ABSTRACT

Title of Dissertation: Spin Dynamics,
Pulsed Decoupling,
and Deep Reinforcement Learning
for Quantum Sensing

Jner Tzern Oon
Doctor of Philosophy, Physics, 2025

Dissertation Directed by: Professor Ronald Walsworth
Department of Physics

Quantum sensing faces challenges in moving from proof-of-concept demonstrations to practical technologies, with discrepancies between idealized theoretical frameworks and experimental realities presenting a key obstacle. This dissertation explores these nuances — with a focus on nitrogen-vacancy (NV) centers in diamond — through four projects spanning experiment, theory, simulations, and algorithmic optimization.

We characterize Ramsey envelope modulation effects in ^{15}NV diamond magnetometry, showing that magnetic field misalignments produce envelope effects that degrade sensitivity. Next, we study the breakdown of Average Hamiltonian Theory (AHT) in experimental regimes, introduce exact methods to calculate a sensor response to a target signal that are valid beyond AHT, and establish symmetries that guarantee AHT convergence. With Ensemble Cluster Sampling (ECS), we address overfitting in dynamical decoupling by training algorithms on heterogeneous parameter distributions rather than idealized systems. Finally, we present TEMPO, an

open-source Python package for accessible pulse sequence simulations.

These studies underline the need for continued collaboration between quantum sensing theory and laboratory applications, while maintaining an eye on modern algorithms and software.

Spin Dynamics, Pulsed Decoupling, and Deep Reinforcement Learning for Quantum Sensing

by

Jner Tzern Oon

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2025

Advisory Committee:
Professor Ronald Walsworth, Chair/Advisor
Professor John Cumings, Dean's Representative
Professor Yanne Chembo
Professor Victor Yakovenko
Professor Ian Spielman

© Copyright by
Jner Tzern Oon
2025

Preface

This dissertation represents the 4 projects that I led during my graduate studies, dabbling between quantum sensing experiment, simulation, and theory. Along the way, we also contributed original software for publication and explored deep reinforcement learning for quantum control.

Here I provide a brief timeline of those projects.

I began my PhD working on NV-diamond experiments for magnetic field sensing. In this publication (link), we found that Ramsey magnetometry experiments exhibited envelope modulation effects that negatively affected magnetic field sensitivity. We developed a model that predicts the properties of these modulations (based on perturbation theory), and validated these predictions experimentally across a range of magnetic fields.

Next, I began studying how concepts in nuclear magnetic resonance are used to improve pulse sequences, for both dynamical decoupling and quantum sensing. After studying a theoretical approach (Average Hamiltonian Theory or AHT) that was seminal for development of state-of-the-art sensing protocols, we discovered that AHT produces inaccurate results in experimental regimes where it is often applied. In this work (link), we documented not only the breakdown of AHT in sensing regimes, but also presented 1) conditions for AHT to converge (thus retaining its ability for accurate predictions of quantum dynamics), and 2) an exact alternative method for calculating a dynamical decoupling sequence's impact on sensitivity.

My PhD comes to an end with my longest-running project, based on applying deep reinforcement learning (RL) to synthesize dynamical decoupling protocols. This includes the development of robust environments for algorithm training, and assessments of RL methods for the

goal of preserving quantum states.

These works were accompanied by internal development of software for matrix-based quantum evolution, resulting in TEMPO: an open-source Python package for simulations of pulse sequence dynamics (see links to Arxiv, Documentation, and Github). TEMPO provides two main benefits to existing approaches. First, it provides significant speedups in simulation (wall-clock) times compared to naive implementations. Second, the package is structured to be highly accessible, including for experimentalists with existing familiarity with pulse generator hardware and arbitrary waveform generators.

I believe that academic research is driven by collaboration, and I deeply value mentorship. During my PhD, I've had the privilege of mentoring both undergraduate and graduate students, one-on-one and in small collaborations - they were essential to any success I've had in graduate school. I am proud to have "equally contributed to" multiple manuscripts alongside some of them, though their impact on my life is beyond measure.

Dedication

To my parents, who spared no sweat as they placed brick after brick.

Acknowledgments

Your attention is a commodity – I don't plan to waste it.

This period taught me that happiness sits at the intersection of 3 things:

radical acceptance,

recognition of suffering,

and an internalization of unity.

You might also call them **love, compassion, and family.**

These insights didn't come from any one religion or philosophy, but through repeated questioning of fundamental connectors that link us all – precisely what physics was meant to study.

Take, for example, the claim that "life is suffering" – this sentiment can both inspire and depress. On one hand, this reality is shared by all, representing the ultimate case study for humanity's ability to join as one. Yet it presents a sharp reminder that we as a species continue to impose unnecessary pain on our fellow people.

A few realizations emerge from this dual perspective.

The **practice of love** has been relegated to romantic exhibition.

The **art of compassion** is missing from our educational system.

The **concept of family** is limited to a handful.

And oppression is, unfortunately, the most adopted human language.

As the page flips to a new chapter, I'd like to pen some thank yous on the blank canvas. It stings me to know that words can never fully capture the many who have touched my life, and made me part of theirs.

Despite that, **I will try.**

My father Oon Ee Heng, your dedication leaves no doubt that perseverance is in my blood.

Thank you for raising me, so I know to stand.

My mother Lim Yit Lee, your sacrifices are silent, but your care is heard across oceans.

Thank you for my wings, so I know to soar.

My siblings—Oon Lin Chern, Oon Yuen Tzern, Oon Qian Chern, and Oon Ming Tzern—you ensured that my childhood is one born of love.

Thank you for building the nest, so I know home.

My brother Reza Ebadi, your awareness of life's impermanence hones my clarity of purpose.

Thank you for inching closer to the end with me.

My guru Victoria Kovalchuk, you help us maintain the question "why" with unrelenting passion.

Thank you for walking the path with me.

My mentor Sebastian Carrasco, perspective is your wand, and compassion is your Patronus.

Thank you for lending me your shoulder and your ear.

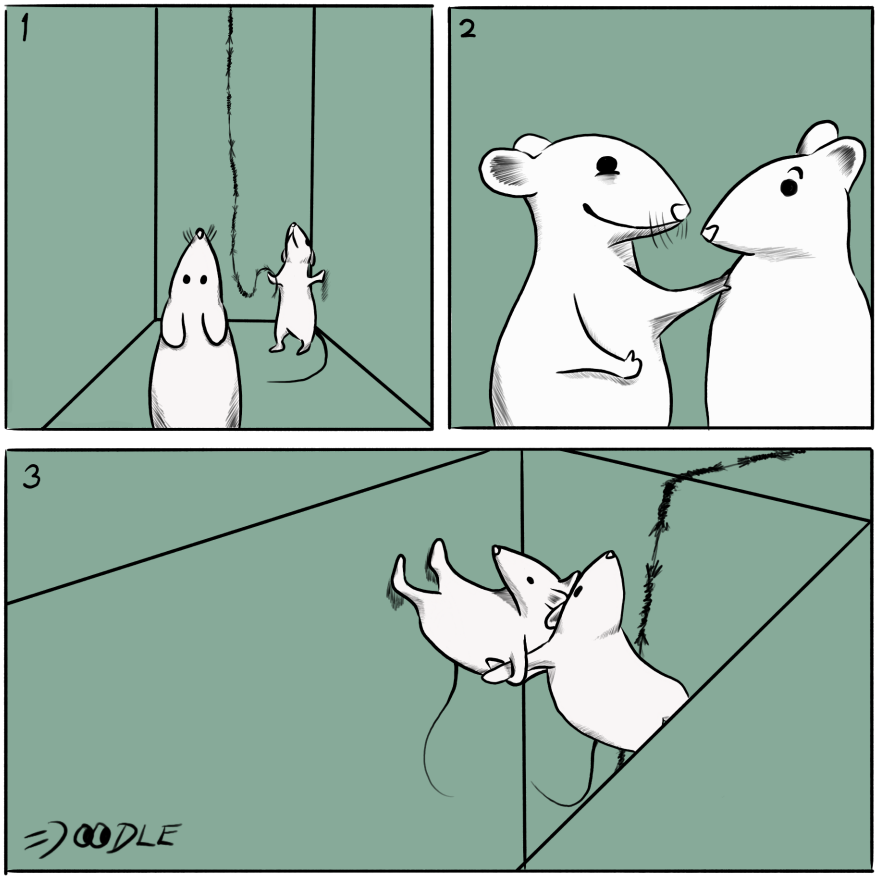
My champion Sohyoun Yun-Cárcamo, the ferocity of your care can move nations.

Thank you for compelling me forward.

My second home the Alexanders, your living room radiates with childhood laughter.

Thank you for loving thy neighbor.

My friend Paige Kunkle, who I miss dearly.



My PhD traveled like a wave – droplets gathered in Boston, encountered a pandemic storm, and landed in College Park. I still remember the day Ron adopted me into his group. It was shortly after I was turned down by the professor I'd entered graduate school to work with, and a few summer months into my time with Ron's lab. I had just finished my presentation at our weekly group meeting, in a dimly lit conference room within the CfA building. My mentors—Connor Hart and Matthew Turner—sat with me around a large table, after the other group members left the room. The conversation started with a question from Ron:

"Are you having fun?"

And the rest is history.

During my 6 years in Ron's group, I had no shortage of good people I could rely upon for sound mentorship. My initial mentors Connor and Matthew were, simply put, excellent scientists. Their attention to detail and passion for the development of quantum sensing impressed upon me the sheer creativity, technical precision, and fun that could come with research. Better yet, this rare combination came bundled with humans who truly care to uplift others. I also felt a welcoming spirit from Jennifer Schloss, Aakash Ravi, and other members of Ron's group at Harvard.

Our move to UMD in 2020 came alongside a disease that shook the global definition of work, and forced many to inspect their priorities. Here I met a new flock of graduate students, colleagues, and friends who shared this new reality. I remember the gatherings at my apartment for food and board games, the long conversations in lab that drifted beyond academia, and the uncertainties we shared for the future. Here I name the people who helped shape those years.

The sharp Jiashen, I remember fond conversations and our raised glasses in Seattle.

The warm Zechuan, the first student I personally mentored and supervised.

The considerate Sai, whose meticulous nature personifies a careful physicist-practitioner.

The philosopher Jeyson, who lives by the message of gratitude.

The blossoming George, whose humanity outshines his Witt.

The ever-welcoming Chen, a buddy who shared more than an optical table.

The empathetic Sonja, who reminds me the joy of collaboration.

The curious Niko, who sits straight even when life comes with weights.

The list continues with people like Johannes, Kristine, Kevin, Shantam, and more recently, Katrijn, Justin, Andrew, Smriti, and other bright new faces of Ron's group.

To all of you, **I wish you happiness.**

If I can ever be helpful—as a scientist, mentor, friend, or listener—don't hesitate to call.

Love is an act of rebellion.

It sets fire to the ropes of division knotted into our upbringing, culture, and classrooms. It's the one herb we have to treat the plague of forced segregation.

I've practiced love every day for the past 10 years.

This was not required for my PhD. Yet it forms the pillar to my success. A foundational brick.

And thanks to my partner Lakshmi Kumar, I was not alone in these daily drills.

We inched towards *doctorates* for seven years in split states. We wished each other goodnight before cutting the call. We kissed goodbye at 4 am before leaving for planes and trains.

But a commitment to **love** meant recognizing the things we share – the things that **unite** us.

We shared fears for our relationships, our careers, and floating homes in a foreign country.

We spent 10 years growing a **family** bound by more than blood.

Because even *experts* have questions, even *professors* cry, and even *doctors* feel pain.

Love is blind to titles, race, and color. Which makes one observation forever baffling.

How beautiful she is to me.

Table of Contents

Preface ii

Dedication iv

Acknowledgements v

Table of Contents x

List of Tables xii

List of Figures xiii

List of Abbreviations xix

Introduction 1

Chapter 1: Ramsey Envelope Modulation in NV Diamond Magnetometry 6
- 1.1 Introduction . . . 6
- 1.2 Electron Ramsey Envelope Modulation (EREEM) . . . 10
- 1.3 Experimental Methods . . . 15
- 1.4 Measured EREEM Properties . . . 16
- 1.5 Modulation Amplitude and Impact on Magnetometry . . . 20
- 1.6 Double-quantum Ramsey . . . 23
- 1.7 Conclusion . . . 25
- 1.8 Supplementary Notes . . . 27
  - 1.8.1 Theoretical Details . . . 27
  - 1.8.2 Experimental Details . . . 31
  - 1.8.3 Bias Field Misalignment Angle . . . 32
  - 1.8.4 Microwave Frequency Calibration . . . 34
  - 1.8.5 Measured Ramsey Time Series Fit Details . . . 35
  - 1.8.6 Ramsey Measurements Across Misalignment Angles . . . 37
  - 1.8.7 QuTiP Simulations . . . 38
  - 1.8.8 Normalized Inverse Sensitivity . . . 41
  - 1.8.9 Transverse Hyperfine Parameter . . . 41
  - 1.8.10 Double-quantum Envelope Amplitude . . . 43

Chapter 2: Beyond Average Hamiltonian Theory for Quantum Sensing 45

2.1 Introduction . . . 46
2.2 Average Hamiltonian Theory . . . 47
2.3 AHT Breakdown . . . 48
2.4 Magnus Expansion Convergence . . . 51
2.5 Exact Sensor Coupling Factor . . . 54
2.6 Retaining Convergence with Rapid Echoes . . . 56
2.7 Outlook . . . 59
2.8 Supplementary Notes . . . 61
2.8.1 Magnus Expansion . . . 61
2.8.2 Discrete Time . . . 62
2.8.3 Magnus expansion converges for rapid echoes . . . 64
2.8.4 Numerical simulation with random echo sequences . . . 65

Chapter 3: Ensemble Cluster Sampling for Heterogeneous Dynamical Decoupling 67
3.1 Introduction . . . 67
3.2 Ensemble Cluster Sampling (ECS) . . . 69
3.3 Heterogeneous Dynamical Decoupling . . . 72
3.4 Ensemble Dephasing Timescale . . . 77
3.5 Supplementary: Genetic Algorithm for Dynamical Decoupling Sequence Search . 81
3.5.1 GA Results . . . 86
3.5.2 Optimizing GA parameters . . . 88
3.5.3 Deviations from Hamiltonian parameters . . . 89
3.5.4 Outlook . . . 91

Chapter 4: TEMPO: A Python Package for Time Evolution of Pulse Sequences in QuTiP 94
4.1 Summary . . . 94
4.2 Statement of Need . . . 95
4.3 Usage . . . 96
4.4 Efficient Simulation Times . . . 97

Conclusion 100

List of Tables

List of Figures

1.1 (a) Experimental setup schematic. Diamond containing an ensemble (≈ 0.3 ppm) of ^{15}NV centers is probed using 532 nm green laser excitation through a microscope objective, with red fluorescence collected back along the same path. A gold Ω-shaped waveguide is placed above the NV layer for microwave (MW) delivery. A pair of ring magnets apply a bias magnetic field at two distinct orientations, aligned and misaligned with the symmetry axis for one class of NV centers. Top right inset: diamond crystal axes and bias field orientations relative to a single NV quantization axis. Bottom right inset: NV electronic energy level diagram of ground and excited states. (b) Example Ramsey time series measured for bias fields aligned (gray) and misaligned (blue) to one class of NV centers. Effect of electron Ramsey envelope modulation (EREEM) is clearly observable for data with the misaligned field. Upper inset: Ramsey pulse sequence diagram. 7

1.2 (a) Energy level diagram for the ^{15}NV ground state, with splittings due to Zeeman and hyperfine interactions assuming an aligned bias field. The frequency of an applied microwave field ν_{+1} is centered between the two $m_s = 0 \leftrightarrow m_s = +1$ electronic resonances split by the NV hyperfine interaction. (b) Effective magnetic field vectors on the ^{15}N nuclear spin due to a misaligned bias magnetic field, for each electronic spin state $|m_s\rangle$. The nuclear spin experiences an effective magnetic field consisting of an electronic-spin-dependent contribution $\vec{\beta}(m_s)$, and a component $\vec{\beta}_{\text{ind}}$ that is independent of the electronic spin state. (c) Example ^{15}NV Ramsey free induction signal due to a bias magnetic field of magnitude $B = 100$ G at a misalignment of 22°, showing envelope modulation. The amplitude modulation $\chi(\tau)$ is indicated, showing oscillations between $\chi_{\text{min}} = \cos(\Phi_{0,+1})$ and 1, with a characteristic period $2\pi/\omega_0$. (d) Amplitude spectrum of the Ramsey signal from (c), revealing two peaks split by the envelope beat frequency $\omega_0 = 2\pi \times 254$ kHz. (e) Effective magnetic field vectors on the ^{14}N nuclear spin. For all three electronic spin states, the corresponding effective fields on the nuclear spin are nearly parallel due to a dominant nuclear quadrupolar field. 13

1.3 Ramsey experimental data and envelope properties from model fits. (a) Example Ramsey time series at field magnitude $B = 90.08(2)$ G and misalignment angle $\theta = 16.79(9)^\circ$. Data is shown in black and the fit in blue. The observed slow beating is an example of EREEM. (b) Estimates of envelope beat frequency ω_0 extracted from fits to experimental data, collected across a range of magnetic field magnitudes and orientations. Theoretical predictions for ω_0 are shown in solid lines for each magnetic field. After allowing the transverse hyperfine constant $A_\perp$ to vary as a fitting parameter, updated ω_0 predictions at each field are shown with dashed lines. (c) Estimates of the relative contrast at amplitude modulation nodes, χ_{min}, for the magnetic field configurations used in (b). Solid lines indicate theoretical predictions with $A_\perp$ fixed. Inset: Magnified view of χ_{min} values for misalignment angles $20^\circ < \theta < 40^\circ$. 17

1.4 Calculations of expected Ramsey magnetic field sensitivity with and without envelope modulation (EREEM), and at different bias magnetic field configurations. (a) Inverse sensitivity η^{-1} vs Ramsey free evolution time τ without EREEM, given by Eq. (1.9) for fixed dead time $T_D = 5\,\mu$s and dephasing time $T_2^* = 5\,\mu$s. The optimal inverse sensitivity η_{opt}^{-1} is indicated, corresponding to a free evolution time τ_{opt} indicated in all subfigures with a vertical dashed line. (b) Relative inverse sensitivity with EREEM $\eta_{\text{opt}}/\widetilde{\eta}$, normalized to the optimal sensitivity and free evolution time established in (a), for two distinct magnetic field misalignment angles $\theta = 10^\circ, 20^\circ$ at the fixed magnitude $B = 100\,$G. Adjusted optimal evolution times $\widetilde{\tau}_{\text{opt}}$ are indicated by symbols $\diamond$. (c) 2D color plot of the normalized inverse sensitivity at a fixed misalignment angle $\theta = 10^\circ$, across bias field magnitudes $0 < B < 200\,$G. 21

1.5 Comparison between experimental single-quantum (SQ) and double-quantum (DQ) Ramsey measurements under a misaligned bias magnetic field. Representative SQ and DQ Ramsey data collected at a bias field magnitude $B \approx 50\,$G and misalignment angle $\theta \approx 35^\circ$, in both time (a) and frequency (b) domains. The inset in (b) shows the effective magnetic fields $\vec{\beta}_{\text{ind}} + \vec{\beta}(\pm 1)$ coupled to the nuclear spin, due to the electronic spin states $|m_s = \pm 1\rangle$ participating in the DQ protocol. The effective fields are nearly antiparallel ($\Phi_{-1,+1} \approx \pi$), resulting in no observable envelope modulation in the DQ Ramsey signal ($\chi_{\text{min}} = |\cos \Phi_{-1,+1}| \approx 1$). (c) Calculation of relative contrast at amplitude modulation nodes χ_{min} for both SQ and DQ Ramsey signals, over a range of bias magnetic field magnitudes and misalignment angles. For SQ Ramsey, calculated values of $\chi_{\text{min}} = |\cos \Phi_{0,+1}| = |\cos \phi_{+1}|$ are shown, obtained using Eq. (1.5). For DQ Ramsey, $\chi_{\text{min}} = |\cos \Phi_{-1,+1}| = |\cos(\phi_{+1} - \phi_{-1})|$. 26

1.6 Transition frequency difference $\Delta_{\pm 1}$ between $m_s = 0 \leftrightarrow +1$ and $m_s = 0 \leftrightarrow -1$ resonances as a function of misalignment angle. Calculation is performed at 90 G. (a) Numerical results of $\Delta_{\pm 1}$ using the full NV ground state Hamiltonian (solid red line) and the approximate formula from perturbation theory $2\gamma_e B \cos\theta$ (dashed blue line). (b) Absolute and percentage deviation of $\Delta_{\pm 1}$ between the two calculation methods. 33

1.7 Example bootstrap fit result for the EREEM beat frequency. Starting with a single Ramsey time series, $10,000$ additional datasets are created by randomly sampling with replacement from the initial time series. Estimates of the EREEM beat frequency ω_0 are extracted using the MATLAB `lsqcurvefit` function for each dataset. The distribution supports the assumption of normality used by the standard interval method for confidence interval (CI) construction. 36
1.8 Experimental Ramsey time series data obtained using a purposefully detuned microwave driving frequency, with its corresponding amplitude spectrum displayed as an inset. The larger detuning enables clear separation among all frequency components. The fit result for the EREEM beat frequency $\omega_0/2\pi$ is 273.3 ± 1.4 MHz, which overlaps with the confidence interval (CI) of the corresponding beat frequency fit result, 273.6 ± 0.8 kHz, obtained when the microwave frequency is centered between the hyperfine splitting. 37
1.9 Experimental single-quantum Ramsey time series measurements (left column) and corresponding amplitude spectra (right column) at various bias field misalignment angles θ, at a bias field magnitude of $90.08(2)$ G. 38
1.10 Calculation and comparison of Ramsey envelope (EREEM) beat frequency ω_0 as a function of bias magnetic field magnitude and misalignment angle, using analytical approximation and QuTiP numerical simulations. (a) ω_0 obtained using numerical simulations (circles) and analytical perturbation theory (solid line), for magnetic fields between 40 G and 140 G. (b) Absolute differences between ω_0 estimates from QuTiP numerical simulations and analytical approximation, for the field configurations shown in (a). (c) Percentage deviation between ω_0 estimates from QuTiP numerical simulations and analytical approximation. 39
1.11 2D color plots of the normalized inverse sensitivity with respect to the Ramsey free evolution time τ and bias magnetic field magnitude B, at 2 distinct misalignment angles $\theta = 20°$ (a) and $54.7°$ (b) 41
1.12 Results for the transverse hyperfine parameter $A_\perp$ as a fit parameter to the Ramsey envelope (EREEM) beat frequency ω_0, extracted from experimental data, for a range of magnetic fields, with error bars showing the 95% confidence interval, shown in black. Estimated corrections to values of $A_\perp$ are depicted in red, using the discrepancy between QuTiP numerical simulations and theoretical predictions of ω_0. 42
1.13 Calculated double-quantum (DQ) Ramsey relative contrast at amplitude modulation nodes $\chi_{\text{min}} = |\cos(\Phi_{-1,+1})|$, for magnetic field magnitudes B up to 800 G and misalignment angles up to $90°$. Dashed contour lines indicate magnetic field configurations where χ_{min} equals 0.9 and 0.5, representing 10% and 50% contrast loss at envelope nodes, respectively. 43

2.1 Fig 1 (a) WAHUHA pulse sequence and toggling frame representations. (b) Effective field corresponding to leading order average Hamiltonian. (c) WAHUHA time series $\langle \sigma_x(t) \rangle$ and corresponding power spectra, obtained using the first order average Hamiltonian (gray) vs ground truth unitary evolution (red). 49

2.2 Fig 2 (a) State fidelity between exact and AHT evolution for varying sequence duration t, where $\mathcal{U}^m(t)$ is calculated for $m = 1, 3$ and compared to the ground truth unitary $U_0(t)$. (b) Spectral norm of Magnus expansion terms $|\bar{H}^m|$ for $m = 1, 3, 5, 7$ for the WAHUHA protocol, for varying sequence duration t. The region of guaranteed Magnus series convergence $t < r_c/\pi$ is shaded in gray. 52
2.3 (a) An initial state ψ is transformed by a pulse sequence unitary propagator to $U_0\psi$. In the presence of a perturbation field V (i.e., a target signal field to be measured), the final state is $U_1\psi$. The operator Ω generates a unitary operator $U_\epsilon = e^{-i\Omega t}$ such that $U_1 = U_\epsilon U_0$. (b) Exact coupling factor α_ϵ for the WAHUHA pulse sequence, as sequence duration t is increased. 56
2.4 (a) Square wave perturbation with amplitude ϵ, changing sign with each toggling frame. Each pulse P_i is applied to the quantum sensor (e.g., NV electronic spins in diamond) before the i-th frame. For an initial sensor spin Hamiltonian S_z, the resulting toggling frame Hamiltonians for pulse sequences XY8, WAHUHA+Echo and DROID-60 are shown. (b) For each sequence, the exact coupling factor α_ϵ is plotted as a function of sequence duration t. Calculations for XY8 and DROID-60 agree with predictions obtained from first-order AHT, i.e., $\bar{H}^{(1)}$. 58

3.1 (a) Conceptual illustration of ensemble cluster sampling (ECS), where n_C clusters are drawn from distributions of Hamiltonian parameters.(b) Distributions of nearest-neighbor diamond lattice spacing d and corresponding coupling strengths J, for defect concentration $C = 2$ ppm. (c) Average variance of the propagator under a Ramsey sequence as a function of the number of clusters N_{MC}, averaged over 5000 samples per point. The inset shows example variance distributions. Saturation approaching $N_{MC} = 100$ sets the minimum cluster count for simulation use. 70
3.2 (a) Schematic of deep reinforcement learning approach using proximal policy optimization (PPO). The policy network $\pi_a(s_t)$ decides the pulse action a_t at integer step t, which is applied to the ECS environment followed by a free evolution period τ. The resulting state s_{t+1} is fed back to the network to decide the next action. Network weights are updated periodically using the ensemble reward $\mathcal{R}$. (b) Decoupling score histograms for best-performing RL PPO sequences from point-sample training (light blue) and ECS training (dark blue). (c) The beam search algorithm starts with a random seed sequence—a beam—and generates all neighbor sequences with up to k pulse changes. The top $n_B = 100$ pulse sequences are selected as new beams to repeat the search. (d) Histograms for best-performing beam search sequences for point-sample (green) and ECS training (dark green). Reference sequence Angle-12 from Peng et al.[1] is plotted in both (b) and (d) as black histograms. 73
3.3 (a) Ramsey decay curves show ensemble sz projection decreasing with free evolution time τ, for varying defect concentration C with central spin simulations. (b) Fitted decay timescales are plotted as a function of C, on a log-log plot. A reference line from Bauch et al. is shown in black. 78

