $\Delta^{\tilde{\mathcal{P}}}$.

Once this EDM-switch correlation has been introduced, there are two generic ways that

it can result in a systematic. Let's refer back to Eq. 4.4:

$$\bar{\phi}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = \frac{\int dV \, \epsilon (\rho^{\mathrm{nr}} \phi_0^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} + \delta\rho^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} \delta\phi^{\mathrm{nr}})}{N^{\mathrm{nr}}}. \tag{4.14}$$

The first situation in which a systematic can occur is that the overall measured molecular phase may depend on laser detunings, $d\phi_0^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = \frac{\partial\phi_0}{\partial\Delta}\Delta^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$. For the most part, ACME I systematic errors arose in this way. The second situation is somewhat more complicated to understand and requires three ingredients in the context of an $\mathcal{E}^{\mathrm{nr}}$-related effect. (1) The phase of the molecules must have a spatial dependence, $\delta\phi^{\mathrm{nr}} \neq 0$. For example, in the presence of a constant ambient magnetic field gradient $\frac{d\mathcal{B}}{dz}$, the spatially-dependent part of the non-reversing phase is $\delta\phi^{\mathrm{nr}} = -g\mu_B \frac{d\mathcal{B}}{dz} z\tau$. (2) There must be a non-reversing electric field gradient that generates a spatially-dependent correlated detuning. We will write the spatially-dependent part of the correlated detuning as $\delta\Delta^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$, in analogy to the way in which $\delta\rho^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ and $\delta\phi^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ are defined. For example, if there is a constant non-reversing electric field gradient $\frac{d\mathcal{E}^{\mathrm{nr}}}{dz}$, then $\delta\Delta^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = D\frac{d\mathcal{E}^{\mathrm{nr}}}{dz}z$. (3) Finally, the molecular population density that is either prepared or probed must depend on laser detunings, $\delta\rho^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = \frac{\partial\rho}{\partial\Delta}\delta\Delta^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$. These three conditions together are sufficient to generate a non-zero contribution to the overall measured phase $\bar{\phi}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ in Eq. 4.4.

In ACME II, we have systematic errors corresponding to both of these kinds of situation. In the following sections, I will first review all of the systematic errors first seen in ACME I that are still relevant to ACME II, and then I will describe systematic errors that are new to the ACME II measurement.

4.5 Review of AC Stark shift systematics

In ACME I, the leading systematic arose from AC Stark shifts coupled to non-reversing electric fields [128]. See [129, 149] for detailed elaboration. Here, I will briefly review the model of how AC Stark shifts in the preparation laser beams can produce an EDM systematic error proportional to $\mathcal{E}^{\mathrm{nr}}$.

We prepare and read out the spin projection of molecules in the H state by optical
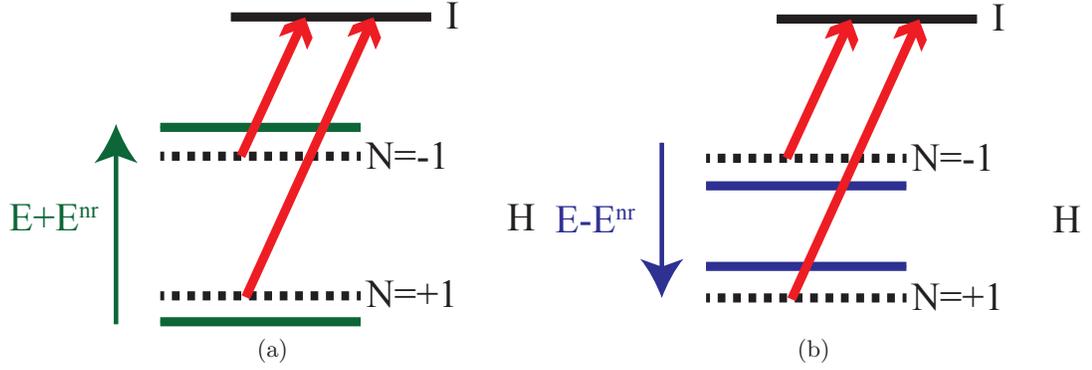
Figure 4.2: (a) Energy levels in $\tilde{\mathcal{E}} = +1$ state. With $\mathcal{E}^{\mathrm{nr}} > 0$, the $\tilde{\mathcal{N}} = -1$ state is blue-detuned, while the $\tilde{\mathcal{N}} = +1$ state is red-detuned. (b) Energy levels in $\tilde{\mathcal{E}} = -1$ state. With $\mathcal{E}^{\mathrm{nr}} > 0$, the $\tilde{\mathcal{N}} = -1$ state is red-detuned, while the $\tilde{\mathcal{N}} = +1$ state is blue-detuned. Thus changing either $\tilde{\mathcal{N}}$ or $\tilde{\mathcal{E}}$ reverses the detuning, $\Delta = \tilde{\mathcal{N}}\tilde{\mathcal{E}} \, D\mathcal{E}^{\mathrm{nr}}$, where D is the dipole moment in $H(J = 1)$.

pumping. In the tails of the laser beams, the Rabi frequency might be low enough that molecules are not optically pumped, but large enough to induce AC Stark shifts that effectively perturb the spin orientation of the molecules after preparation or before readout.

Instead of the basis $|H, J = 1, M = \pm 1\rangle$ and $|I, J = 1, M = 0\rangle$, it is convenient to work in the dressed-state basis $\{|D\rangle, |B_-\rangle, |B_+\rangle\}$, where $|D\rangle$ is the "dark" state (some superposition of $M = \pm 1$ in the H state), $|B_+\rangle$ is the higher-energy bright state, and $|B_-\rangle$ is the lower-energy bright state. If the Rabi frequency of the excitation light is Ω and the detuning is Δ, then the dressed-state basis Hamiltonian in the presence of the laser beams is [129, Eq. 65]

$$
\tilde{H} = \begin{bmatrix} 0 & -i\dot{\chi}^* \kappa_+ & -i\dot{\chi}^* \kappa_- \\ i\dot{\chi}\kappa_+ & E_{B-} & -\frac{i}{2}\frac{\dot{\Omega}\Delta - \Omega\dot{\Delta}}{\Delta^2 + \Omega^2} \\ i\dot{\chi}\kappa_- & \frac{i}{2}\frac{\dot{\Omega}\Delta - \Omega\dot{\Delta}}{\Delta^2 + \Omega^2} & E_{B+} \end{bmatrix}, \tag{4.15}
$$

where $\kappa_\pm^2 = \frac{1}{2}\left(1 \pm \frac{\Delta}{\sqrt{\Delta^2 + \Omega^2}}\right)$ and $E_{B\pm} = \frac{1}{2}\left(\Delta \pm \sqrt{\Delta^2 + \Omega^2}\right)$. The significance of $\dot{\chi}$ will be explained shortly. In general, $\Delta = \Delta_{\mathrm{det.}} - i\gamma/2$ is actually a complex "detuning" that incorporates the effect of the decay rate γ of the excited state B_+, as well as the physical detuning $\Delta_{\mathrm{det.}}$. Beware that including this complex decay rate causes the Hamiltonian $\tilde{H}$ to be non-Hermitian, so that $E_{B\pm}$ are complex and the population within the three-level

system is not conserved (as should indeed be the case for a system with decays to other states). Furthermore, we cannot have $\Delta = 0$ even when the resonance condition $\Delta_{\text{det.}} = 0$ is satisfied.

The factor $\dot{\chi}$ characterizes (typically small) changes to the polarization of the excitation light. If we parametrize the polarization as [129, Eq. 14]

$$\hat{\epsilon} = -e^{-i\theta}\cos\Theta\hat{\epsilon}_{+1} + e^{+i\theta}\sin\Theta\hat{\epsilon}_{-1}, \tag{4.16}$$

then $\dot{\chi} = \dot{\Theta} - i(\dot{\theta} - g\mu_B\mathcal{B}\tilde{\mathcal{B}})$, where the Zeeman precession rate has the same effect as a constant linear polarization gradient in the tails of the preparation and readout beams because it induces precession of the molecular angular momentum alignment with respect to the laser polarization.

Let us first consider an unrealistically simple case, in which both the Rabi frequency and detuning are constant in the tails of the laser beams. Then the effect of the AC Stark shift is equivalent to perturbing the polarization defining the prepared state according to

$$\hat{\epsilon}' = \hat{\epsilon} + d\chi\Pi i\hat{z} \times \hat{\epsilon}^*, \tag{4.17}$$

where $\Pi = \sum_{\pm} \kappa_{\mp}^2 e^{-iE_{B\pm}t} \int_0^t dt' \frac{\dot{\chi}(t')}{d\chi} e^{iE_{B\pm}t'}$ and $d\chi = d\Theta - i(d\theta - g\mu_B\mathcal{B}\tilde{\mathcal{B}}t)$. In terms of the parametrization in Eq. 4.16, the effective change in the laser polarization due to the actual change in polarization and the dynamical evolution of the molecular state is

$$
\begin{aligned}
d\theta_{\text{eff}} &= -d\Theta\text{Im}\Pi + (d\theta - g\mu_B\mathcal{B}\tilde{\mathcal{B}}t)\text{Re}\Pi \\
d\Theta_{\text{eff}} &= -d\Theta\text{Re}\Pi - (d\theta - g\mu_B\mathcal{B}\tilde{\mathcal{B}}t)\text{Im}\Pi.
\end{aligned}
\tag{4.18}
$$

The simplest case occurs if $\dot{\chi} = d\chi/t$ is constant during the optical pumping in the tails, so that

$$\Pi = \sum_{\pm} \kappa_{\mp}^2 e^{-iE_{B\pm}t/2}\text{sinc}(E_{B\pm}t/2). \tag{4.19}$$

The effective polarization errors in Eq. 4.18 modify the measured phase ϕ according to

$$\phi \to \phi + d\theta_{\text{ref}} - \frac{1}{2}(d\theta_X + d\theta_Y) - \tilde{\mathcal{P}}_{\text{prep}}\tilde{\mathcal{P}}_{\text{read}}d\Theta_{\text{ref}}(d\Theta_X - d\Theta_Y) + \mathcal{O}(d\epsilon)^3. \qquad (4.20)$$

In the present discussion, we will be concerned with the second term, $d\theta_{\text{ref}}$, an effective change in the linear polarization angle occuring in the refinement beam due to an AC Stark shift in the "falling tail" of the laser. From Eq. 4.18, we see that this can arise from a change in laser ellipticity, $d\Theta$, across the tail of the laser beam. In particular, we are concerned with the effect of $\Delta^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ on the measured EDM-correlated frequency, $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$, so we Taylor expand Eq. 4.18 in $\Delta_{\text{det.}}$ and divide by the precession time:

$$d\omega_{\mathcal{E}^{\text{nr}}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = -\frac{1}{\tau}d\Theta\frac{\partial\text{Im}\Pi}{\partial\Delta_{\text{det.}}}\Delta^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}. \qquad (4.21)$$

4.5.1 Polarization gradients

In the apparatus, the mounting of both the vacuum windows and the electric field plates results in non-uniform stress-induced birefringence. The combined effect of the window and first field plate that the laser passes through (i.e., before it interacts with the molecules) is equivalent to a single birefringent element with an effective fast axis and birefringence. There is no reason that this effective fast axis will be exactly perpendicular or parallel to the laser polarization axis, so the laser is imprinted with a polarization ellipticity as it passes through the birefringent material. Furthermore, and more problematically, the imprinted birefringence is non-uniform across the laser beam due to the stress gradients across the vacuum windows and field plates.

We have precisely measured the polarization of our lasers after passing through both vacuum windows and field plates, in an effort led by Vitaly Andreev [150, 198]. We found that the birefringence axis is aligned with the optical table to within $\approx 5°$. This near-alignment is fortuitous since the imprinted ellipticity on the laser beam scales with the misalignment of the laser polarization (nominally $\hat{x}$ for the refinement beam) and the birefringence axis. However, this is not particularly surprising since the vacuum windows and field plates are

both rectangular, with symmetry axes along $\hat{x}$ and $\hat{y}$.

For a small misalignment angle β between the laser polarization and birefringence axis, and small effective retardance δ, the ellipticity is found[2] by $\Theta - \frac{\pi}{4} = -\beta\delta$. Thus, at leading order, $d\Theta = -d\beta\,\delta - \beta\,d\delta$, where $d\beta$ and $d\delta$ are small changes in the birefringence axis and retardance across the laser beam, respectively.

Vitaly Andreev has carefully measured ellipticity gradients and found that, under normal conditions, $\frac{d\Theta}{dx} < 10^{-3}/$mm when the laser passes through both windows and both fields plates. Unfortunately, direct polarimetry of the laser light in between the field plates is not possible under realistic experimental conditions (e.g., under vacuum). If we suppose, for point of argument, that the region of the laser beam where the AC Stark shift occurs is $\sim$1 mm, then we can expect $d\Theta \sim 10^{-3}$.[3]

4.5.2 Behavior of $\mathrm{Im}[\Pi(\Delta_{\mathrm{det}})]$ and quantitative estimate

It may seem intuitive that physical quantities should be at most quadratic in Δ_{det} on resonance, $\Delta_{\mathrm{det}} = 0$, so that in an ideal measurement where $\Delta_{\mathrm{det}} = 0$, the factor $\frac{\partial \mathrm{Im}[\Pi]}{\partial \Delta_{\mathrm{det}}}$ would vanish. In fact, precisely the opposite is true; the imaginary part of Π is completely odd in the detuning, and $|\frac{\partial \mathrm{Im}[\Pi]}{\partial \Delta_{\mathrm{det}}}| > 0$ even on resonance.

To see this, let's take $\dot{\chi} = \dot{\Theta}$ since we are only concerned with ellipticity gradients here. Then $\dot{\chi}$ and $d\chi$ are real, while a bit of algebra shows that $\kappa_{\pm}^2(-\Delta_{\mathrm{det}}, \gamma, \Omega) = [\kappa_{\mp}^2(+\Delta_{\mathrm{det}}, \gamma, \Omega)]^*$ and, likewise, $E_{B\pm}(-\Delta_{\mathrm{det}}, \gamma, \Omega) = E_{B\mp}^*(+\Delta_{\mathrm{det}}, \gamma, \Omega)$. It follows (due to the sum over $\pm$ in the definition of Π) that $\Pi(-\Delta_{\mathrm{det}}, \gamma, \Omega) = \Pi^*(+\Delta_{\mathrm{det}}, \gamma, \Omega)$. Let us now decompose $\Pi = \Pi_{\Delta_{\mathrm{det}}} + \Pi_0$ into a part that is an odd function of detuning and a part that is an even function of detuning. Then we can find that $\mathrm{Im}[\Pi] = \frac{1}{2i}(\Pi - \Pi^*) = \Pi_{\Delta_{\mathrm{det}}}$ and $\mathrm{Re}[\Pi] = \frac{1}{2}(\Pi + \Pi^*) = \Pi_0$. Thus, in the case that $\dot{\chi} = \dot{\Theta}$ is real, the imaginary component of Π is an odd function of detuning and the real component is an even function of detuning.

---

2. See [150, Eq. 4.2] for the formalism. The conversion from the normalized Stokes parameter S/I to our ellipticity parameter is $\Theta = \frac{\pi}{4} - \frac{1}{2}\frac{S}{I}$.

3. Parameters of the AC Stark shift model can't be matched to the experimental conditions to better than a factor of a few since the exact molecular dynamics in the tails of the laser beams are not perfectly understood. In this instance, the relevant length scale is not rigorously defined, so we can only indicate the expected order of magnitude for $d\Theta$.

Imperfection	Corresponding quantity	Typical value	Uncertainty
$\mathcal{E}^{\mathrm{nr}}$	$\Delta^{\tilde{\mathcal{N}\mathcal{E}}}$	-2.6 mV/cm	1.6 mV/cm
$d\Theta$	Optics birefringence	$\sim 10^{-3}$	—
Quantity	**Value**	**Uncertainty**	**Units**
$\partial\omega^{\tilde{\mathcal{N}\mathcal{E}}}/\partial\mathcal{E}^{\mathrm{nr}}$	21	52	$(\mu\mathrm{rad/sec})/(\mathrm{mV/cm})$
$d\omega^{\tilde{\mathcal{N}\mathcal{E}}}_{\mathcal{E}^{\mathrm{nr}}}$	-56	140	$\mu\mathrm{rad/sec}$

Table 4.2: Summary of $\mathcal{E}^{\mathrm{nr}}$ systematic parameters and shift. The experimental imperfection $\mathcal{E}^{\mathrm{nr}}$ leads to a correlated detuning, $\Delta^{\tilde{\mathcal{N}\mathcal{E}}} = D\mathcal{E}^{\mathrm{nr}}$. Together with a change in the laser ellipticity in the refinement region, $d\Theta$, due to a birefringence gradient, $\frac{d\Theta}{dx}$, the correlated detuning can lead to a systematic slope, $d\omega^{\tilde{\mathcal{N}\mathcal{E}}}/d\mathcal{E}^{\mathrm{nr}}$. We compute the systematic slope by comparing $\omega^{\tilde{\mathcal{N}\mathcal{E}}}$ in the presence of a deliberately applied large value of $\mathcal{E}^{\mathrm{nr}} = 150 - 300$ mV/cm, and the value of $\omega^{\tilde{\mathcal{N}\mathcal{E}}}$ under ideal conditions. We then compute a systematic shift, $d\omega^{\tilde{\mathcal{N}\mathcal{E}}}_{\mathcal{E}^{\mathrm{nr}}} = \frac{d\omega^{\tilde{\mathcal{N}\mathcal{E}}}}{d\mathcal{E}^{\mathrm{nr}}}\mathcal{E}^{\mathrm{nr}}$, using the measured systematic slope (consistent with zero) and a typical value of $\mathcal{E}^{\mathrm{nr}}$ obtained via microwave spectroscopy (see Sec. 5.1).

This shows us that, whatever the exact functional form of Π is, we cannot null out the AC Stark shift systematic by tuning the lasers to resonance since even then, $\frac{\partial \mathrm{Im}[\Pi]}{\partial\Delta_{\mathrm{det}}} \neq 0$.

Approximate analytic results are obtained in [149, Sec. 4.3.2] for $\mathrm{Im}[\Pi(\Delta, \gamma, \Omega)]$, for several specific models of the time-dependent Rabi frequency in the falling tail of the laser beam. However, these models should not be expected to match our system perfectly, and we resort instead to the experimental data. In ACME I, ellipticity gradients on the order of $\sim 10^{-2}/\mathrm{mm}$ led, via the AC Stark shift mechanism, to a systematic error effect (after suppression of the initial systematic) of up to $d\omega^{\tilde{\mathcal{N}\mathcal{E}}} \sim \frac{1}{\tau}10^{-6}\frac{\mathrm{rad}}{2\pi\times10\,\mathrm{kHz}}\Delta^{\tilde{\mathcal{N}\mathcal{E}}}$ [129, Sec. 5.2.5]. As we will see in Sec. 5.1, non-reversing electric fields contribute $\Delta^{\tilde{\mathcal{N}\mathcal{E}}}$ on the order of several kHz, so with $\tau \sim 1\,\mathrm{ms}$ and $d\Theta < 10^{-3}$, we could expect $d\omega^{\tilde{\mathcal{N}\mathcal{E}}}_{\mathcal{E}^{\mathrm{nr}}} < 100\,\mu\mathrm{rad/sec}$, which is only smaller than the ACME II sensitivity by a factor of order unity.

4.5.3 Measured $\mathcal{E}^{\mathrm{nr}}$ systematic

To test the contribution of the AC Stark shift systematic seen from ACME I, we deliberately apply a large non-reversing component of the electric field, $\mathcal{E}^{\mathrm{nr}} = 150 - 300\,\mathrm{mV/cm}$, and monitor $\frac{d\omega^{\tilde{\mathcal{N}\mathcal{E}}}}{d\mathcal{E}^{\mathrm{nr}}}$ throughout the EDM data set (typically one superblock measured per day with a deliberately large value of $\mathcal{E}^{\mathrm{nr}}$, and ~ 20 total superblocks per day under ideal conditions). We observed $\frac{d\omega^{\tilde{\mathcal{N}\mathcal{E}}}}{d\mathcal{E}^{\mathrm{nr}}} = 21(52)\frac{\mu\mathrm{rad}}{\mathrm{sec}}/(\mathrm{mV/cm})$, where the quantity in parentheses

is a 1σ statistical uncertainty[4]. The average non-reversing electric field in the preparation and probe region, where $\mathcal{E}^{\mathrm{nr}}$ can generate $\Delta^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ and is therefore liable to produce a systematic shift in $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$, is $\mathcal{E}^{\mathrm{nr}} = -2.6(16)$ mV/cm as described in Sec. 5.1. We therefore infer a systematic shift due to $\mathcal{E}^{\mathrm{nr}}$ of $d\omega_{\mathcal{E}^{\mathrm{nr}}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = \frac{d\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{d\mathcal{E}^{\mathrm{nr}}}\mathcal{E}^{\mathrm{nr}} = -56(140)\,\mu\mathrm{rad/sec}$, consistent with zero. For a summary of parameters related to the $\mathcal{E}^{\mathrm{nr}}$ systematic, see Table 4.2.

During the acquisition of the final dataset, we took four runs ($\sim$days) of data with all refinement and probe laser optics on the west, rather than east, side of the interaction region, and the lasers propagating along $-\hat{z}$ instead of along $+\hat{z}$ as in the ordinary configuration of the apparatus. We distinguish between these configurations by $\hat{k}\cdot\hat{z} = \pm 1$, where $\hat{k}$ denotes the wavevector of preparation and probe beams. At the sensitivity of these four runs with $\hat{k}\cdot\hat{z} = -1$, all measured precession frequencies were consistent with those obtained for $\hat{k}\cdot\hat{z} = +1$. However, we observed a systematic slope of $\frac{d\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{d\mathcal{E}^{\mathrm{nr}}} \approx 700\,\frac{\mu\mathrm{rad}}{\mathrm{sec}}/(\mathrm{mV/cm})$, which would result in an EDM shift of ~ 3 mrad/sec, significantly larger than the statistical sensitivity of the EDM measurement. If this systematic shift were measured sufficiently precisely, then it could be subtracted from $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = \omega_T^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} + d\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} + \cdots$ in order to obtain the contribution to the EDM channel from T-violating physics alone, $\omega_T^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$. However, because the systematic slope is larger in the $\hat{k}\cdot\hat{z}$ configuration, the statistical uncertainty in $d\omega_{\mathcal{E}^{\mathrm{nr}}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ would exceed the statistical uncertainty of $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ for data taken under ideal conditions (373 μrad/sec). As a result, the *uncertainty* in the systematic error contribution due to data taken with $\hat{k}\cdot\hat{z} = -1$ would dominate the overall error budget of the measurement. The observed increase in the $\mathcal{E}^{\mathrm{nr}}$ systematic error is consistent with the fact that the effective ellipticity of the laser while interacting with the molecules is determined only by the first window and field plate that it passes through. If the birefringence axis is less well-aligned with $\hat{x}$, or if the retardance is larger, on one side of the vacuum chamber than the other, then the systematic slope $\frac{d\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{d\mathcal{E}^{\mathrm{nr}}}$ will be proportionally larger for the corresponding $\hat{k}\cdot\hat{z}$ configuration.

In ACME I, the $\mathcal{E}^{\mathrm{nr}}$ systematic error was suppressed in part by aligning the refinement

4. I always give systematic error slopes, shifts, and uncertainties from the average over all four analysis routines, which can disagree by $\sim \pm 10\%$, depending on our various choices of analysis parameters. We have verified that under the same conditions, we obtain the same results to extremely high precision.

laser polarization along the combined birefringence axis of the field plates and vacuum windows. In ACME II, adjusting the refinement laser away from the molecular angular momentum axis prepared by the STIRAP beams (which cannot be adjusted) necessarily incurs a loss in signal. However, the larger problem with suppressing the $\mathcal{E}^{\mathrm{nr}}$ systematic in the $\hat{k} \cdot \hat{z} = -1$ configuration is that the refinement laser polarization must align with the STIRAP-prepared angular momentum alignment in order to suppress the "$P^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$" systematic error. This is discussed in Sec. 4.7. Thus, it is not practical to suppress the $\mathcal{E}^{\mathrm{nr}}$ systematic further for the $\hat{k} \cdot \hat{z} = -1$ configuration.

For these reasons, we only include in the calculation of the EDM value data taken with $\hat{k} \cdot \hat{z} = +1$, where the $\mathcal{E}^{\mathrm{nr}}$ systematic slope is sufficiently small that the overall uncertainty in the systematic error contribution is small compared to the statistical sensitivity of the EDM measurement.

4.6 Systematic errors in the STIRAP lasers

4.6.1 Stark interference

In this section we will consider a separate systematic that arises from molecular dynamics in the presence of laser light. Up to now, we have assumed that all laser transitions involve the dominant $E1$ matrix element. However, in fact all multipole transition matrix elements $E1$, $M1$, $E2\ldots$ will act to drive molecules between the addressed states. Under certain conditions, these various matrix element amplitudes can interfere with each other. Here, I will consider only the $E1$ and $M1$ transition amplitudes, which are dominant. The calculations are fairly involved; see [129, 149] for details. I will merely summarize the concepts here.

For a classical light field, the corresponding operators on the molecular states take the form

$$
\begin{aligned}
O_{E1} &= \hat{\epsilon} \cdot \sum_i \vec{r}_i \\
O_{M1} &= (\hat{k} \times \hat{\epsilon}) \cdot \sum_i \frac{1}{2m_i}(g_L \vec{L}_i + g_S \vec{S}_i),
\end{aligned}
\tag{4.22}
$$

142

where the sums run over all charged particles in the system and m_i is a mass. Note that for both operators, we have a lab-frame vector ($\hat{\epsilon}$ or $\hat{k} \times \hat{\epsilon}$) and a molecule-frame vector ($\vec{r}$ or $g_L \vec{L} + g_S \vec{S}$). By using the usual transformations between different frames, it is possible to compute the transition probability between an initial and final state, $|\langle \psi_f | O_{E1} + O_{M1} | \psi_i \rangle|^2$. This transition probability contains "Stark interference" terms $\langle \psi_f | O_{E1} O_{M1}^* | \psi_i \rangle + c.c.$ that are only first-order in the relative strength of $E1$ and $M1$ matrix elements, $\frac{c_{M1}}{c_{E1}}$. (See [149, Sec. 4.2] for rigorous definitions). We find it convenient to treat this interference as a perturbation to the effective $E1$ polarization:

$$\vec{\epsilon}_{\text{eff}} = \hat{\epsilon} - \frac{c_{M1}}{c_{E1}} i\, \hat{n} \times (\hat{k} \times \hat{\epsilon}). \tag{4.23}$$

The symmetries of the $E1$ and $M1$ operators are expected to require $\frac{c_{M1}}{c_{E1}}$ to be entirely real. The factor $\hat{n} = \tilde{\mathcal{N}}\tilde{\mathcal{E}}\hat{z}$ in the perturbation term ultimately arises from the fact that the $M1$ molecular matrix element $\langle \Lambda_f, \Sigma_f, \Omega | L_0^{\text{mol.}} | \Lambda_i, \Sigma_i, \Omega_i \rangle \propto \Omega = M\tilde{\mathcal{N}}\tilde{\mathcal{E}}$, while the analogous $E1$ matrix element $\langle \Lambda_f, \Sigma_f, \Omega | r_0^{\text{mol.}} | \Lambda_i, \Sigma_i, \Omega_i \rangle$ is not proportional to Ω.

It turns out that when $\hat{k} \cdot \hat{z} \approx \pm 1$, as for the refinement and probe beams, the effective $\tilde{\mathcal{N}}\tilde{\mathcal{E}}$-correlated perturbation to the linear polarization angle is $d\theta_{\text{eff}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = \mathcal{O}(\theta_k^2)$, where θ_k is the angular deviation of the wavevector $\hat{k}$ from the $\hat{z}$-axis. Under ordinary experimental conditions, therefore, the Stark interference has negligible effect on the refinement and probe beams. On the other hand, when $\hat{k} = \hat{y}$ and $\hat{\epsilon} \approx \hat{x}$, as for the STIRAP beams, the effective $\tilde{\mathcal{N}}\tilde{\mathcal{E}}$-correlated perturbation to the linear polarization angle is $d\theta_{\text{eff}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = -\frac{c_{M1}}{c_{E1}} d\Theta$, where $d\Theta$ is the average ellipticity parameter of the laser beam. Note that this average laser ellipticity is of order $\sim 10^{-2}$, an order of magnitude larger than the *change* in ellipticity across the laser profile that we considered for the AC Stark shift systematics. For $\frac{c_{M1}}{c_{E1}} \sim \alpha \sim 10^{-2}$, we expect $d\theta_{\text{eff}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} \sim 10^{-4}$ rad. In the absence of a refinement laser, we would therefore measure $d\omega_{\text{ST}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = \frac{1}{\tau} d\theta_{\text{eff}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} \sim 100$ mrad/sec, a rather dramatic factor of $\sim 10^3$ greater than the statistical sensitivity in the ACME II measurement. As we will see, we observe $|d\omega_{\text{ST}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}| \lesssim 30$ mrad/sec, likely smaller than anticipated due to order-of-unity errors in our estimate of c_{M1}/c_{E1}.

4.6.2 AC Stark shifts in the STIRAP beams

In addition to the Stark interference effect, there are also AC Stark shift systematics, proportional to $\mathcal{E}^{\mathrm{nr}}$, that modulate the phase prepared in the falling tail of the STIRAP laser beams. Counter-intuitively, we actually expect that the largest polarization gradient in the STIRAP beams arises from the intrinsic polarization structure of Maxwell-Gaussian laser beams, due to the tight focus of the beams along $\hat{x}$ (waist size $w_x \approx 150\,\mu$m). The local electric field axis (i.e., polarization) in a gaussian beam nominally polarized along $\hat{x}$ is[149, Sec. 4.3.3]

$$\hat{\epsilon} = \hat{x} + \left(\frac{\lambda}{2\pi}\right)\frac{xy}{w_x^2 w_y^2}\hat{y} - i\left(\frac{\lambda}{2\pi}\right)\frac{x}{w_x^2}\left(1 - \frac{\lambda z}{2\pi w_x^2}\right)\hat{z} + \cdots , \qquad (4.24)$$

where $w_{x(y)}$ is the waist size along the $x(y)$-direction and λ is the wavelength of the laser light. For the 1090 nm STIRAP beam, the ellipticity gradient is $\frac{d\Theta}{dx} \sim 10^{-2}$/mm, significantly larger than that observed due to stress-induced birefringence in the vacuum windows. Nevertheless, the systematic error from Stark interference, together with a non-zero ellipticity, dominates the systematic error from AC Stark shifts coupled to $\mathcal{E}^{\mathrm{nr}}$ and birefringence gradients.

4.6.3 Residual STIRAP systematics after refinement

In order to suppress the enormous Stark interference systematic arising in the STIRAP beams, we reproject the molecular state using the refinement beam. In particular, we optically pump away the bright state of the refinement beam, nominally leaving only the refinement beam dark state remaining. However, we do not successfully optically pump 100% of the bright-state molecules in the refinement region.

Recall that we can characterize the molecular state in the $H(J = 1)$ manifold by the orientation of its angular momentum alignment in the xy-plane, and that this orientation has a one-to-one correspondence with the linear polarization of the preparation light. Let the dark state of the refinement beam have an alignment angle θ_{ref} and the STIRAP-prepared state have an alignment angle θ_{ST}. Under ordinary conditions, the misalignment angle is small, $d\theta_{\mathrm{ST}} \equiv \theta_{\mathrm{ST}} - \theta_{\mathrm{ref}} \ll 1$ in any particular state of the experiment. Due to the Stark

interference and AC Stark effects just described, the STIRAP-prepared alignment angle, $d\theta_{\rm ST} = d\theta_{\rm ST}^{\rm nr} + \tilde{\mathcal{N}}\tilde{\mathcal{E}}d\theta_{\rm ST}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$, has an $\tilde{\mathcal{N}}\tilde{\mathcal{E}}$-correlated component in addition to a non-reversing offset from, e.g., misalignment of the refinement beam polarization. In the absence of a refinement beam, the $\tilde{\mathcal{N}}\tilde{\mathcal{E}}$-correlated component of the STIRAP-prepared phase would directly generate a systematic error, $\omega_{\rm ST}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = d\theta_{\rm ST}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\tau$. However, the refinement beam reprojects the molecular state's angular momentum alignment along the refinement beam polarization axis (the dark state of the laser), attenuating any perpendicular (bright state) component via optical pumping. We define the attenuation factor

$$A_{\rm ref.} = \left(\frac{\partial \omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{\partial \omega_{\rm ST}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}\right)^{-1}. \tag{4.25}$$

Here, $\omega_{\rm ST}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ is the measured $\tilde{\mathcal{N}}\tilde{\mathcal{E}}$-correlated frequency with the refinement laser off, and $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ is the residual value of the EDM-correlated frequency when the refinement laser is on but all experimental conditions are otherwise identical. Note that for powerful suppression of the refinement laser bright state, $A_{\rm ref} \gg 1$. In fact, in the following discussion we will be concerned with the part of $\omega_{\rm ST}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ that *leaks* through to $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$, rather than the part that is attenuated, since the STIRAP systematic contribution to the EDM channel is $d\omega_{\omega-{\rm ST.}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = \frac{d\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{d\omega_{\rm ST}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}\Delta\omega_{\rm ST}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$, where $\Delta\omega_{\rm ST}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ is the value of $\omega_{\rm ST}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ during ordinary EDM data.

When the 1090 nm laser polarization is misaligned with the vacuum window's birefringence axis, the laser is imprinted with a non-zero ellipticity $d\Theta$ and ellipticity gradient $\frac{d\Theta}{dx}$. When we rotate a half-wave plate, which controls the 1090 nm laser's linear polarization angle in the xz-plane, by angle $d\theta_{\lambda/2}$, then we therefore observe an enhanced combination of Stark interference and AC Stark shift systematics, $d\omega_{\rm ST}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = -0.319(4) \frac{\rm rad}{\rm sec} d\theta_{\lambda/2}$. During the EDM data set, approximately once per day, we rotate this half-wave plate by $15°$, creating $\omega_{\rm ST}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} \sim 5\,{\rm rad/sec}$. Under these conditions, we can measure the residual $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ with the refinement laser on to infer $A_{\rm ref}$. In particular, we compute a linear regression of $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ for both ordinary data (where $\omega_{\rm ST}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} \approx 0$) and this "intentional parameter variation" data (where $\omega_{\rm ST}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} \sim 5$ rad/sec) against $\omega_{\rm ST}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$, to compute $\partial \omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial \omega_{\rm ST}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = A_{\rm ref}^{-1} = -3(37) \times 10^{-5}$,

Imperfection	Corresponding parameter	Typical value	Uncertainty
$\Delta\omega_{\mathrm{ST}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$	$d\theta_{\lambda/2}$	-1 mrad/sec	15 mrad/sec
$A_{\mathrm{ref}} < \infty$	P_{ref}	> 2500	—

Quantity	Value	Uncertainty	Units
$\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial\omega_{\mathrm{ST}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$	-3×10^{-5}	37×10^{-5}	—
$d\omega_{\omega-\mathrm{ST}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$	-0.1	1.5	μrad/sec

Table 4.3: Summary of $\omega_{\mathrm{ST}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ systematic parameters and shift. The imperfection, $\Delta\omega_{\mathrm{ST}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$, is a correlated phase (up to a factor of the precession time) prepared by the STIRAP lasers, which arises from a misalignment of the STIRAP 1090 nm laser polarization with the birefringence axis of the vacuum window. This misalignment is controlled by a half-wave plate angle, $d\theta_{\lambda/2}$. Together with a finite refinement laser attenuation factor, $A_{\mathrm{ref}} = A_{\mathrm{ref}}(P_{\mathrm{ref}})$, of the unwanted component of the STIRAP-prepared phase, the correlated component of the STIRAP-prepared phase can lead to a systematic slope, $d\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/d\omega_{\mathrm{ST}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$. We compute the systematic slope by comparing $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ in the presence of a deliberately applied large value of $\omega_{\mathrm{ST}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} \sim 5$ rad/sec, and the value of $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ under ideal conditions. We compute a systematic shift, $d\omega_{\omega-\mathrm{ST}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = \frac{d\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{d\omega_{\mathrm{ST}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}\Delta\omega_{\mathrm{ST}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$, using the measured systematic slope (consistent with zero) and a typical value of $\Delta\omega_{\mathrm{ST}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ obtained by comparing the value of $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ with the refinement laser on vs. off.

consistent with $|A_{\mathrm{ref}}| > 2500$.[5]

In order to compute a systematic error contribution, $d\omega_{\omega-\mathrm{ST}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = \frac{d\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{d\omega_{\mathrm{ST}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}\Delta\omega_{\mathrm{ST}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$, we must measure the value of $\omega_{\mathrm{ST}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ under the conditions of ordinary EDM data. Approximately once per day, we measure the EDM-channel precession frequency with the refinement laser off, but otherwise all experimental conditions optimized. We compute $\Delta\omega_{\mathrm{ST}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = \omega_{\mathrm{ref\,off}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} - \omega_{\mathrm{ref.\,on}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ to remove the effect of an actual EDM from our calculation of $\Delta\omega_{\mathrm{ST}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$, where the subscripts indicate measurements made with the refinement laser off ($A_{\mathrm{ref}} \to 1$) and the refinement laser on ($A_{\mathrm{ref}} \gg 1$), respectively. We find $\Delta\omega_{\mathrm{ST}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = -1(15)$ mrad/sec. Thus the systematic error contribution is $d\omega_{\omega-\mathrm{ST}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = -0.1(15)\,\mu$rad/sec, consistent with zero and far below the statistical sensitivity of the EDM measurement. For a summary of parameters related to the $\omega_{\mathrm{ST}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ systematic, see Table 4.3.

Interestingly, the 2σ bound on $\Delta\omega_{\mathrm{ST}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ is only $|\Delta\omega_{\mathrm{ST}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}| \lesssim 30$ mrad/sec, a factor of approximately three below the estimated magnitude of anticipated Stark interference effects. I expect that the discrepancy arises primarily from the assumption that $c_{M1}/c_{E1} \sim 10^{-2}$.

---

5. To connect this discussion more directly to the "knob" we turn in the lab, namely $d\theta_{\lambda/2}$, simply use the chain rule: $\frac{\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{\partial\omega_{\mathrm{ST}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}} = \frac{\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{\partial d\theta_{\lambda/2}} / \frac{\partial\omega_{\mathrm{ST}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{\partial d\theta_{\lambda/2}}$.

Without detailed calculations of the matrix elements, c_{M1} and c_{E1}, we should not expect our previous estimate to be accurate to better than a factor of order unity.

4.7 Effect of correlated refinement power

In ACME I, there was evidence of an intrinsically $\tilde{\mathcal{N}}\tilde{\mathcal{E}}$-correlated Rabi frequency, $\Omega_r^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$, in the optical pumping preparation laser (most closely analogous to the refinement laser in the ACME II experimental scheme). Although the source of this correlated Rabi frequency was never fully understood, for the moment let us consider one of its effects on our measurement. Recall from the AC Stark shift model that the effective modification to the refinement laser polarization due to AC Stark shift effects is $d\theta_{\text{eff}} = -d\Theta\,\text{Im}[\Pi] + (d\theta - g\mu_B\mathcal{B}\tilde{\mathcal{B}}t)\text{Re}[\Pi]$, where $\Pi = \Pi(\Omega_r, \Delta)$ depends on the molecular dynamics in the falling tail of the laser. Here, the linear polarization gradient in the laser, $d\theta$, and the continuous accumulation of phase from Zeeman precession, $-g\mu_B\mathcal{B}\tilde{\mathcal{B}}t$, have exactly analogous contributions to $d\theta_{\text{eff}}$. In the present discussion we are interested in the term that appears odd under $\tilde{\mathcal{B}}$: $d\theta_{\text{eff}}^{\tilde{\mathcal{B}}} = -g\mu_B\mathcal{B}t\,\text{Re}[\Pi]$. However, recall that Π depends on the Rabi frequency, Ω_r. Therefore, if $\Omega_r = \Omega_r^{\text{nr}} + \Omega_r^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}\tilde{\mathcal{N}}\tilde{\mathcal{E}}$, then this term contributes an effective phase that is correlated with $\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}$:

$$d\theta_{\text{eff}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}} = -g\mu_B\mathcal{B}t\,\text{Re}\left[\frac{\partial\Pi}{\partial\Omega_r}\right]\Omega_r^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}, \tag{4.26}$$

where we've made the reasonable approximation that $\Omega_r^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\Omega_r^{\text{nr}} \ll 1$. The observation of $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}} \propto \mathcal{B}$ in ACME I was one key piece of evidence in support of a non-zero $\Omega_r^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$.

Because of these observations in ACME I, we considered the possibility of a correlated Rabi frequency in the refinement beam during ACME II. To test the effect of such a correlated Rabi frequency, we artificially impose a correlated laser power, $P_{\text{ref}} = P_{\text{ref}}^{\text{nr}} + P_{\text{ref}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}\tilde{\mathcal{N}}\tilde{\mathcal{E}}$, typically with $P_{\text{ref}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/P_{\text{ref}}^{\text{nr}} \approx 0.1$. We initially found a large systematic error slope, $\frac{\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{\partial(P_{\text{ref}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/P_{\text{ref}}^{\text{nr}})}$ up to ± 10 rad/sec (e.g., $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ up to ± 1 rad/sec with $P_{\text{ref}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/P_{\text{ref}}^{\text{nr}} = 0.1$). Contrary to the discussion of Sec. 2.3.5 for an idealized measurement, the systematic slope depends linearly on a "global" polarization angle, i.e., a common rotation of the refinement polarization θ_{ref} and readout polarizations $\theta_{X(Y)}$. The reason for this is that the angular momentum align-
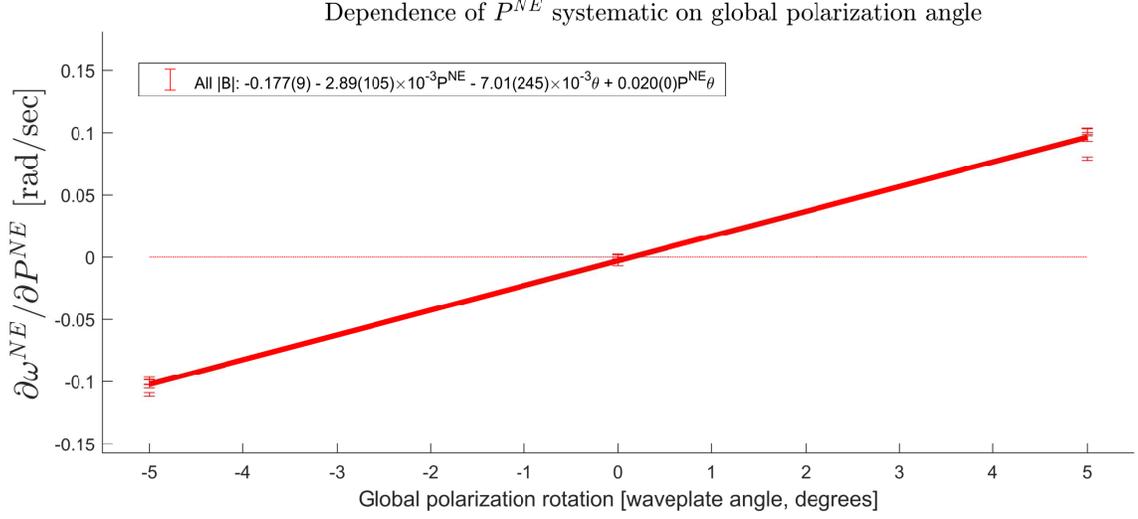
Figure 4.3: $P^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ systematic slope, $\frac{\partial \omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{\partial (P^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/P^{\mathrm{nr}})}$, as a function of the global polarization angle. The x-axis gives the offset of the preparation and readout half-wave plates in degrees. By setting $\Delta\theta_{\mathrm{ref}} = 0$, the systematic slope can be tuned to zero. Data from Runs 55-56.

ment prepared by the STIRAP lasers, at angle $\theta_{\mathrm{ST}}^{\mathrm{nr}}$ with respect to $\hat{x}$, cannot be rotated together with θ_{ref} and $\theta_{X(Y)}$. This fact suggests a form for the systematic error:

$$d\omega_{\delta P}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = \frac{\partial \omega}{\partial(\delta P_{\mathrm{ref}}/P_{\mathrm{ref}})} \frac{P_{\mathrm{ref}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{P_{\mathrm{ref}}^{\mathrm{nr}}} \Delta\theta_{\mathrm{ref}}, \tag{4.27}$$

where $\delta P_{\mathrm{ref}}/P_{\mathrm{ref}}$ denotes a fractional modulation of the refinement laser power and $\Delta\theta_{\mathrm{ref}} \equiv \theta_{\mathrm{ref}} - \theta_{\mathrm{ST}}^{\mathrm{nr}}$. Note that this systematic error contribution is proportional to both $P_{\mathrm{ref}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/P_{\mathrm{ref}}^{\mathrm{nr}}$ and to the misalignment between the refinement and STIRAP axes, $\Delta\theta_{\mathrm{ref}}$.

We can explain the dependence of $d\omega_{\delta P}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ on both the correlated refinement laser power $P_{\mathrm{ref}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/P_{\mathrm{ref}}^{\mathrm{nr}}$ and polarization misalignment $\Delta\theta_{\mathrm{ref}}$ as follows. The refinement laser reprojects the molecular angular momentum alignment from angle θ_{ST} to $\phi_{\mathrm{ref}} = \theta_{\mathrm{ref}} - \Delta\theta_{\mathrm{ref}}/A_{\mathrm{ref}}$, where again A_{ref} is the attenuation factor of the refinement laser. Note that as $A_{\mathrm{ref}} \to \infty$, $\phi_{\mathrm{ref}} \to \theta_{\mathrm{ref}}$. On the other hand, if the refinement laser is off, $A_{\mathrm{ref}} \to 1$, then $\phi_{\mathrm{ref}} \to \theta_{\mathrm{ST}}$. However, the refinement laser attenuation depends on the refinement laser power, $A_{\mathrm{ref}} = A_{\mathrm{ref}}(P_{\mathrm{ref}})$, so

$$\frac{\partial \phi_{\mathrm{ref}}}{\partial(\delta P_{\mathrm{ref}}/P_{\mathrm{ref}})} = \frac{\Delta\theta_{\mathrm{ref}}}{A_{\mathrm{ref}}^2} \frac{\partial A_{\mathrm{ref}}}{\partial(\delta P_{\mathrm{ref}}/P_{\mathrm{ref}})}. \tag{4.28}$$

Imperfection	Corresponding parameter	Typical value	Uncertainty
$P_{\text{ref}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/P_{\text{ref}}^{\text{nr}}$	$\Omega_r^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\Omega_r^{\text{nr}}$	-3.1×10^{-3}	0.3×10^{-3}
$\Delta\theta_{\text{ref}}$	STIRAP/refinement alignment	$\sim 0.1°$	—

Quantity	Value	Uncertainty	Units
$\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial(P_{\text{ref}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/P_{\text{ref}}^{\text{nr}})$	24	35	mrad/sec
$d\omega_{\delta P}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$	-74	107	μrad/sec

Table 4.4: Summary of $P_{\text{ref}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ systematic parameters and shift. The imperfection, $P_{\text{ref}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/P_{\text{ref}}^{\text{nr}}$, is a correlated refinement laser power, or possibly an equivalent correlated refinement laser Rabi frequency (which might not arise from actual laser power correlations). Together with misalignment, $\Delta\theta_{\text{ref}}$, of the refinement laser polarization relative to the STIRAP-prepared phase orientation, this correlated laser power can lead to a systematic slope, $\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial(P_{\text{ref}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/P_{\text{ref}}^{\text{nr}})$. The systematic slope is minimized by setting $\Delta\theta_{\text{ref}} \approx 0$. We compute the systematic slope by comparing $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ in the presence of a deliberately applied large value of $P_{\text{ref}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/P_{\text{ref}}^{\text{nr}} = 0.1$, and the value of $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ under ideal conditions. We compute a bound on the systematic shift, $d\omega_{\delta P}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = \partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial(P_{\text{ref}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/P_{\text{ref}}^{\text{nr}}) \times P_{\text{ref}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/P_{\text{ref}}^{\text{nr}}$, using the measured systematic slope (consistent with zero) and a bound on the typical value of $P_{\text{ref}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/P_{\text{ref}}^{\text{nr}}$, which is found as follows. We calibrate the effect of a deliberately applied large value of $P_{\text{ref}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/P_{\text{ref}}^{\text{nr}}$ on $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}}$ in the presence of a large magnetic field, $\mathcal{B} = 26$ mG. We then infer a bound on $P_{\text{ref}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/P_{\text{ref}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ under ideal conditions using the observed value of $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}}$ for all data with $\mathcal{B} = 26$ mG in the EDM data set and the calibrated slope $\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}}/\partial(P_{\text{ref}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/P_{\text{ref}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}})$, which arises due to well-understood AC Stark shift effects.

With $d\omega_{\delta P}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} \equiv \frac{1}{\tau}\frac{\partial\phi_{\text{ref}}}{\partial(\delta P_{\text{ref}}/P_{\text{ref}})}\frac{P_{\text{ref}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{P_{\text{ref}}^{\text{nr}}}$ and $\frac{\partial\omega}{\partial(\delta P_{\text{ref}}/P_{\text{ref}})} \equiv \frac{1}{\tau}\frac{\partial\phi_{\text{ref}}}{\partial(\delta P_{\text{ref}}/P_{\text{ref}})}$, we obtain Eq. 4.27.

In order to minimize the overall contribution from this systematic error, we minimize $\Delta\theta_{\text{ref}}$. To do this, we amplify $\frac{d\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{d(\Delta\theta_{\text{ref}})}$ by deliberately applying a large $P_{\text{ref}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/P_{\text{ref}}^{\text{nr}}$. Then in this configuration, $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} \approx 0$ where $\Delta\theta_{\text{ref}} \approx 0$. Since $P_{\text{ref}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/P_{\text{ref}}^{\text{nr}} \ll 1$ under ordinary conditions, the systematic contribution is then proportional to two small quantities.

Under optimized conditions, $\Delta\theta_{\text{ref}} \approx 0$, we directly measure the systematic error slope $\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial(P_{\text{ref}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/P_{\text{ref}}^{\text{nr}}) = 24(35)$ mrad/sec by performing a linear regression on $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ under ordinary conditions and conditions with a deliberately applied $P_{\text{ref}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/P_{\text{ref}}^{\text{nr}} = 0.1$. Data with a large value of $P_{\text{ref}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/P_{\text{ref}}^{\text{nr}}$ is taken approximately once per day. This systematic error slope is consistent with zero.

Next, we must measure the intrinsic correlated power (or equivalent correlated Rabi frequency), $\Delta P_{\text{ref}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$, under ordinary conditions. To do this, we use the AC Stark shift effect of Eq. 4.26. The intentional parameter variation data with $P_{\text{ref}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/P_{\text{ref}}^{\text{nr}} = 0.1$ and $\mathcal{B} = 26$ mG (corresponding to $\phi^{\tilde{\mathcal{B}}} \approx \frac{\pi}{4}$ rad) creates a large $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}} = d\theta_{\text{eff}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}}/\tau$. Given the

known, large $P_{\text{ref}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/P_{\text{ref}}^{\text{nr}}$ for this intentional parameter variation data, we compute $d\omega_{\delta P}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}} = \frac{P_{\text{ref}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{P_{\text{ref}}^{\text{nr}}} \frac{\mathcal{B}}{\text{mG}} \times 168(4) \frac{\text{mrad}}{\text{sec}}$. Using this relation, for ordinary EDM data at $\mathcal{B} = 26$ mG (where the AC Stark shift effect is amplified), we compute $P_{\text{ref}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/P_{\text{ref}}^{\text{nr}} = \omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}}[\text{mrad/sec}]/(26 \times 168)$, which is valid under the assumption that the only source of $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}}$ arises due to the AC Stark shift effect with a non-zero intrinsic $\Omega_r^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ (effective $P_{\text{ref}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$). We find $P_{\text{ref}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/P_{\text{ref}}^{\text{nr}} = -3.1(3)\times10^{-3}$, quite inconsistent with zero and much larger than the values (consistent with zero) obtained from continuously monitored mesaurements of the refinement laser power[6], $P_{\text{ref}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/P_{\text{ref}}^{\text{nr}}[\text{logged}] = -1.5(14) \times 10^{-5}$. The inferred shift due to this systematic would therefore be $d\omega_{\delta P}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = \partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial(P_{\text{ref}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/P_{\text{ref}}^{\text{nr}}) \times P_{\text{ref}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/P_{\text{ref}}^{\text{nr}} = -74(107)\,\mu\text{rad/sec}$, consistent with zero. We have decided to include the uncertainty in this systematic error contribution in our systematic error budget, but have not subtracted the shift of $-74\,\mu\text{rad/sec}$ because we have only indirect evidence that the measured value of $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}} \neq 0$ arises entirely due to $P_{\text{ref}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} \neq 0$ (or an equivalent $\Omega_r^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} \neq 0$). For a summary of parameters related to the $P^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ systematic, see Table 4.4.

4.8 Effects of a correlated $\mathcal{B}^{\tilde{\mathcal{E}}}$

A well-understood method of generating an EDM systematic arises through two simultaneous imperfections. If a component of the applied magnetic field is correlated with the $\tilde{\mathcal{E}}$ switch, such that $\mathcal{B}_z = \mathcal{B}\mathcal{B}^{\tilde{\mathcal{B}}} + \mathcal{B}^{\tilde{\mathcal{E}}}\tilde{\mathcal{E}}$, then we would expect a resulting $\tilde{\mathcal{E}}$-correlated Zeeman precession, $\omega^{\tilde{\mathcal{E}}} = -g\mu_B\mathcal{B}^{\tilde{\mathcal{E}}}$. If, furthermore, the g-factor does not perfectly reverse between $\tilde{\mathcal{N}} = \pm1$ states, $g = g^{\text{nr}} + \Delta g\tilde{\mathcal{N}}$, then we will have an EDM systematic shift,

$$d\omega_{\Delta g}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = -\Delta g\mu_B\mathcal{B}^{\tilde{\mathcal{E}}} = \frac{\Delta g}{g}\omega^{\tilde{\mathcal{E}}}. \tag{4.29}$$

An $\tilde{\mathcal{E}}$-correlated magnetic field could arise, for example, due to a finite impedance between the electric field plates. In this case, current would flow between the field plates in a direction determined by the sign of the applied voltage, i.e., the $\tilde{\mathcal{E}}$ state. This current could generate a magnetic field component $\mathcal{B}^{\tilde{\mathcal{E}}}$.

---

6. I quote the central value here as the median, since the values among superblocks are non-normally distributed.

The second ingredient necessary to cause an EDM systematic $d\omega_{\Delta g}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ is a difference in the g-factor between $\tilde{\mathcal{N}} = \pm 1$. Due to perturbations to the molecular states, we know that such a g-factor difference exists[146, Sec. 2.1.3][7]:

$$g(J,\tilde{\mathcal{N}},\mathcal{E}) = g(J) + \eta(J)\mathcal{E}\tilde{\mathcal{N}}, \qquad (4.30)$$

where in ACME I, it was found that $\eta(J = 1) = -0.79(1)$ nm/V, consistent with the value measured in the ACME II data set, $\eta(J = 1) = -0.82(1)$ nm/V. Thus $\Delta g = \eta(J)\mathcal{E}$. Recall that η is computed from the $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{B}}}$ channel, $\eta = -\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{B}}}/(\mu_B\mathcal{E}\mathcal{B})$. The statistical sensitivity of η in the ACME II dataset is not much improved over ACME I in large part because the sensitivity of the measured value of η improves at higher magnetic field $\mathcal{B}$, but in ACME II we take only a small fraction of data at higher magnetic fields, $\mathcal{B}_{\frac{\pi}{4}}$ such that $\phi \approx \frac{\pi}{4}$, and no data at the highest ACME I magnetic fields, where $\phi \approx \frac{\pi}{2}$. Furthermore, the data we do take at $\mathcal{B}_{\frac{\pi}{4}}$ has a higher level of noise than our ordinary data, as discussed in Sec. 4.15.2.

To confirm the systematic error model Eq. 4.29, we applied a deliberately large $\mathcal{B}^{\tilde{\mathcal{E}}}$ and observed a systematic error slope $\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial\omega^{\tilde{\mathcal{E}}} = 1.5(2) \times 10^{-3}$ when $\mathcal{E} = 80$ V/cm, exactly as expected: at this electric field, $\Delta g = \eta\mathcal{E} = -6.4 \times 10^{-6}$ so that $\Delta g/g = 1.5 \times 10^{-3}$.

We also monitor $\omega^{\tilde{\mathcal{E}}}$ throughout the final data set and observe a value consistent with zero. The overall systematic error contribution requires the slope $\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial\omega^{\tilde{\mathcal{E}}}$ to be adjusted in proportion to $\mathcal{E}$ for each subset of the entire data set, consistent with the model for $\eta \neq 0$. The resulting systematic error shift averaged over the entire data set is $d\omega_{\Delta g}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = -0.7(5)\,\mu$rad/sec, consistent with zero and several orders of magnitude smaller than the statistical sensitivity of the EDM measurement.

The systematic error shift computed instead using the value of η measured in each superblock, rather than using the systematic error slope $\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial\omega^{\tilde{\mathcal{E}}}$ measured under conditions of a deliberately applied $\mathcal{B}^{\tilde{\mathcal{E}}}$, is consistent with the shift quoted above, with comparable uncertainty: $d\omega_{\Delta g}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}[\eta] = 1.0(6)\,\mu$rad/sec. However, this value is merely a sanity check and is not used in the systematic error budget. For a summary of parameters related to the $\omega^{\tilde{\mathcal{E}}}$

---

7. Components beyond those written here – for example, a term that depends on $\tilde{\mathcal{N}}$ but *not* $\mathcal{E}$ – are small and typically neglected in our analysis.

Imperfection	Corresponding parameter	Typical value	Uncertainty
$\omega^{\tilde{\mathcal{E}}}$	$\mathcal{B}^{\tilde{\mathcal{E}}}$	$790\,\mu\text{rad/sec}$	$420\,\mu\text{rad/sec}$
η	$\mathcal{N}$-dependent g-factor	-0.82 nm/V	0.01 nm/V
Quantity	**Value**	**Uncertainty**	**Units**
$\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial\omega^{\tilde{\mathcal{E}}}$	1.5×10^{-3}	0.2×10^{-3}	—
$d\omega_{\Delta g}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$	1.0	0.6	$\mu\text{rad/sec}$

Table 4.5: Summary of $\omega^{\tilde{\mathcal{E}}}$ systematic parameters and shift. The imperfection, $\omega^{\tilde{\mathcal{E}}}$, is a precession frequency that is correlated with $\tilde{\mathcal{E}}$, which can arise due to a correlated magnetic field, $\mathcal{B}^{\tilde{\mathcal{E}}}$. Together with the non-zero difference between the g-factors of the $\tilde{\mathcal{N}} = \pm 1$ states (characterized by the molecular constant η), this correlated precession frequency can "leak" into the EDM channel to produce a systematic slope, $\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial\omega^{\tilde{\mathcal{E}}}$. We compute the systematic slope by comparing $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ in the presence of a deliberately applied large value of $\mathcal{B}^{\tilde{\mathcal{E}}} \approx 1.3$ mG (producing $\omega^{\tilde{\mathcal{E}}} \approx 50$ rad/sec), and the value of $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ under ideal conditions. Because the systematic slope depends on $\mathcal{E}$, we compute a weighted average, $d\omega_{\Delta g}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$, of the systematic shifts under different conditions throughout the EDM data set. For a given value of $\mathcal{E}$, the systematic shift, $d\omega_{\Delta g}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = \frac{\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{\partial\omega^{\tilde{\mathcal{E}}}}\omega^{\tilde{\mathcal{E}}}$, is computed using the measured systematic slope when $\mathcal{E} = 80$ V/cm (value of slope shown in table, consistent with zero), adjusted for the relevant magnitude of the electric field, and the average value of $\omega^{\tilde{\mathcal{E}}}$ measured for the subset of EDM data with the relevant value of $\mathcal{E}$.

systematic, see Table 4.5.

4.9   Frequency-contrast correlations

We have observed a well-understood correlation between $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ and certain components of the contrast magnitude, namely $|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}}$ and $|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ (see Fig. 4.4). This correlation is a necessary consequence of the large $\phi^{\tilde{\mathcal{B}}}$ and ϕ^{nr} phase components. We can see this by supposing that $|\mathcal{C}|^{\tilde{s}} \ll |\mathcal{C}|^{\text{nr}}$ for any combination of switches $\tilde{s}$, but that $|\mathcal{C}|^{\tilde{s}} \neq 0$ precisely either due to an intrinsic offset to the parity component or due to random measurement noise. Then by Taylor expanding the defining relation $\phi = \frac{\mathcal{A}}{2|\mathcal{C}|}$ and considering only the resulting component that is odd under $\tilde{\mathcal{N}}\tilde{\mathcal{E}}$, we find

$$\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = \frac{\mathcal{A}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{2\tau|\mathcal{C}|^{\text{nr}}} - \omega^{\tilde{\mathcal{B}}}\frac{|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}}}{|\mathcal{C}|^{\text{nr}}} - \omega^{\text{nr}}\frac{|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{|\mathcal{C}|^{\text{nr}}} - \cdots , \qquad (4.31)$$

where in fact there is a correction term for every frequency component $\omega^{\tilde{s}}$. However, the frequency components due to Zeeman precession and non-reversing offset angles between
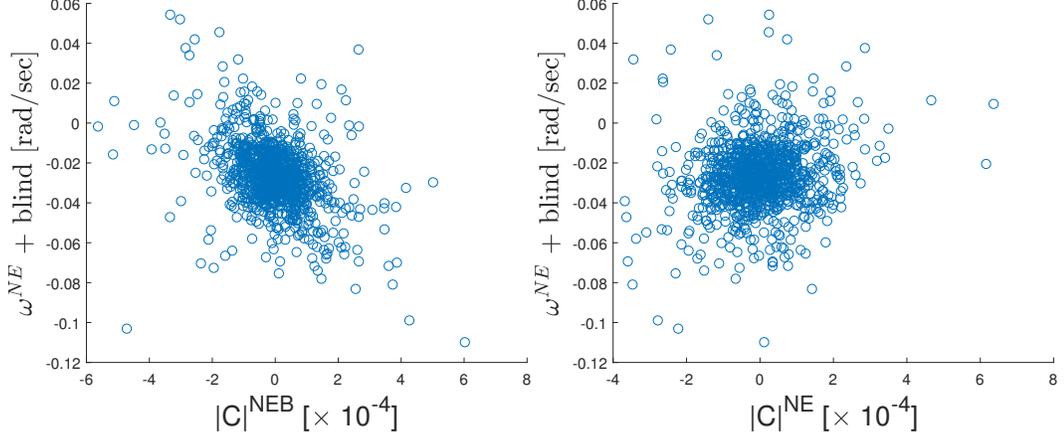
152

Figure 4.4: Scatter plot data showing correlations between $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ with $|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}}$ and $|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ for all superblocks in the EDM data set. Error bars omitted for clarity.

the readout and preparation bases, $\omega^{\tilde{\mathcal{B}}}$ and ω^{nr}, respectively, are much larger than any other frequency components. As a result, only the terms explicitly appearing in Eq. 4.31 are statistically significant at the ACME II sensitivity.

Let us consider, for the moment, the possibility that there is no intrinsic correlation of the contrast magnitude with the experimental switches listed. However, in any given experiment we will typically have $\langle |\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}}\rangle \sim \langle |\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}\rangle \sim \sigma_{|\mathcal{C}|} \neq 0$, i.e., the average measured value will be on the same order as the uncertainty in the contrast (averaged over the entire dataset). This will have the effect of introducing shifts in $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ that are proportional to $\omega^{\tilde{\mathcal{B}}}$ and ω^{nr}, respectively.

When exploring correlations in the dataset, we discovered that the average Pearson correlation coefficient of $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ vs. $|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}}$ is the fairly remarkable value of -0.45, which is significant at the 15σ level, due to the linear dependence of $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ on $|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}}$ seen in Eq. 4.31. Likewise, the correlation coefficient between $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ and $|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ is 0.08, which is significant at 2.7σ. No other correlations between $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ and contrast channels was significant, as expected based on Eq. 4.31 and the relative size of other frequency parity components of ω compared to $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}}$ and ω^{nr}.

As expected, the average values of $|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}}$ and $|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ are consistent with zero over the entire data set, $\langle |\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}}\rangle = -2.4(77) \times 10^{-6}$ and $\langle |\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}\rangle = -8.5(45) \times 10^{-6}$.

To subtract or not?

It should perhaps *not* be obvious that the "true" value of the EDM is obtained by subtracting the contributions due to $\omega^{\tilde{\mathcal{B}}}|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}}/|\mathcal{C}|^{\mathrm{nr}}$ and $\omega^{\mathrm{nr}}|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/|\mathcal{C}|^{\mathrm{nr}}$. Doing so would be equivalent to calculating the EDM from $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} \to \frac{[\mathcal{A}\mathrm{sign}(\mathcal{C})]^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{2\tau|\mathcal{C}|^{\mathrm{nr}}}$ alone, rather than $\left[\frac{\mathcal{A}\mathrm{sign}(\mathcal{C})}{2\tau|C|}\right]^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$. Each choice has distinct merits. In ACME I, it was observed that (due to the different state preparation method used there; see [146, Appendix E]), components of the contrast were susceptible to significant correlations; e.g., $\langle|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}\rangle \neq 0$. In this case, $\frac{\mathcal{A}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{2|\mathcal{C}|^{\mathrm{nr}}}$ would be $\tilde{\mathcal{N}}\tilde{\mathcal{E}}$-correlated even if the physical molecular phase ϕ were completely uncorrelated with any switches (an *uncorrelated* phase, together with a *correlated* contrast, necessarily implies a compensatingly correlated asymmetry). Thus $\phi_{\mathrm{biased}}^{\mathrm{EDM}} = \frac{\mathcal{A}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{2|\mathcal{C}|^{\mathrm{nr}}}$ would give a biased estimate of the EDM value.

However, we have every indication that $|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ and $|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}}$ are *not* biased away from 0 on average at the level of our statistical sensitivity: they are only non-zero at a level that is expected from statistical noise alone, and we know of no mechanism by which we should expect a bias in $|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ or $|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}}$ in ACME II.

In this case, a sufficiently large "systematic slope" such as $\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ couples to a statistically random offset $|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ to introduce a shift in $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$. An estimate of $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ using Eq. 4.31 is unbiased in the sense that the shift due to $\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ has a random magnitude and direction. However, for any given run of the experiment, the shift will move $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ away from its ideal, long-run average value of $\frac{\mathcal{A}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{2\tau|\mathcal{C}|^{\mathrm{nr}}}$ (in the case that no components of the contrast are intrinsically correlated with experimental switches).

Based on this reasoning, we have decided to subtract the contributions to $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ from the statistically significant "systematic slopes" $\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ and $\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}}$ and the average values of $|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ and $|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}}$ over the EDM data set. This choice reflects judgment that the "true," long-run average values of $|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ and $|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}}$ are likely to be closer to 0 than to their measured values (which are, once again, consistent with 0 at the expected statistical level).

Since we know that $\omega^{\tilde{\mathcal{B}}}$ depends on $\mathcal{B}$, and furthermore ω^{nr} can generally depend on the experimental configuration, we separately compute systematic slopes $\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ and

Imperfection	Corresponding parameter	Typical value	Uncertainty
$\langle|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\mathcal{B}}\rangle$, $\langle|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}\rangle$	Statistical fluctuations	$\sim 10^{-5}$	—
$\omega^{\mathcal{B}}$, ω^{nr}	Zeeman, phase offsets	$\lesssim 10^{3}$ rad/sec	—

Quantity	Value	Uncertainty	Units
$\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}(\mathcal{B})}$	~ 50	—	rad/sec
$d\omega_{|\mathcal{C}|}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$	77	125	μrad/sec

Table 4.6: Summary of $|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}}$ and $|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ systematic parameters and shifts. The imperfections, $\langle|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}}\rangle$ and $\langle|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}\rangle$, are correlated components of the contrast magnitude, which arise due to statistical noise in the contrast measurement. Together with large precession frequency components, $\omega^{\mathcal{B}}$ and ω^{nr} (arising from Zeeman precession and global phase offsets, respectively), these statistical fluctuations produce systematic slopes, $\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}}$ and $\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$. We compute the systematic slopes from a linear fit of $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ vs. $|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}}$ and $|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ using data under ideal conditions throughout the EDM data set. Because the systematic slopes may depend on $\mathcal{E}$ and $\mathcal{B}$, we compute a weighted average, $d\omega_{|\mathcal{C}|}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$, of the systematic shifts with each value of $\mathcal{E}$ and $\mathcal{B}$. For a given value of $\mathcal{E}$ and $\mathcal{B}$, a systematic shift, $d\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = \frac{\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{\partial|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}(\mathcal{B})}}\omega^{\mathrm{nr}(\tilde{\mathcal{B}})}$, is computed using the measured linear fit and $\langle|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}(\tilde{\mathcal{B}})}\rangle$ measured in the corresponding subset of the EDM dataset. Representative values of $\langle|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}(\tilde{\mathcal{B}})}\rangle$ and the systematic slopes are given in the table.

$\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}}$, as well as offset values $\langle|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}\rangle$ and $\langle|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}}\rangle$, separately for each value of $\mathcal{B}$ and $\mathcal{E}$ throughout the data set. We then propagate the respective (eight) systematic shifts and corresponding uncertainties in proportion to the relative contributions of each $\mathcal{B}$, $\mathcal{E}$ configuration to the statistical sensitivity of the EDM measurement. The resulting total shift, from both correlations in all $\mathcal{B}$ and $\mathcal{E}$ states, is $d\omega_{|\mathcal{C}|}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = 77(125)\,\mu$rad/sec, consistent with zero. For a summary of parameters related to the $|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}}$ and $|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ systematics, see Table 4.6.

4.10 Magnetic field gradient systematics

The final set of systematic errors that contribute a shift to $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ arise from ambient magnetic field gradients in the interaction region, which do not reverse under any experimental switches. We can amplify the effect of these ambient gradients by deliberately applying large values of, e.g., $(\partial\mathcal{B}/\partial z)^{\mathrm{nr}}$ and $(\partial\mathcal{B}/\partial y)^{\mathrm{nr}}$, where I'll typically drop the nr superscript in this section. As in previous sections, the only important cartesian component of $\vec{\mathcal{B}}$ is $\mathcal{B}_z$, and we continue to write $\mathcal{B} \equiv \mathcal{B}_z$ for shorthand. See Fig. 4.5 for each systematic prior to
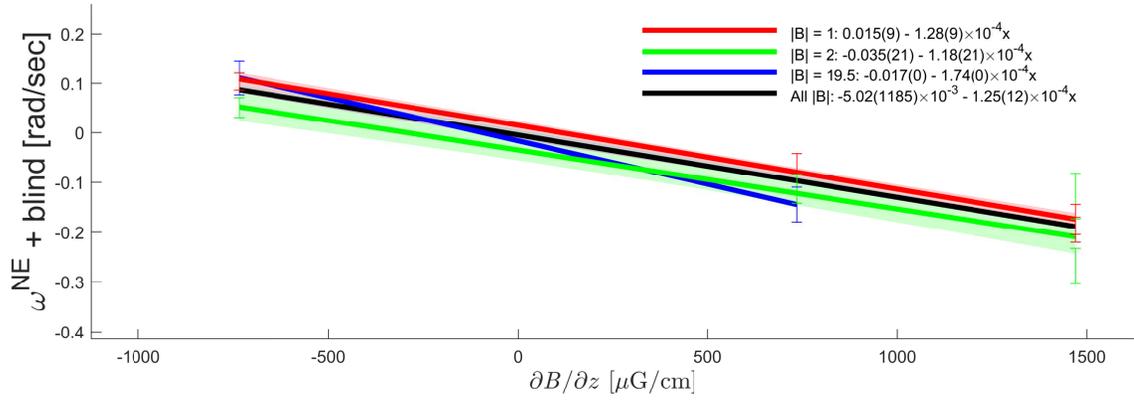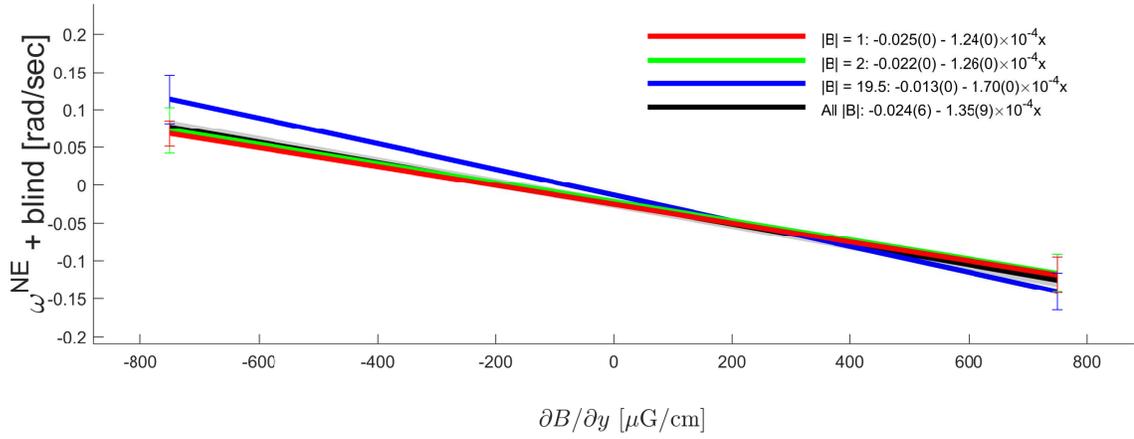
Figure 4.5: (a) $\partial \mathcal{B}/\partial z$ systematic before suppression. (b) $\partial \mathcal{B}/\partial y$ systematic before suppression. Both systematic slopes are comparable. Data from Runs 150, 151, and 153.

suppression.

Let us consider the effect of $\partial\mathcal{B}/\partial z$ first. We initially observed a systematic slope $\frac{\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{\partial(\partial\mathcal{B}/\partial z)} \sim 100\,\mathrm{mrad/(mG/cm)}$, though the systematic slope shifted by up to $\sim 50\%$ over many-hour to many-day timescales. We can understand this effect using the formalism of Sec. 4.2.1. If there is a spatially-dependent, $\tilde{\mathcal{N}}\tilde{\mathcal{E}}$-correlated component of the molecular distribution that is prepared or read out, $\delta\rho(\vec{x})$, then a spatially-dependent distribution of the molecular phase induced by a magnetic field gradient, $\delta\phi(\vec{x})$, will contribute to a systematic shift in the average measured phase of $d\bar{\phi} = \int dV\, \epsilon\delta\rho^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}\delta\phi^{\mathrm{nr}}/N^{\mathrm{nr}}$, where $\epsilon(\vec{x})$ is the detection efficiency as a function of the molecular position in the readout region. As we saw previously, in the special case of a constant phase gradient $\phi = \phi_0 + \frac{\partial\phi}{\partial z}z$, the phase measured over the entire population distribution is modulated by $d\bar{\phi}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = \frac{\partial\phi}{\partial z}z_{\mathrm{CM}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$, where $z_{\mathrm{CM}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ is an $\tilde{\mathcal{N}}\tilde{\mathcal{E}}$-correlated shift in the center of mass of the detected molecular population distribution. Here, $\frac{\partial\phi}{\partial z} = -g\mu_B\frac{\partial\mathcal{B}}{\partial z}\tau$ is proportional to the applied magnetic field gradient along z.