3.4 Statistics for genetic algorithm repeated with 50 seeds. **Left:** Best fitness values $\bar{\mathcal{F}}_{best}(n_G)$ averaged over 50 seeds, as a function of generation number n_G. **Right:** Best fitness value across all 50 seeds, plotted at each generation. Black dashed line shows fitness of sequence B under these conditions. Hamiltonian parameters: $N = 3,\ J_{int} = 1$ kHz, $\tau = 2\mu s,\ \alpha = 10$ kHz, **+5% pulse rotation error.** GA parameters: $N_{\text{pop}} = 501,\ N_G = 4000,\ N_T = 3, \mathbf{P}_C = 0.5,\ p_C = 0.1,\ \mathbf{P}_M = 0.5,\ p_M = 0.06$ 87
3.5 Statistics for genetic algorithm repeated with 50 seeds. **Left:** Best fitness values $\bar{\mathcal{F}}_{best}(n_G)$ averaged over 50 seeds, as a function of generation number n_G. **Right:** Best fitness value across all 50 seeds, plotted at each generation. Black dashed line shows fitness of sequence B under these conditions. Hamiltonian parameters: $N = 3,\ J_{int} = 1$ kHz, $\tau = 2\mu s,\ \alpha = 10$ kHz, **no pulse angle error**. GA parameters: $N_{\text{pop}} = 101,\ N_G = 4000,\ N_T = 8, \mathbf{P}_C = 0.5,\ p_C = 0.1,\ \mathbf{P}_M = 0.5,\ p_M = 0.06$ 88
3.6 Fitness graphs for genetic algorithm executions with individual pulse mutation probability p_M (denoted as P_MUTATE_INDEP) varied between 0.01 and 0.15, each repeated over 50 seeds. **Left:** Each trace shows the best fitness values $\bar{\mathcal{F}}_{best}(n_G)$ averaged over 50 seeds, as a function of generation number n_G. **Right:** Best average fitness values at the last generation (no. 400). Standard deviation across all 50 repetitions are shown as error bars. Black dashed line shows fitness of sequence B under these conditions. Hamiltonian parameters: $N = 3,\ J_{int} = 1$ kHz, $\tau = 2\mu s,\ \alpha = 10$ kHz, +5% pulse rotation error. GA parameters: $N_{\text{pop}} = 101,\ N_G = 400,\ N_T = 3, \mathbf{P}_C = 0.5,\ p_C = 0.1,\ \mathbf{P}_M = 0.5,\ p_M = 0.06$. 89
3.7 For the pulse sequence with the highest fitness from Fig. 3.4 (right), plots show fitness when Hamiltonian parameters are changed in the presence of +5% pulse error. Top left: α and τ are varied. Top right: J_{int} and τ are varied. Bottom left: Rotation angle error and α are varied. Bottom middle: Rotation angle error and J_{int} are varied. Bottom right: Rotation angle error and τ are varied. White dotted line indicates parameters used for GA execution and pulse sequence search. Green dashed lines mark fitness contours for this pulse sequence. 90
3.8 Fitness of DROID-60 sequence over a range of Hamiltonian parameters, in the presence of pulse error +5%. Left: α and τ are varied. Right: J_{int} and τ are varied. White dotted line indicates parameters used for GA execution and pulse sequence search 91
3.9 Improvement of pulse sequence fitness relative to DROID-60 sequence, over a range of Hamiltonian parameters, in the presence of pulse error +5%. Top left: α and τ are varied. Top right: J_{int} and τ are varied. Bottom left: Rotation angle error and α are varied. Bottom middle: Rotation angle error and J_{int} are varied. Bottom right: Rotation angle error and τ are varied. White dotted line indicates parameters used for GA execution and pulse sequence search. Pink dashed lines mark the fitness contour for DROID-60 (no improvement). Minimum colorscale value is manually capped at 0, which indicates no fitness improvement. 92
3.10 Same data as in Fig. 3.9 but with the colorscale edited so the minimum fitness improvement is not capped at 0. 92

4.1 Outline of steps in TEMPO used to perform time evolution due to a pulse sequence. First, each pulse recipe contains the functional form of a time-dependent Hamiltonian $H_i(t) = \hat{O}_i \cdot f_i(\mathbf{x}, t)$. Pulse recipe(s) are used to create individual pulse(s) $P_1, \ldots, P_n$, with individual parameter settings $\mathbf{x_1}, \ldots, \mathbf{x_n}$, respectively. The pulses are organized into a sequence along with an optional static Hamiltonian. Next, the pulse sequence is provided to the Evolver, with an initial system state $\psi(t_0)$ and an array of evaluation times $(t_0, t_1, \ldots, t_m)$. Time evolution returns the state $\psi(t_0), \ldots, \psi(t_m)$ at these times; if operators $E_0, \ldots, E_j$ are provided, the operator expectation values $E_0(t_0), \ldots, E_j(t_0)$ are calculated instead. 96

4.2 (a) Average simulation wall-clock time T_S as a function of pulse sequence duration τ without using TEMPO, for number of pulses $n = 5, 10, 20, 50$, shown in shades of blue. Average simulation times using TEMPO are shown in red at high pulse number count $(n = 50)$. Best-fit curves are displayed as dashed lines, with slope values plotted in (b) for measurements with TEMPO in red and without TEMPO (via QuTiP's mesolve function) in black. (c) Example visualization of pulse sequence segmentation performed by TEMPO, where $n = 3$ pulses result in $2n - 1 = 5$ time segments for solver evaluation. All measurements of T_S are average values over 20 repeated runs. 98

List of Abbreviations

NV	Nitrogen-Vacancy Center
NV^-	Nitrogen-Vacancy Center - negatively charged
AHT	Average Hamiltonian Theory
EREEM	Electron Ramsey Envelope Modulation
ESEEM	Electron Spin Echo Envelope Modulation
ECS	Ensemble Cluster Sampling
TEMPO	Time-dependent Evolution of Multiple Pulse Operations
PPO	Proximal Policy Optimization
RL	Reinforcement Learning
DD	Dynamical Decoupling
DQ	Double-Quantum
SQ	Single-Quantum
CPMG	Carr-Purcell-Meiboom-Gill
NMR	Nuclear Magnetic Resonance
EPR	Electron Paramagnetic Resonance
ESR	Electron Spin Resonance
CW-ODMR	Continuous-Wave Optically Detected Magnetic Resonance
QuTiP	Quantum Toolbox in Python
MW	Microwave
GA	Genetic Algorithm
DC	Direct Current
AC	Alternating Current
T_2^*	Ramsey Decay Time
T_2	Spin Echo Decay Time
DEAP	Distributed Evolutionary Algorithms in Python

Introduction

The second quantum revolution promises transformative advances in sensing, computing, and communication by harnessing quantum mechanical phenomena such as superposition, entanglement, and coherence[2,3]. At the forefront of this movement, quantum sensing platforms exploit these effects to achieve unprecedented measurement precision, potentially surpassing classical limits by orders of magnitude[4,5,6]. However, the transition from proof-of-concept demonstrations to practical quantum technologies faces a fundamental challenge: the substantial gap between idealized theoretical frameworks and the complex realities of experimental implementations[7].

This dissertation addresses aspects of this theory-practice divide through four studies that span experimental characterization, theoretical analysis, computational method development, and algorithmic optimization. While quantum sensing theory often assumes perfect control, homogeneous systems, and ideal operating conditions, real experiments exhibit parasitic effects, ensemble heterogeneities, and at times, a breakdown of foundational theoretical approximations[8]. The work presented here tackles these challenges, providing both fundamental insights and practical tools to advance quantum sensing from laboratory demonstrations toward robust, scalable technologies.

Quantum sensors based on solid-state spin systems, particularly nitrogen-vacancy (NV)

centers in diamond, have emerged as leading platforms for magnetic field sensing with applications ranging from materials characterization to biomedical imaging[9,10]. These systems offer the combination of optical initialization and readout, microwave control, and operation under ambient conditions. However, their practical implementation reveals complications absent from idealized theoretical treatments. These challenges are not merely technical inconveniences but barriers that limit the scalability and reliability of quantum sensing technologies. Addressing them requires moving beyond the idealized frameworks that continue to inform quantum sensing development toward approaches that explicitly account for experimental complexities.

This thesis presents four studies that collectively advance our understanding and capabilities in quantum sensing under realistic conditions. Each study addresses a specific aspect of the theory-practice gap while contributing tools and insights with broader applicability.

Chapter 1 investigates an unexpected experimental phenomenon in ^{15}NV diamond magnetometry: electron Ramsey envelope modulation (EREEM). When bias magnetic fields are misaligned with the NV quantization axis (a common scenario in practical sensing applications) Ramsey magnetometry measurements exhibit pronounced envelope modulations that reduce sensitivity. Through theoretical modeling and experimental validation across various field configurations, this work characterizes the physical origins of EREEM and quantifies its impact on sensing performance. The study demonstrates that this effect can be understood through perturbation theory analysis of the electron-nuclear hyperfine coupling in transverse magnetic fields, providing both insight and practical guidance for optimizing sensing protocols.

Chapter 2 examines the breakdown of Average Hamiltonian Theory (AHT), a theoretical framework widely used in quantum sensing protocol development. While AHT has been instrumental in designing dynamical decoupling and sensing sequences, this work reveals that the

conditions guaranteeing Magnus expansion convergence conflict with typical operating parameters in solid-state quantum sensing. Through analysis of convergence criteria and development of exact computational methods, the study identifies when AHT predictions become unreliable and provides alternative approaches for characterizing sensor performance. The discovery of rapid echo structures that preserve AHT accuracy beyond guaranteed convergence opens new avenues for pulse sequence engineering while establishing exact methods valid regardless of convergence conditions.

Chapter 3 addresses the challenge of ensemble heterogeneity in quantum control optimization. Traditional algorithmic approaches, including machine learning methods, typically optimize control sequences for idealized single-instance systems, leading to overfitting that fails to capture the parameter variations encountered in more realistic experimental ensembles. This work introduces Ensemble Cluster Sampling (ECS), a simulation framework that incorporates experimentally motivated parameter distributions directly into the optimization process. By applying deep reinforcement learning and beam search algorithms to ECS environments rather than point-sample systems, the study demonstrates improvements in pulse sequence robustness and generalization. The framework enables the discovery of dynamical decoupling protocols inherently designed for heterogeneous systems while providing tools for simulating ensemble dephasing dynamics across realistic parameter ranges.

Chapter 4 presents TEMPO, an open-source Python package designed to make pulse sequence simulations more accessible and efficient for the quantum sensing community. Building on the QuTiP framework, TEMPO addresses computational bottlenecks that often limit the scope and scale of pulse sequence optimization studies. Through time segmentation that preserves only active pulses in each interval, the package achieves simulation times largely independent of pulse

sequence complexity while maintaining compatibility with existing QuTiP workflows. The software bridges the gap between theoretical pulse sequence design and experimental implementation by incorporating structures familiar to experimentalists working with pulse generator hardware.

While each chapter addresses distinct challenges, several connecting themes emerge that reflect broader issues in quantum sensing development. The transition from idealized to realistic system modeling appears throughout this work, whether in characterizing unexpected experimental effects (EREEM), analyzing the breakdown of theoretical approximations (AHT), accounting for ensemble heterogeneities (ECS), or developing practical computational tools (TEMPO). Similarly, the interplay between theoretical understanding and experimental validation drives each investigation, emphasizing the importance of grounding theoretical advances in experimental reality.

The methodological approaches also reflect a common philosophy: rather than treating experimental complications as obstacles to be minimized, this work embraces them as fundamental aspects of quantum sensing that must be understood and addressed. This perspective leads to solutions that are inherently robust - protocols that suppress multiple sources of error, exact theoretical methods that remain valid across all operating regimes, optimization frameworks that explicitly account for heterogeneity, and computational tools designed for practical experimental workflows.

The implications extend beyond the specific quantum sensing platforms and applications studied here. The fundamental challenges of ensemble heterogeneity, theoretical approximation breakdown, and unexpected experimental phenomena appear across quantum technologies. The frameworks and insights developed in this thesis provide templates for addressing similar challenges in quantum computing, quantum communication, and other emerging quantum appli-

cations.

The work presented in this thesis establishes foundations for several promising research directions. The exact methods developed for beyond-AHT analysis could be extended to higher-dimensional systems and more complex noise environments, potentially revealing new classes of robust pulse sequences. The ECS framework provides a platform for exploring quantum control in increasingly realistic systems, with applications ranging from quantum error correction to many-body quantum simulation. The computational tools and insights from TEMPO could inform the development of next-generation quantum control software that seamlessly integrates theoretical design with experimental implementation.

Perhaps most importantly, this work demonstrates that the theory-practice gap in quantum sensing is not an insurmountable barrier but rather a rich source of scientific insight and technological innovation. By systematically characterizing and addressing the discrepancies between idealized models and experimental reality, we can develop quantum sensing technologies that are not only more robust and reliable but also capable of exploiting phenomena that purely theoretical approaches might overlook.

The path from laboratory demonstrations to practical quantum technologies requires more than incremental improvements in sensitivity or coherence. It demands a fundamental understanding of how quantum systems behave under realistic conditions and the development of theoretical and computational frameworks capable of handling this complexity. This thesis contributes to that understanding while providing concrete tools and insights to navigate the translation of quantum sensing advances into real-world applications.

Chapter 1: Ramsey Envelope Modulation in NV Diamond Magnetometry

This chapter has been adapted from our previously published work,

***Ramsey Envelope Modulation in NV Diamond Magnetometry**,*

with authors Jner Tzern Oon, Jiashen Tang, Connor A. Hart, Kevin S. Olsson, Matthew J. Turner, Jennifer M. Schloss, and Ronald L. Walsworth,

published in Physical Review B[11].

My contributions include early experimental development and validation, primary theory and simulation responsibilities, manuscript preparation, and supervision of all project components. I am proud to share a statement of equal contribution with Jiashen Tang, who led experimental efforts on a setup with renewed capabilities for validation across a broader magnetic field range. Together we drafted the manuscript, with assistance from Connor A. Hart, Kevin S. Olsson, Matthew J. Turner, Jennifer M. Schloss, and Ronald L. Walsworth. The work was performed under the supervision of Ronald L. Walsworth.

1.1 Introduction

Ensembles of negatively charged nitrogen-vacancy (NV) centers in diamond are a leading quantum sensing platform, particularly for applications in magnetometry. The NV center has a magnetically sensitive electronic triplet ground state with spin $S = 1$ that can be optically

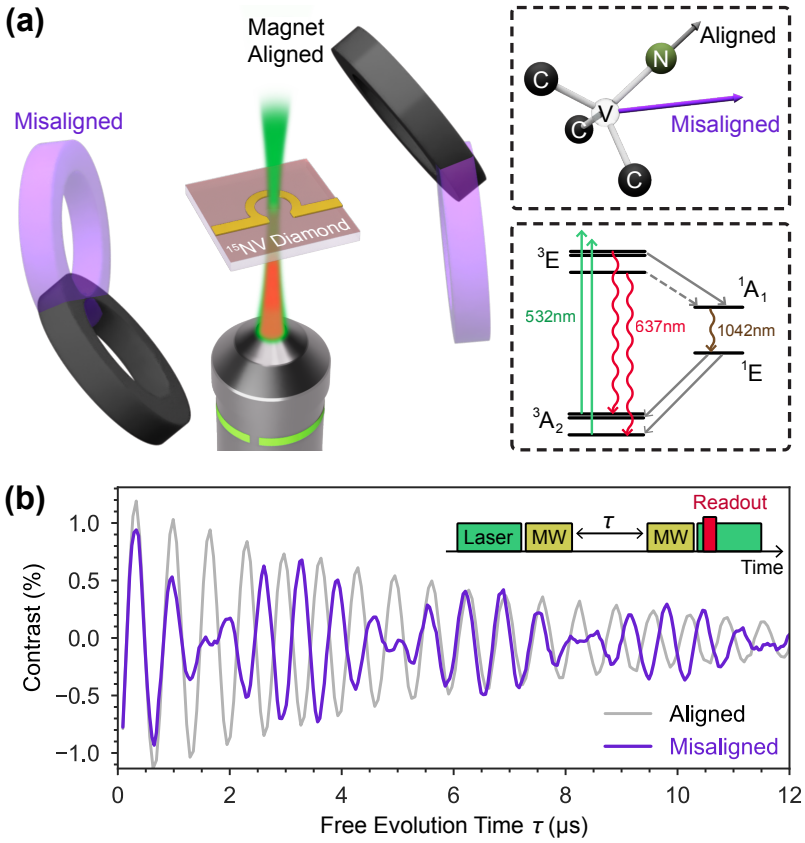


Figure 1.1: (a) Experimental setup schematic. Diamond containing an ensemble (≈ 0.3 ppm) of ^{15}NV centers is probed using 532 nm green laser excitation through a microscope objective, with red fluorescence collected back along the same path. A gold Ω-shaped waveguide is placed above the NV layer for microwave (MW) delivery. A pair of ring magnets apply a bias magnetic field at two distinct orientations, aligned and misaligned with the symmetry axis for one class of NV centers. Top right inset: diamond crystal axes and bias field orientations relative to a single NV quantization axis. Bottom right inset: NV electronic energy level diagram of ground and excited states. (b) Example Ramsey time series measured for bias fields aligned (gray) and misaligned (blue) to one class of NV centers. Effect of electron Ramsey envelope modulation (EREEM) is clearly observable for data with the misaligned field. Upper inset: Ramsey pulse sequence diagram.

initialized and read out, and coherently manipulated using microwave fields, while operating at ambient conditions.

Demonstrations of sensing or imaging of static and broadband (DC) magnetic fields have predominantly used continuous-wave optically detected magnetic resonance (CW-ODMR) techniques. However, the achievable volume-normalized magnetic sensitivity in CW-ODMR is con-

strained by competing effects of the optical and microwave fields applied during sensing[9]. Alternatively, pulsed measurement protocols such as Ramsey interference magnetometry can be employed to measure DC magnetic fields[12,13]. By separating spin control and readout from the sensing interval, pulsed measurements enable the use of increased optical and microwave intensities to improve sensitivity. As a result, some of the best volume-normalized DC sensitivities reported to date in NV ensembles (the sensitivity achieved per unit volume of interrogated diamond) were produced using Ramsey-based sensing protocols.[14,15].

Beyond advancements in sensing protocols, the optimization of diamond material properties provides a crucial path towards improvements in volume-normalized sensitivity. In particular, ^{15}NV centers found in ^{15}N-enriched diamond provide practical advantages over the naturally abundant ^{14}NV due to the nuclear spin $I = 1/2$ of ^{15}N, as compared to $I = 1$ for ^{14}N. For sensing, this difference translates to increased signal contrast during optical readout while driving a single ^{15}NV hyperfine resonance, as the nuclear spin population is only distributed between two states. Both ^{15}NV hyperfine-split electronic resonances can also be driven simultaneously with the same Rabi nutation rate by tuning the frequency of the applied microwave field to the midpoint of the splitting, enabling more uniform spin control. In addition, the two-level nuclear spin system simplifies quantum logic protocols that exploit the coupled electron-nuclear system for enhanced sensing[16,17].

The ease with which Ramsey magnetometry can be implemented with ^{15}N-enriched diamond depends on the bias magnetic field commonly applied to break the ground state electronic spin degeneracy, associated with spin sublevels $m_s = \pm 1$. In practice, the bias field magnitude and orientation is often constrained by the desired sensing modality or system to be studied. For example, full vector reconstruction of magnetic fields in three dimensions typically requires a

bias field oriented to produce a unique projection onto each class of NV centers across the four crystal axes[8,18,19,20,20]. This approach ensures that the resonances associated with each class of NV centers are non-degenerate and individually addressable with microwave control. Alternatively, the bias field may be applied to spectrally overlap two or more NV classes to increase the number of spins participating in sensing, improving sensitivity[8,21,22,23,24].

In the presence of such misaligned fields (Fig. 1.1), we observe envelope modulations in ^{15}NV Ramsey measurements, which negatively impacts sensitivity if left unaddressed. The physical origin of this behavior can be attributed to the electron-nuclear hyperfine coupling of the ^{15}NV center in the presence of a transverse magnetic field. This effect resembles the well-known electron spin echo envelope modulation (ESEEM), which has received extensive study for over half a century in NMR systems and more recently in solid state defects[25,26,27]. However, the analogous effect on a Ramsey measurement and the resulting impact on NV magnetic sensing has yet to be detailed.

In this paper, we characterize this effect, which we refer to as electron Ramsey envelope modulation or EREEM. First, we model the Ramsey envelope properties by considering an NV electronic spin coupled to its native nitrogen nuclear spin. The resulting EREEM predictions are compared to experimental results, showing good agreement for ^{15}NV ensembles across a range of magnetic field magnitudes and misalignments. We then discuss the impact of EREEM on NV-diamond magnetic sensitivity, considering typical operating conditions used for magnetometry. Finally we study EREEM in the context of double-quantum (DQ) protocols, which leverage superpositions of NV electronic spin states $|m_s = \pm 1\rangle$ for magnetometry. We demonstrate dramatic suppression of envelope modulations in DQ Ramsey measurements. These results provide further motivation for the use of DQ sensing schemes, in addition to their documented robustness

to strain gradients and temperature drift.[15,28,29,30,31].

1.2 Electron Ramsey Envelope Modulation (EREEM)

This section presents a derivation of EREEM properties, described by a simple vector model of the effective magnetic field on the nitrogen nuclear spin - which, importantly, is dependent on the NV electronic spin state. First, the ^{15}NV center is modeled by an electronic spin system ($S = 1$) coupled to the native ^{15}N nuclear spin ($I = 1/2$). Under the application of a bias magnetic field $\vec{B} = (B_x, B_y, B_z)$, the ground state Hamiltonian H can be written as[32]

$$\frac{H}{\hbar} = D\hat{S}_z^2 - \gamma_e \vec{B} \cdot \vec{S} - \gamma_n \vec{B} \cdot \vec{I} + \vec{S} \cdot \mathbf{A} \cdot \vec{I}. \tag{1.1}$$

The vectors $\vec{S} = (\hat{S}_x, \hat{S}_y, \hat{S}_z)$ and $\vec{I} = (\hat{I}_x, \hat{I}_y, \hat{I}_z)$ contain the electronic and nuclear spin operators, respectively, with corresponding gyromagnetic ratios $\gamma_e = 2\pi \times -2.8024\,$MHz/G and $\gamma_n = 2\pi \times -431.6\,$Hz/G. The room temperature zero field splitting $D \approx 2\pi \times 2.87\,$GHz[33] sets the electron quantization axis along $\hat{z}$. The NV hyperfine interaction is described by the diagonal tensor $\mathbf{A} = \begin{pmatrix} A_\perp & 0 & 0 \\ 0 & A_\perp & 0 \\ 0 & 0 & A_{||} \end{pmatrix}$, with transverse and longitudinal components $A_\perp = 2\pi \times 3.65\,$MHz and $A_{||} = 2\pi \times 3.03\,$MHz, respectively[34]. This ground state energy level structure is depicted in Fig. 1.2(a), for a magnetic field of magnitude B aligned along the $\hat{z}$ direction.

The C_{3v} symmetry of the NV center allows us to restrict the magnetic field to the x-z plane without any loss in generality[35]. For this study, we consider bias magnetic fields $B < 200\,$G. Within this field regime, the zero field splitting term $D\hat{S}_z^2$ sets the dominant energy scale in the Hamiltonian, allowing us to treat contributions not commuting with S_z as non-secular perturbations. Accurate to second order in perturbation theory, a leading order correction to the

secular Hamiltonian can be obtained[27,35,36]. After transforming into a frame resonant with the two electronic transitions $m_s = 0 \leftrightarrow +1$ and $0 \leftrightarrow -1$[30], the following Hamiltonian under the rotating wave approximation is found:

$$\begin{aligned}\frac{\widetilde{H}}{\hbar} =& A_{||}\hat{S}_z\hat{I}_z - \gamma_n B_z \hat{I}_z \\ &- (1-2\kappa)\gamma_n B_x \hat{I}_x - 3\kappa\gamma_n B_x \hat{S}_z^2 \hat{I}_x.\end{aligned} \tag{1.2}$$

Here, a dimensionless factor $\kappa \equiv \frac{\gamma_e A_\perp}{\gamma_n D} \approx 8.26$ describes an effective amplification of the bare nuclear spin response to a transverse magnetic field B_x, by a factor of $1 - 2\kappa \approx -15.5$.

The contributions to Eq. (1.2) can be separated into two categories. The first category consists of terms that depend on $\hat{S}_z$ or equivalently m_s. The sum of these terms can be described as an effective vector magnetic field $\vec{\beta}(m_s)$ on the nuclear spin. The terms that do not contain $\hat{S}_z$ can be represented by a spin-independent effective field $\vec{\beta}_{\text{ind}}$, with a constant coupling to the nuclear spin regardless of the electronic spin state m_s. The resulting Hamiltonian can thus be summarized as

$$\frac{\widetilde{H}_n}{\hbar} = -\gamma_n \left(\vec{\beta}_{\text{ind}} + \vec{\beta}(m_s)\right) \cdot \vec{I}. \tag{1.3}$$

For a given electronic spin state $|m_s\rangle$, the nuclear spin precesses around an effective magnetic field $\vec{\beta}_{\text{ind}} + \vec{\beta}(m_s)$ with a Larmor frequency

$$\omega_{m_s} = \left|\gamma_n\left(\vec{\beta}_{\text{ind}} + \vec{\beta}(m_s)\right)\right|. \tag{1.4}$$

For $m_s = 0$ in particular, there are no terms in Eq. (1.2) that depend on the electronic spin state, such that $\vec{\beta}(0) = 0$ and $\omega_0 = \left|\gamma_n \beta_{\text{ind}}\right|$. These field vectors are visualized in Figure 1.2(b). For

simplicity, the coordinate system is rotated so the spin-independent field $\vec{\beta}_{\text{ind}} = \beta_{\text{ind}}\hat{z}'$ now lies along the newly defined z'-axis[36]. In this frame, the angle between $\vec{\beta}_{\text{ind}}$ and an effective field $\vec{\beta}_{\text{ind}} + \vec{\beta}(m_s)$ is given by

$$\phi_{m_s} = \tan^{-1}\left(\frac{\beta_{x'}(m_s)}{\beta_{\text{ind}} + \beta_{z'}(m_s)}\right), \tag{1.5}$$

where $\beta_{z'}(m_s)$ and $\beta_{x'}(m_s)$ denote components of $\vec{\beta}(m_s)$ parallel and perpendicular to $\hat{z}'$, respectively. For distinct electronic spin states $|i\rangle$ and $|j\rangle$, we define the angle between the two corresponding effective fields on the nuclear spin as $\Phi_{i,j}$, with the example $\Phi_{0,+1} = \phi_{+1}$ shown in Fig. 1.2(b).

Using this vector description, we derive the expected Ramsey envelope modulation as a function of the free evolution time τ. For an initial superposition of electronic spin states $|i\rangle$ and $|j\rangle$, the Ramsey signal $S_{i,j}(\tau)$, up to an overall phase, is given by

$$\begin{aligned} S_{i,j}(\tau) = &\cos\left(\Phi_{i,j}\right)\sin\left(\frac{\omega_i\tau}{2}\right)\sin\left(\frac{\omega_j\tau}{2}\right) \\ &- \cos\left(\frac{\omega_i\tau}{2}\right)\cos\left(\frac{\omega_j\tau}{2}\right). \end{aligned} \tag{1.6}$$

As an example, Fig. 1.2(c) shows the expected Ramsey response $S_{0,+1}(\tau)$ for a bias field of magnitude $B = 100\,\text{G}$ misaligned from the NV axis by an angle $\theta = 22°$. The Ramsey signal oscillation is modulated by an envelope with a slow characteristic beat frequency ω_0, i.e., an example of EREEM. The corresponding amplitude spectrum is shown in Fig. 1.2(d), revealing two peaks centered around $\omega_{+1}/2 \approx A_{||}/2 = 2\pi \times 1.515\,\text{MHz}$, with a frequency splitting of magnitude ω_0. Experimentally, this behavior can be produced by tuning the frequency ν_{+1} of the applied microwave $\pi/2$ pulses to be centered between the two hyperfine-split resonances, resulting in detunings of equal magnitude $\approx A_{||}/2$ from each transition.

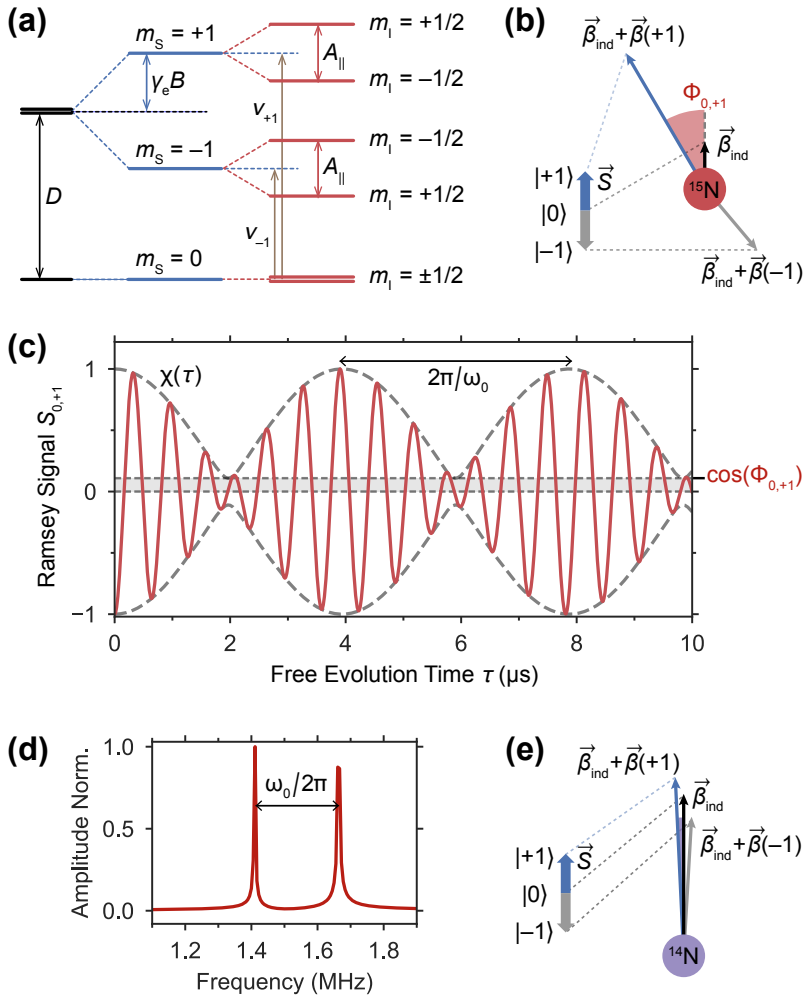


Figure 1.2: (a) Energy level diagram for the ^{15}NV ground state, with splittings due to Zeeman and hyperfine interactions assuming an aligned bias field. The frequency of an applied microwave field ν_{+1} is centered between the two $m_s = 0 \leftrightarrow m_s = +1$ electronic resonances split by the NV hyperfine interaction. (b) Effective magnetic field vectors on the ^{15}N nuclear spin due to a misaligned bias magnetic field, for each electronic spin state $|m_s\rangle$. The nuclear spin experiences an effective magnetic field consisting of an electronic-spin-dependent contribution $\vec{\beta}(m_s)$, and a component $\vec{\beta}_{\text{ind}}$ that is independent of the electronic spin state. (c) Example ^{15}NV Ramsey free induction signal due to a bias magnetic field of magnitude $B = 100\,\text{G}$ at a misalignment of 22°, showing envelope modulation. The amplitude modulation $\chi(\tau)$ is indicated, showing oscillations between $\chi_{\text{min}} = \cos(\Phi_{0,+1})$ and 1, with a characteristic period $2\pi/\omega_0$. (d) Amplitude spectrum of the Ramsey signal from (c), revealing two peaks split by the envelope beat frequency $\omega_0 = 2\pi \times 254\,\text{kHz}$. (e) Effective magnetic field vectors on the ^{14}N nuclear spin. For all three electronic spin states, the corresponding effective fields on the nuclear spin are nearly parallel due to a dominant nuclear quadrupolar field.

At any given time τ, the maximum Ramsey signal contrast is corrected by a multiplicative factor $\chi(\tau)$, due to this envelope modulation. This amplitude modulation factor $\chi(\tau)$ oscillates as a function of τ between values $\chi_{\text{min}} = |\cos(\Phi_{0,+1})|$ and $\chi_{\text{max}} = 1$, indicating points of minimum and maximum contrast, respectively. The depth of this modulation can be inferred from the angle $\Phi_{0,+1}$ between the participating effective fields on the nuclear spin.

To connect this vector model to the expected EREEM behavior in experimentally realistic conditions, we first consider a magnetic field aligned with the NV axis. Since $B_x = 0$, the Hamiltonian in Eq. (1.2) consists only of nuclear spin contributions along $\hat{I}_z$. In this case, the effective fields coupled to the nuclear spin, $\vec{\beta}_{\text{ind}}$ and $\vec{\beta}_{\text{ind}} + \vec{\beta}(+1)$, are parallel such that $\Phi_{0,+1} = 0$. No Ramsey envelope modulation should be observed, as the amplitude modulation factor remains constant: $\chi_{\text{min}} = |\cos(\Phi_{0,+1})| = \chi_{\text{max}} = 1$. However, in the presence of a magnetic field not aligned with the NV axis ($B_x \neq 0$), the nuclear spin experiences an enhanced transverse magnetic field determined by the factor κ. The effective fields are no longer aligned, $\Phi_{0,+1} > 0$, which should lead to observable envelope modulation. At bias magnetic fields where $\vec{\beta}_{\text{ind}}$ and $\vec{\beta}_{\text{ind}} + \vec{\beta}(+1)$ are orthogonal, $\chi_{\text{min}} = 0$ and the Ramsey signal contrast at modulation nodes is maximally suppressed.

This vector model can be readily extended to the ^{14}NV center, with some modifications. Besides straightforward changes to the physical constants $A_\perp$, $A_{||}$, and γ_n, an additional nuclear quadrupolar interaction term $Q\hat{I}_z^2$ contributes to the Hamiltonian in Eq. (1.1), with quadrupolar coupling constant $Q = 2\pi \times -4.945\,\text{MHz}$[37,38]. This inclusion dramatically changes the Ramsey envelope properties, by contributing a large quantizing field of magnitude $|Q/\gamma_n| \approx 16\,000\,\text{G}$ to $\vec{\beta}_{\text{ind}}$[27]. In the small magnetic field regime $B \ll |Q/\gamma_n|$, the effective fields $\vec{\beta}_{\text{ind}} + \vec{\beta}(m_s)$ coupled to the nuclear spin are dominated by the spin-independent contribution $\vec{\beta}_{\text{ind}}$, which is visualized in

Fig. 1.2(e). The small angle between these field vectors results in $\chi_{\text{min}} \approx 1$, and hence suppressed EREEM for ^{14}NV.

1.3 Experimental Methods

The measurements for this study utilize a custom-built microscope setup and a $100\,\mu$m-thick, ^{15}N-enriched CVD diamond layer (NV $T_2^* = 5\,\mu$s, [N] ≈ 3 ppm, $> 99.995\%$ ^{12}C), grown by Element Six Ltd. on a $2\times2\times0.5\,\text{mm}^3$ high-purity diamond substrate, as shown in Fig. 1.1(a). Post-growth treatment via electron irradiation and annealing increases the NV concentration to approximately 0.3 ppm.

Initialization of the NV ensemble electronic spin states is accomplished via pulsed 532 nm excitation, generated by a continuous-wave laser gated by an acoustic optical modulator (AOM). The light is focused onto the NV layer using a microscope objective, which is also used to route the outgoing NV fluorescence for spin state readout. Microwave pulses are synthesized by signal generators and controlled by switches for single- or double-quantum control of the NV spin states. The microwave drive fields are delivered through an Ω-shaped planar waveguide, fabricated onto a sapphire substrate. The bias magnetic field is applied using two identical permanent ring magnets equally spaced from the diamond sample. The field magnitude is manually adjusted by varying the separation between the magnets. To control the field misalignment angle from the target NV axis, two automated rotation stages are used to adjust the yaw and pitch of the magnet pair with a nominal accuracy of 0.047°. The bias magnets provide a homogeneous field magnitude of up to $\sim$150 G over an illumination spot size of $\sim20\,\mu$m in diameter on the NV-rich layer. Additional details regarding the experimental setup are provided in the Supplementary

Material[36].

To accurately determine the bias field magnitude B and misalignment angle θ, pulsed optically detected magnetic resonance (pulsed-ODMR) spectroscopy is employed to probe the NV ground state spin resonances. First, the field is aligned to a single NV axis by adjusting the magnets until the pulsed-ODMR spectra of the other three misaligned NV classes overlap. Using weak microwave π-pulses of duration $\sim 1\,\mu$s each[39], the resonance spectrum of the aligned NV axis is recorded for both electronic transitions $m_s = 0 \leftrightarrow +1$ and $m_s = 0 \leftrightarrow -1$. These transitions are separated by $\Delta_{\pm 1} = 2\gamma_e B$, which is used to estimate the bias field magnitude B. The magnets are then rotated away from the NV quantization axis, and Ramsey measurements are performed at a range of misalignment angles θ. At each position, the ODMR spectrum is again recorded to measure $\Delta_{\pm 1}$. This frequency difference is determined by the field projection along the NV axis, $\Delta_{\pm 1} = 2\gamma_e B \cos\theta$, which is then used to estimate θ. Additional details are provided in the Supplemental Material[36]. Field misalignment angles of up to $\theta \approx 40^\circ$ can be accessed, limited by the geometrical constraints of the setup.

1.4 Measured EREEM Properties

To study the properties of electron Ramsey envelope modulation (EREEM), we perform a series of single-quantum Ramsey experiments involving the electronic basis states $|m_s = 0\rangle$ and $|m_s = +1\rangle$ only (similar results are expected for experiments using $|m_s = 0\rangle$ and $|m_s = -1\rangle$). The envelope properties are extracted using fits to the measured Ramsey time series, repeated at various magnetic field magnitudes and misalignment angles, with results shown in Fig. 1.3. The Ramsey protocol consists of two microwave $\pi/2$ pulses, spaced by a variable free precession

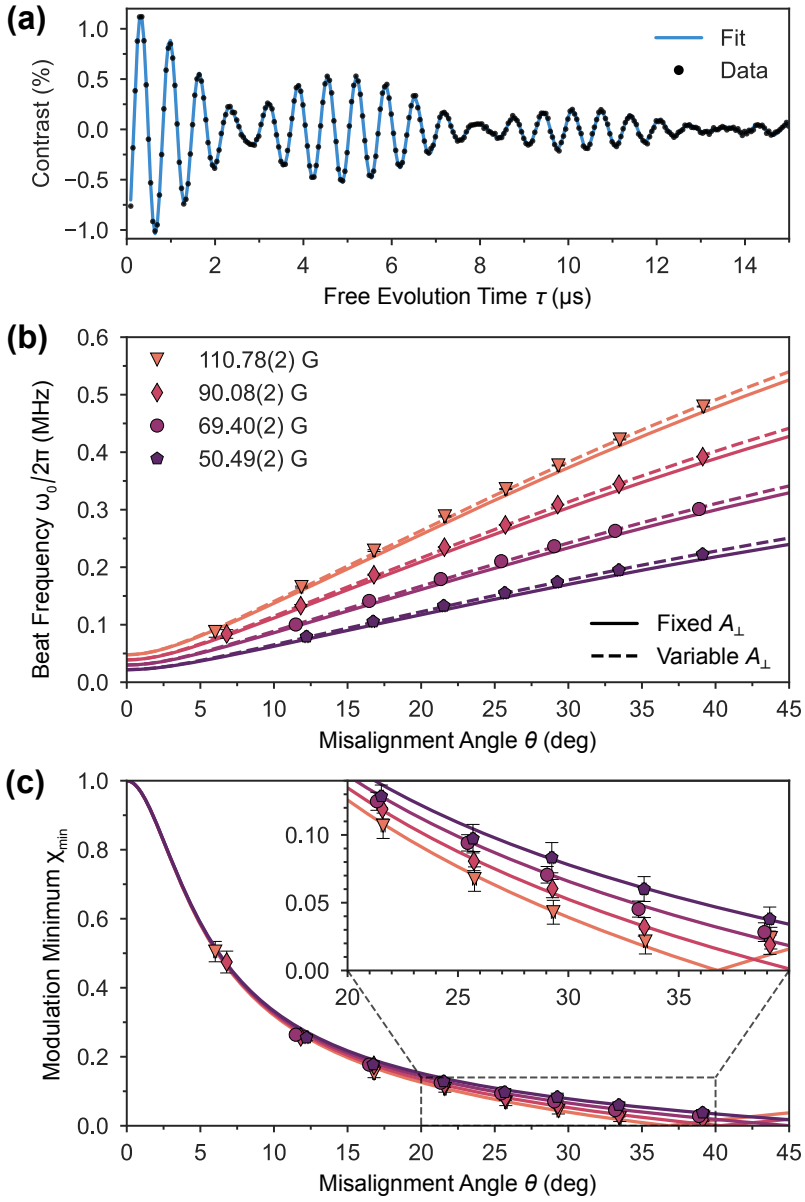


Figure 1.3: Ramsey experimental data and envelope properties from model fits. (a) Example Ramsey time series at field magnitude $B = 90.08(2)$ G and misalignment angle $\theta = 16.79(9)^\circ$. Data is shown in black and the fit in blue. The observed slow beating is an example of EREEM. (b) Estimates of envelope beat frequency ω_0 extracted from fits to experimental data, collected across a range of magnetic field magnitudes and orientations. Theoretical predictions for ω_0 are shown in solid lines for each magnetic field. After allowing the transverse hyperfine constant $A_\perp$ to vary as a fitting parameter, updated ω_0 predictions at each field are shown with dashed lines. (c) Estimates of the relative contrast at amplitude modulation nodes, χ_{min}, for the magnetic field configurations used in (b). Solid lines indicate theoretical predictions with $A_\perp$ fixed. Inset: Magnified view of χ_{min} values for misalignment angles $20^\circ < \theta < 40^\circ$.

interval τ. We tune the frequency ν_{+1} of the applied microwave field to the center of the hyperfine splitting[36], as illustrated in Fig. 1.2(a). This results in equal detunings of approximate magnitude $A_{||}/2 = 2\pi \times 1.515\,\text{MHz}$ from each hyperfine transition, giving the characteristic Ramsey fringe frequency. When exposed to a misaligned bias field, an additional slow beating (EREEM) is observed, as shown by an example in Fig 1.3(a). (The Supplementary Material[36] presents data and analysis for the applied microwave field frequency not being equally detuned from the two hyperfine-split $m_s = 0 \leftrightarrow m_s = +1$ transitions.)

The measured Ramsey signals are fit to a modified form of Eq. (1.6), which incorporates an exponential decay $e^{-(\tau/T_2^*)^p}$ with a characteristic dephasing time T_2^* and stretch factor p. To further adapt the expression to experimental data, an overall amplitude scaling factor, a vertical offset, and phase offsets are all included in the fit function[36]. From the resulting fits, two frequencies ω_0 and ω_{+1} are obtained. The values for the envelope beat frequency ω_0 are plotted in Fig. 1.3(b), with 95% confidence intervals indicated by error bars[36,40]. For each magnetic field magnitude, theoretical predictions for ω_0 are also plotted as solid curves, obtained from the spin-independent contributions to Eqs. (1.2) and (1.3),

$$\begin{aligned} \omega_0 = |\gamma_n \beta_{\text{ind}}| &= |\gamma_n| \sqrt{B_z^2 + (1-2\kappa)^2 B_x^2} \\ &= |\gamma_n|\, B \sqrt{1 + 4(\kappa^2 - \kappa)\sin^2\theta}. \end{aligned} \tag{1.7}$$

As expected from Eq. (1.7), the measured envelope beat frequency ω_0 increases with both the magnetic field magnitude B and misalignment angle θ.

Notably, small differences are observed between theoretical predictions of ω_0 (solid curves in Fig. 1.3(b)) and fits to experimental data, ranging from around 2% to 6% across the measure-

ments presented in Fig. 1.3. To explore this inconsistency, we first perform full-Hamiltonian numerical simulations of Ramsey spin dynamics and compare the observed envelope beat frequency to Eq. (1.7), which was originally obtained using second order perturbation theory. These simulations are conducted using the QuTiP package[41,42] in Python. The NV system is described by the lab frame Hamiltonian from Eq. (1.1) and the pulse sequence is implemented using time-dependent AC magnetic field contributions. For the field configurations considered in Fig. 1.3, strong agreement is observed between Eq. (1.7) and the results of QuTiP simulations, with differences in $\omega_0 \lesssim 1\%$ (see the Supplemental Material[36]). These results, however, do not fully account for the observed discrepancies in ω_0.

Interestingly, the agreement between analytical and experimental results is improved when the enhancement parameter $\kappa = \frac{\gamma_e A_\perp}{\gamma_n D}$ in Eq. (1.7) is allowed to deviate. Besides the transverse hyperfine constant $A_\perp$, other contributions to κ include the well-established gyromagnetic ratios γ_e and γ_n, and the zero field splitting D. We determine D to be 2870.71(4) MHz using pulsed-ODMR measurements, consistent ($< 0.03\%$ deviation) with the value assumed for analytical predictions and QuTiP numerical simulations. In contrast, an experimental determination of $A_\perp$ for ^{15}NV has (to our knowledge) only been reported once, by Felton et al.[34], using EPR studies at higher fields $B \sim 2000$ G. Given the simple relationship between the envelope beat frequency ω_0 and $A_\perp$ at low fields (via κ in Eq. (1.7)), EREEM presents a direct probe of $A_\perp$ in this regime. With this in mind, we conduct a phenomenological fit of Eq. (1.7) to measurements of ω_0 at each magnetic field, with $A_\perp$ as the sole degree of freedom. Using the adjusted $A_\perp$ values at each field, the corresponding values for ω_0 from Eq. (1.7) are shown as dashed lines in Fig. 1.3(b). We obtain values of $A_\perp/2\pi$ between 3.74 MHz and 3.80 MHz across the fields considered here, differing slightly from the previously reported value of 3.65(3) MHz by around 3%. Although

such adjustments to $A_\perp$ reconcile the predictions from Eq. (1.7) with the results of experimental fits, we also observe a dependence on the bias magnetic field B[43,44], detailed further in the Supplemental Material[36]. This behavior warrants a more thorough study of EREEM across an extended range of magnetic fields, and is left as a subject for future work.

Figure 1.3(c) shows excellent agreement between theoretical predictions and experimentally observed values of the relative contrast at amplitude modulation nodes $\chi_{\text{min}} = |\cos(\Phi_{0,+1})|$. Even for modest misalignments of $\sim 10°$, the contrast at the nodes of the envelope modulation is reduced to around 30% of its maximum value. The resulting implications for magnetic field sensitivity are discussed in the following section.

1.5 Modulation Amplitude and Impact on Magnetometry

To study the impact of EREEM on NV magnetic field sensitivity, we first consider a conventional Ramsey magnetometry measurement in the absence of any envelope modulation, using a single-quantum coherence between states $|m_s = 0\rangle$ and either $|m_s = +1\rangle$ or $|m_s = -1\rangle$. After a Ramsey free evolution duration τ, this coherent superposition accumulates a relative phase $\phi \approx \gamma_e b_s \tau$ proportional to a magnetic field signal b_s along the NV axis[45]. The final $\pi/2$ pulse of each Ramsey sequence maps this phase onto changes in the NV electronic spin population, which can be detected using fluorescence contrast measurements at a fixed τ. This optimal working point τ is determined by optimizing the photon shot noise-limited sensitivity[9],

$$\eta = \frac{1}{\gamma_e} \frac{1}{C e^{-(\tau/T_2^*)^p} \sqrt{N}} \frac{\sqrt{\tau + T_D}}{\tau}. \tag{1.8}$$

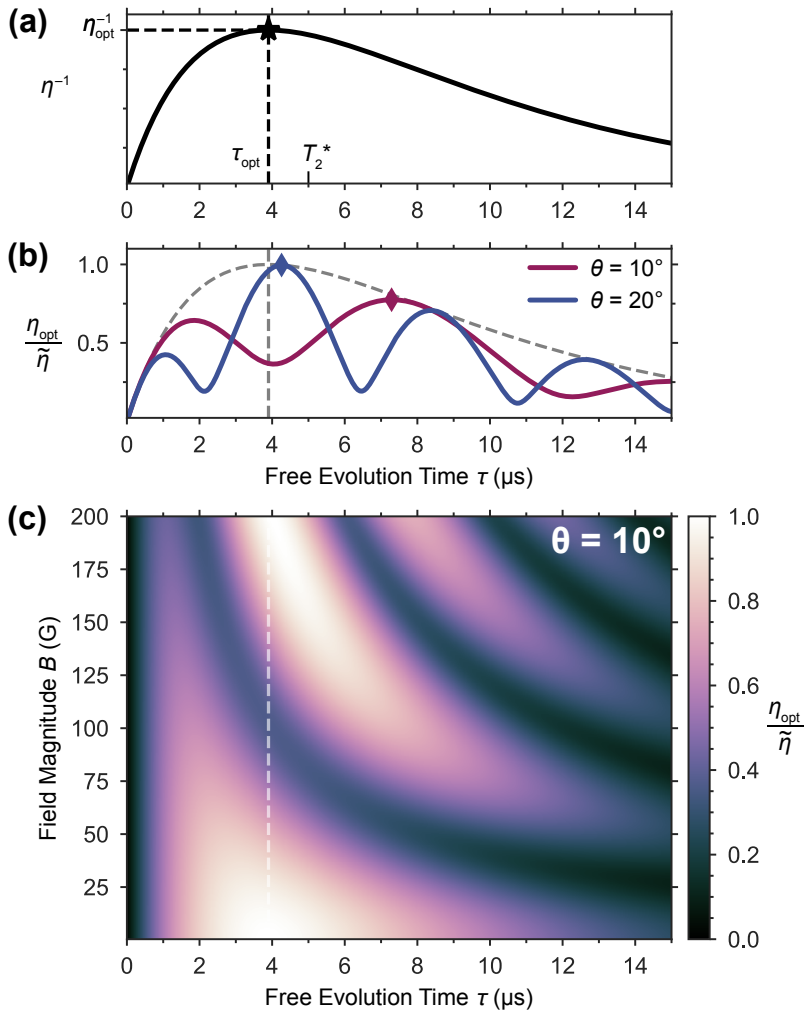


Figure 1.4: Calculations of expected Ramsey magnetic field sensitivity with and without envelope modulation (EREEM), and at different bias magnetic field configurations. (a) Inverse sensitivity η^{-1} vs Ramsey free evolution time τ without EREEM, given by Eq. (1.9) for fixed dead time $T_D = 5\,\mu s$ and dephasing time $T_2^* = 5\,\mu s$. The optimal inverse sensitivity η_{opt}^{-1} is indicated, corresponding to a free evolution time τ_{opt} indicated in all subfigures with a vertical dashed line. (b) Relative inverse sensitivity with EREEM $\eta_{opt}/\widetilde{\eta}$, normalized to the optimal sensitivity and free evolution time established in (a), for two distinct magnetic field misalignment angles $\theta = 10°$, $20°$ at the fixed magnitude $B = 100\,G$. Adjusted optimal evolution times $\widetilde{\tau}_{opt}$ are indicated by symbols $\diamond$. (c) 2D color plot of the normalized inverse sensitivity at a fixed misalignment angle $\theta = 10°$, across bias field magnitudes $0 < B < 200\,G$.

This calculation includes the maximum Ramsey fluorescence contrast C, obtained experimentally, which decays due to NV spin dephasing via a correction factor $e^{-(\tau/T_2^*)^p}$. The average photon number is denoted by N and the dead time T_D represents the time spent outside the Ram-

sey sensing sequence for spin state initialization and readout. A plot of η^{-1} as a function of τ is shown in Figure 1.4(a), setting $p = 1$, $T_2^* = 5\,\mu s$, and $T_D = 5\,\mu s$ according to our experimental conditions. This reveals an optimal sensitivity η_{opt} obtained at a corresponding free evolution time τ_{opt}, which approaches T_2^* in the limit of long overhead time $T_D \gg \tau$.

If EREEM is observed, then the contrast C takes on an additional correction factor due to the amplitude modulation $\chi(\tau)$, which results in an adjusted sensitivity $\widetilde{\eta}$. Normalizing to η_{opt}, the relative inverse sensitivity is therefore given by the following ratio, assuming N, T_2^*, p, and T_D remain unchanged between experiments:

$$\frac{\eta_{\text{opt}}}{\widetilde{\eta}} = \chi(\tau)\sqrt{\frac{\tau_{\text{opt}} + T_D}{\tau + T_D}}\exp\left(\frac{\tau_{\text{opt}} - \tau}{T_2^*}\right). \tag{1.9}$$

As established earlier, $\chi(\tau)$ oscillates between values $\chi_{\text{min}} = |\cos(\Phi_{0,\pm1})|$ and $\chi_{\text{max}} = 1$, at a characteristic beat frequency ω_0. These envelope properties are determined by the magnetic field magnitude B and misalignment angle θ, which in turn affect the relative inverse sensitivity $\eta_{\text{opt}}/\widetilde{\eta}$.

The optimal sensitivity $\widetilde{\eta}_{\text{opt}} = \eta_{\text{opt}}$ is only achieved when an envelope maximum χ_{max} occurs at τ_{opt}, obtained when the beat frequency ω_0 is an integer multiple of $2\pi/\tau_{\text{opt}}$. If this condition is not satisfied, an updated optimal evolution time $\widetilde{\tau}_{\text{opt}}$ is necessary to minimize sensitivity degradation. These two scenarios are depicted in Fig. 1.4(b), which shows $\eta_{\text{opt}}/\widetilde{\eta}$ at two misalignment angles 10° and 20°, for a field magnitude $B = 100\,\text{G}$. The adjusted optimal evolution time $\widetilde{\tau}_{\text{opt}}$ for each case is marked, with a notable reduction in sensitivity seen in the 10° configuration. Similarly, changes in the field magnitude B can affect sensitivity. Fixing the misalignment angle at 10° as an example (see the Supplemental Material for other misalignment angles[36]), Fig. 1.4(c) shows $\eta_{\text{opt}}/\widetilde{\eta}$ across a range of magnetic field magnitudes B and free evolution times τ. Since

the beat frequency ω_0 increases with the field magnitude B, the relative inverse sensitivity $\eta_{\text{opt}}/\widetilde{\eta}$ exhibits faster oscillations with respect to τ for higher fields. The sensitivity is optimized at fields where a Ramsey envelope maximum coincides with τ_{opt}, with the latter indicated by a dashed line in Fig. 1.4(c).

These calculations indicate that the sensitivity loss due to EREEM is highly dependent on changes in the bias field configuration. In practice, the tunability and control of such parameters depend on the specific sensing modality or application. For example, in experiments where an equal bias field projection on multiple NV axes is desired, the misalignment angle θ is highly constrained. Separately, there may be restrictions on the applied field magnitude B, for example, during studies of paramagnetic systems[18].

1.6 Double-quantum Ramsey

As described in the previous section, envelope modulation (EREEM) in single-quantum (SQ) Ramsey experiments depends on the bias magnetic field configuration, and can result in significant magnetic field sensitivity loss. Alternatively, double-quantum (DQ) Ramsey protocols, which exploit the full NV spin-1 system, can circumvent the deleterious effects of EREEM. In fact, we observe a dramatic reduction of envelope modulation while using DQ coherence magnetometry. This behavior is illustrated in Fig. 1.5(a,b), which depicts measured SQ and DQ Ramsey free induction decay signals and their corresponding amplitude spectra, at the same magnetic field configuration.

The DQ Ramsey protocol employs dual-tone microwave pulses with frequencies resonant with both the electronic transitions $m_s = 0 \leftrightarrow +1$ and $m_s = 0 \leftrightarrow -1$, often referred to as

DQ pulses. Besides this change to the applied pulses, the DQ Ramsey sequence mirrors the SQ protocol and consists of a pair of DQ pulses separated by a free evolution interval τ. The first DQ pulse prepares an equal superposition of the electronic spin states $|+1\rangle$ and $|-1\rangle$. After the interval τ, a second DQ pulse maps the relative phase accumulated by these basis states onto the NV spin population, which is then read out optically.

The lack of EREEM in the observed DQ signal can be understood by referring back to the vector model established in Sec. 1.2. The expected DQ Ramsey response can be described by Eq. (1.6) after substituting the electronic basis states denoted by i and j with -1 and $+1$, respectively. The effective nuclear magnetic fields associated with the electronic spin states $|m_s = \pm 1\rangle$ are then given by $\vec{\beta}_{\text{ind}} + \vec{\beta}(\pm 1)$ as depicted in Fig. 1.5(b). These field vectors are nearly antiparallel ($\Phi_{-1,+1} \approx \pi$), resulting in negligible envelope modulation $\chi_{\text{min}} = |\cos(\Phi_{-1,+1})| \approx 1$ for a DQ Ramsey measurement.