The usual suspect for the source of $\tilde{\mathcal{N}}\tilde{\mathcal{E}}$-dependence is $\mathcal{E}^{\mathrm{nr}}$. After seeing the magnetic field gradient systematic, Cris Panda quickly realized that a non-zero $\mathcal{E}^{\mathrm{nr}}$ can cause an $\tilde{\mathcal{N}}\tilde{\mathcal{E}}$-correlated center of mass of the molecular distribution by the following mechanism (see Fig. 4.7):

1. The molecules' ballistic trajectories generate a very strong position-velocity correlation, $z \approx \frac{v_z}{v_x}L$, in the interaction region at a distance $L \approx 1.1$ m downstream of the source (correlation coefficient ~ 0.98).

2. The detuning Δ for a particular velocity class is proportional to its transverse velocity v_z due to the Doppler shift $\Delta_{\mathrm{Dop}} = v_z/\lambda$, where $\lambda = 703$ nm is the transition wavelength.

3. Since the readout beams are not fully saturated, there is a special velocity class (value of v_z) for which $\Delta = \Delta_0 + \Delta_{\mathrm{Dop}} = 0$, which is preferentially detected, where Δ_0 includes all contributions to the detuning other than the Doppler shift. In particular, the velocity class on resonance, for a given value of Δ_0, is $v_z^{\mathrm{res}} = -\Delta_0\lambda$.
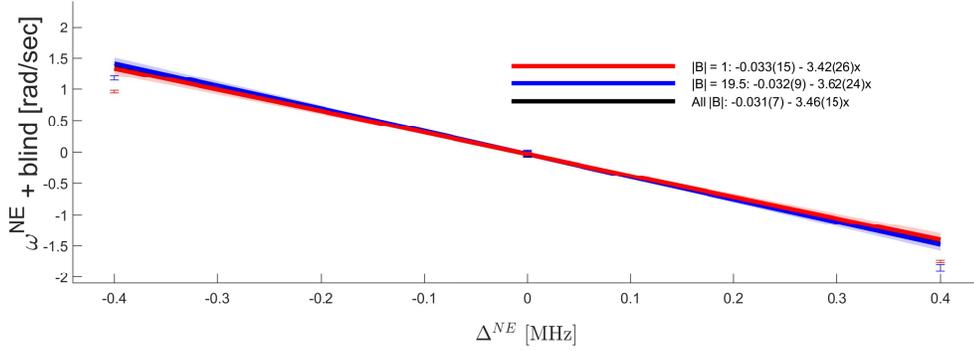
Figure 4.6: $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ vs. $\Delta^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ (due to deliberately applied $\mathcal{E}^{\mathrm{nr}}$) in the presence of $\frac{\partial \mathcal{B}}{\partial z} = 0.74$ mG/cm. The slope is consistent with the model of an $\tilde{\mathcal{N}}\tilde{\mathcal{E}}$-correlated shift in the molecular center of mass, $z_{\mathrm{CM}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$, probed by the readout beam due to a combination of the position-velocity correlation in the probe region and the Doppler shift $\Delta_{\mathrm{Dop}} \propto v_z$, for a readout laser linewidth of 2.5 MHz (1σ). Data from Run 156.

4. Due to the correlation between position and velocity, there is therefore a preferred position class (value of z) that is most efficiently detected: $z_{\mathrm{res}} \approx -\Delta_0 \lambda L / v_x$.

5. $\mathcal{E}^{nr} \sim 5$ mV/cm generates $\Delta_0^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} \sim 2\pi \times 5$ kHz in the cleanup and readout beams due to the Ω-doublet structure of the H state, so $z_{\mathrm{res}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} \sim -100\,\mu\mathrm{m}$.

6. Therefore, with a typical gradient $\partial \mathcal{B}/\partial z \sim 0.8$ mG/cm applied during systematic checks, we would expect up to $\phi^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} \sim 300\,\mu\mathrm{rad}$ due to this effect, in the limit of poor laser saturation.

In summary, there is a center-of-mass shift in the molecular population, which is correlated with $\tilde{\mathcal{N}}\tilde{\mathcal{E}}$, and this shift couples to a $\mathcal{B}$ gradient, producing a correlated average measured phase. Note that the simple estimates above apply to the limiting case of *poor* readout laser saturation. In the opposite limit of perfect readout laser saturation, there can be no systematic effect because the readout efficiency is independent of $\Delta^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ (generated by $\mathcal{E}^{\mathrm{nr}}$) and Δ_{Dop}. Prior to suppression of the systematic, measurements of $\frac{\partial \omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{\partial \Delta^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}$ in the presence of a deliberately applied $\frac{\partial \mathcal{B}}{\partial z} = 0.74\,\mathrm{mG/cm}$, where $\Delta^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ is generated via a deliberately applied $\mathcal{E}^{\mathrm{nr}}$, gave $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}\,[\mathrm{rad/sec}] = -3.46(2) \times \Delta^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}[2\pi \times \mathrm{MHz}]$.

We can validate the model with Monte Carlo trajectory simulations of the molecular trajectories in our apparatus, to compare with the observed magnitude of the systematic

error. We will follow approximately the chain of logic above, using the same numbered steps (though the correspondence is not exact between the detailed steps below and the individual lines in the heuristic model just discussed). The only complication, when computing the actual expected contribution of the systematic error, is that the lineshape of the transition must be considered.

First, (1) the simulations verify that $v_z\,[\mathrm{m/s}] \approx 1.2 \times z\,[\mathrm{cm}]$. (2-3) The LHS of this relationship, a velocity, has a one-to-one correspondence with a Doppler shift, $\Delta_{\mathrm{Dop}} = \frac{v_z}{\lambda}$, where $\lambda = 703$ nm is the wavelength of the $H \leftrightarrow I$ transition. (4) The RHS of the same relationship, a position, has a one-to-one correspondence with a Zeeman precession phase in the presence of a magnetic field gradient, $\phi = -\tau \times g_H \mu_B \frac{\partial \mathcal{B}}{\partial z} z$. With a typical systematic check gradient of 0.8 mG/cm, we therefore obtain a correlated phase and detuning (i.e., Doppler shift), $\frac{\Delta_{\mathrm{Dop}}}{2\pi \times \mathrm{MHz}} \approx -0.06\,\frac{\phi}{\mathrm{mrad}}$, where the sign arises from the negative g-factor of the H state.

To proceed with our model of the systematic effect using the Monte Carlo simulations, we must consider the effect of the readout transition lineshape, which was not addressed in the simplest version of our model above (i.e., the numbered list of steps). We assume a detected population fraction of the ensemble, $\epsilon(\Delta_{\mathrm{Dop}}) \propto \exp[-(\Delta_{\mathrm{Dop}} - \Delta_0)^2/(2\sigma_\Delta^2)]$, that is approximately normally distributed as a function of Δ_{Dop} with $\sigma_\Delta = 2\pi \times 2.5$ MHz (based on Doppler scans of the readout transition), relative to the overall detuning Δ_0. In the Monte Carlo simulation, we compute a weighted average of measured phases, $\bar{\phi} = \sum_i w_i \phi_i / \sum_i w_i$, where ϕ_i is determined by the local Zeeman precession (a function of the molecule's position in the interaction region), and the weight factor $w_i = \epsilon(\Delta_{\mathrm{Dop}})$ is determined by the molecule's Doppler shift (a function of its transverse velocity). We then find a linear dependence of the average measured phase on small detunings, $\frac{\bar{\phi}}{\mathrm{mrad}} = -4(1) \times \frac{\Delta_0}{2\pi \times \mathrm{MHz}} \leftrightarrow \frac{\omega}{\mathrm{rad/sec}} = -4(1) \times \frac{\Delta_0}{2\pi \times \mathrm{MHz}}$. Here, I've determined the uncertainty by variations in the simulation results over reasonable parameter ranges (e.g., assumed position and velocity distributions of the molecular source).

Next, we relate the detuning Δ_0 above to $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ as follows: (5) The detuning Δ_0 becomes correlated with $\tilde{\mathcal{N}}\tilde{\mathcal{E}}$ in the presence of $\mathcal{E}^{\mathrm{nr}}$, via $\Delta_0 \to \Delta^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = D\mathcal{E}^{\mathrm{nr}}$. (6) Then we expect $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}\,[\mathrm{rad/sec}] = -4(1) \times \Delta^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}[2\pi \times \mathrm{MHz}]$ for the systematic model considered here, in

excellent agreement with the observed value of $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ [rad/sec] $= -3.46(2) \times \Delta^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}[2\pi \times \text{MHz}]$.

With $\Delta_0^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} \sim 5$ kHz (a typical value due to the ambient value of $\mathcal{E}^{\text{nr}}$), we therefore expect $d\phi^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} \sim 20\,\mu\text{rad} \leftrightarrow d\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} \approx 20\,\text{mrad/sec}$. This is a large but incomplete fraction of the observed $\frac{\partial \mathcal{B}}{\partial z}$ systematic before suppression ($d\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} \sim 100$ mrad/sec with ~ 1 mG/cm applied $\frac{\partial B}{\partial z}$). Hence, we concluded that there were likely other contributions to the systematic shift associated with a large magnetic field gradient, $\frac{\partial \mathcal{B}}{\partial z}$.

$\frac{d\mathcal{E}^{nr}}{dz}$ component

Since the model just discussed explains only *part* of the overall magnetic field gradient systematic slope, $\partial \omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial(\partial \mathcal{B}/\partial z)$, we write $d\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = \omega_0^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} + \alpha_0(\mathcal{E}^{nr})\frac{\partial \mathcal{B}}{\partial z} + \alpha_1 \frac{\partial \mathcal{B}}{\partial z}$, where $\alpha_1 \neq \alpha_1(\mathcal{E}^{nr})$ is some constant that is independent of $\mathcal{E}^{\text{nr}}$ and is not explained by the systematic error model just discussed. In particular, before suppression, $\alpha_1 \sim 100\,(\text{mrad/sec})/(\text{mG/cm})$, but as mentioned previously its value would drift by $\sim 50\%$ from day to day. We understand this contribution to the systematic slope, which is independent of $\mathcal{E}^{\text{nr}}$, as arising from the following mechanism (see Fig. 4.8).

We once again need to identify a correlated shift in the center of mass, $z_{\text{CM}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$, that couples to the phase gradient $\frac{\partial \phi}{\partial z}$ induced by the applied magnetic field gradient $\frac{\partial \mathcal{B}}{\partial z}$. This time, we consider the effect of a non-reversing electric field *gradient* $\frac{\partial \mathcal{E}^{\text{nr}}}{\partial z}$, which we can measure to be on the order of $\sim 5\,\text{mV/cm}^2$ (see Sec. 5.1.1). This field gradient generates a correlated detuning gradient $\frac{\partial \Delta^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{\partial z} \sim 2\pi \times 5\,\text{kHz/cm}$.

Consider the effect of this gradient on the STIRAP state transfer. At any particular location, the value of $\mathcal{E}^{\text{nr}}(z)$ arising from the gradient $\frac{\partial \mathcal{E}^{\text{nr}}}{\partial z}$ creates a correlated two-photon STIRAP detuning $\delta^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = D\mathcal{E}^{\text{nr}}(z) = D\frac{\partial \mathcal{E}^{\text{nr}}}{\partial z}z$. Recall that the STIRAP beams travel vertically, along $\hat{y}$, and are spread out along z to address the entire molecular population. Since the efficiency η of the STIRAP state transfer depends on the two-photon detuning $\delta \equiv (\Delta_{690} - \Delta_{1090})/2$, we find $\eta(z) = \eta_0 + \frac{\partial \eta}{\partial \delta}\frac{\partial \delta^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{\partial z}z$, where η_0 is the state transfer efficiency at $z = 0$. In the simplified model that the population density of molecules entering the STIRAP beams is uniform along z with half-width a, the shift in the center-of-mass is $z_{\text{CM}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = \int_{-a}^{a} dz\, \eta(z)z / \int_{a}^{a} dz\, \eta(z) = \frac{a^2}{3\eta_0}\frac{\partial \eta}{\partial \delta}\frac{\partial \delta^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{\partial z}$.

To estimate the numerical size of this shift, we use a simple gaussian 2-photon lineshape
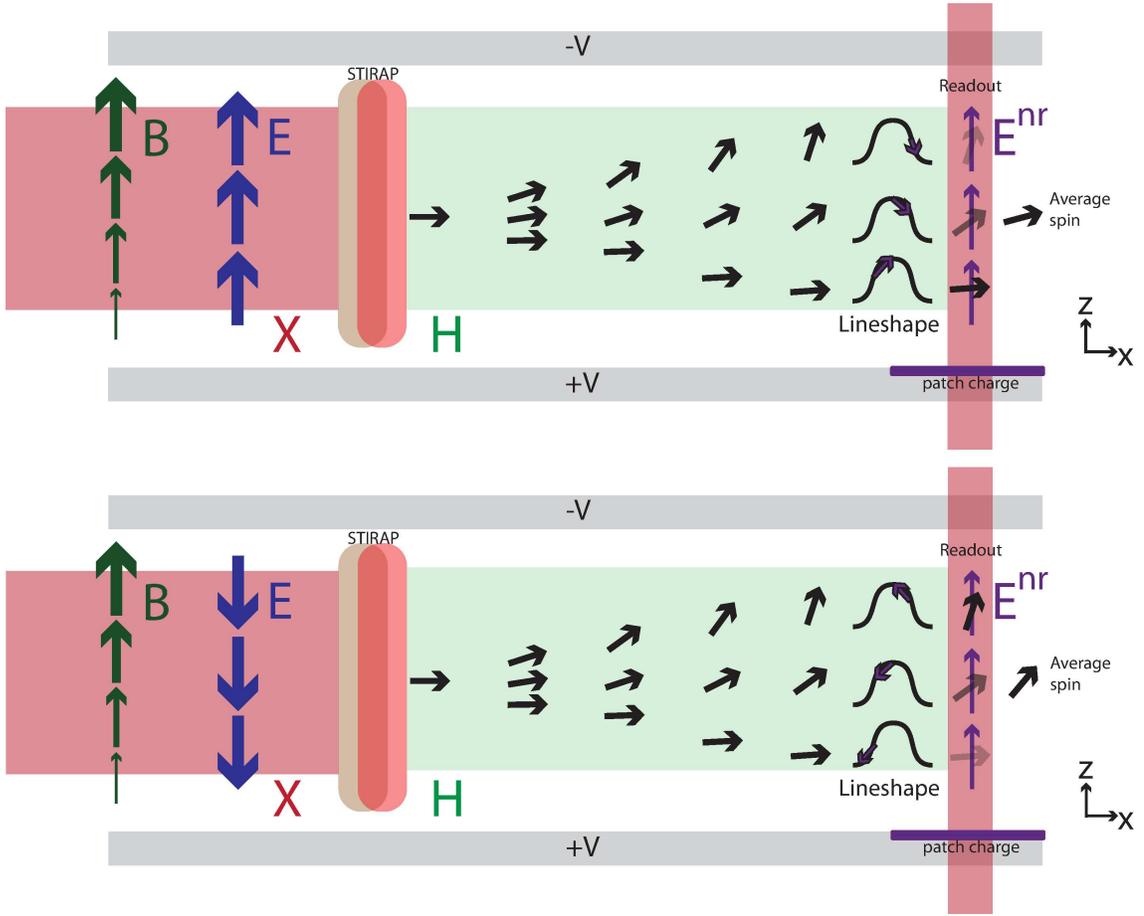
Figure 4.7: Representation of $\partial \mathcal{B}/\partial z \times \mathcal{E}^{\mathrm{nr}}$ systematic. The top (bottom) panel depicts the case where $\tilde{\mathcal{E}} = +1(-1)$; both panels show $\tilde{\mathcal{N}} = +1$. A magnetic field gradient, $\partial \mathcal{B}/\partial z$, is indicated by dark green arows. Molecules are prepared via STIRAP lasers in the H electronic state, uniformly along z (horizontal green band), from the initial X state population (horizontal pink band). Black arrows in precession region represent spin alignment of H state molecules. If the readout beam is on resonance with molecules that have $v_z = 0$, then it will be shifted off resonance (in opposite directions) for molecules with $\pm v_z$. The non-reversing electric field $\mathcal{E}^{\mathrm{nr}}$ (purple arrows, shown as originating from a patch charge on an electric field plate) also shifts the effective detuning, but acts on all molecules in the same way (purple arrows on the molecular lineshapes: right-pointing for $\tilde{\mathcal{E}} = +1$ and left-pointing for $\tilde{\mathcal{E}} = -1$). These effects combine so that one region of space is preferentially addressed (black arrows in readout region; opacity represents readout efficiency), and the magnetic field in this region disproportionately determines the measured phase. Because the detuning associated with $\mathcal{E}^{\mathrm{nr}}$ reverses with $\tilde{\mathcal{N}}$ and $\tilde{\mathcal{E}}$, the effect on the measured phase appears in the EDM channel, $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$.
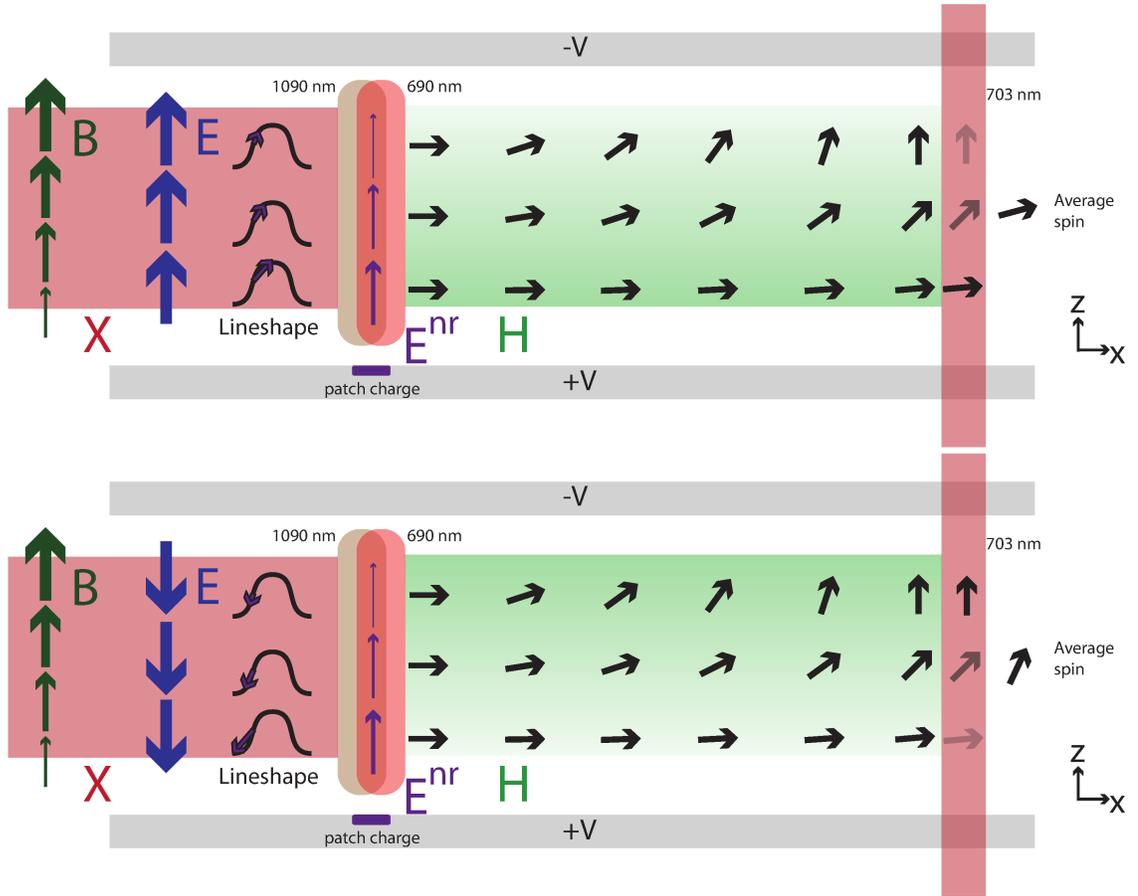
Figure 4.8: Representation of $\partial \mathcal{B}/\partial z \times \frac{\partial \mathcal{E}^{\mathrm{nr}}}{\partial z}$ systematic. The top (bottom) panel depicts the case where $\tilde{\mathcal{E}} = +1(-1)$; both panels show $\tilde{\mathcal{N}} = +1$. A magnetic field gradient, $\partial \mathcal{B}/\partial z$, is indicated by dark green arrows. A non-reversing electric field, $\mathcal{E}^{\mathrm{nr}}$ (purple arrows in STIRAP region), causes a detuning, preferentially preparing molecules where $\mathcal{E}^{\mathrm{nr}}$ cancels an overall two-photon detuning in the STIRAP lasers (purple arrows on lineshape). Due to the non-reversing electric field gradient, $\partial \mathcal{E}^{\mathrm{nr}}/\partial z$, this cancellation occurs either for molecules along $+\hat{z}$ or $-\hat{z}$, depending on the $\tilde{\mathcal{E}}$ state (compare top, bottom). Thus, the prepared population density in the H state is non-uniform along z (horizontal green band, shown with opacity gradient), despite the uniform initial X state population (horizontal pink band). Black arrows in precession region represent spin alignment of H state molecules, which is non-uniform along z due to $\partial \mathcal{B}/\partial z$. Opacity of spin arrows in readout region represent density of molecular population that is probed by the readout beam, and is non-uniform along z due to the gradient of molecular population prepared via STIRAP. Therefore, the average measured spin orientation changes in $\tilde{\mathcal{N}}\tilde{\mathcal{E}} = +1$ vs. $\tilde{\mathcal{N}}\tilde{\mathcal{E}} = -1$, leading to a non-zero value of $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$.
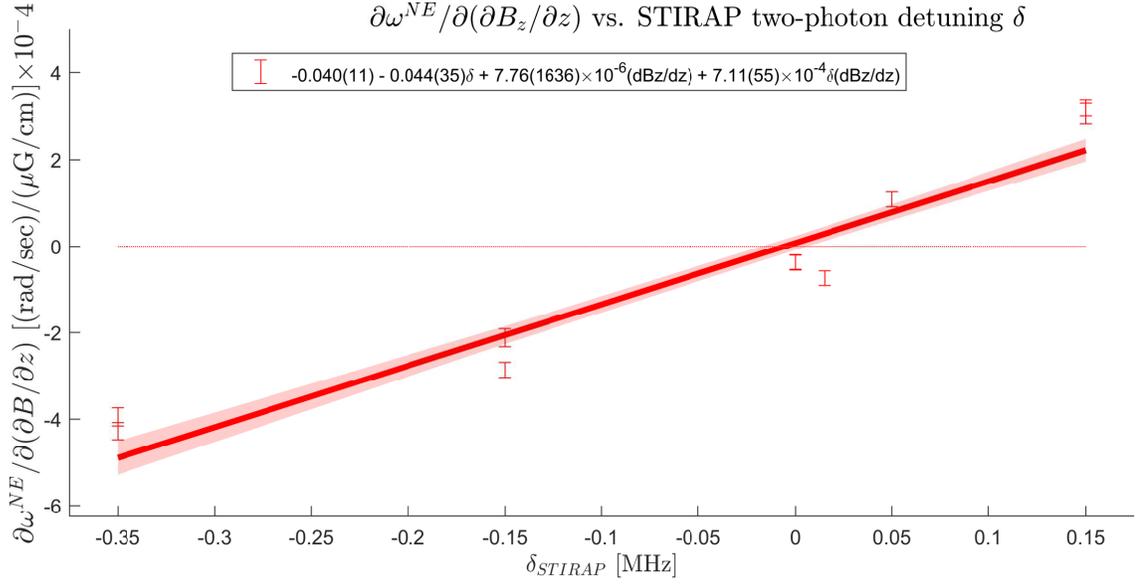
Figure 4.9: Linear dependence of systematic slope $\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial(\partial\mathcal{B}/\partial z)$ on two-photon STIRAP detuning δ. Here (Run 213), the detuning of the 690 nm laser was adjusted; we have verified that detuning the 1090 nm laser has an equal and opposite effect on the systematic slope as expected. By adjusting the two-photon detuning, we can tune the overall systematic slope $\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial(\partial\mathcal{B}/\partial z)$ to zero. A fit is shown, for the model $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = \omega_0^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} + \frac{\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{\partial\delta}\delta + \frac{\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{\partial(\partial\mathcal{B}/\partial z)}(\partial\mathcal{B}/\partial z) + \frac{\partial^2\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{\partial\delta\partial(\partial\mathcal{B}/\partial z)}\delta(\partial\mathcal{B}/\partial z)$, which has a clear dependence on the product $\delta(\partial\mathcal{B}/\partial z)$ but neither δ nor $\partial\mathcal{B}/\partial z$ alone.

model, $\eta(\delta) = \eta_0 \exp\left(-\frac{\delta^2}{2\sigma_{\mathrm{ST}}^2}\right)$ so that $\frac{\partial\eta}{\partial\delta} = -\eta_0 \frac{\delta}{\sigma_{\mathrm{ST}}^2} \exp\left(-\frac{\delta^2}{2\sigma_{\mathrm{ST}}^2}\right)$, where σ_{ST} is the 1σ two-photon linewidth. Note that in the limit of perfect STIRAP laser saturation, $\sigma_{\mathrm{ST}} \to \infty$, the transfer efficiency is independent of detuning and therefore no systematic effect can occur. For a finite STIRAP two-photon linewidth, and small two-photon detunings $\delta \ll \sigma_{\mathrm{ST}}$,

$$z_{\mathrm{CM}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} \approx \frac{a^2}{3}\frac{\delta}{\sigma_{\mathrm{ST}}^2} D \frac{\partial \mathcal{E}^{\mathrm{nr}}}{\partial z}. \tag{4.32}$$

This gives us a "knob" to tune out the systematic slope $\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial(\partial\mathcal{B}/\partial z)$, namely the two-photon detuning δ; see Fig. 4.9. By adjusting this detuning to $\delta \approx 0$, we can set the systematic slope $\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial(\partial\mathcal{B}/\partial z) = 0$. In particular, we apply a deliberate $\frac{\partial\mathcal{B}}{\partial z} \sim 1\,\mathrm{mG/cm}$ and adjust δ until the systematic contribution $d\omega_{\partial\mathcal{B}/\partial z}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = \partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial(\partial\mathcal{B}/\partial z) \times \partial\mathcal{B}/\partial z$ is consistent with zero. Then, the systematic contribution is proportional to small quantities $\delta/\sigma_{\mathrm{ST}}$ and the ambient magnetic field gradient, $(\partial\mathcal{B}/\partial z)_{\mathrm{ambient}} \approx -4\,\mu\mathrm{G/cm}$. Furthermore, we measure the ambient non-reversing magnetic field gradient using *in situ* magnetometers in the interaction region. By applying compensating field gradients during the EDM data set, we obtain a value for the ambient magnetic field gradient after compensation, $(\partial\mathcal{B}/\partial z)_{\mathrm{comp}} = -0.5(60)\,\mu\mathrm{G/cm}$. The uncertainty in the remaining gradient is much larger than the central value due to possible systematic offsets in the measurement; see Sec. 5.2.5. Nevertheless, with the magnetic field gradient cancelled to the level of our measurement precision, and δ set such that the systematic slope is consistent with zero, the remaining systematic error contribution is proportional to two small quantities.

Before describing the calculation of the systematic error mean and uncertainty, it is important to point out that using δ to minimize the systematic slope is equivalent to using the $\partial\mathcal{E}^{\mathrm{nr}}/\partial z$ contribution to *cancel* the $\mathcal{E}^{\mathrm{nr}}$ contribution. We do not minimize each contribution independently. By adjusting the STIRAP detuning, we can zero out the systematic. Note that we expect the STIRAP detuning to only affect the $\frac{\partial\mathcal{E}^{nr}}{\partial z}$ systematic, so when the total $\frac{\partial\mathcal{B}}{\partial z}$ systematic is set to 0, the two sources must have equal and opposite contributions.

To measure the remaining contribution of the $\partial\mathcal{B}/\partial z$ systematics, we measure $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ under conditions of a deliberately large value of $\partial\mathcal{B}/\partial z$ every few hours in the EDM data set. We compute a linear regression against data under ideal conditions to measure the

systematic slope $\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial(\partial\mathcal{B}/\partial z) = -5.9(31)\,(\mu\text{rad/sec})/(\mu\text{G/cm})$ during the EDM data set. As mentioned previously, the ambient magnetic field gradient is measured using magnetometers in the interaction region; the details of this measurement are described in Sec. 5.2.5. We find $(\partial\mathcal{B}/\partial z)_{\text{comp}} = -0.5(60)\,\mu\text{G/cm}$ after compensation with applied magnetic field gradients. This would give a systematic error contribution $d\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}_{\partial\mathcal{B}/\partial z} = 3(35)\,\mu\text{rad/sec}$. However, we calculate the EDM systematic shift and uncertainty due to both $\partial\mathcal{B}/\partial z$ and $\partial\mathcal{B}/\partial y$ together. The systematic errors that depend on $\partial\mathcal{B}/\partial y$ are described below.

$\partial B/\partial y$ systematics

Before suppression of the $\partial\mathcal{B}/\partial z$ systematic, we saw a comparable systematic slope

$$\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial(\partial\mathcal{B}/\partial y) \sim 100\,(\text{mrad/sec})/(\text{mG/cm}). \tag{4.33}$$

We can understand this dependence using exactly the same models developed to understand the $\partial\mathcal{B}/\partial z$ systematics. In particular, because the readout lasers travel along $\hat{z}$ and the STIRAP lasers travel along $\hat{y}$, the same mechanisms that generate a systematic slope $\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial(\partial\mathcal{B}/\partial z)$ will generate the slope $\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial(\partial\mathcal{B}/\partial y)$ with the interchange of the roles of the STIRAP and readout beams. Informally, we simply substitute $y \leftrightarrow z$ and STIRAP$\leftrightarrow$readout.

Consider first the $\mathcal{E}^{\text{nr}}$ contribution to the $\partial\mathcal{B}/\partial z$ systematic slope, in which the position-velocity correlation in the molecular beam creates a position-detuning correlation due to the Doppler shift. This shifts the average position of resonant molecules and, due to the finite saturation of the readout transition, shifts the center of mass of probed molecules. The $\tilde{\mathcal{N}}\tilde{\mathcal{E}}$-correlated shift in the center of mass, $z^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}_{\text{CM}}$, is proportional to the $\tilde{\mathcal{N}}\tilde{\mathcal{E}}$-correlated detuning (unrelated to the Doppler shift) arising from $\mathcal{E}^{\text{nr}}$. This center-of-mass shift combines with the precession frequency gradient associated with the magnetic field gradient in the apparatus, $d\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = z^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}_{\text{CM}} \times \frac{\partial\omega}{\partial z}$.

The analogous effect occurs in the STIRAP beams provided $\partial\mathcal{B}/\partial y \neq 0$. Explicitly, the correlation between y and v_y creates a position-detuning correlation due to the Doppler shift in the STIRAP beams, which travel along $\hat{y}$. This shifts the average y-position of resonant

molecules and, due to the finite saturation of the STIRAP transition, shifts the center of mass y_{CM} of molecules transferred to the H state manifold. The $\tilde{\mathcal{N}}\tilde{\mathcal{E}}$-correlated shift in the center of mass, $y_{\mathrm{CM}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$, is proportional to the $\tilde{\mathcal{N}}\tilde{\mathcal{E}}$-correlated two-photon detuning (unrelated to the Doppler shift) arising from $\mathcal{E}^{\mathrm{nr}}$. This center-of-mass shift combines with the precession frequency gradient associated with the magnetic field gradient in the apparatus, $d\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = y_{\mathrm{CM}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} \times \frac{\partial\omega}{\partial y}$.

Now consider the $\frac{\partial\mathcal{E}^{\mathrm{nr}}}{\partial z}$ contribution to the $\partial\mathcal{B}/\partial z$ systematic slope, in which the $\tilde{\mathcal{N}}\tilde{\mathcal{E}}$-correlated detuning gradient along z, together with an overall offset in the two-photon detuning δ and incomplete saturation of the STIRAP transition, leads to a shift in the $\tilde{\mathcal{N}}\tilde{\mathcal{E}}$-correlated center of mass $z_{\mathrm{CM}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$. This center-of-mass shift combines with the precession frequency gradient associated with the magnetic field gradient in the apparatus, $d\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = z_{\mathrm{CM}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} \times \frac{\partial\omega}{\partial z}$.

The analogous effect occurs in the readout beams provided $\partial\mathcal{B}/\partial y \neq 0$. Any $\frac{\partial\mathcal{E}^{\mathrm{nr}}}{\partial y}$ will cause an $\tilde{\mathcal{N}}\tilde{\mathcal{E}}$-correlated detuning gradient along y. Together with an overall offset in the readout detuning Δ and incomplete saturation of the readout transition, this leads to a shift in the $\tilde{\mathcal{N}}\tilde{\mathcal{E}}$-correlated center of mass $y_{\mathrm{CM}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$. This center-of-mass shift combines with the precession frequency gradient associated with the magnetic field gradient in the apparatus, $d\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = y_{\mathrm{CM}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} \times \frac{\partial\omega}{\partial y}$.

Quantitatively, we have measured that $\frac{\partial\mathcal{E}^{nr}}{\partial y}$ is $\sim 10\%$ as large as $\frac{\partial\mathcal{E}^{nr}}{\partial z}$ (see Sec. 5.1.1), and moreover the probe beam is better-saturated than the STIRAP beam. Therefore, $\frac{\partial\mathcal{E}^{nr}}{\partial y}$ effect is expected to be negligible at the ACME II sensitivity.

On the other hand, the $\partial\mathcal{B}/\partial y \times \mathcal{E}^{\mathrm{nr}}$ systematic is *not* necessarily small compared to the $\partial\mathcal{B}/\partial z \times \mathcal{E}^{\mathrm{nr}}$ systematic, as we saw prior to suppressing the total $\partial\mathcal{B}/\partial z$ systematic error contribution. However, the contribution of $\partial\mathcal{B}/\partial y \times \mathcal{E}^{\mathrm{nr}}$ depends implicitly in δ: a non-reversing electric field, $\mathcal{E}^{\mathrm{nr}}$, contributes a correlated two-photon detuning that "moves" the preparation efficiency along the local STIRAP two-photon lineshape (which depends on δ). That said, it is difficult to write down a simple analytic dependence of $\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial(\partial\mathcal{B}/\partial y)$ on the two-photon detuning, δ. We instead rely on the empirical change of the systematic slope when we adjust δ. We find that when we set δ to null the total $\partial\mathcal{B}/\partial z$ systematic, the total $\partial\mathcal{B}/\partial y$ systematic (presumed to be entirely attributable to the $\partial\mathcal{B}/\partial y \times \mathcal{E}^{\mathrm{nr}}$ effect

occuring in the STIRAP lasers) is also consistent with zero, which we verified both before and after the EDM data set.

Recall that prior to suppression of either magnetic field gradient systematic by tuning $\delta \approx 0$, we observed comparable magnitudes of $\partial \omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial(\partial\mathcal{B}/\partial z)$ and $\partial \omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial(\partial\mathcal{B}/\partial y)$; see Fig. 4.5. As just mentioned, after setting $\delta \approx 0$ by minimizing the $\partial \omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial(\partial\mathcal{B}/\partial z)$ systematic slope, the $\partial \omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial(\partial\mathcal{B}/\partial y)$ was verified to also be consistent with zero. We compute a systematic error shift due to $\partial \omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial(\partial\mathcal{B}/\partial y)$ by using the observed ratio of systematic slopes, $R = [\partial \omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial(\partial\mathcal{B}/\partial y)]/[\partial \omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial(\partial\mathcal{B}/\partial z)] \approx 1.1$ taken before suppression of the systematics, and the value of $\partial \omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial(\partial\mathcal{B}/\partial z)$ measured throughout the EDM data set. The value of $(\partial\mathcal{B}/\partial y)_{\text{ambient}} = 1(6)\,\mu\text{G/cm}$ is measured using magnetometers in the interaction region as described in Sec. 5.2.5. The magnetic field gradient along y was not compensated using the magnetic field coils. Therefore, the total systematic shift due to magnetic field gradients is

$$d\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}_{\partial\mathcal{B}} = \frac{\partial \omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{\partial(\partial\mathcal{B}/\partial z)}((\partial\mathcal{B}/\partial z)_{\text{comp}} + R(\partial\mathcal{B}/\partial y)_{\text{ambient}}). \tag{4.34}$$

The uncertainties in $(\partial\mathcal{B}/\partial z)_{\text{comp}}$ and $(\partial\mathcal{B}/\partial y)_{\text{ambient}}$ are independent, so the effective parameter offset is $(\partial\mathcal{B}/\partial z)_{\text{comp}} + R(\partial\mathcal{B}/\partial y)_{\text{ambient}} \approx -1(9)\,\mu\text{G/cm}$. Propagating the uncertainty in the systematic slope and effective parameter offset, we find $d\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}_{\partial\mathcal{B}} = 7.7(689)\,\mu\text{rad/sec}$, consistent with zero and significantly below the statistical sensitivity of the EDM dataset. For a summary of parameters related to the $\partial\mathcal{B}/\partial z$ and $\partial\mathcal{B}/\partial y$ systematics, see Table 4.7.

4.11 Total uncertainty budget

Based on the well-understood mechanisms for the observed systematic slopes, we decided to include in the uncertainty budget contributions from specific systematic check parameters for which the systematic slope was expected and observed to be consistent with zero. See. Table 4.8 for all systematic error shifts and uncertainties included in the EDM error budget.

Due to the presence of magnetic field gradient systematics $d\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}_{\partial\mathcal{B}/\partial z}$ and $d\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}_{\partial\mathcal{B}/\partial y}$, we in-

Imperfection	Corresponding parameter	Typical value	Uncertainty
$\partial\mathcal{B}/\partial z,\ \partial\mathcal{B}/\partial y$	$\mathcal{B}$-field gradient	$-0.5,\ +1.0\ \mu\text{G/cm}$	$6.0\ \mu\text{G/cm}$
$\mathcal{E}^{\text{nr}}$	$\Delta^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$	$-2.6\ \text{mV/cm}$	$1.6\ \text{mV/cm}$
$\partial\mathcal{E}^{\text{nr}}/\partial z$	$\partial\Delta^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial z$	$-7\ \text{mV/cm}^2$	$3\ \text{mV/cm}^2$
δ	$z_{\text{CM}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$	$\sim 30\ \text{kHz}$	—
Quantity	Value	Uncertainty	Units
$\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial(\partial\mathcal{B}/\partial z)$	-5.9	3.1	$\frac{\mu\text{rad/sec}}{\mu\text{G/cm}}$
$d\omega_{\partial\mathcal{B}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$	7.7	68.9	$\mu\text{rad/sec}$

Table 4.7: Summary of $\partial\mathcal{B}/\partial z$ and $\partial\mathcal{B}/\partial y$ systematic parameters and shifts. The imperfections, $\partial\mathcal{B}/\partial z$ and $\partial\mathcal{B}/\partial y$, are ambient (non-reversing) magnetic field gradients. Together with other imperfections, including a non-reversing electric field or its gradient, $\mathcal{E}^{\text{nr}}$ or $\partial\mathcal{E}^{\text{nr}}/\partial z$, and a non-zero two-photon STIRAP detuning, δ, these magnetic field gradients produce systematic slopes, $\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial(\partial\mathcal{B}/\partial z)$ and $\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial(\partial\mathcal{B}/\partial y)$. We compute the ratio of systematic slopes, R, from the linear fit of $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ vs. $\partial\mathcal{B}/\partial z$ and $\partial\mathcal{B}/\partial y$ prior to suppression of both systematic slopes. The value of $\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial(\partial\mathcal{B}/\partial z)$ throughout the EDM data set is computed by comparing $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ in the presence of a deliberately large $\partial\mathcal{B}/\partial z \sim 1\ \text{mG/cm}$, and $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ under ideal conditions. We assume that the systematic slope of $\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial(\partial\mathcal{B}/\partial y)$ throughout the EDM data set is greater by a factor of $R \approx 1.1$, consistent with observations prior to suppression of the systematic slopes. Suppression of the systematic slopes was achieved by tuning the two-photon detuning, δ, to minimize $\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial(\partial\mathcal{B}/\partial z)$. The systematic shift, $d\omega_{\partial\mathcal{B}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = \frac{\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{\partial(\partial\mathcal{B}/\partial z)}((\partial\mathcal{B}/\partial z)_{\text{comp}} + R(\partial\mathcal{B}/\partial y)_{\text{ambient}})$, is computed using the measured slope (consistent with zero) and the values of the gradient along z, $(\partial\mathcal{B}/\partial z)_{\text{comp}}$, after compensation with small deliberately-applied magnetic field gradients, and the ambient gradient along y, $(\partial\mathcal{B}/\partial y)_{\text{ambient}}$. See Sec. 5.2.5 for the determination of these values.

Parameter	Shift	Uncertainty	Section
$\mathcal{E}^{\text{nr}}$	-56	140	4.5.3
$\omega_{\text{ST}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ (via $d\theta_{\lambda/2}$)	0	1	4.6.3
$P_{\text{ref}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$	-	109	4.7
$\mathcal{B}^{\tilde{\mathcal{E}}}$	1	1	4.8
$|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\mathcal{B}}$ and $|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$	77	125	4.9
$\partial\mathcal{B}/\partial z$ and $\partial\mathcal{B}/\partial y$	7	59	4.10
Other $\mathcal{B}$-field gradients total (4)	-	134	4.11
Non-reversing $\mathcal{B}$-field $\mathcal{B}^{\text{nr}}$	-	106	4.11
Transverse $\mathcal{B}$-fields ($\mathcal{B}_x$, $\mathcal{B}_y$)	-	92	4.11
Refinement/readout laser detunings	-	76	4.11
$\tilde{\mathcal{N}}$-correlated laser detuning, $\Delta^{\tilde{\mathcal{N}}}$	-	48	4.11
Total systematic	29	310	
Statistical		373	
Total uncertainty		486	

Table 4.8: Uncertainty budget for the EDM measurement. Parameters that were observed to cause a shift are ordered by introduction in the text; parameters that were included for conservative, due to a nominal similarity to parameters involved in understood systematic errors, are ordered by contribution to the uncertainty. The statistical error bar is computed only from EDM data taken under ideal conditions. All uncertainties are added in quadrature. Hyphens indicate that no shift was subtracted.

cluded limits from possible contributions $|d\omega_{\partial\mathcal{B}_x/\partial x}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}| < 55\,\mu\text{rad/sec}$, $|d\omega_{\partial\mathcal{B}_y/\partial y}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}| < 51\,\mu\text{rad/sec}$, $|d\omega_{\partial\mathcal{B}_y/\partial x}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}| < 66\,\mu\text{rad/sec}$, and $|d\omega_{\partial\mathcal{B}_z/\partial x}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}| < 86\,\mu\text{rad/sec}$. In addition, we included a limit on $|d\omega_{\mathcal{B}^{\text{nr}}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}| < 106\,\mu\text{rad/sec}$, as well as contributions due to transverse fields, $|d\omega_{\mathcal{B}_x}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}| < 73\,\mu\text{rad/sec}$ and $|d\omega_{\mathcal{B}_y}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}| < 50\,\mu\text{rad/sec}$. See Table 4.9 for details.[8]

Due to the role of the two-photon STIRAP detuning δ and effect of non-reversing electric fields $\mathcal{E}^{\text{nr}}$ via correlated detunings in the readout and probe beams, $\Delta^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$, we included limits on the contributions to the EDM due to overall detunings in the refinement laser $|d\omega_{\Delta_{\text{ref}}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}| < 63\,\mu\text{rad/sec}$ and common detunings in both the refinement and probe lasers, $|d\omega_{\Delta}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}| < 36\,\mu\text{rad/sec}$. Furthermore, we included a limit on the contribution from $\tilde{\mathcal{N}}$-correlated detunings, $|d\omega_{\Delta^{\tilde{\mathcal{N}}}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}| < 48\,\mu\text{rad/sec}$. The limits on all systematic slopes ($\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial\Delta_{\text{ref}}$, $\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial\Delta$, and $\partial\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/\partial\Delta^{\tilde{\mathcal{N}}}$) were determined directly by deliberately applying large values of Δ_{ref}, Δ, and $\Delta^{\tilde{\mathcal{N}}}$. The typical parameter values in each case were

8. Here, I quote the average limit obtained by all four sets of analysis codes. The value in Table 4.8 is the average over analysis routines of the quadrature sum of each contribution. Thus the quadrature sum of individual components quoted here differs from the value in the table by a few μrad/sec. A similar discrepancy occurs for the limit on transverse $\mathcal{B}$ fields and laser detunings.

Gradient	Slope [(mrad/sec)/(mG/cm)]	Typical [μG/cm]	Limit [μrad/sec]
$\partial\mathcal{B}_x/\partial x$	$-7.2(52)$	6	55
$\partial\mathcal{B}_y/\partial y$	$0.1(65)$	5	51
$\partial\mathcal{B}_x\partial y$	$-5.0(42)$	10	66
$\partial\mathcal{B}_z/\partial x$	$2.5(168)$	5	86

Offset	Slope [(mrad/sec)/mG]	Typical [μG]	Limit [μrad/sec]
$\mathcal{B}_x$	$5.8(234) \times 10^{-3}$	300	73
$\mathcal{B}_y$	$-59(316) \times 10^{-3}$	150	50
$\mathcal{B}_z^{\mathrm{nr}}$	$-201(1136) \times 10^{-3}$	80	106

Table 4.9: Non-reversing magnetic field gradients and offsets for which no statistically significant systematic slope was observed. Systematic slopes given in units of (mrad/sec)/(mG/cm) and (mrad/sec)/mG for gradients and offsets, respectively. Typical values of gradients and offsets are given in units of μG/cm and μG, respectively. All systematic error limits in μrad/sec. Each systematic slope was computed by deliberately applying a large magnetic field gradient or offset (typically $\sim 1.5\,\mathrm{mG/cm}$ and $\sim 1 - 10$ mG, respectively) and observing the resulting shift in $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$. Typical parameter values were obtained by placing conservative limits on the ambient magnetic field gradients or offsets in the interaction region, as measured twice per day throughout the EDM data set using the fluxgate magnetometers in the interaction region, and supplemented with offline measurements after all EDM data was acquired (see Sec. 5.2.5). The largest field component is $\mathcal{B}_x$, which we believe arises from ambient magnetization in the mu-metal endcaps (positioned along $\pm\hat{x}$ relative to the electric field plates), which is not fully removed by the degaussing procedure. The anomalously small systematic slope in $\partial\mathcal{B}_y/\partial y$ is due to a coincidental near-cancellation in the $\sim 1\sigma$ magnitude slopes, which are all statistically insignificant, obtained from different analysis procedures; for example, my own analysis gives a slope of $3.7(57)\,\mathrm{mrad/sec/(mG/cm)}$. In order to isolate the systematic error contribution due to each designated magnetic field gradient, we compute and remove the expected contribution due to $\partial\mathcal{B}_z/\partial z$, which necessarily arises due to Maxwell's equations (e.g., the nominal $\partial\mathcal{B}_y/\partial y$ applied field gradient also generates a comparable $\partial\mathcal{B}_z/\partial z$, consistent with $\nabla \cdot \vec{\mathcal{B}} = 0$).

Parameter	Slope [(mrad/sec)/MHz]	Offset [kHz]	Limit [μrad/sec]
Δ_{ref}	$-1.1(26)$	20	63
Δ	$-0.1(16)$	20	36
$\Delta^{\mathcal{N}}$	$-0.2(23)$	20	48

Table 4.10: Laser detuning systematic error bounds. All systematic slopes were determined by deliberately applying large detunings in the corresponding lasers and parity components (typically 2 MHz). The typical parameter imperfections, 20 kHz, were determined from the repeatability of setting the resonance condition for the lasers in each state (e.g., $\tilde{\mathcal{N}} = \pm 1$). The overall and $\tilde{\mathcal{N}}$-correlated detunings, Δ and $\Delta^{\tilde{\mathcal{N}}}$, are applied simultaneously to both the refinement and probe lasers. Any systematic slope due to the probe laser detuning, Δ_{probe}, would simply be a linear combination of the systematic slopes due to Δ_{ref} and Δ. Therefore, we do not compute a separate limit for a systematic effect arising from Δ_{probe}.

determined to be 20 kHz, limited by the resolution of the Doppler scans used to set the 703 nm laser on resonance. See Table 4.10 for details.

The central estimate for the total systematic shift is 29 μrad/sec, due to (1) non-reversing electric fields, $\mathcal{E}^{\mathrm{nr}}$, (2) residual $\tilde{\mathcal{N}}\tilde{\mathcal{E}}$-correlated frequency originating in the STIRAP beams, $\omega_{\mathrm{ST}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$, (3) leakage of a possible $\tilde{\mathcal{E}}$-odd Zeeman frequency due to the non-zero g-factor difference between $\tilde{\mathcal{N}}$ states, $\mathcal{B}^{\tilde{\mathcal{E}}}$, (4) correlations of the EDM channel with components of the contrast magnitude, $|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}}$ and $|\mathcal{C}|^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$, and (5) magnetic field gradients, $\partial\mathcal{B}/\partial z$ and $\partial\mathcal{B}/\partial y$. The total systematic uncertainty (from contributions where only a limit was set as well as those where a shift was subtracted) is 310 μrad/sec.

4.12 Asymmetry effects

We observed statistically significant offsets in the frequency components $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{P}}\tilde{\mathcal{R}}} = 5.3(5)$ mrad/sec (for $\mathcal{B} < 26$ mG), and $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}\tilde{\mathcal{P}}\tilde{\mathcal{R}}} = 7.1(1)$ mrad/sec (for $\mathcal{B} = 26$ mG). Here, we will examine the source of non-zero values of $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{P}}\tilde{\mathcal{R}}}$ and $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}\tilde{\mathcal{P}}\tilde{\mathcal{R}}}$. Note that we compute the EDM from $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ only, so that an offset in any other channel cannot be a true systematic error (unless it "leaks" through to produce an offset in $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$).

In ACME I, offsets in $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{P}}\tilde{\mathcal{R}}}$ and $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}\tilde{\mathcal{P}}\tilde{\mathcal{R}}}$ were understood to arise from the following mechanisms. Suppose there is some difference between the X and Y readout laser beams. For example, suppose there is a pointing misalignment θ_{XY} between the X and Y laser beams, along the x-axis. Then, due to the large molecular velocity $v_x \approx 200$ m/s, there will be a differential Doppler shift $\Delta_{XY} = \frac{\theta_{XY}v_x}{\lambda}$, where $\lambda = 703$ nm is the readout transition wavelength. Since the molecular fluorescence F depends on the laser detuning due to incomplete saturation, this will produce an $\tilde{\mathcal{N}}\tilde{\mathcal{E}}$-correlated asymmetry[129, Sec. 5.3]

$$\mathcal{A}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = \frac{1}{F}\frac{\partial^2 F}{\partial\Delta^2}\Delta_{XY}\Delta^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}, \tag{4.35}$$

where $\Delta^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} = D\mathcal{E}^{\mathrm{nr}} \sim 2\pi \times 5$ kHz arises from a non-reversing component of the electric field. Analogous asymmetry effects can occur if the X and Y lasers have different powers.

These asymmetry effects are distinguished from physical molecular phases by the fact
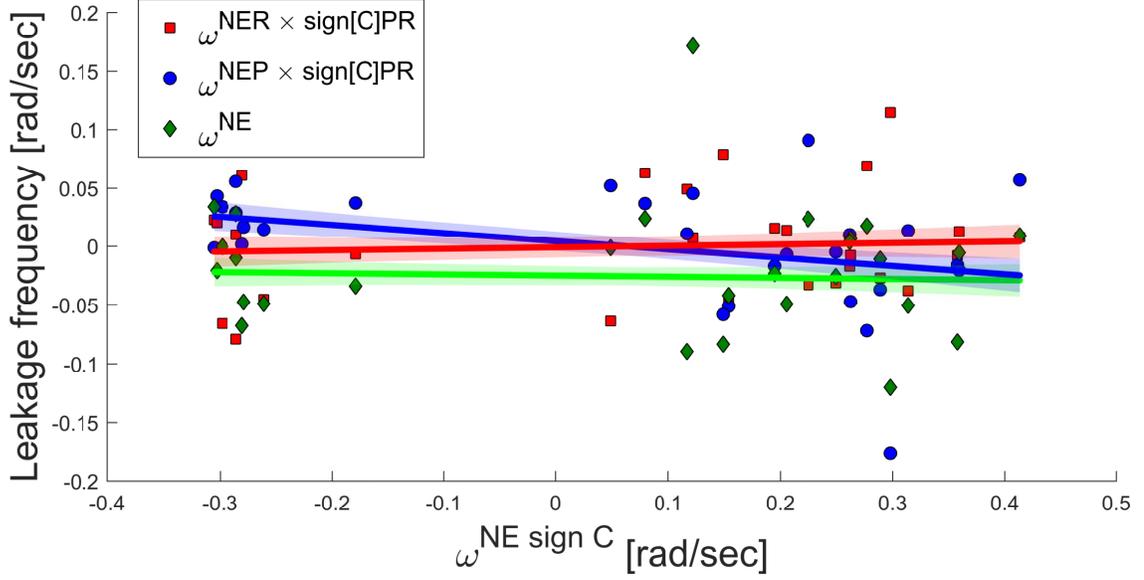
Figure 4.10: Data used to compute "leakage" of a large asymmetry effect, $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\text{sign}(\mathcal{C})}$, into $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$. Taken from Runs 125 and 138, where a large power difference between the X and Y laser beams was deliberately imposed to generate asymmetry effects. Due to the $\tilde{\mathcal{P}}$ switch, $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{R}}\times\text{sign}(\mathcal{C})\tilde{\mathcal{P}}\tilde{\mathcal{R}}}$ is independent of the asymmetry channel $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\times\text{sign}(\mathcal{C})}$. (Here, the factor of $\text{sign}(\mathcal{C})\tilde{\mathcal{P}}\tilde{\mathcal{R}}$ is equal to $\tilde{\mathcal{B}}$ for $\mathcal{B} = 26$ mG and does not depend on any switches otherwise.) Likewise, due to the $\tilde{\mathcal{R}}$ switch, $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{P}}\times\text{sign}(\mathcal{C})\tilde{\mathcal{P}}\tilde{\mathcal{R}}}$ is independent of the asymmetry channel $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\times\text{sign}(\mathcal{C})}$. The blinded EDM channel, $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$, is shown for reference and is also independent of asymmetry effects. Error bars are omitted for clarity on all points. By computing linear regressions of suppressed channels vs. the asymmetry effect channel, where $|\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\times\text{sign}(\mathcal{C})}| \approx 250$ mrad/sec on average (see the figure x-axis), we can limit the suppression factor from the $\tilde{\mathcal{P}}$ switch to at least 30, and the suppression fator from the $\tilde{\mathcal{R}}$ switch to at least 13. Thus the overall suppression of asymmetry effects from both switches combined is at least a factor of ~ 400.

that, at leading order, they do not affect the measured contrast. In particular,

$$\mathcal{A}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} \approx 2|\mathcal{C}|\phi^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}\text{sign}(\mathcal{C}). \tag{4.36}$$

When $\phi^{\tilde{\mathcal{B}}} \approx \frac{\pi}{4}$ (e.g., when $\mathcal{B} = 26$ mG), $\text{sign}(\mathcal{C}) = \tilde{\mathcal{B}}\tilde{\mathcal{P}}\tilde{\mathcal{R}}$. On the other hand, when $\phi^{\tilde{\mathcal{B}}} \approx 0$ (e.g., when $\mathcal{B} = 0.7$, 1.3, or 2.6 mG), $\text{sign}(\mathcal{C}) = \tilde{\mathcal{P}}\tilde{\mathcal{R}}$. Therefore, an offset in $\mathcal{A}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ appears in the frequency channel $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}\tilde{\mathcal{P}}\tilde{\mathcal{R}}}$ or $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{P}}\tilde{\mathcal{R}}}$, depending on the magnitude of the magnetic field.

In ACME I, the asymmetry channels $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}\tilde{\mathcal{P}}\tilde{\mathcal{R}}}$ and $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{P}}\tilde{\mathcal{R}}}$ were consistent with zero throughout the EDM data set. In ACME II, due to the increased sensitivity, comparable asymmetry effects generate measurable offsets in

$$\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\text{sign}(\mathcal{C})} = \{6.5\sigma_{0.7}, \ 6.8\sigma_{1.3}, \ 2.8\sigma_{2.6}, \ 5.5\sigma_{26}\} \tag{4.37}$$

for $\mathcal{B} = \{0.7, \ 1.3, \ 2.6, \ 26\}$ mG, respectively, where σ represents the statistical uncertainty in the frequency channel for each magnetic field magnitude, respectively.

We can make an order-of-magnitude estimate of the pointing misalignment asymmetry effect in ACME II. For a pointing misalignment of $\sim 200\,\mu$rad (limited by the ability to overlap the X and Y lasers on a beam profiler before and after the interaction region; see [147, Sec. 5.4.5]), the differential Doppler shift is $\Delta_{XY} \sim 2\pi \times 10$ kHz. For a gaussian readout transition lineshape with $1\sigma \sim 2\pi \times 3$ MHz, $\frac{\partial^2 F}{\partial \Delta^2} \approx -\frac{1}{\sigma^2}$, we find $\mathcal{A}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} \sim -3 \times 10^{-6}$, which (for precession time $\tau = 1$ ms) leads to $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\text{sign}\mathcal{C}} \sim 1$ mrad/sec. This is on the same order as (in fact, a factor of a few smaller than) the observed asymmetry channels, and larger than the statistical uncertainty of the EDM measurement.

The offsets in the "asymmetry" channels could leak through to the EDM channel only if the suppression from the $\tilde{\mathcal{P}}$ and $\tilde{\mathcal{R}}$ switches is far weaker than expected. To verify that the $\tilde{\mathcal{P}}$ and $\tilde{\mathcal{R}}$ switches suppress asymmetry effects as expected, we use data in which a deliberate power imbalance between the X and Y beams induces a large asymmetry effect ($\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{P}}\tilde{\mathcal{R}}} \sim 300$ mrad/sec). See Fig. 4.10. Casual inspection suggests that the asymmetry effect does not "leak" through to $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{P}}(\tilde{\mathcal{B}})}$ or $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{R}}(\tilde{\mathcal{B}})}$, let alone the doubly-

protected $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ channel. By computing linear regressions of $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{P}}(\tilde{\mathcal{B}})}$ and $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{R}}(\tilde{\mathcal{B}})}$ vs. the asymmetry channel $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{P}}\tilde{\mathcal{R}}(\tilde{\mathcal{B}})}$, we can set bounds on the suppression factor $A_{\tilde{\mathcal{P}}(\tilde{\mathcal{R}})}$ of each switch: $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{P}}\tilde{\mathcal{R}}\tilde{s}(\tilde{\mathcal{B}})} = A_s \times \omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{P}}\tilde{\mathcal{R}}(\tilde{\mathcal{B}})}$, where $s = \tilde{\mathcal{P}}$ or $\tilde{\mathcal{R}}$. In particular, we find bounds $A_P > 30$ and $A_R > 13$. Thus the combined suppression factor is $A_P \times A_R > 390$ and any possible leakage of the asymmetry effects into the EDM channel will be far below the statistical sensitivity. Note that these bounds are limited by the sensitivity of this measurement; we expect both A_P and A_R to significantly exceed these bounds.

Because we have a good model for asymmetry effects and do not expect them to shift the EDM, and we have confirmed that any possible leakage from "asymmetry effect" channels $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{P}}\tilde{\mathcal{R}}}$ and $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{B}}\tilde{\mathcal{P}}\tilde{\mathcal{R}}}$ to the EDM channel $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ must be far below the statistical sensitivity of the EDM measurement, we have not included a limit from this effect in the systematic error budget.

4.13 Statistical uncertainty

As described in Sec. 3.5, the values of $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ in the EDM data set are not normally distributed, and the noise beyond the photon shot noise limit depends on the applied magnetic field $\mathcal{B}$. We will consider the sources of the statistical uncertainty, beyond the shot-noise level, in subsequent sections. However, the nature of these noise sources does not affect the procedure used to calculate the statistical uncertainty of the EDM channel $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ (nor, for that matter, any systematic effects). Therefore, I describe here the last piece of the uncertainty budget in Table 4.8: the statistical sensitivity.

Because the additional noise, beyond the level propagated from the scatter of individual asymmetry measurements within a group (which is consistent with the level expected from shot noise; see Sec. 3.1), depends on $\mathcal{B}$, we compute the mean and uncertainty of $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ separately for each magnetic field magnitude $\mathcal{B}$. As previously mentioned, the mean is determined by a robust M-estimator with Huber weights, and the 1σ uncertainty is computed from the 68.27% coverage region of the bootstrap distribution of the M-estimator mean. In particular, 68.27% of bootstrap samples have an M-estimator mean that lies within $\pm 1\sigma$ of the central (ordinary mean) value of the bootstrap distribution.

In this way, we obtain four values of $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}(\mathcal{B})$ and $\sigma_\mathcal{B}$, for $\mathcal{B} = 0.7$, 1.3, 2.6, and 26 mG (corresponding to 0.5, 1, 2, and 19.5 mA of applied current in the magnetic field coils, precisely). For $\mathcal{B} \leq 2.6$ mG, the resulting value of $\sigma_\mathcal{B}$ is ≈ 1.7 times as large as shot noise, while for $\mathcal{B} = 26$ mG, $\sigma_\mathcal{B}$ is ≈ 3 times as large as shot noise.

Once the sets of values for $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}(\mathcal{B})$ and $\sigma_\mathcal{B}$ have been obtained, we take the average over $\mathcal{B}$ using standard error propagation: $\langle \omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} \rangle = \sum_\mathcal{B} \sigma_\mathcal{B}^{-2} \omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}(\mathcal{B}) / \sum_\mathcal{B} \sigma_\mathcal{B}^{-2}$, and $\sigma_{\text{stat.}} = \left[\sum_\mathcal{B} \sigma_\mathcal{B}^{-2} \right]^{-1/2} = 373 \, \mu\text{rad/sec}$. The resulting statistical uncertainty is slightly larger than the total systematic uncertainty, $\sigma_{\text{syst.}} = 310 \, \mu\text{rad/sec}$.

4.14 Result

All data analysis was completed for two independent analysis routines (mine and Cristian Panda's) before removing the EDM blind. Two additional analysis routines, which were finalized after removing the EDM blind, were consistent with our initial analyses. After removing the systematic shifts (and the analysis blind) from $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$, we find the contribution due to T-violating physics,

$$
\begin{aligned}
\omega_T^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} &= \langle \omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} \rangle - d\omega_{\mathcal{E}^{\text{nr}}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} - d\omega_{\omega-\text{ST}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} - d\omega_{\Delta g}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} - d\omega_{|\mathcal{C}|}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} - d\omega_{\partial \mathcal{B}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}} \\
&= \qquad (-510 \pm 373_{\text{stat.}} \pm 310_{\text{syst.}}) \, \mu\text{rad/sec}.
\end{aligned}
\tag{4.38}
$$

Converting to EDM units, $d_e = -\frac{\hbar \omega_T^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}}{\mathcal{E}_{\text{eff}}}$ with $\mathcal{E}_{\text{eff}} = 78$ GV/cm[161, 162],

$$
d_e = (4.3 \pm 3.1_{\text{stat.}} \pm 2.6_{\text{syst.}}) \times 10^{-30} \, e \cdot \text{cm},
\tag{4.39}
$$

which we interpret as $d_e = (4.3 \pm 4.0) \times 10^{-30} \, e \cdot \text{cm}$ by combining the statistical and systematic uncertainties in quadrature: $\sigma_{d_e} = 4.0 \times 10^{-30} \, e \cdot \text{cm}$. This is improved over the ACME I uncertainty, previously the most stringent measurement of d_e, by a factor of 12[129].

Following the procedure used in ACME I, we interpret this result using the Feldman-Cousins construction[199]. This approach is one (though not the only) method that leads to a limit reported when the measured central value $\langle d_e \rangle$ is sufficiently small compared to
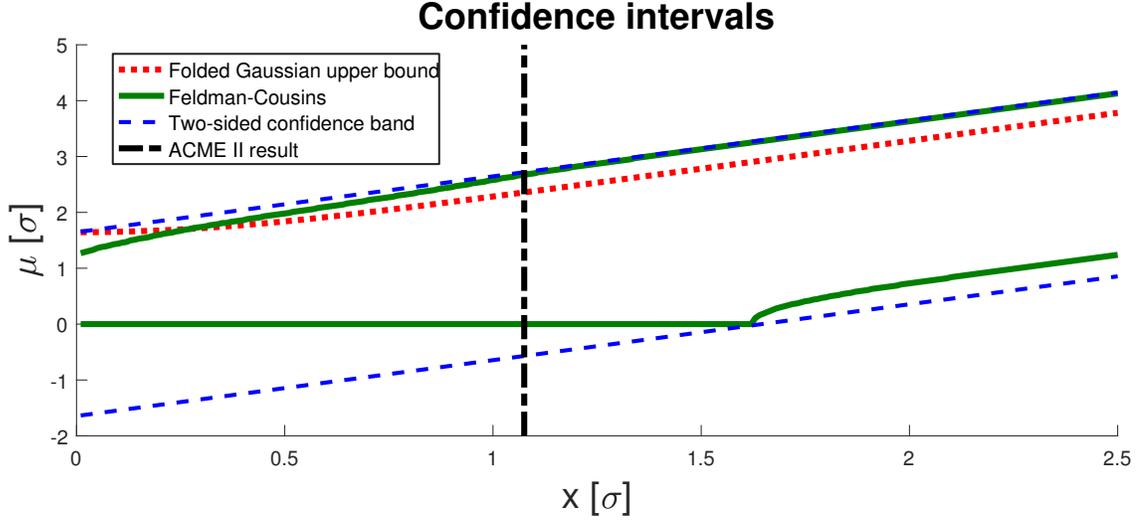
Figure 4.11: 90% confidence intervals in several constructions. The Feldman-Cousins construction (green solid line) gives only an upper bound until $x \approx 1.63\sigma$, after which a central value with two-sided errorbars is reported. The folded gaussian upper bound (dotted red line) has been used in previous experiments and reports a stronger upper bound for $x \gtrsim 0.27\sigma$, but a weaker upper bound otherwise. The two-sided confidence band for a signed central value is shown shown for reference (blue dashed line). The ACME II result, $|\langle d_e \rangle| = 1.1\sigma_{d_e}$, is shown as a vertical dot-dashed black line. The Feldman-Cousins prescription gives a 90% confidence lower limit of 0 so that only an upper bound is reported.

the uncertainty σ_{d_e}, and reports a band (central value with error bars) when the central value is sufficiently large compared to the uncertainty. A more naive approach would be to compute a limit up to some arbitrary threshold of $|\langle d_e \rangle|/\sigma_{d_e}$, and report the central value with symmetric error bars otherwise. However, such a naive approach would not necessarily have well-calibrated confidence intervals; due to the "flip-flopping" between reporting a bound and a central value, the nominal "90%" confidence interval would not contain the true value for 90% of repeated measurements.

We are primarily interested in the *magnitude* of the electron EDM, $x \equiv |\langle d_e \rangle|$, and thus consider the folded gaussian distribution

$$P(x|\mu) = \frac{1}{\sigma\sqrt{2\pi}} \left(\exp\left[-\frac{(x-\mu)^2}{2\sigma^2} \right] + \exp\left[-\frac{(x+\mu)^2}{2\sigma^2} \right] \right), \tag{4.40}$$

where $\mu \equiv |d_e|$ is the true value of the EDM magnitude, and we use the EDM uncertainty $\sigma = \sigma_{d_e}$. The goal is to construct a 90% confidence interval for μ, the true value of $|d_e|$, given the observed (but uncertain) value of $|d_e|$. The details of the construction have been well-

described in previous theses, [151, Appendix D] and [149, Sec. 3.6.2]. The 90% confidence interval is shown as a function of the observed value, in units of the uncertainty, in Fig. 4.11.

Using the Feldman-Cousins construction, we find

$$|d_e| < 1.1 \times 10^{-29} \, e \cdot \text{cm}. \tag{4.41}$$

This is a factor of 8.6 smaller than the ACME I result. The fact that the bound improved by a smaller factor than the uncertainty can be understood from Fig. 4.11: in ACME I, the central value was $\langle |d_e| \rangle_\text{I} = 0.5\sigma_{d_e,\text{I}}$. The resulting upper bound in this case is tighter relative to the sensitivity $\sigma_{d_e,\text{I}}$, compared to the case where the central value is 1.1σ from 0.

The relation $d_e = -\frac{\hbar\omega_T^{\tilde{N}\tilde{\mathcal{E}}}}{\mathcal{E}_\text{eff}}$, which we have used here, assumes that the scalar-pseudoscalar electron-nucleon coupling vanishes, $C_S = 0$. Assuming instead that $d_e = 0$ allows us to set an improved bound on $C_S = +\frac{\omega_T^{\tilde{N}\tilde{\mathcal{E}}}}{W_S}$, where $W_S = -2\pi \times 282$ kHz.[9] Then the Feldman-Cousins construction gives a 90% confidence bound,

$$|C_S| < 7.3 \times 10^{-10}. \tag{4.42}$$

4.15 Noise sources

The various sources of additional noise in the experiment cause two problems. First, superblock measurements of $\omega^{\tilde{N}\tilde{\mathcal{E}}}$ are distributed with a larger standard deviation than expected from shot noise (i.e., in practice, larger than expected from propagating the uncertainty assigned to each "group" within a trace), which reduces the sensitivity of ACME II to d_e. Second, the distribution of $\omega^{\tilde{N}\tilde{\mathcal{E}}}$ is not approximated well by a gaussian beyond ~ 2 standard deviations from the mean of the distribution. We have therefore extended significant effort to understand and, in some cases, suppress all sources of noise in the experiment

---

9. We calculate W_S from [161, 162], but comparison to the literature values requires care because we assume equal electron-proton and electron-neutron couplings. However, some papers assume the electron-neutron coupling vanishes. See [129, Appendix A.2] for useful discussion.

that exceed the shot-noise level.

4.15.1 Noise within pulses

The signature of noise depends on whether it occurs on timescales faster than or slower than the duration of a single molecular pulse in the detection region, which lasts ~ 1 ms. If the noise occurs on faster timescales than this, then group-to-group uncertainties will be underestimated. We can assess this by examining a "clean" channel, such as $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$, which is computed from data collected over much longer timescales than the noise (~ 1 minute). A measure of the excess noise among groups is $\chi^2_{\text{red}}[\text{groups}] \equiv \frac{1}{n-1}\sum_{j=1}^{n}(x_j - \bar{x})^2/\sigma_j^2$, where σ_j is the uncertainty of some quantity (e.g., $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$) assigned to group j, x_j is the estimated value of the same quantity, $\bar{x}$ is the weighted mean of the set of values, and n is the number of groups used. For $\{x_j\}$ normally distributed with correctly assigned uncertainties $\{\sigma_j\}$, $\chi^2_{\text{red}}[\text{groups}] = 1$ on average. When the uncertainties are underestimated, $\chi^2_{\text{red}} > 1$ on average. In the EDM data set, $\chi^2_{\text{red}}[\text{groups}] \approx 1.4$. This result does not depend significantly on $\mathcal{B}$, or on whether the reduced chi-square statistic is evaluated for groups in a block or groups in a superblock. Thus the error bar is inflated by $\sqrt{\chi^2_{\text{red}}[\text{blocks}]} \sim 20\%$ due to the additional noise, beyond the group-level assigned uncertainty (consistent with shot noise), that occurs on $\sim$ kHz or faster time scales. In earlier data sets without rotational cooling, we observed $\chi^2_{\text{red}}[\text{blocks}]$ consistent with 1, which is consistent with the small excess noise observed at the higher statistical sensitivity (i.e., provided the shot noise is sufficiently large, it dominates the uncertainty of the measurement). As we will see, the additional $\sim 20\%$ uncertainty is not the dominant noise source beyond the ACME II group-assigned noise level.

Although we did not eliminate all noise sources causing $\chi^2_{\text{red}}[\text{groups}] \neq 1$, we were able to eliminate several noise sources. The first was a simple data analysis artifact that actually causes $\chi^2_{\text{red}}[\text{blocks}] < 1$, meaning the uncertainty in group-level asymmetries was *over-*estimated. In particular, following the procedure of ACME I, we originally computed the asymmetry of a group from "individual" asymmetries by a sample mean, $\mathcal{A}_j = \sum_{i=1}^{n} \mathcal{A}_{ji}/n$, where $\mathcal{A}_j$ is the value of the j-th grouped asymmetry, and $\mathcal{A}_{ji}$ is the i-th "individual" asymmetry (computed from a single polarization switching cycle) within the j-th group. Here,

n is the number of polarization switching cycles included in each group. The uncertainty in the group, $\sigma_{\mathcal{A}_j}$, was computed from the sample standard deviation of values within the group.

However, assigning the uncertainty to the group, σ_{A_j}, from the standard deviation of values is appropriate only if the true asymmetry values are constant within the group. In particular, the variation in the asymmetry among polarization switching cycles within the same group must be small compared to the statistical sensitivity of individual asymmetry measurements. In ACME II, this assumption is false: velocity dispersion of the molecule beam leads to a locally linear slope of phase vs. time after ablation because later arrivals, on average, are travelling slower. Thus, for reasonable group sizes (e.g., $n = 20$), the standard deviation of measurements can be dominated by a true variation in the asymmetry rather than statistical shot noise. As a result, the uncertainty in the mean of each group would be over-estimated. In early data runs, we initially found $\chi^2_{\mathrm{red}}[\text{groups}] \approx 0.8$ due to this mechanism. Therefore, we assign uncertainties in ACME II from the residuals on a linear regression of individual asymmetry values within a group. Although this is hardly a "true" source of noise, it is important to realize that in future experiments, where the statistical uncertainty in the asymmetry of each polarization switching cycle is even smaller, a more complicated method of assigning group-level uncertainties may be required.

Another noise source (which is actually physical) that we observed to cause an artificially *small* value of $\chi^2_{\mathrm{red}}[\text{groups}]$ was electronic pickup in the DAQ. We found that this originated from the power supply leads to the magnetometers that monitor the magnetic field inside the vacuum chamber. These leads broadcast noise at 80 kHz, which is picked up by the cables from the PMTs to the DAQ (see Fig. 4.12). Recall that the polarization switching frequency is ordinarily 200 kHz. Therefore, the 80 kHz voltage noise will modulate the asymmetry on timescales of a few polarization switching cycles, but will not affect the average asymmetry on the timescale of a group. Because the asymmetry is modulated faster than the duration of a single group ($\sim 100\,\mu\mathrm{s}$), the standard deviation of residuals on the linear fit of individual asymmetry values is inflated relative to the shot noise. This is essentially just another example of assigning the uncertainty to the group by assuming a model of the individual asymmetry values (namely, linear slope) that does not match their
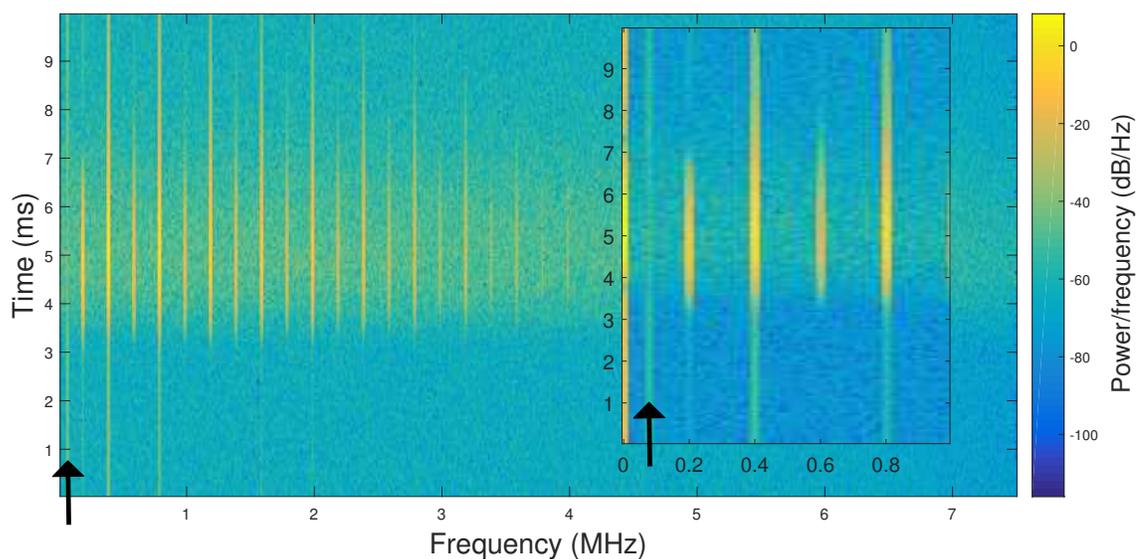
Figure 4.12: Spectrogram of a trace, showing the power spectral density as a function of frequency component (horizontal axis) and time within a trace (vertical axis). The 80 kHz noise is a faint band of yellow toward the left of the plot (emphasized with a black arrow). It is independent of the molecular pulse (speckled yellow in $\sim$ 4–6 ms band). Additional strong components are harmonics of 200 kHz polarization switching frequency. Note that the background signal (before and after the majority of the molecular pulse) has structure at the polarization switching frequency and its harmonics. An inset at right shows frequency components up to 1 MHz and more prominently displays the 80 kHz feature.