The stark difference in envelope modulation behavior between SQ and DQ Ramsey is highlighted by calculations shown in Fig. 1.5(c). For a range of bias field magnitudes B and misalignment angles θ, the relative contrast at amplitude modulation nodes χ_{min} is plotted for both cases (see the Supplemental Material[36] for extended ranges of B and θ). Consistent with the results shown in Fig. 1.3(c), the SQ Ramsey contrast at envelope nodes decays rapidly as a function of θ, nearing zero even for small misalignment angles $\theta \sim 10^\circ$. On the other hand, the DQ contrast is well preserved, with values of $\chi_{\text{min}} \approx 1$ across the fields considered in this work. Compared to SQ Ramsey, DQ Ramsey often provides more than an order of magnitude suppression of Ramsey amplitude modulation.

1.7 Conclusion

In this work, we present a physical model of electron Ramsey envelope modulation (EREEM) and find reasonable agreement with experimental measurements using an ensemble of ^{15}NV centers in diamond. The observed envelope modulation exhibits a characteristic beat frequency and amplitude, dependent on the bias field magnitude and angle with respect to the NV quantization axis. We note a small systematic discrepancy between measurements of the envelope beat frequency and analytical predictions, which can be reconciled using an adjustment to the transverse hyperfine parameter $A_\perp$. These estimates deviate by around 3% from previous EPR measurements conducted at $\sim 2000\,\mathrm{G}$[34], an order of magnitude greater than the magnetic fields considered here. However, these estimates of $A_\perp$ exhibit a dependence on the applied field magnitude, warranting additional measurements across an extended range of magnetic fields in future studies.

For magnetic field sensing modalities requiring misaligned magnetic fields, the integration of ^{15}NV diamond and Ramsey coherence magnetometry is hindered by envelope modulation effects. We show that the resultant loss in sensitivity can be recovered by careful choice of the bias field. However, experimental constraints can limit the tunability of these field parameters. Alternatively, we find that double-quantum coherence magnetometry dramatically suppresses envelope modulation, while providing robustness to strain and temperature changes.

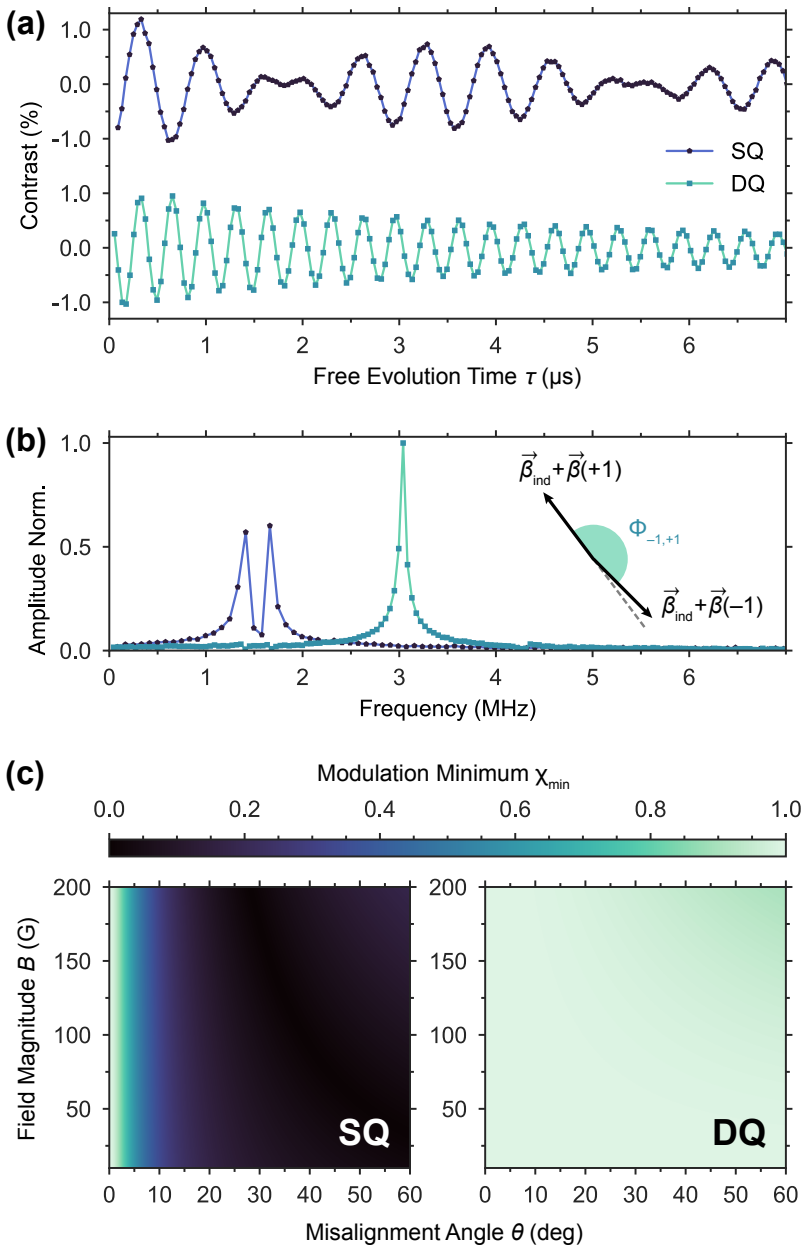


Figure 1.5: Comparison between experimental single-quantum (SQ) and double-quantum (DQ) Ramsey measurements under a misaligned bias magnetic field. Representative SQ and DQ Ramsey data collected at a bias field magnitude $B \approx 50\,\mathrm{G}$ and misalignment angle $\theta \approx 35°$, in both time (a) and frequency (b) domains. The inset in (b) shows the effective magnetic fields $\vec{\beta}_{\mathrm{ind}} + \vec{\beta}(\pm 1)$ coupled to the nuclear spin, due to the electronic spin states $|m_s = \pm 1\rangle$ participating in the DQ protocol. The effective fields are nearly antiparallel ($\Phi_{-1,+1} \approx \pi$), resulting in no observable envelope modulation in the DQ Ramsey signal ($\chi_{\mathrm{min}} = |\cos \Phi_{-1,+1}| \approx 1$). (c) Calculation of relative contrast at amplitude modulation nodes χ_{min} for both SQ and DQ Ramsey signals, over a range of bias magnetic field magnitudes and misalignment angles. For SQ Ramsey, calculated values of $\chi_{\mathrm{min}} = |\cos \Phi_{0,+1}| = |\cos \phi_{+1}|$ are shown, obtained using Eq. (1.5). For DQ Ramsey, $\chi_{\mathrm{min}} = |\cos \Phi_{-1,+1}| = |\cos (\phi_{+1} - \phi_{-1})|$.

1.8 Supplementary Notes

1.8.1 Theoretical Details

This section provides a detailed derivation of the expressions given in Sec. 1.2 of the main text. We start with the NV ground state Hamiltonian H, including the effects of a magnetic field restricted to the x-z plane and the hyperfine interaction between the NV electronic spin and the onsite nitrogen nuclear spin,

$$\frac{H}{\hbar} = DS_z^2 - \gamma_e \left(B_z \hat{S}_z + B_x \hat{S}_x \right) - \gamma_n g_n \left(B_z \hat{I}_z + B_x \hat{I}_x \right) + A_{||} \hat{S}_z \hat{I}_z + A_\perp \left(\hat{S}_x \hat{I}_x + \hat{S}_y \hat{I}_y \right). \tag{1.10}$$

The contributions to H_0 can be separated into secular and non-secular contributions H_S and H_{NS}, respectively. The secular term H_S consists of contributions that commute with operator $\hat{S}_z$. Conversely, the non-secular terms H_{NS} do not commute with $\hat{S}_z$ and are capable of driving transitions within the electronic spin system. This is summarized as follows:

$$\begin{aligned} H &= H_S + H_{NS} \\ \frac{H_S}{\hbar} &\equiv D\hat{S}_z^2 - \gamma_e B_z \hat{S}_z - \gamma_n g_n \left(B_z \hat{I}_z + B_x \hat{I}_x \right) + A_{||} \hat{S}_z \hat{I}_z \\ \frac{H_{NS}}{\hbar} &\equiv -\gamma_e B_x \hat{S}_x + A_\perp \left(\hat{S}_x \hat{I}_x + \hat{S}_y \hat{I}_y \right). \end{aligned} \tag{1.11}$$

We begin by reproducing work done by Childress *et al.*[27] and Myers[35]. The Hamiltonian is expanded as a series determined by order D, the largest energy scale within the Hamiltonian for the magnetic fields $B < 200\,\mathrm{G}$ used in our study. The non-secular terms H_{NS} are treated as a perturbation to the secular Hamiltonian H_S. Using second order perturbation theory, we obtain

the leading order correction to the Hamiltonian $H_{m_s}^{(1)}$ for a given electron spin quantum number m_s,

$$\hat{P}_{m_s} H_S + H_{m_s}^{(1)} = \hat{P}_{m_s} H_{NS} \frac{1}{E_{m_s} - \left(\mathbb{1} - \hat{P}_{m_s}\right) H_S \left(\mathbb{1} - \hat{P}_{m_s}\right)} H_{NS} \hat{P}_{m_s}, \tag{1.12}$$

with projection operator $\hat{P}_{m_s} \equiv |m_s\rangle \langle m_s|$ and identity operator $\mathbb{1}$. An effective nuclear g-tensor $g_n(m_s)$ is defined by evaluating the change in the Hamiltonian with respect to the magnetic field,

$$\begin{aligned} g_n(m_s) &= \frac{d}{dB}\left(\hat{P}_{m_s} H_S + H_{m_s}^{(1)}\right) \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \frac{\gamma_e}{\gamma_n D}(2 - 3|m_s|) \begin{pmatrix} A_\perp & 0 & 0 \\ 0 & A_\perp & 0 \\ 0 & 0 & 0 \end{pmatrix}. \end{aligned} \tag{1.13}$$

This effective nuclear g-tensor is substituted into the expression for H_S. After introducing an enhancement parameter $\kappa = \frac{\gamma_e A_\perp}{\gamma_n D} \approx 8.26$, the following Hamiltonian is obtained:

$$\begin{aligned} H_1 =& D\hat{S}_z^2 - \gamma_e B_z \hat{S}_z + A_{||} \hat{S}_z \hat{I}_z - 3\kappa\gamma_n B_x \hat{S}_z^2 \hat{I}_x \\ &- \gamma_n \left(B_z \hat{I}_z + (1 - 2\kappa) B_x \hat{I}_x\right). \end{aligned} \tag{1.14}$$

After moving into an electronic doubly-rotating frame[30] using the unitary transformation

$$\begin{aligned} R_1(t) &\equiv \left(e^{i\lambda_1 t} |m_s = +1\rangle \langle m_s = +1| + |m_s = 0\rangle \langle m_s = 0| + e^{i\lambda_2 t} |m_s = -1\rangle \langle m_s = -1|\right) \otimes \mathbb{1}_n \\ &= \begin{pmatrix} e^{i\lambda_1 t} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & e^{i\lambda_2 t} \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \end{aligned} \tag{1.15}$$

with $\lambda_{1,2} = D \mp \gamma_e B_z$, we obtain the Hamiltonian $\widetilde{H}_1$ given in Eq. (1.16), which is also given in the main text as Eq. (1.2). Experimentally this Hamiltonian is realized using microwave fields to probe transitions $|m_s = 0\rangle \leftrightarrow |m_s = +1\rangle$ and/or $|m_s = 0\rangle \leftrightarrow |m_s = -1\rangle$, equally detuned from the two hyperfine resonances.

$$\begin{aligned}\widetilde{H}_1 &= R_1(t) H_1 R_1(t)^\dagger + i\frac{dR_1(t)}{dt} R_1(t)^\dagger \\ &= \underbrace{A_{||}\hat{S}_z\hat{I}_z - 3\kappa\gamma_n B_x \hat{S}_z^2 \hat{I}_x}_{\text{depends on } m_s} - \underbrace{\gamma_n\left(B_z\hat{I}_z + (1-2\kappa)B_x\hat{I}_x\right)}_{\text{independent of } m_s}\end{aligned} \tag{1.16}$$

An additional rotation using $R_2(t) = e^{i\phi\hat{I}_y}$ further simplifies the Hamiltonian, with $\phi = \tan^{-1}\left(\frac{(1-2\kappa)B_x}{B_z}\right)$. This rotation diagonalizes the nuclear subsystem that is independent of the electron spin, producing the following Hamiltonian:

$$\begin{aligned}\widetilde{H}_2 &= R_2(t)\widetilde{H}_1 R_2(t)^\dagger + i\frac{dR_2(t)}{dt} R_2(t)^\dagger \\ &= -\gamma_n\left(\vec{\beta}_{\text{ind}} + \vec{\beta}(m_s)\right)\cdot\vec{I}.\end{aligned} \tag{1.17}$$

Equation (1.17), also provided in the main text as Eq. (1.3), consists of an effective magnetic field $\vec{\beta}_{\text{ind}} + \vec{\beta}(m_s)$ interacting with the nuclear spin. This vector sum consists of:

1. a magnetic field along the newly defined z'-axis, *independent* of the electronic spin state,

$$\vec{\beta}_{\text{ind}} = \beta_{\text{ind}}\hat{z}' = \sqrt{B_z^2 + (1-2\kappa)^2 B_x^2}\,\hat{z}', \tag{1.18}$$

2. and a spin-*dependent* field

$$\vec{\beta}(m_s) = \begin{bmatrix} \beta_{x'}(m_s) \\ \beta_{y'}(m_s) \\ \beta_{z'}(m_s) \end{bmatrix} = \frac{1}{\beta_{\text{ind}}} \begin{bmatrix} B_x \left(m_s(1-2\kappa)\frac{A_{||}}{\gamma_n} + m_s^2 3\kappa B_z \right) \\ 0 \\ -m_s \frac{A_{||}}{\gamma_n} B_z + m_s^2 3\kappa(1-2\kappa)B_x^2 \end{bmatrix}. \tag{1.19}$$

An additional rotation generated by $\hat{I}_{y'}$ by an angle ϕ_{m_s} diagonalizes the Hamiltonian, depending on the electronic spin state m_s,

$$\phi_{m_s} = \tan^{-1}\left(\frac{\beta_{x'}(m_s)}{\beta_{\text{ind}} + \beta_{z'}(m_s)} \right). \tag{1.20}$$

From Eq. (1.17), the effective nuclear Larmor precession frequency can be obtained,

$$\begin{aligned} \omega_{m_s} = \left| \gamma_n \left(\vec{\beta}_{\text{ind}} + \vec{\beta}(m_s) \right) \right| &= -\gamma_n \sqrt{\left(\beta_{\text{ind}} + \beta_{z'}(m_s) \right)^2 + \beta_{x'}(m_s)^2} \\ &= -\gamma_n \sqrt{B_x^2 \left(1 - 2\kappa + m_s^2 3\kappa \right)^2 + \left(m_s \frac{A_{||}}{\gamma_n} - B_z \right)^2}. \end{aligned} \tag{1.21}$$

For a given electronic spin state m_s, the nuclear spin precesses at a corresponding frequency ω_{m_s}. Note that when $m_s = 0$, this reduces to the spin-independent Larmor frequency $\omega_0 = |\gamma_n \beta_{\text{ind}}|$.

For a single-quantum (SQ) Ramsey sequence, we effectively operate within a 2-level system consisting of $|m_s = 0\rangle$ and either $|m_s = +1\rangle$ or $|m_s = -1\rangle$. To apply a microwave rotation pulse, we define the unitary operator $\hat{U}_x(\theta) = e^{-i\hat{\sigma}_x \theta/2}$ using the 2-level Pauli spin operator $\hat{\sigma}_x$. The SQ Ramsey sequence consists of two $\pi/2$ pulses separated by a free precession period of

duration τ. The initial state $|\psi(0)\rangle = |m_s = 0\rangle$ evolves under the following unitary rotations:

$$\psi_{SQ}(t) = \hat{U}_x\left(\pi/2\right)\left(e^{-i\widetilde{H}_2\tau/\hbar}\right)\hat{U}_x\left(\pi/2\right)|\psi(0)\rangle\,. \tag{1.22}$$

The Ramsey response S_{SQ} can be represented by a projection $P_0 = |m_s = 0\rangle\langle m_s = 0|$ onto the $m_s = 0$ state,

$$\begin{aligned} S_{SQ}(\tau) &= \mathrm{Tr}\left\{P_0\,|\psi_{SQ}(\tau)\rangle\langle\psi_{SQ}(\tau)|\right\} \\ &= \frac{1}{2}\left(1 - \cos\left(\frac{\omega_0\tau}{2}\right)\cos\left(\frac{\omega_{\pm1}\tau}{2}\right) + \cos\left(\Phi_{0,\pm1}\right)\sin\left(\frac{\omega_0\tau}{2}\right)\sin\left(\frac{\omega_{\pm1}\tau}{2}\right)\right). \end{aligned} \tag{1.23}$$

This response represents the electronic spin population and therefore spans between 0 and 1. For tidiness, this form is slightly modified to give the expression in the main text $S_{i,j}(\tau) = 2S(\tau) - 1$, where $S(\tau)$ denotes the response shown in Eq. (1.23). The double-quantum (DQ) response can be similarly obtained by substituting the 2-level spin operator σ_x in the unitary rotation $\hat{U}(\theta)$ with the appropriate spin-1 operator:

$$S_{DQ} = \frac{1}{2}\left[1 + \cos\left(\frac{\omega_{-1}\tau}{2}\right)\cos\left(\frac{\omega_{+1}\tau}{2}\right) - \cos\left(\Phi_{-1,+1}\right)\sin\left(\frac{\omega_{-1}\tau}{2}\right)\sin\left(\frac{\omega_{+1}\tau}{2}\right)\right] \tag{1.24}$$

1.8.2 Experimental Details

A diode-pumped solid-state laser (Lighthouse Photonics, Sprout-D5W) is employed to generate the continuous-wave 532 nm laser beam. The laser beam is passed through an acousto-optic modulator (AOM) (Gooch & Housego, Model: 3250-220) and the first order diffracted beam is used for the experiments. To create optical pulses for NV ensemble initialization and readout, the

RF driver for the AOM is gated by switches (Mini-Circuits, ZASWA-2-50DR+) using transistor-transistor logic (TTL) pulses (Swabian Instruments, Pulse Streamer 8/2). The beam is focused through a 20×/0.75 NA Nikon objective onto the diamond, illuminating the NV-enriched layer over a spot of $\sim 20\,\mu$m diameter. The illumination intensity is kept below 0.1 mW/μm^2. The outgoing spin-dependent NV fluorescence is collected through the same objective, passed through a $150\,\mu$m pinhole (Thorlabs, P150D) to restrict the collection volume, filtered by a 647 nm long pass filter (Semrock) and finally projected onto an avalanche photodiode (Hamamatsu C10508-01). Photodiode voltage measurements are acquired using a DAQ system (National Instruments, NI USB-6363). The bias magnetic field is generated using two permanent samarium cobalt (SmCo) ring magnets to minimize temperature sensitivity of the bias field. Rotation of the magnet pair for different misalignment angles is controlled using two sets of Thorlabs HDR50 rotation stages with stepper motors. Microwave fields are synthesized from up to two signal generators with built-in IQ mixers (Stanford Research System, SG384); and gated by switches and TTLs to provide either single- or double- quantum pulsed control of the NV ensemble spin states. The Ω-shaped planar waveguide for microwave delivery has a diameter $\approx 150\,\mu$m to ensure field homogeneity over the laser illumination area.

1.8.3 Bias Field Misalignment Angle

When the bias magnetic field is misaligned from the NV axis, both the transverse B_x and longitudinal B_z field components affect the energy levels of $|0\rangle, |\pm 1\rangle$. In the main text, the resonance frequency difference $\Delta_{\pm 1}$ between the two transitions $|m_s = 0\rangle \leftrightarrow |m_s = +1\rangle$ and $|m_s = 0\rangle \leftrightarrow |m_s = -1\rangle$ is used to calculate the misalignment angle θ, via the relationship

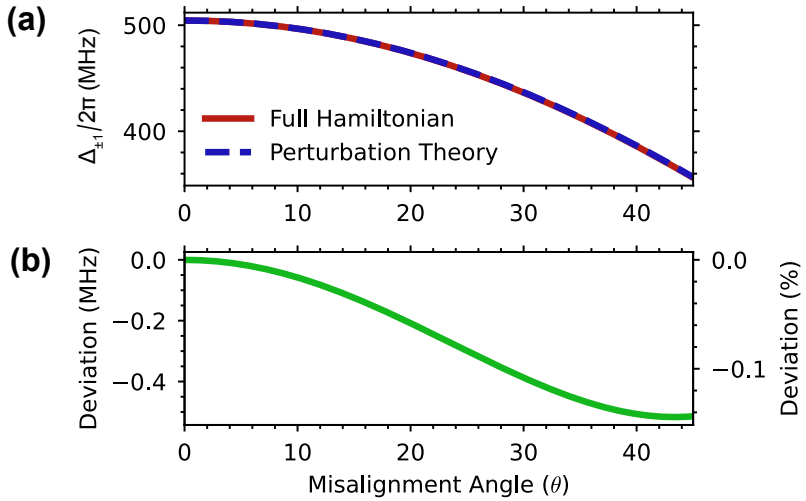


Figure 1.6: Transition frequency difference $\Delta_{\pm1}$ between $m_s = 0 \leftrightarrow +1$ and $m_s = 0 \leftrightarrow -1$ resonances as a function of misalignment angle. Calculation is performed at 90 G. (a) Numerical results of $\Delta_{\pm1}$ using the full NV ground state Hamiltonian (solid red line) and the approximate formula from perturbation theory $2\gamma_e B\cos\theta$ (dashed blue line). (b) Absolute and percentage deviation of $\Delta_{\pm1}$ between the two calculation methods.

$\Delta_{\pm1} \approx 2\gamma_e B_z = 2\gamma_e B\cos\theta$. This approximate relationship is obtained from time-independent perturbation theory. For the small bias field magnitudes $B \lesssim 110\,\mathrm{G}$ used in our study, the large NV zero field splitting $D = 2\pi \times 2870\,\mathrm{MHz}$ dominates the energy scale, $D \gg |\gamma_e B|$. This allows us to treat the transverse field B_x contribution as a perturbation, resulting in the eigenenergies $E_{|0\rangle\to|\pm1\rangle}$ accurate to the second order correction[31]:

$$\frac{E_{|0\rangle\to|\pm1\rangle}}{\hbar} = D + \frac{3\left|\gamma_e B_x\right|^2}{D} \pm \gamma_e B_z. \tag{1.25}$$

Therefore, $\Delta_{\pm1} = \left(E_{|0\rangle\to|+1\rangle} - E_{|0\rangle\to|-1\rangle}\right)/\hbar$ can be accurately modeled by $2\gamma_e B\cos\theta$, up to second order. Using this relationship, the measurements of $\Delta_{\pm1}$ are converted to values of the misalignment angle θ. The error bars for θ in Fig. 1.3 of the main text indicate the corresponding 95% confidence interval (CI), except for cases where the standard deviation σ_θ calculated from error propagation is less than the stage rotation error. In these cases, the nominal accuracy

($\pm 0.047°$) of the stage is used for σ_θ instead.

To verify that the expression for $\Delta_{\pm 1}$ is indeed valid, we compare the measured values of $\Delta_{\pm 1}$ to those obtained by diagonalizing the complete Hamiltonian given in Eq. (1.1). An example plot is shown in Fig. 1.6, for a 90 G bias field magnitude and misalignment angles θ up to 45°. We observe excellent agreement with the results obtained using the full Hamiltonian, with percentage deviation in $\Delta_{\pm 1}$ of under 0.15%. The corresponding error contribution to the angle estimation is less than the dominant error values used for Fig. 1.3 in the main text.

1.8.4 Microwave Frequency Calibration

To set the microwave frequency to the center of the two ^{15}NV hyperfine-split transitions, we perform a two-step calibration procedure combining pulsed-ODMR and Ramsey techniques. First, the two hyperfine transitions are resolved by performing a pulsed-ODMR experiment with low microwave power (π pulse time $\sim$ 900 ns). The resulting spectrum is fit to a sum of two Lorentzians $A - \sum_{i=1}^{2} \frac{C_i}{\pi} \frac{\frac{1}{2}\Gamma_i}{(\nu-\nu_i)^2+(\frac{1}{2}\Gamma_i)^2}$, with free parameters including the contrast C_i, line width Γ_i, frequency ν_i and overall offset A.

The mean of the two Lorentzian center frequencies $\nu_* = (\nu_1 + \nu_2)/2$ is used as the initial value of the applied microwave frequency. Next, a series of Ramsey $\pi/2 - \tau - \pi/2$ experiments are performed to refine this estimate. Fixing τ at a point of maximum contrast, the microwave frequency is varied around ν_* to produce an oscillating magnetometry curve[9]. Choosing a common phase for both microwave $\pi/2$ pulses, the extremum nearest to ν_* is used as the updated center resonance frequency. This calibrated value ν_*^{calib} is obtained by fitting the magnetometry curve to a sinusoidal function and extracting the necessary frequency offset.

1.8.5 Measured Ramsey Time Series Fit Details

For fits to Ramsey time series data, we use a modified form of Eq. (1.6) from the main text:

$$\widetilde{S}_{0,+1}(\tau) = C_0 e^{-(\tau/T_2^*)^p}\Bigg[-\cos\left(\frac{\omega_0}{2}\tau + x_0\right)\cos\left(\frac{\omega_{+1}}{2}\tau + x_{+1}\right) + \cos\left(\Phi_{0,+1}\right)\sin\left(\frac{\omega_0}{2}\tau + x_0\right)\sin\left(\frac{\omega_{+1}}{2}\tau + x_{+1}\right)\Bigg] + B. \tag{1.26}$$

Here, τ represents the free evolution time in a Ramsey sequence. The free fit parameters are the initial contrast C_0, dephasing time T_2^*, stretched exponential factor p, EREEM frequency components ω_0 and ω_{+1}, amplitude modulation factor determined by $\Phi_{0,+1}$, phase offsets x_0 and x_{+1}, and a vertical offset B.

For each estimate of ω_0 or $\chi_{\text{min}} = \cos\left(\Phi_{0,+1}\right)$ reported in Fig. 1.3(b,c) of the main text, its corresponding confidence interval (CI) is constructed using the standard interval method. For each parameter estimate $\hat{\theta}$, its error bar represents a range bounded between $\hat{\theta} \pm z^{(0.95)}\hat{\sigma}$, where $\hat{\theta}$ represents the point estimate of the parameter of interest with standard error $\hat{\sigma}$, and $z^{(0.95)}$ denotes the 95^{th} percentile of the normal deviate. Both $\hat{\theta}$ and $\hat{\sigma}$ are calculated using maximum likelihood estimation via the MATLAB `lsqcurvefit` function. However, the exact CI can differ from the standard interval approximation if the measurement error is not normally distributed. In such scenarios, a bootstrap confidence interval produces a more accurate CI estimation[40]. To test this approach, we applied a bootstrap sampling method to four randomly selected data points from Fig. 1.3(b,c) in the main text. The distribution of fitted parameters resembles a normal distribution with $\sim 95\%$ data points located within $\hat{\theta} \pm z^{(0.95)}\hat{\sigma}$, which supports the reporting of CI estimates obtained by the standard interval method in the main text. A bootstrap fitting example for the

envelope beat frequency ω_0 is shown in Fig. 1.7, for a bias field magnitude $B = 90.08(2)$ G and a misalignment angle $\theta = 25.72(9)^\circ$.

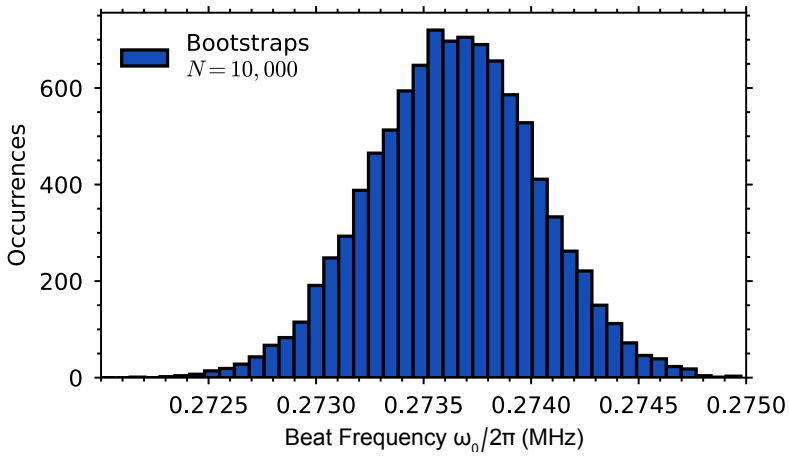


Figure 1.7: Example bootstrap fit result for the EREEM beat frequency. Starting with a single Ramsey time series, $10,000$ additional datasets are created by randomly sampling with replacement from the initial time series. Estimates of the EREEM beat frequency ω_0 are extracted using the MATLAB `lsqcurvefit` function for each dataset. The distribution supports the assumption of normality used by the standard interval method for confidence interval (CI) construction.

A fit to Eq. (1.26) is appropriate when the microwave driving frequency is equally detuned from the two ^{15}NV hyperfine-split transitions, $\left|\delta_{m_I=-\frac{1}{2}}\right| = \left|\delta_{m_I=\frac{1}{2}}\right|$. Alternatively, the EREEM beat frequency ω_0 can be extracted when these detunings are not equal $\left|\delta_{m_I=-\frac{1}{2}}\right| \neq \left|\delta_{m_I=\frac{1}{2}}\right|$. This approach comes at the expense of fitting accuracy due to the presence of extra frequency tones in the Ramsey signal. The Ramsey signal exhibits frequency splittings of magnitude ω_0 centered at both $\left|\delta_{m_I=-\frac{1}{2}}\right|$ and $\left|\delta_{m_I=\frac{1}{2}}\right|$, resulting in four distinct tones. To resolve all frequency components and to avoid ambiguity in the frequency assignments, the following condition is enforced: $\left|\delta_{m_I=-\frac{1}{2}}\right| - \left|\delta_{m_I=\frac{1}{2}}\right| \gg \omega_0$. This alternative measurement is conducted at four randomly selected field configurations, sampled from those shown in Fig. 1.3 of the main text. A four-tone EREEM function modified from Eq. (1.26) is used for fits to the time series data. The extracted envelope beat frequencies overlap with the CI of the corresponding data at the same

field configurations in the main text. Example data is shown in Fig. 1.8 for a bias field magnitude $B = 90.08(2)$ G and misalignment angle $\theta = 25.72(9)°$.

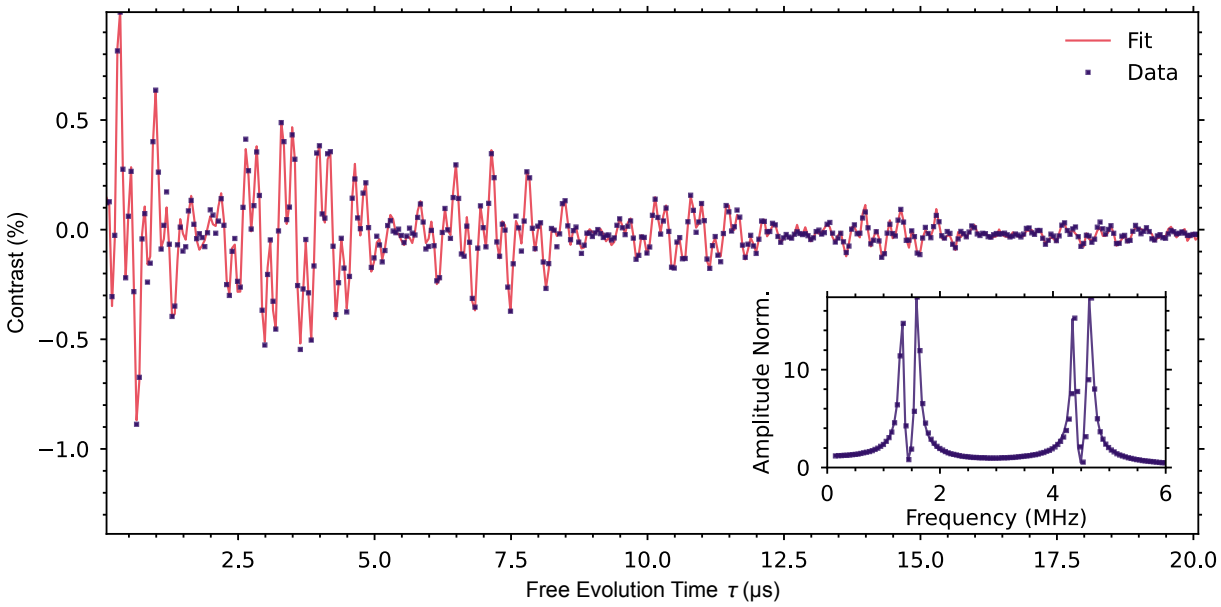


Figure 1.8: Experimental Ramsey time series data obtained using a purposefully detuned microwave driving frequency, with its corresponding amplitude spectrum displayed as an inset. The larger detuning enables clear separation among all frequency components. The fit result for the EREEM beat frequency $\omega_0/2\pi$ is 273.3 ± 1.4 MHz, which overlaps with the confidence interval (CI) of the corresponding beat frequency fit result, 273.6 ± 0.8 kHz, obtained when the microwave frequency is centered between the hyperfine splitting.

1.8.6 Ramsey Measurements Across Misalignment Angles

Figure 1.9 shows experimental Ramsey time series data acquired at a range of bias field misalignment angles θ for a field magnitude of $B = 90.09(2)$ G, in both time and frequency domains. When the magnetic field is aligned to the NV axis ($\theta = 0$), the measured time series oscillate with a single frequency, represented by an individual peak in the amplitude spectrum. After introducing a misalignment, a beating (EREEM) is observed that results in a frequency splitting of magnitude ω_0, which increases as a function of θ. At each field configuration (B, θ),

the time series measurements are fit to Eq. 1.26 to extract the envelope modulation properties. The results of these fits are shown in Fig. 1.3 of the main text, which includes measurements at magnetic field magnitudes $B = 50.49(2)$ G, $69.40(2)$ G, and $110.78(2)$ G.

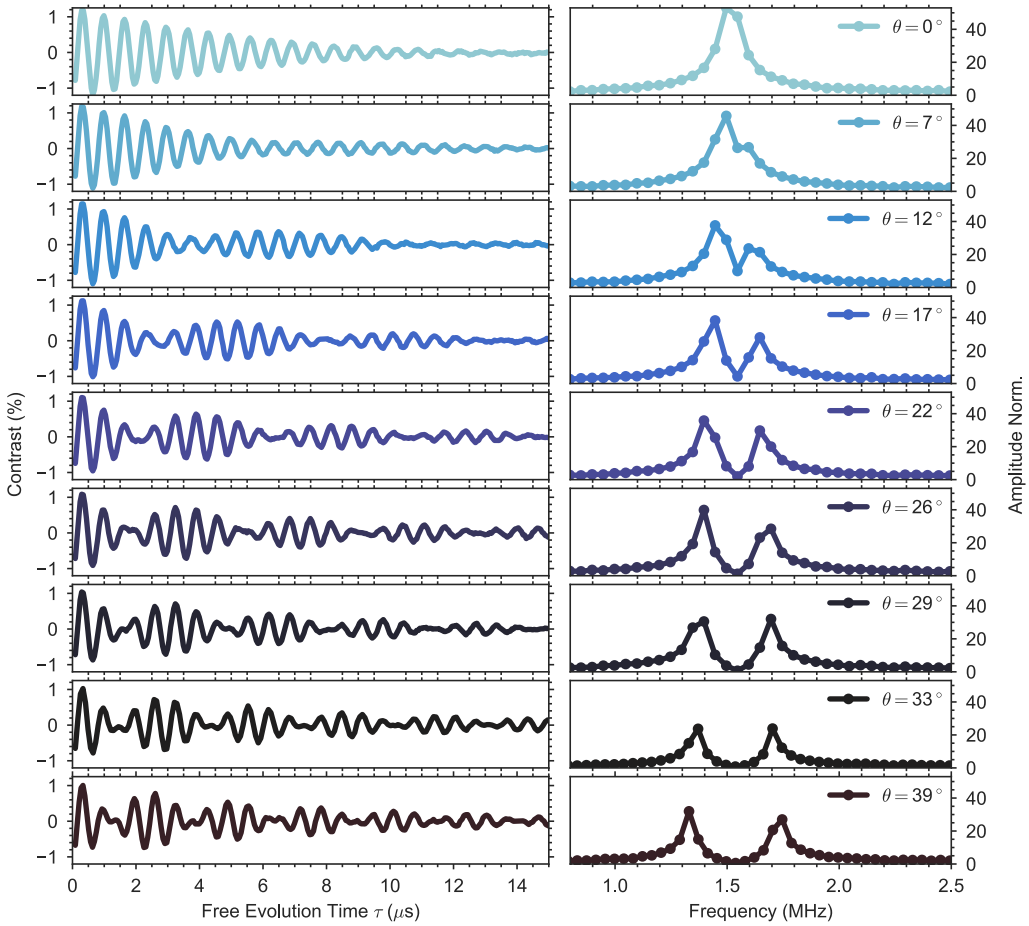


Figure 1.9: Experimental single-quantum Ramsey time series measurements (left column) and corresponding amplitude spectra (right column) at various bias field misalignment angles θ, at a bias field magnitude of $90.08(2)$ G.

1.8.7 QuTiP Simulations

As discussed in Sec. 1.4 of the main text, systematic differences are observed between experimental estimates of the Ramsey envelope beat frequency ω_0, and theoretical predictions from Eq. (1.7) of the main text obtained using second order perturbation theory. This equation is

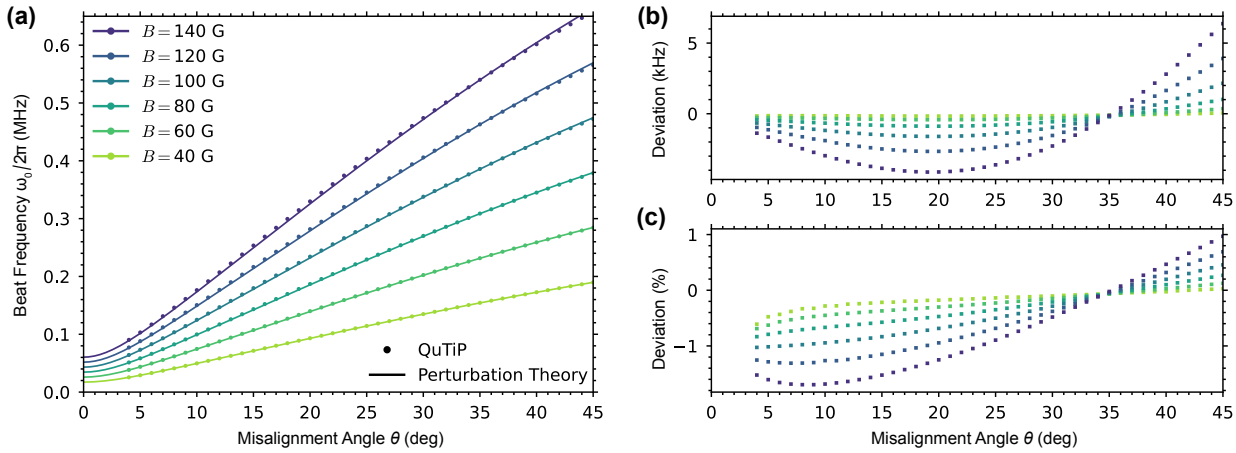


Figure 1.10: Calculation and comparison of Ramsey envelope (EREEM) beat frequency ω_0 as a function of bias magnetic field magnitude and misalignment angle, using analytical approximation and QuTiP numerical simulations. (a) ω_0 obtained using numerical simulations (circles) and analytical perturbation theory (solid line), for magnetic fields between $40\,\mathrm{G}$ and $140\,\mathrm{G}$. (b) Absolute differences between ω_0 estimates from QuTiP numerical simulations and analytical approximation, for the field configurations shown in (a). (c) Percentage deviation between ω_0 estimates from QuTiP numerical simulations and analytical approximation.

reproduced below for clarity.

$$\begin{aligned}\omega_0 = |\gamma_n \beta_{\text{ind}}| &= |\gamma_n| \sqrt{B_z^2 + (1-2\kappa)^2 B_x^2} \\ &= |\gamma_n|\, B \sqrt{1 + 4(\kappa^2 - \kappa)\sin^2\theta}.\end{aligned} \tag{1.7}$$

The observed difference increases with the bias magnetic field misalignment angle. To study this discrepancy, we conduct numerical simulations of Ramsey spin dynamics for a range of bias field configurations, using the QuTiP package in Python. For a coupled NV electron-nuclear system initially described by the lab frame Hamiltonian in Eq. (1.1) of the main text, applied microwave pulses are modeled using time-dependent contributions $\Omega(t)\cos(\omega_e t + \varphi)\hat{S}_x$. The AC magnetic field amplitude $\Omega(t)$ is chosen to yield a 20 MHz Rabi frequency during pulses. The duration of each pulse is calibrated so a $\pi/2$ rotation is performed. The phase φ is used to toggle between pulse rotation axes.

To evaluate the state evolution during the Ramsey sequence, we employ the master equation solver function $mesolve$ provided by the QuTiP module. To adjust the accuracy of the ordinary differential equation solver used for these calculations, absolute and relative tolerance values are supplied to the function, denoted by $atol$ and $rtol$, respectively. These user-defined tolerance values are determined empirically by comparing the accuracy of the NV system dynamics after several Rabi cycles, with values $atol = rtol = 10^{-9}$ chosen for all Ramsey simulations. The system state is recorded at regular 10 ns time steps during the Ramsey free evolution interval, but the solver is allowed to perform up to 10^6 additional evaluations within each time step. Note that all timescales and frequencies are scaled by 6 orders of magnitude for simplicity of use (e.g., setting time = 1 represents 1 μs).

For each field configuration, the lab frame Hamiltonian H is first diagonalized to obtain eigenenergies and eigenstates. The pair of frequencies for transitions between $|m_s = 0, m_I = +1/2\rangle \leftrightarrow |1, +1/2\rangle$ and $|0, -1/2\rangle \leftrightarrow |1, -1/2\rangle$ are calculated. The mean of these two values is used as the frequency ω_e of the pulses. Starting with an initial state $|0, -1/2\rangle$, the Ramsey pulse sequence is applied and the final population in the $m_s = 0$ state is determined. Pulse sequences are simulated for a range of free evolution times τ, and the resulting time series is fit to Eq. (1.23) to extract the parameters ω_0, $\omega_{\pm1}$ and $\Phi_{0,\pm1}$. In Fig. 1.10, the values of ω_0 extracted from numerical simulations are compared to the analytical predictions given by Eq. (1.7). For the field configurations considered in the main text, the difference between these calculations remains around or below 1%.

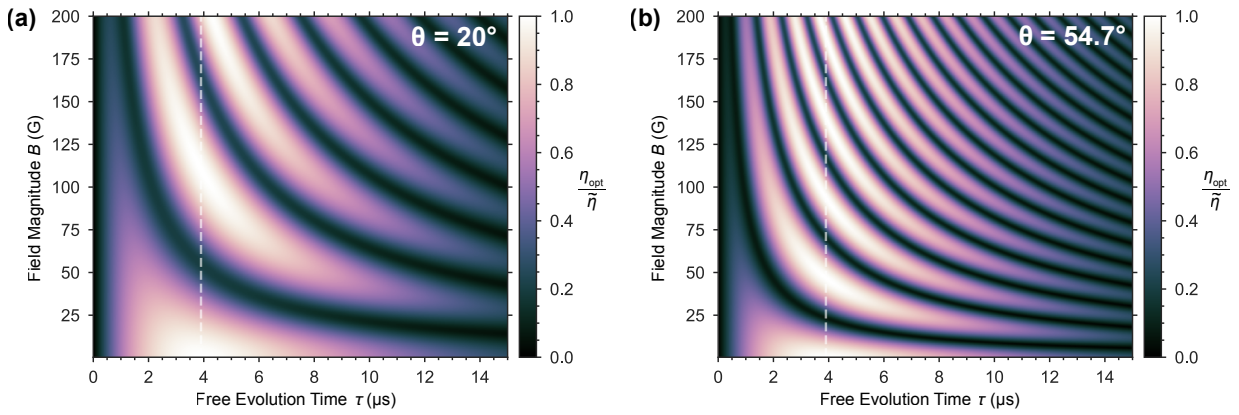


Figure 1.11: 2D color plots of the normalized inverse sensitivity with respect to the Ramsey free evolution time τ and bias magnetic field magnitude B, at 2 distinct misalignment angles $\theta = 20°$ (a) and $54.7°$ (b)

1.8.8 Normalized Inverse Sensitivity

In Fig. 1.4(c) of the main text, a 2D color plot of the relative normalized inverse magnetic field sensitivity is shown, evaluated using Eq. (1.9). These calculations are repeated here for two other misalignment angles $\theta = 20°, 54.7°$ and are shown in Fig. 1.11. The experimental timescales assumed here ($T_D = 5\,\mu s$, $T_2^* = 5\mu s$) are the same as those used in the main text. In particular, the $\theta = 54.7°$ case can describe a bias magnetic field oriented to have equal projections onto all four NV quantization axes.

1.8.9 Transverse Hyperfine Parameter

As discussed in Sec. 1.4 of the main text, we perform fits of Eq. (1.7) to experimental measurements of ω_0 while allowing the transverse hyperfine parameter $A_\perp$ to vary as a degree of freedom. We repeat this process at a series of bias field magnitudes, extending beyond those considered in the main text. The resulting fits for $A_\perp$ are shown in Fig. 1.12 using black markers, with error bars representing a 95% confidence interval (CI). We observe an overall decrease of

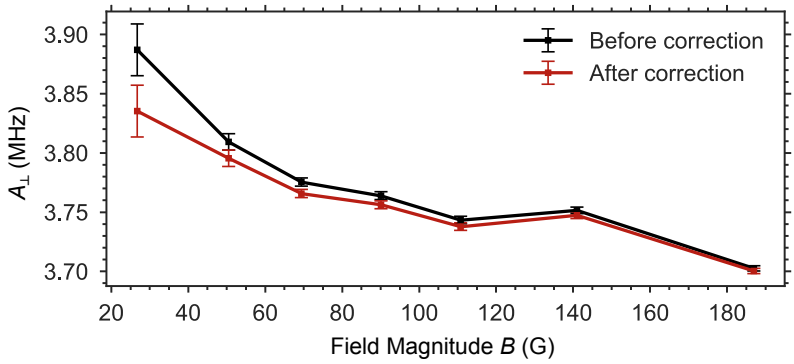


Figure 1.12: Results for the transverse hyperfine parameter $A_\perp$ as a fit parameter to the Ramsey envelope (EREEM) beat frequency ω_0, extracted from experimental data, for a range of magnetic fields, with error bars showing the 95% confidence interval, shown in black. Estimated corrections to values of $A_\perp$ are depicted in red, using the discrepancy between QuTiP numerical simulations and theoretical predictions of ω_0.

$A_\perp$ as the field magnitude is increased.

Due to the field-dependent discrepancy of ω_0 between QuTiP simulations and the predictions of perturbation theory, additional corrections are necessary to precisely determine $A_\perp$. However, as shown in Fig. 1.10, the degree to which these predictions differ depends on the misalignment angle θ. As an initial attempt to estimate these corrections, we fit Eq. (1.7) to the estimates of ω_0 obtained using QuTiP simulations, at the specific magnetic field magnitudes and misalignment angles studied in experiment. The absolute difference between these fits and the originally assumed value for $A_\perp$ is used to update the experimentally determined $A_\perp$ estimates, shown in Fig. 1.12 as red points. We expect that future studies can further refine the estimation of $A_\perp$ by increasing the sampling of misalignment angles at each magnetic field magnitude. In addition, improvements to the electronic dephasing time T_2^* can enhance the measurement precision of ω_0 (and therefore $A_\perp$) at lower fields. For example, this may be achieved by optimizing the diamond sample or using techniques such as spin-bath driving[31]. Alternatively, an extension of these studies to higher bias field magnitudes will allow direct comparisons to the previously

conducted EPR study[34].

1.8.10 Double-quantum Envelope Amplitude

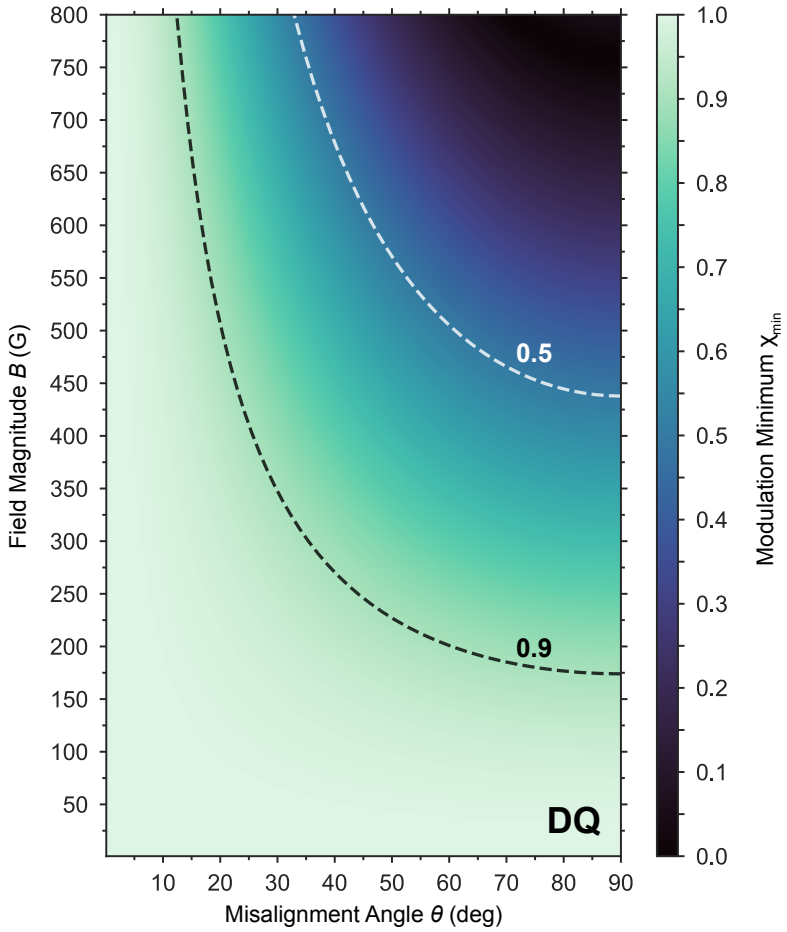


Figure 1.13: Calculated double-quantum (DQ) Ramsey relative contrast at amplitude modulation nodes $\chi_{\min} = |\cos(\Phi_{-1,+1})|$, for magnetic field magnitudes B up to 800 G and misalignment angles up to $90°$. Dashed contour lines indicate magnetic field configurations where $\chi_{\min}$ equals 0.9 and 0.5, representing 10% and 50% contrast loss at envelope nodes, respectively.

In Fig. 1.5(c) of the main text, 2D color plots of the relative contrast at amplitude modulation nodes $\chi_{\min}$ are shown for both single-quantum (SQ) and double-quantum (DQ) Ramsey. The calculations for DQ Ramsey are reproduced here, with the range of magnetic field magnitudes and misalignment angles extended up to 800 G and $90°$, respectively. For the magnetic

field magnitudes considered in this study $B \lesssim 200\,G$, the contrast at envelope modulation nodes is well-preserved ($\chi_{\text{min}} \approx 1$). However, the calculations illustrated in Fig. 1.13 show that for a sufficiently large magnetic field, the DQ Ramsey signal can exhibit significant envelope modulations.

Chapter 2: Beyond Average Hamiltonian Theory for Quantum Sensing

This chapter has been adapted from the preprint titled

***Beyond Average Hamiltonian Theory for Quantum Sensing**,*

with authors Jner Tzern Oon, Sebastian C. Carrasco, Connor A. Hart, George Witt, Vladimir S. Malinovsky, and Ronald Walsworth,

posted on the Arxiv[46].

My contributions include project conception, experimental validation, theory and simulation developments, manuscript preparation, and supervision of all project components. I am proud to share a statement of equal contribution with Sebastian C.Carrasco, who was essential to developing the proof that rapid echo structure results in Magnus expansion convergence, on top of active contribution to numerical simulations and theoretical insights. Connor A. Hart worked with me on experimental validation of AHT breakdown, and provided guidance during early stages of the project. George Witt assisted with development of the proof and early feedback on the manuscript. I drafted the initial manuscript, with assistance from Sebastian C. Carrasco, Connor A. Hart, George Witt, Vladimir S. Malinovsky, and Ronald Walsworth. The work was performed under the supervision of Vladimir S. Malinovsky and Ronald L. Walsworth.

2.1 Introduction

The Magnus expansion offers a general solution to a set of non-autonomous linear differential equations involving linear operators[47]. As a result, it is naturally suited for studying a wide range of phenomena within physics, finding frequent use in classical mechanics[48], quantum field theory[49], atomic physics[50,51], and elsewhere[52]. Considering its ubiquity, studies of the conditions under which the Magnus expansion is valid have widespread implications[53].

Notably, the Magnus expansion forms the mathematical foundation of average Hamiltonian theory (AHT)[54], which was initially developed for studies of nuclear magnetic resonance (NMR)[55,56]. By expressing the solution of the time-dependent Schrödinger equation in an exponential form, AHT approximates time evolution via an effective time-independent Hamiltonian obtained using the Magnus expansion[47]. Besides providing accurate numerical predictions in the perturbative limit, AHT retains hermiticity at any truncation order, while often preserving qualitative physical properties of system dynamics[57]. Over several decades, AHT has seen prolific use in the development of pulse sequences for decoupling of spin-spin interactions; and is a powerful tool for both analytical and numerical predictions of closed quantum dynamics[58,59,60]. Experimentally, such pulse sequences typically involve the time-dependent application of resonant or near-resonant radio frequency, microwave, and/or optical fields suitable for manipulation of the internal (e.g., spin) states of the system of interest.

Recently, the application of AHT has been extended to the field of quantum sensing, where its use ranges from semi-quantitative analysis of pulse sequences for dipolar decoupling[14,61] to forming the mathematical foundation of novel sensing sequences for strongly interacting spin systems[62,63,64]. Notably, research into the latter has contributed to experimental demonstrations

of improved coherence times and magnetic field sensitivity for dense ensembles of nitrogen-vacancy (NV) centers in diamond[65,66].

Here, we show that the conditions that guarantee Magnus expansion convergence conflicts with common operating conditions for magnetometry with solid-state spins (including NV-diamond), leading to a breakdown in AHT. Beyond the regime of guaranteed convergence, we demonstrate that specific symmetries allow pulse sequences to retain Magnus expansion accuracy. Moreover, we present an exact method to evaluate the sensor response to a target classical field, which bypasses the limitations of AHT. We demonstrate the utility of this exact method on widely used pulse sequences for both broadband (DC) and narrowband (AC) magnetometry. This work opens up possibilities for pulse sequence engineering beyond the AHT regime; identification of further symmetries that preserve AHT convergence; and development of numerical techniques that accurately characterize sensor performance.