Figure 4.13: Each line represents the fluorescence summed over all traces within a block, as a function of the polarization cycle index in a $150\,\mu$s wide region near the falling edge of the molecular pulse. There is a consistent short-timescale spike in voltage corresponding to a dip in the fluorescence signal, which arises from electronic pickup due to the flash lamp discharge.
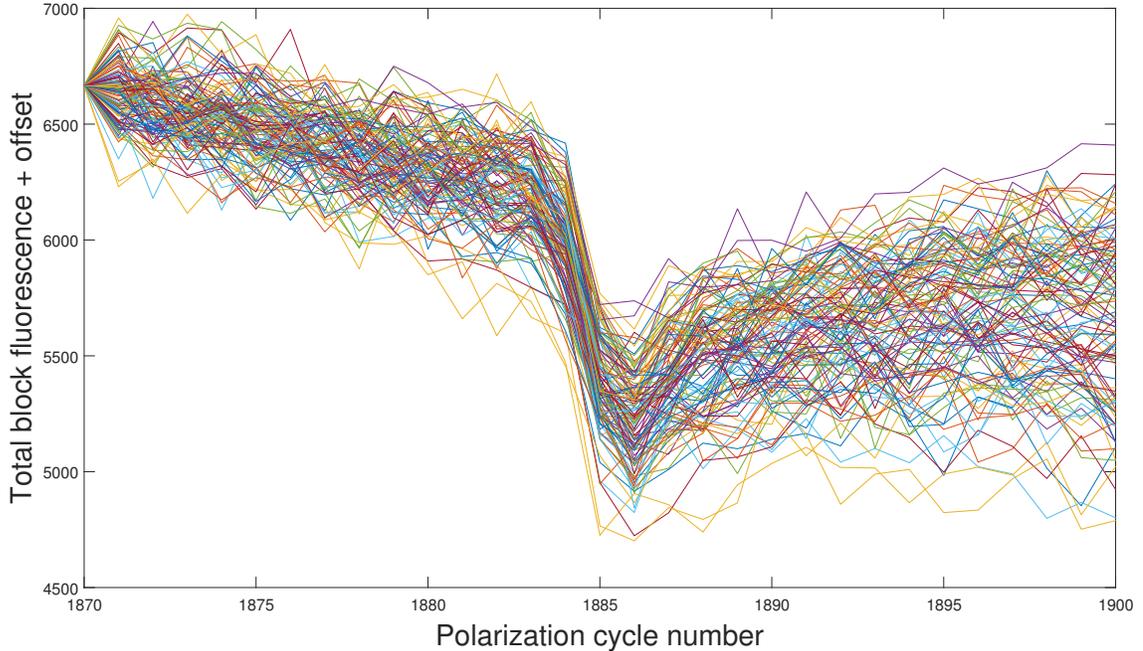
actual time structure (namely, linear slope with a modulation at 80 kHz). This mechanism also reduced the value of χ^2_{red}[groups] to ~ 0.8 in early runs, before we discovered the source of the 80 kHz broadcast noise. After discovering that the noise source was the magnetometer cables, we implemented a protocol of turning off the magnetometers whenever we acquire data, which eliminated the problem. This motivated the data analysis flag for anomalous peaks in the Fourier transform of a trace.

The last source of noise within a molecular pulse that we identified and eliminated is electronic noise broadcasted from the flash lamp discharge in the YAG ablation laser, which is picked up by the DAQ across the room. The flash lamps rapidly discharge at 200 Hz in order to maintain the designed thermal properties of the YAG crystal, even though we only Q-switch the laser at 50 Hz (which triggers the laser pulse). As a result, there were initially voltage spikes observable on the fluorescence signal, $\sim 10\,\mu$s in duration and equivalent to a photoelectron rate of $\sim 10^7$/sec (see Fig. 4.13). This was first identified from anomalous short-timescale noise in the measurement contrast as a function of time after ablation.

Improving the electronic shielding around the YAG laser via a Faraday cage reduced the electronic pickup to below an observable level.

4.15.2 Zeeman phase noise between pulses

To assess noise that occurs on time scales longer than ~ 1 ms, we evaluate the chi-square statistic from distinct blocks or superblocks, rather than distinct groups: $\chi^2_{\text{red}}[\text{SB}] = \frac{1}{n-1} \sum_{j=1}^{n} (x_j - \bar{x})^2 / \sigma_j^2$, where here n is the number of superblocks examined, x_j is the average value of x obtained for the j-th superblock, $\bar{x}$ is the weighted sample mean of x over all superblocks, and σ_j is the nominal uncertainty assigned to the j-th superblock. When not explicitly stated otherwise, going forward χ^2_{red} refers to $\chi^2_{\text{red}}[\text{SB}]$. When a quantity such as $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ varies on a time scale *slower* than ~ 1 ms, then $\chi^2_{\text{red}}[\text{groups}] = 1$ but $\chi^2_{\text{red}}[\text{SB}] > 1$, because the value of $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ will be approximately constant with respect to the group index (time after ablation), but it will vary from measurement to measurement.

The reduced chi-square statistic for the EDM channel, $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$, throughout the EDM data set, depends on the magnetic field magnitude $\mathcal{B}$ as discussed previously. Here, we will understand the source of this $\mathcal{B}$-dependence. In investigating possible sources of noise, it was useful to simulate EDM data under different well-controlled conditions. In particular, to understand the $\mathcal{B}$-dependence of the excess noise, I created simulations with the ability to vary the molecular pulse shape, arrival time, average velocity, and velocity dispersion, among many other parameters[10]. We discovered that including part-per-thousand trace-to-trace average molecular velocity fluctuations in the simulations reproduced excessive $\chi^2_{\text{red}} > 1$ values at larger magnetic fields such as $\mathcal{B} = 26$ mG. In the simplest model, we can even understand this while neglecting velocity dispersion entirely. Then the accumulated Zeeman phase in any state is $\phi_{\text{Zeeman}} = -g\mu_B \mathcal{B} \times \tau = -g\mu_B \mathcal{B} \times \frac{L}{v}$, where $L \approx 20$ cm is the length of the precession region and $v \approx 200$ m/s is the molecular forward velocity. If the standard deviation of the forward velocity among distinct traces is σ_v, then the corresponding phase

---

10. I highly recommend writing analogous simulations to understand future ACME data. Of course, we have to think of the right parameters to include as variables in the simulation. We'll see that the $\mathcal{B}$-independent component of the noise arose due to variation in experimental parameters that we did not initially realize could change in our system. For this reason, the simulations are not the silver bullet we would wish them to be.
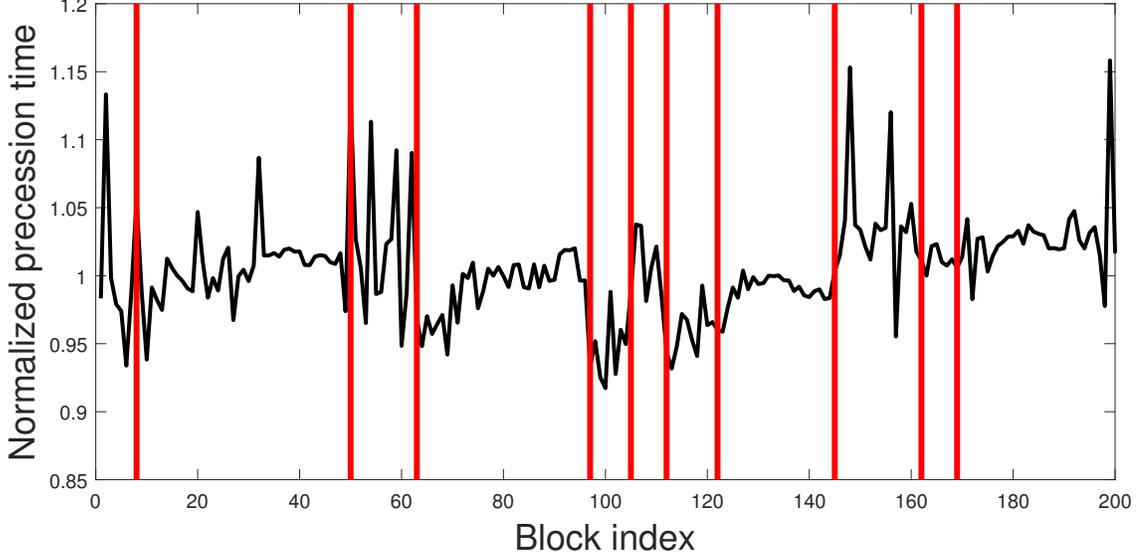
Figure 4.14: Average precession time measurements for 200 blocks in Run 260, normalized to the median value over the entire run to emphasize fractional deviations. Vertical red lines denote blocks where the ablation laser position ("spot") was changed. Depending on the local properties of the ablation target, the precession time will vary by $\sim 1 - 10\%$ between blocks when the ablation spot is not moved, and in some cases (not shown) changes up to $\sim 20\%$ when the ablation spot is moved.

noise is $\sigma_\phi \approx |g|\mu_B \mathcal{B}\frac{L}{v}\frac{\sigma_v}{v} = |\phi_{\text{Zeeman}}|\frac{\sigma_v}{v}$. With $\mathcal{B} = 26$ mG, $|\phi_{\text{Zeeman}}| \sim 1$ rad. In a typical group with high signal, the uncertainty in the asymmetry is $\sigma_{\mathcal{A}}^{\text{grp}} \sim 10^{-3}$, so we see that approximately part-per-thousand velocity fluctuations between traces will increase χ_{red}^2 by a factor of order unity.

Note that $\sigma_v^{\text{traces}}/v \sim 10^{-3}$, the fractional change in velocity between traces, corresponds to $\sigma_v^{\text{pulses}}/v \sim 5 \times 10^{-3}$, the fractional change in v between pulses assuming random noise, or $\sigma_v^{\text{pulses}}/v \sim 4 \times 10^{-5}$ assuming a linear drift in time. We expect that the random noise model more closely describes the operation of the beam source on these fast time scales (~ 0.5 seconds), though we know that the average molecular velocity produced in a single ablation laser position can also drift over many minutes. Alternatively, $\sigma_v^{\text{traces}}/v \sim 10^{-3}$ corresponds to $\sigma_v^{\text{blocks}}/v \sim 10^{-4}$, the fractional change in the average velocity between *blocks* (64 traces) assuming random noise, or a $\sigma_v^{\text{blocks}}/v \sim 6\%$ assuming a linear drift between blocks. An example of the noise in the measurement precession time, from block to block in a typical EDM data run, is shown in Fig. 4.14. We see that, depending on the ablation conditions, block-to-block fractional velocity variations of $\sim 1 - 10\%$ are typical. In light of this, I'd

argue that it is somewhat surprising how *small* the trace-to-trace velocity variations must be in order to account properly for the observed phase noise. Unfortunately, we have no way of directly measuring the velocity aside from computing $\phi^{\tilde{\mathcal{B}}}$, which requires acquiring an entire block of data.[11]

In summary, order-of-magnitude estimates, supported by simulated data with adjustable velocity noise, show that fractional molecular velocity variations of $\sim 10^{-3}$ between traces leads to $\chi^2_{\mathrm{red}} \sim 5$ at $\mathcal{B} = 26$ mG. Note that the exact value of χ^2_{red} arising from this mechanism depends on the distribution of velocity fluctuations as well as the group-assigned uncertainty, which varies by a factor of ~ 1.5 among runs due to depletion of the ablation targets and corresponding changes in the fluorescence signal. The observed value of $\chi^2_{\mathrm{red}} \approx 8$ when $\mathcal{B} = 26$ mG is thus well explained. At $\mathcal{B} = 2.6$ mG, the contribution of velocity fluctuations to the observed $\chi^2_{\mathrm{red}} \approx 3$ is negligible, so another noise source must be accountable.

4.15.3 Triggering noise between pulses

Throughout the ACME II data acquisition and analysis, we were unable to determine the source of the excess noise at $\mathcal{B} \leq 2.6$ mG, where $\chi^2_{\mathrm{red}} \approx 3$. However, we did notice that the noise was exacerbated ($\chi^2_{\mathrm{red}} \sim 10$) under either of two conditions: (1) the delay between the beginning of a polarization bin and the integrated sub-bin used to compute S_X or S_Y (see Fig. 4.15) differs by a non-zero time, δt, between the X and Y quadratures, even when the difference in the delay was as small as a single DAQ sample (62.5 ns); or (2) the beginning of the integrated sub-bin lies in the middle of the sharp rise in the fluorescence signal. These conditions only affect the noise of the dataset–not consistent offsets for any measurement channel–and the additional noise appears in every asymmetry channel. In particular, the EDM mean is consistent for all reasonable choices of the integration sub-bins for both X and Y.

A full-time, dedicated investigation of the excess noise at low $\mathcal{B}$-fields was pursued after completion of the ACME II analysis, primarily by Cole Meisenhelder, Mohit Verma, and

---

11. We have investigated alternative ways to measure the velocity, such as creating a "notch" in the molecular pulse by briefly turning off the STIRAP lasers. However, due to the large velocity dispersion in the molecular beam ($\sim 10\%$), the sensitivity of this measurement could not improve on our usual method using Zeeman precession.
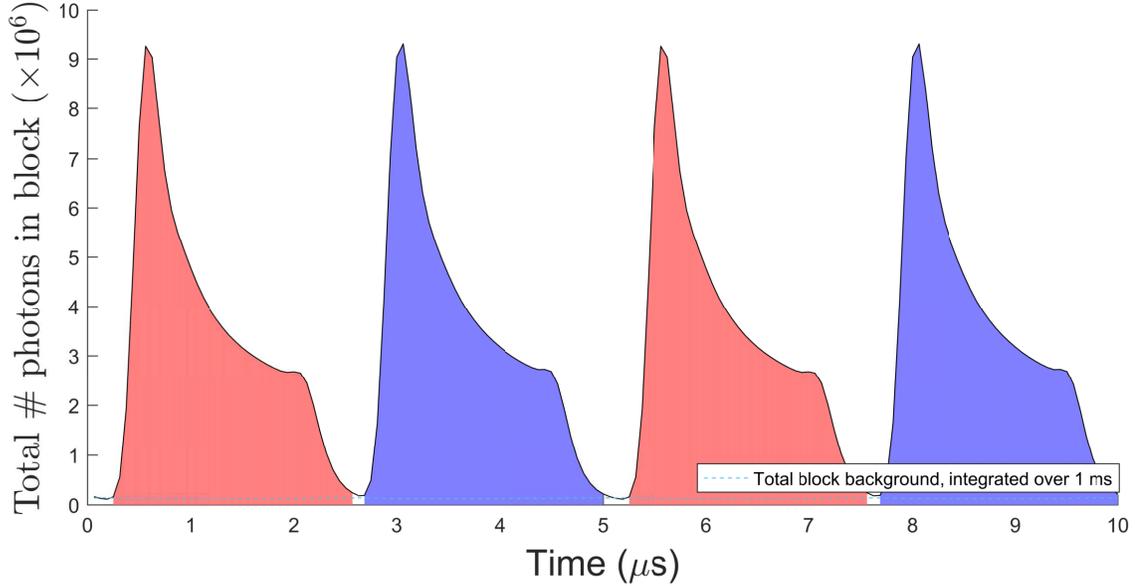
Figure 4.15: Polarization switching cycle from an example block (260.7.0). The total signals S_X and S_Y for an entire block are shown as a function of time within the polarization switching cycle. The region highlighted red is used to compute S_X and the region highlighted blue is used to compute S_Y in the EDM data set. For reference, the dashed line shows the total background signal integrated over a 1 ms duration (approximately the duration of the high-signal region of the molecular pulse).

Cristian Panda. A full explanation for the mechanism has recently been achieved; details of the measurements performed will be found in Cris Panda's thesis (in preparation at the time of this writing). For completeness, I also describe the basic principle here.

We can begin to understand this behavior by observing the asymmetry as a function of time *within* a sub-bin, and as a function of the relative offset between X and Y bins: see Fig. 4.16. There, we compute the time-dependent asymmetry

$$\mathcal{A}(t, \delta t) = \frac{S_X(t) - S_Y(t + \delta t)}{S_X(t) - S_Y(t + \delta t)} \tag{4.43}$$

for all values of t within a single polarization switching cycle. Here, the reference fluorescence signal is averaged over an entire run to suppress the effect of shot noise, which is not of interest in the current discussion. We see that, when $\delta t = 0$, the asymmetry is approximately independent of time within the bulk of the fluorescence region. However, near the turn-on and turn-off points of the excitation laser ($\approx 0.2 \, \mu s$ and $\approx 2.1 \, \mu s$ in Fig. 4.16), the asymmetry is a strong function of time within the polarization bin. This must be the case to
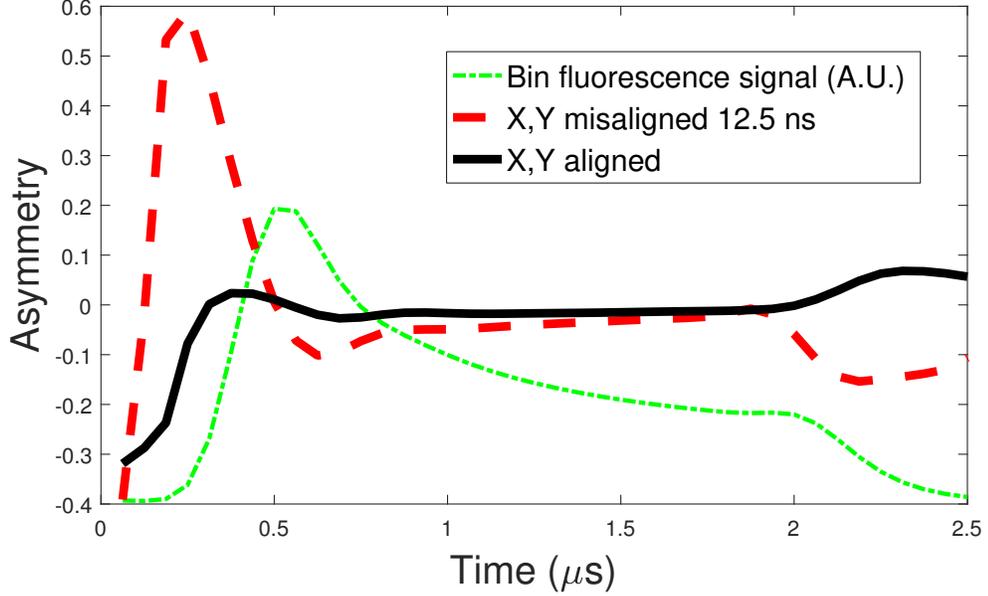
Figure 4.16: Reference fluorescence signal (dot-dashed green curve) and asymmetry as a function of time within a polarization switching cycle (solid black curve), when the X and Y bins are approximately aligned ($\delta t = 0$). When the X and Y curves are misaligned by a fraction of a DAQ sample ($\delta t = 12.5$ ns), the asymmetry displays a much stronger dependence on time within the polarization bin (dashed red curve).

some extent because the background region has a different asymmetry value (dominated by stray light and electronic offsets in the PMT signal) than the high-signal region (dominated by the molecular phase).

However, we also see that this effect is exacerbated if the X and Y signals are computed with a relative time offset, $\delta t \neq 0$. We can understand the origin of the dependence of the asymmetry on the relative X and Y bin offset, $\frac{\partial \mathcal{A}}{\partial(\delta t)}$, as follows. Suppose, in the simplest case, that $S_Y(t) = S_X(t)$ so that the asymmetry is $\mathcal{A}(t) = 0$ for all t. Then $S_Y(t + \delta t) \approx S_X(t) + S'_X(t)\delta t$ for sufficiently small δt, and hence $\mathcal{A}(t, \delta t) \approx -\frac{S'_X}{2S_X}\delta t$. We see that the asymmetry in this case is proportional to the local slope of the fluorescence signal, $S'_X(t)$, and to the misalignment δt between the X and Y bins. Note that in Fig. 4.16, the difference between the red and black curves, $\frac{\partial \mathcal{A}}{\partial(\delta t)}\delta t$, is proportional to the local slope of the fluorescence signal (green curve).

We are nearly ready to consider the calculated asymmetry used to compute the EDM, which is obtained from integrating S_X and S_Y over sub-bins. In order to obtain the highest SNR possible, we typically integrate the signal over a region that begins to display a non-
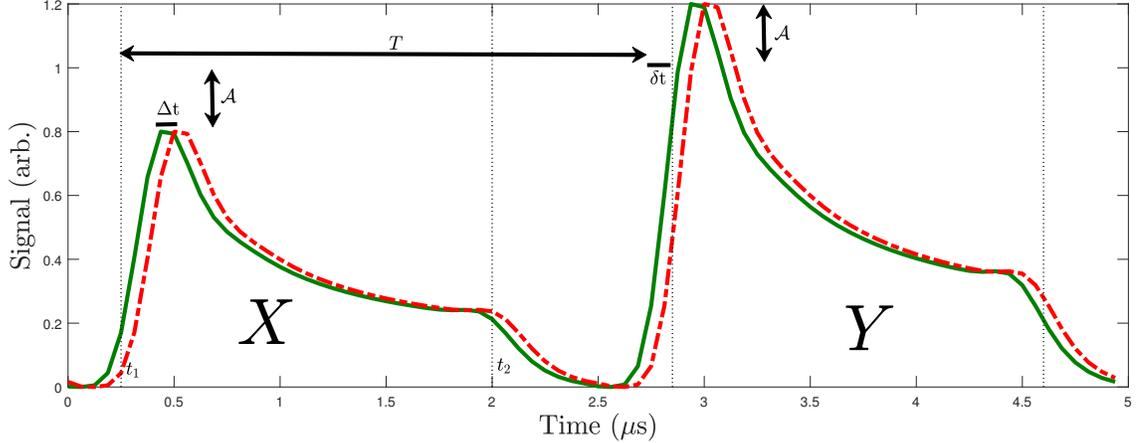
Figure 4.17: A reference fluorescence signal (solid green) and a signal with a global timing offset (dot-dashed red), $\Delta t = 62.5$ ns. The X signal is integrated from t_1 to t_2, denoted by dotted vertical lines. The Y signal is integrated from $t_1 + T + \delta t$ to $t_2 + T + \delta t$ (also shown with dotted vertical lines). Note that the functional form of the signal is periodic with period T, up to a difference in amplitude in the X and Y bins due to the non-zero asymmetry (shown here with $\mathcal{A} = 0.2$, independent of time). As a result, the integrated regions of the X and Y bins are out of phase, with a timing offset of $\delta t = 100$ ns (e.g., the beginning of the integration region for X occurs when the signal is below half its peak, while the beginning of the integration region for Y begins when the signal is above half its peak).

negligible asymmetry vs. time dependence, near the edges of the high-signal region (see Fig. 4.15). Let the actual time delay between the X and Y bins be $T \approx 2.5\,\mu$s, and denote the signal in a bin pair by $S(t)$. We interpret approximately the first half of $S(t)$ to correspond to $S_X(t)$ and the second half to correspond to $S_Y(t)$. It will be necessary to relax our simplifying assumption that $\mathcal{A}(t) = 0$ from the previous discussion, in order to fully describe the ways in which timing noise can manifest as asymmetry noise. In particular, recall that even with $\delta t \approx 0$ (i.e., X and Y bins are aligned with respect to each other), the "local" asymmetry is a function of time within the polarization switching cycle (see Fig. 4.16, black curve). Therefore, a global timing offset in the signal, Δt, leads to a change in the average measured asymmetry, due to the varying value of $\mathcal{A}(t)$ near the beginning and end of the sub-bin used to compute the measured asymmetry. We will see that fluctuations in this global timing offset can lead to noise in the measured asymmetry, through this mechanism independently of noies due to fluctuations in the relative timing, δt, between X and Y bins.

187

With these observations in mind, we will build up a model of how the measured asymmetry depends on timing noise. Suppose that we integrate the X bin from $t_1 + \Delta t$ to $t_2 + \Delta t$ and the Y bin from $t_1 + \Delta t + T + \delta t$ to $t_2 + \Delta t + T + \delta t$, where Δt is some global offset in the signal relative to an arbitrary reference (see Fig. 4.17). Throughout the ACME II data set, there were uncontrolled timing offsets, $\Delta t \leq 100$ ns, due to DAQ triggering errors, which were understood and fixed by Cole Meisenhelder after the acquisition of the EDM data. These timing offsets randomly vary from pulse to pulse (not only trace to trace–this is critical, as we'll see later). We suppose that the fluorescence waveform is periodic with period T, up to the possibility of different amplitudes in the X and Y bins, so that $S(t + T) \propto S(t)$. The proportionality constant is found by the condition that $[S(t) - S(t+T)]/[S(t) + S(t+T)] = \mathcal{A}(t)$ is the "local" asymmetry within the polarization cycle. This model allows us to connect the integrated signal in the X bin, $\int_{t_1+\Delta t}^{t_2+\Delta t} dt\, S(t)$, to the integrated signal in the Y bin, $\int_{t_1+\Delta t+T+\delta t}^{t_2+\Delta t+T+\delta t} dt\, S(t)$, and construct the asymmetry $\mathcal{A}_{\text{meas}}$ in the usual way. In the simplest case of a constant timing offset Δt for every pulse within a trace, we find:

$$\mathcal{A}_{\text{meas}}(\Delta t, \delta t) = \frac{\int_{t_1+\Delta t}^{t_2+\Delta t} dt [(1 + \mathcal{A}(t))S(t) - (1 - \mathcal{A}(t+\delta t))S(t+\delta t)]}{\int_{t_1+\Delta t}^{t_2+\Delta t} dt [(1 + \mathcal{A}(t))S(t) + (1 - \mathcal{A}(t+\delta t))S(t+\delta t)]}. \tag{4.44}$$

Here, the prefactors $1 \pm \mathcal{A}(t)$ adjust the relative amplitudes of S_X and S_Y based on the time-dependent asymmetry $\mathcal{A}(t)$. Note, as a sanity check, that for a constant asymmetry $\mathcal{A}$ within the sub-bin, the expected result $\mathcal{A}_{\text{meas}}(\Delta t, \delta t = 0) = \mathcal{A}$ is obtained. It is clear that $\mathcal{A}_{\text{meas}}(\Delta t, \delta t)$ is a rather complicated function in the general case that $\mathcal{A}(t)$ and $S(t)$ are not constant. In the actual situation applicable to ACME II data, there is a distribution of global timing offsets, Δt, among pulses within a trace, making the dependence of the integrated asymmetry $\mathcal{A}_{\text{meas}}$ on the timing offsets significantly more difficult to understand.

Nevertheless, it is fairly clear that $\frac{\partial \mathcal{A}_{\text{meas}}}{\partial(\Delta t)} \neq 0$ even in the simplest case of a fixed timing offset for the entire trace. This dependence of the computed asymmetry on an overall timing offset arises from at least two sources: (1) the physical dependence of the asymmetry on the time within the sub-bin, $\mathcal{A}(t)$; and (2) the artificial dependence of the computed asymmetry on time within the sub-bin, $\frac{\partial \mathcal{A}}{\partial(\delta t)} \delta t$, due to a timing misalignment δt between the regions of

the X and Y bins used to compute the signals from each phase quadrature, S_X and S_Y.

The latter noise source can be suppressed by careful adjustment of the sub-bins that we integrate over–if necessary, even to timing precision greater than a single DAQ sample, by interpolating the measured signal between samples. The former noise source can be suppressed by choosing very wide integration sub-bins that extend reasonably far into the regions of zero signal, where the dependence of the physical asymmetry on time is rather small. Note, however, that due to a finite dead time between X and Y pulses in ACME II, it is not feasible to completely eliminate the dependence of the asymmetry on time within the sub-bin.

For typical traces and analysis parameters, I find a dependence $\frac{\partial \mathcal{A}_{\text{meas}}}{\partial(\Delta t)} \sim 3 \times 10^{-6}/\text{ns}$. Typically, the shot-noise limit for a single trace is $\sigma_{\mathcal{A}} \sim 10^{-4}$, so the contribution to the asymmetry uncertainty from timing jitter is comparable to shot noise provided $\Delta t \sim 30$ ns for a trace. Indeed, this estimate is quite close to the actual trace-to-trace timing noise throughout ACME II arising from the DAQ triggering errors, as we will see in the following section. By monitoring the timing of the DAQ trigger directly, we have determined that pulse-to-pulse variations in the global timing offset of approximately $\Delta t \sim 60$ ns were typical throughout the ACME II data set.

After eliminating the DAQ triggering error (thus setting Δt to be a constant, which we can take to be zero) and adjusting the relative timing of the X and Y bins via the AOMs that control the probe laser light (thus setting $\delta t = 0$), Cris Panda and Cole Meisenhelder have acquired a small EDM data set, involving superblock switches exactly as in the ACME II data, and demonstrated a value of χ^2_{red} consistent with 1 in the EDM channel, $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$.

4.15.4 Retroactive elimination of excess noise?

Having understood the noise mechanism–namely, an asymmetry dependence $\frac{\partial \mathcal{A}_{\text{meas}}}{\partial(\Delta t)}$ together with an overall timing noise jitter Δt that varies from pulse to pulse–it is natural to wonder whether we can eliminate it in the ACME II systematic error and EDM data by modifying the data analysis. While we can do so for timing jitter on the *trace* level, I have not found a way to eliminate the noise arising from timing jitter on the pulse level.

In a bit more detail, we can measure the timing jitter on an individual trace and
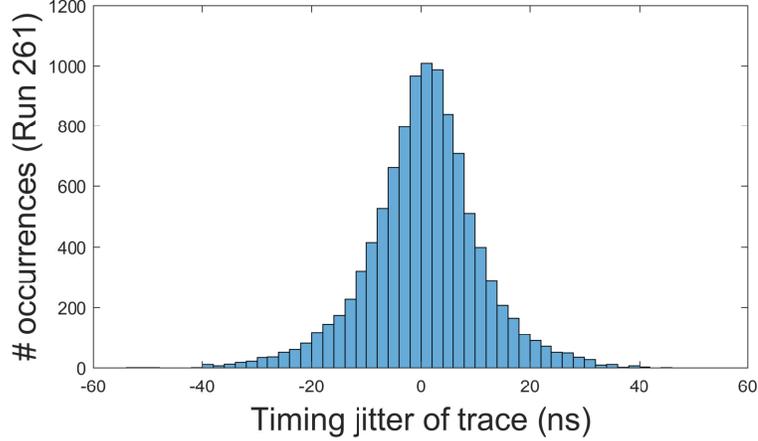
Figure 4.18: Distribution of trace-level timing offsets Δt from Run 261.

correct the integration sub-bins for that jitter. Specifically, we can measure an average fluorescence signal for the X bins, $S_{\mathrm{ref}}(t)$, over some relatively large data set so that $S_{\mathrm{ref}}(t)$ has negligible contribution from photon shot noise, and compare to the actual signal measured in the X bins for a given trace, $S(t)$. We normalize the signals such that $\int_{t_{\mathrm{start}}}^{t_{\mathrm{end}}} dt S_{\mathrm{ref}}(t) = \int_{t_{\mathrm{start}}}^{t_{\mathrm{end}}} dt S(t) = 1$; i.e., the integrated signal from the beginning to the end of the full bin is unity. This common normalization allows us to compare $S(t)$ to $S_{\mathrm{ref}}(t)$ directly.

We suppose that $S(t) = S_{\mathrm{ref}}(t + \Delta t)$, i.e., that the signal of a given trace is offset in time by Δt compared to the reference. For small timing offsets, we Taylor expand $S(t) = S_{\mathrm{ref}}(t) + \frac{dS_{\mathrm{ref}}(t)}{dt}\Delta t + \epsilon(t)$, where $\epsilon(t)$ is the residual for our model. Note that for a well-defined (low-noise) reference signal $S_{\mathrm{ref}}(t)$, the local slope $\frac{dS_{\mathrm{ref}}(t)}{dt}$ can be calculated to high precision. I compute the value of Δt that minimizes the sum of squares of residuals, $\frac{\partial}{\partial(\Delta t)} \sum [S(t) - S_{\mathrm{ref}}(t) - \frac{dS_{\mathrm{ref}}(t)}{dt}\Delta t]^2 = 0$, to infer the actual timing offset of the signal $S(t)$ relative to $S_{\mathrm{ref}}(t)$. Direct inspection of the offsets between trace-level signals $S(t)$ and the reference signal $S_{\mathrm{ref}}(t)$ verifies that the value of Δt assigned in this way accurately describes the timing offset of the data. A distribution of Δt values obtained for a run of EDM data is shown in Fig. 4.18. Unsurprisingly, the distribution is not simply described by a typical form (e.g., Gaussian or Lorentzian), but it is centered around $\Delta t \approx 0$ with standard deviation $\sigma_{\Delta t} \approx 10$ ns.

Once the timing offset of a trace is determined with the above-described procedure, we

190

can adjust the integration sub-bins accordingly for each trace. For example, suppose we find that $\Delta t = t_{\text{DAQ}} = 62.5$ ns, the time of a single DAQ sample. In this case, the actual trace is shifted to the left in a signal-vs-time plot. We should therefore integrate S_X and S_Y from one sample earlier than for our default sub-bins; e.g., $S_X \equiv \sum_{i=i_1}^{i_2} S_i \rightarrow \sum_{i=i_1-1}^{i_2-1} S_i$, where S_i is the signal on the i-th DAQ sample, i_1 is the first sample typically included in the sub-bin, and i_2 is the last sample typically included in the sub-bin. For a timing offset smaller than a DAQ sample, $|\frac{\Delta t}{t_{\text{DAQ}}}| < 1$, we instead compute S_X from a weighted sum over samples. E.g., for $\Delta t > 0$, we compute

$$S_X \rightarrow \sum_{i=i_1}^{i_2} S_i + \frac{\Delta t}{t_{\text{DAQ}}}(S_{i_1-1} - S_{i_2}). \tag{4.45}$$

Except for the sub-bins adjusted in this way, the asymmetry and all subsequent data analysis can proceed as normal. I have directly confirmed that this procedure suppresses the effect of trace timing offsets on the calculated asymmetry, $\frac{\partial \mathcal{A}}{\partial(\Delta t)}$, by at least a factor of 20 (e.g., by analyzing the same trace with all data shifted by one DAQ sample).

Unfortunately, this procedure does not work at all to suppress pulse-to-pulse timing noise. For example, the asymmetry computed in this way from a trace $S(t) \equiv [S_{\text{ref}}(t + \Delta t) + S_{\text{ref}}(t - \Delta t)]/2$, which has components with timing offsets in opposite directions, differs from the asymmetry computed from a reference trace $S_{\text{ref}}(t)$ by a comparable amount ($\sim 3 \times 10^{-6} \, \Delta t/\text{ns}$) as in the naive analysis where no effort is made to adjust for timing jitter. Because in the ACME II data set we only have access to traces, summed over 25 pulses that have random and independent timing offsets, we do not appear to be able to compensate for the timing jitter by adjusting the integration sub-bin on a trace-to-trace basis.

Chapter 5

Electric and magnetic fields

90% bug free!

Minecraft

In the ACME experiment, control of electric and magnetic fields is paramount for suppression of systematic errors. In this chapter, I describe the generation of electric and magnetic fields in the interaction region, as well as our measurements of these fields and their imperfections.

5.1 Electric field measurements

In ACME II, we apply electric fields up to 140 V/cm, with a 4.5 cm separation between the field plates, requiring voltages up to ± 315 V. Due to the larger field plate separation compared to ACME I, this exceeds the ± 200 V capabilities of the previously-used power supplies. Therefore, we use custom-built low-noise voltage supplies based on the PA89A power op-amp, built by Jim MacArthur (Chief Engineer of the Harvard Instrument and Design Lab). These supplies can generate ± 500 V, controlled by a 20-bit "BiasDAC" (also designed and built by MacArthur). The PA89A op-amps are powered by Acopian P01HA30 and N01HA30 regulated power supplies, which provide ± 600 V. As we will see, differential voltage offsets between the two low-noise voltage supplies contribute $\mathcal{E}^{\tilde{\mathcal{L}}} \sim 5$ mV/cm, but the effect of these non-reversing fields on the EDM measurement is suppressed by the $\tilde{\mathcal{L}}$ switch, compared to a true non-reversing electric field $\mathcal{E}^{\mathrm{nr}}$. Directly monitoring the differential

voltage between the field plates for a few days, we observed up to $25\,\mu\mathrm{V/hr}$ drifts in the differential voltage between the plates, which doesn't reverse with the electric field or power supply lead switches, corresponding to $5\,(\mu\mathrm{V/cm})/\mathrm{hr}$ drifts in the non-reversing electric fields (see [200]). We view this as an upper bound on the field plate voltage drift, since the Fluke 8846A digital multimeters used to measure the field plate voltages are specified to drift up to $25\,\mu\mathrm{V/hr}$ under the applicable conditions. The lead switch, $\tilde{\mathcal{L}}$, is implemented via a home-built high-voltage circuit based on the Pickering 101-2-A-12/6 mercury-wetted relay; for details, see the internal ACME document [201].

We can create a map of the electric field, $\mathcal{E}(\vec{x})$, between the refinement and probe lasers, as a function of position in the interaction region, for each experimental state[1]. By comparing the field measured in different experimental states, we can infer (most crucially) the applied and non-reversing components of the electric field in the interaction region. This is especially important because, as we have seen in Ch. 4, non-reversing electric fields $\mathcal{E}^{\mathrm{nr}}$ in the STIRAP and probe regions are essential for several distinct systematic error contributions. The method used to map the electric field in ACME II is nearly identical to the method of ACME I, which has been documented in detail elsewhere [129, 146, 149]. For clarity, I will describe the method briefly here. The microwave measurement of the electric fields was implemented in ACME II primarily by Adam West.

The electric field measurement relies on microwave spectroscopy of the H state, which is susceptible to large Stark shifts due to the small Ω-doublet splitting ($a = 181(1)$ kHz, with splitting $2a$ in $J = 1$[146, Appendix D]) and non-negligible electric dipole moment ($D_{J=1} = 2\pi \times 2.13(2)$ MHz/(V/cm)[127]). Therefore, microwave spectroscopy of the H state manifold with $\sim$kHz resolution allows for $\approx$mV/cm resolution of electric fields, as we will see. Suppose an electric field is applied, $\mathcal{E} \neq 0$, and molecules are prepared in the $J = 1$ manifold with angular momentum aligned at angle ϕ with respect to the x-axis. We excite these molecules via microwaves addressing a well-defined-parity state $H(J = 2, M = 0, P)$. In the ACME apparatus, we use microwaves propagating along $-\hat{x}$, from

---

1. As we will see, the protocol to measure the electric field intrinsically gives spatial resolution along the x-direction. However, by repeating the measurement with the STIRAP or probe laser beams clipped, we can also obtain crude spatial resolution along y and z.

outside the "dump" vacuum chamber and along the molecular beam axis. To understand the measurement scheme, consider first the simplest case, when the molecular phase is prepared along $\phi = \frac{\pi}{4}$, so that an equal number of molecules are projected into the X and Y quadratures by the readout lasers. Then an ordinary asymmetry measurement, in the absence of microwave depletion, will yield $\mathcal{A} = \frac{S_X - S_Y}{S_X + S_Y} = 0$. However, now suppose that the microwaves polarized along $\hat{y}$ resonantly address the $P = +1$ state in $J = 2$ and drive a π pulse, transferring all population from the Y quadrature to the $J = 2$ manifold. Then the asymmetry measurement will yield $\mathcal{A} = 1$, since the $H(J = 1)$ ThO molecules only have population remaining in the X quadrature.

This suggests the following measurement protocol. We prepare the molecules via STI-RAP with angular momentum alignment approximately along $\hat{x}$, and subsequently reproject the state along $\phi = \frac{\pi}{4}$ via the refinement laser. This reprojection reduces the experimental signal by a factor of 2 but is necessary because the microwaves can only have polarization in the yz-plane due to their propagation along $-\hat{x}$. We then excite the molecules with a pulse of microwaves with angular frequency $\omega = \omega_0 + \delta$, where $\omega_0 \approx 2\pi \times 40$ GHz is the resonant frequency of the $H(J = 1, M = \pm 1) \leftrightarrow H(J = 2, \tilde{\mathcal{P}})$ transition in the absence of electric fields. Here, $\delta \ll \omega_0$ is a small detuning from the zero-field resonance.

For an applied electric field of the simple form $\mathcal{E}(\tilde{\mathcal{E}}) = \mathcal{E}\tilde{\mathcal{E}} + \mathcal{E}^{\mathrm{nr}}$, the linear Stark shift is $\langle H_{\mathrm{Stark}} \rangle = D_{J=1}|\mathcal{E}(\tilde{\mathcal{E}})|\tilde{\mathcal{N}} = D_{J=1}(\mathcal{E}\tilde{\mathcal{N}} + \mathcal{E}^{\mathrm{nr}}\tilde{\mathcal{N}}\tilde{\mathcal{E}})$. This expression can be generalized for additional switches; for example, an electric field component that reverses with the power supply lead switch contributes to a Stark shift parity component $\langle H_{\mathrm{Stark}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}\tilde{\mathcal{L}}} \rangle = D_{J=1}\mathcal{E}^{\tilde{\mathcal{L}}}$. We set the combination of microwave power and pulse time such that a π pulse occurs for $\delta = \langle H_{\mathrm{Stark}} \rangle$, and $\mathcal{A}$ is maximized when the microwaves are on resonance. The pulse time is $T_\pi \sim 50\,\mu$s, which is a small fraction of the ~ 1 ms precession time in the interaction region and sets the linewidth of the microwave transition, $\delta\omega \approx 2\pi \times 1.42/T_\pi \approx 2\pi \times 28$ kHz (FWHM). We time the microwave pulse, relative to the ablation laser pulse, such that a significant number of molecules are addressed at all positions in the ≈ 20 cm region between the refinement and probe lasers, and subsequently continuously read out the asymmetry as a function of arrival time after the microwave pulse. The molecules that arrive in the probe region first are those closest to the probe when the microwaves arrive, while those that
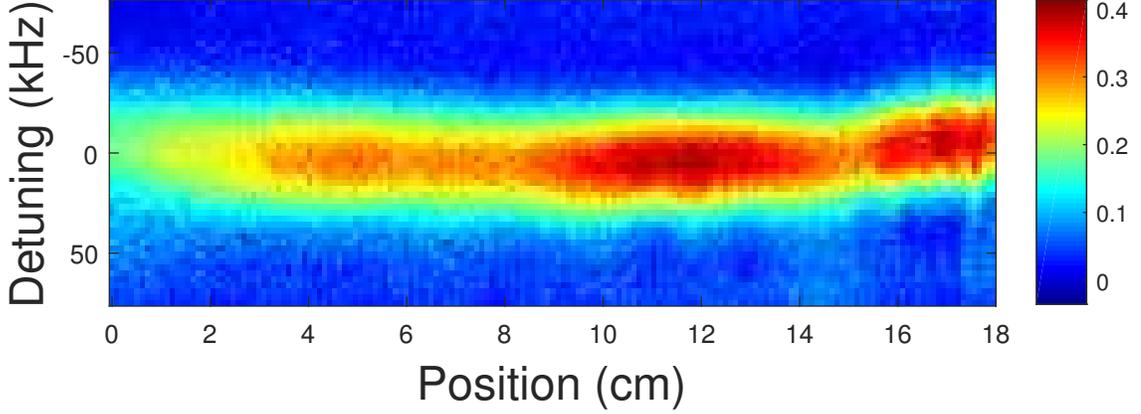
Figure 5.1: A map of asymmetry vs. detuning and approximate position within the interaction region (refinement beam at $x \approx 0$, probe beam at $x \approx 20$ cm). The red and yellow band shows the dependence of the microwave resonance x.

arrive last are those farthest from the probe (i.e., toward the refinement laser). Assuming a typical molecular velocity of $v \approx 200$ m/s, we can interpret the molecular arrival time as proportional to the x-position in the interaction region.

With this method, we can create a 2D map of the asymmetry as a function of the detuning δ and position x within the interaction region, in each experimental state; see Fig. 5.1. At each x-position (vertical slice of Fig. 5.1), we fit the asymmetry vs. detuning to find the resonant frequency $\delta_{\mathrm{res.}}(x) = \tilde{\mathcal{N}} D_{J=1} |\mathcal{E}(\tilde{\mathcal{E}})|(x) + \delta_0$, up to an overall constant offset detuning δ_0. We then compute parity sums over different experimental conditions to compute, e.g., $\mathcal{E}^{\mathrm{nr}}(x) = \delta_{\mathrm{res.}}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}(x)/D_{J=1}$, the non-reversing electric field as a function of position within the interaction region. More generally, we typically compute all parity components of $\mathcal{E}(x)$ with the switches $\tilde{\mathcal{N}}$, $\tilde{\mathcal{E}}$, and $\tilde{\mathcal{L}}$. A careful analysis shows that $\mathcal{E}^{\tilde{s}} = \delta_{\mathrm{res.}}^{\tilde{s}}/D_{J=1}$ for $\tilde{s} \neq \tilde{\mathcal{N}}\tilde{\mathcal{E}}$.

See Fig. 5.2 for all components as a function of position along the beam line. Several features are worth noting: there is a significant (several mV/cm) offset in the channels $\mathcal{E}^{\mathrm{nr}}$, $\mathcal{E}^{\tilde{\mathcal{L}}}$, and $\mathcal{E}^{\tilde{\mathcal{N}}}$. The non-reversing fields $\mathcal{E}^{\mathrm{nr}}$ are expected due to patch potentials on the electric field plates, which do not reverse with the nominal supply voltages. The lead-switch-correlated field $\mathcal{E}^{\tilde{\mathcal{L}}}$ is expected from the ~ 10 mV power supply voltage offsets, which contribute oppositely to the applied electric field in the $\tilde{\mathcal{L}} = +1$ configuration compared to $\tilde{\mathcal{L}} = -1$. We do not have a strong model for $\mathcal{E}^{\tilde{\mathcal{N}}} \sim$mV/cm, since the electric field should

195
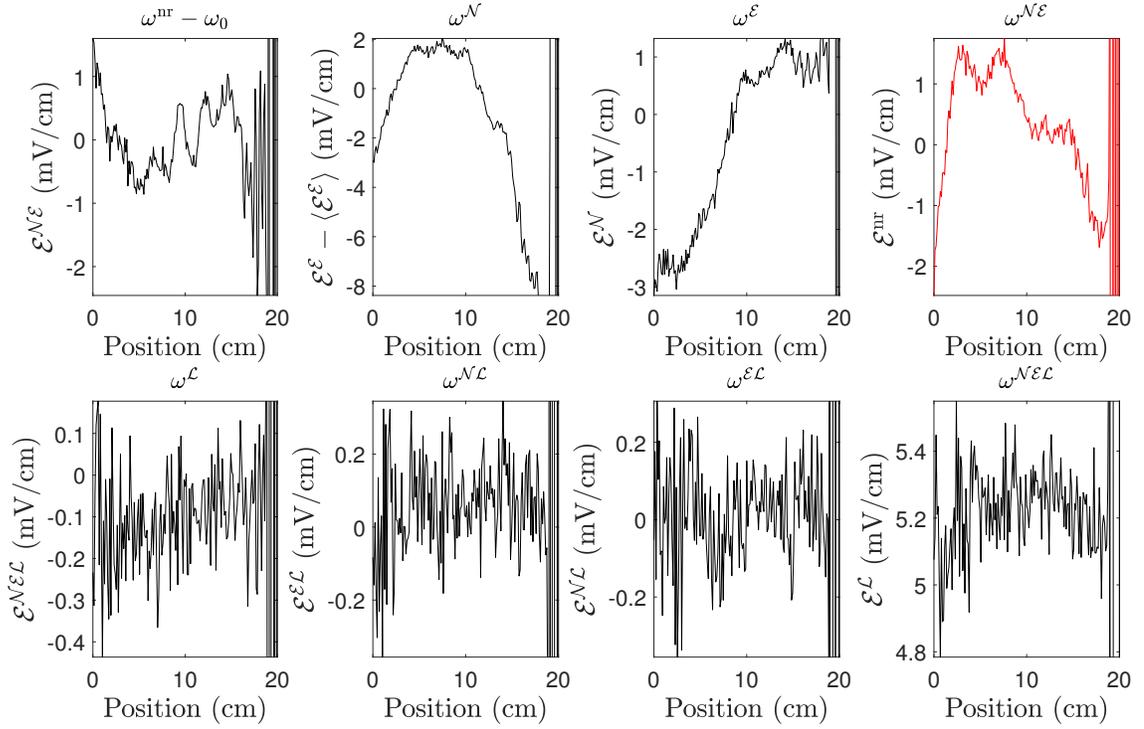
Figure 5.2: Switch-parity components of electric field measurements $\mathcal{E}(x)$ with switches $\tilde{\mathcal{N}}, \tilde{\mathcal{E}}$, and $\tilde{\mathcal{L}}$. Subplot titles denote the frequency component from which the electric field parity component is extracted (e.g., $\mathcal{E}^{\mathrm{nr}}$ is inferred from the component of the resonance frequency $\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$). As expected, $\mathcal{E}^{\mathrm{nr}}$ (top-right, red) and $\mathcal{E}^{\tilde{\mathcal{L}}}$ are non-zero. The average value of ω^{nr} is removed from the top-left plot to isolate $\mathcal{E}^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ from the zero-field resonant microwave frequency, $\omega_0 \approx 40$ GHz. Similarly, the average reversing electric field, $\langle \mathcal{E}^{\mathcal{E}} \rangle \approx 80$ V/cm, is subtracted from the measured value of $\mathcal{E}^{\mathcal{E}}(x)$ for illustration of its variation along x. We see ≈ 10 mV/cm variation in the applied electric field between the refinement and probe regions due to bowing of the field plates. There is a small x-dependent offset in $\mathcal{E}^{\tilde{\mathcal{N}}}$, which is not well-understood. Data taken 2/17/18.

not be affected at all by which $\tilde{\mathcal{N}}$ state is prepared and probed by the lasers. We also routinely perform measurements by addressing both $P = \pm 1$ states in $H(J = 2)$, which produce consistent results as expected.

Furthermore, the applied electric field demonstrates a clear (though small) curvature, and varies by ~ 10 mV/cm over the length of the interaction region. This curvature arises physically from the "bowed" geometry of the electric field plates (see [149, Sec. 6.1]): $d_{\mathrm{FP}} \approx 4.5\,\mathrm{cm} + 0.1\,\mu\mathrm{m} \times (\frac{x}{\mathrm{cm}})^2$, where d_{FP} is the separation between the east and west field plates. For $x \approx \pm 10$ cm, the maximal change in the applied electric field is therefore expected to be $\sim 80\,\mathrm{V/cm} \times \frac{10\,\mu\mathrm{m}}{4.5\,\mathrm{cm}} \approx 18$ mV/cm, in reasonable agreement with the microwave spectroscopy data. We have strong evidence that this bowing is a result of the manner in which the electric field plates are clamped and held by their mounting structure.

We performed a microwave measurement to determine $\mathcal{E}^{\mathrm{nr}}$ approximately every two weeks in the period in which the EDM data set was acquired. Qualitatively, the structure of $\mathcal{E}^{\mathrm{nr}}(x)$ is consistent over time. In particular, $\mathcal{E}^{\mathrm{nr}}(x)$ always has a larger value (not necessarily magnitude) in the center of the interaction region (near $x \approx 10$ cm), and lies in the range $|\mathcal{E}^{\mathrm{nr}}| \lesssim 5$ mV/cm. We compute a typical value of $\mathcal{E}^{\mathrm{nr}}$ from the mean value, across all measurements, averaged over the preparation ($x \approx 0$) and probe ($x \approx 20$ cm) regions, where $\mathcal{E}^{\mathrm{nr}}$ is known to be capable of producing systematic errors. The uncertainty is computed by adding in quadrature the statistical uncertainty of the measurement and the difference between values in the preparation and probe regions alone, to reflect our uncertainty about which region is most relevant to the systematic errors. The statistical uncertainty in the measurement provides a negligible contribution. Computed in this way, we obtain $\Delta \mathcal{E}^{\mathrm{nr}} = -2.6(16)$ mV/cm.

5.1.1 Electric field gradients

Although we do not directly use any measured non-reversing electric field gradients, for example $\partial \mathcal{E}^{\mathrm{nr}}/\partial z$, to compute systematic error contributions or uncertainties in the EDM measurement, we periodically measure both $\partial \mathcal{E}^{\mathrm{nr}}/\partial z$ and $\partial \mathcal{E}^{\mathrm{nr}}/\partial y$ throughout the period of the EDM data set, by performing microwave measurements with artificially shifted molecular centers of mass. In particular, we can block either half of the STIRAP lasers (in the
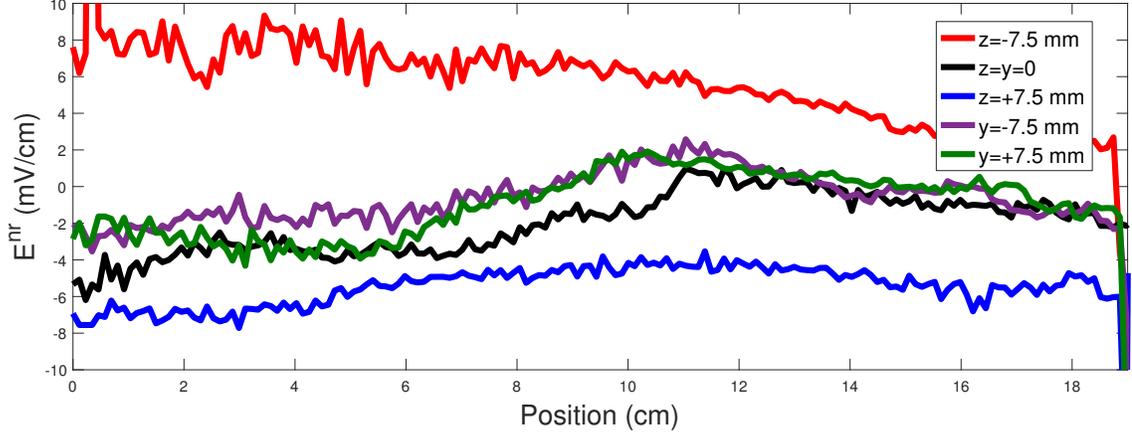
Figure 5.3: Non-reversing electric field, $\mathcal{E}^{\mathrm{nr}}$, for molecular population under ordinary conditions (black), and shifted along z (red and blue) or y (purple and green). The gradient $\partial\mathcal{E}^{\mathrm{nr}}/\partial y$ is negligible, while the gradient $\partial\mathcal{E}^{\mathrm{nr}}/\partial z$ is as large as -10 mV/cm^2. Data from 1/8/18.

regions $z > 0$ or $z < 0$) to shift the center of mass of molecules prepared in the H state along $\hat{z}$. Alternatively, we can block half of the probe lasers (in the regions $y > 0$ or $y < 0$) to shift the center of mass of molecules that are detected along $\hat{y}$. Using Monte Carlo simulations of the molecular trajectories for our experimental geometry (see Sec. 6.1.5 for details), we estimate that the center-of-mass shifts are $\langle z \rangle \approx \langle y \rangle \approx \pm 7.5$ mm, where $\langle z \rangle$ and $\langle y \rangle$ are the center of mass along z or y, respectively, of the molecular beam when the corresponding half of the preparation or readout lasers is blocked. This allows us to measure gradients of $\mathcal{E}^{\mathrm{nr}}$ across all three axes, throughout the interaction region.

We do not observe any significant change of $\mathcal{E}^{\mathrm{nr}}$ for $\langle y \rangle = \pm 7.5$ mm, so that we conclude $|\partial\mathcal{E}^{\mathrm{nr}}/\partial y| \lesssim 1$ mV/cm^2, a negligible value for our purposes (see Fig. 5.3). This is not particularly surprising, given the approximate translational symmetry of the field plates: we expect $\mathcal{E}_y \approx 0$ over the volume occupied by molecules, which would imply $\partial\mathcal{E}_y/\partial z \approx \partial\mathcal{E}_z/\partial y \approx 0$ in every experimental state, and thus $\partial\mathcal{E}^{\mathrm{nr}}/\partial y \approx 0$. Of course, this argument is merely heuristic and does not suggest how large the (inevitably non-zero) "correction terms" to $\partial\mathcal{E}^{\mathrm{nr}}/\partial y$ should be.

On the other hand, we observe strikingly clear gradients $|\partial\mathcal{E}^{\mathrm{nr}}/\partial z| \approx 10$ mV/cm^2 in the most extreme cases, by comparing measured values of $\mathcal{E}^{\mathrm{nr}}$ under conditions with $\langle z \rangle = \pm 7.5$ mm. To understand why the gradients are of this order of magnitude, we can consider a

couple of toy models.

First, model the electric field plates as infinite parallel plate capacitors separated by distance $L = 4.5$ cm. For consideration of $\mathcal{E}^{\mathrm{nr}}$, we are free to assume both plates are held at ground voltage, since the patch potentials generating $\mathcal{E}^{\mathrm{nr}}(\vec{x})$ are independent of the voltage applied to the field plates. Now suppose there is a line of deposited charge with linear density λ deposited on one plate along the line $y = 0$ and $z = -L/2$. Using the method of images and the standard results for the electric field from a line charge, we can find that the resulting electric field and gradient have the relationship $\frac{\partial \mathcal{E}^{\mathrm{nr}}}{\partial z}(z = 0) \approx \mathcal{E}^{\mathrm{nr}}(z = 0)/(L/5)$, where I've assumed that we probe the electric field at $y = 0$.

Alternatively, consider a point charge deposited on one plate at $x = y = 0$ and $z = -L/2$. Once again using the method of images and standard field results, we would find $\frac{\partial \mathcal{E}^{\mathrm{nr}}}{\partial z}(z = 0) \approx \mathcal{E}^{\mathrm{nr}}(z = 0)/(L/6)$. Therefore, electric field gradients on the order of $\frac{\partial \mathcal{E}^{\mathrm{nr}}}{\partial z} \sim (3$–$10) \times \frac{\mathcal{E}^{nr}}{L}$ seem physically reasonable. Since $L = 4.5$ cm in ACME II, we expect to observe $\frac{\partial \mathcal{E}^{\mathrm{nr}}}{\partial z}[\mathrm{mV/cm}^2] \sim (3$–$10) \times \frac{\mathcal{E}^{nr}[\mathrm{mV/cm}]}{L[\mathrm{cm}]} \sim (1$–$2) \times \mathcal{E}^{\mathrm{nr}}[\mathrm{mV/cm}]$. In other words, by numerical coincidence, gradients $\partial \mathcal{E}^{\mathrm{nr}}/\partial z$ measured in units of $\mathrm{mV/cm}^2$ are expected to be comparable to (or a factor of ~ 2 larger than) fields $\mathcal{E}^{\mathrm{nr}}$ measured in mV/cm, just as observed.

5.2 Magnetic field characterization

The magnetic field control in ACME II consists of two core parts: shielding that nulls fields from outside of the apparatus, and sets of coils that apply fields in the interaction region volume. We will consider each of these parts in turn.

5.2.1 Ambient magnetic field reduction

The magnetic shielding consists of five layers of mu-metal magnetic shielding, in a cylindrical geometry. Each cylinder is divided into four pieces, consisting of two flat "endcaps" and two open half-cylinders that partially nest within each other. Holes are cut out of the mu-metal where lasers or light pipes for fluorescence collection must pass between the interaction region and the outside of the interaction region chamber. The cylinders span lengths from 34"–52" (inner to outer), and diameters from 30"–34". The mu-metal shielding is identical

to that used in ACME I, aside from some slight modifications to the holes required for light pipes and the addition of holes that allow laser beams to propagate vertically through the interaction region for STIRAP. Details of the mu-metal shielding design can be found in [144, Sec. 5.4]. Emil Kirilov has measured the attenuation of external fields to be $\sim 10^5$. In other words, external fields of order 1 G are attenuated to $\sim 10\,\mu$G inside the mu-metal shields. As we will see, we are limited to significantly larger ambient magnetic fields than this due to "stuck" magnetic domains in the innermost layer of the mu-metal shield, rather than leakage of the external field through the five layers of mu-metal.

In addition to the mu-metal shielding, in ACME II we installed room-scale passive magnetic field cancellation coils to null the external field in the vicinity of the interaction region, which significantly relaxes the performance requirements of the mu-metal shielding. We refer to these as the "room coils." Although the room coils are not necessary at the ACME II sensitivity, we treat the implementation of this system as a proof-of-principle demonstration that we can use some combination of external coils and mu-metal to suppress external fields, rather than rely on mu-metal alone. In particular, in a future EDM measurement, we may determine that it is technically preferable to use only three or four layers of mu-metal, in which case nulling the fields external to the mu-metal would be more critical. This project was developed by Zack Soule and Aaron Markowitz, undergraduates who worked for a short time in the group, with significant assistance from Elizabeth Petrik West and Adam West. A summary of the design and results are available in the internal ACME document [202]. Briefly, using six magnetic field coils spanning approximately the ceiling, floor, and each wall of the laboratory, with $\sim 100 - 1000$ amp-turns each, we can null the magnetic field in the center of the lab to ~ 1 mG, significantly smaller than earth's field. Unfortunately, the region of interest is ~ 1 m^3, defined by the size of the outermost mu-metal shield. Over approximately this volume, using four three-axis GMR magnetometers, the optimized field cancellation suppresses typical magnetic field component values to $|\mathcal{B}_i| \lesssim 20$ mG, about a factor of five smaller than when the room coils are off. We confirmed that, as completely expected, whether the room coils are on or off has no effect on the value of the EDM (though we did not perform this check at the statistical sensitivity of the EDM measurement).
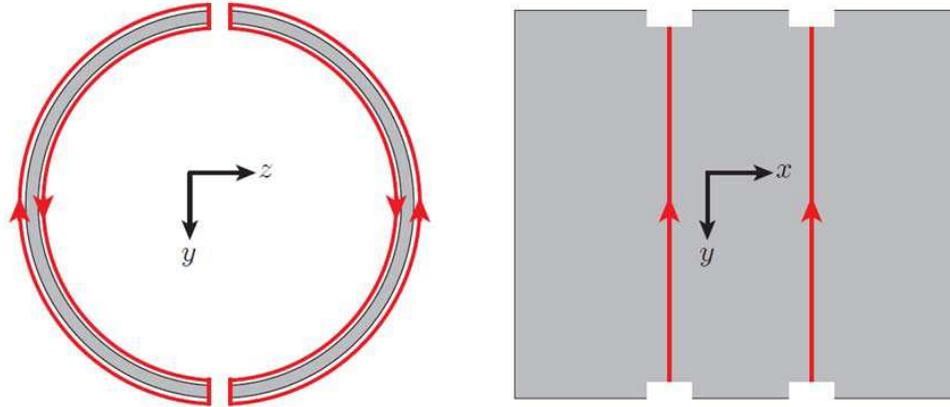
Figure 5.4: Geometry of the degaussing ribbon cables. Arrows denote direction of current flow. Figure by Adam West.

5.2.2 Degaussing

Because mu-metal is a ferromagnetic material, it is critical to ensure that the magnetization of the mu-metal does not create a large magnetic field inside the interaction region, negating its shielding functionality. The standard solution to this problem is to "degauss" the magnetic shields by driving an oscillating magnetic field through them such that the magnetization follows the magnetic field along a hysteresis loop. As the strength of the applied magnetic field is slowly ramped down, the corresponding magnetization likewise decreases. The degaussing setup in ACME II was designed by Brendon O'Leary, and implemented by myself and Adam West.

Geometry

For the degaussing geometry, see Fig. 5.4. In detail, each hemicylindrical mu-metal shield is wrapped with a pair of connected ribbon cables to supply a degaussing field. Each hemicylinder has two Molex Mini Fit Jr. connectors, corresponding to IN and OUT signals. IN connects to one wire of a ribbon cable, which then wraps around the shield. The ribbon wraps around two full times, and is then connected back to itself such that the IN wire makes electrical contact with the wire adjacent to it. This geometry repeats so that each wire connects to its succeeding wire after two loops around the shield. Every ribbon cable has 25 wires, so each ribbon accounts for 50 windings.

The last wire in this ribbon then connects to a second ribbon wrapped around the shield hemicylinder in the same configuration, but offset in the x-direction. There are thus 100 windings around each shield hemicylinder. The last wire from the second ribbon on a given hemicylinder connects to the first wire on the first ribbon of the opposing hemicylinder. The configuration of the second hemicylinder is identical, again with 100 total windings. All 200 windings reinforce the magnetic field driven inside the mu-metal, which circulates around the magnetic shields along $\pm\hat{x}$ (in the hemicylinders) or $\pm\hat{z}$ (in the endcaps).

Adjacent wires within each ribbon cable were initially connected using a standard ribbon connector, but before the EDM data set we found that a connector on the second-innermost shield had broken due to being squeezed between adjacent shields. As a result, we replaced all connectors with solder joints. This has the potential disadvantage of not allowing us to simply remove the cables from the shields while they're assembled, but under normal circumstances doing so is not necessary. Distinct ribbon cables are connected using Molex Mini Fit Jr. connectors.

We designate the innermost shield as "shield 1," and the outermost shield as "shield 5." The handedness of the circulating current reverses from one shield to the next. This ensures, for example, that the magnetic field produced inside shield 2 has contributions in the same direction from the current flowing through the degaussing wires around shields 1, 2, and 3.

Degaussing pulse

Since the hemicylinders are of unequal diameter, the resulting resistance and inductance of the degaussing coils increase from inside to outside. As a result, the same voltage or current inputs will not magnetize each shield to exactly the same degree. In addition, referring to the known resistance and inductance for each coil is useful for trouble-shooting the assembly and electrical connections of the degaussing system. The coils have resistances in the range 103–143 Ω and inductances in the range 63–91 mH.

The degaussing waveform, shown in Fig. 5.5, typically lasts one second and consists of 100 Hz oscillations modulated by a rapidly-rising and slowly-decaying envelope function, with a peak typically at 6.2 V. The waveform is generated by a DAQ analog output and
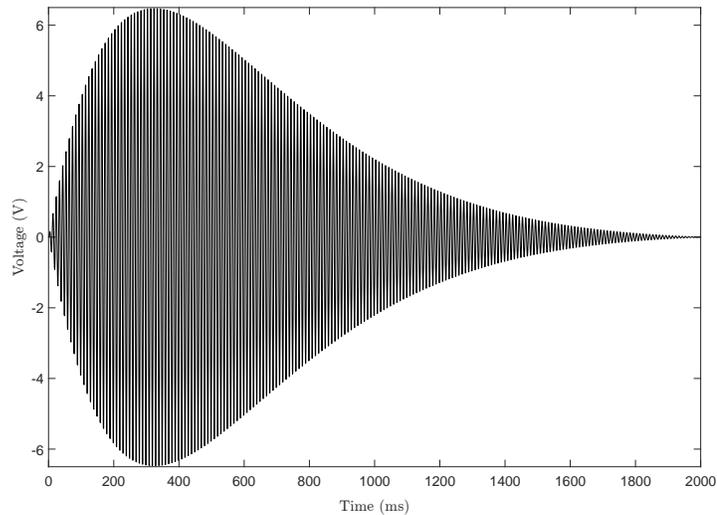
Figure 5.5: A waveform used for degaussing, shown here with two-second duration, 100 Hz oscillations, and 6.2 V maximum input to the current amplifiers.

is distributed to five separate current amplifiers. Four of these amplifiers are Kepco BOP 100-1M models, while one is the Kepco BOP 72-6M model, with maximum voltages of 100 and 72 V, and currents of 1 and 6 A, respectively. Due to the impedance of the degaussing coils, for every supply the voltage limit is reached at corresponding currents below the current limit. Since the outer-most shield is the least important for the experiment, we degauss it with the lowest-voltage supply, namely the BOP 72-6M. These current amplifiers generate 100 mA/V input. The current output is therefore typically 620 mA, contributing a maximum of 124 amp-turns of degaussing current to each shield.[2]

These currents are fed into a relay array designed by Brendon O'Leary, which allows us to completely disconnect the degaussing cables from the current supplies when no current should be applied. In normal operation, we open and close all relays simultaneously. The relay outputs are filtered by a bandpass circuit, which passes frequency components in the range of approximately 10–1300 Hz (with exact values depending on the specific coil impedance), to suppress DC currents in the degaussing coils and high-frequency noise from

2. To the best of my knowledge, the number of windings in the ACME I degaussing geometry is not clearly documented, but I believe it only consisted of a single loop of a ribbon cable, or 25 turns. With the typical degaussing current of ≈ 1 A, this corresponds to only 25 amp-turns, while the "hard degauss" used after the conclusion of the ACME I EDM data set, to remove an ambient $\mathcal{B}_y$ magnetic field, only contributed ≈ 2.5 A, for a total of ≈ 60 amp-turns.

the relay switching.

During initial tests of the degaussing system, we found that the ambient magnetization in the interaction region was consistently increased to $\mathcal{B}_z \sim 100\,\mu\mathrm{G}$ by the application of a 26 mG uniform magnetic field along $\hat{z}$, and reduced by the degaussing pulse just described to $\mathcal{B}_z < 40\,\mu\mathrm{G}$. The efficacy of the degaussing was quite insensitive to the number of degaussing pulses sequentially applied, the direction of the applied magnetic field, and the current amplitude of the degaussing pulse within a reasonable range. In particular, degaussing has no observable effect with peak degaussing currents below 1 mA, and reduces fields to a consistent level below 40 μG, provided the peak current exceeds 20 mA. The 620 mA currents typically used greatly exceed this threshold. Degaussing is always performed in the absence of any applied magnetic fields.

We also explored what effect the shape of the degaussing pulse has on the degaussing efficacy. There was no clear dependence of the ambient field on pulse duration beyond approximately 50 ms, corresponding to five 100 Hz current oscillations. Under normal experimental conditions, we only degauss when the magnetic field is switched, which requires several seconds of settling time. Therefore, we typically use a one-second-long degaussing pulse, which far exceeds the threshold required for degaussing to be effective but is still a relatively small fraction of the time required to perform a $\tilde{\mathcal{B}}$ switch. Furthermore, we did not observe any dependence of the ambient field on the oscillation frequency, from $\approx 50-200$ Hz. In principle, lower frequencies more efficiently permeate the mu-metal shielding (a point we will, alas, have to return to), while higher frequencies allow for a larger number of complete oscillations to fit within a given pulse duration. The choice of 100 Hz current oscillation is an arbitrary compromise between these competing advantages, under the constraint of using pulse durations that do not exceed ≈ 1 second.

5.2.3 Magnetic field measurement protocol

The sets of magnetic field coils were designed by Amar Vutha and Brendon O'Leary and are shown in Fig. 5.6. The most important coils apply $\mathcal{B}_z$, optimized for uniformity between the refinement and probe volumes of the interaction region. All coils are made from resin-coated copper wire connectorized with Molex Mini Fit Jr. connectors, except for the x coils
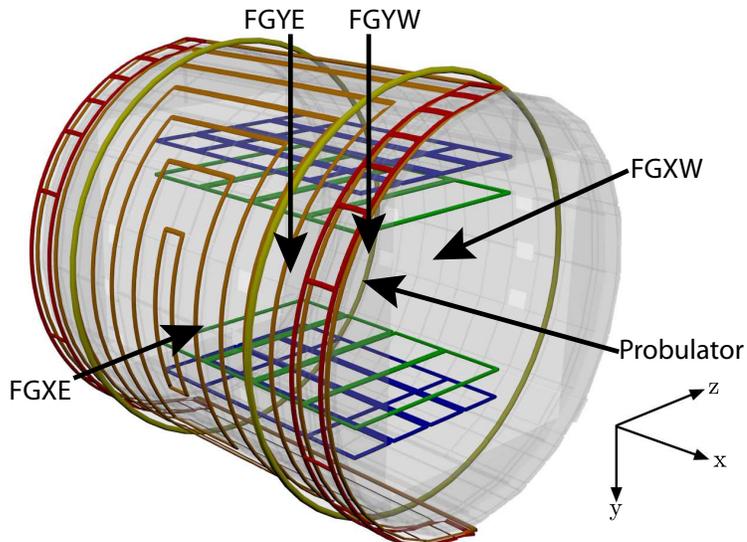
Figure 5.6: Approximate design of magnetic field coils. Green and blue "y" coils apply $\mathcal{B}_y$, $\partial\mathcal{B}_y/\partial x$, $\partial\mathcal{B}_y/\partial y$, and $\partial\mathcal{B}_y/\partial z$; yellow "$x$" coils apply $\mathcal{B}_x$ and $\partial\mathcal{B}_x/\partial x$; orange "main" coils apply $\mathcal{B}_z$ and $\partial\mathcal{B}_z/\partial x$. Red "side" coils create flatter profile of $\mathcal{B}_z$ near the refinement and probe regions, and create $\partial\mathcal{B}_z/\partial z$. All coils shown only around one hemicylinder are mirrored by symmetric coils on the opposing hemicylinder but are omitted here for clarity. Five possible positions of magnetometers are shown, with translation and rotation possible around the axes denoted by arrows.

(yellow in Fig. 5.6), which are ribbon cables connected with Molex Micro Fit connectors.

To calibrate the applied magnetic field, we perform "probulation." For an overview of this process in ACME I, see [149, Sec. 6.3]; the procedure in ACME II is essentially identical but will be described here for completeness. We simultaneously insert Bartington MAG-03 three-axis fluxgate magnetometers into the interaction region in three locations: one along the molecular beam path, and two vertically along $\hat{y}$, offset by 5.5 cm along $-\hat{x}$ and by 8.9 cm along $\pm\hat{z}$[3]. We refer to the magnetometer along the beam line as the "probulator" and the magnetometers along $\hat{y}$ as FGYW(E), for "fluxgate Y west (east)." Thus FGYW is at $z = +8.9$ cm and FGYE is at $z = -8.9$ cm, where here we take the origin, $x = y = z = 0$, to be the position along the molecular beam path centered between the preparation and probe lasers.

In addition to these magnetometer locations, we can insert magnetometers into pockets in the vacuum chamber at $x = y = 0$ and $z \approx \pm20$ cm. We refer to magnetometers in

3. The vertical magnetometers were designed with an offset along $\hat{x}$ so that the horizontal magnetometers could be brought as close as possible to the field plates in the $x = 0$ horizontal position.

	x [cm]	y [cm]	z [cm]	Rotation axis	Fields measured
Probulator	± 12	0	0	x	$\mathcal{B}_y$, $\mathcal{B}_z$
FGYW	-5.5	± 7	+8.9	y	$\mathcal{B}_x$, $\mathcal{B}_z$
FGYE	-5.5	± 7	-8.0	y	$\mathcal{B}_x$, $\mathcal{B}_z$
FGXW	0	0	+20	z	$\mathcal{B}_x$, $\mathcal{B}_y$
FGXE	0	0	-20	z	$\mathcal{B}_x$, $\mathcal{B}_y$

Table 5.1: Positions of all magnetometers, relative to the center of the interaction region. The rotation axes (identical to the translation axes) of each magnetometer, and the corresponding components of the magnetic field that can be measured without an electronic offset, are also indicated.

these locations as FGXW(E) in analogy to FGYW(E) because they can be translated along a horizontal, rather than vertical, axis.[4] However, the FGXW(E) magnetometers are not used during probulation. The positions of all five possible magnetometer positions are given in Table 5.1, and represented in Fig. 5.6.

The probulator apparatus was designed by Elizabeth Petrik West, built by Brendon O'Leary, and implemented in ACME II primarily by Adam West. It is mounted to an automated Velmex translation stage and scans ± 12 cm along the molecular beam line. Inside the interaction region, it is constrained to move snugly along a tube. The three fluxgates within the probulator (each measuring the field along one cartesian direction) are housed in a 5.4 cm long cylinder and separated from each other by 1.5 cm, with the fluxgate that lies along the cylindrical symmetry axis in the center. The FGY magnetometers are both mounted to a single automated Zaber translation stage above the interaction region and scan ± 7 cm vertically. All three fluxgates are also mounted on Zaber T-NM17A04-S rotation stages, with unidirectional repeatability of $0.1° \approx 2$ mrad and step accuracy of $0.25° \approx 4$ mrad.

The fluxgates are expected to be centered on the cylindrical symmetry axis within 2 mm. Due to the non-zero clearance of the fluxgate housings in their constraining tube, we expect that the housings can be oriented at angles up to ~ 5 mrad away from their nominal axes. In addition, the fluxgate axes are specified to be orthogonal only to $0.5° \approx 9$

---

4. This notation may be confusing: while the FGY magnetometers translate–sensibly enough–along $\hat{y}$, the FGX magnetometers translate along $\hat{z}$. The "X" is supposed to suggest "horizontal." Because all of the internal ACME documentation uses this unfortunate naming convention, to avoid even more confusion I use the same names throughout this chapter.

mrad. All of these imperfections can contribute to an offset or scatter in magnetic field measurements, as we consider at length in Sec. 5.2.6.

Each magnetometer can be rotated, via the Zaber rotation stages, around one axis. We define $\tilde{\mathcal{F}} = +1$ to be the initial state of the magnetometers, and $\tilde{\mathcal{F}} = -1$ to be the state with all magnetometers rotated about one of their axes by $180°$. This rotation allows us to identify electronic offsets in the fluxgate magnetometer measurements. Using the usual parity sum notation, the electronic offsets appear in the channel $\mathcal{B}_i^{\tilde{\mathcal{F}}}$, where $\mathcal{B}_i(\tilde{\mathcal{F}})$ represents the nominal magnetic field along a cartesian direction, $\hat{i} \in \{\hat{x}, \hat{y}, \hat{z}\}$, measured in a given $\tilde{\mathcal{F}}$ state. For example, the probulator is rotated about the x-axis, and therefore the electronic offsets in $\mathcal{B}_y$ and $\mathcal{B}_z$ can be distinguished from physical magnetic fields. Note that, in this case, the electronic offset in the $\mathcal{B}_x$ magnetometer channel cannot be distinguished from physical fields. For a list of the offset-free magnetic fields that can be measured in each magnetometer position, refer to Table 5.1.

Furthermore, we distinguish applied and ambient fields via the application of a nominal field along opposite directions, $\tilde{\mathcal{B}} = \pm 1$. In this chapter only, we do not only consider the $\tilde{\mathcal{B}}$ switch to denote the orientation of the field along $\hat{z}$, but instead define $\tilde{\mathcal{B}}$ to denote the direction of current applied to any magnetic field coils of interest. For example, we calibrate the "y" magnetic field coils by applying $\vec{\mathcal{B}}(\tilde{\mathcal{B}}) = \pm|\mathcal{B}_y|\tilde{\mathcal{B}}\hat{y}$.

More concretely, we decompose each field component into its parity components separately for each axis: $\mathcal{B}_i = \mathcal{B}_i^{\mathrm{nr}} + \tilde{\mathcal{B}}\mathcal{B}_i^{\tilde{\mathcal{B}}} + \tilde{\mathcal{F}}\mathcal{B}_i^{\tilde{\mathcal{F}}} + \tilde{\mathcal{B}}\tilde{\mathcal{F}}\mathcal{B}_i^{\tilde{\mathcal{B}}\tilde{\mathcal{F}}}$. Broadly speaking, we interpret these four components as follows:

- $\mathcal{B}_i^{\mathrm{nr}}$ is the magnetic field recorded by a fluxgate channel that does not reverse with rotation or the applied field direction. This is dominated by an electronic offset.

- $\mathcal{B}_i^{\tilde{\mathcal{B}}}$ is the magnetic field recorded by a fluxgate channel that reverses with the applied field direction but not a nominal rotation. This is usually dominated by imperfections in the applied rotation but could also have contributions from magnetic field gradients that couple to translation of the fluxgate as the fluxgate housing rotates. In other words, only systematic errors appear in this channel.

- $\mathcal{B}_i^{\tilde{\mathcal{F}}}$ is the magnetic field recorded by a fluxgate channel that reverses with rotation

but not a nominal applied field. This is dominated by the ambient magnetic field in the interaction region caused by, e.g., magnetization of the mu-metal shields.

- $\mathcal{B}_i^{\tilde{\mathcal{B}}\tilde{\mathcal{F}}}$ is the magnetic field recorded by a fluxgate channel that reverses with rotation and with a nominal applied field. This is dominated by the applied field.

For the fluxgate channel along the cylindrical symmetry axis of a magnetometer, the components $\mathcal{B}_i^{\mathrm{nr}}$ and $\mathcal{B}_i^{\tilde{\mathcal{F}}}$ cannot be disentangled, and likewise for $\mathcal{B}_i^{\tilde{\mathcal{B}}}$, $\mathcal{B}_i^{\tilde{\mathcal{B}}\tilde{\mathcal{F}}}$, since the fluxgate rotation should have no physical effect on the field measured along the rotation axis. A detailed model of the result of imperfections on these parity components is given in Sec. 5.2.6.

In practice, we often rotate the fluxgate magnetometers by 90° instead of 180°. Although the extra orientation states are not necessary to extract any quantities of interest, this affords us the ability to cross-check results obtained via distinct fluxgates in each three-axis magnetometer. For example, the fluxgate in the probulator that is oriented along $+\hat{z}$ in an initial state becomes oriented along $-\hat{y}$ when the device is rotated about $\hat{x}$ by 90°, and can thus measure both $\mathcal{B}_z$ and $\mathcal{B}_y$. We find that the two redundant measurements of the applied and ambient fields along each cartesian direction consistently agree, as expected.[5] Therefore, we report values averaged over the two redundant fluxgates in these cases.

By repeating measurements in both $\tilde{\mathcal{B}} = \pm 1$ and all rotation states independently at many spatial positions ($x \in [-12, +12]$ cm for the probulator and $y \in [-7, +7]$ cm for the FGY magnetometers), we can map out the magnetic field over the regions of greatest interest in the interaction region. It would be excessive to examine here every parity component, axis, magnetometer location, and field configuration, but I will point out the most essential results.

5.2.4 Applied magnetic fields

Consider first the applied magnetic fields, which appear in the $\mathcal{B}^{\tilde{\mathcal{F}}\tilde{\mathcal{B}}}$ or $\mathcal{B}^{\tilde{\mathcal{B}}}$ parity channels. The magnetic field measured by the probulator along the molecular beam axis, when a

---

5. When magnetic field gradients are applied, the distance between the fluxgates needs to be accounted for carefully, however.
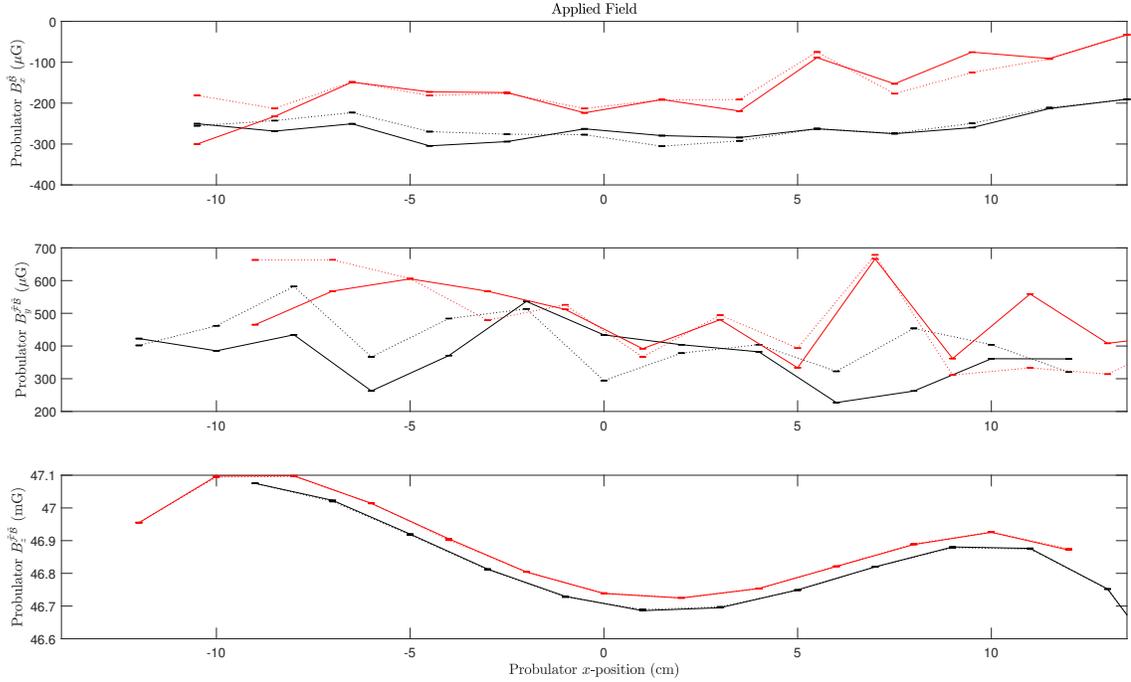
Figure 5.7: Applied magnetic field $\mathcal{B}_z$, measured by probulator in all three axes. Red and black curves correspond to distinct fluxgates (or simply distinct sets of rotation states, $0°$ and $180°$ vs. $90°$ and $270°$, in the case of $\mathcal{B}_x$). Solid and dashed correspond to separate sets of measurements, which agree to excellent precision.

nominal $\mathcal{B}_z$ is applied, is shown in Fig. 5.7. The ratio of current through the "side" and "main" coils, $I_{\text{side}}/I_{\text{main}} = 5.2$, is optimized for maximum uniformity of the field component of interest, $\mathcal{B}_z$. The relatively large offset component measured along $\mathcal{B}_y$ is consistent with the degree of non-orthogonality between fluxgate axes, as we will see in more detail in Sec. 5.2.6.

As an example of another magnetic field coil configuration that produces a less optimal field, consider the effect of an applied $\partial \mathcal{B}_y/\partial y$ as measured by FGYW, which translates along the y-axis, shown in Fig. 5.8. Here, there are significant structure and offset fields along $\mathcal{B}_x$ and $\mathcal{B}_z$, in addition to the desired applied gradient, $\partial \mathcal{B}_y/\partial y$. Note that this gradient can be measured to high precision by translating the magnetometer, despite the electronic offset in the $\hat{y}$-aligned fluxgate, which cannot be removed by rotation.

The applied magnetic fields and gradients arising from each nominal magnetic field coil configuration, per unit of applied current through the coils, is given in Table 5.2.
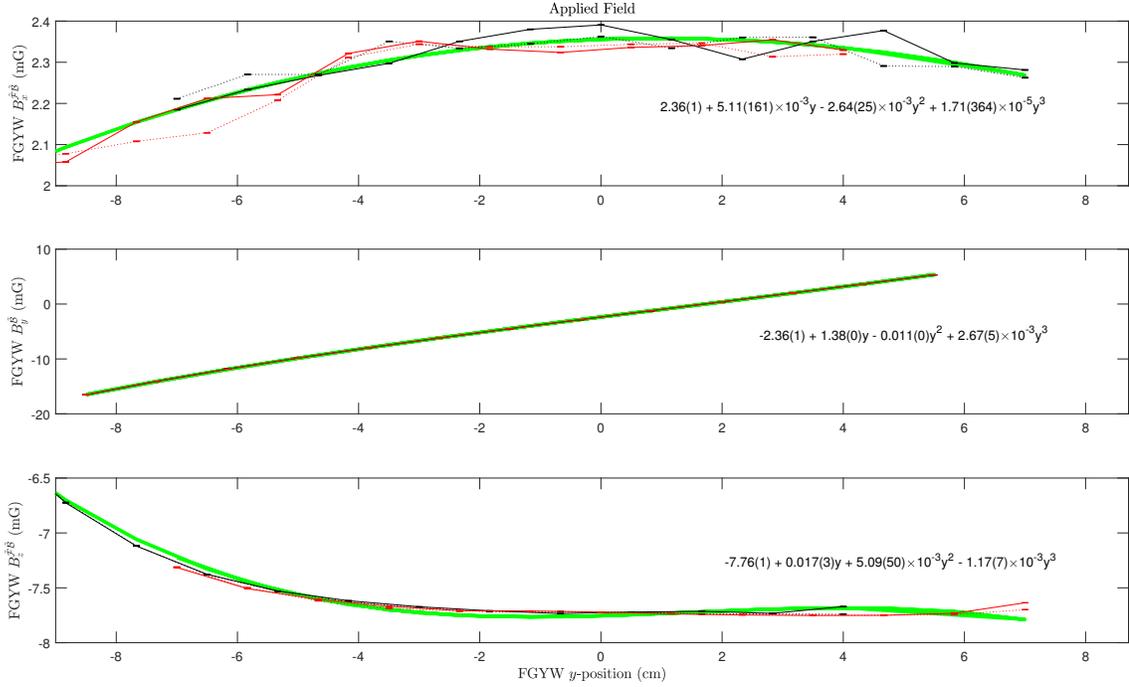
209

Figure 5.8: Applied magnetic field $\partial \mathcal{B}_y / \partial y$, measured by the vertical magnetometer FGYW. Due to the significant spatial dependence arising from these coils, cubic fits are shown (green).

Coil configuration	Field (gradient) [μG(/cm)/mA]
$\mathcal{B}_z$	1337
$\partial \mathcal{B}_z / \partial z$	16.1
$\partial \mathcal{B}_z / \partial y$	7.5
$\mathcal{B}_x$	714
$\mathcal{B}_y$	258
$\partial \mathcal{B}_y / \partial y$	14.1
$\partial \mathcal{B}_x / \partial x$	33.7
$\partial \mathcal{B}_y / \partial x$	14.7
$\partial \mathcal{B}_z / \partial x$	12.4

Table 5.2: Primary applied magnetic fields or gradients in each of nine nominal applied magnetic field configurations. Due to Maxwell's equations, the $\partial \mathcal{B}_y / \partial y$ coils also apply $\partial \mathcal{B}_z / \partial z = -8.5 \, (\mu \text{G/cm})/\text{mA}$, which we sometimes exploit when performing applied $\partial \mathcal{B}_z / \partial z$ systematic checks. All constant field configurations are measured using the probulator with large applied currents. (The applied field in $\mathcal{B}_x$ dominates the electronic offset, which cannot be removed by rotation of the probulator fluxgate in the calibration of the $\mathcal{B}_x$ coils.) Gradients along x are measured via translation of the probulator, and gradients along y are measured by averaging the gradients measured via translation of FGYW and FGYE. Gradients along z are measured by comparing the field measured in FGYW and FGYE and assuming a constant gradient over their 17.8 cm separation. Full magnetic field configurations can be found in the Data\B-fields\Probulation Data folder of the "ACME Storage" hard drive.

5.2.5 Ambient magnetic fields

The ambient magnetic fields, $\mathcal{B}^{\tilde{\mathcal{F}}}$ (for axes affected by rotation) or $\mathcal{B}^{\mathrm{nr}}$ (for axes unaffected by rotation), which appear in the ordinary EDM experiment as non-reversing magnetic field components $\mathcal{B}_i^{\mathrm{nr}}$, are of equal concern to us as the applied magnetic fields summarized in Table 5.2. Most especially, ambient magnetic field *gradients*, $\partial \mathcal{B}_z^{\mathrm{nr}}/\partial z$ and $\partial \mathcal{B}_z^{\mathrm{nr}}/\partial y$, can contribute to systematic errors as described in Sec. 4.10.

When we performed probulation prior to beginning our campaign to identify systematic errors in the experiment, ambient fields in the apparatus were consistently below $\approx 50\,\mu\mathrm{G}$ in all channels and locations (probulator, FGYW, and FGYE). Unfortunately, we later realized that the ambient magnetization of the inner mu-metal shields changed dramatically some time between January and November 2017, when we observed ambient magnetic fields as large as $\sim 300\,\mu\mathrm{G}$, as measured by FGYW(E) and FGXW(E). At this point, we had been regularly taking systematic error data and were unwilling to open the vacuum chamber to atmosphere as is required for inserting the probulation magnetometer. That the ambient magnetic fields could change so significantly was quite surprising and we undertook significant effort to reduce them back to the levels that were observed prior to taking any systematic error data.[6]

Several problems became apparent. First, we quickly discovered that a connector for the degaussing coils on shield 2 (second-innermost) had become disconnected, due to being squeezed between the second and third shield layers. Fixing this connector had no effect on the ambient fields, however.

We then attempted alternate methods of degaussing, partly inspired by the discovery in ACME I of an ambient magnetic field $\mathcal{B}_y$ that could only be removed by applying larger-than-typical degaussing currents. In particular, we performed longer-timescale degaussing (up to ten minutes), and degaussing with higher currents via a variable transformer (i.e., a Variac), as well as different deguassing sequences among the shields (e.g., 5-4-3-2-1-2-3-4-5 rather than degaussing all five shields simultaneously), all with no effect.

---

6. This was an all-hands-on-deck effort. Thanks to Cristian Panda, Xing Wu, Jonathan Haefner, Daniel Ang, and Cole Meisenhelder for their role in these investigations, and especially in disassembling and re-assembling the shields too many times to count.

We then undertook a thorough search for magnetic materials in the interaction region and discovered two culprits that had escaped our attention when assembling the magnetic shields. First, the ribbon cable connectors had become magnetized up to ~mG levels.[7] Second, we discovered that the mechanical supports for the magnetic shield hemicylinders (inherited from ACME I) used screws made from stainless steel rather than a non-magnetic material like brass. We replaced the offending screws with brass, and replaced the ribbon connectors with solder joints and non-magnetic Molex connectors. Unfortunately, these precautions did not significantly reduce the ambient magnetic fields.

By reassembling the magnetic shields with one or several shields removed, we determined by process of elimination that the ambient field originates from the innermost shield. We also installed ribbon cables around the endcaps of shield 1 in order to degauss them, under the hypothesis that the degaussing cables around the hemicylinders were not effective at degaussing the endcaps. This had no effect.

We are indebted to Blayne Heckel for useful discussions in which he suggested much lower-frequency degaussing than our usual 100 Hz waveform. We can understand this suggestion based on the skin depth δ of the magnetic field that we wish to drive through the mu-metal, which is merely $\delta \sim 200\,\mu\mathrm{m}/\sqrt{\omega/(2\pi \times 100\,\mathrm{Hz})}$ [203]. This is a factor of ≈ 8 smaller than the thickness of each shield layer (1/16"). Therefore, it is somewhat surprising that a 100 Hz degaussing waveform can effectively degauss the mu-metal shields at all. Degaussing, instead, with 5 Hz oscillations (for which the skin depth is ~ 1 mm) for a duration of ten minutes, we reduced the ambient magnetization to at most $\approx 150\,\mu\mathrm{G}$, still significantly larger than the ambient fields observed prior to taking any systematic error data. Unfortunately, we were unable to reduce the ambient magnetic fields throughout the interaction region any further than this.

Prior to taking the EDM data set, we measured the magnetic field gradients $\partial \mathcal{B}_z/\partial z$ and $\partial \mathcal{B}_z/\partial y$ using the *in situ* FGYW and FGYE magnetometers. By applying compensating gradients (in particular, $\partial \mathcal{B}_z/\partial z \approx 4\,\mu\mathrm{G/cm}$) using the magnetic field coils, we were able

---

7. This oversight was due to an inadequate protocol to screen magnetic materials: we always wave a material in front of a fluxgate magnetometer prior to inserting it into the interaction region, to make sure it is not magnetized. However, it is necessary to first attempt to magnetize the material with a strong permanent magnet. This initial step ensures that the material is non-magnetic, rather than simply unmagnetized.

to suppress both gradients of interest to at most 1 μG/cm, as measured on the fluxgate magnetometers. (We will consider the uncertainty on this value shortly.) These compensating fields were not adjusted at all throughout the EDM data set. The magnetic fields were re-measured using fluxgate magnetometers FGYW(E) and FGXW(E) twice daily during the period in which the EDM data set was acquired.

A more precise measurement of the magnetic fields that act upon the ThO molecules is obtained by probulation, since the magnetometer can translate exactly along the molecular beam axis. Of course, the initial probulation data could not be relied upon to indicate the magnetic field during the EDM data set, since the ambient fields had clearly become much larger in the interim, so we planned to acquire another probulation data set once the EDM data acquisition had been completed. Unfortunately, at approximately the time that we opened the vacuum chamber to probulate the magnetic fields, the FGY magnetometers in the interaction region clearly indicated a sudden change in the ambient magnetic fields.

Even though it was now impossible to measure the magnetic fields as they were during the EDM data set, we proceeded with the post-run probulation measurement. In order to obtain the maximum information about ambient field gradients, we took measurements with the probulator both centered on the molecular beam axis and translated vertically by $\approx$0.8 cm along $-\hat{y}$. Due to the tight clearance between the probulator guide tube and the electric field plates, we were unable to translate the probulator along $\pm\hat{z}$. After completing this measurement, we removed the probulation equipment and found that the ambient magnetic fields were, once again, changed.

We are now confident that this change was caused by bumping the interaction region structure with the table that the probulator equipment is mounted on. Vibrations caused by this slight collision propagated through the interaction region structure to the shields, randomly magnetizing certain domains in the mu-metal. These domains should be removable by annealing the shields, but we are unable to do that without shipping the shields to an outside facility. It is highly likely that the same kind of process accounts for every discrete change in the ambient fields, including the first one known to have occurred some time during 2017.

In order to bolster our confidence that both probulation data sets were likely good
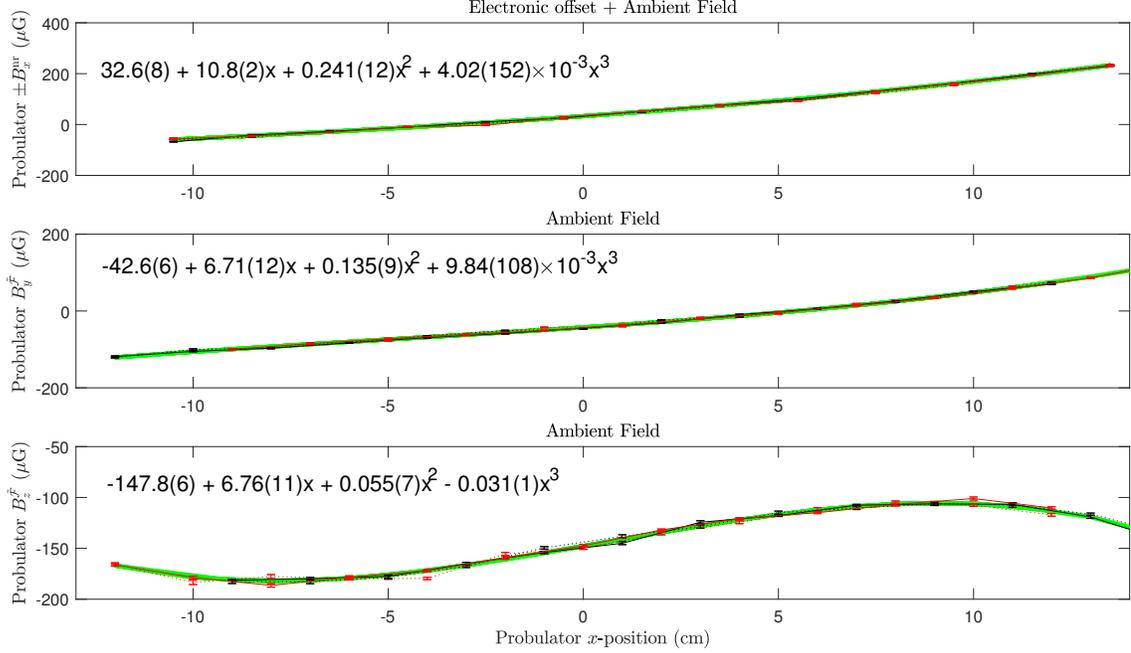
Figure 5.9: Ambient magnetic field measured by the probulator in the "vented" field configuration.

proxies for the ambient magnetic fields during the EDM data set, we performed probulation once again. We refer to the various magnetic field configurations for which probulation data is available as "initial" (prior to the EDM data set), "vented" (immediately after the EDM data set), and "jostled" (subsequent to completing the "vented" field measurements). Because the ambient fields measured in the initial configuration were significantly smaller than during run-time and both the vented and jostled configurations, we disregard it for purposes of constructing a conservative possible range of ambient magnetic fields throughout the EDM data set.

For an example of a probulator measurement of the ambient fields, see Fig. 5.9, which gives measurements from the probulator in the "vented" configuration. All measurements are qualitatively similar. Using cubic trend lines extracted from the data (which always provide an excellent fit), I summarize the data for the FGYW, FGYE, and probulator measurements in Figs. 5.10-5.11. We see that the run-time fields inferred from *in situ* magnetometers give reasonable agreement with the range of fields measured via direct probulation of the the vented and jostled configurations. This lends some degree of confidence that the *in situ* measurements are reasonably representative of the actual fields in the region of the

214
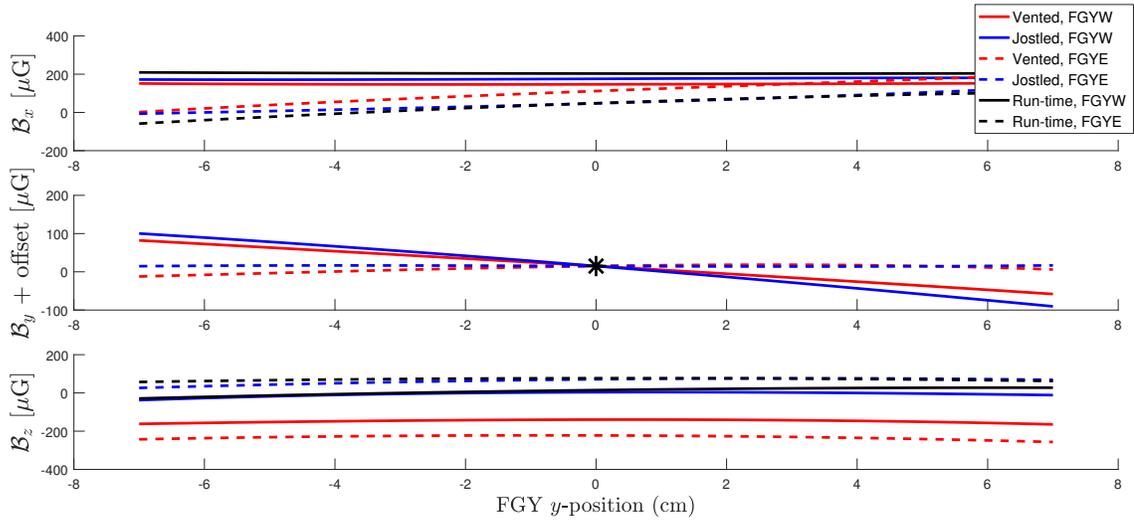
Figure 5.10: Measurements of magnetic fields using FGYW and FGYE during EDM data set ("run-time," black), and from probulation data in vented (red) and jostled (blue) configurations. All lines show cubic fits to ambient fields, which are excellent fits to the data, to avoid excessive clutter. Measurements with FGYW ($+z$ positions) are solid, those with FGYE ($-z$ positions) are dashed. Measurements of $\mathcal{B}_y$ were not recorded from FGYW and FGYE throughout the EDM data set due to the electronic offset, which cannot be removed via rotation of the fluxgate magnetometers. However, we obtain a run-time value of $\mathcal{B}_y$ from the FGXW and FGXE magnetometers ($\ast$). For purposes of the plot, we arbitrarily set the (possibly drifting) electronic offset for $\mathcal{B}_y$ measurements so that all lines cross the measured value of $\mathcal{B}_y$ throughout the EDM data set.
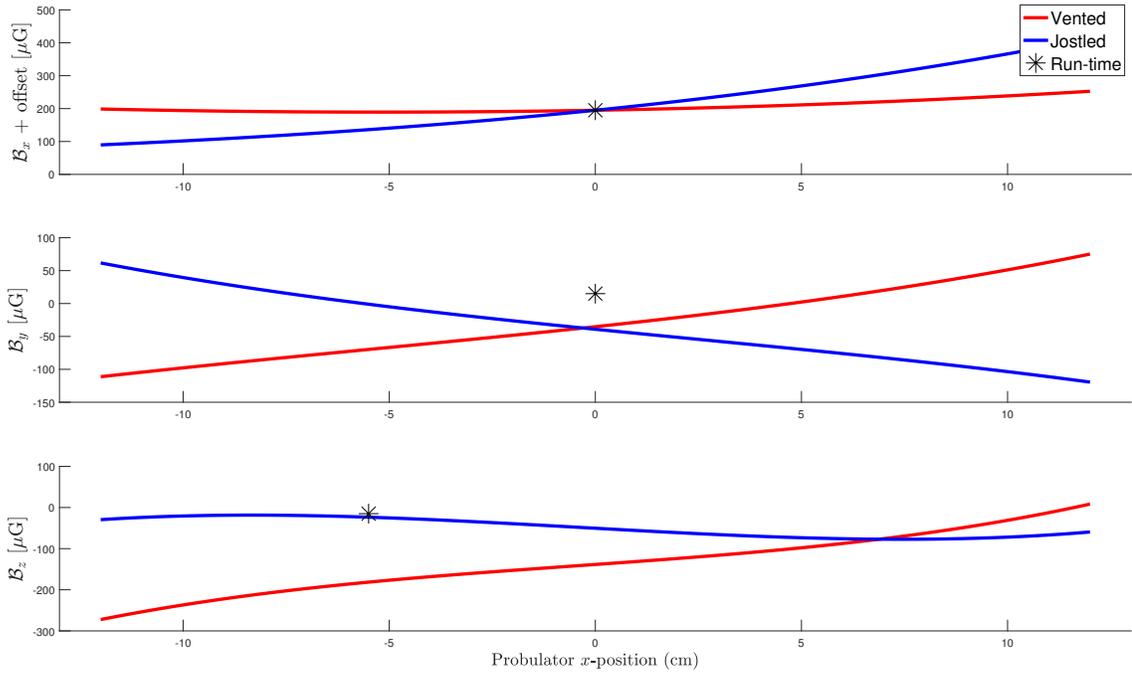
Figure 5.11: Measurements of magnetic fields using the probulator, for vented (red) and jostled (blue) configurations. Run-time fields, measured using *in situ* FGX (for $\mathcal{B}_x$ and $\mathcal{B}_y$) and FGY (for $\mathcal{B}_z$) magnetometers, are also shown ($*$). For purposes of the plot, I assume that the (possibly drifting) electronic offsets in $\mathcal{B}_x$ are such that both post-run configurations agree with the run-time fields at $x = 0$.

molecular beam thoughout the EDM data set.

Ambient magnetic field gradients for the EDM data set

I now describe in detail our method of determining the ambient $\partial \mathcal{B}_z/\partial z$ and $\partial \mathcal{B}_z/\partial y$ gradients present throughout the EDM data set, in order to set a careful bound on the systematic error contribution arising from magnetic field gradients.

Consider first the average value of $\partial \mathcal{B}_z/\partial y$. Using the FGY magnetometers scanned over ± 7 cm, during the period of the EDM data set, we measured $\partial \mathcal{B}_z/\partial y \approx 1\,\mu\text{G/cm}$. We can compare this with the value of $\partial \mathcal{B}_z/\partial y$ obtained by several alternative methods. In the "vented" ambient field configuration, we measured fields both along the center of the molecular beam line and offset by ≈ 0.8 cm vertically, along $-\hat{y}$. Comparing the results for $\mathcal{B}_z$ in these two sets of positions, we can infer $\partial \mathcal{B}_z/\partial y \approx 4(2)\,\mu\text{G/cm}$, where the uncertainty is the root-mean-square (rms) spread in the value of the extrapolated gradient. This result agrees quite well with the gradient extrapolated along the vertical scan of FGY magnetometers in this configuration, $\partial \mathcal{B}_z/\partial y \approx 4(2)\,\mu\text{G/cm}$, where once again the uncertainty is an r.m.s. spread in local gradient values (i.e., there is a slight curvature to the ambient field). This reinforces that the measurements inferred from FGY are consistent with the local field along the molecular beam axis, at least in the vented configuration.

We performed different measurements in the "jostled" configuration to further reinforce our confidence in the measured value of $\partial \mathcal{B}_z/\partial y$ obtained from the FGY magnetometers during the EDM data set. From Maxwell's equations, we know that $\partial \mathcal{B}_z/\partial y = \partial \mathcal{B}_y/\partial z$, so that by translating the FGX magnetometers along $\pm\hat{z}$ and measuring $\mathcal{B}_y$, we can infer $\partial \mathcal{B}_z/\partial y$. The FGX magnetometers do not have associated translation stages[8], so this measurement was performed by hand, which led to much larger angular fluctuations of the fluxgate orientation than is present for typical measurements. As a result, we were able to determine only that $\partial \mathcal{B}_z/\partial y \approx 6(5)\,\mu\text{G/cm}$, consistent with measurements inferred directly from FGY vertical scans, $\partial \mathcal{B}_z/\partial y = 5(2)\,\mu\text{G/cm}$ in this field configuration.

---

8. Our omission of translation stages for the FGX magnetometers is mainly due to space constraints on the optics tables next to the interaction region. However, if it were a priority to use horizontal magnetometer translation stages, then we could surely find a way to fit them in.

Now consider the average value of $\partial \mathcal{B}_z/\partial z$. Based on measurements with FGY magnetometers throughout the EDM data set, we infer $\partial \mathcal{B}_z/\partial z \approx -0.5\,\mu\mathrm{G/cm}$. (Recall that we apply a fixed compensating magnetic field gradient in order to achieve this low value.) In the vented field configuration, we can use $\nabla \cdot \vec{\mathcal{B}} = 0$ to infer $\partial \mathcal{B}_z/\partial z$ in the region of the molecular beam. In particular, we measure $\partial \mathcal{B}_x/\partial x$ directly using the ordinary probulator scan along $\hat{x}$, and we infer $\partial \mathcal{B}_y/\partial y$ by comparing the two sets of measurements in which the probulator is offset along $-\hat{y}$ by 0.8 cm. Then $\partial \mathcal{B}_z/\partial z = -\partial \mathcal{B}_x/\partial x - \partial \mathcal{B}_y/\partial y \approx 1(2)\,\mu\mathrm{G/cm}$. By comparison, the value inferred by comparing $\mathcal{B}_z$ measured with FGYW and FGYE in this field configuration is $\partial \mathcal{B}_z/\partial z \approx 0(2)\,\mu\mathrm{G/cm}$, consistent with the results obtained in the region of the molecular beam.

In the "jostled" field configuration, we obtained results for $\partial \mathcal{B}_z/\partial z \approx -3.5\,\mu\mathrm{G/cm}$ from the FGY magnetometers. We can compare this to results from the scans along $\pm \hat{z}$ of the FGX magnetometers to obtain a direct measurement of $\partial \mathcal{B}_z/\partial z \approx -3(2)\,\mu\mathrm{G/cm}$ at distances of $|z| \approx 20$ cm from the molecular beam axis, consistent with the results from the FGY magnetometers. However, there is another method that gives results slightly inconsistent with this value. In particular, we can measure $\partial \mathcal{B}_y/\partial y$ using vertical scans of the FGY magnetometers (albeit offset from the molecular beam axis) and $\partial \mathcal{B}_x/\partial x$ using the horizontal scan of the probulator. Then $\partial \mathcal{B}_z/\partial z = -\partial \mathcal{B}_x/\partial x - \partial \mathcal{B}_y/\partial y \approx -7(2)\,\mu\mathrm{G/cm}$, not clearly consistent with the other measurements. The discrepancy is no larger than $\approx 4\,\mu\mathrm{G/cm}$.

With all of these cross-checks in mind, we can estimate the uncertainty in the measured values of $\partial \mathcal{B}_z/\partial y$ and $\partial \mathcal{B}_z/\partial z$ that are inferred from FGY magnetometer measurements throughout the EDM data set. In almost every case, different methods of measuring these gradients give results consistent with those inferred from the FGY *in situ* magnetometers, and in one case the discrepancy is at most $4\,\mu\mathrm{G/cm}$. We therefore conservatively assume that the possible difference between a gradient measured via the FGY magnetometers and the actual gradient along the molecular beam axis is at most $4\,\mu\mathrm{G/cm}$ in either component, $\partial \mathcal{B}_z/\partial y$ or $\partial \mathcal{B}_z/\partial z$, independently. Furthermore, in any magnetic field scan, the largest variation in these gradients between the center of the interaction region and either the preparation or probe regions is $6\,\mu\mathrm{G/cm}$, due to quadratic components of the magnetic

field in some cases. We therefore take the possible variation in magnetic field gradients, $\partial \mathcal{B}_z/\partial z + R\partial \mathcal{B}_z/\partial y$, where $R \approx 1.1$ (see Sec. 4.10), to be $\approx 6\,\mu\text{G}/\text{cm}$. We treat both contributions to the uncertainty in the gradient–i.e., a discrepancy between the *in situ* measurements vs. alternative methods to measure the magnetic field, and variations in the magnetic field over the interaction region volume–as uncorrelated and so add their contributions in quadrature, $\sqrt{2 \times 4^2 + 6^2}\,\mu\text{G}/\text{cm} < 9\,\mu\text{G}/\text{cm}$.

Using the central values of the magnetic field gradients measured throughout the EDM data set, we therefore determine $\Delta(\partial \mathcal{B}_z/\partial z + R\partial \mathcal{B}_z/\partial y) = -1(9)\,\mu\text{G}/\text{cm}$, with the uncertainty set in approximately equal parts by the possible variation in the magnetic field gradient over the interaction region ($6\,\mu\text{G}/\text{cm}$) and the possible error in the FGY magnetometer measurements (independently in each gradient component, for an overall contribution of approximately $\sqrt{2} \times 4\,\mu\text{G}/\text{cm}$).

5.2.6 Model of probulation systematic errors

Here, I will document all known contributions to systematic errors in the magnetometer measurements. The $\mathcal{B}^{\tilde{\mathcal{B}}}$ fluxgate magnetometer channels (independent of fluxgate rotation state $\tilde{\mathcal{F}}$), which consists entirely of systematic errors, is useful in understanding various contributions to both measurement scatter and offsets.

Rotating fluxgate channels

We will calculate the measured field in each rotation state, $\tilde{F} = \pm 1$, for a fluxgate pointing along unit vector $\hat{e}_1$ when the magnetic field has components $\mathcal{B}_1$, $\mathcal{B}_2$ and $\mathcal{B}_3$, as well as first- and second-order gradients. (We use arbitrary unit vectors, rather than $\{\hat{x}, \hat{y}, \hat{z}\}$, because the orientation of the fluxgates in the lab are not all identical.) Let the rotation axis be $\hat{e}_3$. We will allow the fluxgate to be misaligned by small angles ϕ toward $\hat{e}_2$ and β toward $\hat{e}_3$. The fluxgate may be displaced by distance L from the rotation axis $\hat{e}_3$ at an angle α with respect to the $\hat{e}_1$ axis in the $\tilde{F} = +1$ state. Finally, a nominal $180°$ rotation may correspond to an actual rotation of $\pi + \delta$, where δ is some small angle (over- or under-shoot of $180°$ rotation). We will ignore any terms that are bound to be smaller than $\sim 1\,\mu\text{G}$.

We will first consider the measured fields in one of the two fluxgate axes that are ideally

affected by the $\tilde{\mathcal{F}}$ switch, i.e., one of the "rotating" fluxgate channels. In particular, we consider the component measured nominally along $\hat{e}_1$ for clarity. It is easiest to build up this model step-by-step. First, note that an ideal measurement would give

$$[\text{ideal}]$$
$$\mathcal{B}(\tilde{\mathcal{F}} = +1) \quad = \quad \mathcal{B}_1$$

$$\mathcal{B}(\tilde{\mathcal{F}} = -1) \quad = \quad -\mathcal{B}_1.$$

(5.1)

Now suppose that a rotation occurs with angle $\pi + \delta$ such that the angle θ for a fluxgate of interest, relative to its nominal axis, is $\theta(\tilde{\mathcal{F}} = +1) = -\delta/2$ and $\theta(\tilde{\mathcal{F}} = -1) = \pi + \delta/2$. Then

$$[+\text{rotation error}]$$
$$\mathcal{B}(\tilde{\mathcal{F}} = +1) \quad = \quad \mathcal{B}_1(1 - \tfrac{\delta^2}{8}) - \mathcal{B}_2\tfrac{\delta}{2}$$

$$\mathcal{B}(\tilde{\mathcal{F}} = -1) \quad = \quad -\mathcal{B}_1(1 - \tfrac{\delta^2}{8}) - \mathcal{B}_2\tfrac{\delta}{2}.$$

(5.2)

Here, I neglect terms of order $\mathcal{B}_i\delta^3$. Applied fields are typically of order 50 mG or smaller, while small angles are of order 10 mrad, so $\mathcal{B}\delta \sim 500\,\mu\text{G}$, $\mathcal{B}\delta^2 \sim 5\,\mu\text{G}$, and $\mathcal{B}\delta^3 \sim 0.05\,\mu\text{G}$, with the last term utterly negligible. For the same reasons, I will drop all terms that are cubic in any small angle going forward.

We can now consider that a fluxgate lying primarily along $\hat{e}_1$ actually has a small component along $\hat{e}_3$ due to a small rotation angle β. Then the measurement will be such that $\mathcal{B}_{1,2} \to \mathcal{B}_{1,2}(1 - \tfrac{\beta^2}{2})$ and an additional rotation-independent component $\mathcal{B}_3\beta$ is measured. At the relevant order, we have

$$[+\text{longitudinal component}]$$
$$\mathcal{B}(\tilde{\mathcal{F}} = +1) \quad = \quad \mathcal{B}_1(1 - \tfrac{\delta^2}{8} - \tfrac{\beta^2}{2}) - \mathcal{B}_2\tfrac{\delta}{2} + \mathcal{B}_3\beta$$

$$\mathcal{B}(\tilde{\mathcal{F}} = -1) \quad = \quad -\mathcal{B}_1(1 - \tfrac{\delta^2}{8} - \tfrac{\beta^2}{2}) - \mathcal{B}_2\tfrac{\delta}{2} + \mathcal{B}_3\beta.$$

(5.3)

We should also consider an electronic offset, which is naturally rotation-independent:

$$[+\text{electronic offset}]$$

$$\mathcal{B}(\tilde{\mathcal{F}} = +1) \;\; = \;\; \mathcal{B}_1(1 - \tfrac{\delta^2}{8} - \tfrac{\beta^2}{2}) - \mathcal{B}_2\tfrac{\delta}{2} + \mathcal{B}_3\beta$$

$$+ \mathcal{B}_{\text{offset}} \tag{5.4}$$

$$\mathcal{B}(\tilde{\mathcal{F}} = -1) \;\; = \;\; -\mathcal{B}_1(1 - \tfrac{\delta^2}{8} - \tfrac{\beta^2}{2}) - \mathcal{B}_2\tfrac{\delta}{2} + \mathcal{B}_3\beta$$

$$+ \mathcal{B}_{\text{offset}}.$$

If the fluxgate is offset by L from the rotation axis in a direction that makes angle α with respect to $\hat{e}_1$ in the e_1e_2-plane, then gradients will couple to its rotation-dependent position, $x_1 = \tilde{\mathcal{F}}L\cos\alpha$ and $x_2 = \tilde{\mathcal{F}}L\sin\alpha$. The largest gradients we apply are of order $\frac{\partial\mathcal{B}_i}{\partial x_j} \sim 5$ mG/cm, while the largest displacement we expect is ~ 2 mm. Therefore, we expect $\frac{\partial\mathcal{B}_i}{\partial x_j}L \sim 1$ mG at most, and $\frac{\partial\mathcal{B}_i}{\partial x_j}\theta \sim 10\,\mu$G, while $\frac{\partial\mathcal{B}_i}{\partial x_j}L\theta^2 \sim 0.1\,\mu$G, with θ here denoting a small angle (δ or α). We therefore ignore terms involving gradients that are second-order in small angles:

$$[+\text{gradients}]$$

$$\mathcal{B}(\tilde{\mathcal{F}} = +1) \;\; = \;\; \mathcal{B}_1(1 - \tfrac{\delta^2}{8} - \tfrac{\beta^2}{2}) - \mathcal{B}_2\tfrac{\delta}{2} + \mathcal{B}_3\beta$$

$$+ \mathcal{B}_{\text{offset}}$$

$$+ \tfrac{\partial\mathcal{B}_1}{\partial x_1}L\cos\alpha - \tfrac{\delta}{2}\tfrac{\partial\mathcal{B}_2}{\partial x_1}L\cos\alpha + \beta\tfrac{\partial\mathcal{B}_3}{\partial x_1}L\cos\alpha$$

$$+ \tfrac{\partial\mathcal{B}_1}{\partial x_2}L\sin\alpha - \tfrac{\delta}{2}\tfrac{\partial\mathcal{B}_2}{\partial x_2}L\sin\alpha + \beta\tfrac{\partial\mathcal{B}_3}{\partial x_2}L\sin\alpha \tag{5.5}$$

$$\mathcal{B}(\tilde{\mathcal{F}} = -1) \;\; = \;\; -\mathcal{B}_1(1 - \tfrac{\delta^2}{8} - \tfrac{\beta^2}{2}) - \mathcal{B}_2\tfrac{\delta}{2} + \mathcal{B}_3\beta$$

$$+ \mathcal{B}_{\text{offset}}$$

$$+ \tfrac{\partial\mathcal{B}_1}{\partial x_1}L\cos\alpha + \tfrac{\delta}{2}\tfrac{\partial\mathcal{B}_2}{\partial x_1}L\cos\alpha - \beta\tfrac{\partial\mathcal{B}_3}{\partial x_1}L\cos\alpha$$

$$+ \tfrac{\partial\mathcal{B}_1}{\partial x_2}L\sin\alpha + \tfrac{\delta}{2}\tfrac{\partial\mathcal{B}_2}{\partial x_2}L\sin\alpha - \beta\tfrac{\partial\mathcal{B}_3}{\partial x_2}L\sin\alpha.$$

We will also consider higher-order gradients. The largest curvatures we produce are of order $\sim 50\,\mu$G/cm^2, so we might expect corrections of $\sim 2\,\mu$G due to field curvatures together with ~ 2 mm position errors. Any second-order gradients with respect to $\mathcal{B}_2$ or $\mathcal{B}_3$, which are only measured in the presence of alignment errors, are negligble. Further, no curvatures of the form $\frac{\partial^2\mathcal{B}_1}{\partial x_i \partial x_j}$ are measured even on the $50\,\mu$G/cm^2 level except when

$i = j = 1$. We will therefore measure

$$[+\text{curvature}]$$

$$\mathcal{B}(\tilde{\mathcal{F}} = +1) = \mathcal{B}_1(1 - \tfrac{\delta^2}{8} - \tfrac{\beta^2}{2}) - \mathcal{B}_2\tfrac{\delta}{2} + \mathcal{B}_3\beta$$

$$+\mathcal{B}_{\text{offset}}$$

$$+\tfrac{\partial \mathcal{B}_1}{\partial x_1}L\cos\alpha - \tfrac{\delta}{2}\tfrac{\partial \mathcal{B}_2}{\partial x_1}L\cos\alpha + \beta\tfrac{\partial \mathcal{B}_3}{\partial x_1}L\cos\alpha$$

$$+\tfrac{\partial \mathcal{B}_1}{\partial x_2}L\sin\alpha - \tfrac{\delta}{2}\tfrac{\partial \mathcal{B}_2}{\partial x_2}L\sin\alpha + \beta\tfrac{\partial \mathcal{B}_3}{\partial x_2}L\sin\alpha$$

$$+\tfrac{1}{2}\tfrac{\partial^2 \mathcal{B}_1}{\partial x_1^2}L^2\cos^2\alpha$$

$$(5.6)$$

$$\mathcal{B}(\tilde{\mathcal{F}} = -1) = -\mathcal{B}_1(1 - \tfrac{\delta^2}{8} - \tfrac{\beta^2}{2}) - \mathcal{B}_2\tfrac{\delta}{2} + \mathcal{B}_3\beta$$

$$+\mathcal{B}_{\text{offset}}$$

$$+\tfrac{\partial \mathcal{B}_1}{\partial x_1}L\cos\alpha + \tfrac{\delta}{2}\tfrac{\partial \mathcal{B}_2}{\partial x_1}L\cos\alpha - \beta\tfrac{\partial \mathcal{B}_3}{\partial x_1}L\cos\alpha$$

$$+\tfrac{\partial \mathcal{B}_1}{\partial x_2}L\sin\alpha + \tfrac{\delta}{2}\tfrac{\partial \mathcal{B}_2}{\partial x_2}L\sin\alpha - \beta\tfrac{\partial \mathcal{B}_3}{\partial x_2}L\sin\alpha$$

$$-\tfrac{1}{2}\tfrac{\partial^2 \mathcal{B}_1}{\partial x_1^2}L^2\cos^2\alpha.$$

The last correction we will make is a "global" rotation by a small angle ϕ, allowing for the fact that the overall orientation of the fluxgate may be miscalibrated to lie partially along the laboratory $\hat{e}_2$-axis (averaged over both $\tilde{\mathcal{F}}$ states). This rotation has the effect of substituting $\tfrac{\delta}{2} \to \tfrac{\delta}{2} - \tilde{\mathcal{F}}\phi$. Keeping only the relevant orders, we have

$$[+\text{global rotation}]$$

$$\mathcal{B}(\tilde{\mathcal{F}} = +1) \quad = \quad \mathcal{B}_1(1 - \tfrac{\delta^2}{8} - \tfrac{\beta^2}{2} - \tfrac{\phi^2}{2} + \tfrac{\delta\phi}{2}) - \mathcal{B}_2(\tfrac{\delta}{2} - \phi) + \mathcal{B}_3\beta$$

$$+\mathcal{B}_{\text{offset}}$$

$$+\frac{\partial\mathcal{B}_1}{\partial x_1}L\cos\alpha - (\tfrac{\delta}{2} - \phi)\frac{\partial\mathcal{B}_2}{\partial x_1}L\cos\alpha + \beta\frac{\partial\mathcal{B}_3}{\partial x_1}L\cos\alpha$$

$$+\frac{\partial\mathcal{B}_1}{\partial x_2}L\sin\alpha - (\tfrac{\delta}{2} - \phi)\frac{\partial\mathcal{B}_2}{\partial x_2}L\sin\alpha + \beta\frac{\partial\mathcal{B}_3}{\partial x_2}L\sin\alpha$$

$$+\tfrac{1}{2}\frac{\partial^2\mathcal{B}_1}{\partial x_1^2}L^2\cos^2\alpha$$

$$(5.7)$$

$$\mathcal{B}(\tilde{\mathcal{F}} = -1) \quad = \quad -\mathcal{B}_1(1 - \tfrac{\delta^2}{8} - \tfrac{\beta^2}{2} - \tfrac{\phi^2}{2} - \tfrac{\delta\phi}{2}) - \mathcal{B}_2(\tfrac{\delta}{2} + \phi) + \mathcal{B}_3\beta$$

$$+\mathcal{B}_{\text{offset}}$$

$$+\frac{\partial\mathcal{B}_1}{\partial x_1}L\cos\alpha + (\tfrac{\delta}{2} + \phi)\frac{\partial\mathcal{B}_2}{\partial x_1}L\cos\alpha - \beta\frac{\partial\mathcal{B}_3}{\partial x_1}L\cos\alpha$$

$$+\frac{\partial\mathcal{B}_1}{\partial x_2}L\sin\alpha + (\tfrac{\delta}{2} + \phi)\frac{\partial\mathcal{B}_2}{\partial x_2}L\sin\alpha - \beta\frac{\partial\mathcal{B}_3}{\partial x_2}L\sin\alpha$$

$$-\tfrac{1}{2}\frac{\partial^2\mathcal{B}_1}{\partial x_1^2}L^2\cos^2\alpha.$$

Using this last result, we can compute the components of a measurement that are even and odd under $\tilde{F}$:

$$\mathcal{B}_1^{\text{nr}} \quad \equiv \quad [\mathcal{B}(\tilde{\mathcal{F}} = +1) + \mathcal{B}(\tilde{\mathcal{F}} = -1)]/2$$

$$= \quad \mathcal{B}_1\tfrac{\delta\phi}{2} - \mathcal{B}_2\tfrac{\delta}{2} + \mathcal{B}_3\beta$$

$$+\mathcal{B}_{\text{offset}}$$

$$+\frac{\partial\mathcal{B}_1}{\partial x_1}L\cos\alpha + \phi\frac{\partial\mathcal{B}_2}{\partial x_1}L\cos\alpha$$

$$+\frac{\partial\mathcal{B}_1}{\partial x_2}L\sin\alpha + \phi\frac{\partial\mathcal{B}_2}{\partial x_2}L\sin\alpha$$

$$(5.8)$$

$$\mathcal{B}_1^{\tilde{\mathcal{F}}} \quad \equiv \quad [\mathcal{B}(\tilde{\mathcal{F}} = +1) - \mathcal{B}(\tilde{\mathcal{F}} = -1)]/2$$

$$= \quad \mathcal{B}_1(1 - \tfrac{\delta^2}{8} - \tfrac{\beta^2}{2} - \tfrac{\phi^2}{2}) + \mathcal{B}_2\phi$$

$$-\tfrac{\delta}{2}\frac{\partial\mathcal{B}_2}{\partial x_1}L\cos\alpha + \beta\frac{\partial\mathcal{B}_3}{\partial x_1}L\cos\alpha$$

$$-\tfrac{\delta}{2}\frac{\partial\mathcal{B}_2}{\partial x_2}L\sin\alpha + \beta\frac{\partial\mathcal{B}_3}{\partial x_2}L\sin\alpha$$

$$+\tfrac{1}{2}\frac{\partial^2\mathcal{B}_1}{\partial x_1^2}L^2\cos^2\alpha.$$

Any of these terms might be even or odd under $\tilde{\mathcal{B}}$, and imperfections could behave as offsets (when consistent across measurements) or scatter (when fluctuating between measurements). It is important to realize that scatter is generated even in $\tilde{\mathcal{B}}$-even channels by

applied magnetic fields coupled to non-repeatable angles; however, any fixed angle offset can only enter into the $\tilde{\mathcal{B}}$-odd channel, provided the applied fields are perfectly reversing.

This has assumed that the fluxgate channel is primarily along $\hat{e}_1$. The same model specifies the systematic error contributions to the $\hat{e}_2$ channel, provided the substitutions $\mathcal{B}_1 \to \mathcal{B}_2$, $\mathcal{B}_2 \to -\mathcal{B}_1$, and $\mathcal{B}_3 \to \mathcal{B}_3$ are made.[9]

Non-rotating fluxgate channel

We use a similar model for the longitudinal fluxgate channel, which ideally does not change under fluxgate rotations. In particular, we will allow for small angles into the $e_1 e_2$-plane, fluxgate rotation errors, electronic offsets, and displacement from the rotation axis. As before, it is easiest to add each complication step by step. The measured field would ideally be

$$\text{[ideal]}$$
$$\mathcal{B}(\tilde{\mathcal{F}} = +1) \quad = \quad \mathcal{B}_3$$
$$\mathcal{B}(\tilde{\mathcal{F}} = -1) \quad = \quad \mathcal{B}_3. \tag{5.9}$$

If the fluxgate is oriented by a small angle β into the $e_1 e_2$-plane, and makes an angle ϕ (not necessarily small) with respect to the $\hat{e}_1$ axis within this plane, then

$$\text{[+orientation error]}$$
$$\mathcal{B}(\tilde{\mathcal{F}} = +1) \quad = \quad \mathcal{B}_3(1 - \tfrac{\beta^2}{2}) + \beta \cos \phi \mathcal{B}_1 + \beta \sin \phi \mathcal{B}_2$$
$$\mathcal{B}(\tilde{\mathcal{F}} = -1) \quad = \quad \mathcal{B}_3(1 - \tfrac{\beta^2}{2}) - \beta \cos \phi \mathcal{B}_1 - \beta \sin \phi \mathcal{B}_2. \tag{5.10}$$

If there are rotation errors such that the angle in the $e_1 e_2$-plane is $\phi \to \phi - \delta/2$ in the $\tilde{\mathcal{F}} = +1$ state but $\phi \to \phi + \pi + \delta/2$ in the $\tilde{\mathcal{F}} = -1$ state, corresponding to a total rotation of $\pi + \delta$, then some trigonometric identities yield

---

9. The negative sign appears in this mapping because the coordinate system $\{\hat{e}_1, \hat{e}_2, \hat{e}_3\}$ is assumed to be right-handed and the $\hat{e}_2$ axis can be put in the position of the $\hat{e}_1$ axis by a $-\pi/2$ rotation about $\hat{e}_3$, which sends $\hat{e}_1 \to -\hat{e}_2$.

$$[\text{+rotation error}]$$

$$\mathcal{B}(\tilde{\mathcal{F}} = +1) = \mathcal{B}_3(1 - \tfrac{\beta^2}{2}) + \beta(\cos\phi + \tfrac{\delta}{2}\sin\phi)\mathcal{B}_1 + \beta(\sin\phi - \tfrac{\delta}{2}\cos\phi)\mathcal{B}_2 \quad (5.11)$$

$$\mathcal{B}(\tilde{\mathcal{F}} = -1) = \mathcal{B}_3(1 - \tfrac{\beta^2}{2}) - \beta(\cos\phi - \tfrac{\delta}{2}\sin\phi)\mathcal{B}_1 - \beta(\sin\phi + \tfrac{\delta}{2}\cos\phi)\mathcal{B}_2.$$

An electronic offset adds to both states the same way:

$$[\text{+electronic offset}]$$

$$\mathcal{B}(\tilde{\mathcal{F}} = +1) = \mathcal{B}_3(1 - \tfrac{\beta^2}{2}) + \beta(\cos\phi + \tfrac{\delta}{2}\sin\phi)\mathcal{B}_1 + \beta(\sin\phi - \tfrac{\delta}{2}\cos\phi)\mathcal{B}_2$$
$$+ \mathcal{B}_{\text{offset}} \quad (5.12)$$

$$\mathcal{B}(\tilde{\mathcal{F}} = -1) = \mathcal{B}_3(1 - \tfrac{\beta^2}{2}) - \beta(\cos\phi - \tfrac{\delta}{2}\sin\phi)\mathcal{B}_1 - \beta(\sin\phi + \tfrac{\delta}{2}\cos\phi)\mathcal{B}_2$$
$$+ \mathcal{B}_{\text{offset}}.$$

Finally, we will allow the channel to be displaced from the rotation axis by a distance L in a direction that makes angle α (not necessarily small) with the $\hat{e}_1$ axis in the e_1e_2-plane. This allows the following coupling to first-order gradients:

$$[\text{+gradients}]$$

$$\mathcal{B}(\tilde{\mathcal{F}} = +1) = \mathcal{B}_3(1 - \tfrac{\beta^2}{2}) + \beta(\cos\phi + \tfrac{\delta}{2}\sin\phi)\mathcal{B}_1 + \beta(\sin\phi - \tfrac{\delta}{2}\cos\phi)\mathcal{B}_2$$
$$+ \mathcal{B}_{\text{offset}}$$
$$+ \tfrac{\partial\mathcal{B}_1}{\partial x_1}\beta\cos\phi L\cos\alpha + \tfrac{\partial\mathcal{B}_2}{\partial x_1}\beta\sin\phi L\cos\alpha + \tfrac{\partial\mathcal{B}_3}{\partial x_1}L\cos\alpha$$
$$+ \tfrac{\partial\mathcal{B}_1}{\partial x_2}\beta\cos\phi L\sin\alpha + \tfrac{\partial\mathcal{B}_2}{\partial x_2}\beta\sin\phi L\sin\alpha + \tfrac{\partial\mathcal{B}_3}{\partial x_2}L\sin\alpha \quad (5.13)$$

$$\mathcal{B}(\tilde{\mathcal{F}} = -1) = \mathcal{B}_3(1 - \tfrac{\beta^2}{2}) - \beta(\cos\phi - \tfrac{\delta}{2}\sin\phi)\mathcal{B}_1 - \beta(\sin\phi + \tfrac{\delta}{2}\cos\phi)\mathcal{B}_2$$
$$+ \mathcal{B}_{\text{offset}}$$
$$+ \tfrac{\partial\mathcal{B}_1}{\partial x_1}\beta\cos\phi L\cos\alpha + \tfrac{\partial\mathcal{B}_2}{\partial x_1}\beta\sin\phi L\cos\alpha - \tfrac{\partial\mathcal{B}_3}{\partial x_1}L\cos\alpha$$
$$+ \tfrac{\partial\mathcal{B}_1}{\partial x_2}\beta\cos\phi L\sin\alpha + \tfrac{\partial\mathcal{B}_2}{\partial x_2}\beta\sin\phi L\sin\alpha - \tfrac{\partial\mathcal{B}_3}{\partial x_2}L\sin\alpha.$$

No second- or higher-order gradients are significant; as discussed in the previous section, only a curvature along the main measurement axis can be significant at the μG level, but in the present case there is negligible translation along the measurement axis upon rotation. We can then take the reversing and non-reversing components of the field along the fluxgate rotation axis using the expression above:

$$
\begin{aligned}
\mathcal{B}_3^{\text{nr}} \quad \equiv \quad & [\mathcal{B}(\tilde{\mathcal{F}} = +1) + \mathcal{B}(\tilde{\mathcal{F}} = -1)]/2 \\
= \quad & \mathcal{B}_3(1 - \tfrac{\beta^2}{2}) + \beta\tfrac{\delta}{2}\sin\phi\mathcal{B}_1 - \beta\tfrac{\delta}{2}\cos\phi\mathcal{B}_2 \\
& \qquad + \mathcal{B}_{\text{offset}} \\
& + \tfrac{\partial\mathcal{B}_1}{\partial x_1}\beta\cos\phi L\cos\alpha + \tfrac{\partial\mathcal{B}_2}{\partial x_1}\beta\sin\phi L\cos\alpha \\
& + \tfrac{\partial\mathcal{B}_1}{\partial x_2}\beta\cos\phi L\sin\alpha + \tfrac{\partial\mathcal{B}_2}{\partial x_2}\beta\sin\phi L\sin\alpha
\end{aligned}
\tag{5.14}
$$

$$
\begin{aligned}
\mathcal{B}_3^{\tilde{\mathcal{F}}} \quad \equiv \quad & [\mathcal{B}(\tilde{\mathcal{F}} = +1) - \mathcal{B}(\tilde{\mathcal{F}} = -1)]/2 \\
= \quad & \beta\cos\phi\mathcal{B}_1 + \beta\sin\phi\mathcal{B}_2 \\
& \tfrac{\partial\mathcal{B}_3}{\partial x_1}L\cos\alpha + \tfrac{\partial\mathcal{B}_3}{\partial x_2}L\sin\alpha.
\end{aligned}
$$

Again, these terms can generically be even or odd under $\tilde{\mathcal{B}}$, and they may appear as offsets or as scatter depending on the mechanisms involved. The components odd under $\tilde{\mathcal{F}}$ consist entirely of systematic errors. The systematic error contribution to $\mathcal{B}_3^{\text{nr}}$ is overwhelmingly dominated by $\mathcal{B}_{\text{offset}}$.

5.2.7 Effect of imperfections

When applying a nominal $\mathcal{B}_z$, we have $\mathcal{B}_z \sim 45$ mG, $\frac{\partial\mathcal{B}_i}{\partial x_j} \sim 1$ mG/cm for multiple gradients, and $\frac{\partial^2\mathcal{B}_z}{\partial z^2} \sim 10\,\mu$G/cm^2, which is small enough to neglect. At most, $\mathcal{B}_{x(y)} \sim 500\,\mu$G.

There are 10 meaningful parity terms (neglecting $\mathcal{B}^{\tilde{\mathcal{F}}}$ and $\mathcal{B}^{\tilde{\mathcal{B}}\tilde{\mathcal{F}}}$ along the rotation axes) for each magnetometer (probulator, FGYE, FGYW), giving thirty field components in all for a given magnetic field. It is tedious and unhelpful to walk through every contribution to each of these field components in turn. We will begin by crudely characterizing the components, beginning with the probulator. First, suppose that the fluxgates are centered perfectly on the rotation axis, so $L \to 0$. This is a valid assumption because the gradients in the molecular beam line are sufficiently small (there are larger $\frac{\partial\mathcal{B}_y}{\partial y}$ gradients in the vicinity

Parity	xyz	Offset (μG)	Exp.	μG$_{\text{pp}}$	Exp.	Notes
$\mathcal{B}^{\tilde{\mathcal{B}}\tilde{\mathcal{F}}}$	$\mathcal{B}_x$	-	-	50	100	Not meaningful channel
	$\mathcal{B}_y$	0	± 400	100	100	
	$\mathcal{B}_z$	46.8×10^3	$\sim 50 \times 10^3$	1	1	**Applied field channel**
$\mathcal{B}^{\tilde{\mathcal{F}}}$	$\mathcal{B}_x$	-	-	-	-	Not meaningful channel
	$\mathcal{B}_y$	-50	± 50	100	100	
	$\mathcal{B}_z$	-25	± 50	10	1	
$\mathcal{B}^{\tilde{\mathcal{B}}}$	$\mathcal{B}_x$	-200	± 200	50	100	
	$\mathcal{B}_y$	100	± 200	100	100	
	$\mathcal{B}_z$	2	± 4	4	1	
$\mathcal{B}^{\text{nr}}$	$\mathcal{B}_x$	-80	± 50	80	100	Electronic offsets
	$\mathcal{B}_y$	150	± 50	100	100	Electronic offsets
	$\mathcal{B}_z$	150	± 50	1	1	Electronic offsets

Table 5.3: Representative field offsets and scatter (μG$_{\text{pp}}$) along the probulator scan, with a nominal applied $\mathcal{B}_z$. Offsets are expected ("Exp.") in $\mathcal{B}_y^{\tilde{\mathcal{B}}\tilde{\mathcal{F}}}$ up to $\sim 400\,\mu$G due to fluxgate non-orthogonality; in $\mathcal{B}^{\tilde{\mathcal{B}}}$ up to $\sim 200\,\mu$G due to specified rotation inaccuracy; in $\mathcal{B}^{\tilde{\mathcal{F}}}$ up to $\sim 50\,\mu$G due to genuine ambient fields; and in $\mathcal{B}^{\text{nr}}$ up to $50\,\mu$G due to electronic offsets (though measured electronic offsets exceed this specification). The offset errors in $\mathcal{B}_z^{\tilde{\mathcal{B}}\tilde{\mathcal{F}}}$ and $\mathcal{B}_z^{\tilde{\mathcal{B}}}$ are suppressed by an additional factor of $\sim 10^{-2}$ because the effect of rotation errors along the applied field direction is only second order in the rotation error. The largest offset, in $\mathcal{B}_z^{\tilde{\mathcal{B}}\tilde{\mathcal{F}}}$, is naturally dominated by the genuine applied field. Scatter is expected in all parity terms up to $\sim 100\,\mu$G from both non-repeatability of the rotation stage and small changes in magnetometer orientation with respect to the nominal x-axis due to moving of the magnetometer in the probulator "guide tube"—but these are again suppressed by a factor of $\sim 10^{-2}$ in $\mathcal{B}_z$ relative to $\mathcal{B}_y$ and $\mathcal{B}_z$. All gradients along x (the probulator scan axis) are negligible except $\frac{\partial \mathcal{B}_x}{\partial x} \sim 0.1$ mG/cm, which is necessarily generated together with a comparable $\frac{\partial \mathcal{B}_z}{\partial z}$ from the applied field. These data are from the "initial" configuration, prior to collection of systematic error data. See Fig. 5.7 for an example of probulator measurements in this configuration.

of the FGY magnetometers) that their contributions to the measured field can be neglected.[10]

The data are consistent with rotation non-repeatability and fluxgate non-orthogonality that are well within their specified limits.

There are so many potential parameters of interest (e.g., the offset, slope, scatter, or deviations from linearity in three field components, measured by up to two separate channels in three magnetometers, across every applied field configuration) that it is extremely difficult to present our results in a way that is both tractable and detailed. As a compromise, I have

---

10. The exception is a small gradient $\frac{\partial \mathcal{B}_x}{\partial x}$ generated by necessity due to a non-zero $\frac{\partial \mathcal{B}_z}{\partial z}$. However, this is too small to couple non-trivially to any imperfections.

visually examined all plots obtained from the magnetic field scans, amounting to at least 540 traces (four meaningful parity components for each reversible lab axis and two for the magnetometer rotation axis, measured by 3 magnetometers over 9 configurations, and plotted separately for measurements performed with pairs of rotation states $\{0°, 180°\}$ and $\{90°, -90°\}$). In doing so, I failed to identify any outstanding anomalies. A summary of representative values of the magnetic field offsets and scatter, which in most cases arise due to systematic errors in the measurement, is given in Table 5.3. The systematic error model for the probulator measurement previously discussed, which includes (most importantly) the effects of rotation errors and non-orthogonality of the fluxgates within a single three-axis magnetometer, accounts for the measurement offsets and scatter reasonably well. As noted previously, a summary of the primary applied field component for each nominal magnetic field coil configuration is given in Table 5.2.

Chapter 6

Toward an improved measurement

This chapter will deal with an eclectic assortment of considerations relevant to improving the ACME experiment even further, toward a target EDM sensitivity of $10^{-30}\, e \cdot \mathrm{cm}$. The next-generation version of the ACME experiment will be referred to as "ACME III." This will involve improved control over systematic error parameters, in particular birefringence gradients that can lead to AC Stark shift systematics and magnetic field gradients. It will also include improved statistics via various routes to increase the flux of useful molecules through the interaction region and to increase the detection efficiency of molecules already in the interaction region.

6.1   Molecular focusing with electric or magnetic fields

In buffer gas beam experiments, molecular divergences from the source are typically on the order of $\sim 45°$ FWHM. For a detection region with transverse size on the order of ~ 1 cm at a distance of ~ 1 m from the source, only one in $\sim 10^{-4}$ molecules will have the required transverse velocity to pass through the detection region. A molecular lens, in analogy to an optical lens, focuses the diverging molecular beam as it moves into the

detection region. This section will discuss a system for designing a molecular lens, along with the limitations in the ideal and practical performance. In the context of the ACME experiment, we are pursuing the possibility of either a magnetostatic lens or an electrostatic lens, with both options likely realized using the long-lived Q electronic state of ThO. Xing Wu has now characterized the Q state in detail and is actively considering both routes. In this section, I will describe our approach to analyzing the gain in useful signal due to a lens. The details of the beamline geometry and achievable electric or magnetic potential depth are still indeterminate, and so the quantitative results here are merely representative.

6.1.1 Optical lens analogy

In a molecular lens, a molecule is subject to a potential field $U(x, y, z)$, and the trajectory it traces through space is affected. If we select the x-direction as the molecular beam axis (to be consistent with the coordinates used in the ACME experiment), then we can draw an analogy with geometric optics with an "optical axis" along $\hat{x}$. We will use the Lagrangian perspective in order to make the analogy clear. The optical Lagrangian expresses Fermat's principle, namely that a light ray takes the path of least time between two points:

$$\mathcal{L}_{\text{optics}} = n(x, y, z), \tag{6.1}$$

where $n(x, y, z)$ is the index of refraction at a particular point in space. A mechanical Lagrangian is of the form

$$\mathcal{L}_{\text{mechanical}} = \frac{1}{2}m(v_x^2 + v_y^2 + v_z^2) - U(x, y, z), \tag{6.2}$$

where m is the particle mass. Consider the extreme limit of the molecular analog to the "paraxial approximation," $v_y, v_z \to 0$, so

$$\begin{aligned} \mathcal{L}_{\text{mechanical}} &= \tfrac{1}{2}mv_x^2 - U(x, y, z) \\ &= \tfrac{1}{2}mv_0^2 + mv_0 dv + \tfrac{1}{2}m(dv)^2 - U(x, y, z), \end{aligned} \tag{6.3}$$

230

where $v_x \equiv v_0 + dv$ for an initial forward velocity v_0 and change in the velocity dv. We are free to ignore the overall offset to the Lagrangian, $\frac{1}{2}mv_0^2$. If the largest energy scale of the problem is U_0, then the largest possible change to the forward velocity is given by energy conservation: $\frac{1}{2}m(v_0 + dv)^2 - \frac{1}{2}mv_0^2 = U_0 \Rightarrow \frac{dv}{v_0} \approx U_0/(mv_0^2)$. In other words, the fractional change in the forward velocity is suppressed by the ratio of initial kinetic energy to available potential energy. For a molecular lens, the available potential energy is $U_0 \sim k_B \times 1$ K, while the ThO beam in ACME has $\sim k_B \times 600$ K of kinetic energy. Therefore, we can treat $dv \approx 0$ and write $\mathcal{L}_{\mathrm{mechanical}} \to -U(x, y, z)$. We identify the effective index of refraction of the molecular lens, $n_{\mathrm{eff}}(x, y, z) = -U(x, y, z)$, up to a constant. Of course the $v_y, v_z \to 0$ limit is not perfectly achieved, and corrections to the Lagrangian occur at the order of $\mathcal{O}(\theta U_0)$, where $\theta = \sqrt{(v_y^2 + v_z^2)/v_x^2}$ is the molecular divergence angle. However, all molecules that we are able to redirect with the ~ 1 K depth potentials that are technically available via Stark or Zeeman shifts have extremely small divergence angles, $\theta \ll 1$, at all points in their trajectories.

In conclusion, we can replace a mechanical treatment of the molecules with an optical treatment as long as $v_y^2 + v_z^2 \ll v_x^2$ and $U_0 \ll mv_x^2$. This is useful because we no longer think in terms of time evolution, but instead parametrize trajectories by their positions along the molecular beam axis, just as we can consider the position of a light ray along the optical axis. Further, the phase-space of geometric optics is only four-dimensional (y, p_y, z, p_z) rather than six-dimensional (also including x, p_x) since the x coordinate takes the role of time in an ordinary Hamiltonian. This makes the analysis significantly simpler and allows us to build intuition for the behavior of the molecular trajectories.

6.1.2 Conservation laws

Etendue

The optical Lagrangian can be recast in terms of a Hamiltonian, which has an associated phase-space. An analogue to Liouville's theorem then tells us that etendue (phase-space volume) is conserved, $\frac{d}{dt} \int n^2 dS \cos\theta \, d\Omega = 0$, where dS is a differential area element (in the yz-plane for our purposes) of a ray bundle, and $d\Omega$ is a differential solid angle element

subtended by its outgoing rays. For a surface with total area S and divergence angles up to half-angle α, in some region with constant index n, the total etendue is $G_{\mathrm{et}} = \pi S\, n^2 \sin^2 \alpha \approx \pi S\, n^2\, \alpha^2$.