2.2 Average Hamiltonian Theory

To describe a general pulse sequence applied to a quantum system, we consider a series of n control pulses with unitary operators $P_1, \cdots, P_n$. Each pulse is immediately followed by free evolution of the quantum system for duration $\tau_1, \ldots, \tau_n$, respectively. With each applied pulse, an initial Hamiltonian H_0 is rotated to be

$$H_i = (P_i \ldots P_1)^\dagger H_0 (P_i \ldots P_1), \tag{2.1}$$

such that state evolution during the i-th free evolution frame is given by the piecewise constant Hamiltonian H_i. For a total sequence duration $t = \tau_1 + \tau_2 \ldots \tau_n$, AHT approximates the unitary

propagator as [1]

$$\mathcal{U}^{(m)}(t) = \exp\left(-i\mathcal{H}^{(m)}t\right), \tag{2.2}$$

where $\mathcal{H}^{(m)}$ denotes a time-independent effective Hamiltonian described by the Magnus expansion,

$$\mathcal{H}^{(m)} = \bar{H}^{(1)} + \bar{H}^{(2)} \cdots + \bar{H}^{(m)}. \tag{2.3}$$

For reference, the first two terms in the expansion are provided in Eq. (2.4). The first order contribution $\bar{H}^{(1)}$ describes the mean of the interaction frame Hamiltonians H_i, weighted by the duration of each corresponding free evolution interval τ_i. Higher order terms consist of summations over increasing numbers of nested commutators.

$$\begin{aligned} \bar{H}^{(1)} &= \frac{\sum_i^n H_i\tau_i}{\sum_i^n \tau_i} \\ \bar{H}^{(2)} &= \frac{1}{2i\sum_{i=1}^n \tau_j} \sum_{j=1}^{n}\sum_{i=1}^{j-1} \left[H_j, H_i\right]\tau_j\tau_i \end{aligned} \tag{2.4}$$

2.3 AHT Breakdown

To characterize the effect of a specific pulse sequence on sensor performance, we define the sensor coupling factor $\gamma = \partial\Delta/\partial V$, which represents a first order shift in the sensor energy level difference Δ with respect to a target (classical) signal field V to be measured[4]. Sensitivity

[1] Multiplicative factors of the reduced Planck constant $\hbar$ are dropped in the exponent for simplicity. All Hamiltonian operators have non-angular frequency units in this work.

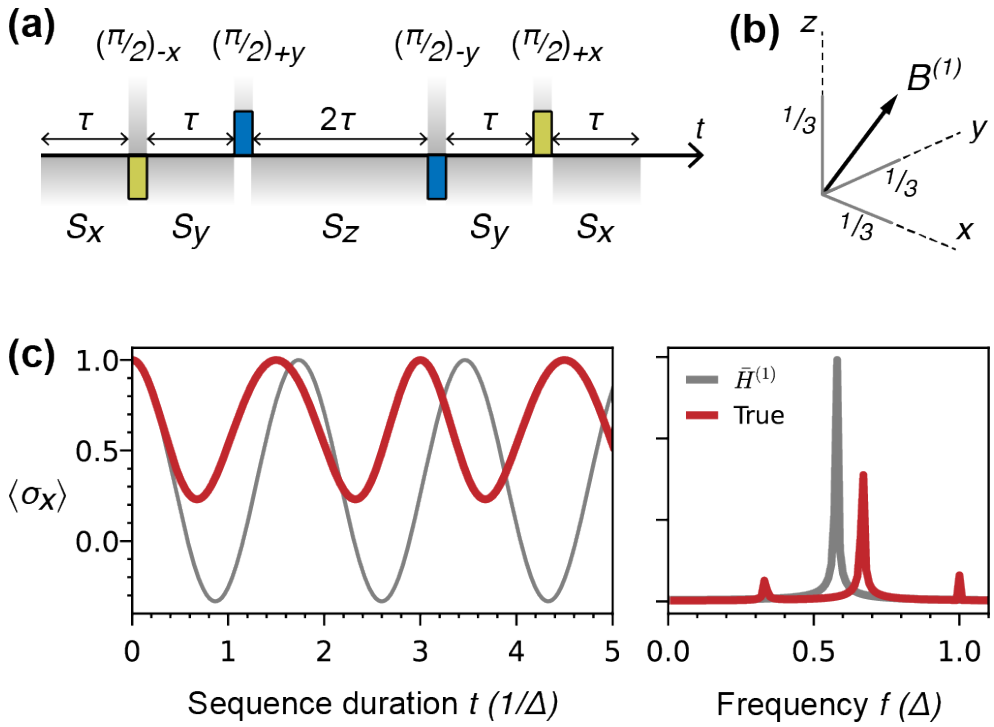


Figure 2.1: Fig 1 (a) WAHUHA pulse sequence and toggling frame representations. (b) Effective field corresponding to leading order average Hamiltonian. (c) WAHUHA time series $\langle \sigma_x(t) \rangle$ and corresponding power spectra, obtained using the first order average Hamiltonian (gray) vs ground truth unitary evolution (red).

to changes in V scales linearly with respect to γ, roughly following

$$\eta \propto \frac{1}{\gamma \sqrt{T_d}},$$

where T_d refers to the characteristic timescale over which a measurement observable decays (for example, due to dephasing or decoherence).

For accurate assessments of sensitivity, reliable estimates of γ are crucial. For example, when using electronic spins for magnetic field measurements, γ represents an "effective" gyromagnetic ratio, and can be defined with respect to the electron gyromagnetic ratio $\gamma_e \approx -2.8 \cdot 10^{10}$ Hz/T. We therefore introduce a dimensionless coupling factor $\alpha = \gamma/\gamma_e$, which

represents a suppression of the bare electron response that depends on the sensing protocol employed. In other words, α is fixed for a given sensing protocol, even if the sensor platform (e.g., electronic spins) is changed.

For a given sensing protocol, previous approaches utilized AHT to estimate α[64,65]. To evaluate the sensor response due to a target Hamiltonian H_0, the first order average Hamiltonian $\mathcal{H}^{(1)}$ is projected onto the relevant basis operators, e.g., the spin operators S_x, S_y, S_z. The components of α (normalized by H_0) are calculated using

$$a_{x,y,z} = \frac{\mathrm{Tr}(\mathcal{H}^\dagger S_{x,y,z})}{\mathrm{Tr}(H_0^\dagger H_0)}, \tag{2.5}$$

such that $\alpha = \sqrt{a_x^2 + a_y^2 + a_z^2}$.

To provide intuition for α, we consider the highly influential WAHUHA pulse sequence, which was originally conceived by Waugh et al.[67] to decouple homonuclear magnetic dipolar interactions in solid-state NMR. As shown in Fig 2.1(a), the WAHUHA protocol consists of four $\pi/2$-rotation pulses applied along $-S_x, S_y, -S_y, S_x$, where $S_{x,y,z}$ denote spin operators along the x, y, z directions. The pulses are accompanied by intervals of free evolution, expressed in units of time τ. By spending equal time along all three orthogonal magnetic field directions, WAHUHA generates a vanishing first order average Hamiltonian $\bar{H}^{(1)} = 0$ for spins coupled by dipolar interactions.

Applying WAHUHA to a target magnetic field $\vec{B}_0 = B_0\hat{z}$ such that $H_0 = -\gamma_e B_0 S_z$, Eq. (2.4) reveals the first order average Hamiltonian,

$$\mathcal{H}^{(1)} = -\gamma_e \frac{B_0}{3}(S_x + S_y + S_z). \tag{2.6}$$

Here, $\mathcal{H}^{(1)}$ is described by an effective field $\vec{B}^{(1)} = (B_0/3, B_0/3, B_0/3)$, as illustrated in Fig. 2.1(b). Setting $\mathcal{H} = \mathcal{H}^{(1)}$ in Eq. (2.5) results in $\alpha = |\vec{B}^{(1)}|/|\vec{B}_0| = 1/\sqrt{3}$. Indeed, α denotes the fractional length of the effective field associated with $\mathcal{H}$, relative to the corresponding field of the target Hamiltonian H_0 [2].

The leading order result in Eq. (2.6) predicts state precession around the effective field $\vec{B}^{(1)}$ with a single-tone frequency $\Delta/\sqrt{3}$, where $H_0 = \Delta S_z$. In Fig 1(c), we show the resulting spin dynamics due to $\mathcal{U}^{(1)}$ and compare to the time evolution using the exact unitary,

$$U_0 = \left(e^{-iH_0\tau_n}P_n\right)\cdots\left(e^{-iH_0\tau_2}P_2\right)\left(e^{-iH_0\tau_1}P_1\right). \tag{2.7}$$

Here, we assume instantaneous pulses applied to an initial state prepared along S_x. The time-dependent state projection onto S_x is plotted as a function of time t, evolved separately using unitaries $\mathcal{U}^{(1)}$ (gray curve) and U_0 (red curve). Immediately, we observe significant differences. In contrast to the single frequency oscillation $\Delta/\sqrt{3}$ predicted by $\bar{H}^{(1)}$, the true time series exhibits beatings with multiple tones, with none at the frequency $\Delta/\sqrt{3}$. This behavior is also clearly seen in the accompanying power spectra in Fig. 2.1(c). The amplitude of the WAHUHA fringes also greatly differ from AHT predictions.

2.4 Magnus Expansion Convergence

To investigate the discrepancy between the AHT prediction of time evolution and the exact dynamics, we include higher order contributions to calculate $\mathcal{U}^m$, for the same target Hamiltonian

[2] Alternatively, Ref. [68] calculates α by considering the largest difference between the eigenvalues of $\bar{H}^{(1)}$, which gives the same result for a two-level system.

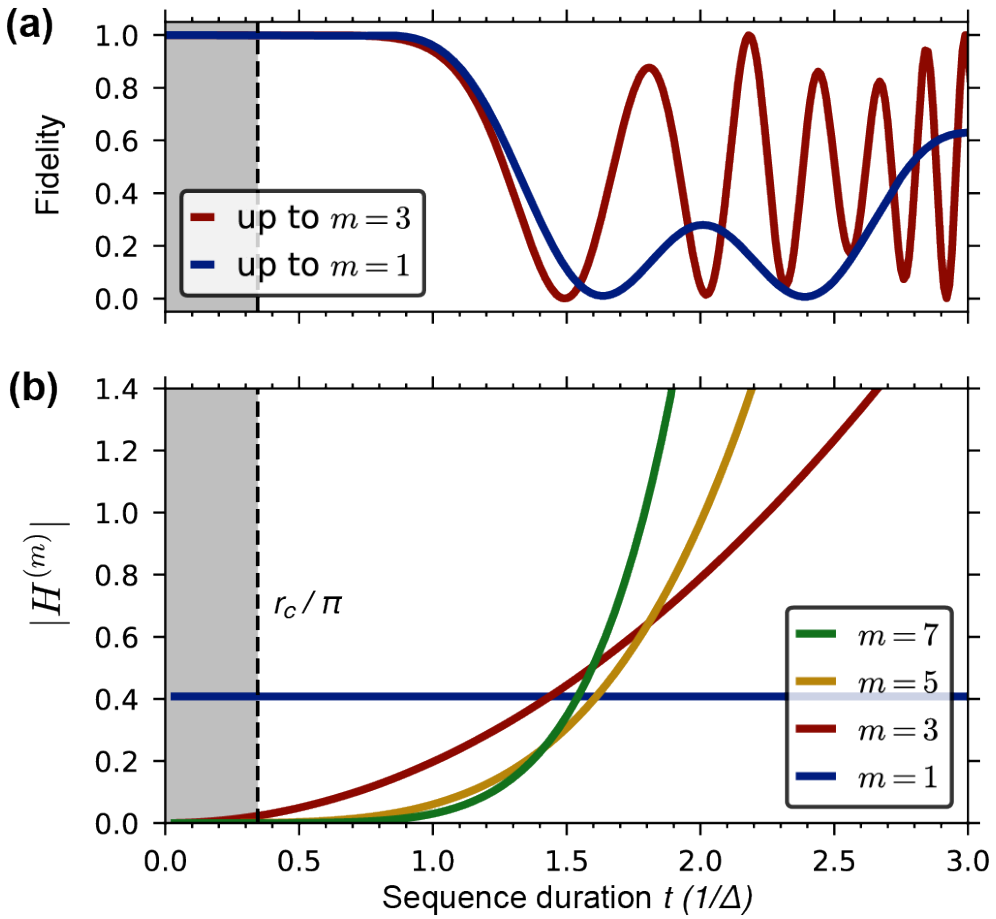


Figure 2.2: Fig 2 (a) State fidelity between exact and AHT evolution for varying sequence duration t, where $\mathcal{U}^m(t)$ is calculated for $m = 1, 3$ and compared to the ground truth unitary $U_0(t)$. (b) Spectral norm of Magnus expansion terms $|\bar{H}^m|$ for $m = 1, 3, 5, 7$ for the WAHUHA protocol, for varying sequence duration t. The region of guaranteed Magnus series convergence $t < r_c/\pi$ is shaded in gray.

and initial state $|+x\rangle$ as above. Fig. 2.2(a) shows a comparison of state fidelity between exact and AHT evolution[69], Notably, as the sequence length t is extended, $\mathcal{U}^{(3)}$ produces fringes with increasing precession frequency, which suggests that the energy scale of $\bar{H}^{(3)}$ is increasing with time t. This observation prompts us to consider the conditions where the Magnus expansion is expected to be valid.

Typically, sufficient conditions for convergence of the Magnus expansion are established using a radius of convergence r_c[52,53]. For the pulse sequences described here, this condition can

be summarized as

$$\sum_{i}^{n} |H_i|\tau_i < \frac{r_c}{2\pi}, \tag{2.8}$$

where $|\cdots|$ represents the spectral norm. The pulses generate a Hamiltonian that is piecewise constant during each free evolution interval, and the spectral norm $|H_i| = \Delta/2$ remains unchanged. Using the upper bound $r_c \approx 1.087$ from Refs.[70,71], the condition in Eq. (2.8) simplifies to

$$\Delta \cdot t < \frac{r_c}{\pi} \approx 0.346. \tag{2.9}$$

In order to guarantee convergence, the energy scale Δ of the interaction frame Hamiltonian and the total sequence duration t can be chosen so their product does not exceed this threshold.

Although Eq. (2.9) assures convergence for the series expansion as a whole, it is common practice to truncate the series to the first few terms due to the increasing complexity of the higher order Magnus expansion. To compare the relative contributions from these leading terms, the spectral norms $|\bar{H}^{(m)}|$ for the WAHUHA sequence are plotted in Fig 2.2(b) as a function of sequence duration t, for m up to 7[36,72]. As expected from Eq. (2.4), $|\bar{H}^{(1)}|$ is constant with respect to the sequence duration t, since the weighted average of the toggling frame Hamiltonians is unchanged. However, the norms of higher order contributions increase with t, and eventually dominate. Specifically, we observe that $|\bar{H}^{(m)}| \propto (\Delta \cdot t)^{m-1}$, which follows from the m nested summations (or more generally, time integrals) that are present in $\bar{H}^{(m)}$.

With these results in mind, we consider the typical operating regime for sensing, with NV centers in diamond as an example. These electronic spin defects are generally dominated by magnetic disorder within the diamond host, with spin resonance linewidths ranging from $\sim 10^4$ Hz to upwards of 10^6 Hz, which inform the typical magnitude of observable changes in Δ. More-

over, the presence of NV hyperfine interactions with both onboard nitrogen and neighboring ^{13}C nuclear spins in the diamond lattice often result in additional 10^6 Hz-scale frequency detunings during control pulses, which may even be exploited to improve signal contrast[66]. Optimal sequence durations t for sensing are typically limited by dephasing/decoherence times that frequently exceed 10^{-6} s. In short, pulsed sensing experiments in these systems routinely operate in a regime where Magnus series convergence is not guaranteed.

2.5 Exact Sensor Coupling Factor

Considering the limitations of AHT outside the radius of convergence, we propose a method to calculate the exact (dimensionless) coupling factor, denoted here as α_ϵ. First, we consider deviations from the ground-truth unitary evolution by introducing a perturbation field V (i.e., a target signal field to be measured) to the internal Hamiltonian H_0 in Eq. (2.7). The resulting unitary evolution operator is given by

$$U_1 = \left(e^{-i(H_0+V_n)\tau_n} P_n\right) \cdots \left(e^{-i(H_0+V_1)\tau_1} P_1\right). \tag{2.10}$$

We choose V to be a small static (DC) perturbation field such that $V_k = \epsilon S_z$, with ϵ as a tunable scalar coefficient. Next, we define U_ϵ such that $U_1 = U_\epsilon U_0$, or equivalently,

$$U_\epsilon = U_1 U_0^\dagger. \tag{2.11}$$

Note that $U_1 = U_0$ in the absence of a signal field ($\epsilon = 0$), reducing U_ϵ to the identity operator. Therefore, any non-trivial rotation generated by U_ϵ is due to a nonzero signal contribution from

V.

Since U_ϵ is unitary, it can be expressed in an exponential form $U_\epsilon = e^{-i\Omega(t)}$, where $\Omega(t)$ denotes an effective time-dependent operator to be determined. After diagonalizing U_ϵ to give the operator $\mathcal{D}(t) = \mathcal{P}^{-1}U_\epsilon\mathcal{P}$, we evaluate the principal logarithm of its eigenvalues along the diagonal of $\mathcal{D}(t)$ to obtain the exponent

$$\Omega(t) = i\mathcal{P}\log[\mathcal{D}(t)]\mathcal{P}^{-1}. \tag{2.12}$$

To ensure that $\Omega(t)$ represents the perturbative response, it is necessary to tune ϵ, for example, by ensuring $|\epsilon t| \ll 1$[73]. In line with AHT, we then extract an effective Hamiltonian H_ϵ by setting $\Omega(t) = H_\epsilon t$. However, we allow H_ϵ to retain a functional dependence on the sequence duration t. After normalizing the projection of $H_\epsilon(t)$ with respect to the perturbation, we obtain the exact coupling factor

$$\alpha_\epsilon(t) = \frac{\sqrt{2\sum_{i\in\{x,y,z\}}\mathrm{Tr}(H_\epsilon^\dagger S_i)^2}}{\sqrt{|\mathrm{Tr}(V^\dagger V)|}}. \tag{2.13}$$

As a check, we apply this method to a Ramsey sequence, which consists of a single toggling frame Hamiltonian $H_1 = S_z$. In this case, Eq. (2.10) simplifies to $U_\epsilon = e^{-i(S_z+\epsilon S_z)\tau}e^{+i(S_z)\tau} = e^{-i\epsilon S_z\tau}$ to reveal the effective Hamiltonian $H_\epsilon(t) = \epsilon S_z$, resulting in an exact coupling factor of $\alpha_\epsilon = 1$ that agrees with AHT.

For the WAHUHA sequence, we plot $\alpha_\epsilon(t)$ as a function of sequence duration t in Fig 3, along with individual projections along S_x, S_y, S_z, for a nonzero initial detuning of 1 MHz. First, we note that $\alpha_\epsilon(t) = 1/\sqrt{3}$ as the sequence duration approaches zero, agreeing with the predictions due to leading order AHT in this regime. This result is expected, as we have shown

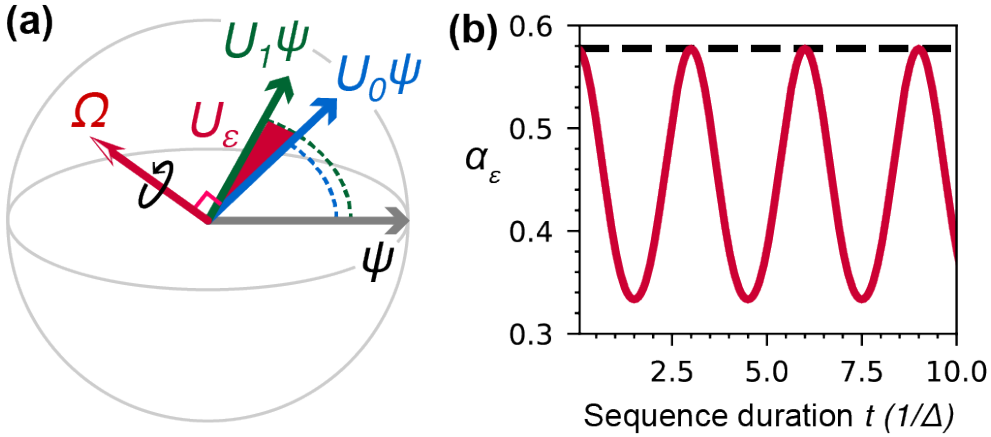


Figure 2.3: (a) An initial state ψ is transformed by a pulse sequence unitary propagator to $U_0\psi$. In the presence of a perturbation field V (i.e., a target signal field to be measured), the final state is $U_1\psi$. The operator Ω generates a unitary operator $U_\epsilon = e^{-i\Omega t}$ such that $U_1 = U_\epsilon U_0$. (b) Exact coupling factor α_ϵ for the WAHUHA pulse sequence, as sequence duration t is increased.

in Fig. 2.2 that AHT yields the correct unitary propagator for sufficiently small timescales. However, as a function of sequence duration, $\alpha_\epsilon(t)$ exhibits oscillations between values of $1/3$ and $1/\sqrt{3}$. Therefore, achieving the dimensionless coupling factor $1/\sqrt{3}$ (predicted by $\mathcal{H}^{(1)}$) requires careful choice of the sequence duration with respect to the energy scale of the interaction frame Hamiltonian Δ.

2.6 Retaining Convergence with Rapid Echoes

Next, we extend this method to determine α_ϵ in response to an oscillating (AC) target signal field. For simplicity, we describe the target AC field as a square wave perturbation that is present during the free evolution intervals, as illustrated in Fig. 2.4(a). By substituting $V_k = (-1)^k \epsilon S_z$ into Eq. (2.10) and following the steps in Eqs. (2.11-2.13), we determine the effective time-dependent operator Ω due to this AC perturbation, and the corresponding α_ϵ.

We then apply this analysis to three pulse sequences XY8, DROID-60 and WAHUHA+Echo[63].

Their toggling frame transformations are shown in Fig 2.4(a), for an initial static magnetic field along $+S_z$. All three sequences compensate each free precession interval around a given axis with an equal period in the opposite direction. Such "echo" structures generate a null first order average Hamiltonian, $\bar{H}^{(1)} = 0$, suppressing contributions from the static field but retaining a response to the target AC magnetic field.

In particular, XY8 and DROID-60 are commonly employed for high sensitivity AC magnetometry in NV ensembles[66]. In the case of XY8, which consists of a regular series of π pulses, all toggling frame Hamiltonians commute ($[H_i, H_j] = 0$ for all i, j) to give $\alpha_\epsilon = 1$, analogous to the Ramsey example discussed above. In contrast, both DROID-60 and WAHUHA+Echo cycle between three orthogonal field orientations to give the first-order AHT prediction $\alpha = 1/\sqrt{3}$ in response to an AC target signal field[65]. For each sequence, the exact coupling factor α_ϵ is shown in 2.4(b) as a function of sequence duration. Although all three sequences retain a response to AC magnetic fields, only AHT predictions for XY8 and DROID-60 align with the exact calculation ($\alpha_\epsilon(t) \approx \alpha$). Here, exiting the radius of convergence established in Eq. (2.9) does not signal a breakdown of the AHT prediction. However, in the case of WAHUHA+Echo, α_ϵ clearly oscillates as t is increased, deviating from the predictions of AHT.

We identify a common property of XY8 and DROID-60 that ensures the accuracy of AHT beyond the radius of convergence, namely the presence of a "rapid" echo structure. Specifically, we refer to rapid echoes as having pairs of nearest-neighbor toggling frame Hamiltonians with opposite sign, such that

$$H_{2l} = -H_{2l-1} \tag{2.14}$$

for all positive integers l. For any sequence that obeys this rapid echo structure, we show that all

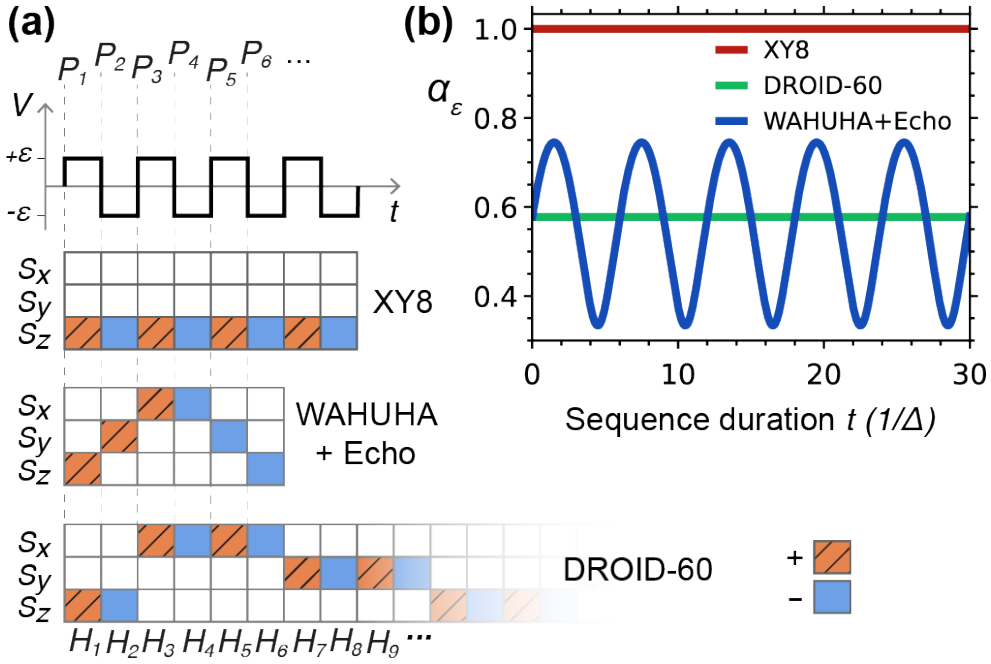


Figure 2.4: (a) Square wave perturbation with amplitude ϵ, changing sign with each toggling frame. Each pulse P_i is applied to the quantum sensor (e.g., NV electronic spins in diamond) before the i-th frame. For an initial sensor spin Hamiltonian S_z, the resulting toggling frame Hamiltonians for pulse sequences XY8, WAHUHA+Echo and DROID-60 are shown. (b) For each sequence, the exact coupling factor α_ϵ is plotted as a function of sequence duration t. Calculations for XY8 and DROID-60 agree with predictions obtained from first-order AHT, i.e., $\bar{H}^{(1)}$.

Magnus expansion contributions are zero (see Supplementary Material[36]). The inclusion of a small AC signal field only introduces a $\sim(\epsilon t)^m$ contribution into each term $\bar{H}^{(m)}$; thus the system dynamics can be effectively described using only $\bar{H}^{(1)}$. To summarize, Hamiltonian dynamics that commute at all times (such as with XY8 and Ramsey) and the presence of rapid spin echoes (DROID-60 and XY8) are examples of special cases where the Magnus expansion remains accurate beyond the sufficient radius of convergence.

2.7 Outlook

Recent applications of average Hamiltonian theory (AHT) range from robust dynamical decoupling and sensing protocols to engineering a wide array of Hamiltonians for probes of many-body physics[74,75]. In this work, we characterize the breakdown of AHT for approximating features of time series data and the sensor coupling factor, illustrated by example pulse sequences used for sensing with solid-state spins. We also analyze sufficient convergence criteria for the Magnus expansion for a general two-level system.

Using exact methods to determine the sensor coupling factor, we show that rapid echoes allow the Magnus approximation to remain accurate well beyond its guaranteed radius of convergence. This beyond AHT result is merely one example of structures that validate the Magnus approximation, opening up opportunities for Hamiltonian engineering principles without the increasing computational cost of higher order Magnus expansion terms[68,76].

These exact methods are valid regardless of Magnus expansion convergence, and can be readily adapted to more complex systems and sources of noise. Example future areas of application may include finite pulse errors, rotation angle errors, and interacting systems, which can be major contributors to dephasing and decoherence. In addition, frequency domain analysis can complement existing filter function approaches, which are frequently employed to study dynamical decoupling protocols across quantum information science and related fields[77,78,79]. We foresee future use of beyond AHT techniques in designing optimized measurement protocols that adhere to leading order Magnus approximations in the presence of dephasing sources, for open quantum simulations, and as potential reward metrics for optimal control and machine learning approaches[1,64,80]. Extending this beyond AHT approach to higher-dimensional systems

may also be relevant to sensing protocols that leverage qudit structures[15,30,81]. However, in such cases we expect the contributions to the radius of convergence in Eq. (2.9) to remain pertinent.