The conservation of G_{et} gives one fundamental performance limit for a lens design: we cannot focus any phase-space volume from our source that is larger than the phase-space volume of trajectories that would pass between the input and output boundaries of the target. One of the most important consequences of etendue conservation is that decreasing the divergence of a bundle of rays, regardless of the method, must increase the spatial extent and vice versa. The design of optical elements that most efficiently transform a bundle of incident rays into the optimal bundle of outgoing rays, subject to the major constraint of etendue conservation, is the domain of non-imaging optics[204].

It is useful to know how strongly the etendue limit constrains the gain that could be achieved with a molecular lens. In ACME, the initial molecule distribution has cylindrical symmetry, essentially defined by a radial velocity distribution and 3 mm radius "skimmer" aperture that attenuates the neon gas load from the beam source into the beam line. The phase-space volume of trajectories with transverse velocity up to $v_\perp$ is therefore $\pi {v_\perp}^2 \times \pi r^2$. Without a lens, the viable velocity classes to reach the detection region are approximately those with $v_\perp \leq 2$ m/s ($\theta = \frac{v_\perp}{v_x} \approx 10^{-2}$). Thus the phase-space volume of trajectories from the source that can reach the detector, in units where[1] $m_{\mathrm{ThO}} = 1$, is $\sim 4 \times 10^{-4} \frac{\mathrm{m2}}{\mathrm{s}}$.

The maximum target phase-space volume is defined by trajectories that would travel between two boundaries (in particular, the collimating aperture and the detection region). If these boundaries are square, then the trajectories that survive are decoupled in the y and z directions. Consider only the y-direction, for example, and we will assume for simplicity (and a bit conservatively) that the detector area is also bounded to the same area as the collimator, but is separated by the field plate length L. Starting from the first collimator, at a given initial y-position of $y(0) = y_0$, there is some minimum velocity $v_-(y_0)$ that reaches position $y(L) = -d_{\mathrm{coll}}/2$ and some maximum velocity $v_+(y_0)$ that reaches position

1. Recall that phase-space volume in one dimension has units of [momentum]×[length], but throughout this section I use units of [velocity]×[length], which is more natural to the design on a lens for a molecule with fixed mass.

$y(L) = +d_{\text{coll}}/2$. Both functions $v_{\pm}(y_0)$ are linear in the starting position y_0 for a given forward velocity v_x and length L, so the allowed trajectories form a parallelogram in (y, v_y) phase-space at the collimator position. Two of the corners of this phase-space parallelogram, defining the minimum velocity for a molecule starting at $y_0 = -d_{\text{coll}}/2$ and the maximum velocity for a molecule starting at $y_0 = +d_{\text{coll}}/2$, are $v_{\pm}(\pm d_{\text{coll}}/2) = 0$. The opposing corners of the phase-space parallelogram, defining the maximum velocity for a molecule starting at $y_0 = -d_{\text{coll}}/2$ and the minimum velocity for a molecule starting at $y_0 = +d_{\text{coll}}/2$, is obtained from a diagonal trajectory passing between $y(0) = \pm d_{\text{coll}}/2$ and $y(L) = \mp d_{\text{coll}}/2$ over a distance L, giving $v_{\pm}(\mp d_{\text{coll}}/2) = \frac{d_{\text{coll}}}{L} v_x$ in the small-angle approximation. The area of this parallelogram is $\frac{d_{\text{coll}}^2}{L} v_x$ and the area of the four-dimensional phase-space volume is $\left(\frac{d_{\text{coll}}^2}{L} v_x \right)^2$.

In ACME II, $d_{\text{coll}} = 2.4$ cm and $L = 43$ cm, while $v_x \approx 180$ m/s. Therefore, the maximum phase-space area (again, letting $m_{\text{ThO}} = 1$) that we could obtain without increasing the collimator spacing is $\sim 6 \times 10^{-2} \frac{\text{m}^2}{\text{s}}$, exceeding what we currently obtain by a factor of ~ 150. If we allow a larger spatial extent at the downstream end of the field plates, this could of course increase the admissible etendue even without changing the "fixed collimator" on the upstream end of the field plates. We will see that the anticipated gain in flux from a molecular lens is $\lesssim 15$, so the etendue limit is not a significant constraint.

In particular, there are likely minimal gains to be obtained from using the full power of non-imaging optics, which is designed to transfer the maximal initial number of rays to the region of phase-space that can be detected. An imaging element such as a lens, while not achieving this goal, is suitable because the phase-space volume that we are capable of detecting is orders of magnitude larger than the phase-space volume of molecules that we actually detect without any molecular focusing.

Energy

The velocity classes that we can turn around in a lens (or other focusing object) are set by the potential energy of the electric or magnetic interaction with the molecule. For a potential energy of U_0, we expect to be able to focus molecules with velocity up to $v_{\max} \approx \sqrt{2U_0/m}$. The corresponding gain in flux would be $\sim \left(\frac{v_{\max}}{v_0} \right)^2$, where v_0 is the maximum velocity

233

that can be captured with no focusing (in ACME, $v_0 \approx 2$ m/s as stated earlier). For a 1-kelvin potential depth with ThO, $v_{\max} = 8.2$ m/s, giving a maximum gain of ~ 17. Since $U_0 \propto v_{\max}^2$, and likewise the number of momentum states that can be captured are proportional to $v_{\max}^2$ (the area of the circle in (v_y, v_z) phase-space with radius $v_{\max}$), we expect the gain to be limited by $G \leq 17 \times U_0[\mathrm{K}]$

In principle, it is possible to turn around *any* velocity class with any potential depth provided a lens profile can be constructed in an explicitly x-dependent way. This will be explored more in subsequent sections. Unfortunately, given reasonable engineering constraints, the performance over a "simple" lens can only be improved by $\sim 30\%$.

6.1.3 Collimating an ideal source

To help build intuition for the design of a molecular lens, we'll consider the problem of collimating a point source, which would be the ideal limit (from the perspective of molecular focusing) of our beam source.

Consider the simplest case of a point source at a distance l from the beginning of a lens, defined to be at position $x = 0$. For a molecule with initial divergence angle $\theta \ll 1$, the position and velocity upon entry are $y(x = 0, \theta) = l\theta$ and $v_y(x = 0, \theta) = v_x\theta$. We would like to find a class of potentials $U(x, y)$ such that all trajectories emerge at $x = L$ collimated, i.e., $v_y(x = L, \theta) = 0$ independent of θ.

It is fairly easy to show that $U(x, y) = U(y) = \frac{1}{2}m\omega^2 y^2$ suffices for some particular length L; we'll work this out in detail in Sec. 6.1.4. In particular, all molecules undergo oscillatory motion. The harmonic oscillator potential has the special property that the period of oscillation is independent of the amplitude, so all molecules turn around at the same time, and thus the same x-position for a given forward velocity.

As we will see, it's actually possible to generalize this result to the case in which $\omega = \omega(x)$, so that the strength of the transverse harmonic potential depends on the longitudinal position within the lens. For clarity, I will defer the details of this until I've described the basic approach to the molecular lens design.
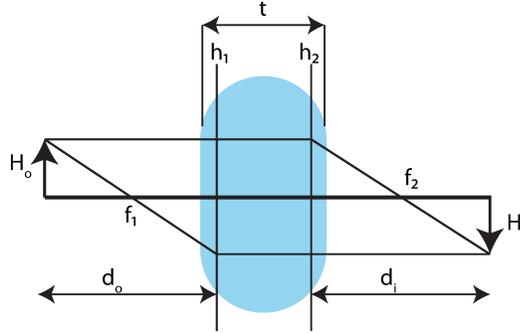
Figure 6.1: Diagram of a thick lens. Here, the lens is drawn as symmetric so that $f_1 = f_2$ and $d_o = d_i$.

6.1.4 Thick lens model

We will build on the geometric optics analogy for a molecular lens by considering the operation of a thick optical lens. First, the thin lens equation is

$$\frac{1}{d_i} + \frac{1}{d_o} = \frac{1}{f}, \tag{6.4}$$

where d_i is the image-lens distance, d_o is the object-lens distance, and f is the focal length of the lens. The magnification is $M = \frac{d_i}{d_o}$. (We don't care whether the image is inverted, so I always take $M > 0$).

For a thick lens, we will see that similar formulas hold but the corresponding quantities are defined as follows [205, Sec. 6.1]; see Fig. 6.1. The focus f is the position from which a point source is collimated through the lens. If we extend the path of an initally diverging ray straight through the lens, and do likewise in the reverse direction for the outgoing horizontal ray, then those two lines intersect at a point. The collection of all such intersection points define a plane, known as the first principal plane, at $x = h_1$. There is a second focal point on the "downstream" side of the lens, with a corresponding "second principal plane" at $x = h_2$. We then interpret d_o and d_i, respectively, as the distance between the object and "first principal plane," and between the image and the "second principal plane." Conventionally, the focal lengths f_1 and f_2 are also measured relative to their respective principal planes. Note that typically $f_1 \neq f_2$ if the potential used to generate an effective molecular lens does not have reflection symmetry about $x = L/2$ for

a lens with length L (a case that we will explore).

To compute the magnification, note that $M \equiv \frac{H_i}{H_o}$, where $H_{o(i)}$ denotes the transverse position of a ray at the object or image plane. Let θ_o be the angle between the horizontal and the line passing between a point on the object and the focus, while θ_i is the analogous angle for the image. Then by construction, $H_i = \theta_o f_1$ and $H_o = \theta_i f_2$, so $M = \frac{\theta_o}{\theta_i} \frac{f_1}{f_2}$. Now notice that $(H_i + H_o)\theta_o = d_o$, and likewise for d_i, so that $M = \frac{f_1}{f_2} \frac{d_i}{d_o}$. (Reference to Fig. 6.1 may be helpful.) This reduces to the thin-lens magnification equation when $f_1 = f_2$, as occurs for an ordinary optical lens with air on both sides.

We would like to obtain the analogue to the thin-lens equation as well. We note that $\theta_o d_o = \theta_i d_i = H_1 + H_2$, and $\theta_o f_1 = \theta_i (d_i - f_2) = H_2$. From these, we find $d_i f_1 + d_o f_2 = d_o d_i$. In the special case that $f_1 = f_2$, this can be simplified to the usual thin-lens equation.

Harmonic potential as a thick lens

It is clear from the discussion of collimating a spatial point source, with the logic reversed, that a set of horizontal rays will converge on some focal point if incident on a potential $U(y) = \frac{1}{2} m \omega^2 y^2$ from either the left or right, so there exist some focal lengths $f_{1(2)}$ relative to principal planes at $x = h_{1(2)}$. The simplest case is a uniform harmonic potential between $x = 0$ and $x = L$, and zero potential outside these bounds. By reflection symmetry about $x = L/2$, we already know that $f_1 = f_2 \equiv f$ and $h_1 = L - h_2 \equiv h$. We consider the "inputs" to be R (the inner radius of the lens), $U_{\max}$ (the potential depth of the lens at the wall), m (the molecular mass), v (the forward velocity), M (the magnification), and d_o (the distance between the object and first principal plane). By straightforward application of Newton's laws, we can find the useful relations

- $\omega^2 = 2U_{\max}/(mR^2)$

- $k \equiv \frac{\omega}{v}$ is the wavenumber of molecule oscillations in the potential (typically ~ 0.4 cm^{-1})

- $d_i = M d_o$

- $f = \frac{d_i}{M+1}$

236

- $l = \sqrt{f^2 - k^{-2}}$ is the distance between the focal point and the physical input aperture of the lens

- $h = \sqrt{l^2 + k^{-2}} - l$

- $L = k^{-1}\left(\frac{\pi}{2} - \sin^{-1}\left[\frac{l}{f}\right]\right)$

In practice, the only variables of interest for optimization are the lens radius R, the distance from the object to the first principal plane d_o, and the magnification M. It is interesting to note that, other things equal, there is a maximum inner radius R_{max} at which the distance between the focal point and physical input aperture vanishes, $l \to 0$. For $R > R_{\text{max}}$, the focal point is *inside* the physical lens, so that a beam will only be collimated if it is converging (as opposed to parallel) at the entrance to the lens. Naturally, I ignore this regime of parameter space in my simulations. We usually find that the gain is optimized with respect to R by allowing $R \to R_{\text{max}}$ so that as many molecules are admitted into the lens as possible.

For any other geometry with a harmonic potential that depends on the longitudinal position, $U(x,y) = \frac{1}{2}m\omega(x)^2 y^2$, it is not typically possible to find algebraic expressions for f_2 and f_1. However, the focal lengths can be found by direct simulation of single trajectories that are initially horizontal with $\vec{v} = +v_x\hat{x}$ and $\vec{v} = -v_x\hat{x}$; the positions at which these trajectories cross the x-axis are the focal points.

A homogeneous molecular lens with a quadratic potential can therefore be analyzed fairly straightforwardly in terms of the thick-lens equation. I have found that, although this helped build my intuition somewhat, this approach did not facilitate any quantitative benefit over numerically optimizing a lens with respect to the radius and length.

6.1.5 Parameter space and optimization strategy

In ACME, we have many constraints on the molecular beam geometry. We require that no molecules hit the field plates in order to prevent patch potentials from forming; we could modify the skimmer aperture; we can vary the dimensions of the collimators at the entrance to the field plates, as well as the source-lens and lens-field plate distances; we can design

the collection optics to accommodate different "detection volumes"; and of course we can implement different lens geometries.

In order to make the design process tractable, we want to eliminate several degrees of freedom. Changing the skimmer diameter or position within reasonable ranges should not significantly affect the optimal lens design, even if it affects the final flux (depending on the distribution of molecules in the zone of freezing, which we don't understand particularly well; see Sec. 6.2). Therefore, I do not consider alternate skimmer dimensions when optimizing the lens geometry.

We can also pull the field plates farther apart to reduce the number of trajectories that intersect them, at the relatively minor cost of applying larger voltages to obtain the same electric field, and the more significant cost of potentially reduced optical collection efficiency. While initially optimizing over the lens parameters, we can assume completely open collimators at the interaction region, only maximizing the number of trajectories that pass through some detection region. After an optimal lens design is obtained, we can decrease the collimator size until molecules no longer hit the field plates. This approach decouples the lens and collimator specifications. The final gain, in which no molecules are permitted to hit the field plates, can be up to 20% lower than when the collimators are left completely open[2].

For a 1-kelvin potential depth, it is possible to turn around ThO molecules with transverse velocity $v_\perp \approx 8$ m/s. If the source has a 3 mm radius at a distance of 20 cm from the lens, then the opening aperture must be nearly an inch in diameter to admit all energetically accessible molecules. (The situation is actually somewhat worse than this because a trajectory with $v_\perp \approx 8$ m/s can only be turned around if it enters near the center of the lens, not near the walls.) A larger-diameter lens must necessarily be longer because the corresponding oscillation frequency ω decreases with the inner lens radius. Given reason-

---

2. This result assumes a 5 cm field plate separation, slightly larger than the 4.5 cm separation used in ACME II. However, the presumed detection volume is no larger than ACME II, and this change makes little substantial difference in the calculated flux gain due to the lens. The main effect of a slightly larger plate separation is the complete elimination of otherwise rare events in which molecules intersect the field plates for the largest practical collimator spacing. If we decide against any redesign of the electric field plate geometry, then the conclusions of this section will not be significantly modified (in particular, the focusing potential energy used in the end is likely to alter the results by a larger margin).

able engineering constraints, I find that we should place the lens as close to the source as possible. The default skimmer-lens distance I have used is 25 cm so that we plausibly have space to prepare the focusing state. Xing Wu has recently explored more aggressive ways to reduce the source-lens distance to as little as 20 cm, including by performing rotational cooling inside the beam source vacuum chamber. I likewise minimize the distance between the lens and interaction region field plates at around 80 cm, based on practical constraints such as the size of the magnetic shielding. This could plausibly be reduced further, for example by eliminating one or two of the five layers of magnetic shielding. However, the lens-field plate distance affects the number of detected molecules only weakly[3].

For the quantitative results of simulations discussed here, I assume a maximum potential energy of 1.4 K, consistent with estimates for the magnetic lens made by Adam West. However, the exact potential depth that we will use–even whether we use a magnetic or electric lens–is still to be determined. To a good approximation, the number of molecules focused into the detection region is linear in the potential depth.

I compute a gain over the ACME II flux assuming perfect preparation into and out of the focusing state (only because the degree of imperfection is not yet known). The detection area is assumed to be 3 cm in diameter (comparable to the cloud size at the end of the field plates in ACME II). The whole procedure is repeated for detection in a dump region, 1 m downstream of the field plates, with "gains" normalized to the number of molecules that reach *that* detection region in the ACME II geometry. These simulations account for the possibility that we will perform optical cycling detection in ACME III in a separate vacuum chamber outside of the magnetic shields. The assumption that this will occur 1 m downstream is fairly arbitrary; as the optical cycling project develops, the requirements for the detection region will be much clearer. Nevertheless, under these assumptions, a "gain" of 30 in the dump region is a gain of only ~ 10 compared to current ACME II levels because the "baseline" with no lens is smaller (due to the reduced solid angle of a detector 1 m downstream of the field plates). I've chosen to use this convention partly because it

---

3. For a given geometry, the total distance from the source to the detector *can* affect the "gain" over the no-lens configuration significantly, simply because the solid angle subtended by the detection area is reduced if the detector is farther from the molecular source. This is important to keep in mind if comparing the advantage of a molecular lens to the geometry optimized for *no* lens.
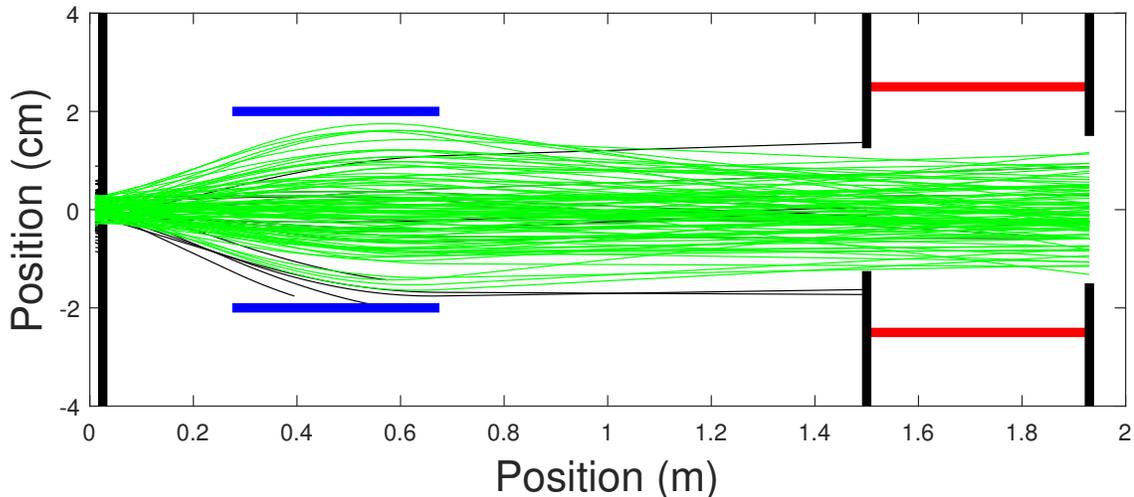
Figure 6.2: Example trajectories with detection in the interaction region. Detected trajectories are in green, undetected in black. No molecules hit the field plates in this geometry. Thick vertical black lines are collimators (skimmer, IR collimators, and detection boundary), hotizontal blue lines show the inner lens radii, and horizontal red lines show the field plates (5 cm apart in this simulation).

makes clearer what the "pain vs. gain" curve is for implementing a particular lens design; if we detect in the dump, then a lens will be much more essential even if it cannot contribute an increased flux over ACME II levels. In the same spirit, however, it is important to realize that the state transfer into and out of the focusing state will not be perfect. For the plausible one-way transfer efficency of 80%, the gains reported here would need to be scaled down to 65% of their nominal values, to account for the losses when transferring both into and out of the focusing state.

Trajectory simulation

The trajectory simulations are written in MATLAB, and closely follow the approach used by Adam West in earlier trajectory simulations for ACME.

I model the source as a disc, uniformly occupied up to a small thickness and finite diameter. The longitudinal velocities are normally distributed with average value $\bar{v}_{||} = 180$ m/s and standard deviation $\sigma_{v_{||}} = 17$ m/s. The transverse velocities are normally distributed with average $\bar{v}_\perp = 0$ and standard deviation $\sigma_{v_\perp} = 52$ m/s, so that the FWHM divergence angle is 39 degrees [152]. A molecule path is "successful" if it passes within the

skimmer aperture and through some target detection region. I also include a collimating aperture upstream of the field plates to prevent depositing molecules on the field plates, though this doesn't significantly affect the optimal lens design. Molecules are also marked "unsuccessful" if they hit the inner wall of the focusing element.

I create an array of particle positions, velocities, and arrival times in the source, at the skimmer and other apertures (e.g., 50 K radiation shield), at a sequence of small steps within the lens, at the field plate collimator, at the end of the field plates, and in the detection region. It is convenient to propagate particles by position rather than time steps, though the time to traverse to the next position must be calculated. For each step within the lens, I assume an approximately constant radial acceleration and use the corresponding kinematic equations of evolution. Step sizes are typically ~ 1 mm, or up to 1 cm for rough optimization of a lens geometry. (Either step size is sufficient for all practical purposes.) I index each particle at each position along the beam line to record whether its radial position is valid (e.g., within the relevant collimator).

It is instructive to examine trajectory plots including color-coded "good" trajectories, "bad" trajectories with reasonably small transverse velocities, and trajectories that hit the field plates. Extremely large transverse velocities will never be focused by the lens and might as well not be plotted. The collimator, lens, and field plate boundaries are all shown. See Fig. 6.2.

A minor practical point about trajectory simulation: in principle, there should be a forward (or backward) momentum impulse as a molecule enters or exits a lens, or for that matter moves through a longitudinally inhomogeneous lens, consistent with energy conservation. In practice, including this impulse makes a negligible difference on the trajectories because $v_\parallel \gg v_\perp$ and $\frac{1}{2}mv_\parallel^2 \gg U_{\max}$. I therefore typically "turn off" these forward impulses, but have the option of including them in the simulation.

Adam West's simulations of a magnetic Halbach array found that fields at the walls of the magnets could be $1.5 - 1.9$ T. The Zeeman shift is $-G_\parallel \frac{\Omega}{J(J+1)} \mu_{\mathrm{B}} \mathcal{B} M$, where $|G_\parallel| = g_S \Sigma + g_L \Lambda$[146, Sec. 2.5]. For the $Q^3 \Delta_2$ state, $|G_\parallel| = 2$ and $g_Q(J=2) \equiv G_\parallel \frac{\Omega}{J(J+1)} = \frac{2}{3}$[4].

---

4. Xing Wu has since made careful measurements of $|G_\parallel| = 2.061(4)$. Details are available in the internal ACME lab log[206].
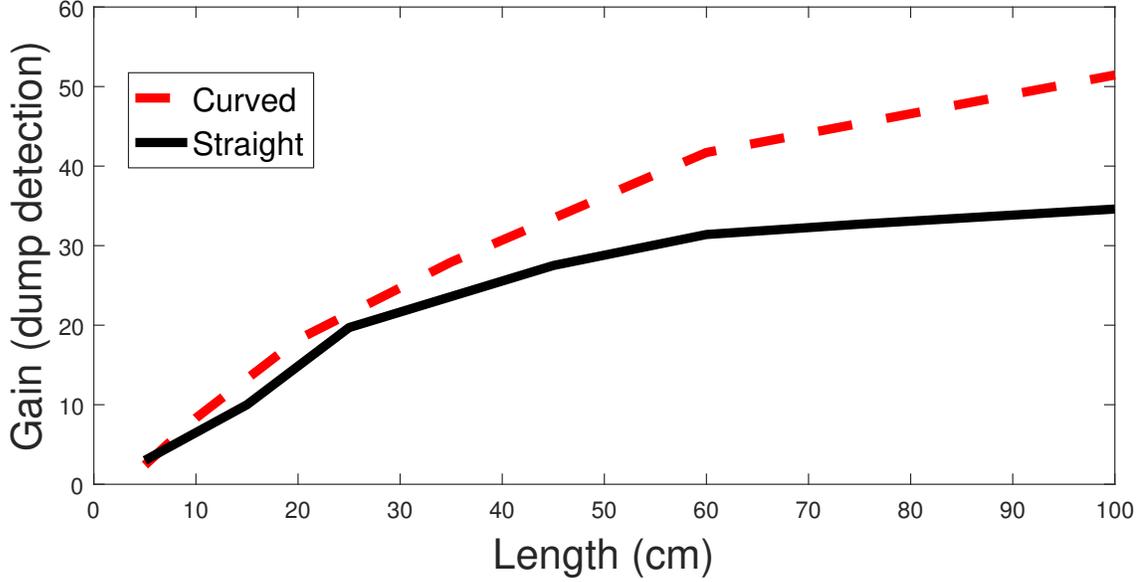
Figure 6.3: Gain with a curved and straight magnetic lens for detection in the dump region, as a function of lens length. The distances between the source and lens entrance, as well as between the lens exit and interaction region, are kept fixed.

For a focusing state of $J = 2$, $M = 2$, the total magnetic moment is $g_Q(J = 2)M = \frac{4}{3}\mu_{\mathrm{B}} \leftrightarrow \frac{4}{3} \times 0.67\,\mathrm{K/T} \approx 0.9\,\mathrm{K/T}$. Assuming we can obtain a magnetic field (averaged over azimuthal angles) of ~ 1.6 T, we could have thus a ~ 1.4 K potential depth. This is what I've assumed in all the following results. In a reasonable regime, the achievable gain scales proportionally with the potential depth, as discussed previously.

Adam West had studied the possibility of using the $X(J = 2, M = 2)$ state for an electrostatic lens, but we have since realized that the linear Stark shift in the Q state, due to the Ω-doublet structure, would allow for a large linear Stark shift at lower applied electric fields. Xing Wu has measured a molecule-frame dipole moment of $D_Q \approx 4.07$ Debye and has estimated an achievable 1.8 K potential depth with an electric hexapole configuration[207]. In this scenario, the gains would be $\approx 1.8/1.4 \approx 30\%$ higher than the values quoted here.

6.1.6 Inhomogeneous potentials

Now that we have discussed the basic approach to simulating the performance of a molecular lens, let us return to the consideration of the case in which the strength of the harmonic potential depends on the longitudinal position within the lens. Any x-dependent potential of

the form $U(x, y) = \frac{1}{2}m\omega(x)^2 y^2$ maintains the relationship $v(x, \theta_1)/\theta_1 = v(x, \theta_2)/\theta_2$ over the evolution of molecular trajectories originating in a point source, for any initial divergence angles θ_1 and θ_2. Thus all trajectories from a point source have vanishing transverse velocity $v_y(x = L, \theta) = 0$ for the same lens length L, for any initial divergence angle θ. This is because at any particular position[5], $a_y = -\omega(x)^2 y$. Therefore $y(t + dt, \theta) = y(t, \theta) + v_y(t, \theta)dt - \frac{1}{2}\omega(x)^2 y(t, \theta)dt^2$, and $v_y(t + dt, \theta) = v_y(t, \theta) - \omega(x)^2 y(t, \theta)dt$ for a small time interval dt. Given that $v_y(t = 0, \theta) \propto \theta$ and $y(t = 0, \theta) \propto \theta$, it follows that at early times $y(dt, \theta) \propto \theta$ and $v_y(dt, \theta) \propto \theta$. This reasoning proceeds iteratively up until any finite time, and we find that $y(t, \theta) \propto \theta$, $v_y(t, \theta) \propto \theta$, where the proportionality constant at any time is generically a complicated function of $\omega(x)^2$. I have confirmed this reasoning (which clearly does not constitute a rigorous proof) by simulating trajectories for a point source with many different functional forms of $\omega(x)^2$, and a focal point is always produced at some distance from the input of the lens.

A straight lens with $\omega^2 \neq \omega(x)^2$ is simplest to construct and understand, so we would have to see a significant improvement in flux to justify using an x-dependent potential. Here, we will explore this possibility. To do so, I assume that in practice we would build a magnetic lens using many longitudinal segments, all of which can produce a maximum B-field at the wall corresponding to a fixed potential depth U_0. The harmonic oscillation frequency varies with the inner radius as $\omega(x)^2 \propto R(x)^{-2}$, so let $\omega(x)^2 \equiv \omega_0^2 \frac{R_0^2}{R(x)^2}$. Note that for an electrostatic lens, the potential depth falls off as $1/R(x)$ (linear Stark shift) or $1/R(x)^2$ (quadratic Stark shift) so that $\omega(x)^2 \propto R(x)^{-3}$ or $\omega(x)^2 \propto R(x)^{-4}$, respectively.

One particularly interesting choice of $R(x)$ satisfies the condition that the most extreme trajectory from a point source, with $\theta = \theta_{\max}$, enters the magnetic lens at the wall and follows the path of the wall until its exit. Since $F = -m\omega(x)^2 r$, this ensures that the restoring force on this trajectory is always maximized; a larger wall radius would reduce the force while a smaller one would obstruct the molecule. We can find the profile of such a lens by solving Newton's second law for the extreme trajectory, $\ddot{r}(t) = \frac{d^2 r}{dx^2} v_x^2 = -\omega_0^2 \frac{R_0^2}{R(x)^2} r(x)$.

---

5. For the conclusions here to hold exactly, we must neglect any acceleration along x that arises due to the x-dependent potential. However, since the x-velocity of trajectories within a molecular lens changes negligibly compared with the forward velocity, this is a reasonable approximation.

Enforcing $R(x) = r(x)$, we obtain the differential equation

$$\frac{d^2 R(x)}{dx^2} = -\frac{\omega_0^2 R_0^2}{v_x^2} \frac{1}{R(x)}.$$

(6.5)

This can be solved in Mathematica with initial conditions $R(0) = l\theta_{\max}$, $R'(0) = v_x \theta_{\max}$. The analytical result is fairly complicated, but can be implemented directly into simulations to generate a lens profile. For reasonable parameters, the resultant trajectories (and likewise, the lens walls) look more or less indistinguishable from parabolic. The entrance aperture has radius $l\theta_{\max}$, and the exit aperture is typically larger. The length of the lens could be approximately optimized by finding the position at which the tangent line of the lens profile intersects the center of the detection region, since a molecule exiting the lens at that position (originating from a point source) would follow the same path.

Under some conditions, using this sort of potential can give $\sim 35\%$ more molecules entering the detection region relative to a lens of constant radius. If we were willing to construct arbitrarily large and complicated lenses, then we should be able to much more nearly saturate the etendue limit[6] with this method. Indeed, the advantage of a curved lens over a straight lens tends to grow with the lens length; see Fig. 6.3.

The advantage of a curved lens over a straight lens is modest only because of engineering constraints. If we want to capture molecules with an initial divergence angle of $\theta = 0.05$ rad, then the entrance aperture (assumed to be a distance of ~ 25 cm from the source) must be ~ 2 cm in radius to provide enough distance for the molecule to turn around. The corresponding length is then $\sim 2\,\text{cm}/0.05 \sim 40$ cm. This is, more or less, the minimal length to remove a single potential depth of transverse energy. A curved lens must therefore be at least ~ 80 cm in order to turn around a significantly higher transverse velocity than a straight lens. While this is possible to construct, it is probably not worthwhile to do so.
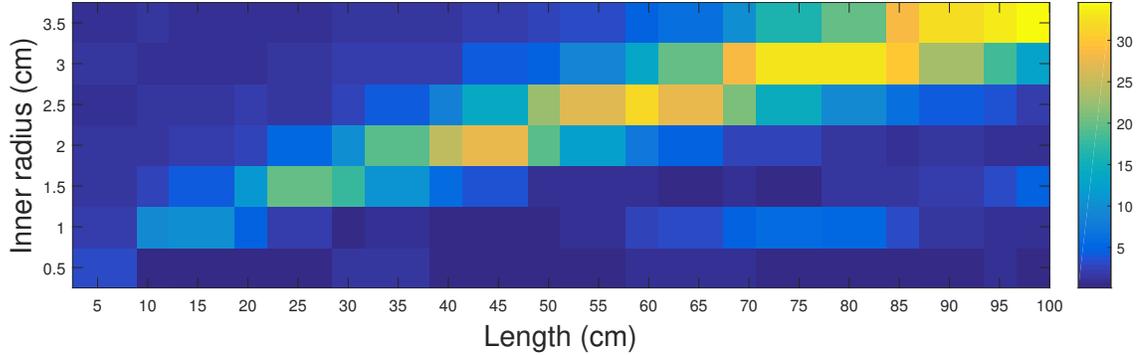
Figure 6.4: Gains for dump region detection with a curved lens. The inner radius of the lens (vertical axis) is proportional to the maximum designed angle θ_c that parametrizes the curved lens geometry, and the horizontal axis gives the length of the lens, up to 1 m. With a 45 cm long lens, gains of 27 (compared to detection in the dump region with no lens) could be obtained if molecules were allowed to hit the field plates. Optimizing the interaction region collimator spacing to prevent collisions with the field plates only reduces the expected gain by $\sim 10\%$, to $G = 24$. Longer lenses have large output apertures, making collimation through the field plates much more difficult.

Location	Type	Length (cm)	$r\ [\theta_c]$	G (no lens)	G (ACME II)
IR	Straight	45	2 cm	10.5	10.5
IR	Curved	35	0.05 rad	13.7	13.7
Dump	Straight	45	2 cm	21.6	7.9
Dump	Curved	35	0.05 rad	26.7	9.7

Table 6.1: Optimized performance with straight and curved lenses, with detection in the IR and dump. Gains over both the no-lens case (for the same geometry, aside from the length of the lens itself) and ACME II levels are given. Curved lenses tend to slightly outperform straight lenses and are more compact for the same gain, but also much more difficult to construct.

Results

I scan over design parameters as in Fig. 6.4 to find expected gains in the following four scenarios: detection in the interaction region with a straight or curved lens, and detection in the dump region with a straight or curved lens. After optimizing the collimator spacing to prevent collisions with the field plates, I obtain the results in Table 6.1. We see that curved lenses tend to somewhat outperform straight lenses (20-30%) and need not be as long. I have found that returns are very diminishing for using much longer lenses because the output molecule cloud is large and it becomes difficult to collimate the beam through the field plates. Note that the optimal design parameters are essentially identical for detection in either the dump or IR since the beam is highly collimated in both cases (see Fig. 6.2).

I have also considered the possibility of a non-imaging focusing element, which does not significantly outperform the "imaging" molecular lens. See Appendix E for details.

It is essential to realize that the results summarized in Table 6.1 apply under the assumptions of a 1.4 K potential depth lens and perfect transfer into and out of the molecular focusing state. This latter assumption is certainly incorrect. Provisionally assuming a $\approx 80\%$ transfer efficiently in each step, the gains computed here should be reduced by a factor of ≈ 0.65. For example, the straight lens with detection in the interaction region would have a gain of approximately 6.8, when state transfer inefficiencies are included. Furthermore, certain geometric assumptions made here could turn out to be merely approximate. For example, as mentioned previously, Xing Wu is exploring the possibility of performing rotational cooling in the beam source vacuum chamber, which would allow the lens to be placed closer to the ablation cell, further improving the gain over the ACME II molecular flux.

6.2 Modeling the ablation source

In ACME, we have always assumed an effective molecular source with a diameter of 7 mm at a "zone of freezing," after which no collisions occur, 2 cm downstream of the cell aperture.

---

6. Namely, the maximum possible gain of 150, based on the volume of the region in phase-space (i.e., the etendue) that evolves from the entrance of the field plates to the exit without colliding with the walls.

However, these parameters have never been measured very precisely before. In ACME III, we could conceivably obtain an improved molecular flux by expanding the skimmer radius, provided there are transversely slow molecules just outside the current skimmer radius of 3 mm. The distance between the zone of freezing and skimmer is assumed to be ~ 1 cm.

6.2.1 Effective source model

In order to characterize the source, we will use a very simple model (which is consistent with, but not strongly supported by, data presented later in this section). Assume that at some position x_0 along the beam line, collisions freeze out. Beyond this position, trajectories are entirely ballistic. The transverse velocities $v_y, v_z \sim N(0, \sigma_v^2)$ are normally distributed with zero mean and standard deviation σ_v. Consistent with a Maxwell-Bolzmann distribution, velocities in orthogonal directions are assumed to be uncorrelated with each other, $\mathrm{cov}(v_y, v_z) = 0$. The transverse spatial distribution is also modelled as a normal distribution, $y, z \sim N(0, \sigma_r^2)$, again with no correlation between orthogonal axes. Upstream of the zone of freezing, trajectories are "scrambled" by collisions, so we assume that $v_{y(z)}$ and $y(z)$ are uncorrelated with each other at the zone of freezing. This last assumption is especially unlikely to hold, since contours of constant density are curved and any boosting should happen perpendicular to these contours. However, the assumption that positions and velocities are uncorrelated at the zone of freezing gives reasonable results and is much simpler to work with.

Downstream of the zone of freezing, both transverse positions y and z evolve completely independently. I will discuss the dimension y for concreteness. At a distance $x - x_0$ downstream of the zone of freezing, the transverse position and velocity are given by $y \to y' = y + \frac{(x-x_0)}{v_x} v_y$ and $v_y \to v_y' = v_y$. In other words, y' and v_y' are linear combinations of the independent normally distributed variables y and v_y (assuming a fixed velocity v_x along the beam axis). By definition, y' and v_y' are therefore jointly normal random variables. It turns out that the full distribution of a pair of jointly normal random variables, each of which have vanishing mean, is characterized by only $\sigma_{y'}^2, \sigma_{v_y'}^2$, and the correlation coefficient ρ between y' and v_y'[208, Sec. 4.7]. We can directly compute $\sigma_{v_y'}^2 = \sigma_{v_y}^2$ and

$\sigma_{y'}^2 = \sigma_y^2 + \frac{(x-x_0)^2}{v_x^2}\sigma_{v_y}^2$. Further,

$$\rho = \left[1 + \frac{v_x^2}{(x-x_0)^2}\frac{\sigma_y^2}{\sigma_{v_y^2}}\right]^{-1/2}. \tag{6.6}$$

The resulting bivariate normal distribution has several nice properties. For example, $\int dy f(y, v_y) \equiv f(v_y)$ is a normal distribution in the velocity along y with variance $\sigma_{v_y}^2$, and analogously for $f(y)$. This will be important for our measurement of $\sigma_y^2, \sigma_{v_y}^2$, and $\rho(x_{\text{skimmer}})$, which determines everything we'd like to know about the effective source at the skimmer position (to the extent that this model accurately describes the molecular trajectories at the "zone of freezing").

6.2.2 Measurements

We used a 512 nm laser beam driving the $X \to I$ transition to perform Doppler scans and absorption measurements at many positions in the beam box. The laser is split into two paths. One beam probes the in-cell absorption just in front of the cell nozzle and the other beam is translated outside the cell over a 2×1 cm area in the $x \times y$ directions. We perform out-of-cell Doppler scans on-axis every 3 mm between the nozzle and skimmer. In addition, we measure the out-of-cell density, normalized to in-cell production, every 2 mm vertical step and 3 mm horizontal step along the beam axis. We have found that the peak density occurs at the cell nozzle and is transversely narrow there, as expected. Farther from the cell, the density along the beam line becomes smaller and the beam becomes transversely larger. Most of the analysis code to extract densities from the raw data was written by Xing Wu.

The beam width and transverse velocity width extracted from the density scans and Doppler scans, respectively, are shown in Fig. 6.5. As expected, the transverse velocity width rapidly increases just outside the cell, as the molecules are boosted out of the cell nozzle, but levels off for $x \gtrsim 1$ cm. Unfortunately, the signal-to-noise ratio from the absorption data significantly deteriorates for $x > 15$ mm. Perhaps counterintuitively, however, the transverse spatial width of the beam does *not* rapidly increase between $x \approx 1$ cm and $x \approx 2$ cm. This is in fact expected from our model, in which $\sigma_y^2(x) = \sigma_y^2(x_0) + \frac{(x-x_0)^2}{v_x^2}\sigma_{v_y}^2$.
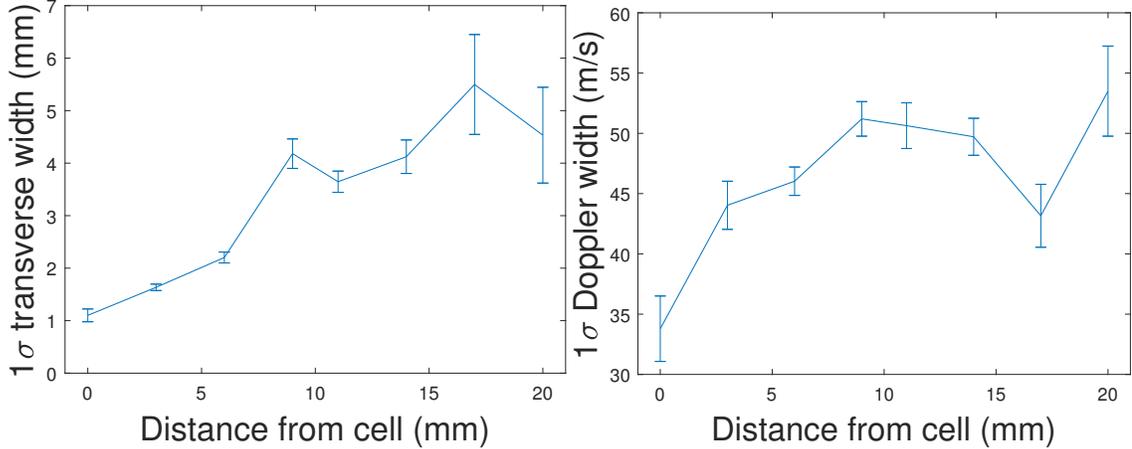
Figure 6.5: Measured beam width (left) and transverse velocity width (right) at each x-position. Transverse spatial widths are obtained from gaussian fits to absorption data as a function of the vertical position of the absorption laser, for each longitudinal position. Doppler widths are obtained from gaussian fits to laser frequency scans with the absorption laser vertically centered on the molecular beam, for each longitudinal position. Both plots are expected to be nearly flat between $x \approx 1$ cm and $x \approx 2$ cm; see discussion in main text.

Note that for $v_x \approx 200$ m/s, $\sigma_{v_y} \approx 50$ m/s, $\sigma_y(x_0) \approx 4$ mm, and $x_0 \approx 1$ cm, we anticipate $\sigma_y(20\,\text{mm}) \approx 4.7$ mm, indistinguishable from $\sigma_y(10\,\text{mm}) \approx 4$ mm within the noise of the measurement. On the other hand, sufficiently far downstream (in this case, $x \gtrsim 3$ cm), our model predicts a much more dramatic dependence of σ_y on the longitudinal position, $\frac{d\sigma_y}{dx} \to \frac{\sigma_{v_y}}{v_x}$, corresponding to a FWHM divergence of[7] $2.355\frac{\sigma_{v_y}}{v_x} \approx 33°$. This far-downstream prediction may match our intuition better than the intermediate regime, which occurs near the zone of freezing.

Both transverse widths (namely, spatial and velocity) increase rapidly until around 1 cm and then level off, suggesting the longitudinal position of the zone of freezing is $x_{\text{zof}} \approx 1\,\text{cm}$. The actual "zone of freezing" is not necessarily a plane at a well-defined longitudinal position, so I won't take great pains to define it rigorously. Combining the data at $x =$9, 11, 14, 17, and 20 mm, we obtain a downstream Doppler width of $1\sigma = 50 \pm 1$ m/s. The beam size, of course, is expected to continue growing past the zone of freezing due to the transverse velocity width, but as we just explained, it only does so only negligibly very near the zone of freezing according to both our model and the data. I'll therefore take

---

7. Nick Hutzler has measured $\sigma_{v_y} \approx 75$ m/s, with correspondingly larger FWHM divergence angles, in a different molecular beam source and with a different ablation cell geometry [146, p. 3.3.4]
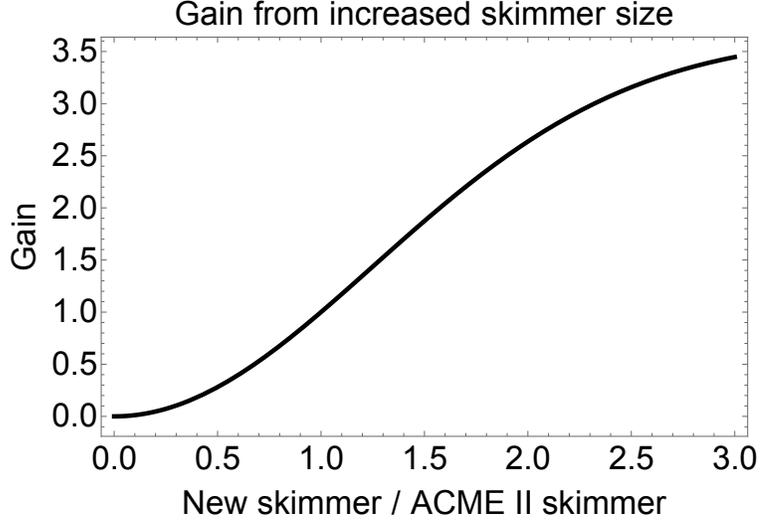
Figure 6.6: Gain over ACME II as a function of the fractional increase in skimmer size. Doubling the skimmer size gives a gain of approximately 2, while increasing the skimmer size by a factor of 2.5 increases the flux by over a factor of 3. This does not include effects of the lens, and assumes positions and velocities are uncorrelated at the zone of freezing.

the size of the beam "at the zone of freezing" to be the weighted average of measured sizes at 9 and 11 mm, giving $\sigma_r(x = x_{\text{zof}}) = 3.8 \pm 0.2$ mm. Note that $2\sigma_r(x = x_{\text{zof}})$ is fairly close to the 7 mm usually quoted as the "size" of the source.

6.2.3 Potential gains from increased skimmer size

In the ACME II geometry, we can only currently accept molecules with a transverse velocity of approximately 2 m/s or less. With a magnetic lens, this acceptance could increase to about 8 m/s. Over a distance of 1 cm, a molecule with transverse velocity of even 10 m/s increases its transverse position by only 0.5 mm, assuming a 200 m/s forward velocity. Therefore, to a good approximation, the molecules of interest to us remain at the same transverse position between the zone of freezing and the skimmer. Currently, the skimmer has a radius of 3 mm, or $\sigma_r(x = x_{zof})/1.3$. As a result, if we assume that transverse positions and velocities are uncorrelated at the zone of freezing, then about 73% of low-divergence molecules are currently removed from the beam by the skimmer[8]. Recovering the 73% of transversely slow molecules would constitute a gain of about $1/0.27 = 3.7$. With a skimmer

---

8. To perform these calculations, remember that the radial population distribution is $f(r) \propto r \exp(-r^2/(2\sigma_r^2))$, rather than a normal distribution as in the transverse cartesian directions.

radius of about $2.5\sigma_r(x = x_{zof}) = 9.5$ mm, we would expect a signal gain of greater than 3. This would come at a potentially very significant engineering cost of pumping the increased neon background gas downstream of the skimmer. The large skimmer vs. small skimmer gain is borne out by trajectory simulations with no lens included, for the model considered in this section. See Fig. 6.6 for the expected gain vs. skimmer size, normalized to ACME II.

In the most pessimistic model, consider instead the case that the positions and velocities are maximally correlated at the zone of freezing, so the molecules with $v_\perp = \sigma_v/5 = 10$ m/s lie at transverse positions of $r = \sigma_r/5 = 0.8$ mm. (Recall that we can't easily infer the correlation between position and velocity, since we measure only $f(v_z)$ and $f(y)$, with no simultaneous information about position and velocity along a single axis.) These molecules are already admitted to the IR, so expanding the skimmer would not give a gain in flux. More realistically, we might (or might not) have a partial correlation between positions and velocities due to the boosting effect in the buffer gas beam, so a gain somewhere between these two models–in which the positions and velocities are either fully correlated or not correlated at all–is expected. Which case is more realistic is best determined by measurements with a larger skimmer, of course, but we could also refine our model by comparing to direct simulation Monte Carlo trajectories. We have not yet made measurements of the downstream flux using a larger skimmer diameter to distinguish between these two cases.

6.3   Silicon photomultipliers for improved quantum efficiency

One of the significant improvements of ACME II over ACME I was the new detection wavelength (512 nm vs. 690 nm), where PMTs with higher quantum efficiency could be obtained. In particular, the Hamamatsu R7600U-300 used in ACME II has a quantum efficiency of $\approx 25\%$ at 512 nm. We have made some preliminary investigations into using silicon multipliers (SIPMs) to achieve a quantum efficiency of 40-50%, for a molecule-flux-equivalent gain between a factor of 1.5 and 2. In this section, I will briefly describe the various considerations involved in changing the photodetection technology used in ACME.

One primary impediment to using silicon photomultipliers rather than PMTs is the

	$\epsilon_{\text{det.}}$	F	$\epsilon_{\text{det.}}/F$	Dark p.e. rate (cps)
PMT	0.25	1.2	0.20	$\sim 6 \times 10^3$
Silicon photomultiplier	0.4	1	0.4	$\sim 9 \times 10^7$
Avalanche photodiode	0.80	2	0.4	$\sim 2 \times 10^9$

Table 6.2: Approximate photodetection efficiencies and excess noise factor. Dark photoelectron (p.e.) count rates are given in counts per second (cps) at room temperature and scaled for ≈ 1 in^2 detector areas. Typical values are merely representative; see [209–211] for full data sheets.

higher "dark current," an electronic signal indistinguishable from photoelectrons that originates due to thermal processes rather than molecular fluorescence. The time-averaged dark current is associated with current shot noise, analogously to a photoelectron current. If the dark current exceeds the fluorescence signal current, then the signal-to-noise ratio (SNR) will be dominated by the dark current shot noise rather than photoelectron shot noise.

We will return to this issue of the dark current shortly. For the moment, consider the case that the dark current is negligible, either because it is reduced through some technical effort or because it is dominated by the signal current. In this case, the photon shot noise limit is determined by the total number of detected photons: the uncertainty in the measured phase is $\sigma_\phi = (2|\mathcal{C}|\sqrt{N \times \epsilon/F})^{-1}$, where N is the total number of molecules, ϵ is the photon detection efficiency, and $F \geq 1$ is the photodetector "excess noise factor" arising from the statistical nature of the photodetector's intrinsic gain process. Here, ϵ includes contributions from both finite optical collection efficiency (e.g., due to light collection from a region with less than 4π solid angle) and from the finite photon detection efficiency of the photodetector itself: $\epsilon = \epsilon_{\text{coll.}} \times \epsilon_{\text{det.}}$.

For light collection at 512 nm, the photon detection efficiency is up to $\epsilon_{\text{det.}} \approx 0.25$ as stated previously. On the other hand, silicon photomultipliers can have up to $\epsilon_{\text{det.}} \approx 0.5$ if operated at significant "overvoltage" (explained below), and avalanche photodiodes (APDs) have significantly higher efficiencies, $\epsilon_{\text{det.}} \approx 0.8$. However, it is important to consider the effect of excess noise factors, which vary by up to a factor of 2 among different types of detector. For PMTs, $F \approx 1.2$; for silicon photomultipliers, $F \approx 1$; for avalanche photodiodes, $F \approx 2$. The origin of these excess noise factors is discussed in detail in Appendix F.

See Table 6.2 for a representative summary of the efficiency, excess noise, and dark

count rate of the photodetectors we've considered. Due to the combination of photodetector efficiency $\epsilon_{\text{det.}}$ and excess noise factor F, the potential shot-noise limit from silicon photomultipliers and avalanche photodiodes is comparable. However, the dark current from avalanche photodiodes is approximately an order of magnitude larger than that of silicon photomultipliers (and both are several orders of magnitude larger than for PMTs). As a result, we decided to investigate the possibility of switching from PMTs to SiPMs, but not to APDs.

6.3.1 SiPM cooling

However, even the SiPM has a higher dark current (and associated dark current shot noise) than the ACME II signal level, which is $\sim 3 \times 10^5$ detected photons per pulse, and ~ 2 ms of useful signal in each pulse, for signal count rates of $\sim 10^8$ photoelectron counts per second (cps). With 8 detectors, this corresponds to $\sim 10^7$ cps/detector. As we see in Table 6.2, the typical dark current for a 1 in^2 area SiPM detector is 9×10^7 cps, about an order of magnitude larger than the signal per detector in ACME II. If the signal in ACME III is $\sim 100\times$ as large as in ACME II, as anticipated, then this dark current is unlikely to limit the signal-to-noise ratio we can achieve. However, if the signals are not actually this large, then it may be necessary to reduce the dark current.

The solution to the high dark current is to cool the detector, which reduces the dark current by a factor of 2 for every $7 - 10°$ C[212–215]. Thus we can expect the SiPM dark current to be overwhelmed by ACME II signals provided the detectors are cooled below $\approx -15°$ C. Unfortunately, because this temperature lies below the dew point (and indeed below the freezing point of water), the SiPMs must be operated in either vacuum or an environment flushed with dry nitrogen gas, in order to prevent condensation and ice accumulation on the detector surfaces.

Because SiPM devices are constructed to be have dimensions at most 6 mm $\times$ 6 mm, we will need to use an array of SiPM devices in order to obtain ≈ 1 in^2 detection areas. We prefer to use manufacturer-provided (likely SensL or Hamamatsu) PCB boards to power and read out the signals from these arrays rather than design our own (which would be significantly more complex than operating a single detector). However, PCB is an electric

insulator, so direct cooling of an out-of-the-box SiPM array using a thermo-electric cooler (TEC) would be difficult. For this reason, Adam West designed a small vacuum chamber where we can directly cool a small copper box with a TEC to $\approx -20°$ C. The SiPM detector can be operated inside this copper box and radiatively cooled. Light can be conveyed to the SiPM via a light pipe that enters a hole in the top of the cold copper box. This project was further developed by Shadi Fadaee, an undergraduate working briefly in the group, but the cooling of the box and corresponding temperature-dependent SiPM properties have not yet been well characterized.

As mentioned previously, we expect that ACME III will have higher signal count rates by a factor of ~ 100 compared to ACME II[9]. As a result, it might not be necessary to cool SiPMs at all, and if some cooling is necessary then it might not be necessary to cool below the dew point, significantly simplifying the technical inconvenience of cooling the detectors. However, due to a likely redesign of the collection optics for optical cycling detection[10] (one of the most important statistical upgrades being pursued for ACME III), it may be necessary to increase the detector area by a non-negligible factor (~ 3). If this is the case, then the number of detectors used will increase, and the signal per detector will correspondingly decrease. Thus it is reasonable to expect that significant cooling might still be required in ACME III if SiPM arrays replace PMTs for photon detection. However, as of now, it is still unclear whether any cooling at all would be required.

6.3.2 Amplifier electronics

The other technical challenge of using SiPMs instead of PMTs, in addition to that of the dark current discussed previously, is that SiPMs have a large intrinsic capacitance, which makes low-noise signal amplification difficult. For example, a SensL $24\,\mathrm{mm} \times 24\,\mathrm{mm}$ J-

---

9. The increased count rate would arise as follows: $10\times$ from the molecular lens, $10\times$ from optical cycling detection, $2\times$ from improved quantum efficiency of silicon photomultipliers. If one or several of these improvements do not pan out, then an order-of-magnitude increase in statistical sensitivity would require only $\sim 30\times$ higher count rates due to the demonstrated suppression of excess noise from ACME II, which increases the signal-to-noise ratio by $\sim \sqrt{3}$ with no increase in the signal count rate.

10. In particular, we may change our detection protocol to detect fluorescence in a "dump" region, downstream of the interaction region, where it may be easier to optically cycle photons without facing critical limitations in sensitivity due to a background of scattered light. The fluorescence collection would be designed from scratch in this case.

series array has a total capacitance of 16 nF[210]. Given this, suppose that we were to measure the signal voltage drop across a resistor R, due to the signal current I. Suppose we want to maximize R subject to the constraint that the detection bandwidth is 5 MHz, a factor of 3.6 greater than the characteristic decay rate of the electronic I state (not to be confused with the signal current). This bandwidth therefore allows us to resolve molecular dynamics during the experimental readout. Then the maximum viable resistance is only $R = (2\pi \times 5\,\text{MHz} \times 16\,\text{nF})^{-1} = 2\,\Omega$.

Now consider the noise on this measurement. The current shot noise, measured in units of $\text{A}/\sqrt{\text{Hz}}$, is $\delta I = G\sqrt{2q_e I_0}$, where I_0 is the unamplified photoelectron current and $G \approx 6 \times 10^6$ is the intrinsic gain factor of the SiPM. For a photoelectron count rate of 10^7 cps on a single $\approx 1\,\text{in}^2$ detector, $I_0 = 1.6 \times 10^{-12}$ A and $\delta I = 4\,\text{nA}/\sqrt{\text{Hz}}$. The corresponding voltage noise, measured across a resistance of $2\,\Omega$, is then $\delta V = 8\,\text{nV}/\sqrt{\text{Hz}}$.

We would like this voltage noise to be small compared to the input voltage on a subsequent voltage amplifier. For example, the SRS SR445A has an input voltage noise of $6.4\,\text{nV}/\sqrt{\text{Hz}}$, quite comparable to the signal shot noise from an individual SiPM array. Since all noise sources are added in quadrature, we would like the amplifier electronic noise to be at least $\sim 3\times$ smaller than the signal shot noise. As a result, it is necessary to have a larger effective resistance, or transimpedance, without sacrificing additional bandwidth.

This is a fairly standard problem; the standard solution is a transimpedance amplifier, in which the effective resistance (or transimpedance) that converts the signal current to a voltage is set by the negative feedback resistor on an op-amp circuit. In a naive analysis, the output of the transimpedance amplifier is determined only by the resistance R in the op-amp feedback network, but not by the capacitance of the photodetector. As a result, the transimpedance R can be set to relatively large values (e.g., $R \geq 50\,\Omega$) and the corresponding output voltage noise $\delta V = R\,\delta I$ can be increased far beyond the input voltage of a subsequent amplifier stage. But this would be too good to be true; the photodetector's capacitance can contribute to instability of the op-amp, rendering the circuit useless if not dealt with carefully. Some useful discussions can be found in [216–220].

The challenge in our application is that the photodetector capacitance is relatively large, and the bandwidth requirements are also relatively high (compared to typical tran-

simpedance applications). I designed a transimpedance amplifier based on the high-gain-bandwidth-product op-amp OPA847 ($f_{\mathrm{GBW}} = 5$ GHz), which Xinyi Chen has built and tested. Unfortunately, the results were somewhat inconsistent across measurement conditions, and the circuit's voltage output often oscillates. We subsequently identified a different circuit design that was recently developed for comparable use requirements (detector capacitance of ≈ 30 nF, bandwidth of ≈ 4 MHz, transimpedance $R = 3.9\,\mathrm{k}\Omega$, and transimpedance amplifier noise of $0.7\,\mathrm{nV}/\sqrt{\mathrm{Hz}}$)[221, 222]. This design was tested in our lab and found to meet all these specifications, which are sufficient for our needs. Going forward, we will rely on this demonstrated design to ensure stability of the SiPM amplifier electronics.

Note that, as with the cooling requirements, in ACME III the circuit requirements might be relaxed due to the higher shot-noise-limited current I_0, and correspondingly larger voltage noise δV for a given transimpedance R. However, this benefit might once again be mitigated if a larger total detector area is needed for photon cycling detection.

One last hurdle to using SiPM arrays in place of PMTs is that the gain on each SiPM detector is sensitive to the overvoltage and temperature. As a result, the supplied overvoltage and temperature will both have to be carefully stabilized in order to avoid excess noise associated with fluctuations in the gain between detectors. To date, we have not made any serious effort to design or test systems to achieve this high level of stability.

In summary, the SiPM could provide a factor of ≈ 2 increased effective photodetection efficiency $\epsilon_{\mathrm{det.}}/F$ compared to PMTs. However, SiPMs are best operated in a higher-signal regime, where dark current shot noise and amplifier noise are negligible compared to the signal shot noise. These difficulties can be overcome by cooling the detectors and by careful electronic design, both of which should be feasible. Alternatively, they are automatically overcome by sufficiently increasing the signal count rate. Therefore, we can reasonably expect SiPMs to be a more straightforward alternative to PMTs in ACME III, where the photoelectron count rate will be larger by a factor of ~ 100, compared to ACME II.

6.4 Phase measurement with optical cycling

In ACME II, only $\epsilon \approx 0.05$ of photons emitted in the detection region are actually detected, due to imperfect optical collection of photons ($\approx 20\%$) and quantum efficiency of the PMTs ($\approx 25\%$). Thus, for a quantum-projection-limited measurement with phase sensitivity $\sigma_\phi = \frac{1}{2\mathcal{C}\sqrt{N\epsilon}}$, where N is the total number of molecules and $\mathcal{C}$ is the measurement contrast, achieving perfect photon detection ($\epsilon = 1$) would provide the same improvement in sensitivity as a 20-times greater flux of molecules.[11]

The obvious ways to increase the total photon detection efficiency, ϵ, are to improve the optical collection (which was indeed done between ACME I and ACME II, by optimizing the lens configurations collecting fluorescence and by replacing fiber bundles with thick light pipes for a higher effective optical collection area) and to use a higher-quantum-efficiency detector (the subject of Sec. 6.3). There is, however, another method that is complementary to these two: to detect fluorescence from an optical cycling transition so that many photons are scattered from each molecule. Suppose, for instance, that exactly n photons are emitted by each molecule. As long as the probability to detect each molecule is still small, i.e., $n\epsilon \ll 1$, then we may expect the signal-to-noise ratio (SNR) to improve as $\sigma_\phi \approx \frac{1}{2\mathcal{C}\sqrt{N \times n\epsilon}}$. Obviously, this result will not hold when $n\epsilon > 1$, since then the phase sensitivity would be better than the quantum projection limit. In this section, I will consider a relatively general framework for computing the phase sensitivity σ_ϕ for a phase measurement via laser-induced fluorescence (LIF) detection with an optical cycling transition, for a realistic model of a partially open optical cycling process and finite photon detection efficiency [12]. In the following discussion, we let $\mathcal{C} \to 1$.

6.4.1 Phase measurements with LIF detection

The approach we will use to determine the effect of optical cycling on a precision phase measurement using LIF detection extends beyond the narrow domain of the ACME exper-

---

11. We will see that this estimate should actually be modified by the excess noise factor $f \approx 1.25$ of the PMTs used in ACME II, so that the maximum possible improvement in effective molecular flux due to improved detection efficiency would be a factor of 25.

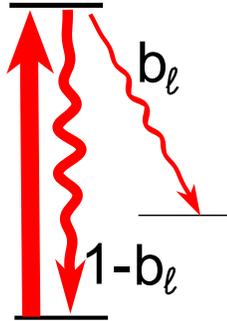12. The core of this work has been published in [223].

Figure 6.7: Simple schematic of an optical cycling transition, with probability of decay b_ℓ to a dark state, which is not addressed by any laser, and decay probability of $1 - b_\ell$ to the ground state, which is subsequently driven again to a short-lived excited state.

iment. Many atomic and molecular experiments use the same basic approach as ACME; namely, a quantum state is read out by LIF, in which population is driven to a short-lived state and the resulting fluorescence photons are detected. As described previously, due to geometric constraints on optical collection and technological limitations of photodetectors, the majority of emitted photons are typically undetected, reducing the experimental signal. Optical cycling transitions can be exploited to overcome these limitations, by scattering many photons per particle (atom or molecule); see Fig. 6.7. In the limit that many photons from each particle are detected, the signal-to-noise ratio (SNR) may be limited by the quantum projection (QP) noise, also often referred to as atom or molecule shot noise. LIF detection with photon cycling is commonly used in ultra-precise atomic clock [224, 225] and atom interferometer [226] experiments to approach the QP limit.

Molecules possess additional features, beyond those in atoms, that make them favorable probes of fundamental symmetry violation [119, 122, 128, 132, 133, 227] and fundamental constant variation [225, 228–232], as well as promising platforms for quantum information and simulation [233–237]. Many molecular experiments that have been proposed, or which are now being actively pursued, will rely on optical cycling to enhance measurement sensitivity while using LIF detection [119, 128, 132, 133, 227, 232]. Due to the absence of selection rules governing vibrational decays, fully closed molecular optical cycling transitions cannot be obtained: each photon emission is associated with a non-zero probability of decaying to a dark state that is no longer driven to an excited state by any lasers. However, for some molecules many photons can be scattered using a single excitation laser, and up

to $\sim 10^6$ photons have been scattered using multiple repumping lasers to return population from vibrationally excited states into the optical cycle. This has enabled, for example, laser cooling and magneto-optical trapping of molecules [238–240]. Furthermore, some precision measurements rely on atoms in which no simply closed optical cycle exists [96, 111]; our discussion here will be equally applicable to such species.

These considerations motivate a more careful, general study of LIF detection for precision measurement under the constraint of imperfectly closed optical cycling. Some consequences of loss during the cycling process have been considered in [241]. However, the effect of the statistical nature of the cycling process on the optimal noise performance has not been previously explored. In particular, the number of photons scattered before a particle decays to an unaddressed "dark" state, and therefore ceases to fluoresce, is governed by a statistical distribution rather than a fixed finite number. We will see that due to the width of this distribution, a naive cycling scheme inflates the noise above the QP limit. In particular, we find that in addition to the intuitive requirement that many photons from every particle are detected, to approach the QP limit it is also necessary that the probability of each particle exiting the cycling transition (via decay to a "dark" state outside the cycle) is negligible during detection. If this second condition is not satisfied, so that each particle scatters enough photons that it is very likely to have been optically pumped into a dark state, then the SNR is decreased by a factor of $\sqrt{2}$ below the QP limit.

6.4.2 Model of the phase measurement

Analogously to the description in Sec. 2.2.2, consider an ensemble of N particles in an effective two-level system, in a state of the form

$$|\psi\rangle = \frac{e^{-i\phi}|\uparrow\rangle + e^{i\phi}|\downarrow\rangle}{\sqrt{2}}. \tag{6.7}$$

The relative phase ϕ is the quantity of interest in this discussion. It can be measured, as seen previously, by projecting the wavefunction onto an orthonormal basis

$$\{|X\rangle \propto |\uparrow\rangle + |\downarrow\rangle, |Y\rangle \propto |\uparrow\rangle - |\downarrow\rangle\} \tag{6.8}$$

such that $|\langle X|\psi\rangle|^2 = \cos^2(\phi)$ and $|\langle Y|\psi\rangle|^2 = \sin^2(\phi)$. In the LIF technique, this can be achieved by driving state-selective transitions, each addressing either $|X\rangle$ or $|Y\rangle$, through an excited state that subsequently decays to a ground state and emits a fluorescence photon. This light is detected, and the resulting total signals, S_X and S_Y, are associated with each state. The measured value of the phase, $\tilde{\phi}$, is computed from the observed values of S_X and S_Y. In the absence of optical cycling, the statistical uncertainty of the phase measurement is $\sigma_{\tilde{\phi}} = \frac{1}{2\sqrt{N\epsilon}}$, where the photon detection efficiency lies in the range $0 < \epsilon \leq 1$. Note that $N\epsilon$ is the average number of detected photons; hence, this result is often referred to as the "photon shot noise limit." In the ideal case of $\epsilon = 1$, the QP limit $\sigma_{\tilde{\phi}} = \frac{1}{2\sqrt{N}}$ is obtained. This scaling is derived as a limiting case of our general treatment below, where the effects of optical cycling are also considered.

We suppose that the phase is projected onto the $\{|X\rangle, |Y\rangle\}$ basis independently for each particle. Repeated over the ensemble of particles, the total number of particles N_X projected along $|X\rangle$ is drawn from a binomial distribution, $N_X \sim B(N, \cos^2 \phi)$, where $x \sim f(\alpha_1, \cdots, \alpha_k)$ denotes that the random variable x is drawn from the probability distribution f parametrized by $\alpha_1, \cdots, \alpha_k$, and $B(\nu, \rho)$ is the binomial distribution for the total number of successes in a sequence of ν independent trials that each have a probability ρ of success. Therefore, $\overline{N_X} = N \cos^2 \phi$ and $\sigma_{N_X}^2 = N \cos^2 \phi \sin^2 \phi$, where $\bar{x}$ is the expectation value of a random variable x and σ_x is its standard deviation over many repetitions of an experiment. We define the number of photons scattered from the i-th particle to be n_i, where a "photon scatter" denotes laser excitation followed by emission of one spontaneous decay photon, and define $\overline{n_i} = \bar{n}$ (the average number of photons scattered per particle) and $\sigma_{n_i} = \sigma_n$. Note that these quantities are assumed to be the same for all particles (i.e., independent of i). We will evaluate $\bar{n}$ and σ_n for a realistic model of an optical cycling process in Sec. 6.4.4. We define d_{ij} to be a binary variable indexing whether the j-th photon scattered from the i-th particle is detected. Therefore, $d_{ij} \sim B(1, \epsilon)$, and it follows that $\overline{d_{ij}} = \epsilon$ and $\sigma_{d_{ij}}^2 = \epsilon(1 - \epsilon)$. The gain g of each photoelectron is then drawn from a distribution with mean G and standard deviation σ_G. Conventionally, a photodetector's variance is characterized by its excess noise factor $f \equiv 1 + \frac{\sigma_G^2}{G^2}$. We define g_{ij} to be the signal gain of the j-th photon scattered from the i-th particle if that photon is detected. In the case that

the j-th photon scattered from the i-th particle is not detected, g_{ij} is undefined.

We define the signal of the measurement of a particular quadrature $|X\rangle$ or $|Y\rangle$ from the ensemble, when projecting onto that quadrature, to be the total number of signal electrons flowing from the photodetector. For example, the signal S_X from particles projected along $|X\rangle$ is

$$S_X = \sum_{i=1}^{N_X} \sum_{j=1}^{n_i} d_{ij} g_{ij}. \tag{6.9}$$

Explicitly, among N total particles, N_X are projected by the excitation light onto the $|X\rangle$ state and the rest are projected onto $|Y\rangle$. The i-th particle projected onto $|X\rangle$ scatters a total of n_i photons, and we count each photon that is detected (in which case $d_{ij} = 1$), weighted by the detector gain g_{ij}. The right-hand side of Eq. 6.9 depends on ϕ implicitly through N_X, and we use this dependence to compute $\tilde{\phi}$, the measured value of ϕ. Because N_X, n_i, d_{ij}, and g_{ij} are all statistical quantities, the extracted value $\tilde{\phi}$ has a statistical uncertainty. The QP limit is achieved when the only contribution to uncertainty arises from N_X due to projection onto the $\{|X\rangle, |Y\rangle\}$ basis.

6.4.3 Calculation of statistical moments

We now turn to the question of how to compute the average phase and uncertainty in the phase, given the physical model just described. It will be necessary to compute the average signals $\overline{S_X}$ and $\overline{S_Y}$, as well as the variances $\sigma_{S_X}^2$ and $\sigma_{S_Y}^2$ in order to propagate uncertainty to the measured phase.

We can compute $\overline{S_X}$ by repeated application of Wald's lemma ([242, 243]), $E[\sum_{i=1}^{m} x_i] = E[m]E[x_1]$, where $E[x] \equiv \bar{x}$ will prove to be a convenient alternative notation in some of the long derivations ahead. I will also define the number of *detected* photons emitted by the i-th projected molecule to be $n_i^{(d)} = \sum_{j=1}^{n_i} d_{ij}$. Then

$$
\begin{aligned}
E[S_X] &= E\left[\sum_{i=1}^{N_X}\sum_{j=1}^{n_i} d_{ij}g_{ij}\right]\\
&= E\left[N_X\right]E\left[\sum_{j=1}^{n_1} d_{1j}g_{1j}\right]\\
&= E\left[N_X\right]E\left[\sum_{j=1}^{n_1^{(d)}} g_{1j}\right]\\
&= E\left[N_X\right]E\left[n_1^{(d)}\right]E\left[g_{11}\right]\\
&= E\left[N_X\right]E\left[\sum_{j=1}^{n_1} d_{1j}\right]E\left[g_{11}\right]\\
&= E\left[N_X\right]E\left[n_1\right]E\left[d_{11}\right]E\left[g_{11}\right]\\
&= N\cos^2\phi\,\bar{n}\epsilon G.
\end{aligned}
\tag{6.10}
$$

That is, the expected signal from projecting onto the $|X\rangle$ state is (as could be anticipated) simply the product of the average number of particles in $|X\rangle$, $N\cos^2\phi$, the number of photons scattered per particle, $\bar{n}$, the probability of detecting each photon, ϵ, and the average gain per photoelectron, G.

We compute the variance in S_X by repeated use of the law of total variance [244], $\mathrm{Var}[a] = E[\mathrm{Var}[a|b]] + \mathrm{Var}[E[a|b]]$, where $E[a|b]$ denotes the mean of a conditional on a fixed value of b and, analogously, $\mathrm{Var}[a|b]$ denotes the variance of a conditional on a fixed value of b. This gives

$$
\begin{aligned}
\mathrm{Var}[S_X] &= \mathrm{Var}\left[\sum_{i=1}^{N_X}\sum_{j=1}^{n_i} d_{ij}g_{ij}\right]\\
&= E[N_X]\mathrm{Var}\left[\sum_{j=1}^{n_1} d_{1j}g_{1j}\right] + \mathrm{Var}[N_X]E\left[\sum_{j=1}^{n_1} d_{1j}g_{1j}\right]^2\\
&= E[N_X]\mathrm{Var}\left[\sum_{j=1}^{n_1^{(d)}} g_{1j}\right] + \mathrm{Var}[N_X]E\left[\sum_{j=1}^{n_1^{(d)}} g_{1j}\right]^2\\
&= E[N_X]\left(E\left[n_1^{(d)}\right]\mathrm{Var}[g_{11}] + \mathrm{Var}\left[n_1^{(d)}\right]E\left[g_{11}\right]^2\right)\\
&\qquad + \mathrm{Var}[N_X]E[n_1]^2 E[d_{11}]^2 E[g_{11}]^2\\
&= E[N_X]\left(E\left[n_1\right]E[d_{11}]\mathrm{Var}[g_{11}] + E\left[g_{11}\right]^2\left(E[n_1]\mathrm{Var}[d_{11}] + \mathrm{Var}[n_1]E[d_{11}]^2\right)\right)\\
&\qquad + \mathrm{Var}[N_X]E[n_1]^2 E[d_{11}]^2 E[g_{11}]^2\\
&= N\cos^2\phi\,\bar{n}\epsilon G^2\left(f - 1 + 1 - \epsilon + \tfrac{\sigma_n^2}{\bar{n}}\epsilon + \bar{n}\epsilon\sin^2\phi\right)\\
&= N\cos^2\phi\,\bar{n}\epsilon^2 G^2\left(\tfrac{f}{\epsilon} + \tfrac{\sigma_n^2}{\bar{n}} - 1 + \bar{n}\sin^2\phi\right).
\end{aligned}
\tag{6.11}
$$

The results for S_Y are identical, with the substitution $\cos^2\phi \leftrightarrow \sin^2\phi$. Many atomic clocks [245–249] and some molecular precision measurement experiments [128, 227] measure both

S_X and S_Y, while others detect only a single state [96, 111, 122, 132]. In what follows, we assume the case particular to ACME, namely that both states are probed. We will examine the case of detecting only one state, under the convenient (but perhaps unrealistic) assumption that there exists some means of normalizing for variations in N, separately.

In the regime $\phi = \pm\frac{\pi}{4} + \delta\phi$, where $\delta\phi \ll 1$, sensitivity to small changes in phase, $\delta\phi$, is maximized. In this case, we define the measured phase deviation $\delta\tilde{\phi}$ by $\tilde{\phi} = \pm\frac{\pi}{4} + \delta\tilde{\phi}$. This is related to measured quantities via the asymmetry $\mathcal{A} = \frac{S_X - S_Y}{S_X + S_Y} = \mp\sin(2\delta\tilde{\phi}) \approx \mp2\delta\tilde{\phi}$. When $N \gg 1$, the average value of $\tilde{\phi}$ computed in this way is equal to the phase ϕ of the two-level system, similarly to the case when there is no optical cycling (see Sec. 2.2.2)

The uncertainty in the asymmetry,

$$\sigma_\mathcal{A} \approx \frac{1}{N}\sqrt{\sigma_{S_X}^2 + \sigma_{S_Y}^2 - 2\sigma_{S_X,S_Y}^2}, \tag{6.12}$$

can be computed to leading order in $\delta\phi$ from σ_{S_X}, σ_{S_Y}, and the covariance $\sigma_{S_X,S_Y}^2 = \overline{S_X S_Y} - \overline{S_X}\,\overline{S_Y}$ using standard error propagation [250]. We relate $\sigma_\mathcal{A}$ to the uncertainty in the measured phase by $\sigma_\mathcal{A} = 2\sigma_{\tilde{\phi}}$. This relationship defines the statistical uncertainty in $\tilde{\phi}$, the measured value of ϕ, for the protocol described here. The covariance, $\sigma_{S_X,S_Y}^2 = -\frac{N}{4}\bar{n}^2\epsilon^2$, can be calculated directly using the same methods already described. The calculation is fairly complicated, so it is instructive to work it out explicitly. We have already calculated $\overline{S_X}$ and $\overline{S_Y}$. The first term in the expression for the covariance, $\overline{S_X S_Y}$, can be computed by Wald's equation after the appropriate grouping of terms:

$$
\begin{aligned}
E[S_X S_Y] &= E[\sum_{i=1}^{N_X}\sum_{j=1}^{n_i}\sum_{k=1}^{N_Y}\sum_{l=1}^{n_k} d_{ij}g_{ij}d_{kl}g_{kl}] \\
&= E[\alpha_{11} + \cdots + \alpha_{N_X N_Y}],
\end{aligned}
\tag{6.13}
$$

where $\alpha_{ij} = \sum_{j=1}^{n_i}\sum_{l=1}^{n_k} d_{ij}g_{ij}d_{kl}g_{kl}$. Then Wald's equation gives

$$E[S_X S_Y] = E[N_X N_Y]E[\alpha_{11}]. \tag{6.14}$$

The second factor is simple:

$$E[\alpha_{11}] = E\left[\sum_{j=1}^{n_1}\sum_{l=1}^{n_1} d_{1j}g_{1j}d_{1l}g_{1l}\right]. \tag{6.15}$$

Since l and j are independent, this expression reduces to $E[\alpha_{11}] = E[\sum_{j=1}^{n_1} d_{1j}g_{1j}]^2 = \bar{n}^2\epsilon^2 G^2$, where the RHS is obtained by repeatedly applying Wald's lemma just as we saw when calculating $E[S_X]$.

To calculate $E[N_X N_Y]$, we note that $N_Y = N - N_X$, where N is the total number of particles. Then $E[N_X N_Y] = NE[N_X] - E[N_X^2]$. We already know that $E[N_X] = N\cos^2\phi$, while the second term can be obtained from $\mathrm{Var}[N_X] = E[N_X^2] - E[N_X]^2 = N\cos^2\phi\sin^2\phi$.

Combining these expressions and simplifying, we find

$$\mathrm{cov}(S_X, S_Y) = -N\cos^2\phi\sin^2\phi\,\bar{n}^2\epsilon^2 G^2. \tag{6.16}$$

This result can be understood as follows: the photon scattering and detection processes for particles projected onto $|X\rangle$ and $|Y\rangle$ are independent, so the covariance between signals S_X and S_Y only arises from quantum projection. In the simplest case of perfectly efficient, noise-free detection and photon scattering, e.g., $\epsilon = 1$, $\bar{n} = 1$, $\sigma_n = 0$, and $G = 1$, the quantum projection noise leads to signal variances $\sigma_{S_X}^2 = \sigma_{S_Y}^2 = \frac{N}{4}$. The covariance is negative because a larger number of particles projected onto $|X\rangle$ is associated with a smaller number of particles projected onto $|Y\rangle$. The additional factor of $\bar{n}^2\epsilon^2 G^2$ for the general case accounts for the fact that both signals S_X and S_Y are scaled by $\bar{n}\epsilon G$ when on average $\bar{n}$ photons are scattered per particle, a proportion ϵ of those photons are detected, and each detected photoelectron is amplified by a factor of G.

The uncertainty in the measured phase, computed using the procedure just described, has the form

$$\sigma_{\tilde{\phi}} = \frac{1}{2\sqrt{N}}\sqrt{F}, \tag{6.17}$$

where we have defined the "excess noise factor" F given in this phase regime by

$$F = 1 + \frac{1}{\bar{n}}\left(\frac{f}{\epsilon} - 1\right) + \frac{\sigma_n^2}{\bar{n}^2}. \tag{6.18}$$

264

Direct computation shows that the definition of the measured phase $\tilde{\phi}$ considered here saturates the Cramer-Rao bound [251], which specifies a lower bound on the uncertainty of a statistical quantity that can be computed from measured statistical variables, when $N \gg 1$, provided S_X and S_Y are distributed as bivariate normal variables in this case[13]. In other words, no alternative definition of $\tilde{\phi}$ computed from S_X and S_Y would give an unbiased estimator of ϕ with smaller measurement uncertainty $\sigma_{\tilde{\phi}}$. This is a nice sanity check because it assures us that we are not simply choosing an unwise definition of the measured phase, $\tilde{\phi}$, that under-utilizes the information contained in S_X and S_Y for the model considered here.

It is instructive to evaluate Eq. 6.18 in some simple limiting cases. For example, consider the case when exactly one photon is scattered per particle so that $\bar{n} = 1$ and $\sigma_n = 0$. (This is typical for experiments with molecules, where optical excitation essentially always leads to decay into a dark state.) In this case, $F = \frac{1}{\epsilon}$ and the uncertainty in the phase measurement is $\sigma_{\tilde{\phi}} = \frac{1}{2\sqrt{N}\epsilon}$, as stated previously. Alternatively, as $\bar{n} \to \infty$, $F \to 1 + \left(\frac{\sigma_n}{\bar{n}}\right)^2$. This is in exact analogy with the excess noise of a photodetector whose average gain is $\bar{n}$ and whose variance in gain is σ_n^2 [252]. By inspection, the ideal result of $F \to 1$ can be achieved only if $\frac{\sigma_n}{\bar{n}} \to 0$, and either $\epsilon/f \to 1$ or $\bar{n} \to \infty$.

6.4.4   Model of optical cycling process

The results up to now, and in particular Eq. 6.18, do not depend on any particular model of the optical cycling process. We now consider a realistic model of optical cycling to compute $\bar{n}$ and σ_n^2 in terms of more practical quantities. We define the branching fraction to dark states, which are lost from the optical cycle, to be b_ℓ. We assume that each particle interacts with the excitation laser light for a time T, during which the scattering rate of a particle in the optical cycle is r. Therefore, an average of rT photons would be scattered in the absence of decay to dark states, i.e. when $b_\ell = 0$. (All of our results hold for a time-dependent scattering rate $r(t)$, with the substitution $rT \to \int r(t)dt$.) Note that in

---

13. In particular, I have shown that if N_X and N_Y were *Poisson*-distributed rather than binomially distributed, then the Cramer-Rao bound would be saturated, but I have not gone through the derivation for our exact case. However, I find it extremely implausible that there is a significant difference between the Poisson-distributed and binomial-distributed cases for N_X and N_Y.

the limit $rT \to \infty$, $1/b_\ell$ photons are scattered per particle on average. Recall that the number of photons scattered from the i-th particle, when projected to a given state, is n_i. We define the probability that a particle emits exactly n_i photons to be $P(n_i; rT, b_\ell)$. This probability distribution can be computed by first ignoring the decay to dark states. For the case where $b_\ell = 0$, the number of photons emitted in time T follows a Poisson distribution with average number of scattered photons rT. For the more general case where $b_\ell > 0$, we assign a binary label to each photon depending on whether it is associated with a decay to a dark state. Each decay is characterized by a Bernoulli process, and we use the conventional labels of "successful" (corresponding to decay to an optical cycling state) and "unsuccessful" (corresponding to decay to a dark state) for each outcome.

For concreteness, we have assumed here that "unsuccessful" decays, i.e., those that populate dark states, emit photons with the same detection probability as all successful decays.[14] Before we provide the general formula for $P(n_i; rT, b_\ell)$, let's consider a specific example. For instance, $P(n_i = 3; rT, b_\ell)$ is the probability that there are *exactly* three photons scattered from a particle subject to scattering rate r for a total interaction time T, when the probability of each decay leading to a dark state is b_ℓ. We will denote a sequence of events with a series of symbols, where ✓ denotes a successful event (decay to a bright state), ✗ denotes an unsuccessful event (decay to a dark state), and □ denotes either a successful or unsuccessful event. For example, two successful events followed by an unsuccessful event is denoted by ✓✓✗, and the probability of exactly this sequence of events occurring is denoted by $P(\checkmark\checkmark\text{✗}; rT, b_\ell)$. Then

$$
\begin{aligned}
P(n_i = 3; rT, b_\ell) &= \quad P(\checkmark\checkmark\square; rT, b_\ell) \\
&\quad + P(\checkmark\checkmark\text{✗}\square; rT, b_\ell) \\
&\quad + P(\checkmark\checkmark\text{✗}\square\square; rT, b_\ell) \\
&\quad + \cdots .
\end{aligned}
\tag{6.19}
$$

In detail, we sum the probabilities for mutually exclusive scenarios. First, there can be exactly 3 physical decays (top line: two decays to the bright state followed by a decay

---

14. The opposite case, in which decays to dark states are always undetected, can be worked out with the same approach and leads to similar conclusions, as described below.

to either a bright or dark state). Alternatively, there can be exactly four events in the Poisson process such that the first two decays are to bright states and the third is to a dark state (second line); the fourth "event" does not physically occur, regardless of whether it's labelled successful or unsuccessful, because there can be no physical decays once a particle has occupied a dark state. Next, we consider the scenario in which there are exactly five events in the Poisson process such that the first two decays are to bright states and the third is to a dark state (third line), with any possibilities for the fourth and fifth event. We continue in this way for every possible number of events in the Poisson process greater than or equal to three.

With this example in mind, we note that in general, $P(n_i;\, rT, b_\ell)$ is the probability that (I) there are exactly n_i events in the Poisson process such that the first $n_i - 1$ are "successful" (and the last may be successful or unsuccessful; see the top line in the expression for $P(n_i = 3;\, rT, b_\ell)$ above), or (II) there are more than n_i events such that the first $n_i - 1$ are successful and the n_i-th is unsuccessful (see the second and lower lines in the expression for $P(n_i = 3;\, rT, b_\ell)$ above).

It is useful to note that the number of decays in the case that $b_\ell \to 0$ is given by the Poisson distribution, $P(n_i;\, rT, 0) = e^{-rT}(rT)^{n_i}/n_i!$. Now we will consider $P(n_i;\, rT, b_\ell)$ for general values of b_ℓ. For the simplest case in which no decays occur, $n_i = 0$, we can set $b_\ell = 0$ without loss of generality since then the branching fraction to dark states cannot be relevant. Thus $P(n_i = 0;\, rT, b_\ell) = P(n_i = 0;\, rT, 0) = e^{-rT}$. That is, the probability that a particle doesn't scatter *any* photons exponentially decreases with the product of scattering rate and time, since any excited-state population exponentially decays in time.

For $n_i > 0$, condition (I) contributes

$$P(n_i;\, rT, 0)(1 - b_\ell)^{n_i-1} = (1 - b_\ell)^{n_i-1}e^{-rT}\frac{(rT)^{n_i}}{n_i!}. \tag{6.20}$$

Explicitly, the probability that exactly n_i photons are emitted and all are detected is the product of the probability that exactly n_i photons are emitted, and the probability that the first $n_i - 1$ decays each independently lead back to the bright state (recall that we assume that decays to both the bright and dark states are detected with the same probability).

Condition (II) contributes

$$
\begin{aligned}
\sum_{k=1}^{\infty} P(n_i + k;\, rT,\, 0)(1 - b_\ell)^{n_i-1} b_\ell &= e^{-rT} b_\ell (1 - b_\ell)^{n_i-1} \sum_{k=1}^{\infty} \frac{(rT)^{n_i+k}}{(n_i+k)!} \\
&= e^{-rT} b_\ell (1 - b_\ell)^{n_i-1} \frac{e^{rT}\gamma(n_i+1, rT)}{n_i!} \qquad (6.21) \\
&= b_\ell (1 - b_\ell)^{n_i-1} \frac{\gamma(n_i+1, rT)}{n_i!},
\end{aligned}
$$

where $\gamma(n_i + 1, rT) = \int_0^{rT} dx\, x^{n_i} e^{-x}$ is a lower incomplete gamma function. The sum is computed (line 1 to line 2) via Mathematica, allowing us to re-express the result in terms of an integral with no closed form[15]. To understand why the LHS expresses the contribution to $P(n_i;\, rT, b_\ell)$ due to condition (II), consider a single term with a fixed value of k. The probability that exactly $n_i + k$ photons are emitted in the Poisson process such that the first $n_i - 1$ lead to the bright state and the n_i-th leads to a dark state is the product of the the probability that exactly $n_i + k$ photons are emitted, the probability that the first $n_i - 1$ decays lead back to the bright state, and the probability that the next decay leads to a dark state. It is important to remember that the k events after the decay to a dark state do not physically occur; they are merely artifacts that allow us to calculate $P(n_i;\, rT, b_\ell)$, but do not correspond to physical photon emissions. In the example of Eq. 6.19, they are represented by $\square$ symbols in lines two and three, and arise from the Poisson process for photon emissions that we consider *before* labeling decays as successful or unsuccessful.

Therefore, the probability that more than one decay occurs, combining conditions (I) and (II), is

$$
P(n_i > 0;\, rT, b_\ell) = \frac{e^{-rT}}{1 - b_\ell} \frac{[rT(1 - b_\ell)]^{n_i}}{n_i!} + \frac{b_\ell}{1 - b_\ell} \int_0^{rT} dx\, e^{-x} \frac{[x(1 - b_\ell)]^{n_i}}{n_i!}. \qquad (6.22)
$$

In order to compute $\bar{n}$ and σ_n^2, we need to compute the central moments:

$$
\bar{n}_i = \sum_{n_i=0}^{\infty} n_i P(n_i;\, rT, b_\ell), \qquad (6.23)
$$

---

15. A sensible person may wonder whether it's really much better to have a non-analytic result in the form of this integral rather than in the form of a sum. However, the fact that this sum can be re-expressed as an integral is the first of two computational miracles that will allow us to get an analytical result for the excess noise factor $F(rT, b_\ell, \epsilon/f)$.

$$\overline{n_i^2} = \sum_{n_i=0}^{\infty} n_i^2 P(n_i; rT, b_\ell). \tag{6.24}$$

First, the mean:

$$
\begin{aligned}
\bar{n} &= \sum_{n_i=0}^{\infty} n_i P(n_i; rT, b_\ell) \\[2mm]
&= \frac{e^{-rT}}{1-b_\ell} \sum_{n_i=1}^{\infty} \frac{[rT(1-b_\ell)]^{n_i}}{n_i!} n_i + \frac{b_\ell}{1-b_\ell} \int_0^{rT} dx \, e^{-x} \sum_{n_i=1}^{\infty} \frac{[x(1-b_\ell)]^{n_i}}{n_i!} n_i \\[2mm]
&= \frac{e^{-rT}}{1-b_\ell} e^{rT} e^{-rTb_\ell} rT(1-b_\ell) + \frac{b_\ell}{1-b_\ell} \int_0^{rT} dx \, e^{-x} e^x e^{-xb_\ell} x(1-b_\ell) \\[2mm]
&= rT e^{-rTb_\ell} + b_\ell \int_0^{rT} dx \, e^{-xb_\ell} x \\[2mm]
&= rT e^{-rTb_\ell} + \frac{1-e^{-rTb_\ell}(rTb_\ell+1)}{b_\ell} \\[2mm]
&= \frac{1-e^{-rTb_\ell}}{b_\ell}.
\end{aligned}
\tag{6.25}
$$

Let's take the steps slowly: in the second line, we have written out $P(n_i; rT, b_\ell)$ explicitly and pulled factors out of the sum that are independent of n_i. In the second term, we have interchanged the order of the summation and integration. We then use the closed form for the sums (line 3), which allows the integrand in the second term to be simplified, and thus the integral to be completed[16]. Exactly analogous steps allows us to compute $\overline{n_i^2}$ and therefore $\sigma_n^2 = \overline{n_i^2} - (\bar{n}_i)^2$. The result is

$$\sigma_n^2 = \frac{1 - b_\ell + e^{-rTb_\ell} b_\ell (2rTb_\ell - 2rT + 1) - e^{-2rTb_\ell}}{b_\ell^2}. \tag{6.26}$$

Recall that in order to get this result, we have assumed that the decays to dark states are detected at the same rate as decays to bright states. If we instead assume that decays to dark states are never observed, then the average number of observable photons emitted per particle is reduced, $\bar{n}_i \to \bar{n}_i b_\ell$, and likewise $\overline{n_i^2} \to \overline{n_i^2} b_\ell$. However, note that this implies $\sigma_n^2 \to b_\ell(\overline{n_i^2} - b_\ell(\bar{n}_i)^2) \not\propto \sigma_n^2$. As a result, there is no simple substitution to express subsequent results in the case that decays to dark states are never observed. However, that case can be worked out using the same methods, and all results are qualitatively similar. In all

---

16. The fact that the integral of the sum has a closed form, while the integral of the summand does not, is the second miracle allowing us to get an analytic result for F—*a priori*, we don't have the right to expect any such thing.

that follows, we will only consider the case that decays to dark states and bright states are detected with equal likelihood.

6.4.5 Excess noise factor in general case

We now plug in the expressions for $\bar{n}$ and σ_n^2 to Eq. 6.18 and, after quite a bit of simplifying algebra, find

$$F = 1 + \frac{1}{1 - e^{-rTb_\ell}} \left(\frac{b_\ell}{\epsilon/f} + \frac{1 - 2b_\ell + 2b_\ell e^{-rTb_\ell}(1 - rT(1 - b_\ell)) - e^{-2rTb_\ell}}{1 - e^{-rTb_\ell}} \right). \qquad (6.27)$$

This is the fundamental result of this section: the excess noise factor in Eq. 6.27 allows us to compute the uncertainty in the measured phase, $\sigma_{\tilde{\phi}} = \frac{1}{2\sqrt{N}}\sqrt{F}$, as a function of the branching fraction to dark states b_ℓ, photon detection efficiency ϵ, photodetector excess noise factor f, and combination of scattering rate and scattering time rT. The behavior of the SNR (proportional to $1/\sqrt{F}$) arising from Eq. 6.27 is illustrated in Fig. 6.8.

To understand the implications of this result, we consider several special cases, summarized in Table 6.3. We first consider the simple case when cycling is allowed to proceed until all particles decay to dark states, i.e., $b_\ell rT \to \infty$. We refer to this as the case of "cycling to completion." In this case, for the generically applicable regime $\frac{\epsilon}{f} \leq \frac{1}{2}$ we find $F \geq 2$, even as the transition becomes perfectly closed ($b_\ell \to 0$). We can understand this result intuitively as follows. As the optical cycling proceeds, the number of particles that exit the optical cycle after each photon scatter is proportional to the number of particles that are currently in the optical cycle, $\frac{dP}{dn_i} \propto P$. Hence, we expect $P(n_i; rT \to \infty, b_\ell) \propto e^{-\alpha n_i}$ for some characteristic constant α. The width σ_n of this exponential distribution is given by the mean $\bar{n}$; that is, $\sigma_n \approx \bar{n}$. Therefore, we should expect that cycling to completion reduces the SNR by a factor of $\sqrt{F} = \sqrt{1 + (\sigma_n/\bar{n})^2} \to \sqrt{2}$ compared to the ideal case of $F = 1$, which requires $\frac{\sigma_n}{\bar{n}} = 0$.

In a bit more detail, we can derive $P(n_i; rT \to \infty, b_\ell)$ exactly and see that this heuristic argument holds. The probability that the particle lands in a dark state after exactly n_i photons are emitted, provided cycling proceeds arbitrarily long, is $P(n_i; rT \to \infty, b_\ell) =$
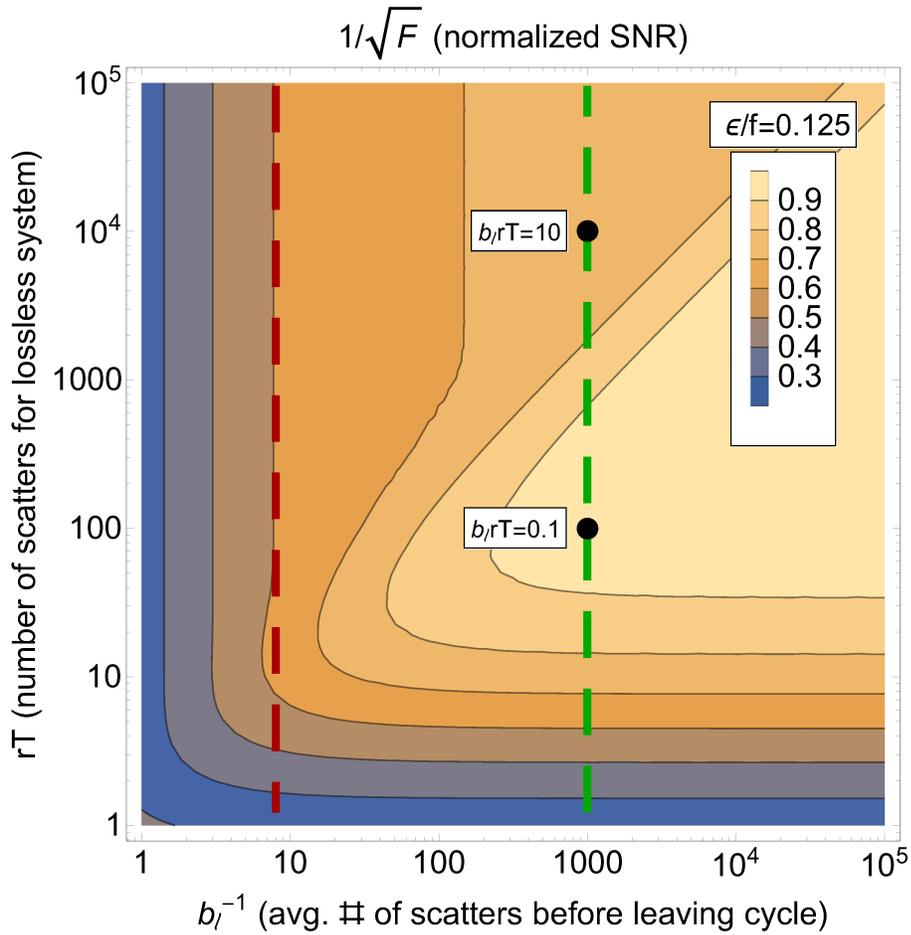
Figure 6.8: $1/\sqrt{F}$, the SNR resulting from Eq. 6.17, normalized to the ideal case of the QP limit ($F = 1$). This plot assumes an ambitious value of $\epsilon/f = 0.125$, which would result from $\approx 50\%$ optical collection efficiency, $\approx 25\%$ photodetector quantum efficiency, and negligible photodetector excess noise ($f \approx 1$). Note that, without optical cycling, such a configuration would represent an improvement in the signal-to-noise ratio over ACME II of a factor of ≈ 1.8. When few photons per particle can be detected, i.e., when $(\epsilon/f)/b_\ell \ll 1$ (far left of plot), cycling to very deep completion ($b_\ell rT \gg 1$, top of plot) does not significantly affect the SNR. Even when one photon per particle can be detected on average, i.e., when $(\epsilon/f)/b_\ell = 1$ (dashed red line), the SNR never exceeds roughly 60% of its ideal value. By further closing the optical cycle, i.e. such that $(\epsilon/f)/b_\ell \gg 1$ (right of dashed red line), the SNR can be improved to near the optimal value given by the QP limit. However, to reach this optimal regime, the number of photons that would be scattered in the absence of dark states, rT, must be small compared to the average number that can be scattered before a particle exits the optical cycle, $1/b_\ell$. For example, with $1/b_\ell = 1,000$ (green dashed line) and $rT = 100$ so that $b_\ell rT = 0.1$ (lower circle), the SNR is more than 30% larger than in the case when $rT = 10,000$ and $b_\ell rT = 10$ (upper circle). Cycling through the I electronic state in ThO is expected to yield $b_\ell = 0.09$, and thus for the assumed value of ϵ/f we would have $(\epsilon/f)/b_\ell \approx 1.4$, where the SNR behaves qualitatively similarly as along the dashed red line.

$(1-b_\ell)^{n_i-1}b_\ell$. We can find closed forms for $\bar{n} = \sum_{n_i=1}^{\infty} P(n_i; rT \to \infty, b_\ell)n_i$ and analogously for $\overline{n^2}$ using Mathematica to see that, in this case, $\bar{n} = 1/b_\ell$ and $\sigma_n = \sqrt{1-b_\ell}/b_\ell$. We can understand the full distribution $P(n_i; rT \to, b_\ell)$ better by rewriting it as $P(n_i; rT \to, b_\ell) = \frac{1}{\bar{n}-1}\exp\left[-\ln[\bar{n}/(\bar{n}-1)]n_i\right] = A\exp(-\alpha \times n_i)$, so the probability of scattering n_i photons before being lost to a dark state is given as a decaying exponential in n_i as anticipated, with decay rate $\alpha = \ln[\bar{n}/(\bar{n}-1)]$ and normalization constant $A = (\bar{n}-1)^{-1}$.

Surprisingly, this reduction in SNR can be partially recovered for an imperfectly closed optical cycle, by choosing a finite cycling time, $rT < \infty$, to minimize $\sigma_{\tilde{\phi}}$. The best limiting case, as found from Eq. 6.27, preserves the condition that many photons are detected per particle, $rT\epsilon/f \gg 1$, but additionally requires that the probability of decaying to a dark state remains small, $rTb_\ell \ll 1$. In this case, photon emission is approximately a Poisson process for which $\left(\frac{\sigma_n}{\bar{n}}\right)^2 \approx \frac{1}{rT} \ll 1$, and the excess noise factor, F, does not have a significant contribution from the variation in scattered photon number, σ_n. The optimal value of rT for a finite proportion of decays to dark states, b_ℓ, and detection efficiency, ϵ, lies in the intermediate regime and can be computed numerically.

A special case of "cycling to completion," which must be considered separately, occurs when every particle scatters exactly one photon, corresponding to parameter values $b_\ell = 1$ and $rT \gg 1$ so that $\bar{n} = 1$ and $\sigma_n = 0$. As we have already seen, in this case there is no contribution to the excess noise arising from variation in the scattered photon number, and hence the SNR is limited only by photon shot noise: $F = \frac{f}{\epsilon}$.

In atomic physics experiments with essentially completely closed optical cycles, $b_\ell \approx 0$, the limit $b_\ell rT \to \infty$ is not obtained even for very long cycling times where $rT \gg 1$. Instead, in this case $b_\ell rT \to 0$ and hence $F \to 1 + \frac{f}{rT\epsilon}$, which approaches unity as the probability to detect a photon from each particle becomes large, $\frac{\epsilon}{f}rT \gg 1$. Therefore, the reduction in the SNR associated with the distribution of scattered photons does not occur in this limit of a completely closed optical cycle.

6.4.6   Extensions to alternative situations

One may also consider how the additional noise due to optical cycling combines with other noise sources in the detection process. For example, similar derivations can be performed

	Condition	Sub-conditon	F
1a	$b_\ell rT \to \infty$		$2 + b_\ell(\frac{f}{\epsilon} - 2)$
1b	$b_\ell rT \to \infty$	$\epsilon/f \le 0.5$	≥ 2
2a	$b_\ell rT \to 0$		$1 + \frac{f}{rT\epsilon} + \frac{1}{2}b_\ell(\frac{f}{\epsilon} - 2)$
2b	$b_\ell rT \to 0$	$\frac{\epsilon}{f}rT \to \infty$	1
3a	$b_\ell \to 1$		$\frac{f}{\epsilon}\frac{1}{1-e^{-rT}}$
3b	$b_\ell \to 1$	$rT \to \infty$	$\frac{f}{\epsilon}$

Table 6.3: The excess noise factor F in some special cases. (1a) All particles are lost to dark states during cycling. (1b) With all particles lost and realistic detection efficiency, $\epsilon/f \le 0.5$, $F \ge 2$. (2a) No particles are lost to dark states. (2b) No particles are lost, but many photons per particle are detected. The QP limit is reached. (3a) No more than one photon can be scattered per particle. (3b) Exactly one photon is scattered per particle and the photon shot noise limit is reached.

assuming a statistical distribution of N or ϕ to obtain qualitatively similar but more cumbersome results. I did not manage to glean any insights, beyond those already described, from these more involved calculations and omit them here.

We now address the question of how the excess noise factor F differs when the signal is only detected from one state, rather than from both X and Y. This could be important because several experiments of interest detect only a single state, assumed here to be $|X\rangle$ without loss of generality [96, 111, 122, 132]. In such experiments, the asymmetry $\mathcal{A} = \frac{S_X - S_Y}{S_X + S_Y}$ cannot be constructed and it is necessary to measure the phase ϕ in a different way. Following the standard approach, we use Eq. 6.10 to define $\tilde{\phi}$ by $S_X = N \cos^2 \tilde{\phi} \bar{n} \epsilon G$. We assume once again the regime $\phi = \pm\frac{\pi}{4} + \delta\phi$, with $\delta\phi \ll 1$, and as before define $\tilde{\phi} = \pm\frac{\pi}{4} + \delta\tilde{\phi}$ so that $S_X = N\bar{n}\epsilon G(\frac{1}{2} - \delta\tilde{\phi})$. The prefactor $N\bar{n}\epsilon G$ must be calibrated independently. For example, suppose one applies a known change $\Delta\phi$ to the phase and measures the contrast $\mathcal{C} = dS_X/d(\Delta\phi)$, with $\bar{\mathcal{C}} = -N\bar{n}\epsilon G$. We assume that the uncertainty of this calibration is small, so that $N\bar{n}\epsilon G$ may be treated as a known parameter. In the context of any given experiment, this assumption may be quite dubious: in particular, the number of particles N can often fluctuate between the calibration and the phase measurement steps of an experimental protocol. Nevertheless, we will proceed with this assumption to understand the qualitative distinction between observing one vs. two states in the phase measurement. We find that the uncertainty in the measured phase still has the form $\sigma_{\tilde{\phi}} = \frac{1}{2\sqrt{N}}\sqrt{F}$ under these assumptions, but with a larger excess noise factor:

$$F = 1 + \frac{2}{\bar{n}} \left(\frac{f}{\epsilon} - 1 \right) + 2 \frac{\sigma_n^2}{\bar{n}^2}. \tag{6.28}$$

As with the case when both states are measured, this definition of the measured phase $\tilde{\phi}$ saturates the Cramer-Rao bound provided $N \gg 1$. That is, no alternative definition of $\tilde{\phi}$ computed from only S_X and the calibration factor $N\bar{n}\epsilon G$ would give an unbiased estimator of ϕ with smaller measurement uncertainty $\sigma_{\tilde{\phi}}$, reassuring us that the excess noise we've computed does not simply arise due to a poor choice of how to calculate $\tilde{\phi}$.

Therefore, using the results for $\bar{n}$ and σ_n applicable to the same optical cycling process considered previously, we find that when only a single state $|X\rangle$ is detected, the excess noise factor takes the form

$$F = 1 + \frac{2}{1 - e^{-b_\ell rT}} \left(\frac{b_\ell f}{\epsilon} + \frac{1 - 2b_\ell + 2b_\ell e^{-b_\ell rT}(1 - rT(1 - b_\ell)) - e^{-2b_\ell rT}}{1 - e^{-b_\ell rT}} \right). \tag{6.29}$$

Comparing to Eq. 6.27, we see that second term in the expression for F is simply inflated by a factor of 2 compared to the case where both states can be detected.

6.4.7 Implications for ACME

Let us finally zoom back in to the problem at hand: how much will optical cycling improve the SNR in a future ACME measurement? In ACME II, we have $\epsilon \approx 0.05$ due to an optical collection efficiency of $\approx 20\%$ and PMT quantum efficiency of $\approx 25\%$. Further, the PMTs have an excess noise factor of $f \approx 1.25$ (based on measurements by Cristian Panda; see Sec. 3.2.1), so $\epsilon/f \approx 0.04$. Thus the excess noise factor is $F = \frac{f}{\epsilon} \approx 25$ in ACME II. Therefore, by a combination of optical cycling detection and improved photon detection efficiency, we could hope for at most a factor of 25 improvement in effective molecular flux (in terms of the effect on the SNR).[17]

---

17. This treatment relies on the assumption that in ACME II, each molecule emits a photon with near-unit probability. This is valid: we have estimated from measurements that with 65 mW of power (beam waists $w_x \approx 2$ mm and $w_y \approx 4$ cm, transition dipole moment $d \approx 1.8$ D [174]), approximately 99% of the molecules in the molecular beam fluoresce [253]. The typical probe laser power in the ACME II data set was slightly higher, $P_{\text{probe}} \approx 75$ mW (based on continuously logged measurements with a photodiode). Note that the refinement laser power was typically much higher, $P_{\text{ref}} \approx 620$ mW, in order to ensure efficient refinement of the STIRAP-prepared phase, $A_{\text{ref}} \gg 1$.
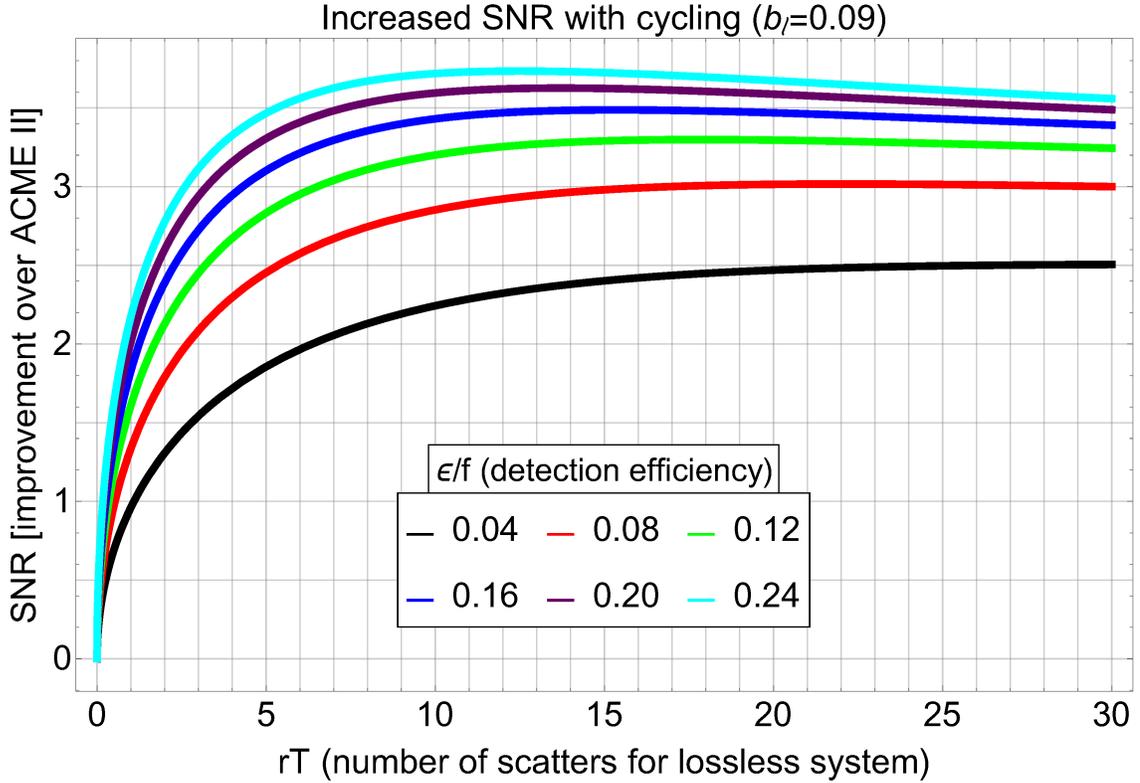
Figure 6.9: Gain in SNR between ACME II and ACME III, via a cycling scheme on $X \leftrightarrow I$, where $b_\ell \approx 0.09$. Distinct curves represent different possible values of ϵ/f that may be applicable to ACME III, with the lowest curve showing no improvement over the ACME II photon detection efficieincy ϵ or PMT excess noise factor f. We see that the optimal combination of scattering rate and scattering time, rT, is typically finite such that not all molecules decay to dark states (and this effect is more pronounced for large ϵ/f). With no improvement in photon detection, optical cycling (average of 11 scattered photons per molecule) yields an increase in SNR by a factor of ≈ 2.5. On the other hand, a factor of 6 improvement in photon detection, $\epsilon/f \to 0.24$ (cyan curve), together with optical cycling, yields a total increase in SNR of $\approx 3.7\times$ (i.e., using optical cycling represents a factor of ≈ 1.5 improvement in this case over a configuration with the same improved detection but no cycling, which alone gives an improvement in SNR of $\approx 2.5\times$). In all cases, as $rT \to 0$ (no photons are scattered), the SNR also approaches zero. For $b_\ell = 0.09$, on average one photon per molecule is scattered when $rT \approx 1.05$. For this value of rT, the improvement in SNR, compared to ACME II, is unity for $\epsilon/f = 0.04$ and larger than unity for $\epsilon/f > 0.04$, due to the improved detection efficiency (since approximately one photon is scattered per molecule in ACME II).

The most promising optical cycling transition is $X(v = 0) \leftrightarrow I(v = 0)$, which has a branching ratio to dark states of $b_\ell \approx 0.09$[174]. Ongoing work to implement optical cycling detection with this transition is being led by Daniel Ang. Here, I will assume that no repumping lasers are implemented to address loss to other electronic states (dominantly H and Q) or higher vibrational levels.

The gain in effective molecular flux is $G_{\mathrm{cyc}} = F_{\mathrm{ACME\,II}}/F_{\mathrm{ACME\,III}}$, where $F_{\mathrm{ACME\,II}} \approx 25$ and $F_{\mathrm{ACME\,III}}$ is the excess noise factor that will be obtained in ACME III. As usual, the improvement in the SNR scales as $\sigma_{\tilde{\phi}} \propto G_{\mathrm{cyc}}^{-1/2}$. We see in Fig. 6.9 that the effective molecular gain from cycling with $b_\ell = 0.09$ (≈ 11 scattered photons per molecule) is $G_{\mathrm{cyc}} \approx 6$ provided the effective photon detection efficiency ϵ/f is not improved.

On the other hand, we may perform cycling detection in a separate vacuum chamber downstream of the interaction region, where far more efficient optical collection is possible (e.g., collecting closer to $\approx 75\%$ of photons rather than $\approx 25\%$). Further, as we saw in Sec. 6.3, it may be possible to use silicon photomultipliers with $\epsilon_{\mathrm{det.}}/f \approx 0.4$ rather than $\epsilon_{\mathrm{det.}}/f \approx 0.2$ as is the case with the PMTs used in ACME II. Therefore, suppose rather ambitiously that $\epsilon/f \to 0.24$ in ACME III. Then we see in Fig. 6.9 that the effective molecular flux can be increased by a factor of ≈ 14 with optical cycling detection, compared to ACME II. While this would be an extremely significant gain, note that in this case cycling itself contributes only a factor of ≈ 2 effective flux beyond the level achieved simply by improving the effective detection efficiency.

This trade-off is generic: the better the effective detection efficiency ϵ/f, the less beneficial is using optical cycling detection compared to scattering a single photon per molecule. Conversely, using optical cycling detection diminishes the marginal benefit of a fixed increase in ϵ/f. Critically, we cannot compute the overall gain in ACME III vs. ACME II by simply multiplying the gain from cycling vs. not cycling, and the gain from improved vs. unimproved detection. The actual marginal benefit of one upgrade, keeping the other upgrades fixed, should be well-understood to inform how many resources are devoted to getting either optical cycling or an improved detection technology to work.

6.5 Outlook

In this chapter, I've considered only a few possible improvements to the ACME apparatus to achieve greater statistical sensitivity to d_e: an electrostatic or magnetostatic molecular lens, modifications to the molecular beam source "skimmer" collimating aperture, an alternative photodetector technology, and optical cycling detection. These broadly cover the candidates for improving statistical sensitivity in the next-generation ACME measurement, although a few other low-priority ideas also exist (e.g., implementing a "load lock" to the molecular beam source to allow replacement of used targets before their molecular yields significantly degrade)[18]. One other upgrade to the statistical sensitivity, which has already been implemented by Cole Meisenhelder and Cristian Panda, is eliminating the source of additional noise as described in Sec. 4.15.3. This contributes an improved statistical sensitivity to d_e by a factor of ≈ 1.7 compared to ACME II.

In addition to the improvements in statistical sensitivity, we must suppress all known systematic error contributions to a level below the target ACME III statistical sensitivity. To this end, it will be advantageous to redesign the magnetic shielding in the interaction region, which we have observed to have undesirably large residual magnetization ($\sim 100\,\mu\mathrm{G}$), which is not removed by degaussing and changes significantly when the apparatus is jostled. Because we have observed systematic errors proportional to residual magnetic field gradients in ACME II, it will be important to have better control over residual fields in ACME III than previously anticipated. Methods to improve the magnetic shielding design are currently being explored by James Chow.

Furthermore, it will be necessary to suppress systematic errors arising from birefringence gradients together with $\mathcal{E}^{\mathrm{nr}}$. In ACME I, these effects were suppressed in part by aligning the state preparation laser polarization with the birefringence axis of the vacuum windows and electric field plates; however, in ACME II we cannot do this because the refinement laser polarization must match the STIRAP-prepared angular momentum alignment in order

---

18. To be clear, the skimmer modification is itself one of these low-priority ideas. However, at least one of the molecular lens or optical cycling detection will almost certainly have to be operational before an additional measurement is made. These projects are the primary focus of Xing Wu and Daniel Ang, respectively.

to suppress the $P^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$ systematic slope, $d\omega^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}/dP^{\tilde{\mathcal{N}}\tilde{\mathcal{E}}}$. In turn, the STIRAP-prepared angular momentum alignment cannot be made to match the birefringence axis because the STIRAP beams travel along $\hat{y}$, constraining the projection of the effective 1090 nm laser polarization in the xy-plane. If, as was the case in ACME I, all lasers were to propagate along $\pm\hat{z}$, then the refinement laser polarization could generically be aligned with the optical birefringence axis in order to suppress the $\mathcal{E}^{\text{nr}}$ systematic error. Cole Meisenhelder is currently exploring options to perform STIRAP with an alternative intermediate state, which could allow laser powers sufficiently low to be suitable for propagation through the field plates.

With these anticipated upgrades to the statistical sensitivity and systematic error control, up to an order-of-magnitude additional improvement in sensitivity to d_e may be achievable using the ACME beam of ThO molecules in the next five years, probing even deeper into high-mass and small-CP-violation regions of the parameter space of theories beyond the Standard Model.

Appendix A

Calculations with the spherical basis

This is intended as a handy reference for the definition of the spherical basis and useful identities, including some discussion of the composition of multiple spherical tensors and the correspondence between higher-rank tensors in spherical and cartesian bases. These results have been useful in various molecular calculations, especially for systematic error models. The reader is encouraged to double-check all identities before using them, for understanding if not assurance of accuracy.

Relationship to cartesian vectors

Definitions of spherical basis. See, e.g., [254, Sec. 4.2.1] Sec. 8.1.

$$
\begin{aligned}
\hat{u}_{\pm} &= \mp \tfrac{1}{\sqrt{2}}(\hat{x} \pm i\,\hat{y}) \\
\hat{u}_0 &= \hat{z}
\end{aligned}
\tag{A.1}
$$

$$
\begin{aligned}
\hat{x} &= \tfrac{1}{\sqrt{2}}(\hat{u}_- - \hat{u}_+) \\
\hat{y} &= \tfrac{i}{\sqrt{2}}(\hat{u}_- + \hat{u}_+) \\
\hat{z} &= \hat{u}_0
\end{aligned}
\tag{A.2}
$$

279

Conjugation

$$\hat{u}_\pm^* = -\hat{u}_\mp$$
$$\hat{u}_0^* = \hat{u}_0 \tag{A.3}$$

$$\hat{u}_q \cdot \hat{u}_{q'}^* = \delta_{qq'} \tag{A.4}$$

Components

Defining $v_q = \vec{v} \cdot \hat{u}_q$, we have the expansion

$$
\begin{aligned}
\vec{v} &= v_-\hat{u}_-^* + v_+\hat{u}_+^* + v_0\hat{u}_0 \\
&= -v_-\hat{u}_+ - v_+\hat{u}_- + v_-\hat{u}_0.
\end{aligned} \tag{A.5}
$$

$$
\begin{aligned}
v_x &= \tfrac{1}{\sqrt{2}}(v_- - v_+) \\
v_y &= \tfrac{i}{\sqrt{2}}(v_- + v_+) \\
v_z &= v_0
\end{aligned} \tag{A.6}
$$

$$
\begin{aligned}
v_\pm &= \mp\tfrac{1}{\sqrt{2}}(v_x \pm i v_y) \\
v_0 &= v_z
\end{aligned} \tag{A.7}
$$

For real cartesian components v_x, v_y, v_z:

$$
\begin{aligned}
v_\pm^* &= -v_\mp \\
v_0^* &= v_0
\end{aligned} \tag{A.8}
$$

For real spherical components $v_\pm$, v_0:

$$
\begin{aligned}
v_x^* &= +v_x \\
v_y^* &= -v_y \\
v_z^* &= v_z
\end{aligned} \tag{A.9}
$$

Identities

$$\vec{a} \cdot \vec{b} = -a_-b_+ - a_+b_- + a_0b_0 = (-1)^q a_q b_{-q} \tag{A.10}$$

$$\vec{v}^* \quad = \quad v_-^* \hat{u}_- + v_+^* \hat{u}_+ + v_0^* \hat{u}_0$$

$$\quad = \quad -v_-^* \hat{u}_+^* - v_+^* \hat{u}_-^* + v_0^* \hat{u}_0 \tag{A.11}$$

$$\vec{v}^* \cdot \vec{w} = +v_-^* w_- + v_+^* w_+ + v_0^* w_0 \tag{A.12}$$

$$\vec{v}^* \cdot \vec{v} = |v_-|^2 + |v_+|^2 + |v_0|^2 \tag{A.13}$$

$$(\vec{a} \times \vec{b})_\pm = \pm i(a_0 b_\pm - b_0 a_\pm) \tag{A.14}$$

$$(\vec{a} \times \vec{b})_0 = i(a_- b_+ - a_+ b_-) \tag{A.15}$$

$$\hat{u}_\pm \cdot \vec{a} = \mp i \hat{u}_\pm \cdot (\hat{u}_0 \times \vec{a}) \tag{A.16}$$

Application to E1 transitions

We most often use the spherical basis to evaluate electric dipole transitions, which involve a perturbing Hamiltonian of the form $\vec{\epsilon} \cdot \vec{r}$. We would like to express $\vec{r}$ in terms of components of a spherical tensor $T_q^{(1)}$, where I'll drop the rank-1 superscript for convenience. The spherical tensor components are defined to correspond to spherical harmonics Y_q^k as

$$T_0 \equiv r\sqrt{\frac{4\pi}{3}} Y_0^1 = z \tag{A.17}$$

$$T_- \equiv r\sqrt{\frac{4\pi}{3}} Y_-^1 = \frac{x - iy}{\sqrt{2}} \tag{A.18}$$

$$T_+ \equiv r\sqrt{\frac{4\pi}{3}} Y_+^1 = -\frac{x + iy}{\sqrt{2}}. \tag{A.19}$$

We can solve for $x = \frac{1}{\sqrt{2}}(T_- - T_+)$ and $y = \frac{i}{\sqrt{2}}(T_- + T_+)$. Note that $r_q = T_q$ are

equivalent notations. Then

$$\vec{r} = -T_- \hat{u}_+ - T_+ \hat{u}_- + T_0 \hat{u}_0. \tag{A.20}$$

Therefore, E1 transitions involve Hamiltonian terms proportional to

$$\vec{\epsilon} \cdot \vec{r} = -\epsilon_- T_+ - \epsilon_+ T_- + \epsilon_0 T_0 \tag{A.21}$$

$$\vec{\epsilon}^* \cdot \vec{r} = +\epsilon_+^* T_+ + \epsilon_-^* T_- + \epsilon_0^* T_0, \tag{A.22}$$

where $\vec{\epsilon}^* \cdot \vec{r}$ is the appropriate Hamiltonian for an absorbed, rather than emitted, photon in an E1 transition. For example, if $\epsilon_+ = 1$ and $\epsilon_0 = \epsilon_- = 0$, then the light is σ^+-polarized and an absorbed photon drives states to higher angular momentum projections along the z-axis (T_+ is the relevant operator), while emitted photons drive to states of lower angular momentum projection (T_- is the relevant operator). For linearly polarized light, $\vec{\epsilon} = \vec{\epsilon}^*$, so the distinction is irrelevant.

Composition of spherical tensors

Spherical tensors can be combined using Clebsch-Gordan coefficients:

$$T_q^k = \sum_{q_1, q_2} \langle k_1 k_2; q_1 q_2 | k q \rangle A_{q_1}^{k_1} B_{q_2}^{k_2}. \tag{A.23}$$

Thus two rank-1 tensors can be combined to form a rank-2 tensor, a vector, and a scalar. When combining more than two vectors, we can get multiple resulting tensors of the same rank. For example, when combining three vectors, we first take the tensor product of two vectors to obtain tensors of rank 0, 1, and 2. We then take the tensor product of each of these with the remaining vector, producing tensors of rank {1}, {0, 1, 2}, and {1, 2, 3}, for a total of one scalar, three vectors, two rank-2 tensors, and one rank-3 tensor. Together, it can be checked that these contain the correct number of independent components. Many useful identities are worked out in [255, Sec. 3.2].

Some useful expressions apply for rank-2 Cartesian tensors T_{ik} ($i, k = x, y, z$). From

[255, Sec. 3.2.2]:

$$
\begin{aligned}
T_{ik} &= E\delta_{ik} + A_{ik} + S_{ik} \\
E &= T_{ii}/3 \\
A_{ik} &= (T_{ik} - T_{ki})/2 \\
S_{ik} &= (T_{ik} + T_{ki})/2 - \delta_{ik}T_{jj}/3
\end{aligned}
\tag{A.24}
$$

Then the spherical components are

$$
\begin{aligned}
T_{00} &= E \\
T_0^1 &= A_{xy} \\
T_{\pm 1}^1 &= \mp(A_{yz} \pm iA_{zx})/\sqrt{2} \\
T_0^2 &= S_{zz} \\
T_{\pm 1}^2 &= \mp\sqrt{2/3}(S_{zx} \pm iS_{zy}) \\
T_{\pm 2}^2 &= \sqrt{1/6}(S_{xx} - S_{yy} \pm 2iS_{xy})
\end{aligned}
\tag{A.25}
$$

One must be careful when combining tensors in the special case that multiple tensors are identical. Some light is shed on this in the lecture notes [256]. The appropriate tensor products of spherical vectors involving no spins or differential operators are given by [255, Sec. 3.2]

$$
\{\cdots\{\{V \otimes V\}_{l_2} \otimes V\}_{l_3} \cdots \otimes V\}_{l_n m_n} = \sqrt{\frac{4\pi}{2l_n + 1}}|V|^n Y_{l_n m_n}(\theta, \phi)\Pi_{i=2}^n C_{10 l_{i-1} 0}^{l_i 0}, \tag{A.26}
$$

where $C_{l_1 m_1 l_2 m_2}^{lm}$ is a Clebsch-Gordan coefficient. In its simplest form, this is the familiar result that products of position vectors correspond to spherical harmonic operators.

Appendix B

Further details of rotational cooling

In Sec. 2.6, we considered a fairly detailed model of rotational cooling and described the experimental performance of our scheme. Here, I describe more detailed considerations, especially the power requirements and optimal geometry.

B.1 Power requirements of rotational cooling

We have previously assumed that optical pumping is driven to saturation during rotational cooling. However, we are actually power-limited, reducing the optical pumping efficiency slightly. We use HL6750MG diodes from Thorlabs, which are specified for 50 mW output power. This power is then attenuated through an optical isolator ($\sim 10\%$ loss) and a series of optics including two beam splitters (to couple to the wavemeter and locking cavity), so that an additional $\sim 10\%$ is lost. We typically have $\sim 65\%$ coupling into a PM fiber (but sometimes only 50% for diodes with very bad spatial modes and no more than 80% for diodes with truly extraordinary spatial modes). Thus, we have up to ~ 25 mW of power in the rotational cooling region. We often operate at lower diode currents, extending the diode lifetime, so that we only get $\sim 12 - 15$ mW coupled through the fiber.

B.1.1 Optical setup

In order to use our available power as efficiently as possible, we optimize the optical setup. See Fig. B.1 for the geometry. The quarter-wave plates ensure that subsequent passes have
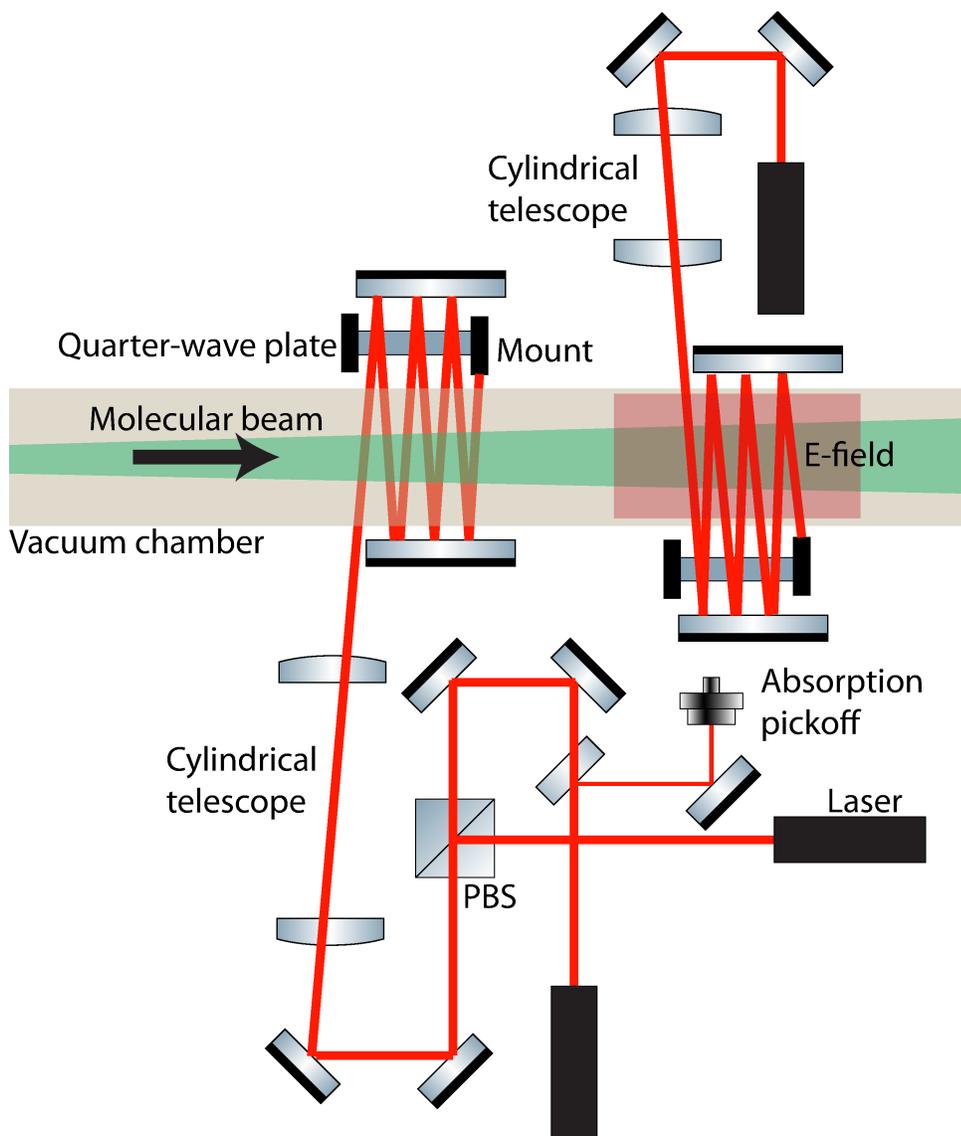
Cylindrical
telescope

Quarter-wave plate   Mount

Molecular beam

E-field

Vacuum chamber

Absorption
pickoff

Cylindrical
telescope

Laser

PBS

Figure B.1: Optics setup for rotational cooling.

orthogonal linear polarizations. The two lower lasers and $J3$ and $J2$ (pumping $J' = 3 \rightarrow$ $J = 2 \rightsquigarrow J' = 1, 3$ and $J' = 2 \rightarrow J = 1 \rightsquigarrow J' = 0, 2$, respectively). The upper laser is $J1$ (pumping $J' = 1 \rightarrow J = 1 \rightsquigarrow J' = 0, 1, 2$).

If a pair of mirrors responsible for the multiple passes are not parallel to each other, then the angle and distance between subsequent laser passes will not be constant. If the mirrors are parallel to each other but not to the molecular beam, then the distance between subsequent laser passes will be constant, but the angle will not be, introducing a double-peak to the Doppler profile. We can effectively broaden our laser by including a small intentional angle between the mirrors and beam line. Any molecule with a transverse velocity will thus see a reduced detuning for one pass and an increased detuning for the next pass, with an effect that is advantagous on net. The Doppler shift is $\delta f = \frac{v}{c} \sin \theta$, where θ is the angle between the laser and the normal to the beam line. Our molecular beam has a forward velocity of approximately 180 m/s, so a 1-degree angular deviation corresponds to a relative Doppler shift of 4.5 MHz [146, Sec. 3.3.2]. In order to minimize the angular deviations among beam passes, we first make sure that all passes are roughly equidistant by eye and then fine-tune using a Doppler scan. Ensuring that subsequent passes are equidistant in the middle of the vacuum chamber also requires us to place the mirrors equidistant from the center of the beam line.

An alternative way to obtain a double-peaked Doppler profile is to detune the laser with respect to the resonance of a molecule with no transverse velocity. If the laser is red-detuned, for example, then odd passes will be more resonant with with molecules approaching the laser, while even passes (counterpropagating to the odd passes) will be more resonant with molecules retreating from the laser. If the laser is blue-detuned, then the opposite is true. This method of obtaining a double peak has the advantage of being insensitive to forward velocity spread, but the difference is small. In practice, we have both slightly detuned lasers and slightly misaligned mirrors.

The Doppler profile of our molecular beam is determined by the collimator geometry. The 1σ Doppler width at 690 nm is 2.6 MHz [257]. Based on simple geometric constraints, the maximum transverse velocity of any molecule is 2.5 m/s given a field plate separation of 4.5 cm, corresponding to $\Delta = 2\pi \times 3.6$ MHz at 690 nm [258]. I independently calculated

$\Delta = 2\pi \times 3.9$ MHz with a similar approach but slightly different geometric assumptions. For conservatism, I will assume that rotational cooling is saturated if optical pumping is efficient for all molecules with $\Delta \leq 2\pi \times 4$ MHz.

There are several frequency scales: $\gamma = 2\pi \times 300$ kHz, the decay rate of the C state [148, Sec. 3.1]; $\Delta = 2\pi \times 4$ MHz, the maximum detuning from resonance; time-of-flight bandwidth $\omega_p = 1/T_p$, where $T_p = \sigma_x/v$ is the time allotted to pump a molecule with forward velocity $v = 180$ m/s [146, Sec. 3.3.2] travelling the horizontal $1\sigma_x$ width of the optical pumping lasers; and Ω, the Rabi frequency.

B.1.2   Optimal laser profile

Vertical shaping

The Rabi frequency at a particular point (for example, $x = \sigma_x$, $y = \sigma_y$, where σ_i is the gaussian width of the laser along $\hat{\imath}$) depends on the beam shape and power. A 2D Gaussian beam with fixed power P has an intensity profile

$$I(x,y) = \frac{P}{2\pi\sigma_x\sigma_y} \exp\left(-\frac{x^2}{2\sigma_x^2}\right) \exp\left(-\frac{y^2}{2\sigma_y^2}\right). \tag{B.1}$$

For a fixed molecular beam half-height (i.e., measured from the center of the beamline) h, the laser intensity is maximized at the edge of the molecular beam provided $\sigma_y = h$. The molecular source has an effective diameter of 7 mm and is constrained by the fixed collimators to have a maximum diameter of 3.9 cm at the end of the interaction region field plates (taking into account the 0.3 cm safety margin between the extrema of the beam and each plate) since we have chosen a plate spacing of 4.5 cm [258]. The distance between the source and the end of the field plates is approximately 154 cm [259]. We thus expect the parts of the molecular beam that will ultimately be detected in the interaction region to have a height h at distance d from the source given by $h = 0.35\,\text{cm} + \frac{3.9-0.7}{2}\frac{d}{154}$. The distance between the source and the end of Step 1 of rotational cooling is about 15 cm, while the distance to the end of Step 2 is about 25 cm. These correspond to molecular beam heights of $h_1 = 0.51$ cm and $h_2 = 0.61$ cm, so the 2σ heights should be 1.02 cm and 1.22 cm, respectively.

Rayleigh length

Determining the optimal horizontal width is much more subtle. First, if the horizontal width is too small, then the Rayleigh length $z_R = 4\pi\sigma^2/\lambda$ can be smaller than the path length as the laser makes many passes, and the beam will significantly expand. Over the path length we use (approximately 10 passes, 15" long each), this is significant for roughly $4\sigma < 1$ mm. This limits the total number of passes that is achievable. At a given relative beam separation (e.g., 4σ between passes), the horizontal width can be set by maximizing the number of passes.

Number of polarization switches

However, the number of passes is not the only optimization condition for the horizontal beam width because only a few passes are required provided the transitions are saturated. To see this, we compute the probability that a molecule has not been either lost or pumped to its target state after n decays. If the branching ratio back to the state addressed by the laser is b, then the probability of being lost or transferred to the target in exactly i decays is $b^{i-1}(1-b) = (\frac{1}{b} - 1)b^i$. Then the probability of being lost or transferred after at least n decays is

$$
\begin{aligned}
(\tfrac{1}{b} - 1)\sum_{i=n}^{\infty} b^i &= (\tfrac{1}{b} - 1)\frac{b^n}{1-b} \\
&= b^{n-1}.
\end{aligned}
\tag{B.2}
$$

For the transitions of interest here, typically $b \sim 0.1^{1/2}$, so we see that only about 10% of molecules require more than 2 decays in order to be either lost or transferred to the target state. Therefore, in the regime that the laser transition is saturated, only about two polarization switches ought to be necessary in total. This estimate is actually conservative because many decays within a given pass return to a bright state and can be pumped again without switching polarizations. Since a well-shaped beam allows for 5 or 6 polarization switches, we see that the number of polarization switches is not limiting.
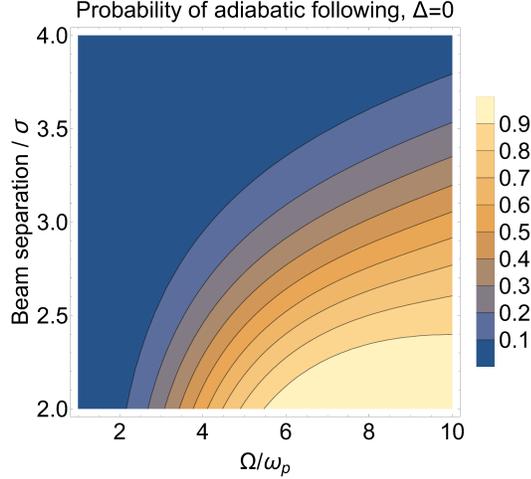
Figure B.2: Adiabatic following on resonance. For $\Omega/\omega_p \sim 5$, we require the maxima of subsequent laser passes to be separated by $\sim 4\sigma$ in order to have negligible adiabatic following. Here, each subsequent contour represents an additional 10% of dark-state molecules that adiabatically follow. This is computed for a three-level system, which differs significantly from our many-level optical pumping systems. Any quantitative interpretation of these results should therefore be taken loosely.

Adiabatic following

In order to set the beam separation, we must understand the molecular dynamics between laser passes. If the driving field is too large, or if the beams are not sufficiently separated, then the molecules in the dark state of one pass will adiabatically follow to the dark state of the next pass, defeating the purpose of polarization switching. Using a simplified three-level model with two Gaussian beams, we can compute the proportion of molecules that adiabatically follow the dark state of the first beam into the dark state of the second beam. See [149, Sec. 4.3] for a similar framework. Adiabatic following is most severe on resonance, so that is the only condition I'll consider here. The required beam separation to achieve a given proportion of molecules that adiabatically follow then depends only on the ratio of Rabi frequency to the frequency corresponding to the timescale of molecular propagation through the lasers, Ω/ω_p. See Fig. B.2 for results. Over a reasonable range of Rabi frequencies such that several Rabi oscillations occur within a beam pass, subsequent laser beams should be separated by $\sim 4\sigma_x$. Given a fixed interaction length, we want to minimize the beam separation for a better "duty cycle" of optical pumping.

Because adiabatic following depends only on Ω/ω_p, the beam waist has a non-trivial

effect on the acceptable beam separation: $\Omega^2 \propto 1/\sigma_x$, but $\omega_p \propto 1/\sigma_x$. Thus $\Omega/\omega_p \propto \sqrt{\sigma_x}$, and minimal adiabatic following is favored by smaller beams.

Magnetic remixing

It is common in optical pumping schemes to use magnetic fields to remix dark sublevels into bright sublevels. The time scale for remixing is typically given by $\tau = 1/(2\pi\vec{\mu} \cdot \vec{\mathcal{B}})$, so if $\tau \ll L/v$ where L is the interaction length and v is the molecule velocity, then a single polarization can suffice to pump out both dark and bright states. For this to work, the magnetic field must be neither parallel nor perpendicular to the laser polarization. This is treated in much more subtle and rigorous detail in [260].

Since the X state of ThO is essentially entirely composed of $^1\Sigma$ character, the magnetic moment is extremely small [159]. However, due to molecular rotation, there is still a magnetic moment on the order of the nuclear magneton μ_N [261]. Because the interaction length available to us for rotational cooling is so large, the condition $\tau \ll L/v$ is possible to satisfy using magnetic fields on the order of 1 G. Unfortunately, this argument breaks down because the Zeeman shift is small compared to the AC Stark shift from the laser field, and the quantization axis is set almost entirely by the laser polarization. In this regime, a dark state $|D\rangle$ evolves into a time-dependent state with bright component, $|\Psi(t)\rangle = |D\rangle + \epsilon(t)|B\rangle$, where $\epsilon(t) = -i\int_0^t dt' \langle B|\vec{\mu} \cdot \vec{\mathcal{B}}|D\rangle e^{-i\Delta E \, t'}$. AC Stark shifts produce an energy shift $\Delta E = \Omega/2$ on resonance, so the steady-state (non-oscillating) contribution to the bright state amplitude component has magnitude $2\langle B|\vec{\mu} \cdot \vec{\mathcal{B}}|D\rangle/\Omega \sim \vec{\mu} \cdot \vec{\mathcal{B}}/\Omega$. For $\mu = \mu_N$, $\mathcal{B} \approx 0.5$ G, and $\Omega \sim 1$ MHz, this amplitude is extremely small and the magnetic field does not induce a significant enough bright-state population to pump out of. Therefore, magnetic remixing is not sufficient and polarization switching must be used.

Saturation condition

We want to maximize the time that the molecules spend in the excited state so that it has the most opportunity to decay into the target state. Using a simple two-level model, it is straightforward to compute that the probability of occupying the excited state after time t, with detuning Δ and Rabi frequency Ω, is

$$P(t) = \frac{\Omega^2}{\Delta^2 + \Omega^2} \sin^2(\sqrt{\Omega^2 + \Delta^2}\, t/2). \tag{B.3}$$

To be conservative, I'll consider an interaction length from $-\sigma_x$ to $+\sigma_x$ at a vertical distance at $y = \sigma_y$ and treat the beam as a flat-top with intensity $I = P/(2\pi e \sigma_x \sigma_y)$. The intensity is related to the electric field $\mathcal{E}$ by $I = cn\epsilon_0 \mathcal{E}^2/2$, where c is the speed of light, $n = 1$ is the index of refraction of vacuum, and ϵ_0 is the permittivity of space. This field is related to the Rabi frequency by $\Omega = |\vec{d} \cdot \vec{\mathcal{E}}|$, where $\vec{d}$ is the transition dipole moment. The transition moment between the X and C states in ThO is $|d_{X-C}| = 0.52 \pm 0.08\, ea_0$ [148, Sec. 3.3.3]. Any transition between particular rotational and magnetic levels must also include normalized Hönl-London factors and Clebsch-Gordan coefficients.

As long as $\sqrt{\Omega^2 + \Delta^2}(\sigma/v) > \pi$, we can average the time spent in an excited state over a molecular beam pass such that the average excited-state population is $\frac{1}{2}\frac{\Omega^2}{\Delta^2 + \Omega^2}$ over a duration $2\sigma/v$. As a conservative example, with $\sqrt{\Omega^2 + \Delta^2} = 2\pi \times 1$ MHz and $\sigma = 0.25$ mm, the LHS of this condition is $\sim 3\pi$. Thus the expected time T that a molecule spends in the excited state is

$$T = \sum_{i=1}^{N} \frac{\sigma_i}{v} \frac{\Omega_i^2}{\Delta^2 + \Omega_i^2}, \tag{B.4}$$

where I've allowed the possibility of N passes with different beam widths and Rabi frequencies.

If we require n decay times in the excited state, then we need $T \geq n\tau$, where $\tau = 2\pi/\gamma$. Based on the discussion in Sec. B.1.2, we should require about 2 decays, and we ought to allot 2τ per decay in order to have a high probability of decaying. Therefore, I'll let $n = 4$. Since $\tau = 490 \pm 40$ ns [148, Sec. 3.1], we need

$$vT = \sum_{i=1}^{N} \frac{\Omega_i^2}{\Delta^2 + \Omega_i^2} \sigma_i \geq 350\, \mu\text{m}. \tag{B.5}$$

Using the relations above, we can compute $\Omega^2 = d^2 P/(c\epsilon_0 \pi e \sigma_x \sigma_y) \equiv \zeta/\sigma_x$, where I've defined the constant ζ. Then it's easy to work out that the condition above is

$$F \equiv \sum_{i=1}^{N} \frac{\sigma_{x,i}}{\sigma_{x,i} + \zeta/\Delta^2} \frac{\zeta}{\Delta^2} \geq 350\,\mu\text{m}, \tag{B.6}$$

where F is a figure of merit (with units of length) for the optical pumping geometry.

To leading order in Ω^2/Δ^2, the LHS reduces to $N\zeta/\Delta^2$ (in other words, it becomes independent of $\sigma_{x,i}$). However, it's easy to show that the LHS scales at higher order with $\sigma_{x,i}$, provided N is fixed.

We therefore have the following competing constraints on setting $\sigma_{x,i}$:

- Very small beam waists will lead to small Rayleigh lengths and thus constrain the number of passes, favoring larger beams.

- More passes can be fit with smaller beam waists (provided the effect of the Rayleigh length is not dominant), favoring smaller beams.

- Adiabatic following is more severe for larger ratios $\Omega/\omega_p \propto \sqrt{\sigma_x}$, favoring smaller beams.

- The figure of merit F increases slightly with $\sigma_{x,i}$, provided the number of passes is fixed, favoring larger beams.

We can handle all of these conditions by maximizing F in a given optical geometry, taking account of the beam divergence, with the beam separation as a free parameter. If we then fix the separation under the constraint that adiabatic following is negligible, we will have a near-optimal configuration. It's possible that the optimum occurs when there is a small degree of adiabatic following but many passes; however, this should not be a dominant effect.

Beam divergence

The simplest optical configuration is that shown in Fig. B.1: an aspheric lens is used to collimate light in the horizontal direction from an optical fiber. Since we only use one lens, we cannot necessarily reach the minimal divergence at a given beam waist set by diffraction. In the configuration shown in Fig. B.3, we can freely choose the focal length f of the asphere
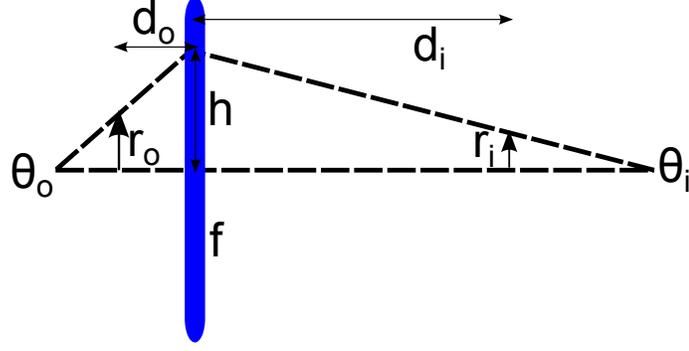
Figure B.3: Parameters determining the beam waist and divergence at a given distance from the optical fiber

and the object-lens distance d_o to set the image-lens distance d_i and waist size r_i as desired. Then θ_i is fixed by geometry.

Specifically, the object size r_0 is set by the mode field diameter of the optical fiber, $MFD = 4\sigma$ and the divergence is set by the numerical aperture $NA = \sin \theta_o$. For Thorlabs P1-630PM-FC patch cables, $NA = 0.12$ and $MFD = 4.2\,\mu\text{m}$. Letting $r = 2\sigma$, we know that $r_o = MFD/2 = 2.1\,\mu\text{m}$.

The image and object distance are related by the thin-lens equation,

$$\frac{1}{f} = \frac{1}{d_o} + \frac{1}{d_i}. \tag{B.7}$$

The image size r_i is determined in terms of θ_i by conservation of etendue, $r_o \sin \theta_o = r_i \sin \theta_i$. The last condition we need to determine θ_i is based on the lens geometry: $h = r_o + d_o \tan \theta_o = r_i + d \tan \theta_i$. For a given desired waist size r_i at distance d_i, we can compute the necessary focal length f and object distance d_o, along with the consequential waist divergence θ_i:

$$\theta_i = \sin^{-1}\left(\frac{r_0 \sin \theta_o}{r_i}\right)$$

$$d_o = \frac{d_i \tan \theta_i + r_i}{\tan \theta_o} - r_o \tag{B.8}$$

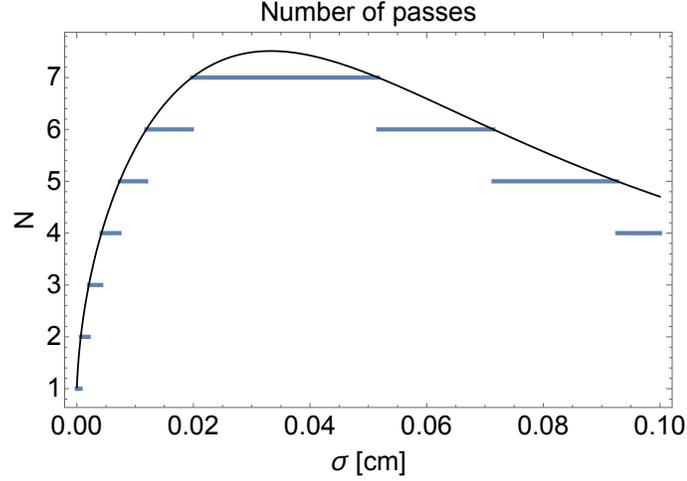$$f = \frac{1}{\frac{1}{d_o} + \frac{1}{d_i}}.$$

293

Figure B.4: The number of passes that can fit on the $\lambda/4$ for a beam 1.25 cm tall, as a function of the 1σ focal size along x.