2.8 Supplementary Notes

2.8.1 Magnus Expansion

The Magnus expansion is an approximate solution to a linear differential equation of the form

$$\psi'(t) = O(t)\psi(t)\,, \tag{2.15}$$

where the solution is given by the expression

$$\psi(t) = \exp\left(\Omega(t)\right). \tag{2.16}$$

For the pulse sequences considered in the main text, $O(t) = -iH_k$ during the k-th period of free evolution ($(k-1)\tau \leq t < k\tau$, with τ being the duration of each period). H_k denotes the toggling frame Hamiltonian during this interval. To evaluate the m-th order term in the Magnus series, Ω_m, we follow the steps outlined in Arnal et al. 2018[72],

$$\Omega_m = \frac{1}{m}\sum_{\sigma}(-1)^{d_b}\frac{d_a!d_b!}{m!}A(\sigma) = \sum_{\sigma} g(\sigma)A(\sigma)\,, \tag{2.17}$$

where

$$A(\sigma) = \int_0^t dt_1 \int_0^{t_1} dt_2 \cdots \int_0^{t_{m-1}} dt_m O(t_{\sigma(1)})O(t_{\sigma(2)})\cdots O(t_{\sigma(m)})\,. \tag{2.18}$$

Here, σ is a permutation of $\{1,\ldots,m\}$, while d_a and d_b are the number of ascents and descents in a permutation σ, respectively.

2.8.2 Discrete Time

We consider a series of toggling frame Hamiltonian operators that are constant during periods of duration τ. For a sequence of total duration $n\tau$, Eq. (2.18) can be expressed by discrete sums over these n sub-intervals,

$$A(\sigma) = \sum_{x_1=1}^{n} \sum_{x_2=1}^{x_1} \cdots \sum_{x_m=1}^{x_{m-1}} \int_{(x_1-1)\tau}^{x_1\tau} dt_1 \int_{(x_2-1)\tau}^{\min[t_2,x_2\tau]} dt_2 \cdots \int_{(x_m-1)\tau}^{\min[t_{m-1},x_m\tau]} dt_m O(t_{\sigma(1)}) \cdots O(t_{\sigma(m)}) \,. \tag{2.19}$$

Setting $A(\sigma) = \sum_{\vec{x}} a(\vec{x}, \sigma)$ with

$$a(\vec{x}, \sigma) = \int_{(x_1-1)\tau}^{x_1\tau} dt_1 \int_{(x_2-1)\tau}^{\min[t_2,x_2\tau]} dt_2 \cdots \int_{(x_m-1)\tau}^{\min[t_{m-1},x_m\tau]} dt_m O(t_{\sigma(1)}) \cdots O(t_{\sigma(m)}) \tag{2.20}$$

$$= O_{x_{\sigma(1)}} \cdots O_{x_{\sigma(m)}} \int_{(x_1-1)\tau}^{x_1\tau} dt_1 \int_{(x_2-1)\tau}^{\min[t_2,x_2\tau]} dt_2 \cdots \int_{(x_m-1)\tau}^{\min[t_{m-1},x_m\tau]} dt_m \tag{2.21}$$

$$= \mathcal{O}_{\vec{x},\sigma} f(\vec{x}) \,, \tag{2.22}$$

where $O_k = O((k-1)\tau \leq t < k\tau)$, the function $f(\vec{x})$ is the result of the nested integrals, and $O_{x_{\sigma(1)}} \cdots O_{x_{\sigma(m)}} = \mathcal{O}_{\vec{x},\sigma}$. We can evaluate the individual contributions for an arbitrary sub-interval denoted by $\vec{x} = (x_1, \ldots, x_m)$. During this sub-interval, $O(t_{\sigma(1)}) \cdots O(t_{\sigma(m)}) = \mathcal{O}_{\vec{x},\sigma}$ is constant and can be taken out of the integral. Since $x_i \geq x_{i+1}$, the sequence $\vec{x}$ is in descending order. Therefore, for a pair of neighboring indices x_i and x_{i+1}, the corresponding nested integrals

in $a(\vec{x},\sigma)$ fall under one of two cases:

$$\int_{(x_i-1)\tau}^{\min[t_{i-1},x_i\tau]} dt_i \int_{(x_{i+1}-1)\tau}^{\min[t_i,x_{i+1}\tau]} dt_{i+1} = \begin{cases} \displaystyle\int_{(x_i-1)\tau}^{\min[t_{i-1},x_i\tau]} dt_i \int_{(x_{i+1}-1)\tau}^{x_{i+1}\tau} dt_{i+1}, & \text{if } x_i > x_{i+1} \\ \displaystyle\int_{(x_i-1)\tau}^{\min[t_{i-1},x_i\tau]} dt_i \int_{(x_{i+1}-1)\tau}^{t_i} dt_{i+1}, & \text{if } x_i = x_{i+1} \end{cases} . \tag{2.23}$$

When $x_i > x_{i+1}$, the bounds of the integrals are independent, and can be evaluated separately. However, if $x_i = x_{i+1}$, we combine and evaluate the integrals that correspond to a subsequence containing only repeated indices. For k repeated indices, i.e., $x_{i+1} = x_{i+2} = \cdots = x_{i+k}$, we obtain the following result:

$$\int_{(x_{i+1}-1)\tau}^{x_{i+1}\tau} dt_{i+1} \int_{(x_{i+2}-1)\tau}^{t_{i+1}} dt_{i+2} \cdots \int_{(x_{i+k}-1)\tau}^{t_{i+k-1}} dt_{i+k} = \frac{\tau^k}{k!} . \tag{2.24}$$

When applied to the entire set of indices $\vec{x}$, each unique integer x_j with k_j repeated entries introduces a multiplicative factor $\tau^{k_j}/(k_j!)$ during evaluation of the corresponding time integrals. For example, for a particular set of $m = 11$ indices $\vec{x}' = (4,4,3,3,3,3,3,2,1,1,1)$, we obtain

$$a(\vec{x}',\sigma) = \underbrace{\frac{\tau^{11}}{2!\,5!\,1!\,3!}}_{f(\vec{x}')} \quad \underbrace{O_{x'_{\sigma(1)}} O_{x'_{\sigma(2)}} \cdots O_{x'_{\sigma(11)}}}_{\mathcal{O}_{\vec{x}',\sigma}} , \tag{2.25}$$

where τ_i describes any time between $i\tau \le t < (i+1)\tau$ when $O(t)$ is constant. As a result, the Magnus expansion can be expressed as

$$\begin{aligned} \Omega_m &= \sum_{\sigma,\vec{x}} g(\sigma) a(\vec{x},\sigma) \\ &= \sum_{\sigma,\vec{x}} g(\sigma) f(\vec{x}) \mathcal{O}_{\vec{x},\sigma} . \end{aligned} \tag{2.26}$$

2.8.3 Magnus expansion converges for rapid echoes

We have that

$$a(\vec{x},\sigma)=\sum_{\vec{x}} f(\vec{x})\mathcal{O}_{\vec{x},\sigma}\,, \tag{2.27}$$

where $\vec{x}=(x_1,x_2,\ldots,x_m)$ is a vector of indices with $x_j\geq x_{j+1}$.

We define rapid echos as having each odd pulse followed by an even pulse that generates a toggling frame Hamiltonian of opposite sign, such that $O_{2j-1}=-O_{2j}$. To show that $a(\vec{x},\sigma)$ is zero for an arbitrary sequence of rapid echoes, we first consider the case where x_1 is an even number, repeats k times ($x_1=x_2=\cdots=x_k$), and is then followed by the index x_{k+1} that does not belong to the same echo ($x_{k+1}\neq x_1$ and $x_{k+1}\neq x_1-1$), namely

$$\vec{x}_0=(\underbrace{x_1,\ldots,x_1}_{k\text{ repetitions}},x_{k+1},\ldots)\,. \tag{2.28}$$

We shall see that $a(\vec{x}_0,\sigma)$ is cancelled by sets of indices of the form

$$\vec{x}_l=(\underbrace{x_1,\ldots,x_1}_{k-l\text{ repetitions}},\underbrace{x_1-1,\ldots,x_1-1}_{l\text{ repetitions}},x_{k+1},\ldots)\,. \tag{2.29}$$

Specifically, $a(\vec{x}_0,\sigma)$ is cancelled by all the other sets of indices that start with k pulses belonging to the same echo. To show this, we notice that

$$f(\vec{x}_l)=\frac{1}{l!(k-l)!}f(x_{k+1},\ldots)=\frac{1}{k!}\binom{k}{l}f(x_{k+1},\ldots)\,, \tag{2.30}$$

which results in

$$\sum_{l=0}^{k} f(\vec{x}_l)\mathcal{O}_{\vec{x}_l,\sigma} = \frac{1}{k!} f(x_{k+1},\ldots) \sum_{l=0}^{k} (-1)^l \binom{k}{l} \mathcal{O}_{\vec{x}_0,\sigma}\,. \tag{2.31}$$

The factor $(-1)^l$ comes from exchanging each O_{x_1-1} for $-O_{x_1}$ as $\mathcal{O}_{\vec{x}_l,\sigma}$ is converted to $\mathcal{O}_{\vec{x}_0,\sigma}$. Applying the identity

$$\sum_{j=0}^{i} (-1)^j \binom{i}{j} = 0\,, \tag{2.32}$$

we obtain

$$\sum_{l=0}^{k} f(\vec{x}_l)\mathcal{O}_{\vec{x}_l,\sigma} = 0\,. \tag{2.33}$$

This implies that the total contribution from all set of indices $\vec{x}$ starting with k pulses in the same echo is equal to zero. Therefore, since the sum $\sum_{\vec{x}} f(\vec{x})\mathcal{O}_{\vec{x},\sigma}$ can be sub-divided according to the number of starting pulses belonging to the same echo, we immediately see that

$$\sum_{\vec{x}} f(\vec{x})\mathcal{O}_{\vec{x},\sigma} = \sum_{\substack{\vec{x} \\ k=1}} f(\vec{x})\mathcal{O}_{\vec{x},\sigma} + \sum_{\substack{\vec{x} \\ k=2}} f(\vec{x})\mathcal{O}_{\vec{x},\sigma} + \cdots = 0\,. \tag{2.34}$$

Under these conditions, we conclude that the Magnus series contribution at arbitrary order is zero, i.e.,

$$\Omega_m = \sum_{\sigma,\vec{x}} g(\sigma)a(\vec{x},\sigma) = \sum_{\sigma} g(\sigma) \sum_{\vec{x}} f(\vec{x})\mathcal{O}_{\vec{x},\sigma} = 0\,, \tag{2.35}$$

2.8.4 Numerical simulation with random echo sequences

As a sanity check, we generate 10^4 random sequences of $n = 20$ with a rapid echo structure, compute the Magnus expansion up to order $m = 5$, and calculate the spectral norm of each order. We find that all the spectral norms are around 10^{-16}, limited by machine precision (for double-

precision numbers) and therefore zero for practical considerations. We conclude that no rapid echo sequences have non-zero Magnus expansion for these parameters, consistent with the result in the previous section.

The random echo sequences are described by the most general Hamiltonian (up to a multiplicative constant) for a two-level system, $H(\theta, \phi) = \sin\theta\cos\phi\sigma_x + \sin\theta\sin\phi\sigma_y + \cos\theta\sigma_z$. To generate a random Hamiltonian, we select a random value of θ within $[0, \pi]$ and a random value of ϕ within $[0, 2\pi]$. A sequence of M rapid echoes ($n = 2M$) is generated using M random Hamiltonians, each repeated with the opposite sign. Consequently, the sequence is $H(\theta_1, \phi_1)$, $-H(\theta_1, \phi_1)$, ..., $H(\theta_M, \phi_M)$, and $-H(\theta_M, \phi_M)$.

Chapter 3: Ensemble Cluster Sampling for Heterogeneous Dynamical Decoupling

The work described in this chapter is currently being organized as a manuscript for journal submission.

My contributions include project conception, numerical simulations and development of ECS environments for algorithm training, genetic algorithm implementation, manuscript preparation, and supervision of all project components. I expect to share a statement of equal contribution with George Witt, who has spent years working alongside me to develop and supervise all project components, in addition to leading efforts on algorithmic approaches, implementation, and results. Connor A. Hart worked with me to develop the project idea, and provided guidance during early years of the project. The work was performed under the supervision of Ronald L. Walsworth.

3.1 Introduction

Ensemble quantum systems form the backbone of precision measurement, from atomic clocks, to magnetic resonance imaging to solid-state sensors based on nitrogen-vacancy (NV) centers in diamond[9,82]. Their sensitivity hinges on maintaining quantum coherence across many identical subsystems, where averaging across the ensemble enables improvements in signal-to-

noise[4,83]. However, preserving coherence across large ensembles is fundamentally limited by heterogeneities: variations in local electromagnetic fields, many-body interactions, and control parameters across the measurement region[6,84,85]. Addressing these variations is key to transitioning from proof-of-concept demonstrations to scalable quantum technologies[86].

Dynamical decoupling (DD) techniques tackle decoherence by applying timed control pulses to suppress unwanted interactions. While analytic approaches based on average Hamiltonian theory have produced workhorse sequences like CPMG and XY8[56,87,88,89], they rely on idealized assumptions that often break down in experimental conditions[46]. Recent algorithmic approaches, including those based on machine learning, show promise for system-tailored protocol engineering[1,90]. However, these methods face a core challenge: training on single-instance systems leads to overfitting that fails to capture the heterogeneities encountered in real experimental ensembles.

We introduce Ensemble Cluster Sampling (ECS), a scalable simulation framework that directly incorporates experimentally motivated parameter distributions into quantum control optimization. Rather than simulating full many-body dynamics—computationally intractable for realistic ensemble sizes—ECS represents the ensemble as a collection of smaller, finite-size clusters, each sampling from distributions of Hamiltonian parameters. This approach exploits the local nature of dominant interactions in systems like NV centers, where dipolar couplings limit the effective correlation length[91]. By evolving each cluster in parallel and evaluating DD protocols across the full parameter distribution, ECS enables the discovery of pulse sequences that are inherently robust to ensemble heterogeneities.

We apply ECS to two distinct algorithmic search methods: deep reinforcement learning using proximal policy optimization and a heuristic beam search technique[92,93]. Training these

algorithms on ECS environments rather than single-instance systems yields improvements in decoupling fidelity when pulse sequences are evaluated on ensemble distributions. Beyond DD sequence discovery, ECS enables direct simulation of ensemble dephasing dynamics: our simulations of NV-diamond systems reproduce the expected concentration-dependent scaling behavior, with our idealized treatment placing upper bounds on ensemble dephasing times for given defect concentrations. These results establish ECS as a broadly applicable tool for quantum control in ensemble systems[94], with implications extending beyond the DD applications shown here.

3.2 Ensemble Cluster Sampling (ECS)

Realistic quantum ensembles exhibit significant heterogeneity across the detection volume. In NV diamond magnetometry, for example, variations in local magnetic fields, strain environments, defect separations, and microwave control amplitudes can span orders of magnitude across a macroscopic sample[9,35,95]. Similarly, atomic vapor magnetometers face challenges from optical pumping inhomogeneities, magnetic field gradients, and spin-exchange broadening[96], while trapped ion sensors contend with variations in trap frequencies, laser intensities, and local electric fields across the ion crystal[97,98,99,100]. Traditional approaches to quantum control design either ignore this heterogeneity entirely—optimizing sequences for idealized, uniform systems—or use approximate methods such as mean-field theories that cannot capture the full distribution of local environments[31].

The key insight that enables Ensemble Cluster Sampling is that dominant interactions in many quantum platforms are inherently local. For dipolar-coupled spin systems like NV centers, interactions scale as $(\text{distance})^{-3}$, ensuring that each spin is primarily influenced by its

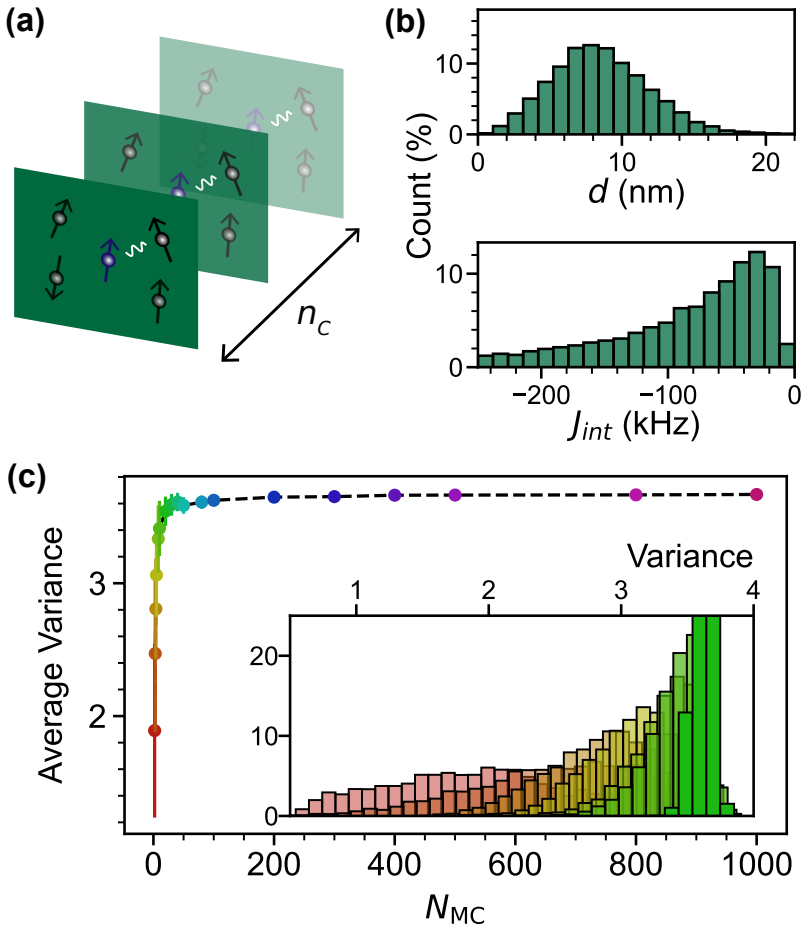


Figure 3.1: (a) Conceptual illustration of ensemble cluster sampling (ECS), where n_C clusters are drawn from distributions of Hamiltonian parameters.(b) Distributions of nearest-neighbor diamond lattice spacing d and corresponding coupling strengths J, for defect concentration $C = 2\,\text{ppm}$. (c) Average variance of the propagator under a Ramsey sequence as a function of the number of clusters N_{MC}, averaged over 5000 samples per point. The inset shows example variance distributions. Saturation approaching $N_{MC} = 100$ sets the minimum cluster count for simulation use.

nearest neighbors rather than the global ensemble state. Similarly, in atomic vapor systems, spin-exchange collisions are predominantly local processes, while trapped ion crystals exhibit nearest-neighbor Coulomb interactions that dominate dynamics[82]. As illustrated in Fig. 3.1(a), ECS assumes the ensemble as a collection of n_C low-dimensional clusters, each capturing its local environment while remaining computationally tractable for numerical simulations.

To implement ECS for NV diamond systems, we define a cluster Hamiltonian $H(\vec{x})$ that

captures the dominant dipolar interactions and local disorder. For a pair of electronic spins i and j displaced by $\vec{r}_{ij}$, we describe the dipolar coupling strength by

$$J_{ij} = \frac{\mu_0 \hbar}{4\pi} \frac{\gamma_e^2}{r_{ij}^3} \tag{3.1}$$

where γ_e is the electron gyromagnetic ratio, μ_0 is the vacuum permeability, and $\hbar$ is the reduced Planck constant. Under the secular approximation[101], the cluster Hamiltonian takes the form:

$$\begin{aligned} H(\vec{x}) = &\sum_{i<j} J_{ij} \left[\alpha S_z^{(i)} S_z^{(j)} - \beta \left(S_x^{(i)} S_x^{(j)} + S_y^{(i)} S_y^{(j)} \right) \right] \\ &+ \sum_i \Delta_i S_z^{(i)}, \end{aligned} \tag{3.2}$$

where $S_{x,y,z}^{(i)}$ are spin-1/2 operators, Δ_i represents local magnetic disorder (or any static shift), and scalar modifiers α and β account for specific interaction forms[91].

Each cluster is initialized by randomly populating a diamond lattice region with defects at concentration C, then selecting a representative spin pair for evolution. This process generates realistic distributions of coupling strengths as shown in Fig. 3.1(b), while maintaining computational tractability via small cluster sizes. Additional heterogeneity is incorporated by sampling local magnetic disorder and control field variations from Gaussian distributions, creating a statistical ensemble that captures the essential physics of local environments across the full ensemble parameter space.

A key advantage of ECS is that only a single parameter, n_C, requires tuning to guarantee ensemble convergence. We choose n_C by studying the convergence of ensemble statistics under representative control sequences. We evaluate the variance of the time evolution propagator under

a Ramsey sequence consisting of two $\pi/2$-rotation pulses separated by a free evolution interval. For each value of n_C, we compute the propagator U across all clusters and calculate the sum of matrix element variances. As shown in Fig. 3.1(c), this variance saturates as n_C increases beyond approximately 100 clusters, indicating convergence of the ensemble representation[1]. The accompanying inset shows the variance distributions with increasing n_C, transitioning from broad distributions at small cluster numbers to narrow Gaussians at larger values, in line with the central limit theorem[103,104].

This saturation behavior provides a simple yet practical convergence check: once the variance plateaus, additional clusters yield diminishing statistical improvements while increasing computational cost.

3.3 Heterogeneous Dynamical Decoupling

Next we apply ECS[2] Our first approach employs deep reinforcement learning (RL)[105], illustrated in Fig. 3.2(a). We formulate sequence discovery as a sequential decision problem. At each step t, the algorithm selects from a library of $p = 9$ possible pulses (actions), which consists of $\pi/2$- and π-rotation pulses about $\pm x$, $\pm y$ axes in addition to a null action choice. After the chosen pulse action a_t is applied to the ensemble, all clusters evolve freely under their respective Hamiltonians $H(\vec{x})$ for time τ before the next pulse. The system state s_t at each step contains the expectation values and covariance matrix for a complete basis set of operator observables. After the pulse a_t and free evolution, the updated state s_{t+1} is input to the algorithm to decide

[1] Alternative metrics such as the Kullback-Leibler divergence between cluster distributions can similarly quantify convergence[102].

[2] Note minor modifications were made to how interactions are generated. We sample from Gaussian distributions of defect separations before converting to interaction strengths to eliminate time spent on lattice generation.

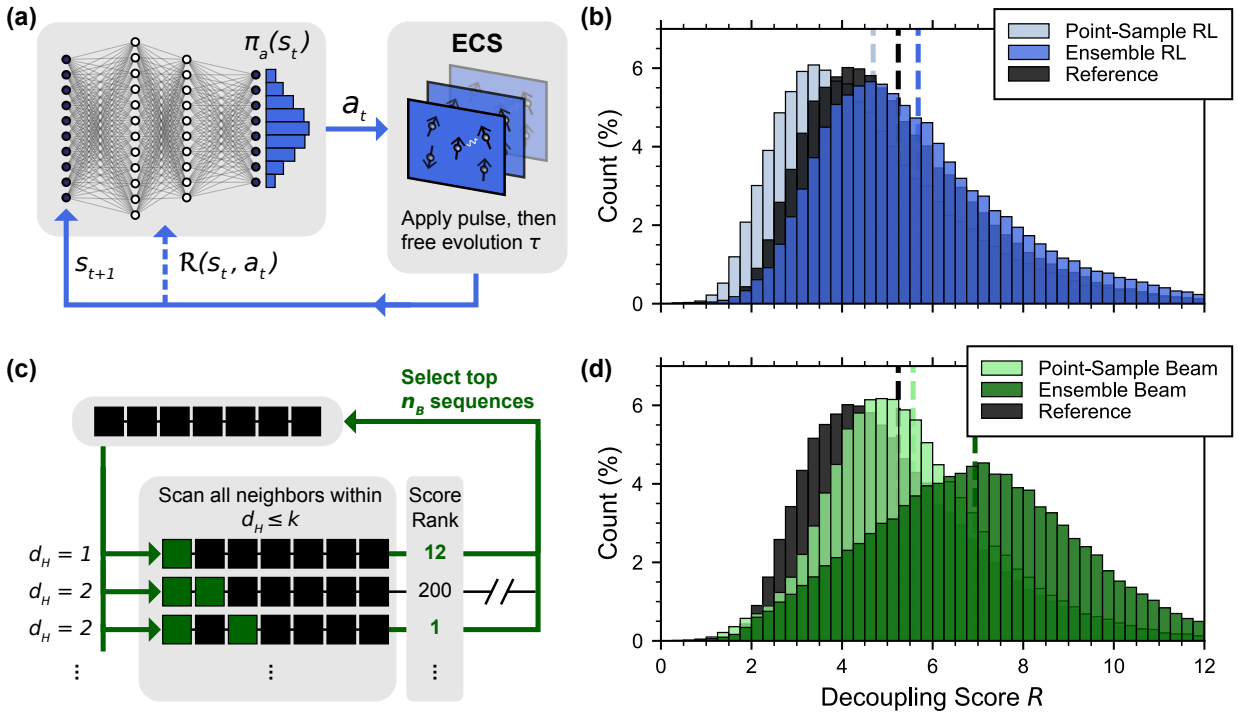


Figure 3.2: (a) Schematic of deep reinforcement learning approach using proximal policy optimization (PPO). The policy network $\pi_a(s_t)$ decides the pulse action a_t at integer step t, which is applied to the ECS environment followed by a free evolution period τ. The resulting state s_{t+1} is fed back to the network to decide the next action. Network weights are updated periodically using the ensemble reward $\mathcal{R}$. (b) Decoupling score histograms for best-performing RL PPO sequences from point-sample training (light blue) and ECS training (dark blue). (c) The beam search algorithm starts with a random seed sequence—a beam—and generates all neighbor sequences with up to k pulse changes. The top $n_B = 100$ pulse sequences are selected as new beams to repeat the search. (d) Histograms for best-performing beam search sequences for point-sample (green) and ECS training (dark green). Reference sequence Angle-12 from Peng et al.[1] is plotted in both (b) and (d) as black histograms.

the next pulse action a_{t+1}. This iteration continues until a sequence of length L is constructed: $(a_1, \tau, a_2, \tau, \ldots a_L, \tau)$.

The challenge in dynamical decoupling sequence discovery lies not just in the exponentially large search space (p^L possible sequences) but in finding sequences that perform robustly across the heterogeneous parameter distributions. Traditional approaches either optimize for non-generalizable dynamics[1] or rely on restrictive assumptions about noise spectra and bath properties. We define an evaluation (reward) metric $\mathcal{R}$ that quantifies pulse sequence performance

across the ensemble. For the i-th cluster evolved under U_i, we first calculate $F_i = \mathrm{Tr}\,|U_i|\,/2^N$, representing the decoupling fidelity with respect to ideal (trivial) unitary evolution. Next, a log-transform is used to amplify differences in fidelity, which is crucial for RL given the sparseness of high-performing sequences in the search space. This decoupling score R_i is then averaged across all clusters to give the ensemble reward:

$$\mathcal{R} = \big\langle R_i \big\rangle_i = \big\langle -\ln\left(1 - F_i\right) \big\rangle_i \tag{3.3}$$

Note that even high-performing DD sequences are not guaranteed to show steady improvement in decoupling fidelity during construction, so we modify our $\mathcal{R}$ to be sparse—rewards are only provided at the last pulse. As an example of a modern RL framework with proven success in discrete sequence optimization, we employ Proximal Policy Optimization (PPO), a policy-gradient method that maintains stable learning in sparse reward environments by limiting policy updates through gradient clipping[92].