Since this is based only on geometric optics, we ought to check that the computed parameters don't violate the diffraction limit, $\theta_i > r_i/z_R$, where z_R is the Rayleigh range. It's straightforward to check that in the range of interest with $\sigma \in (100\,\mu\mathrm{m}, 1\,\mathrm{mm})$, this approach doesn't violate the diffraction limit.

B.1.3   Computing the required power

Figure of merit

In order to use the figure of merit F, we need to compute the number of passes the beam can make. Taking account of the beam divergence θ, the 1σ width of the Nth pass of a beam with focused width σ_0 is $\sigma_N = \sigma_0 + \frac{N-1}{2}\theta L$, where $L = 27$ cm is the path length between passes. Thus if N beams are uniformly spaced by $4\sigma_N$ (to avoid significant adiabatic following), the total interaction length is $L_{\mathrm{int}} = 4N\sigma_0 + 2N(N-1)\theta L$. The beam interaction length is limited by the 1" diameter quarter-wave plate used for polarization switching. For a beam 1.25 cm tall, the unclipped length along the wave plate is just over $l = 2$ cm. We can solve for N by setting $L_{\mathrm{int}} = l$ and then rounding down to the nearest integer. The number of passes achievable as a function of focused beam size is shown in Fig. B.4.

Given this, and the size of the ith beam, we can compute F as a function of the focused beam size σ_o and compare to 350 μm (the estimated requirement to saturate optical pump-

ing). It turns out that variations in beam size don't affect the efficiency much within a range that allows a fixed number of passes, and the optimal size occurs with the maximum number of passes. Further, the profile of $F(\sigma)$ doesn't depend strongly on ζ/Δ^2. These features are consequences of the limit $\zeta/\Delta^2 \ll \sigma$. Since the figure of merit F doesn't vary strongly in the range $\sigma \in (200, 500)\,\mu$m (where 7 laser passes can be made), for any fixed value of ζ/Δ^2, it is useful to compute $F(\zeta/\Delta^2, \sigma = 350\,\mu$m). In the range of interest, we find that $F \propto \zeta/\Delta^2$ to a good approximation.

Using these results, we find that $F > 350\,\mu m$ when $\zeta/\Delta^2 > 30\,\mu$m. For reference, if $\Delta = 2\pi \times 1$ MHz, $P = 1$ mW, $\sigma_y = 1.25$ cm, and $d = 1\,ea_0$, then $\zeta/\Delta^2 = 580\,\mu$m. It is then straightforward to scale ζ/Δ^2 according to the actual experimental parameters. All of this assumes that the detuning is fixed for all passes, so the mirrors must be parallel to the beamline. If we angle the mirrors so that every other pass preferentially addresses a different velocity class, to leading order this has the effect of letting $N \rightarrow N/2$ and $\Delta \rightarrow \Delta/2$, so that $F \rightarrow 2F$.[1] The saturation condition is then equivalent to requiring $F = 700\,\mu$m when using $\Delta = 2\pi \times 2$ MHz to account for angled mirrors, leading in turn to the requirement $\zeta/\Delta^2 > 60\,\mu$m.

Transition dipole moment

The most uncertain input parameter to the constant ζ is d, the transition dipole moment. We have measured $|d_{X-C}| = 0.52 \pm 0.08\,ea_0$ [148, Sec. 3.3.3]. The transition strength between given rotational levels is then modified by the normalized Hönl-London factor, Franck-Condon factor FC $= 0.84$, and Clebsch-Gordan coefficients. For the transitions $J' = 3 \rightarrow J = 2$ and $J' = 2 \rightarrow J = 1$ in the absence of an electric field, the contribution of squared Clebsch-Gordan coefficients is $1/3$ on average, assuming a uniform distribution over sublevels. For $J' = 1 \rightarrow J = 1$, only $m = \pm 1$ are excited and the squared Clesbch-Gordan coefficients contribute a factor of $1/2$ for each laser polarization. The contribution of Hönl-London factors is given in Table 2.4. All contributions together are shown in Table

---

1. This is a variation of artificially broadening the laser linewidth to more efficiently use the available power. It is of course not entirely ideal because opposite polarizations see different average detunings and thus pump with different efficiencies, but the overall pumping efficiency can be increased somewhat.

	$3 \to 2$	$2 \to 1$	$1 \to 1$
$(\text{CG})^2$	$1/3$	$1/3$	$1/2$
HL	$2/5$	$1/3$	$1 - \eta_0$
FC	0.84	0.84	0.84
$(\text{CG})^2 \times \text{HL} \times \text{FC}$	0.11	0.093	0.21
$d^2 = d^2_{X-C} \times (\text{CG})^2 \times \text{HL} \times \text{FC}\ [(ea_0)^2]$	0.030	0.025	0.057

Table B.1: Computation of the transition dipole moments used in rotational cooling. The value $\eta_0 = 1/2$ is used.

B.1.

Saturation power

Using these results, we can compute the necessary power P such that $\zeta/\Delta^2 = 60\,\mu\text{m}$ using $\Delta = 2\pi \times 2$ MHz, $\sigma_0 = 350\,\mu\text{m}$, $\sigma_y = 1.25$ cm, and the transition moments shown in Table B.1. We then find $P_{3 \to 2} = 14$ mW, $P_{2 \to 1} = 16$ mW, and $P_{1 \to 1} = 7$ mW. Note that this is not an ordinary "saturation power." Instead, it's the power required for the most detuned molecule at the vertical edge of the laser beam to be in the excited state for 4 decay times (also assuming a flat top beam that slightly underestimates the average intensity). The majority of molecules will be either closer to the center of the beam or less detuned, and therefore require less power than these values.

With 7 passes and the optical geometry as described in previous sections, we find that the $1/e^2$ saturation powers are 8 mW, 7 mW, and a few mW for $J1$, $J2$, and $J3$, respectively (the SNR for the $J3$ laser saturation curve is poor because the gain is so small, and we can only resolve a few percent change in molecular signal). Thus with 10-15 mW for each laser, we can fully saturate rotational cooling. This appears to be consistent with our estimates of the power required to efficiently optically pump the most outlying molecules (in both position and velocity), which lead to power requirements a factor of ~ 2 more stringent than for the molecular ensemble overall.

B.2 Decay with coherences

In the foregoing treatment, we have assumed that the coherences in excited states, which arise during optical pumping, don't have any effect on branching ratios. Using the density

matrix formalism, I show here that this assumption is correct, as would be naively expected.

Note that while this calculation is relatively straightforward and has a trivial result, I include it here to serve as a reference about similar calculations using the density operator. The lecture notes by John Preskill [262] give an especially useful overview of the density operator approach.

B.2.1 Density matrices and the master equation

The density operator encodes all available information about an ensemble of quantum mechanical states, and is the natural formalism to describe the interactions of a system with its environment. It is usually expressed as a matrix in a particular basis; as usual, it is simplest (but not necessary) to use an orthonormal basis. Diagonal entries are the probabilities for a system to be found in a particular state of the basis, while off-diagonal entries provide information about coherences.

The density operator evolves in time under the action of a "superoperator," which maps operators to operators. A physical time-evolution process will map a density operator to another density operator, so the superoperator must be trace-preserving, hermiticity-preserving, and positive to ensure reasonable behaviors of probabilities. Further, the superoperator is usually taken to be linear in its argument, which simplifies (but might not be necessary for) the ensemble interpretation of the density operator.

To further constrain the equation of motion for the density operator, we also require that the time-evolution process be memoryless. In particular, terms like $\int_{-\infty}^{t} dt'\, f(t')\rho(t')$ in an equation of motion are consistent with all previous conditions. Physically, these terms arise from information leaking into the environment and then later leaking backing into the system. The memorylessness (or Markov) condition removes these contributions to the evolution of the density operator.

The most general equation of motion for the density operator consistent with these conditions is called the Lindblad equation, or sometimes the "master equation":

$$\dot{\rho} = -i[H,\rho] + \sum_{k}^{N^2-1} C_k \rho C_k^\dagger - \frac{1}{2}\{\rho, C_k^\dagger C_k\}. \tag{B.9}$$

As usual, $[A, B]$ represents a commutator, and $\{A, B\}$ represents an anticommutator. N is the dimension of the Hilbert space describing the system. The first term usually describes unitary evolution for a given Hamiltonian (though H is not unique and cannot always be identified as the Hamiltonian since transformations of H and each C_k exist that leave the equation of motion unchanged), the transformation of ρ under each C_k describes "dissipative" effects, and the anticommutator enforces properties of the time-evolution superoperator like norm-preservation. Each of the $N^2 - 1$ C_k operators is called a Lindblad operator, or sometimes a Krauss operator. There are N^2 Krauss operators in total, with the last one accounting for the Hamiltonian and anticommutator terms.

B.2.2 Describing decays

We will take H as the usual Hamiltonian and let each C_k describe a separate decay channel. Note that for N distinct states, there are only $N^2 - N < N^2 - 1$ pairings of distinct states, so we never need to describe multiple decays in the same Lindblad operator. All molecules that decay to unwanted states are assumed to fall into the same "dump" state for simplicity. The decay $A \rightsquigarrow B$ is described by an operator C^{BA} with matrix elements $C_{ij}^{BA} = \sqrt{\gamma_{BA}}\,\delta_{iB}\delta_{jA}$ in the obvious orthonormal basis.

I will show that a state s uncoupled by the Hamiltonian to other states accumulates in a "classical" way, provided each Lindblad operator couples s to at most one other state (there is another less intuitive condition that we will see, which is satisfied for our usual Lindblad operators):

$$
\begin{aligned}
\dot{P}_s &= \dot{\rho}_{ss} \\
&= \sum_i \left(\gamma_{si} P_i - \gamma_{is} P_s\right),
\end{aligned}
\tag{B.10}
$$

where P_i is the population of the i-th state. The equation above is just a classical rate equation including decays both into and out of state s.

Explicitly, we assume that $H_{is} = H_{si} = \epsilon_s \delta_{is}$ so that s is not coupled by the Hamiltonian to any other states. Then the commutator in the first term of the Lindblad equation for $\dot{\rho}_{ss}$ is

$$
\begin{aligned}
[H, \rho]_{ss} &= H_{si}\rho_{is} - \rho_{si}H_{is} \\
&= \epsilon_s \delta_{is}\rho_{is} - \epsilon_s \delta_{is}\rho_{si} \\
&= \epsilon_s \rho_{ss} - \epsilon_s \rho_{ss} \\
&= 0.
\end{aligned}
\tag{B.11}
$$

We now consider the effect of the Lindblad operators on the change in population of state s. Consider the contribution from just one Lindblad operator C, and assume that C couples s to at most one other state q so that $C_{si} = c_{sq}\delta_{iq} + c_{ss}\delta_{is}$ and $C_{is} = c_{sq}\delta_{iq} + c_{ss}\delta_{is}$. Then we need to compute the contribution to the evolution of ρ involving C,

$$
C_{si}\rho_{ij}(C^\dagger)_{js} - \tfrac{1}{2}\rho_{si}(C^\dagger)_{ij}C_{js} - \tfrac{1}{2}(C^\dagger)_{si}C_{ij}\rho_{js}
\tag{B.12}
$$

$$
= C_{si}\rho_{ij}C_{sj}^* - \tfrac{1}{2}\rho_{si}C_{ji}^*C_{js} - \tfrac{1}{2}C_{is}^*C_{ij}\rho_{js}.
$$

It is easiest to find the coefficient of each possible entry in ρ. For example, consider the coefficient of ρ_{ab} where $a, b \neq s, q$. Only the first term above can possibly contribute, but its contribution is $C_{sa}\rho_{ab}C_{sb}^* = 0$ since $C_{sa} = 0$. Similarly, it is easy to see that the coefficient of ρ_{ss} is $-|C_{qs}|^2$ and the coefficient of ρ_{qq} is $|C_{sq}|^2$.

The coefficient of ρ_{sq} is $(1/2)(C_{sq}^*C_{ss} - C_{qs}C_{qq}^*)$ and the coefficient of ρ_{qs} is its complex conjugate. Therefore there are no contributions from coherences (off-diagonal terms) as long as $C_{sq}^*C_{ss} - C_{qs}C_{qq}^* = 0$, which is achieved (for example) when $C_{ss} = C_{qq} = 0$.

We define $|C_{ij}|^2 = \gamma_{ij}$. Then, the contribution to $\dot{P}_s$ from the Lindblad operator coupling s to q is $\gamma_{sq}P_q - \gamma_{qs}P_s$. Summing over Lindblad operators then recovers the rate equation above, with no coherent effects. We can even generalize to the case in which state s is coupled to other states by the Hamiltonian, in which case

$$
\dot{P}_s = -i[H, \rho]_{ss} + \sum_i \left(\gamma_{si}P_i - \gamma_{is}P_s \right),
\tag{B.13}
$$

and the contribution of decays is still described in the manner of a classical rate equation.

Appendix C

Comments on the parity sum

C.1   Parity sum decomposition

In Sec. 3.1, I described how we use the "parity sum" formalism for data analysis. In particular, for some quantity $X(s_1, \cdots, s_n)$ measured with a configuration of binary switches $s_1, \cdots, s_n$ each taking value ± 1, we write

$$X(s_1, \cdots, s_n) = X^{[\mathrm{nr}]} + s_1 X^{s_1} + \cdots + s_n X^{s_n} + s_1 s_2 X^{s_1 s_2} + \cdots + (s_1 \cdots s_n) X^{s_1 \cdots s_n}. \quad \text{(C.1)}$$

We interpret a term $X^{s_a \cdots s_z}$ as being the contribution to a measured value of X that is odd under switches $s_a, \cdots, s_z$ and even under all other switches.

A bit more concretely, this means that the value of X measured in a configuration with all switches set to $s_i = +1$ is given by

$$X(+, +, \cdots, +) = X^{[\mathrm{nr}]} + X^{s_1} + X^{s_2} + \cdots + X^{s_n} + X^{s_1 s_2} + \cdots X^{s_1 \cdots s_n}, \quad \text{(C.2)}$$

and reversing the sign of switch s_i introduces a relative sign on the RHS for any term in which s_i appears explicitly as a superscript. We can write a parity component as

$$X^{s_a s_b \cdots s_z} = \frac{1}{2^n} \sum_{s'_1, \cdots, s'_n = \pm 1} X(s'_1, \cdots, s'_n) s'_a \cdots s'_z. \quad \text{(C.3)}$$

If X has a component that reverses sign under s_a, then the terms on the RHS of Eq. C.3

300

with $s'_a = \pm 1$ reinforce each other; otherwise these terms cancel. The same argument holds for all reversing components up to s'_z; therefore, this formula extracts the contribution to the measured value of X that is odd under the reversal of any switch in $\{s_a, \cdots, s_z\}$. For any $s_i \notin \{s_a, \cdots, s_z\}$, a contribution to X that is odd under s_i will cancel in the sum over $s'_i = \pm 1$, while a contribution even under s_i will reinforce. We therefore know that Eq. C.3 extracts the contribution to the measured value of X that reverses sign with $s_a, \cdots, s_z$ but no other switch. The prefactor of 2^n is necessary because a total of 2^n signed states are included in the sum. (Showing explicit agreement between Eqs. C.1-C.3 is rather tedious.)

C.2   Parity sum as a special case of least squares

The parity sum as a data analysis tool is just a special case of least squares regression (see, e.g., [263]). In the general case, we have n observations $y_1, \cdots, y_n$ and fit to a model $\vec{y} = X\vec{\beta} + \vec{\epsilon}$, where the residuals $\vec{\epsilon}$ are normally distributed for each observation (though possibly with different variances). Here, X is an $n \times k$ matrix, and $\vec{\beta}$ is a k-dimensional vector, specifying the k "variables" in the model. The entry X_{ij} describes the value of the j-th variable taken in the i-th observation of y. Written this way, it's clear that an estimate $\hat{\beta}$ of the parameter values $\vec{\beta}$ can be obtained with $\hat{\beta} = X^{-1}\vec{y}$ provided X is invertible. Even when X is not invertible, the parameters can be estimated by $\hat{\beta} = (X^T X)^{-1} X^T \vec{y}$.

Note that this estimate of $\vec{\beta}$ might have larger uncertainty than necessary: in particular, if some observations are much noisier than others, then it will be advantageous to perform generalized (rather than ordinary) least squares regression, for which the relative uncertainties of (and covariances among) different measurements is taken into account. However, we do *not* want to weight certain experimental states more than others, since this could introduce bias to our results (if, for instance, some experimental configurations systematically have less signal than others, while also systematically giving a larger precession frequency in the EDM channel).

To take our usual parity sum approach in ACME as an example, $\vec{\beta}$ is the vector of "parity components," and X is a matrix of switch states over all measured states of the experiment. For example, we might have a single switch s, in which case $\vec{\beta}^T = [X^{[\mathrm{nr}]}\ X^s]$.

If we take two measurements, with $s = +1$ followed by $s = -1$, then

$$X = \begin{bmatrix} +1 & +1 \\ +1 & -1 \end{bmatrix}. \tag{C.4}$$

This perspective makes it quite clear that we can "generalize" the parity sum approach (for example, to the case of a switch with three states) by simply performing ordinary least squares with some appropriate model to distinguish physically distinct contributions to the phase (or contrast, frequency, etc.).

Appendix D

AC Stark and Zeeman shifts

In this appendix, we will consider a few specific scenarios for generating systematic error terms, distinguished from each other by the intermediate states through which $M = \pm 1$ levels are perturbed. I'll drop factors of order unity throughout this initial treatment. We use the notation developed in Sec. 4.3. As a reminder, we take Δ to be a characteristic energy splitting betwen $|H, J = 1, M = \pm 1\rangle$ and some relevant intermediate state. The three types of intermediate states we'll consider occur in:

- The Ω-doublet structure in $|H, J = 1\rangle$, with splitting $\Delta \sim 100$ MHz

- Higher rotational levels in $|H\rangle$, $\Delta \sim 10$ GHz

- The $|Q\rangle$ state or other electronic states, $\Delta \sim 10$ THz

Let's first estimate a conservative order of magnitude for the second-order perturbations involved. We will have terms at the order of $c^{(2)} \sim \frac{V^2 t}{\Delta}$. Suppose the perturbations V arise from AC Stark or Zeeman shifts. If the AC fields are produced by the molecules flying through DC field gradients, then V represents the amplitude by which the Stark or Zeeman energy changes throughout the molecule's trajectory. Under normal conditions, we have magnetic field gradients of $\sim 10\,\mu\text{G/cm}$, giving up to $100\,\mu\text{G}$ field difference between the preparation and readout regions, in addition to $\sim 100\,\mu\text{G}$ or smaller offset fields. Then $V \sim g_{H(Q)}\mu_B \times 100\,\mu\text{G} \sim 2\pi \times 10^{-2}(1)$ kHz, where I have supposed that interactions mediated by the Q or other electronic states involve g-factors of order unity. The largest electric field

gradients are produced by the curvature of the field plates, which give misalignments up to $\sim 10\,\mu\text{m}/10\,\text{cm} \sim 10^{-4}$ rad. This in turn generates up to $\mathcal{E}_x \sim 10^{-4} \times 100\,\text{V/cm} \sim 10$ mV/cm fields that vary across the molecules' flight, and a characteristic Stark shift of $\sim 2\pi \times 10$ kHz. In all cases, the precession time is $t \sim 1/\text{kHz}$.

We consider here contributions to the second-order perturbation of the molecular state, $c^{(2)}$, involving $\mathcal{B}_\perp^2$, $\vec{\mathcal{E}}_\perp \cdot \vec{\mathcal{B}}_\perp$, and $\mathcal{E}_\perp^2$. In all cases, the $\mathcal{E}_\perp^2$ term dominates, but other terms may have a more concerning behavior with respect to our experimental switches. Higher-order corrections are suppressed further by factors of $\sim \mathcal{E}_\perp(\mathcal{B}_\perp)/\Delta \leq 10^{-4}$. Any possible mixing through excited vibrational levels should be qualitatively similar to mixing through the Ω-doublet or rotationally excited states, but the couplings are suppressed by much larger detunings, $\Delta \sim 1$ THz. Similarly, spin-orbit mixing with $^1\Pi_1$ and $^3\Pi_1$ states modifies the effective electric and magnetic dipole moments of the H state, but does not act on the lab-frame quantities J, M, Ω and thus has no effect on the qualitative behavior of Stark and Zeeman perturbations.

In ACME II, we are sensitive to $\sim \mu$rad phases, so we are concerned with perturbation coefficients of order $c^{(2)} \sim 10^{-6}$ or larger. Here, I'll examine all terms that can generate $c^{(2)} > 10^{-7}$ by the most naive order-of-magnitude estimate. Below is a fairly exhaustive list of the couplings between $M = \pm 1$ states that might arise due to time-dependent electric and magnetic fields. I always consider the largest numerical contribution to the perturbation for purposes of estimating the order of magnitude.

1. Various effects within the H state Ω-doublet with $J = 1$

    (a) Second-order mixing, $c^{(2)} \sim \frac{\mathcal{E}_\perp^2 t}{\Delta_\mathcal{N}} \sim 10^{-3}$

    (b) Corrections due to time-dependent parallel (i.e., $\hat{z}$-aligned) fields, $c^{(2)} \sim \frac{D^2 \mathcal{E}_\perp^2 t}{\Delta_\mathcal{N}} \times \frac{D \delta \mathcal{E}_z}{\Delta_N} \sim 10^{-7}$

    (c) Transitions involving the "wrong" Ω state due to incomplete molecular polarization, $c^{(2)} \sim \frac{D^2 \mathcal{E}_\perp^2 t}{\Delta_\mathcal{N}} \times \left(\frac{\Delta_\Omega}{D|\mathcal{E}|}\right)^2 \sim 10^{-9}$

    (d) Higher-order corrections to second-order mixing coefficients, $c^{(2)} \sim \frac{D^2 \mathcal{E}_\perp^2 t}{\Delta_\mathcal{N}} \times \frac{t}{\Delta_\mathcal{N}} \sim 10^{-8}$

(e) Corrections to the splitting in the Ω-doublet, $c^{(2)} \sim \frac{D^2 \mathcal{E}_\perp^2 t}{\Delta_\mathcal{N}} \times \frac{g_H \mu_B |\mathcal{B}|}{\Delta_\mathcal{N}} \sim 10^{-8}$

   i. Note: the effects of both $\tilde{\mathcal{P}}$ states of $M = 0$ further suppress corrections involving $\Delta_\Omega \neq 0$

2. Transitions involving $J = 2$ of the H state

   (a) Transitions between the $J = 1$ manifold and $J = 2$, $c^{(2)} \sim \frac{D^2 \mathcal{E}_\perp^2 t}{B_{\text{rot}}} \sim 10^{-5}$

   (b) Mixing with $J = 2$ due to the ordinary Stark shift, transitions within the $J = 2$ doublet, and mixing back to $J = 1$: $c^{(2)} \sim \left(\frac{D |\mathcal{E}|}{B_{\text{rot}}} \right)^2 \times \frac{D^2 \mathcal{E}_\perp^2 t}{\Delta_\mathcal{N}} \sim 10^{-7}$

   (c) Mixing with $J = 2$ due to the ordinary Stark shift, transition within the $J = 2$ doublet, and transition back to $J = 1$: $c^{(2)} \sim \left(\frac{D |\mathcal{E}|}{B_{\text{rot}}} \right) \times \frac{D \mathcal{E}_\perp}{\Delta_\mathcal{N}} \times \frac{D \mathcal{E}_\perp}{B_{\text{rot}}} \sim 10^{-12}$

3. Transitions involving the Q or other electronic states, $c^{(2)} \sim \frac{D^2 \mathcal{E}_\perp^2 t}{A} \sim 10^{-7}$

   (a) Note: the spin-uncoupling interaction that mixes H with Q only applies to $J \geq 2$ states

   (b) Selection rules of the Stark interaction actually prohibit coupling H and Q, and other states (e.g., A) are too far away to contribute effects at the 10^{-7} level.

   (c) Zeeman interactions with the Q or other electronic states are too weak to contribute at the level of interest

Based on these estimates, I consider cases 1(a-b) and 2(a-b) in greater detail. Note that only case 1(a) is large enough to still be significant if a factor of $g_H \mu_B \mathcal{B}_\perp$ is substituted for a factor of $D \mathcal{E}_\perp$, so I'll ignore magnetic fields in cases 2(a-b).

Matrix elements

We start by computing the matrix elements for transverse electric and magnetic fields $\mathcal{E}_{x,y}$ and $\mathcal{B}_{x,y}$.

Stark matrix elements

The Stark interaction is $\vec{D} \cdot \vec{\mathcal{E}}$ and has matrix elements (see [146, Eq. 2.25] or Eq. 2.9)

$$\langle JM\Omega; \Lambda S\Sigma|\vec{D} \cdot \vec{\mathcal{E}}|J'M'\Omega'; \Lambda'S'\Sigma'\rangle = \delta_{SS'}\delta_{\Sigma\Sigma'}(-1)^{M'-\Omega}D_{\Omega-\Omega'}\mathcal{E}_{M'-M}$$

$$\times \sqrt{(2J+1)(2J'+1)}\begin{pmatrix} J & 1 & J' \\ -\Omega & (\Omega-\Omega') & \Omega' \end{pmatrix}\begin{pmatrix} J & 1 & J' \\ -M & (M-M') & M' \end{pmatrix},$$

where I am using the shorthand $\mathcal{E}_{M'-M} \equiv T^1_{M'-M}(\mathcal{E})$ and $D_{\Omega-\Omega'} \equiv \langle\Lambda|T^1_{\Omega-\Omega'}(D)|\Lambda'\rangle$.

Note that in all cases presently of interest, we have transitions only within the H state, so $\delta_{SS'}\delta_{\Sigma\Sigma'} = 1$. Since $\Delta\Omega = \pm1, 0$ for vector operators and only $\Omega = \pm1$ states exist in the H manifold, we will always have $\Omega' = \Omega$ so that $D_{\Omega-\Omega'} \equiv D_\parallel$. Then

$$\langle JM\Omega; \Lambda S\Sigma|\vec{D} \cdot \vec{\mathcal{E}}|J'M'\Omega'; \Lambda'S'\Sigma'\rangle \to (-1)^{M'-\Omega}D_\parallel\mathcal{E}_{M'-M}$$

$$\times \sqrt{(2J+1)(2J'+1)}\begin{pmatrix} J & 1 & J' \\ -\Omega & 0 & \Omega \end{pmatrix}\begin{pmatrix} J & 1 & J' \\ -M & (M-M') & M' \end{pmatrix}.$$

1(a-b): Transitions within $J = 1$ manifold

For transverse electric fields causing transitions with the $J = 1$ manifold, we need $J = J' = 1$, $M = 0$, $M' = \pm1$, and $\Omega = \pm1$, up to complex conjugation of the matrix element depending on whether we're driving from or to $M = 0$. Then we compute (with a change in priming convention),

$$\langle 1, 0, \Omega; \Lambda S\Sigma|\vec{D} \cdot \vec{\mathcal{E}}|1, M, \Omega; \Lambda S\Sigma\rangle = -\frac{M\Omega}{2} \times D_\parallel\mathcal{E}_M. \tag{D.1}$$

The usual Stark shift in $J = 1$ due to z-aligned fields, important for case 1(b), has a similar form except for the substitution $\mathcal{E}_z \leftrightarrow -\mathcal{E}_M$:

$$\langle 1, M, \Omega; \Lambda S\Sigma|\vec{D} \cdot \vec{\mathcal{E}}|1, M, \Omega; \Lambda S\Sigma\rangle = +\frac{M\Omega}{2} \times D_\parallel\mathcal{E}_0. \tag{D.2}$$

2(a): Transitions between $J = 1$ and $J = 2$ due to perpendicular fields

Here, we let $J' = 1$, $J = 2$, $M = 0$, $M' = \pm 1$, and $\Omega = \pm 1$. Then we consider

$$\langle 2, 0, \Omega; \Lambda S\Sigma | \vec{D} \cdot \vec{\mathcal{E}} | 1, M, \Omega; \Lambda S\Sigma \rangle = -\frac{1}{2\sqrt{5}} \times D_\| \mathcal{E}_M. \tag{D.3}$$

Also consider $M = 2M'$ (and other quantum numbers the same), for

$$\langle 2, 2M, \Omega; \Lambda S\Sigma | \vec{D} \cdot \vec{\mathcal{E}} | 1, M, \Omega; \Lambda S\Sigma \rangle = -\frac{\sqrt{3}}{5} \times D_\| \mathcal{E}_M. \tag{D.4}$$

2(b): Mixing between rotational levels and transitions within $J = 2$

Nick Hutzler works out the matrix elements for mixing between adjacent rotational levels due to z-aligned fields [146, Eq. 6.19]:

$$\langle 2, M, \Omega; \Lambda S\Sigma | \vec{D} \cdot \vec{\mathcal{E}} | 1, M, \Omega; \Lambda S\Sigma \rangle = +\frac{1}{2}\sqrt{\frac{3}{5}} \times D_\| \mathcal{E}_0. \tag{D.5}$$

We will also need M-changing matrix elements with $J = J' = 2$. Here, I still assume $M = \pm 1$:

$$\langle 2, 0, \Omega; \Lambda S\Sigma | \vec{D} \cdot \vec{\mathcal{E}} | 2, M, \Omega; \Lambda S\Sigma \rangle = -\frac{M\Omega}{2\sqrt{3}} \times D_\| \mathcal{E}_M \tag{D.6}$$

$$\langle 2, 2M, \Omega; \Lambda S\Sigma | \vec{D} \cdot \vec{\mathcal{E}} | 2, M, \Omega; \Lambda S\Sigma \rangle = +\frac{M\Omega}{\sqrt{6}} \times D_\| \mathcal{E}_{-M} \tag{D.7}$$

Zeeman matrix elements, 1(a): Transitions within $J = 1$ manifold

The only Zeeman matrix element we need to consider here is of the form

$$\langle 1, 0, \Omega; \Lambda S\Sigma | g_L \mu_B \vec{L} \cdot \vec{\mathcal{B}} + g_S \mu_B \vec{S} \cdot \vec{\mathcal{B}} | 1, M, \Omega; \Lambda S\Sigma \rangle. \tag{D.8}$$

Let $\vec{v} = \vec{S}, \vec{L}$ for convenience. Nick Hutzler shows that [146, Eqs. 2.34-2.35]

$$\langle JM\Omega; \Lambda S\Sigma|g_v\mu_B\vec{v}\cdot\vec{\mathcal{B}}|J'M'\Omega'; \Lambda'S'\Sigma'\rangle = (-1)^{M'-\Omega}g_v\mu_B\mathcal{B}_{M'-M}$$

$$\times \sqrt{(2J+1)(2J'+1)}\begin{pmatrix} J & 1 & J' \\ -\Omega & 0 & \Omega' \end{pmatrix}\begin{pmatrix} J & 1 & J' \\ -M & (M-M') & M' \end{pmatrix}\langle \Lambda S\Sigma|T_0^1(\vec{v})|\Lambda'S'\Sigma'\rangle,$$

where the last factor $\langle \Lambda S\Sigma|T_0^1(\vec{v})|\Lambda'S'\Sigma'\rangle$ is complicated to calculate in detail since many perturbations are involved at the precision that we care about (see Nick Hutzler's thesis for details in the H state). However, the result in the end is that it can be folded into an effective dipole moment $\mu_\parallel$. Note that $\mu_\parallel$ is defined to "carry" a factor of Ω, so

$$\langle JM\Omega; H|g_L\mu_B\vec{L}\cdot\vec{\mathcal{B}} + g_S\mu_B\vec{S}\cdot\vec{\mathcal{B}}_{Zeeman}|J'M'\Omega'; H\rangle = (-1)^{M'-\Omega}\Omega\mu_\parallel\mathcal{B}_{M'-M}$$

$$\times \sqrt{(2J+1)(2J'+1)}\begin{pmatrix} J & 1 & J' \\ -\Omega & 0 & \Omega' \end{pmatrix}\begin{pmatrix} J & 1 & J' \\ -M & (M-M') & M' \end{pmatrix}.$$

We will restrict attention to $\Omega' = \Omega$, $M' = \pm 1$, $M = 0$, $J = J' = 1$:

$$\langle 1, 0, \Omega; H|g_L\mu_B\vec{L}\cdot\vec{\mathcal{B}} + g_S\mu_B\vec{S}\cdot\vec{\mathcal{B}}_{Zeeman}|1, M', \Omega; H\rangle = -\frac{M}{2}\times \mu_\parallel\mathcal{B}_M. \tag{D.9}$$

This is similar to the case for z-aligned magnetic fields [146, Eq. 2.39]

$$\langle 1, M, \Omega; H|g_L\mu_B\vec{L}\cdot\vec{\mathcal{B}} + g_S\mu_B\vec{S}\cdot\vec{\mathcal{B}}_{Zeeman}|1, M, \Omega; H\rangle = +\frac{M}{2}\times \mu_\parallel\mathcal{B}_0. \tag{D.10}$$

Mixing within the $J = 1$ manifold

Now that we have all necessary matrix elements in hand, we will consider the effect of perturbations that mix states $M = +1 \leftrightarrow M = -1$ within the $J = 1$ manifold, via the $M = 0$ intermediate states in $J = 1$. I will assume fields that are linear in time, $D_\parallel\mathcal{E}_x = \nu_{\mathcal{E}x}t$, etc. This has the advantage of giving us terms arising from an average

offset and gradient, so that the resulting effect should be reasonably characteristic of an arbitrary perturbation. Note that we do not necessarily expect the resulting perturbations to be rotationally symmetric about the z-axis. I also assume changes in the $\hat{z}$-aligned fields that are linear in time, $D_\parallel \mathcal{E}_z = \nu_{\mathcal{E}z} t$ and analogously for $\mathcal{B}$ so that the diagonal energy shift is given by $-\frac{M}{2}(\nu_{\mathcal{B}z} + \Omega \nu_{\mathcal{E}z}) t \equiv \dot{\omega}_z t$. This will modify the effective detunings such that $\omega_{\tilde{\mathcal{P}}M} t \to \omega_{\tilde{\mathcal{P}}M} t - \dot{\omega}_z \int_0^t \tau d\tau = \omega_{\tilde{\mathcal{P}}M} t - \frac{1}{2}\dot{\omega}_z t^2$. Here, a state in $M = 0$ is labeled by its parity $\tilde{\mathcal{P}}$ and a state with $M = \pm 1$ is simply labeled by its M quantum number, so $\omega_{\tilde{\mathcal{P}}M} = E_M - E_{\tilde{\mathcal{P}}} \approx -\tilde{\mathcal{N}} D_\parallel |\mathcal{E}|/2$. The electric field is in absolute value brackets here to emphasize that it is the primary applied field (not a perturbation) and I've neglected higher-order corrections to the detuning, $\omega_{\tilde{\mathcal{P}}M}$. The factor of $\tilde{\mathcal{N}}$ comes from $M\Omega$ in the expression for the Stark shift, together with the fact that the applied electric field is odd under $\tilde{\mathcal{E}}$ (and using $\Omega = M\tilde{\mathcal{N}}\tilde{\mathcal{E}}$). Further, let $\nu_{\mathcal{E}x} + \nu_{\mathcal{E}y} + \nu_{\mathcal{B}x} + \nu_{\mathcal{B}y} \equiv \nu_\perp$.

Our initial and final states are labelled by $|\tilde{\mathcal{N}}, M\rangle$ and our intermediate states will be $M = 0$, labelled by $\tilde{\mathcal{P}}$. We therefore compute matrix elements

$$
\begin{aligned}
V_{\tilde{\mathcal{P}}M} &= \tfrac{1}{\sqrt{2}}(\langle J = 1, M = 0, \Omega = 1| + \tilde{\mathcal{P}}\langle J = 1, M = 0, \Omega = -1|)V \\
&\quad \times |J = 1, M, \Omega = M\tilde{\mathcal{N}}\tilde{\mathcal{E}}\rangle \\
&= -\tfrac{1}{\sqrt{2}}(\langle J = 1, M = 0, \Omega = 1| + \tilde{\mathcal{P}}\langle J = 1, M = 0, \Omega = -1|)\nu_\perp \qquad \text{(D.11)} \\
&\quad \times |J = 1, M, \Omega = M\tilde{\mathcal{N}}\tilde{\mathcal{E}}\rangle t \\
&\equiv \alpha_{\tilde{\mathcal{P}}M} t.
\end{aligned}
$$

We can work through the exact form of this matrix element shortly, but we can already find a closed form of $c_{MM}^{(2)}$ and $c_{M(-M)}^{(2)}$ in terms of $\alpha_{\tilde{\mathcal{P}}M}$:

$$
\begin{aligned}
c_{MM}^{(2)}(t) &= -\sum_{\tilde{\mathcal{P}}} \int_0^t dt'' \alpha_{\tilde{\mathcal{P}}M}^* t'' e^{-i(\omega_{\tilde{\mathcal{P}}M} t'' - \dot{\omega}_z t''^2/2)} \int_0^{t''} dt' \alpha_{\tilde{\mathcal{P}}M} t' e^{i(\omega_{\tilde{\mathcal{P}}M} t' - \dot{\omega}_z t'^2/2)} \\
&= -\sum_{\tilde{\mathcal{P}}} |\alpha_{\tilde{\mathcal{P}}M}|^2 \int_0^t dt'' t'' e^{+i(\tilde{\mathcal{N}} D_\parallel |\mathcal{E}| t'' + \dot{\omega}_z t''^2)/2} \int_0^{t''} dt' t' e^{-i(\tilde{\mathcal{N}} D_\parallel |\mathcal{E}| t' + \dot{\omega}_z t'^2)/2}.
\end{aligned}
$$
$$\text{(D.12)}$$

By inspection, a given Ω state can only couple to one component of the $\tilde{\mathcal{P}}$ state, and a factor of $\tilde{\mathcal{P}}$ explicitly appears in the matrix element $\alpha_{\tilde{\mathcal{P}}M}$ for $\Omega = -1$ but not for $\Omega = +1$. However, when we take the absolute magnitude of the matrix element, the (possible) factor

of $\tilde{\mathcal{P}}$ is squared and plays no role. Therefore, $|\alpha_{\tilde{\mathcal{P}}M}| \to |\alpha_M|$ is actually independent of $\tilde{\mathcal{P}}$, and the sum simply yields a factor of 2.

Computing the integral to leading order in $\frac{\dot{\omega}_z t}{D_\parallel |\mathcal{E}|}$ and dropping higher-order terms in $\frac{t}{D_\parallel |\mathcal{E}|}$, we find

$$c_{MM}^{(2)} = -i \frac{\tilde{\mathcal{N}} |\alpha_M t|^2}{D_\parallel |\mathcal{E}|} \left(\frac{4}{3} - 2 \frac{\tilde{\mathcal{N}} \dot{\omega}_z t}{D_\parallel |\mathcal{E}|} \right) t. \tag{D.13}$$

I have collected terms so this amplitude takes the form of a phase accumulating linearly in time.

We use standard conversions between vectors in spherical and cartesian bases,

$$\begin{aligned} v_M &= -\frac{1}{\sqrt{2}}(v_x + iM v_y) \\ v_0 &= v_z. \end{aligned} \tag{D.14}$$

Then, using $\Omega = M\tilde{\mathcal{N}}\tilde{\mathcal{E}}$, we find

$$\begin{aligned} |\alpha_M t|^2 &= \frac{1}{16} |M\tilde{\mathcal{N}}\tilde{\mathcal{E}} D_\parallel (\mathcal{E}_x + iM\mathcal{E}_y) + \mu_\parallel (\mathcal{B}_x + iM\mathcal{B}_y)|^2 \\ &= \frac{1}{16} (D_\parallel^2 \mathcal{E}_\perp^2 + \mu_\parallel^2 \mathcal{B}_\perp^2 + 4M\tilde{\mathcal{N}}\tilde{\mathcal{E}} D_\parallel \mu_\parallel \vec{\mathcal{E}}_\perp \cdot \vec{\mathcal{B}}_\perp). \end{aligned} \tag{D.15}$$

Therefore,

$$c_{MM}^{(2)} = -i \frac{\tilde{\mathcal{N}} D_\parallel^2 \mathcal{E}_\perp^2 + \tilde{\mathcal{N}} \mu_\parallel^2 \mathcal{B}_\perp^2 + 4M\tilde{\mathcal{E}} D_\parallel \mu_\parallel \vec{\mathcal{E}}_\perp \cdot \vec{\mathcal{B}}_\perp}{D_\parallel |\mathcal{E}|} \left(\frac{1}{12} + \frac{1}{8} \frac{M\tilde{\mathcal{N}} \mu_\parallel \delta\mathcal{B}_z + \tilde{\mathcal{E}} D_\parallel \delta\mathcal{E}_z}{D_\parallel |\mathcal{E}|} \right) t, \tag{D.16}$$

where $\delta\mathcal{B}_z$ is the accumulated change in the z-aligned magnetic field relative to $t = 0$, and similarly for $\delta\mathcal{E}_z$. The only higher-order correction that can be at the 10^{-7} level is $\sim \mathcal{E}_\perp^2 \delta\mathcal{E}_z$, which has no M-dependence and thus can't contribute to a systematic. Indeed, there is no term at the sensitivity of interest that can be made proportional to $M\tilde{\mathcal{N}}\tilde{\mathcal{E}}$, so $c_{MM}^{(2)}$ does not give a systematic in this model.

Next, consider

$$\begin{aligned} c_{M(-M)}^{(2)} &= -\sum_{\tilde{\mathcal{P}}} \int_0^t dt'' \alpha_{\tilde{\mathcal{P}}M}^* t'' e^{+i(\tilde{\mathcal{N}} D_\parallel |\mathcal{E}| t'' + \dot{\omega}_{zM} t''^2)/2} \\ &\quad \times \int_0^{t''} dt' \alpha_{\tilde{P}(-M)} t' e^{-i(\tilde{\mathcal{N}} D_\parallel |\mathcal{E}| t' + \dot{\omega}_{z(-M)} t'^2)/2}, \end{aligned} \tag{D.17}$$

which is superficially similar to $c^{(2)}_{M(-M)}$. It is convenient to now explicitly label $\dot{\omega}_{zM}$ with the relevant M state instead of inserting the M-dependence back in at the end. Here, the time-dependent diagonal elements are actually not suppressed by a factor of $D_\parallel|\mathcal{E}|$. This is because the z-aligned fields effectively modify the dominant contribution to $c^{(2)}_{M(-M)}$, which is analogous to the term proportional to $[\exp(i\omega_{ni}t) - 1]/(\omega_{ni}t)$ in the case of a static perturbation, by changing the small detuning ω_{ni} (see Sec. 4.3). Doing the integral with $\dot{\omega}_{z(\pm M)}t$ kept to first order,

$$c^{(2)}_{M(-M)} = -\sum_{\tilde{\mathcal{P}}} \frac{\tilde{\mathcal{N}}(\alpha^*_{\tilde{\mathcal{P}}M}t)(\alpha_{\tilde{P}(-M)}t)}{D_\parallel|\mathcal{E}|} \left(\frac{2i}{3} + \frac{1}{5}(\dot{\omega}_{z(-M)}t^2 - \dot{\omega}_{zM}t^2) \right) t. \qquad \text{(D.18)}$$

At second order in $\dot{\omega}_{z(\pm M)}t^2$, terms that are M-even such as $\dot{\omega}_{zM}t^2 \times \dot{\omega}_{z(-M)}t^2$ will contribute. However, we can actually have time-dependent z-aligned electric fields comparable to ~ 10 kHz, so this small-parameter expansion is not appropriate. To my knowledge, there's no clean form for $c^{(2)}_{M(-M)}$ in the general case, but I expect from the time-independent situation that correction terms should be at most order-unity. Further, I strongly suspect (but haven't proved) that terms with even powers of $\dot{\omega}_z$ will be imaginary while terms with odd powers will be real, and any real terms will be proportional to M because they physically arise from the splitting between initial and final levels. For a static perturbation, we see exactly this behavior by expanding $[\exp(i\omega_{ni}t) - 1]/(\omega_{ni}t)$. From here on, I drop the terms involving $\dot{\omega}_z$ because the M-odd part, due to gradients of $\mathcal{B}_z$, is rather small.

Finally, we will compute $(\alpha^*_{\tilde{\mathcal{P}}M}t)(\alpha_{\tilde{P}(-M)}t)$ directly. Here, there are no "cross-terms" that scale like $\vec{\mathcal{E}}_\perp \cdot \vec{\mathcal{B}}_\perp$ because simultaneous M-reversal and complex conjugation in the matrix elements $\alpha_{\tilde{\mathcal{P}}M}$ looks like mapping $\mathcal{E} \to -\mathcal{E}$ but $\mathcal{B} \to \mathcal{B}$. Therefore, the transition amplitudes for "perturbed by $\mathcal{B}$, then $\mathcal{E}$" and "perturbed by $\mathcal{E}$, then $\mathcal{B}$" exactly cancel. We also see that perturbations due to $\mathcal{E}_x^2$ and $\mathcal{E}_y^2$ appear with opposite sign, essentially because of the relative factor of i appearing between v_x and v_y in the spherical basis expansion. Putting everything together, we obtain

$$c_{M(-M)}^{(2)} = -\frac{i\tilde{\mathcal{N}}}{12}\frac{\mu_\parallel^2(\mathcal{B}_x^2 - \mathcal{B}_y^2) - D_\parallel^2(\mathcal{E}_x^2 - \mathcal{E}_y^2)}{D_\parallel|\mathcal{E}|}t - \frac{M\tilde{\mathcal{N}}}{6}\frac{\mu_\parallel^2\mathcal{B}_x\mathcal{B}_y - D_\parallel^2\mathcal{E}_x\mathcal{E}_y}{D_\parallel|\mathcal{E}|}t. \tag{D.19}$$

Although this has a rather odd form, recall that the perturbation we assumed is not rotationally symmetric about the z-axis, so we can't expect rotational symmetry in the expression for the perturbed state amplitude.

With the result finally in hand, we can note that the real part is M-odd and the imaginary part is M-even, so this perturbation only contributes to the measurement through the term $\phi_{\text{meas.}} \sim (c_+ - c_-)\mu$ and its signature in the measured phase is proportional to the ellipticity of the preparation lasers. (Once again, see Sec. 4.3 for the relevant notation.) Based on these results, we expect that the second-order couplings between opposite-M states, due to time-dependent electric and magnetic fields, will not contribute to an EDM systematic. Therefore, in the next section, where we consider perturbations coupling the EDM spin precession state to the $J = 2$ manifold, I only compute the direct second-order energy shift $c_{MM}^{(2)}$.

Perturbations via $J = 2$

The (possibly M-dependent) energy shifts mediated by the $J = 2$ manifold take the same form that we found previously,

$$c_{MM}^{(2)} = -\frac{i}{12}\sum_{M_2}\frac{|\alpha_{M_2 M}t|^2}{B_{\text{rot}}}t, \tag{D.20}$$

where I'm ignoring the higher-order contributions involving time-dependent z-aligned fields at the precision of interest, and I have replaced the detuning $-\tilde{\mathcal{N}}D_\parallel|\mathcal{E}|/2$ with $-4B_{\text{rot}}$, and we must now sum over $M_2 = 0, 2M$ in the $J = 2$ manifold instead of summing over $\tilde{\mathcal{P}}$ states. Without proceeding further, we can simply note that there is nowhere for a factor of $\tilde{\mathcal{N}}$ to appear, so perturbations that drive $J = 1 \to J = 2 \to J = 1$ cannot generate a systematic error.

However, perturbations due to Stark mixing between $J = 1$ and $J = 2$, followed by tran-

sitions within the $J = 2$ manifold and finally followed by projecting back into $J = 1$, do have the potential to generate $\tilde{\mathcal{N}}$-odd terms. The Stark mixing into and out of $J = 2$ generates a suppression factor of $(\frac{1}{2}\sqrt{\frac{3}{5}} \times D_{\parallel}\mathcal{E}_0)^2/(4B_{\text{rot}})^2 = \frac{3}{20}\frac{D_{\parallel}^2|\mathcal{E}|^2}{(4B_{\text{rot}})^2} \sim \frac{1}{10^2}(\frac{100\,\text{MHz}}{10\,\text{GHz}})^2 \sim 10^{-6}$. We see that the numerical prefactor is sufficiently small that the order-of-magnitude estimate for $c^{(2)} \sim 10^{-7}$ was far too large, and the AC Stark/Zeeman perturbations mediated by $J = 2$ can be neglected after all.

In conclusion, time-dependent $\mathcal{E}$ and $\mathcal{B}$ fields are unable to generate systematic errors at the ACME II sensitivity via any known second-order perturbations to the $|H, J = 1, M = \pm 1\rangle$ states. These calculations were inspired by the observed $\partial\mathcal{B}/\partial z$ systematic error in ACME II, before it was fully understood, and the approach used here may be a useful guide to other systematic error models considered in future measurements.

Appendix E

Nonimaging molecular focusing

We've considered the possibility that a "lens" is not the optimal way to focus the molecular source. The domain of non-imaging optics deals with the design and construction of optical elements that most efficiently map rays from some input region of phase-space to a more convenient "target" region of phase-space. For example, it may be useful to compress the largest number of input rays into the smallest spatial area on some photodetector. In such an application, it is irrelevant whether the rays form an image of their source when they are incident on the detector.

In fact, this is exactly the situation we are in with respect to the molecular "lens": it does not matter whether we image the source onto the area where the molecules are detected. For this reason, I considered possible molecular focusing designs that would mimic optical elements used in the field of non-imaging optics. Ultimately, these are not useful because the number of molecules we can direct into the interaction region is limited by the focusing potential that is technically feasible rather than by inefficiently filling the detector area with molecule trajectories (or "rays").

Analog to the "compound parabolic concentrator," or Winston cone

The conservation of etendue, considered in Sec. 6.1.2, limits the range of molecular divergence angles that can be redirected into the detection area in the interaction region. As a simple example, consider a round source with maximum divergence angle $\alpha \ll 1$ and area S. An ideal optical element that reduces the divergence to $\alpha' < \alpha$ must, necessarily,
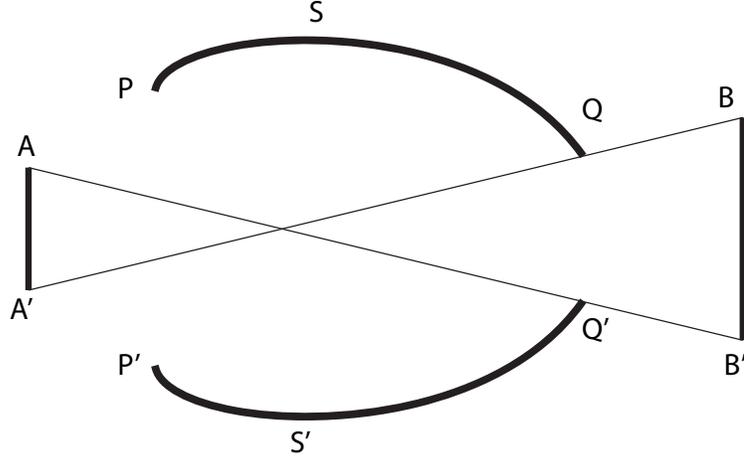
Figure E.1: Crude schematic of the geometry for a "generalized Winston cone" that maps rays between two distant regions

increase the area by a factor α/α'. The book Nonimaging Optics by Roland Winston is a very useful reference for problems of this type [204].

Following the strategies in [204], it is easiest to construct a nonimaging optical element with a surface reflector, which could be approximated by a very high-order electric or magnetic field multipole configuration. We can regard our "focusing" problem in the following simple terms: we have an optical source (for now, in 2D) along the vertical line segment AA' and a desired target along the vertical line segment BB'. The boundaries of our optical element are the curved segments PQ and $P'Q'$. There is a fixed horizontal distance L_1 between the source and optic, and a distance L_2 between the optic and target, based on practical constraints. The maximal divergence angle we'll consider from the source is θ_c.

Let the source (along AA') have radius r_1 and the target (along BB') have radius r_2. We can design a reasonably efficient concentrator for rays originating in AA' to be directed to the target BB'. See Fig. E.1 for a crude representation of the geometry. The concentrator has two segments. The first is a parabolic curve PS, tilted at an angle θ_c with respect to the horizontal axis so that a ray emitted anywhere along the line AA' at the critical (maximum) design angle θ_c will intersect the parabola and be reflected to the focal point, placed at B'. Likewise, the other side of the concentrator has a parabolic segment $P'S'$ with focus at B. Any shallower ray, emitted with $\theta < \theta_c$, will be reflected to a point within the line segment BB' rather than the edges.

The second segment, SQ, is an ellipse with foci at A' and B'. (The other side, of course, has a segment $S'Q'$ with foci at A and B.) This ensures that any ray from the line segment AA' that intersects the curve $S'Q'$ will be reflected to some point on the line segment BB'. Finally, the endpoint Q' is set to lie along the line AB'. Therefore, any ray emitted from AA' at an angle up to θ_c must be reflected once along the path $P'S'Q'$ or PSQ and pass through BB'. However, if θ_c is made too large, then some rays reflected from PSQ will intersect another part of PSQ before reaching BB', and be rejected back past the source (perhaps after many bounces). Indeed, rays like this must exist in order for etendue conservation to hold in the limit $\theta_c \to \frac{\pi}{2}$.

The exact geometry can be expressed in terms of Cartesian coordinates through the following system of equations. I regard the "inputs" to be r_1, r_2, L_1, L_2, and θ_c. I assume that primed coordinates are for $y > 0$ and unprimed coordinates are for $y < 0$, and $x_A = x_{A'} = 0$ defines the origin. A 3D concentrator is obtained by rotating the 2D geometry about the $y = 0$ line. The defining equations will be elaborated upon below.

1. $p = \frac{1}{2} \left(r_2 \sin \theta_c + y_{P'} \sin \theta_c + x_B \cos \theta_c - L_1 \cos \theta_c \right.$

   $\left. + \sqrt{x_B^2 + r_2^2 + y_{P'}^2 - 2x_B L_1 + 2r_2 y_{P'}} \right)$

2. $k = x_B \cos \theta_c + r_2 \sin \theta_c - p$

3. $h = -x_B \sin \theta_c + r_2 \cos \theta_c$

4. $y_{\text{parab}}(x) = \sqrt{2} \sec^2 \theta_c \sqrt{p(p - k + 2x \cos \theta_c - (p + k)\cos(2\theta_c) + h \sin(2\theta_c)}$

   $- x \tan \theta_c - \sec \theta_c (h + 2p \tan \theta_c)$

5. $y_{P'} = r_1 + L_1 \tan \theta_c$

6. $x_{S'} = \frac{1}{4} \left[2 \cot \theta_c (r_1 + p \csc \theta_c) - 2(h \csc \theta_c + p \sec \theta_c) \right.$

   $\left. + \sqrt{2} \cot^2 \theta_c \sqrt{p \sec^6 \theta_c (p - k + (p + k)\cos(4\theta_c) + 4r_1 \cos^2 \theta_c \sin \theta_c - h \sin(4\theta_c)}} \right]$

7. $\sqrt{x^2 + (y_{\text{ellipse}}(x) - r_1)^2} + \sqrt{(x - x_B)^2 + (y_{\text{ellipse}} - r_2)^2}$

   $= \sqrt{x_{S'}^2 - (y_{\text{parab}}(x_{S'}) - r_1)^2} + \sqrt{(x_{S'} - x_B)^2 + (y_{\text{parab}}(x_{S'}) - r_2)^2}$

8. $y_{\text{ellipse}}(x_B - L_2) = -r_1 + \frac{r_1 + r_2}{x_B}(x_B - L_2)$

316

Lines 1-3 give shorthand expressions for the parameters of the parabola. The function $y_{\mathrm{parab}}(x)$, given in line 4, gives the functional form of the parabolic segment defining the reflector from $x = x_{P'}$ to $x = x_{S'}$. It is rather complicated only because the parabola's axis is misaligned from the horizontal axis. Line 5 defines the opening radius of the parabolic section (used implicitly in previous lines), defined so that a ray at angle θ_c from A' intersects P'. Line 6 solves for the endpoint of the parabola along the x-axis, defined so that a ray at angle θ_c from A intersects S'. Line 7 gives an implicit equation for $y_{\mathrm{ellipse}}(x)$, defining the reflector along curve $S'Q'$. (This equation actually has an analytic solution, but it's extremely messy.) Line 8 implicitly defines x_B, which in turn gives $x_Q = x_B - L_2$. These results are obtained by direct geometric analysis, and can be numerically solved together in simulations to construct the reflector geometry.

I do not claim that this reflector is optimal in the sense of allowing the source rays to be directed into a minimal area, but I believe it should more efficiently concentrate rays than an imaging element. However, since maximal spatial compression between the source and target is of limited benefit for molecular focusing in ACME (given the more-limiting finite potential depth), I have not made a thorough study of the performance of this non-imaging element for realistic experimental geometries.

Design for multiple reflections

One of the principles of the "Winston cone" and analogous devices is that each ray from the input bundle experiences at most one reflection before arriving at the target bundle. This is useful for optimal concentration, but it assumes a perfectly reflective surface is available. In our case, the interaction energy is limited so that the projection of the molecule velocity along the surface normal vector, at the point of reflection, must be less than a critical velocity v_c such that $\frac{1}{2}mv_c^2 = U_{\mathrm{max}}$. For a fixed forward velocity, let the critical angle be $\theta_c = v_c/v_x$. If a trajectory has angle θ with respect to the horizontal, then the reflector surface must be angled so that $\theta - \theta_{\mathrm{surf}} \leq \theta_c$ or the molecule will "break through" the potential barrier. Let the surface be given by a function $R(x)$ and let there be a point source at distance l from the opening of the reflector. Suppose that every trajectory from the point source, parametrized by θ, hits the reflector with the critical condition $\theta - \theta_{\mathrm{surf}} = \theta_c$. Therefore,

$\frac{dR}{dx} = \Theta(x) - \theta_c$, where $\Theta(x)$ is a function giving the angle of the trajectory that intersects the reflector at position x. By definition, we must have $R(x) = (l + x)\Theta(x)$. Substituting into the differential equation for $R(x)$, we find

$$\frac{dR}{dx} = \frac{R(x)}{l + x} - \theta_c. \tag{E.1}$$

This can be solved analytically. If the maximum accepted angle is θ_0 (intersecting the reflector at its opening aperture), then we find

$$R(x) = (l + x)\left(\theta_0 - \theta_c \ln\left(1 + \frac{x}{l}\right)\right). \tag{E.2}$$

Unlike the Winston cone, this construction does not guarantee any "good" behavior of the output ray bundle, nor does it guarantee a single reflection per ray. However, this could conceivably solve a practical problem with the reflector design: for example, if a source is at minimum 25 cm from a harmonic lens and we want to admit angles up to 0.05 rad, then the lens aperture must be greater than 1 cm in radius. In practice, we need a factor of ~ 2 margin because a molecule cannot be turned around if it enters the lens already at the wall; therefore, let $R \sim 2$ cm. To turn around such a trajectory, it must travel a transverse distance comparable to the lens radius, corresponding to a longitudinal distance of $\sim 2\,\text{cm}/0.05 \sim 40$ cm. (These estimates are actually quite close to optimal lens length values for the ACME geometry.) In this distance, at most $\sim U_{\max}$ energy can be removed from the transverse velocity. On the other hand, a molecule that repeatedly "skips" along a slanted wall could have energy $U_{\max}$ removed many times in a short distance because the molecule need not cross to the opposite end of the lens.

Unfortunately, I've found from direct simulation that this multiple-reflection design is only comparably as effective as a harmonic potential. We can understand this heuristically: in order to have two useful collisions between a molecule and the wall, we'd want the initial admitted divergence to be at least $\theta_0 = 2\theta_c$. We'll want the reflector to be at least as long as required to "flatten out," where $R'(x) = 0$, which occurs for $L \approx 1.7\,l$ for the given value of θ_0. Here, the reflector will turn around trajectories with $\theta = \theta_c$. Therefore, with $l = 25$ cm, we'll have a 43 cm long reflector that can (energetically speaking) "double-bounce" a

single divergence class, representing negligible improvement over an ordinary curved lens. Further, it is better to have a slightly converging reflector at the end, increasing the length further, and the length needed for the reflector to flatten out grows *exponentially* in the ratio θ_0/θ_c.

Therefore, a useful multiple-reflection design will be impractically long. We cannot make a shorter reflector with a similar design principle because the geometry considered here already enforces a maximally hard turn inward. Decreasing the radius of curvature anywhere, to reduce the reflector length, will cause most molecules to be rejected (i.e., either reflected back toward the source or "break through" the reflecting potential) on their first "bounce."

Velocity point source

As something of an aside, it is also interesting to understand how we could focus a "velocity point source." To this end, given an extended source $-y_{max} < y < +y_{max}$ with a single divergence angle θ, we wish to find a class of potentials $U(x, y)$ ensuring that all molecules turn around at the same position, $v_y(x = L, y_0) = 0$. Here, y_0 indexes the initial position of the trajectory in the source. We know that a position-independent force (e.g., gravity near earth's surface) achieves this, since then the momentum transfer along y is identical for all molecules over their trajectories. As we saw in the case of a spatial point source, this argument holds even if the force varies with the x-position of the molecules. Thus a potential $U(x, y) = U_{\max}y/R(x)$ collimates a "velocity point source," where $R(x)$ is any function of x with units of length.

In 3 dimensions, a harmonic potential is separable, $r^2 = y^2 + z^2$, so a point source can still be collimated perfectly. However, a linear potential is not separable, $r = \sqrt{y^2 + z^2} \neq y + z$, and the reasoning of this section cannot be generalized exactly. (In particular, it is valid only for molecules with vanishing azimuthal velocity, where the centrifugal barrier of the potential in cylindrical coordinates vanishes.) Nevertheless, I expect that a linear potential could be more favorable if the source distribution were extremely extended in space (relative to other length scales) and it were paramount to redirect a narrow range of velocity classes. In ACME, a harmonic potential appears to be more advantageous than a linear potential,

consistent with the fact that our source is reasonably well-concentrated in space (3 mm radius) but not in velocity ($\sim 45°$ FWHM divergence).

Appendix F

Photodetector excess noise factors

We discussed in Sec. 6.3 that the signal-to-noise ratio of a phase measurement is $\sigma_\phi \propto$ $(\epsilon_{\text{det.}}/F)^{-1/2}$, where $\epsilon_{\text{det.}}$ is the detector efficiency and F is the excess noise factor. We will see in this appendix how the excess noise factor comes about and why it varies among different detection technologies.

The uncertainty in the phase measurement can ultimately be traced to statistical noise in the detected number of photoelectrons, which are detected at some rate r. We can understand the *noise* associated with the detection in terms of the power spectral noise density of the photon detection process. This is elaborated upon below.

Power spectral noise density

The power spectral density of a signal $x(t)$ is

$$P \equiv \lim_{T \to \infty} \frac{1}{T} \int_0^T dt\, x(t)^2, \tag{F.1}$$

in analogy to the fact that the physical power associated with a signal (e.g., a voltage $V(t)$) is typically proportional to the average of the square of the signal. It is useful to consider the power associated with the signal component only at angular frequency ω. Let the finite-window Fourier transform be

$$\hat{x}_T(\omega) \equiv \int_0^T dt\, x(t)e^{-i\omega t}, \tag{F.2}$$

321

where I'm assuming the signal obeys the constraint that $x(t < 0) = 0$. Then the power contributed within a band ω to $\omega + \delta\omega$ is

$$P(\omega, \omega + \delta\omega) = \lim_{T\to\infty} \frac{1}{T}|\hat{x}_T(\omega)|^2 \delta\omega. \tag{F.3}$$

We define the power spectral density at frequency ω to be $P(\omega) = \lim_{T\to\infty} \frac{1}{T}|\hat{x}_T(\omega)|^2$. Consider the idealized case that a signal $x(t)$ measures a photodetector current and has the form of a "spike train," where discrete photon arrivals are observed as delta functions of current. In particular, for photon arrivals at times $t_1, t_2, \cdots, t_n$, the signal will be $x(t) = \sum_{i=1}^{n} \delta(t_i)$. If the photon arrival times are generated from a Poisson process of photon arrivals with characteristic arrival rate r, then for $\omega > 0$, one can show that[264, Sec. 5.2] $P(\omega) = r$. Thus the power spectral density is independent of frequency, and the noise arising from the Poisson process is "white."

The case of $\omega = 0$ requires additional care (which I will not derive here in detail), because there is a contribution to $P(\omega = 0)$ from the time-averaged signal, r, that leads to a delta-function in the power spectrum, $P(\omega) = 2\pi r^2 \times \delta(\omega) + r$. We therefore define the power spectral *noise* density to be the power spectrum of the actual signal $x(t)$, which is generated by a stochastic process, minus the expected (non-stochastic) signal $\langle x(t) \rangle = r$. Thus the delta-function contribution to the power spectrum is removed and the power-spectral noise density turns out to be only $P_{\text{noise}}(0) = r$. Therefore, for a Poisson process, the power spectral noise density is $P_{\text{noise}}(\omega) = r$ for all ω.

Introduction to the excess noise factor

The power spectral noise density gives an expression for the frequency-dependent noise associated with the photodetector signal $x(t)$, considered above as a "spike train," or a sequence of "clicks" denoting photon detection events. Consider now an ideal photodetector with gain G detecting photons whose arrival is governed by a Poisson process so that the measured current is a spike train, $I(t) = G \times \sum_{i=1}^{n} q_e \delta(t_i)$, where q_e is the electron charge. Then the power spectral noise density of the current measurement is $\delta I^2(|\omega|) = 2q_e I_0 G^2$, where I_0 is the average (DC) current generated directly by photoelectrons, before

322

amplification by the internal gain G. Here, the factor of 2 appears because we conventionally sum the power spectral noise density over positive and negative Fourier components.

When the gain is not a fixed constant but rather is governed by a statistical distribution with mean $\langle G \rangle$ and variance ΔG^2, then the power spectral noise density of the current becomes $\delta I^2 = 2q_e I_0 \langle G \rangle^2 \times F$, where $F = 1 + \text{Var}(G)/\langle G \rangle^2$ is the excess noise factor (see [265] for details). This factor is derived below for several photodetector gain mechanisms. Note that here, I_0 is the actual photoelectron current, and already includes the effect of the photodetection efficiency $\epsilon \leq 1$. Formally, this is valid because the detected and undetected photons can be thought of as being generated by completely independent Poisson processes. This is known as thinning or splitting the original Poisson process [266].

To compute the excess noise factor F, we must have a model of the stochastic process governing G for any particular detector. In practice, the excess noise factor can be measured by observing the distribution of currents generated by single-photon detection events. For the Hamamatsu R7600U-300 detectors used in ACME II, we have measured excess noise factors of $1.2 - 1.3$ as described in Sec. 3.2.1.

Photomultiplier tubes

In a photomultiplier tube, a photoelectron is generated in an anode, accelerates across a potential, and induces secondary emission of electrons upon colliding with a cathode. Typically, there are about 5 electrons generated in each secondary emission and up to 9 dynode (anode-cathode) stages, for a gain of order $5^9 \sim 10^6$. Generally, the average number of output electrons in an m-stage photomultiplier tube is $\langle G \rangle = \Pi_{k=1}^{m} \langle g_k \rangle$, where $\langle g_k \rangle$ is the average gain of the k-th stage. For purposes of illustration, note that the fluctuations in G are typically dominated by the fluctuations in the gain of the first stage since the variance in subsequent stages is suppressed by the law of large numbers (many electrons independently undergo a gain at the same stage). Modelling all stages subsequent to the first as having perfectly-defined gain, we compute $\langle \Delta G^2 \rangle = \langle \Delta g_1^2 \rangle (G/\langle g_1 \rangle)^2$ and $F = 1 + \langle \Delta G^2 \rangle / G^2 = 1 + \langle \Delta g_1^2 \rangle / \langle g_1 \rangle^2$.

The secondary emission is frequently modelled as a Poisson process, so that $\langle \Delta g_1^2 \rangle = \langle g_1 \rangle$ and $F = 1 + 1/\langle g_1 \rangle$. Taking a reasonable value of $g_1 \approx 5$, we infer $F = 1.2$. For a

photomultiplier with fixed gain at all stages, $\langle g_i \rangle = \langle g \rangle$, the result is only slightly larger, as expected: $F = 1 + (1 - \langle g \rangle^{-m})/(\langle g \rangle - 1) \approx 1 + 1/(\langle g \rangle - 1)$. See [176] for details.

Therefore, the excess noise factor for photomultiplier tubes is typically in the range 1.2-1.4, consistent with what we've observed for the PMTs used in ACME II.

Avalanche photodiodes

An avalanche photodiode (APD) operated in current mode amplifies photoelectrons by accelerating them through a p-n junction operated at reverse bias in a solid-state material such as silicon. Impact ionization generates electron-hole pairs, and the generated electrons continue to accelerate within the material, leading to further amplification. Hole ionization also occurs, typically at a rate about $k \sim 1 - 10\%$ that of electron ionization [176]. I will neglect hole ionization in the simple discussion below. However, it's important to note that when an APD is operated at large gain, the hole ionization process can contribute significantly to the excess noise factor.

Following [267], we will model the ionization process as follows: each photoelectron is exponentially amplified as it traverses the gain material, so $\langle G(x) \rangle = e^{\alpha x}$, where x represents the spatial coordinate of the photoelectron's path through the gain medium. For a gain region of length L, the expected gain is then $\langle G \rangle \equiv \langle G(L) \rangle = e^{\alpha L}$. Within a small region of length dx centered on x, the current generation is subject to shot noise, contributing to the spectral noise density P by an amount $\langle dP(x) \rangle = 2q_e \langle G(L-x) \rangle^2 \langle I(x) \rangle \alpha dx$, where $\langle I(x) \rangle$ is the total current expected at x, and $\langle G(L-x) \rangle$ is the remaining amplification it is expected to undergo. Note that $\langle I(x) \rangle \propto \langle G(x) \rangle$, so we can write $\langle I(x) \rangle = I_0 e^{\alpha x}$. Plugging this in to the expression for $\langle dP \rangle$ and integrating, we find

$$\langle P(L) \rangle = 2q_e \alpha I_0 e^{2\alpha L} \int_0^L dx \, e^{-\alpha x} = 2q_e I_0 \langle G \rangle^2 (1 - 1/\langle G \rangle). \tag{F.4}$$

Note that this is only the noise associated with amplification. Adding this contribution to the shot noise of the photoelectron generation process itself, we see that $F = 2 - 1/\langle G \rangle$. This formula applies to the special case $k = 0$ (i.e., hole ionization is ignored). Similar considerations with $k \neq 0$ lead to additional terms $k(\langle G \rangle - 2 + 1/\langle G \rangle)$. In any case,

ideal operation of an avalanche photodiode typically gives $F \approx 2$. Thus the benefit of the remarkably high detection efficiency ($\epsilon \approx 0.8$) is significantly mitigated by the large excess noise factor.

Silicon photomultipliers

A silicon photomultiplier (SiPM) is an array of small ($\sim 50\,\mu\text{m}$) avalanche photodiode cells [214]. Each cell can be modelled as a capacitor with capacitance C and stored charge Q. This charge can be readily discharged in the presence of a sufficiently large electric field, which arises from a voltage applied across the cell. The critical voltage required to allow electric discharge is the "breakdown voltage" of the cell, and silicon photomultipliers are operated with an applied voltage that exceeds this breakdown voltage by a margin known as the "overvoltage." Under conditions of a modest overvoltage (~ 5 V), when a photoelectron is generated, it triggers an avalanche, causing all of the charge stored in the cell to flow as current. The result is a current spike of highly predictable amplitude. The gain is Q/q_e, typically of order 10^6. Although there is a dead time τ_{dead} associated with the recharging of each cell, consecutive photons are highly unlikely to hit the same microcell, and the maximum counting rate is of order $N_{\text{cells}}/\tau_{\text{dead}}$, which can be of order 100 GHz. Here, N_{cells} is the number of micro-cells in a device. The current output of all cells is summed.

The gain during normal operation is extremely stable ($F \approx 1$) because very uniform voltages can be applied both across distinct cells and from avalanche to avalanche for a given cell. The total charge $Q = C \times V$ that flows in a given avalanche is therefore fixed to high precision.

However, when operated at high over-voltages, an avalanche in one cell can trigger an avalanche in an adjacent cell with probability ϵ_c, effectively doubling the gain at random intervals. This interaction between adjacent cells is known as cross-talk. Chains of cross-talk events can also occur, but since $\epsilon_c \sim 0.1$ typically, I will neglect these chain effects (which enter first at order ϵ_c^2). Another noise mechanism is afterpulsing, where a charge carrier gets "stuck" during the avalanche and is released soon after the cell recharges. The effect of such an event is, once again, to double the gain with probability ϵ_{ap}. Usually, $\epsilon_{\text{ap}} < \epsilon_c$, but we can include both effects to leading order by considering the small probability $\epsilon = \epsilon_{\text{ap}} + \epsilon_c$

to double the gain of the detector.

Finding the excess noise factor due to ϵ, to leading order, is straightforward. Normalizing the gain to the ordinary value of the gain G_0, we compute

$$\frac{\langle G \rangle}{G_0} = 1 \times (1 - \epsilon) + 2 \times \epsilon = 1 + \epsilon. \tag{F.5}$$

The variance, in units of G_0, is the same as the variance in a Bernoulli process with "success" probability ϵ, namely

$$\frac{\Delta G^2}{G_0^2} = \epsilon(1 - \epsilon) \approx \epsilon, \tag{F.6}$$

where I've dropped the term of order ϵ^2 because higher-order effects have already been neglected in the model. We then find that $F = 1 + \epsilon + \mathcal{O}(\epsilon^2) \approx 1.1$ in a typical case. This is in good agreement with the literature on SiPM excess noise factors.

We might wonder how the excess noise factor is affected if we create an array of silicon photomultipliers, some of which have different breakdown voltages than others. For example, suppose we have a 2×1 array of SensL J-series devices. These are specified to have a range of breakdown voltages spanning ± 0.25 V, which translates directly to a difference in over-voltage if sensors are not individually powered. For a typical device, the recommended operating overvoltage is 5 V and the device gain is very linear in the overvoltage [210]. We can therefore assume that half of the photons incident on this array experience gain $(5.25/5)\,G_0 = 1.05\,G_0$ and the other half experience gain $(4.75/5)\,G_0 = 0.95\,G_0$. The variance in gain of this process is given by $\langle \Delta G^2 \rangle / G_0^2 = (1/2)(0.05^2) + (1/2)(0.05^2) = 0.0025$. Ordinary variation in the breakdown voltage of units should therefore contribute negligibly to the excess noise in an array of SiPMs, and in particular is dominated by the intrinsic excess noise in each individual SiPM device due to cross-talk and after-pulsing.

Appendix G

Example header

Here, I show part of an example header, mentioned in Sec. 3.3. The quantities recorded here are used to identify the experimental configuration, any amplified imperfections used to search for systematic errors, and to search for possible correlations in the data set with auxiliary parameters (e.g., temperatures, magnetic fields in the room, etc.). As mentioned previously, there are 6 categories for each trace: (0) Trace #, Start time, End time; (1) Switch times; (2) DAQ properties; (3) Switch states; (4) Instrument setpoints; (5) Logging measurements. Below, I show the part of a header used only to record data for Run 281, Sequence 9, Block 8, Trace 34. As seen in the section for "switch states," during this block there was an applied non-reversing electric field, $\mathcal{E}^{\mathrm{nr}}$, used to measure a possible systematic error during the EDM data set.

```
header  0281.0009.0008.0034.0
Start Time      2018-03-30T13:00:12.262
End Time        2018-03-30T13:00:13.004
==      switch times    ==
total switch time       200     ms
E-fields        200     ms
B-fields        0       ms
Freqs/Powers    100     ms
Waveplates      0       ms
==      DAQ Properties  ==
Human Operator  EDM Student
dt      62.5000000000E-9        s
Records per Trace       25
DAQ Voltage Range       2.000000        V
Number of Channels      8
Conversion Factor Individual    1.6479743648E-5 V
Conversion Factor Summed        2.0599679560E-6 V
```

```
Acquisition Rate           5.000000E+1      Hz
Polarization Switching Frequency         2.000000E+2       kHz
Polarization Switching Deadtime 8.000000E-1      s
Polarization Switching Extra XY Delay    1.000000E-2      s
Polarization Switching XY Swapped        0.000000E+0
Scope Trigger Offset     4.000000E+0      s
Current Sequence Code ID         2355
Ablation Mirror Position X       6.514    arb
Ablation Mirror Position Y       5.2449   arb
==       switch states    ==
E_mag    140      V/cm
690_STIRAP_Detuning      0.4      MHz
dBz/dz_nr        0.25     mA
STIRAP_1090_half_wave_angle      58.5      degrees
P_NE     0        %
Enr      0.3      V/cm
Bz_rev   0.5      mA
R        +1
L        -1
P        -1
B        +1
theta    -1
E        -1
N        -1
==       instrument setpoints     ==
Bx       0.000000
By       0.000000
Bz       0.500000
dBx/dx   0.000000
dBy/dx   0.000000
dBy/dy   0.000000
dBy/dz   0.000000
dBz/dx   0.000000
dBz/dz   0.250000
Bleads   1.000000
field plate east voltage         314.325000
field plate west voltage         -314.325000
guard ring east voltage 314.325000
guard ring west voltage -314.325000
e-field leads    -1.000000
electric field   140.000000
probe half wave angle    -0.500000
prep half wave angle     11.800000
prep quarter wave angle 16.000000
STIRAP 1090 half wave angle      58.500000
N        -1.000000
Pump AOM Frequency       0.000000
P        -1.000000
Pump RF Voltage 0.500000
Probe X RF Voltage       0.550000
Probe Y RF Voltage       0.540000
TiSapph Detuning Upper (N-)      0.000000
TiSapph Detuning Lower (N+)      0.000000
```

```
690 STIRAP Detuning     0.400000
1090 STIRAP Detuning    0.000000
All West East    0.000000
==      logging measurements    ==
TiSapph Beat Note Frequency     2018-03-30T13:00:12.394 7.790430E+8     Hz
STIRAP lambda/2 2018-03-30T13:00:12.400 5.850000E+1     degrees
Pump lambda/2 kz-1      2018-03-30T13:00:12.400 1.000200E+0     degrees
Probe lambda/2  2018-03-30T13:00:12.400 -4.996000E-1    degrees
Pump lambda/2   2018-03-30T13:00:12.400 1.179850E+1     degrees
Probe lambda/2 kz-1     2018-03-30T13:00:12.400 -2.700000E-3    degrees
STIRAP Linear Stage Position    2018-03-30T13:00:12.659 1.804011E+1     mm
Beam Box Cell Top Temperature   2018-03-30T13:00:12.641 1.701600E+1     K
Beam Box Pressure       2018-03-30T13:00:12.690 3.160000E-7     Torr
Stem Pressure   2018-03-30T13:00:12.634 2.460000E-7     Torr
Neon Buffer Gas Flow Rate       2018-03-30T13:00:12.619 3.990000E+1     sccm
Dump Region Pressure    2018-03-30T13:00:12.666 3.300000E-7     Torr
Beam Box Snorkle Temperature    2018-03-30T13:00:12.670 1.728700E+1     K
TiSapph Beat Note Frequency     2018-03-30T13:00:12.694 7.789679E+8     Hz
East Field Plate Voltage        2018-03-30T13:00:12.746 -3.143700E+2    V
West Field Plate Voltage        2018-03-30T13:00:12.742 3.143062E+2     V
Field Plate Voltage Difference  2018-03-30T13:00:12.744 -6.286762E+2    V
Field Plate Voltage Offset      2018-03-30T13:00:12.744 -3.192308E-2    V
Leads Configuration     2018-03-30T13:00:12.744 1.000000E+0     state
STIRAP lambda/2 2018-03-30T13:00:12.800 5.850000E+1     degrees
Pump lambda/2 kz-1      2018-03-30T13:00:12.800 1.000200E+0     degrees
Probe lambda/2  2018-03-30T13:00:12.800 -4.996000E-1    degrees
Pump lambda/2   2018-03-30T13:00:12.800 1.179850E+1     degrees
Probe lambda/2 kz-1     2018-03-30T13:00:12.800 -2.700000E-3    degrees
X North Coil Current    2018-03-30T13:00:12.608 4.835469E-3     mA
X South Coil Current    2018-03-30T13:00:12.608 1.088947E-3     mA
Y +++ Coil Current      2018-03-30T13:00:12.608 6.197840E-3     mA
Y ++- Coil Current      2018-03-30T13:00:12.608 -2.734241E-4    mA
Y +-+ Coil Current      2018-03-30T13:00:12.608 6.197840E-3     mA
Y +-- Coil Current      2018-03-30T13:00:12.608 1.429540E-3     mA
Y -++ Coil Current      2018-03-30T13:00:12.608 4.077617E-4     mA
Y -+- Coil Current      2018-03-30T13:00:12.608 5.857247E-3     mA
Y --+ Coil Current      2018-03-30T13:00:12.608 -1.295203E-3    mA
Y --- Coil Current      2018-03-30T13:00:12.608 -1.295203E-3    mA
Z North Coil Current    2018-03-30T13:00:12.608 2.560644E+0     mA
Z South Coil Current    2018-03-30T13:00:12.608 2.530331E+0     mA
Z East Coil Current     2018-03-30T13:00:12.608 2.541494E-1     mA
Z West Coil Current     2018-03-30T13:00:12.608 7.418784E-1     mA
TiSapph Beat Note Frequency     2018-03-30T13:00:12.994 7.790371E+8     Hz
Integrated Fluorescence Signal  2018-03-30T13:00:12.263 7.9834792368E-4 arb
G14 Room Temperature South      2018-03-30T13:00:12.290 1.8333689360E+1 C
G14 Room Humidity South 2018-03-30T13:00:12.323 5.2738645000E+1 %
GMR1x   2018-03-30T13:00:12.350 1.2099504471E+1 mG
GMR1y   2018-03-30T13:00:12.350 2.3368239403E+1 mG
GMR1z   2018-03-30T13:00:12.350 2.0061016083E+1 mG
GMR2x   2018-03-30T13:00:12.350 -2.7768015862E+1        mG
GMR2y   2018-03-30T13:00:12.350 -3.5614728928E+1        mG
GMR2z   2018-03-30T13:00:12.350 -2.0369708538E+2        mG
GMR3x   2018-03-30T13:00:12.350 2.2541224957E+2 mG
```

```
GMR3y    2018-03-30T13:00:12.350 1.0328531265E+1 mG
GMR3z    2018-03-30T13:00:12.350 1.1771774292E+2 mG
GMR4x    2018-03-30T13:00:12.350 -3.4404158592E+1        mG
GMR4y    2018-03-30T13:00:12.350 2.3727536201E+1 mG
GMR4z    2018-03-30T13:00:12.350 7.8442096710E+0 mG
G14 Room Pressure       2018-03-30T13:00:12.360 7.6183519424E+2 Torr
Beam Box Skimmer Temperature    2018-03-30T13:00:12.397 0.0000000000E+0 K
J162 Room Temperature   2018-03-30T13:00:12.400 1.9600000000E+1 C
J162 Room Humidity      2018-03-30T13:00:12.400 4.7800000000E+1 %
Beam Box Zorb Temperature       2018-03-30T13:00:12.433 6.2839026624E+0 K
Probe Laser Power       2018-03-30T13:00:12.433 8.4751041758E+1 mW
Cleanup Laser Power     2018-03-30T13:00:12.433 6.4172685337E+2 mW
Beam Box 4K Shield Top Temperature      2018-03-30T13:00:12.467 9.5717745872E+0 K
Integrated Fluorescence Signal  2018-03-30T13:00:12.477 7.9834792368E-4 arb
Ion Sweeper Voltage     2018-03-30T13:00:12.503 -5.0438117853E+2        V
1090A Lock Status       2018-03-30T13:00:12.527 1.0000000000E+0
703A Lock Status        2018-03-30T13:00:12.527 1.0000000000E+0
690A Lock Status        2018-03-30T13:00:12.527 1.0000000000E+0
North Cancellation Current      2018-03-30T13:00:12.540 1.4158755000E+1 A
South Cancellation Current      2018-03-30T13:00:12.577 1.5928475000E+1 A
690 STIRAP AOM Frequency        2018-03-30T13:00:12.577 2.3660000000E+2 Hz
1090 STIRAP AOM Frequency       2018-03-30T13:00:12.577 9.3300000000E+1 Hz
G14 Room Temperature North      2018-03-30T13:00:12.660 1.9400000000E+1 C
G14 Room Humidity North 2018-03-30T13:00:12.660 5.4800000000E+1 %
Integrated Fluorescence Signal  2018-03-30T13:00:12.670 7.9834792368E-4 arb
690 Daredevil Error Signal Width        2018-03-30T13:00:12.703 1.0353406819E+0 MHz
690 Gollum Error Signal Width   2018-03-30T13:00:12.703 1.3729084362E+0 MHz
1064 Cavity 1 Error Signal Width        2018-03-30T13:00:12.703 1.1624089512E-1 MHz
1064 Cavity 2 Error Signal Width        2018-03-30T13:00:12.703 2.7063207204E-1 MHz
Probe Laser Power       2018-03-30T13:00:12.737 8.4751041758E+1 mW
Cleanup Laser Power     2018-03-30T13:00:12.737 6.4172685337E+2 mW
GMR1x    2018-03-30T13:00:12.797 1.3064861298E+1 mG
GMR1y    2018-03-30T13:00:12.797 2.3067593575E+1 mG
GMR1z    2018-03-30T13:00:12.797 1.9447445869E+1 mG
GMR2x    2018-03-30T13:00:12.797 -2.9074192047E+1        mG
GMR2y    2018-03-30T13:00:12.797 -3.5710334778E+1        mG
GMR2z    2018-03-30T13:00:12.797 -2.0535492897E+2        mG
GMR3x    2018-03-30T13:00:12.797 2.2578632832E+2 mG
GMR3y    2018-03-30T13:00:12.797 1.0239124298E+1 mG
GMR3z    2018-03-30T13:00:12.797 1.1745524406E+2 mG
GMR4x    2018-03-30T13:00:12.797 -3.5638093948E+1        mG
GMR4y    2018-03-30T13:00:12.797 2.3824334145E+1 mG
GMR4z    2018-03-30T13:00:12.797 7.6076984406E+0 mG
1090A Lock Status       2018-03-30T13:00:12.830 1.0000000000E+0
703A Lock Status        2018-03-30T13:00:12.830 1.0000000000E+0
690A Lock Status        2018-03-30T13:00:12.830 1.0000000000E+0
Pump room monitoring connection 2018-03-30T13:00:12.877 1.0000000000E+0
G14 Room Temperature North      2018-03-30T13:00:12.883 1.9400000000E+1 C
G14 Room Humidity North 2018-03-30T13:00:12.883 5.4800000000E+1 %
J162 Room Temperature   2018-03-30T13:00:12.893 1.9600000000E+1 C
J162 Room Humidity      2018-03-30T13:00:12.893 4.7800000000E+1 %
G14 Room Pressure       2018-03-30T13:00:12.860 7.6183592288E+2 Torr
Beam Box Skimmer Temperature    2018-03-30T13:00:12.897 0.0000000000E+0 K
```

```
Integrated Fluorescence Signal  2018-03-30T13:00:12.913 7.6140020434E-4 arb
Beam Box Zorb Temperature       2018-03-30T13:00:12.933 6.2874638182E+0 K
ULE A Pressure  2018-03-30T13:00:12.943 6.4000000000E-8 Torr
ULE A Temperature       2018-03-30T13:00:12.943 2.7919638231E+1 C
Beam Box 4K Shield Top Temperature      2018-03-30T13:00:12.967 9.4121258072E+0 K
Ion Sweeper Voltage     2018-03-30T13:00:13.003 -5.0437617543E+2        V
```

Bibliography

[1] Peter W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". *Physical Review Letters* 13.16 (Oct. 1964), pp. 508–509. DOI: `10.1103/PhysRevLett.13.508`.

[2] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. "Global Conservation Laws and Massless Particles". *Physical Review Letters* 13.20 (Nov. 1964), pp. 585–587. DOI: `10.1103/PhysRevLett.13.585`.

[3] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons". *Physical Review Letters* 13.9 (Aug. 1964), pp. 321–323. DOI: `10.1103/PhysRevLett.13.321`.

[4] Sheldon L. Glashow. "Partial-symmetries of weak interactions". *Nuclear Physics* 22.4 (Feb. 1961), pp. 579–588. DOI: `10.1016/0029-5582(61)90469-2`.

[5] Steven Weinberg. "A Model of Leptons". *Physical Review Letters* 19.21 (Nov. 1967), pp. 1264–1266. DOI: `10.1103/PhysRevLett.19.1264`.

[6] H. David Politzer. "Reliable Perturbative Results for Strong Interactions?" *Physical Review Letters* 30.26 (June 1973), pp. 1346–1349. DOI: `10.1103/PhysRevLett.30.1346`.

[7] David J. Gross and Frank Wilczek. "Ultraviolet Behavior of Non-Abelian Gauge Theories". *Physical Review Letters* 30.26 (June 1973), pp. 1343–1346. DOI: `10.1103/PhysRevLett.30.1343`.

[8] Kurt Hübner. "The CERN intersecting storage rings (ISR)". *The European Physical Journal H* 36.4 (Mar. 2012), pp. 509–522. DOI: `10.1140/epjh/e2011-20058-8`.

[9]   J. J. Aubert et al. "Experimental Observation of a Heavy Particle J". *Physical Review Letters* 33.23 (Dec. 1974), pp. 1404–1406. DOI: `10.1103/PhysRevLett.33.1404`.

[10]   J. -E. Augustin et al. "Discovery of a Narrow Resonance in e$^+$ e$^-$ Annihilation". *Physical Review Letters* 33.23 (Dec. 1974), pp. 1406–1408. DOI: `10.1103/PhysRevLett.33.1406`.

[11]   M. L. Perl et al. "Evidence for Anomalous Lepton Production in e$^+$ e$^-$ Annihilation". *Physical Review Letters* 35.22 (Dec. 1975), pp. 1489–1492. DOI: `10.1103/PhysRevLett.35.1489`.

[12]   S. W. Herb et al. "Observation of a Dimuon Resonance at 9.5 GeV in 400-GeV Proton-Nucleus Collisions". *Physical Review Letters* 39.5 (Aug. 1977), pp. 252–255. DOI: `10.1103/PhysRevLett.39.252`.

[13]   D. P. Barber et al. "Discovery of Three-Jet Events and a Test of Quantum Chromodynamics at PETRA". *Physical Review Letters* 43.12 (Sept. 1979), pp. 830–833. DOI: `10.1103/PhysRevLett.43.830`.

[14]   G. Arnison et al. "Experimental observation of lepton pairs of invariant mass around 95 GeV/c^2 at the CERN SPS collider". *Physics Letters B* 126.5 (July 1983), pp. 398–410. DOI: `10.1016/0370-2693(83)90188-0`.

[15]   J. J. Aubert et al. "The ratio of the nucleon structure functions F_2^N for iron and deuterium". *Physics Letters B* 123.3-4 (Mar. 1983), pp. 275–278. DOI: `10.1016/0370-2693(83)90437-9`.

[16]   S. Abachi et al. "Observation of the Top Quark." *Physical Review Letters* 74.14 (1995), pp. 2632–2637. DOI: `10.1103/PhysRevLett.74.2632`.

[17]   F. Abe et al. "Observation of Top Quark Production in $\bar{p}p$ Collisions with the Collider Detector at Fermilab". *Physical Review Letters* 74.14 (1995). DOI: `10.1103/PhysRevLett.74.2626`.

[18]   K. Kodama et al. "Observation of tau neutrino interactions". *Physics Letters B* 504.3 (Apr. 2001), pp. 218–224. DOI: `10.1016/S0370-2693(01)00307-0`.

[19] G. Aad et al. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". *Physics Letters B* 716.1 (Sept. 2012), pp. 1–29. DOI: `10.1016/j.physletb.2012.08.020`.

[20] Matthias Bartelmann. "The dark Universe". *Reviews of Modern Physics* 82.1 (Feb. 2010), pp. 331–382. DOI: `10.1103/RevModPhys.82.331`.

[21] Jonathan L. Feng. "Naturalness and the Status of Supersymmetry". *Annual Review of Nuclear and Particle Science* 63.1 (Oct. 2013), pp. 351–382. DOI: `10.1146/annur ev-nucl-102010-130447`.

[22] David J. E. Marsh. "Axion cosmology". *Physics Reports* 643 (2016), pp. 1–79. DOI: `10.1016/j.physrep.2016.06.005`.

[23] Robert H. Sanders and Stacy S. McGaugh. "Modified Newtonian Dynamics as an Alternative to Dark Matter". *Annual Review of Astronomy and Astrophysics* 40.1 (Sept. 2002), pp. 263–317. DOI: `10.1146/annurev.astro.40.060401.093923`.

[24] Michael E. Peskin and Daniel V. Schroeder. *An Introduction to Quantum Field Theory*. Reading: Perseus Books Publishing, 1995.

[25] J. J. Sakurai. *Modern Quantum Mechanics*. Ed. by San Fu Tuan. Reading: Addison-Wesley, 1994.

[26] Jonathan Engel, Michael J. Ramsey-Musolf, and U. van Kolck. "Electric dipole moments of nucleons, nuclei, and atoms: The Standard Model and beyond". *Progress in Particle and Nuclear Physics* 71 (July 2013), pp. 21–74. DOI: `10.1016/J.PPNP.2 013.03.003`.

[27] Takeshi Fukuyama. "Searching for new physics beyond the Standard Model in electric dipole moment". *International Journal of Modern Physics A* 27.16 (June 2012), p. 1230015. DOI: `10.1142/S0217751X12300153`.

[28] Maxim Pospelov and Adam Ritz. "Electric dipole moments as probes of new physics". *Annals of Physics* 318.1 (July 2005), pp. 119–169. DOI: `10.1016/J.AOP.2005.04.0 02`.

[29] Iosif B. Khriplovich and Steve K. Lamoreaux. *CP Violation Without Strangeness*. Verlag: Springer, 1997. ISBN: 978-3-642-64577-8. DOI: `10.1007/978-3-642-60838-4`.

[30] Julian Schwinger. "The Theory of Quantized Fields. I". *Physical Review* 82.6 (June 1951), pp. 914–927. DOI: `10.1103/PhysRev.82.914`.

[31] Gerhart Lüders. "On the Equivalence of Invariance under Time Reversal and under Particle-Antiparticle Conjugation for Relativistic Field Theories". *Det Kongelige Danske Videnskabernes Selskab Matematisk-fysiske Meddelelser* 28.5 (1954), pp. 1–17. URL: `http://inspirehep.net/record/48330/files/mfm-28-5.pdf`.

[32] Wolfgang Pauli. *Niels Bohr and the development of physics*. London: Pergamon Press, 1955.

[33] David J. Griffiths. *Introduction to elementary particles*. Wiley-VCH, 2008. ISBN: 9783527406012.

[34] Yuval Grossman, Yosef Nir, and Riccardo Rattazzi. "CP Violation Beyond the Standard Model". *Advanced Series on Directions in High Energy Physics: Heavy Flavours II*. Ed. by A. J. Buras and M. Lindner. Singapore: World Scientific Publishing Company, June 1998, pp. 755–794. DOI: `10.1142/9789812812667_0011`.

[35] M. E. Pospelov and I. B. Khriplovich. "Electric dipole moment of the W boson and the electron in the Kobayashi-Maskawa model". *Soviet Journal of Nuclear Physics* 53 (1991), pp. 638–640. URL: `https://inis.iaea.org/collection/NCLCollectionStore/_Public/23/084/23084502.pdf`.

[36] S. L. Glashow, J. Iliopoulos, and L. Maiani. "Weak Interactions with Lepton-Hadron Symmetry". *Physical Review D* 2.7 (Oct. 1970), pp. 1285–1292. DOI: `10.1103/PhysRevD.2.1285`.

[37] Darwin Chang, Wai-Yee Keung, and Jiang Liu. "The electric dipole moment of W-boson". *Nuclear Physics B* 355.2 (May 1991), pp. 295–304. DOI: `10.1016/0550-3213(91)90115-E`.

[38]   Maxim Pospelov and Adam Ritz. "CKM benchmarks for electron electric dipole moment experiments". *Physical Review D* 89.5 (Mar. 2014), p. 056006. DOI: `10.1103/PhysRevD.89.056006`.

[39]   Michael Booth. "The Electric Dipole Moment of the W and Electron in the Standard Model" (Jan. 1993). arXiv: `hep-ph/9301293 [hep-ph]`.

[40]   F. Hoogeveen. "The standard model prediction for the electric dipole moment of the electron". *Nuclear Physics B* 341.2 (Sept. 1990), pp. 322–340. DOI: `10.1016/0550-3213(90)90182-D`.

[41]   J. M. Pendlebury et al. "Revised experimental upper limit on the electric dipole moment of the neutron". *Physical Review D* 92.9 (Nov. 2015), p. 092003. DOI: `10.1103/PhysRevD.92.092003`.

[42]   B. Graner et al. "Reduced Limit on the Permanent Electric Dipole Moment of ^{199}Hg". *Physical Review Letters* 116.16 (Apr. 2016), p. 161601. DOI: `10.1103/PhysRevLett.116.161601`.

[43]   Jihn E. Kim. "Light pseudoscalars, particle physics and cosmology". *Physics Reports* 150.1-2 (June 1987), pp. 1–177. DOI: `10.1016/0370-1573(87)90017-2`.

[44]   Diptimoy Ghosh and Ryosuke Sato. "Lepton electric dipole moment and strong CP violation". *Physics Letters B* 777 (Feb. 2018), pp. 335–339. DOI: `10.1016/J.PHYSLETB.2017.12.052`.

[45]   Christopher Smith and Selim Touati. "Electric dipole moments with and beyond flavor invariants". *Nuclear Physics B* 924 (Nov. 2017), pp. 417–452. DOI: `10.1016/J.NUCLPHYSB.2017.09.013`.

[46]   John Paul Archambault, Andrzej Czarnecki, and Maxim Pospelov. "Electric dipole moments of leptons in the presence of Majorana neutrinos". *Physical Review D* 70.7 (Oct. 2004), p. 073006. DOI: `10.1103/PhysRevD.70.073006`.

[47]   Wolfgang Altmannshofer, Roni Harnik, and Jure Zupan. "Low energy probes of PeV scale sfermions". *Journal of High Energy Physics* 2013.11 (Nov. 2013), p. 202. DOI: `10.1007/JHEP11(2013)202`.

[48]   Takeo Moroi and Minoru Nagai. "Probing supersymmetric model with heavy sfermions using leptonic flavor and CP violations". *Physics Letters B* 723.1-3 (June 2013), pp. 107–112. DOI: 10.1016/J.PHYSLETB.2013.04.049.

[49]   Xiao-Gang He, Bruce H. J. McKellar, and Sandip Pakvasa. "CP violating electron-nucleon interactions in multi-Higgs doublet and leptoquark models". *Physics Letters B* 283.3-4 (June 1992), pp. 348–352. DOI: 10.1016/0370-2693(92)90030-8.

[50]   Yuichiro Nakai and Matthew Reece. "Electric dipole moments in natural supersymmetry". *Journal of High Energy Physics* 2017.8 (Aug. 2017), p. 31. DOI: 10.1007/JHEP08(2017)031.

[51]   Cari Cesarotti et al. "Interpreting the Electron EDM Constraint" (Oct. 2018). arXiv: 1810.07736.

[52]   E. Fermi. "Versuch einer Theorie der β-Strahlen. I". *Zeitschrift für Physik* 88.3-4 (Mar. 1934), pp. 161–177. DOI: 10.1007/BF01351864.

[53]   Fred L. Wilson. "Fermi's Theory of Beta Decay". *American Journal of Physics* 36.12 (Dec. 1968), pp. 1150–1160. DOI: 10.1119/1.1974382.

[54]   Csaba Balazs, Graham White, and Jason Yue. "Effective field theory, electric dipole moments and electroweak baryogenesis". *Journal of High Energy Physics* 2017.3 (Mar. 2017), p. 30. DOI: 10.1007/JHEP03(2017)030.

[55]   Rodrigo Alonso et al. "Renormalization group evolution of the Standard Model dimension six operators III: gauge coupling dependence and phenomenology". *Journal of High Energy Physics* 2014.4 (Apr. 2014), p. 159. DOI: 10.1007/JHEP04(2014)159.

[56]   Andrei D. Sakharov. "Violation of CP invariance, C asymmetry, and baryon asymmetry of the universe". *Soviet Physics Uspekhi* 34.5 (May 1991), pp. 392–393. DOI: 10.1070/PU1991v034n05ABEH002497.

[57]   Laurent Canetti, Marco Drewes, and Mikhail Shaposhnikov. "Matter and antimatter in the universe". *New Journal of Physics* 14.9 (Sept. 2012), p. 095012. DOI: 10.1088/1367-2630/14/9/095012.

[58]   Neil Turok. "A critical review of inflation". *Classical and Quantum Gravity* 19.13 (July 2002), p. 305. DOI: 10.1088/0264-9381/19/13/305.

[59]   Andrei Linde. "Inflationary Cosmology". *Lecture Notes in Physics* 738 (2008), pp. 1–54. DOI: 10.1007/978-3-540-74353-8_1.

[60]   Vladimir M. Dubovik and Valentin E. Kuznetsov. "The Toroid Dipole Moment of the Neutrino". *International Journal of Modern Physics A* 13.30 (Dec. 1998), pp. 5257–5277. DOI: 10.1142/S0217751X98002419.

[61]   M. J. Musolf and Barry R. Holstein. "Observability of the anapole moment and neutrino charge radius". *Physical Review D* 43.9 (May 1991), pp. 2956–2970. DOI: 10.1103/PhysRevD.43.2956.

[62]   V. W. Hughes, L. J. Fraser, and E. R. Carlson. "The electrical neutrality of atoms". *Zeitschrift für Physik D* 10.2-3 (June 1988), pp. 145–151. DOI: 10.1007/BF01384848.

[63]   B. Odom et al. "New Measurement of the Electron Magnetic Moment Using a One-Electron Quantum Cyclotron". *Physical Review Letters* 97.3 (July 2006), p. 030801. DOI: 10.1103/PhysRevLett.97.030801.

[64]   T. Aoyama et al. "Revised value of the eighth-order QED contribution to the anomalous magnetic moment of the electron". *Physical Review D* 77.5 (Mar. 2008), p. 053012. DOI: 10.1103/PhysRevD.77.053012.

[65]   A. Góngora-T and Robin G. Stuart. "The charge radius and anapole moment of a free fermion". *Zeitschrift für Physik C* 55.1 (Mar. 1992), pp. 101–105. DOI: 10.1007/BF01558294.

[66]   W. C. Haxton and C. E. Wieman. "Atomic Parity Nonconservation and Nuclear Anapole Moments". *Annual Review of Nuclear and Particle Science* 51.1 (Dec. 2001), pp. 261–293. DOI: 10.1146/annurev.nucl.51.101701.132458.

[67]   A. N. Petrov et al. "Zeeman interaction in ThO $H^3\Delta_1$ for the electron electric-dipole-moment search". *Physical Review A* 89.6 (June 2014), p. 062505. DOI: 10.1103/PhysRevA.89.062505.

[68]   J.M. Brown et al. "Remarks on the signs of g factors in atomic and molecular Zeeman spectroscopy". *Molecular Physics* 98.20 (Oct. 2000), pp. 1597–1601. DOI: `10.1080/0 026897009483366`.

[69]   E. M. Purcell and N. F. Ramsey. "On the Possibility of Electric Dipole Moments for Elementary Particles and Nuclei". *Physical Review* 78.6 (June 1950), pp. 807–807. DOI: `10.1103/PhysRev.78.807`.

[70]   Norman F. Ramsey. "A Molecular Beam Resonance Method with Separated Oscillating Fields". *Physical Review* 78.6 (June 1950), pp. 695–699. DOI: `10.1103/Phys Rev.78.695`.

[71]   L. I. Schiff. "Measurability of Nuclear Electric Dipole Moments". *Physical Review* 132.5 (Dec. 1963), pp. 2194–2200. DOI: `10.1103/PhysRev.132.2194`.

[72]   P. G. H. Sandars. "The electric dipole moment of an atom". *Physics Letters* 14.3 (Feb. 1965), pp. 194–196. DOI: `10.1016/0031-9163(65)90583-4`.

[73]   Eugene D. Commins, J. D. Jackson, and David P. DeMille. "The electric dipole moment of the electron: An intuitive explanation for the evasion of Schiff's theorem". *American Journal of Physics* 75.6 (June 2007), pp. 532–536. DOI: `10.1119/1.2710 486`.

[74]   J. P. Carrico et al. "Electric Dipole Moments of Alkali Atoms. A Limit to the Electric Dipole Moment of the Free Electron". *Physical Review* 174.1 (Oct. 1968), pp. 125–138. DOI: `10.1103/PhysRev.174.125`.

[75]   Z. W. Liu and Hugh P. Kelly. "Analysis of atomic electric dipole moment in thallium by all-order calculations in many-body perturbation theory". *Physical Review A* 45.7 (Apr. 1992), R4210–R4213. DOI: `10.1103/PhysRevA.45.R4210`.

[76]   Ann-Marie Mårtensson-Pendrill and Per Öster. "Calculations of atomic electric dipole moments". *Physica Scripta* 36.3 (Sept. 1987), pp. 444–452. DOI: `10.1088/0031-894 9/36/3/011`.

[77] E. E. Salpeter. "Some Atomic Effects of an Electronic Electric Dipole Moment". *Physical Review* 112.5 (Dec. 1958), pp. 1642–1648. DOI: `10.1103/PhysRev.112.1642`.

[78] D. T. Wilkinson and H. R. Crane. "Precision Measurement of the g Factor of the Free Electron". *Physical Review* 130.3 (May 1963), pp. 852–863. DOI: `10.1103/PhysRev.130.852`.

[79] D. F. Nelson et al. "Search for an Electric Dipole Moment of the Electron". *Physical Review Letters* 2.12 (June 1959), pp. 492–495. DOI: `10.1103/PhysRevLett.2.492`.

[80] J. Goldemberg and Y. Torizuka. "Upper Limit of the Electric Dipole Moment of the Electron". *Physical Review* 129.6 (Mar. 1963), pp. 2580–2581. DOI: `10.1103/PhysRev.129.2580`.

[81] R. E. Rand. "Determination of the Upper Limit to the Electric Dipole Moment of the Electron at High Momentum Transfer". *Physical Review* 140.6B (Dec. 1965), B1605–B1610. DOI: `10.1103/PhysRev.140.B1605`.

[82] E. S. Ensberg. "Experimental Upper Limit for the Permanent Electric Dipole Moment of Rb[85] by Optical-Pumping Techniques". *Physical Review* 153.1 (Jan. 1967), pp. 36–43. DOI: `10.1103/PhysRev.153.36`.

[83] P. G. H. Sandars and E. Lipworth. "Electric Dipole Moment of the Cesium Atom. A New Upper Limit to the Electric Dipole Moment of the Free Electron". *Physical Review Letters* 13.24 (Dec. 1964), pp. 718–720. DOI: `10.1103/PhysRevLett.13.718`.

[84] T. S. Stein et al. "Electric Dipole Moment of the Cesium Atom. A New Upper Limit to the Electric Dipole Moment of the Free Electron". *Physical Review Letters* 19.13 (Sept. 1967), pp. 741–743. DOI: `10.1103/PhysRevLett.19.741`.

[85] J. R. P. Angel, P. G. H. Sandars, and M. H. Tinker. "Observation of a $v \times E$ effect in an electric dipole moment experiment using a reversible atomic beam machine". *Physics Letters A* 25.2 (July 1967), pp. 160–161. DOI: `10.1016/0375-9601(67)90401-X`.

[86]   M. C. Weisskopf et al. "Electric Dipole Moment of the Cesium Atom. A New Upper Limit to the Electric Dipole Moment of the Electron". *Physical Review Letters* 21.24 (Dec. 1968), pp. 1645–1648. DOI: `10.1103/PhysRevLett.21.1645`.

[87]   T. S. Stein et al. "Permanent Electric Dipole Moment of the Cesium Atom. An Upper Limit to the Electric Dipole Moment of the Electron". *Physical Review* 186.1 (Oct. 1969), pp. 39–51. DOI: `10.1103/PhysRev.186.39`.

[88]   S. A. Murthy et al. "New limits on the electron electric dipole moment from cesium". *Physical Review Letters* 63.9 (Aug. 1989), pp. 965–968. DOI: `10.1103/PhysRevLett.63.965`.

[89]   M. Bijlsma, B. J. Verhaar, and D. J. Heinzen. "Role of collisions in the search for an electron electric-dipole moment". *Physical Review A* 49.6 (June 1994), R4285–R4288. DOI: `10.1103/PhysRevA.49.R4285`.

[90]   Cheng Chin et al. "Measurement of an electron's electric dipole moment using Cs atoms trapped in optical lattices". *Physical Review A* 63.3 (Feb. 2001), p. 033401. DOI: `10.1103/PhysRevA.63.033401`.

[91]   Fang Fang and David S. Weiss. "Resonator-enhanced optical guiding and trapping of Cs atoms". *Optics Letters* 34.2 (Jan. 2009), p. 169. DOI: `10.1364/OL.34.000169`.

[92]   Harvey Gould. "Search for an Electric Dipole Moment in Thallium". *Physical Review Letters* 24.20 (May 1970), pp. 1091–1093. DOI: `10.1103/PhysRevLett.24.1091`.

[93]   P. G. H. Sandars and R. M. Sternheimer. "Electric-dipole-moment enhancement factor for the thallium atom, and a new upper limit on the electric dipole moment of the electron". *Physical Review A* 11.2 (Feb. 1975), pp. 473–476. DOI: `10.1103/PhysRevA.11.473`.

[94]   K. Abdullah et al. "New experimental limit on the electron electric dipole moment". *Physical Review Letters* 65.19 (Nov. 1990), pp. 2347–2350. DOI: `10.1103/PhysRevLett.65.2347`.

[95]   Eugene D. Commins et al. "Improved experimental limit on the electric dipole moment of the electron". *Physical Review A* 50.4 (Oct. 1994), pp. 2960–2977. DOI: `10.1103/PhysRevA.50.2960`.

[96]   B. C. Regan et al. "New Limit on the Electron Electric Dipole Moment". *Physical Review Letters* 88.7 (Feb. 2002), p. 071805. DOI: `10.1103/PhysRevLett.88.071805`.

[97]   V. V. Flambaum and I. B. Khriplovich. "New bounds on the electric dipole moment of the electron and on T-odd electron-nucleon coupling". *Journal of Experimental and Theoretical Physics* 89.5 (1985), p. 1505. URL: `http://www.jetp.ac.ru/cgi-bin/dn/e_062_05_0872.pdf`.

[98]   V. A. Dzuba, V. V. Flambaum, and S. G. Porsev. "Calculation of (P,T)-odd electric dipole moments for the diamagnetic atoms ^{129}Xe , ^{171}Yb , ^{199}Hg , ^{211}Rn , and ^{225}Ra". *Physical Review A* 80.3 (Sept. 2009), p. 032120. DOI: `10.1103/PhysRevA.80.032120`.

[99]   W. Happer and A. C. Tam. "Effect of rapid spin exchange on the magnetic-resonance spectrum of alkali vapors". *Physical Review A* 16.5 (Nov. 1977), pp. 1877–1891. DOI: `10.1103/PhysRevA.16.1877`.

[100] R. A. Sen'kov et al. "Reexamination of the Schiff theorem". *Physical Review A* 77.1 (Jan. 2008), p. 014101. DOI: `10.1103/PhysRevA.77.014101`.

[101] M. A. Player and P. G. H. Sandars. "An experiment to search for an electric dipole moment in the ^{3}P$_2$ metastable state of xenon". *Journal of Physics B* 3.12 (Dec. 1970), pp. 1620–1635. DOI: `10.1088/0022-3700/3/12/007`.

[102] T. G. Vold et al. "Search for a Permanent Electric Dipole Moment on the ^{129}Xe Atom". *Physical Review Letters* 52.25 (June 1984), pp. 2229–2232. DOI: `10.1103/PhysRevLett.52.2229`.

[103] M. A. Rosenberry and T. E. Chupp. "Atomic Electric Dipole Moment Measurement Using Spin Exchange Pumped Masers of ^{129}Xe and ^{3}He". *Physical Review Letters* 86.1 (Jan. 2001), pp. 22–25. DOI: `10.1103/PhysRevLett.86.22`.

[104] S. K. Lamoreaux et al. "New constraints on time-reversal asymmetry from a search for a permanent electric dipole moment of ^{199}Hg". *Physical Review Letters* 59.20 (Nov. 1987), pp. 2275–2278. DOI: `10.1103/PhysRevLett.59.2275`.

[105] J. P. Jacobs et al. "Testing time-reversal symmetry using ^{199}Hg". *Physical Review Letters* 71.23 (Dec. 1993), pp. 3782–3785. DOI: `10.1103/PhysRevLett.71.3782`.

[106] J. P. Jacobs et al. "Limit on the electric-dipole moment of ^{199}Hg using synchronous optical pumping". *Physical Review A* 52.5 (Nov. 1995), pp. 3521–3540. DOI: `10.1103/PhysRevA.52.3521`.

[107] M. V. Romalis et al. "New Limit on the Permanent Electric Dipole Moment of ^{199}Hg". *Physical Review Letters* 86.12 (Mar. 2001), pp. 2505–2508. DOI: `10.1103/PhysRevLett.86.2505`.

[108] W. C. Griffith et al. "Improved Limit on the Permanent Electric Dipole Moment of ^{199}Hg". *Physical Review Letters* 102.10 (Mar. 2009), p. 101601. DOI: `10.1103/PhysRevLett.102.101601`.

[109] V. Spevak, N. Auerbach, and V. V. Flambaum. "Enhanced T-odd, P-odd electromagnetic moments in reflection asymmetric nuclei". *Physical Review C* 56.3 (Sept. 1997), pp. 1357–1369. DOI: `10.1103/PhysRevC.56.1357`.

[110] V. V. Flambaum and V. G. Zelevinsky. "Enhancement of nuclear Schiff moments and time-reversal violation in atoms due to soft nuclear octupole vibrations". *Physical Review C* 68.3 (Sept. 2003), p. 035502. DOI: `10.1103/PhysRevC.68.035502`.

[111] R. H. Parker et al. "First Measurement of the Atomic Electric Dipole Moment of ^{225}Ra". *Physical Review Letters* 114.23 (June 2015), p. 233002. DOI: `10.1103/PhysRevLett.114.233002`.

[112] Michael Bishof et al. "Improved limit on the ^{225}Ra electric dipole moment". *Physical Review C* 94.2 (Aug. 2016), p. 025501. DOI: `10.1103/PhysRevC.94.025501`.

[113] E. R. Tardiff et al. "The radon EDM apparatus". *Hyperfine Interactions* 225.1-3 (Jan. 2014), pp. 197–206. DOI: `10.1007/s10751-013-0898-2`.

[114]   P. G. H. Sandars. "Measurability of the Proton Electric Dipole Moment". *Physical Review Letters* 19.24 (Dec. 1967), pp. 1396–1398. DOI: `10.1103/PhysRevLett.19.1396`.

[115]   Edward A. Hinds and P. G. H. Sandars. "Experiment to search for P- and T-violating interactions in the hyperfine structure of thallium fluoride". *Physical Review A* 21.2 (Feb. 1980), pp. 480–487. DOI: `10.1103/PhysRevA.21.480`.

[116]   Dean A. Wilkening, Norman F. Ramsey, and Daniel J. Larson. "Search for P and T violations in the hyperfine structure of thallium fluoride". *Physical Review A* 29.2 (Feb. 1984), pp. 425–438. DOI: `10.1103/PhysRevA.29.425`.

[117]   D. Schropp et al. "New limits on time-reversal invariance from the hyperfine structure of thallium fluoride". *Physical Review Letters* 59.9 (Aug. 1987), pp. 991–994. DOI: `10.1103/PhysRevLett.59.991`.

[118]   D. Cho, K. Sangster, and E. A. Hinds. "Tenfold improvement of limits on T violation in thallium fluoride". *Physical Review Letters* 63.23 (Dec. 1989), pp. 2559–2562. DOI: `10.1103/PhysRevLett.63.2559`.

[119]   L. R. Hunter et al. "Prospects for laser cooling TlF". *Physical Review A* 85.1 (Jan. 2012), p. 012511. DOI: `10.1103/PhysRevA.85.012511`.

[120]   E. B. Norrgard et al. "Hyperfine structure of the $B^3\Pi_1$ state and predictions of optical cycling behavior in the $X \to B$ transition of TlF". *Physical Review A* 95.6 (June 2017), p. 062506. DOI: `10.1103/PhysRevA.95.062506`.

[121]   J. J. Hudson et al. "Measurement of the Electron Electric Dipole Moment Using YbF Molecules". *Physical Review Letters* 89.2 (June 2002), p. 023003. DOI: `10.1103/PhysRevLett.89.023003`.

[122]   J. J. Hudson et al. "Improved measurement of the shape of the electron". *Nature* 473.7348 (May 2011), pp. 493–496. DOI: `10.1038/nature10104`.

[123]   M. R. Tarbutt et al. "Design for a fountain of YbF molecules to measure the electron's electric dipole moment". *New Journal of Physics* 15.5 (May 2013), p. 053034. DOI: `10.1088/1367-2630/15/5/053034`.

[124] D. Kawall et al. "Precision Zeeman-Stark Spectroscopy of the Metastable $a(1)[^3\Sigma^+]$ State of PbO". *Physical Review Letters* 92.13 (Apr. 2004), p. 133007. DOI: `10.1103 /PhysRevLett.92.133007`.

[125] S. Eckel et al. "Search for the electron electric dipole moment using Ω-doublet levels in PbO". *Physical Review A* 87.5 (May 2013), p. 052130. DOI: `10.1103/PhysRevA.8 7.052130`.

[126] A. C. Vutha et al. "Search for the electric dipole moment of the electron with thorium monoxide". *Journal of Physics B* 44.7 (Apr. 2011), p. 079803. DOI: `10.1088/0953 -4075/44/7/079803`.

[127] A. C. Vutha et al. "Magnetic and electric dipole moments of the $H^3\Delta_1$ state in ThO". *Physical Review A* 84.3 (Sept. 2011), p. 034502. DOI: `10.1103/PhysRevA.84.034502`.

[128] J. Baron et al. "Order of magnitude smaller limit on the electric dipole moment of the electron". *Science* 343.6168 (Jan. 2014), pp. 269–72. DOI: `10.1126/science.12 48213`.

[129] J. Baron et al. "Methods, analysis, and the treatment of systematic errors for the electron electric dipole moment search in thorium monoxide". *New Journal of Physics* 19.7 (July 2017), p. 073029. DOI: `10.1088/1367-2630/aa708e`.

[130] J. Lee et al. "An electron electric dipole moment search in the X $^3\Delta_1$ ground state of tungsten carbide molecules". *Journal of Modern Optics* 56.18-19 (Oct. 2009), pp. 2005–2012. DOI: `10.1080/09500340903349930`.

[131] J. Lee et al. "Optical spectroscopy of tungsten carbide for uncertainty analysis in electron electric-dipole-moment search". *Physical Review A* 87.2 (Feb. 2013), p. 022516. DOI: `10.1103/PhysRevA.87.022516`.

[132] Parul Aggarwal et al. "Measuring the electric dipole moment of the electron in BaF". *The European Physical Journal D* 72.11 (Nov. 2018), p. 197. DOI: `10.1140/epjd/e 2018-90192-9`.

[133]   Ivan Kozyryev and Nicholas R. Hutzler. "Precision Measurement of Time-Reversal Symmetry Violation with Laser-Cooled Polyatomic Molecules". *Physical Review Letters* 119.13 (Sept. 2017), p. 133002. DOI: `10.1103/PhysRevLett.119.133002`.

[134]   A. E. Leanhardt et al. "High-resolution spectroscopy on trapped molecular ions in rotating electric fields: A new approach for measuring the electron electric dipole moment". *Journal of Molecular Spectroscopy* 270.1 (Nov. 2011), pp. 1–25. DOI: `10.1016/J.JMS.2011.06.007`.

[135]   D. Kawall. "Searching for the electron EDM in a storage ring". *Journal of Physics: Conference Series* 295.1 (May 2011), p. 012031. DOI: `10.1088/1742-6596/295/1/012031`.

[136]   William B. Cairncross et al. "Precision Measurement of the Electron's Electric Dipole Moment Using Trapped Molecular Ions". *Physical Review Letters* 119.15 (Oct. 2017), p. 153001. DOI: `10.1103/PhysRevLett.119.153001`.

[137]   B. V. Vasil'ev and E. V. Kolycheva. "Measurement of the electric dipole moment of the electron with a quantum interferometer". *Journal of Experimental and Theoretical Physics* 47.2 (1978), p. 243. URL: `http://www.jetp.ac.ru/cgi-bin/dn/e_047_02_0243.pdf`.

[138]   S. Eckel, A. O. Sushkov, and S. K. Lamoreaux. "Limit on the Electron Electric Dipole Moment Using Paramagnetic Ferroelectric $Eu_{0.5}Ba_{0.5}TiO_3$". *Physical Review Letters* 109.19 (Nov. 2012), p. 193003. DOI: `10.1103/PhysRevLett.109.193003`.

[139]   Y. J. Kim et al. "New experimental limit on the electric dipole moment of the electron in a paramagnetic insulator". *Physical Review D* 91.10 (May 2015), p. 102004. DOI: `10.1103/PhysRevD.91.102004`.

[140]   B. J. Heidenreich et al. "Limit on the Electron Electric Dipole Moment in Gadolinium-Iron Garnet". *Physical Review Letters* 95.25 (Dec. 2005), p. 253004. DOI: `10.1103/PhysRevLett.95.253004`.

[141] A. C. Vutha, M. Horbatsch, and E. A. Hessels. "Oriented Polar Molecules in a Solid Inert-Gas Matrix: A Proposed Method for Measuring the Electric Dipole Moment of the Electron". *Atoms* 6.1 (Jan. 2018), p. 3. DOI: `10.3390/atoms6010003`.

[142] A C Vutha, M Horbatsch, and E A Hessels. "Orientation-dependent hyperfine structure of polar molecules in a rare-gas matrix: A scheme for measuring the electron electric dipole moment". *Physical Review A* 98 (2018), p. 32513. DOI: `10.1103/Phy sRevA.98.032513`.

[143] M. Bardeen et al. "Planning the Future of U.S. Particle Physics". *2013 Community Summer Study on the Future of U.S. Particle Physics.* Ed. by Norman Graf, Michael Peskin, and Jonathan Rosner. Minneapolis: SLAC-econf-C130729.2, FERMILAB-CONF-13-648, SLAC-PUB-15960, 2013. URL: `http://www.slac.stanford.edu/ec onf/C1307292/docs/SnowmassBook.pdf`.

[144] Amar Chandra Vutha. "A search for the electric dipole moment of the electron using thorium monoxide". PhD thesis. Yale University, 2011. URL: `http://www.doylegr oup.harvard.edu/files/bufferpubs/theses/acv_thesis.PDF`.

[145] Yulia Vsevolodovna Gurevich. "Preliminary Measurements for an Electron EDM Experiment in ThO". PhD thesis. Harvard University, 2012. URL: `http://www.doy legroup.harvard.edu/files/bufferpubs/theses/yvg_thesis.pdf`.

[146] Nicholas Richard Hutzler. "A New Limit on the Electron Electric Dipole Moment: Beam Production, Data Interpretation, and Systematics". PhD thesis. Harvard University, 2014. URL: `http://www.doylegroup.harvard.edu/files/bufferpubs/th eses/nrh_thesis.pdf`.

[147] Benjamin Norman Spaun. "A Ten-Fold Improvement to the Limit of the Electron Electric Dipole Moment". PhD thesis. Harvard University, 2014. URL: `http://www .doylegroup.harvard.edu/files/bufferpubs/theses/bns_thesis.pdf`.

[148] Paul William Hess. "Improving the Limit on the Electron EDM: Data Acquisition and Systematics Studies in the ACME Experiment". PhD thesis. Harvard University,

2014. URL: `http://www.doylegroup.harvard.edu/files/bufferpubs/theses/pwh_thesis.pdf`.

[149] Brendon R. O'Leary. "In search of the electron's electric dipole moment in thorium monoxide: an improved upper limit, systematic error models, and apparatus upgrades". PhD thesis. Yale University, 2016. URL: `http://www.doylegroup.harvard.edu/wiki/images/5/5e/O%27Leary%2C_Brendon_-_thesis_v3.pdf`.

[150] Vitaly Andreev. *Polarimetry on the Advanced Cold Molecule Electron Electric Dipole Moment Experiment.* 2016. URL: `http://www.doylegroup.harvard.edu/wiki/images/5/5e/ThesisV2012final.pdf`.

[151] Elizabeth Petrik West. "A Thermochemical Cryogenic Buffer Gas Beam Source of ThO for Measuring the Electric Dipole Moment of the Electron". PhD thesis. Harvard University, 2017. URL: `http://www.doylegroup.harvard.edu/wiki/images/a/a2/Elizabeth_West_PhD-20170502.pdf`.

[152] Nicholas R. Hutzler et al. "A cryogenic beam of refractory, chemically reactive molecules with expansion cooling". *Physical Chemistry Chemical Physics* 13.42 (Oct. 2011), p. 18976. DOI: `10.1039/c1cp20901a`.

[153] Wesley C. Campbell et al. "Advanced cold molecule electron EDM". *EPJ Web of Conferences* 57 (Aug. 2013). Ed. by O. Dulieu et al., p. 02004. DOI: `10.1051/epjconf/20135702004`.

[154] E. Kirilov et al. "Shot-noise-limited spin measurements in a pulsed molecular beam". *Physical Review A* 88.1 (July 2013), p. 013844. DOI: `10.1103/PhysRevA.88.013844`.

[155] C. D. Panda et al. "Stimulated Raman adiabatic passage preparation of a coherent superposition of ThO $H^3\Delta_1$ states for an improved electron electric-dipole-moment measurement". *Physical Review A* 93.5 (May 2016), p. 052110. DOI: `10.1103/PhysRevA.93.052110`.

[156] ACME Collaboration. "Improved limit on the electric dipole moment of the electron". *Nature* 562.7727 (Oct. 2018), pp. 355–360. DOI: `10.1038/s41586-018-0599-8`.

[157]   Gerhard Herzberg and J. W. T. Spinks. *Molecular Spectra and Molecular Structure.* 2nd ed. Princeton: D. Van Nostrand Company, Inc., 1950.

[158]   John Brown and Alan Carrington. *Rotational Spectroscopy of Diatomic Molecules.* Cambridge: Cambridge University Press, 2003. ISBN: 9780521810098.

[159]   Jozef Paulovič et al. "Relativistic and correlated calculations on the ground and excited states of ThO". *The Journal of Chemical Physics* 119.2 (July 2003), pp. 798–805. DOI: 10.1063/1.1578053.

[160]   Gunnar Edvinsson and Albin Lagerqvist. "A low-lying $\Omega = 2$ state in the ThO molecule". *Journal of Molecular Spectroscopy* 113.1 (Sept. 1985), pp. 93–104. DOI: 10.1016/0022-2852(85)90123-7.

[161]   Malika Denis and Timo Fleig. "In search of discrete symmetry violations beyond the standard model: Thorium monoxide reloaded". *The Journal of Chemical Physics* 145.21 (Dec. 2016), p. 214307. DOI: 10.1063/1.4968597.

[162]   L. V. Skripnikov. "Combined 4-component and relativistic pseudopotential study of ThO for the electron electric dipole moment search". *The Journal of Chemical Physics* 145.21 (Dec. 2016), p. 214301. DOI: 10.1063/1.4968229.

[163]   L. V. Skripnikov, A. N. Petrov, and A. V. Titov. "Communication: Theoretical study of ThO for the electron electric dipole moment search". *The Journal of Chemical Physics* 139.22 (Dec. 2013), p. 221103. DOI: 10.1063/1.4843955.

[164]   B. W. Shore. "Examples of counter-intuitive physics". *Contemporary Physics* 36.1 (Jan. 1995), pp. 15–28. DOI: 10.1080/00107519508222135.

[165]   David J. Griffiths. *Introduction to Quantum Mechanics.* 2nd ed. Upper Saddle River: Pearson Prentice Hall, 2005.

[166]   Brendon O'leary. *Rotational Cooling.* 2014. URL: http://www.doylegroup.harvard.edu/files/internal/2014_01_16_Rotational_Cooling_In_The_Next_Generation.pdf.

[167] Nicholas Hutzler. *Rotational Temperature, First Data*. 2009. URL: `https://lablog s.gabrielse.physics.northwestern.edu/lablog_EDMexperiment/2009/10/rot ational_temperature_first_d_1.html`.

[168] Nicholas Hutzler. *Rotational Temperatures vs Flow and Position*. 2010. URL: `https ://lablogs.gabrielse.physics.northwestern.edu/lablog_EDMexperiment/20 10/05/rotational_temperatures_vs_flo.html`.

[169] Richard N. Zare. *Angular Momentum: Understanding Spatial Aspects in Chemistry and Physics*. New York: John Wiley and Sons, 1988.

[170] M. Born and R. Oppenheimer. "Zur Quantentheorie der Molekeln". *Annalen der Physik* 389.20 (Jan. 1927), pp. 457–484. DOI: `10.1002/andp.19273892002`.

[171] M. Born and J. R. Oppenheimer. *On the Quantum Theory of Molecules*. URL: `http s://www.ulb.ac.be/cpm/people/bsutclif/bornopn_corr.pdf`.

[172] D. Budker, Derek F. Kimball, and David P. DeMille. *Atomic Physics: An Exploration through Problems and Solutions*. 2nd ed. Oxford: Oxford University Press, 2008.

[173] Tunis Wentink, Jr. and Robert J. Spindler, Jr. "The isoelectronic series ScF through ThO - I notes on the band spectra of TiO, HfO and ThO". *Journal of Quantitative Spectroscopy and Radiative Transfer* 12.11 (Nov. 1972), pp. 1569–1590. DOI: `10.10 16/0022-4073(72)90131-8`.

[174] Damian L. Kokkin, Timothy C. Steimle, and David DeMille. "Branching ratios and radiative lifetimes of the U, L, and I states of thorium oxide". *Physical Review A* 90.6 (Dec. 2014), p. 062503. DOI: `10.1103/PhysRevA.90.062503`.

[175] Zack Lasner. *Updated rotational cooling write-up*. 2016. URL: `https://lablogs.gab rielse.physics.northwestern.edu/lablog_EDMgraduate_2014-Zach_Lasner/2 016/01/updated-rotational-cooling-write-up.html`.

[176] M. Teich, K. Matsuo, and B. Saleh. "Excess noise factors for conventional and super-lattice avalanche photodiodes and photomultiplier tubes". *IEEE Journal of Quantum Electronics* 22.8 (Aug. 1986), pp. 1184–1193. DOI: `10.1109/JQE.1986.1073137`.

[177] Ward Whitt. *Compound Poisson Process*. 2012. URL: http://www.columbia.edu
/~ww2040/6711F12/lect0927.pdf.

[178] David A. van Dyk et al. "Analysis of Energy Spectra with Low Photon Counts
via Bayesian Posterior Simulation". *The Astrophysical Journal* 548.1 (Feb. 2001),
pp. 224–243. DOI: 10.1086/318656.

[179] Taeyoung Park et al. "Bayesian Estimation of Hardness Ratios: Modeling and Com-
putations". *The Astrophysical Journal* 652.1 (Nov. 2006), pp. 610–628. DOI: 10.108
6/507406.

[180] D. Karlis and I. Ntzoufras. "Bayesian analysis of the differences of count data".
Statistics in Medicine 25.11 (June 2006), pp. 1885–1905. DOI: 10.1002/sim.2382.

[181] Jack Hayya, Donald Armstrong, and Nicolas Gressis. "A Note on the Ratio of Two
Normally Distributed Variables". *Management Science* 21.11 (July 1975), pp. 1338–
1341. DOI: 10.1287/mnsc.21.11.1338.

[182] P. G. Moschopoulos and W. B. Canada. "The distribution function of a linear com-
bination of chi-squares". *Computers & Mathematics with Applications* 10.4-5 (Jan.
1984), pp. 383–386. DOI: 10.1016/0898-1221(84)90066-X.

[183] Pennsylvania State University. *Sums of Chi-Square Random Variables*. 2018. URL:
https://onlinecourses.science.psu.edu/stat414/node/171/.

[184] S. S. Shapiro and M. B. Wilk. "An Analysis of Variance Test for Normality (Complete
Samples)". *Biometrika* 52.3/4 (Dec. 1965), p. 591. DOI: 10.2307/2333709.

[185] S. S. Shapiro and R. S. Francia. "An Approximate Analysis of Variance Test for
Normality". *Journal of the American Statistical Association* 67.337 (Mar. 1972),
pp. 215–216. DOI: 10.1080/01621459.1972.10481232.

[186] Maurice G. Kendall and Alan Stuart. *The Advanced Theory of Statistics*. 3rd ed.
London: Charles Griffin & Company, Ltd., 1969.

[187] Gwowen Shieh. "Estimation of the simple correlation coefficient". *Behavior Research
Methods* 42.4 (Nov. 2010), pp. 906–917. DOI: 10.3758/BRM.42.4.906.

[188]   Raymond Kan and Xiaolu Wang. "On the distribution of the sample autocorrelation coefficients". *Journal of Econometrics* 154.2 (Feb. 2010), pp. 101–121. DOI: `10.1016/J.JECONOM.2009.06.010`.

[189]   Wayne A. Fuller. *Measurement Error Models*. New York: John Wiley and Sons, 2009.

[190]   Rand R. Wilcox. *Introduction to Robust Estimation and Hypothesis Testing*. 3rd ed. Amsterdam: Academic Press, 2005.

[191]   Catherine Stuart. *Robust Regression*. 2011. URL: `https://pdfs.semanticscholar.org/f50f/6e74b773ba0b89df34744523bd7c6b148125.pdf`.

[192]   Peter J. Huber. "Robust Estimation of a Location Parameter". *The Annals of Mathematical Statistics* 35.1 (Mar. 1964), pp. 73–101. DOI: `10.1214/aoms/1177703732`.

[193]   Leonard A. Stefanski and Dennis D. Boos. "The Calculus of M-Estimation". *The American Statistician* 56.1 (Feb. 2002), pp. 29–38. DOI: `10.1198/000313002753631330`.

[194]   R. Frühwirth et al. *Vertex reconstruction and track bundling at the LEP collider using robust algorithms*. 1996. URL: `https://web-docs.gsi.de/~ikisel/reco/Methods/Fruehwirth-LEP-CPC96-1996.pdf`.

[195]   Ian Joseph Smallman. "A New Measurement of the Electron Electric Dipole Moment Using Ytterbium Fluoride". PhD thesis. Imperial College London, 2013. URL: `https://www.imperial.ac.uk/media/imperial-college/research-centres-and-groups/the-centre-for-cold-matter/public/Joe-Smallman-2013.compressed.pdf`.

[196]   Bradley Efron. "Bootstrap Methods: Another Look at the Jackknife". *Breakthroughs in Statistics*. Ed. by Samuel Kotz and Norman L. Johnson. New York: Springer, 1992, pp. 569–593. DOI: `10.1007/978-1-4612-4380-9_41`.

[197]   B. Efron and R. Tibshirani. "Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy". *Statistical Science* 1.1 (Feb. 1986), pp. 54–75. DOI: `10.1214/ss/1177013815`.

[198] V. Andreev et al. "A self-calibrating polarimeter to measure Stokes parameters" (Feb. 2017). arXiv: 1703.00963.

[199] Gary J. Feldman and Robert D. Cousins. "Unified approach to the classical statistical analysis of small signals". *Physical Review D* 57.7 (Apr. 1998), pp. 3873–3889. DOI: 10.1103/PhysRevD.57.3873.

[200] Jonathan Haefner. *Single Fluke Electric Field Stability, Gen II.* 2016. URL: https://lablogs.gabrielse.physics.northwestern.edu/lablog_EDMgraduate_2016-Jonathan_Haefner/2016/09/single-fluke-electric-field-stability-gen-ii.html.

[201] Zack Lasner. *E-field lead switching circuit.* 2016. URL: https://lablogs.gabrielse.physics.northwestern.edu/lablog_EDMgraduate_2014-Zach_Lasner/2016/06/e-field-lead-switching-circuit.html.

[202] Adam West. *G14 B-field Cancellation Update.* 2015. URL: https://lablogs.gabrielse.physics.northwestern.edu/lablog_EDMpostdoc_2013-Adam_West/2015/10/g14-b-field-cancellation-update.html.

[203] Kirt Blattenberger. *Conductor Bulk Resistivity & Skin Depths at High Frequencies Microwave RF.* 2018. URL: http://www.rfcafe.com/references/electrical/cond-high-freq.htm.

[204] Roland Winston et al. *Nonimaging optics.* Amsterdam: Elsevier Academic Press, 2005. ISBN: 9780127597515.

[205] Eugene Hecht. *Optics.* 2nd ed. Reading: Addison-Wesley, 1987.

[206] Xing Wu. *Zeeman shift, Helmholtz Coils connected in series.* 2018. URL: https://lablogs.gabrielse.physics.northwestern.edu/lablog_EDMpostdoc_2017-Xing_Wu/2018/06/zeeman-shift-helmholtz-coils-connected-in-series.html.

[207] Xing Wu. *update on Q-C transition.* 2018. URL: https://lablogs.gabrielse.physics.northwestern.edu/lablog_EDMpostdoc_2017-Xing_Wu/2018/05/update-on-q-c-transition.html.

[208] Dimitri P. Bertsekas and John N. Tsitsiklis. *Introduction to Probability*. 1st ed. Belmont: Athena Scientific, 2002.

[209] *Photomultiplier Tubes and Assemblies for Scintillation Counting & High Energy Physics*. Tech. rep. Hamamatsu Photonics, 2017. URL: `https://www.hamamatsu.com/resources/pdf/etd/High_energy_PMT_TPMZ0003E.pdf`.

[210] *J-Series High PDE and Timing Resolution*. Tech. rep. SensL, 2017. URL: `http://sensl.com/downloads/ds/DS-MicroJseries.pdf`.

[211] *Cooled Large Area Red Silicon Photodiode SD 394-70-72-591*. Tech. rep. Luna Optoelectronics, 2016. URL: `http://lunainc.com/wp-content/uploads/2016/07/3947072591.pdf`.

[212] Marco Ramilli. "Characterization of SiPM: Temperature dependencies". *2008 IEEE Nuclear Science Symposium Conference Record*. IEEE, Oct. 2008, pp. 2467–2470. ISBN: 978-1-4244-2714-7. DOI: `10.1109/NSSMIC.2008.4774854`.

[213] W. C. J. Hunter et al. "Measured temperature dependence of scintillation camera signals read out by geiger-Müller mode avalanche photodiodes". *2009 IEEE Nuclear Science Symposium Conference Record (NSS/MIC)*. IEEE, Oct. 2009, pp. 2662–2665. ISBN: 978-1-4244-3961-4. DOI: `10.1109/NSSMIC.2009.5401995`.

[214] *Introduction to the Silicon Photomultiplier*. Tech. rep. SensL, 2011. URL: `https://www.sensl.com/downloads/ds/TN%20-%20Intro%20to%20SPM%20Tech.pdf`.

[215] Y. Wang et al. "Feasibility study of SiGHT: a novel ultra low background photosensor for low temperature operation". *Journal of Instrumentation* 12.02 (Feb. 2017), P02019–P02019. DOI: `10.1088/1748-0221/12/02/P02019`.

[216] Bob Pease. *What's All This Transimpedance Amplifier Stuff, Anyhow?* 2001. URL: `https://www.electronicdesign.com/analog/whats-all-transimpedance-amplifier-stuff-anyhow-part-1`.

[217] Tony Wang and Barry Erhman. *Compensate Transimpedance Amplifiers Intuitively*. 2005. URL: `http://www.ti.com/lit/an/sboa055a/sboa055a.pdf`.

[218] Planet Analog. *Understand and apply the transimpedance amplifier (Part 1 of 2)*. 2007. URL: `https://www.planetanalog.com/document.asp?doc_id=527534&site=planetanalog`.

[219] Akshay Bhat. "Stabilized TIAs Key to Reliable Performance". *Electronic Design* (2011). URL: `https://www.electronicdesign.com/analog/stabilized-tias-key-reliable-performance`.

[220] Robert Keim. *Negative Feedback, Part 1: General Structure and Essential Concepts*. 2015. URL: `https://www.allaboutcircuits.com/technical-articles/negative-feedback-part-1-general-structure-and-essential-concepts/`.

[221] Marco D'Incecco et al. "Development of a Very Low-Noise Cryogenic Preamplifier for Large-Area SiPM Devices". *IEEE Transactions on Nuclear Science* 65.4 (Apr. 2018), pp. 1005–1011. DOI: `10.1109/TNS.2018.2799325`.

[222] Marco D'Incecco et al. "Development of a Novel Single-Channel, 24 cm^2, SiPM-Based, Cryogenic Photodetector". *IEEE Transactions on Nuclear Science* 65.1 (Jan. 2018), pp. 591–596. DOI: `10.1109/TNS.2017.2774779`.

[223] Zack Lasner and D. DeMille. "Statistical sensitivity of phase measurements via laser-induced fluorescence with optical cycling detection". *Physical Review A* 98.5 (Nov. 2018), p. 053823. DOI: `10.1103/PhysRevA.98.053823`.

[224] R. Wynands and S. Weyers. "Atomic fountain clocks". *Metrologia* 42.3 (June 2005), S64–S79. DOI: `10.1088/0026-1394/42/3/S08`.

[225] T. Zelevinsky, S. Kotochigova, and Jun Ye. "Precision Test of Mass-Ratio Variations with Lattice-Confined Ultracold Molecules". *Physical Review Letters* 100.4 (Jan. 2008), p. 043201. DOI: `10.1103/PhysRevLett.100.043201`.

[226] Alexander D. Cronin, Jörg Schmiedmayer, and David E. Pritchard. "Optics and interferometry with atoms and molecules". *Reviews of Modern Physics* 81.3 (July 2009), pp. 1051–1129. DOI: `10.1103/RevModPhys.81.1051`.

[227] Jack Alexander Devlin. "Progress towards a more sensitive measurement of the electron electric dipole moment with YbF". PhD thesis. Imperial College London, 2015. URL: https://spiral.imperial.ac.uk/handle/10044/1/28125.

[228] Mateusz Borkowski. "Optical Lattice Clocks with Weakly Bound Molecules". *Physical Review Letters* 120 (2018). DOI: 10.1103/PhysRevLett.120.083202.

[229] K. Beloy et al. "Effect of α variation on the vibrational spectrum of Sr_2". *Physical Review A* 84.6 (Dec. 2011), p. 062114. DOI: 10.1103/PhysRevA.84.062114.

[230] D. DeMille et al. "Enhanced Sensitivity to Variation of m_e/m_p in Molecular Spectra". *Physical Review Letters* 100.4 (Jan. 2008), p. 043202. DOI: 10.1103/PhysRevLett.100.043202.

[231] A. Shelkovnikov et al. "Stability of the Proton-to-Electron Mass Ratio". *Physical Review Letters* 100.15 (Apr. 2008), p. 150801. DOI: 10.1103/PhysRevLett.100.150801.

[232] Ivan Kozyryev, Zack Lasner, and John M. Doyle. "Enhanced Sensitivity to Ultralight Bosonic Dark Matter in the Spectra of the Linear Radical SrOH" (May 2018). arXiv: 1805.08185.

[233] D. DeMille. "Quantum Computation with Trapped Polar Molecules". *Physical Review Letters* 88.6 (Jan. 2002), p. 067901. DOI: 10.1103/PhysRevLett.88.067901.

[234] L. R. Liu et al. "Building one molecule from a reservoir of two atoms". *Science* 360.6391 (May 2018), pp. 900–903. DOI: 10.1126/science.aar7797.

[235] A. Micheli, G. K. Brennen, and P. Zoller. "A toolbox for lattice-spin models with polar molecules". *Nature Physics* 2.5 (May 2006), pp. 341–347. DOI: 10.1038/nphys287.

[236] Bhuvanesh Sundar, Bryce Gadway, and Kaden R. A. Hazzard. "Synthetic dimensions in ultracold polar molecules". *Scientific Reports* 8.1 (Dec. 2018), p. 3422. DOI: 10.1038/s41598-018-21699-x.

[237] M. L. Wall, K. Maeda, and Lincoln D. Carr. "Realizing unconventional quantum magnetism with symmetric top molecules". *New Journal of Physics* 17.2 (Feb. 2015), p. 025001. DOI: 10.1088/1367-2630/17/2/025001.

[238] J. F. Barry et al. "Magneto-optical trapping of a diatomic molecule". *Nature* 512.7514 (Aug. 2014), pp. 286–289. DOI: 10.1038/nature13634.

[239] Loïc Anderegg et al. "Radio Frequency Magneto-Optical Trapping of CaF with High Density". *Physical Review Letters* 119.10 (Sept. 2017), p. 103201. DOI: 10.1103/PhysRevLett.119.103201.

[240] H. J. Williams et al. "Magnetic Trapping and Coherent Control of Laser-Cooled Molecules". *Physical Review Letters* 120.16 (Apr. 2018), p. 163201. DOI: 10.1103/PhysRevLett.120.163201.

[241] E. Rocco et al. "Fluorescence detection at the atom shot noise limit for atom interferometry". *New Journal of Physics* 16.9 (Sept. 2014), p. 093046. DOI: 10.1088/1367-2630/16/9/093046.

[242] F. Thomas Bruss and James B. Robertson. "'Wald's Lemma' for sums of order statistics of i.i.d. random variables". *Advances in Applied Probability* 23.03 (Sept. 1991), pp. 612–623. DOI: 10.2307/1427625.

[243] Abraham Wald. *Sequential Analysis*. 1st ed. New York: John Wiley and Sons, 1947.

[244] Joseph K. Blitzstein and Jessica Hwang. *Introduction to probability*. 1st ed. Boca Raton: CRC Press, 2015. ISBN: 9781466575578.

[245] S. Weyers et al. "Uncertainty evaluation of the atomic caesium fountain CSF1 of the PTB". *Metrologia* 38.4 (Aug. 2001), pp. 343–352. DOI: 10.1088/0026-1394/38/4/7.

[246] S. R. Jefferts et al. "Accuracy evaluation of NIST-F1". *Metrologia* 39.4 (Aug. 2002), pp. 321–336. DOI: 10.1088/0026-1394/39/4/1.

[247] T. Kurosu et al. "Preliminary Evaluation of the Cs Atomic Fountain Frequency Standard at NMIJ/AIST". *IEEE Transactions on Instrumentation and Measurement* 53.2 (Apr. 2004), pp. 466–471. DOI: 10.1109/TIM.2004.823313.

[248] F. Levi et al. "IEN-CsF1 Accuracy Evaluation and Two-Way Frequency Comparison". *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control* 51.10 (Oct. 2004), pp. 1216–1224. DOI: `10.1109/TUFFC.2004.1350948`.

[249] K. Szymaniec et al. "Evaluation of the primary frequency standard NPL-CsF1". *Metrologia* 42.1 (Feb. 2005), pp. 49–57. DOI: `10.1088/0026-1394/42/1/007`.

[250] Philip Bevington. *Data reduction and error analysis for the physical sciences.* 1st ed. New York: McGraw-Hill, 1969.

[251] C. Radhakrishna Rao. "Information and the Accuracy Attainable in the Estimation of Statistical Parameters". *Breakthroughs in Statistics. Springer Series in Statistics (Perspectives in Statistics).* Ed. by Samuel Kotz and Norman L. Johnson. New York: Springer, 1992, pp. 235–247. DOI: `10.1007/978-1-4612-0919-5_16`.

[252] Glenn Knoll. *Radiation Detection and Measurement.* 4th ed. Hoboken: John Wiley and Sons, 2010.

[253] Cristian Panda. *I-H state data.* 2014. URL: `https://lablogs.gabrielse.physics.northwestern.edu/lablog_graduate_2013-Chris_Panda/2014/09/i-h-state-data.html`.

[254] William J. Thompson. *Angular Momentum.* 1st ed. New York: John Wiley and Sons, 1994.

[255] D. A. Varshalovich, A. N. Moskalev, and V. K. Khersonskii. *Quantum Theory of Angular Momentum.* Singapore: World Scientific, Oct. 1988. ISBN: 978-9971-5-0107-5. DOI: `10.1142/0270`.

[256] Leon van Dommeln. *Quantum Mechanics for Engineers.* 2018. URL: `http://www.eng.fsu.edu/~dommelen/quantum/style_a/contents.html`.

[257] Cristian Panda. *Signals in the Interaction Region - Detected photon number check and Doppler widths.* 2015. URL: `https://lablogs.gabrielse.physics.northwestern.edu/lablog_EDMexperiment/2015/08/signals-in-the-interaction-region-photon-number-check-and-doppler-widths.html`.

[258] Elizabeth Petrik West. *Beamline Reconfiguration Calculations*. 2015. URL: `https://lablogs.gabrielse.physics.northwestern.edu/lablog_EDMgraduate_2009-Elizabeth_Petrik/2015/04/beamline-reconfiguration-calculations.html`.

[259] Cristian Panda. *Simulations of signal gain by increasing collimator spacing and decreasing beamline length in GEN II*. 2015. URL: `https://lablogs.gabrielse.physics.northwestern.edu/lablog_graduate_2013-Chris_Panda/2015/04/simulations-of-signal-gain-by-increasing-collimator-spacing-and-decreasing-beamline-length-in-gen-ii.html`.

[260] D. J. Berkeland and M. G. Boshier. "Destabilization of dark states and optical spectroscopy in Zeeman-degenerate atomic systems". *Physical Review A* 65.3 (Feb. 2002), p. 033413. DOI: `10.1103/PhysRevA.65.033413`.

[261] Stephan P. A. Sauer. "A relation between the rotational g-factor and the electric dipole moment of a diatomic molecule". *Chemical Physics Letters* 297.5-6 (Dec. 1998), pp. 475–483. DOI: `10.1016/S0009-2614(98)01157-9`.

[262] John Preskill. *Caltech Physics 219 Lecture Notes*. 2018. URL: `http://www.theory.caltech.edu/people/preskill/ph229/`.

[263] William H. Greene. *Econometric Analysis*. 5th ed. Upper Saddle River: Prentice Hall, 2003.

[264] Wulfram Gerstner and Werner M. Kistler. *Spiking neuron models: single neurons, populations, plasticity*. Cambridge University Press, 2002. ISBN: 0521813840.

[265] K. M. Van Vliet and L. M. Rucker. "Noise associated with reduction, multiplication and branching processes". *Physica A: Statistical Mechanics and its Applications* 95.1 (Jan. 1979), pp. 117–140. DOI: `10.1016/0378-4371(79)90046-3`.

[266] Kyle Siegrist. *Thinning and Superposition*. URL: `http://www.randomservices.org/random/poisson/Splitting.html`.

[267] R. J. McIntyre. "Multiplication noise in uniform avalanche diodes". *IEEE Transactions on Electron Devices* ED-13.1 (Jan. 1966), pp. 164–168. DOI: `10.1109/T-ED.1966.15651`.