We benchmark our approach by comparing to the best-performing 12-pulse sequence, Angle-12, from a reinforcement learning result by Peng et al.[1]. Shown as the black reference histograms in Fig. 3.2(b) and 3.2(d), Angle-12 was discovered using an alternative RL implementation with a smaller action library of $p = 5$ pulse choices (totaling a search space of $5^{12} \approx 10^8$ sequences, well within brute-force search capabilities), to decouple dipolar interactions within a spin-$1/2$ chain. For fair comparison, we modify our cluster Hamiltonian to use $\alpha = \beta = 1$ in Eq. (3.2), representing purely dipolar-coupled spin pairs, and expand our action library to 9 choices to provide greater sequence design flexibility. In contrast to our ECS approach, the reference sequence was trained on "point-sample" environments—a single cluster of spins with

fixed Hamiltonian parameters. Despite being the optimal sequence within its original training environment, Angle-12 exhibits a broad performance distribution with significant variance when evaluated across realistic parameter heterogeneity—precisely the generalization challenge that motivates ensemble training methods.

Initial PPO training attempts on point-sample systems reproduced this generalization limitation. Figure 3.2(b) illustrates the overfitting problem through reward distributions achieved by sequences discovered using point-sample training (light blue histogram). While these sequences achieved high performance under their specific training conditions, evaluation across the full ECS parameter distribution reveals performance comparable to the literature reference - shown by a low-performance tail and lower average fidelity.

The solution lies in training directly on ECS environments that capture the full parameter heterogeneity from the outset. Rather than optimizing sequences for idealized conditions, we train PPO algorithms on the ensemble clusters sampled from experimentally motivated parameter distributions. This ensemble training approach forces the algorithm to discover sequences with high average fidelity across the parameter space, naturally selecting for robustness. The effectiveness of this approach is demonstrated in Fig. 3.2(b), where the ”Ensemble RL” histogram (dark blue) shows improved performance compared to point-sample training. This improvement reflects the algorithm's discovery of sequence patterns that are inherently robust to parameter variations. Instead of exploiting narrow resonance conditions that fail under parameter changes, ensemble training identifies dynamical decoupling protocols that maintain effectiveness across the sampled range of dipolar coupling strengths, magnetic disorder, and control field imperfections present in realistic experimental conditions. The resulting sequences represent a fundamentally different class of control protocols—optimized not for ideal performance under specific

conditions, but for high average fidelity in the presence of heterogeneity.

The generalizability of ensemble training extends beyond reinforcement learning to other algorithmic search methods. To demonstrate its broad applicability, we also implemented a modified beam search algorithm—a heuristic exploration technique that expands from high-performing sequences by exploring local neighborhoods in sequence space[106]. As shown in Fig 3.2(c), a randomized seed sequence—referred to as a beam—is mutated by generating all possible neighbor sequences with up to k pulses changed (i.e., Hamming distance $d_H \leq k$). The top $n_B = 100$ candidates are selected as new beams and the mutation process repeats on all of these individually, while maintaining a taboo list to prevent redundancy.

Figure 3.2(b) shows that beam search exhibits the same dramatic improvement when trained on ECS environments versus point-sample conditions. The "Point-Sample Beam" distribution (light green) displays the characteristic broad, low-performance tail seen with RL point-sample training, while "Ensemble Beam" (dark green) achieves a narrow, high-performance distribution. Remarkably, the ensemble-trained beam search not only overcomes the generalization challenge but also achieves higher overall rewards than the ensemble-trained RL approach, with peak performance reaching decoupling scores above 10. This method-agnostic improvement reinforces the benefits of ensemble training rather than algorithmic specifics. Both RL and beam search, when exposed to the full parameter heterogeneity during optimization, find robust sequence designs that maintain high decoupling scores across diverse experimental conditions. The superior performance of beam search reflects its systematic exploration of sequence neighborhoods - manual inspection of sequences showed smaller Hamming distances between high-performing pulse sequences, which motivated our use of beam search.

These results establish ensemble training as a general principle for robust quantum con-

trol design that extends beyond the specific DD applications demonstrated here. The core insight—that optimization algorithms must be exposed to realistic parameter heterogeneity during training to achieve practical robustness—applies broadly to quantum control problems where system parameters vary across experimental runs or detection volumes. The ECS framework itself offers immediate applications to other quantum sensing modalities. Atomic vapor magnetometers face similar ensemble heterogeneity from optical pumping variations, magnetic field gradients, and thermal density fluctuations, while trapped ion sensors must contend with variations in trap frequencies, laser intensities, and local electric fields. In each case, ECS offers a computationally tractable representation of the ensemble while capturing the dominant local interactions that govern decoherence. The linear scaling with cluster number makes ECS particularly attractive for exploring parameter spaces that would be unfeasible with full many-body simulations. Beyond sensing applications, ensemble training addresses a key challenge in quantum control: the gap between theoretical optimization under idealized conditions and practical performance in realistic experimental environments. This shift from pursuing ideal performance to ensuring robustness is essential as quantum technologies transition from proof-of-principle demonstrations to practical applications requiring consistent operation across diverse conditions.

3.4 Ensemble Dephasing Timescale

Beyond enabling robust control sequence discovery, ECS provides a framework for direct simulation of ensemble dephasing dynamics that offers unique theoretical insights. In NV diamond systems, ensemble coherence is limited by multiple decoherence mechanisms including dipolar interactions between NV centers, coupling to nuclear spin baths, magnetic field gradients,

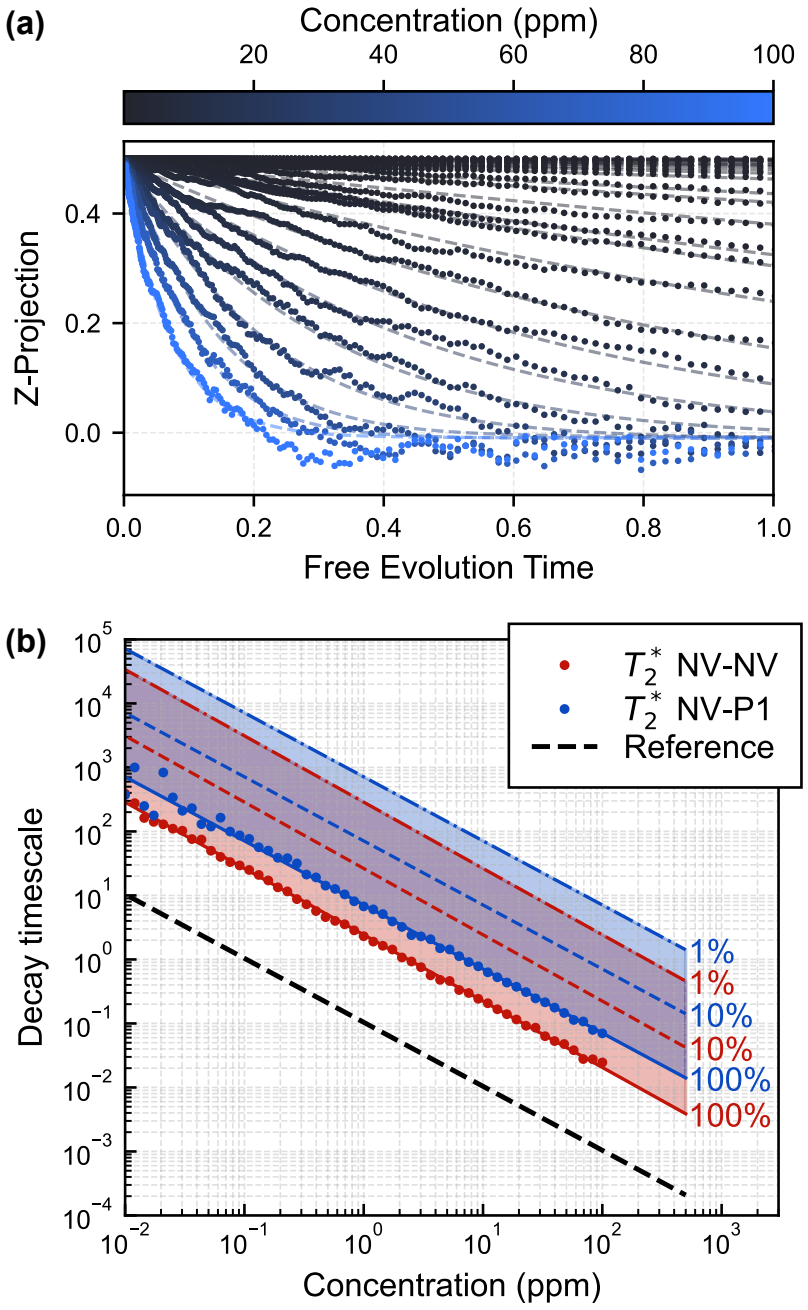


Figure 3.3: (a) Ramsey decay curves show ensemble sz projection decreasing with free evolution time τ, for varying defect concentration C with central spin simulations. (b) Fitted decay timescales are plotted as a function of C, on a log-log plot. A reference line from Bauch et al. is shown in black.

strain inhomogeneities, and control field imperfections[9,31,64]. While experimental studies characterize the combined effect through phenomenological decay times, ECS enables systematic investigation of individual contributions by simulating idealized systems with controlled noise

sources.

Here we examine concentration-dependent dephasing for clusters of spin pairs, where each pair samples realistic coupling strengths from the ensemble distribution. We implement two distinct interaction dynamics prevalent in NV diamond samples[64]: a central spin model, where an NV center couples to a nearby P1 defect (set $\alpha = 1$ in $H(\vec{x})$), and a like-spin pair model, representing NV-NV dipolar interactions ($\alpha = 2$)[91]. The central spin model reflects the experimental scenario where P1 electronic spins dominate magnetic noise while control fields interrogate only the central NV spin. The like-spin model captures direct NV-NV interactions within the ensemble. This systematic isolation of individual decoherence mechanisms demonstrates ECS's unique capability for controlled parameter studies. By aggregating these dynamics across the full parameter space, we establish fundamental upper bounds on ensemble dephasing times that reveal the intrinsic limits imposed by dipolar coupling alone.

We probe these limits using Ramsey interferometry sequences, consisting of two ideal $\pi/2$-pulses separated by a period of free evolution τ. Each cluster evolves a spin pair initially polarized along the $+z$ direction, with interaction strength determined by the distance distribution shown in Fig. 3.1(b) and Eq. (3.1). The central spin model results are shown in Fig. 3.3(a), which contains time series decay curves of ensemble-averaged S_z projection under pure dipolar coupling. Both interaction models yield qualitatively similar decay behavior, differing primarily in absolute timescales, so we present only the central spin curves for visual clarity. Exponential decay fits are used to extract ensemble dephasing times T_2^* as a function of defect concentration C, demonstrating faster decay with increasing defect concentration.

To isolate the effect of dephasing due to dipolar interactions, we implement a near-idealized ECS simulation with perfect experimental conditions, with ideal control pulses and no magnetic

disorder. As shown in Fig. 3.3(b), fitted T_2^* values for both central spin (blue) and like-spin pair (red) models reproduce the expected power-law scaling with defect concentration, demonstrating excellent agreement with experimental trends[31,95]. Besides agreement in scaling behavior, we observe a gap in absolute timescales one to two orders of magnitude longer than experimental measurements. This difference is to be expected - with solely interaction-induced dephasing, these timescales likely suggest upper bounds to T_2^* for a given defect concentration, due to our use of near-ideal unitary evolution with no dissipative dynamics. No realistic quantum control protocol can exceed these limits, making them fundamental benchmarks for experimental performance.

Our analysis of upper bound dephasing times can be extended to scenarios where bulk nitrogen concentration [N] is more readily measured than specific defect populations ([NV] or [P1]). The additional scaling lines in Fig. 3.3(b) show how defect conversion efficiency (fraction of defect population with respect to total nitrogen) affects achievable dephasing times. For example, at 1 ppm total nitrogen with 100% conversion to NV centers, the maximum T_2^* reaches $\sim 3\,\mu$s, but smaller conversion efficiencies of 10% or 1% would yield proportionally longer dephasing times of approximately 30 μs or 300 μs, respectively.

3.5 Supplementary: Genetic Algorithm for Dynamical Decoupling Sequence Search

Nitrogen-vacancy (NV) spin ensembles in diamond provide an advanced magnetic sensing platform at ambient conditions. Improvements in fabrication techniques provide access to diamond material with moderate ($\sim$1 ppm) to high (> 10 ppm) NV densities while maintaining isotopic purity and low ($\gtrsim 10$ kHz) crystal strain gradients. Combining dynamical decoupling techniques and suppression of spin bath induced dephasing, NV coherence magnetometry can operate within the interaction-limited regime, where spin coherences are restricted by interactions between NV sensor spins. By considering the effective average Hamiltonian during a quantum control sequence, previous influential work in Hamiltonian engineering successfully extends NV coherence by suppressing the effects of such dipolar interactions, resulting in sensitivity improvements to AC magnetic fields.

In this section, we search for pulse sequences with the goal of dyanamically decoupling an interacting spin system, by implementing a genetic algorithm (GA) for a spin-1/2 chain that includes the effects of magnetic disorder and pulse errors. We compare our results to a select pulse sequence, namely DROID-60, which was developed by Choi et al.[63] to decouple magnetic dipolar interactions in the presence of magnetic disorder and pulse errors. To optimize the GA, we vary its parameters and repeat the algorithm 50 times. We inspect a pulse sequence that achieves higher fidelity than DROID-60 in the presence of pulse rotation errors, but conclude that this improvement is not present when there are no pulse errors.

For the purpose of magnetic sensing with dense NV ensembles, the rotating frame Hamil-

tonian includes the Zeeman contribution $H_{dis} = \alpha S_z$ for magnetic-field-dependent energy α and dipolar coupling strength H_{dd}. The collective spin operators in x, y, z are denoted as $S_{x,y,z}$ and $\vec{S}^{(i)} = (S_x, S_y, S_z)$ represents a vector containing these 3 spin operators. So far, we only consider nearest-neighbor interactions, and the unit displacement vector between nearest neighbor spins i and j is represented by $\hat{r}_{ij}$ below.

$$H_s = \alpha S_z + \sum_{i \neq j} H_{ij} \tag{3.4}$$

$$= \underbrace{\alpha S_z}_{H_{dis}} + \sum_{i \neq j} \underbrace{-\frac{J_0}{r^3}\left[3\left(\vec{S}^{(i)} \cdot \hat{r}_{ij}\right)\left(\vec{S}^{(j)} \cdot \hat{r}_{ij}\right) - \vec{S}^{(i)} \cdot \vec{S}^{(j)}\right]}_{H_{ij}} \tag{3.5}$$

The general form of the interaction is indicated by H_{ij}. Within the NV spin-1 ground state sublevels, certain transitions are suppressed due to the anharmonicity introduced by the zero-field splitting[91]. Additionally double flip-up or flip-down transitions are suppressed as these terms in the Hamiltonian are rapidly rotating. After applying these approximations, the effective dipole-dipole Hamiltonian in the rotating frame is shown below for spins i and j for aligned vs misaligned quantization axes.

$$H_{ij} \approx \begin{cases} \dfrac{J_0}{r^3}\left(1 - 3\cos^2\theta\right)\left[-\dfrac{1}{2}\begin{pmatrix} |+,0\rangle\langle 0,+| + |0,+\rangle\langle +,0| \\ \\ +|-,0\rangle\langle 0,-| + |0,-\rangle\langle -,0| \end{pmatrix} + S_z^{(i)} S_z^{(j)}\right] & \text{aligned} \\ \underbrace{\dfrac{J_0}{r^3}\left(1 - 3\cos^2\theta\right)}_{J_{int}} S_z^{(i)} S_z^{(j)} & \text{misaligned} \end{cases} \tag{3.6}$$

Kets (and bras) $|m_i \in \{0, \pm\}, m_j \in \{0, \pm\}\rangle$ denote the coupled spin states for the i-th and

j-th spins, with each being spin-1. θ refers to the angle between the spin quantization axis $\hat{z}$ and the displacement vector $\hat{r}$, related by $\cos\theta = \hat{z} \cdot \hat{r}$. We start by studying only the aligned case.

In this section we only study spin-1/2 ensembles. This is because there are established pulse sequences that effectively decouple spin-1/2 dipolar interactions, which we will use as a benchmark while we optimize parameters of the genetic algorithm. To be specific, we study the Lukin sequence B (DROID-60) as our reference[63,65], as it is designed to decouple both dipolar interactions and disorder effectively, while being robust to pulse errors.

In our pulse action library, we allow $\pi/2$ and π pulses along $\pm x$ and $\pm y$. We also include composite pulses with 2 consecutive pulses as individual actions. Including the identity operator (no pulse), the number of unique pulse actions is 22. Thus for a pulse sequence of length $m = 48$, the number of possible pulse sequences total $22^{48} \sim 10^{64}$, which is not tractable using a brute force search.

Given a native static Hamiltonian $H_s = H_{dis} + H_{dd}$, the system undergoes unitary evolution via the operator $U_\tau = \exp\left(-\frac{i}{\hbar} H_s \tau\right)$ during each free evolution interval τ. The overall unitary after applying the pulse sequence is therefore

$$U\left(\{A_m\}\right) = U_\tau A_m U_\tau A_{m-1} \ldots U_\tau A_1 \tag{3.7}$$

We use the fitness function given by $\mathcal{F}_{\text{UNI}}$ shown in Eq. 3.8, for a sequence of m pulse action unitaries $\{A_m\}$. For the sole goal of mitigating preserving a quantum state, the target unitary is the identity operator $U_{tgt} = \mathbb{1}$. The fitness function $\mathcal{F}_{\text{UNI}}(\{A_m\})$ is thus defined as a

comparison between the overall unitary and the target unitary (via inner product), shown below.

$$\text{Fitness: } \mathcal{F}_{\text{UNI}}(\{A_m\}) = -\ln\left(1 - \underbrace{\frac{1}{\mathcal{N}}\text{Tr}\left\{\left|U\left(\{A_m\}\right)U_{tgt}^{\dagger}\right|\right\}}_{\text{Fidelity}}\right) \tag{3.8}$$

Here, $\mathcal{N}$ denotes a normalization factor which depends on the dimensionality of the system. For a spin-1/2 ensemble of n spins, $\mathcal{N} = 2^n$. The log function is used in anticipation of fidelity values near unity. For reference, fidelity values of 3 nines (0.999), 5 nines and 7 nines result in fitness values of $\mathcal{F}_{\text{UNI}} = 6.9, 11.5$ and 16 respectively.

A genetic algorithm is an optimization procedure used to solve problems by maximizing or minimizing a particular fitness function[107], and is inspired by biological processes such as natural selection & mutation. We express each pulse sequence of length m as a genetic individual - an array with m elements denoting the pulse actions.

We outline one iteration of the learning process used for the genetic algorithm approach, to be repeated up to N_G times.

1. Define a population of N_{pop} pulse sequences. For the first generation, initialize the pulse sequences randomly.

2. Evaluate the fitness $\mathcal{F}\left(\{A_m\}\right)$ of all pulse sequences, rank agents by their fitnesses.

3. The pulse sequence with the best fitness value is labeled the "elite" of the current generation. This pulse sequence is included in the next generation of pulse sequences without any changes made.

4. Selection: a process is carried out to replace the remaining population. Currently we

are employing a selection "tournament" which randomly selects N_T sequences from the current generation and keeps the one with the highest fitness. This is repeated $N_{\text{pop}} - 1$ times until the number of pulse sequences total N_{pop}. Typically we use tournament sizes $N_T \sim 10\%$ of N_{pop}.

5. Crossover: The population is split into 2 halves and pairs of pulse sequences are matched up. With a user-defined probability $\mathbf{P}_C$, sequence pairs undergo "crossover", which involves swapping segments within their pulse sequence with each other (helpful pictures). For now, we employ a uniform crossover procedure, which utilizes a flat probability p_C of each pulse participating in the crossover.

6. Mutation: After crossover, each pulse sequence has a chance $\mathbf{P}_M$ of being mutated. If a sequence is targeted for mutation, each pulse within the sequence is resampled from the pulse action library with a uniform probability of p_M.

Note some similarities in the execution between DRL in Peng et al[1] and GA approaches - both handle a population of agents or pulse sequences and rank them based on fitness evaluations. The main functional differences include 1. direct manipulations of the pulse sequences rather than neural network weights, and 2. crossovers between pulse sequences within a generation.

Below I list the parameters to be optimized for the GA approach.

- Maximum number of generations N_G.
- Population N_{pop}.
- Tournament size N_T.
- Sequence crossover probability $\mathbf{P}_C$.

- Pulse crossover probability p_C.
- Sequence mutation probability $\mathbf{P}_M$.
- Pulse mutation probability p_M.

We use the DEAP package in Python to implement the genetic algorithm. This module contains several ready-to-go default scripts for the steps outlined in the GA approach, including selection, crossover and mutation functions.

3.5.1 GA Results

For all the genetic algorithm executions shown here, our spin chain contains $N = 3$ spins. The length of DROID-60 is 48, so the length of our pulse sequence m is fixed to match that. The interaction and disorder strengths are $J_{int} = 1$ kHz and $\alpha = 1$ MHz respectively.

We start by fixing both the Hamiltonian and GA parameters and running the pulse sequence search, repeated 50 times. The results are shown in Fig. 3.4, and the GA parameters listed in the caption. The purpose of the 50 repetitions is to mitigate the dependence of the search results on the initial seed. To clarify, the randomness inherent to this approach can lead to guesses that result in high fitness pulse sequences, but does not necessarily mean that the GA parameters used are better in general. By choosing a set of GA parameters and averaging the performance over many initial seeds, this should better describe the overall performance of the GA using these parameters. This average fitness $\bar{\mathcal{F}}_{best}(n_G)$ is shown on the left of Fig. 3.4, for each generation number n_G. The right figure only shows the best fitness across all 50 repetitions, for each generation n_G. Note: although we look to these average trends while we optimize the GA parameters, we are ultimately concerned with the best pulse sequence that was discovered for our problem. Thus

pulse sequence candidates are chosen from those with the **absolute** best fitness values, as shown on the right plot.

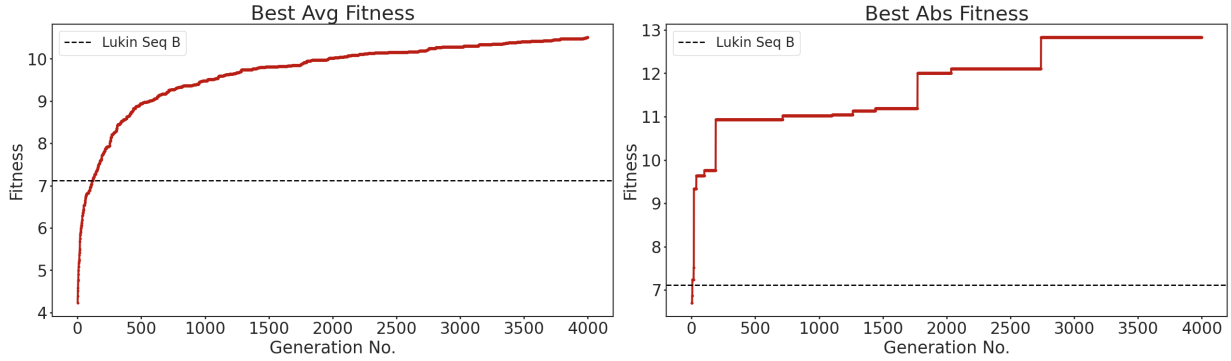


Figure 3.4: Statistics for genetic algorithm repeated with 50 seeds. **Left:** Best fitness values $\bar{\mathcal{F}}_{best}(n_G)$ averaged over 50 seeds, as a function of generation number n_G. **Right:** Best fitness value across all 50 seeds, plotted at each generation. Black dashed line shows fitness of sequence B under these conditions. Hamiltonian parameters: $N = 3,\ J_{int} = 1\,\text{kHz},\ \tau = 2\mu s,\ \alpha = 10\,\text{kHz}$, **+5% pulse rotation error.** GA parameters: $N_{\text{pop}} = 501,\ N_G = 4000,\ N_T = 3, \mathbf{P}_C = 0.5,\ p_C = 0.1,\ \mathbf{P}_M = 0.5,\ p_M = 0.06$

In Fig. 3.4, it appears that the GA is able to find pulse sequences with higher fitness values than Lukin's sequence B (black dashed line). However, this is specific to the Hamiltonian parameters used. In particular, this statement is true in the presence of the 5% rotation error applied to all pulses. Below in Fig. 3.5 I show the results of a comparable execution of the GA but without pulse errors.

We see that, despite the GA's ability to achieve respectable fitness values $\mathcal{F}_{\text{UNI}} \sim 23$ (10 nines), the fitness of sequence B is superior here. The plotted fitness value 30 for sequence B is artificially imposed - I capped the fitness at 30 to avoid infinities during the log evaluations. This fitness $\mathcal{F}_{\text{UNI}} = 30$ corresponds to a fidelity of ~ 13 nines.

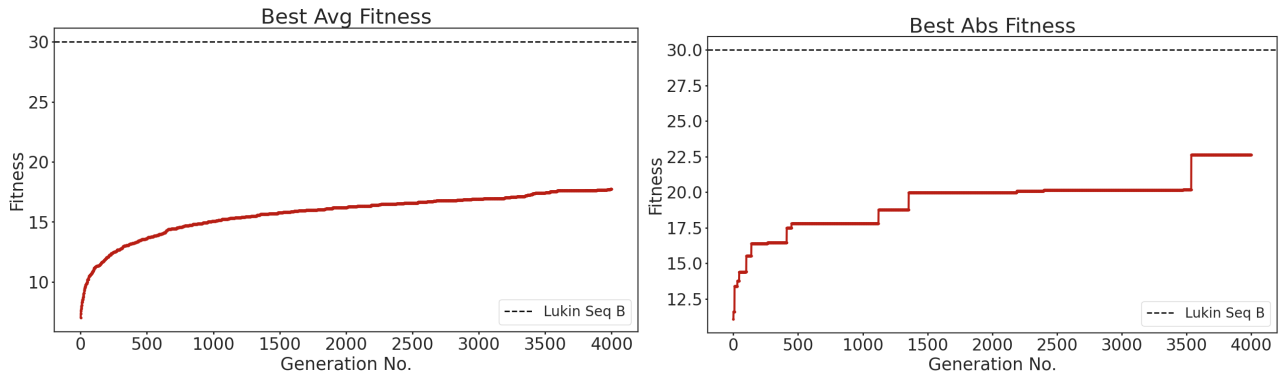


Figure 3.5: Statistics for genetic algorithm repeated with 50 seeds. **Left:** Best fitness values $\bar{\mathcal{F}}_{best}(n_G)$ averaged over 50 seeds, as a function of generation number n_G. **Right:** Best fitness value across all 50 seeds, plotted at each generation. Black dashed line shows fitness of sequence B under these conditions. Hamiltonian parameters: $N = 3$, $J_{int} = 1$ kHz, $\tau = 2\mu s$, $\alpha = 10$ kHz, **no pulse angle error**. GA parameters: $N_{\text{pop}} = 101$, $N_G = 4000$, $N_T = 8, \mathbf{P}_C = 0.5$, $p_C = 0.1$, $\mathbf{P}_M = 0.5$, $p_M = 0.06$

3.5.2 Optimizing GA parameters

We use the figure of merit established earlier to evaluate the performance of our genetic algorithm as we vary its parameters - the best fitness averaged over multiple seeds $\bar{\mathcal{F}}_{best}(n_G)$. An example is shown in Fig. 3.6, where the individual pulse mutation probability p_M is swept. F value of p_M, we obtain one trace of the best $\bar{\mathcal{F}}_{best}(n_G)$ with colors indicating the mutation probability p_M used for the 50 repetitions.

Best performance is observed for $p_M \sim 0.06$. The values for $\bar{\mathcal{F}}_{best}(N_G)$ at the last generation N_G is shown on the right plot with respect to p_M. The values of $\bar{\mathcal{F}}_{best}(N_G)$ at the last generation looks to be largest between $0.04 \leq p_M \leq 0.08$. For a 48-pulse sequence this corresponds to around 2 to 4 individual pulse mutations on average. Note that error bars indicate standard deviations across the 50 averages.

We repeat this analysis while varying GA parameters. The parameters that have shown to be most impactful towards the performance of the GA are p_M and p_C, the probabilities for

individual mutation and individual crossover respectively.

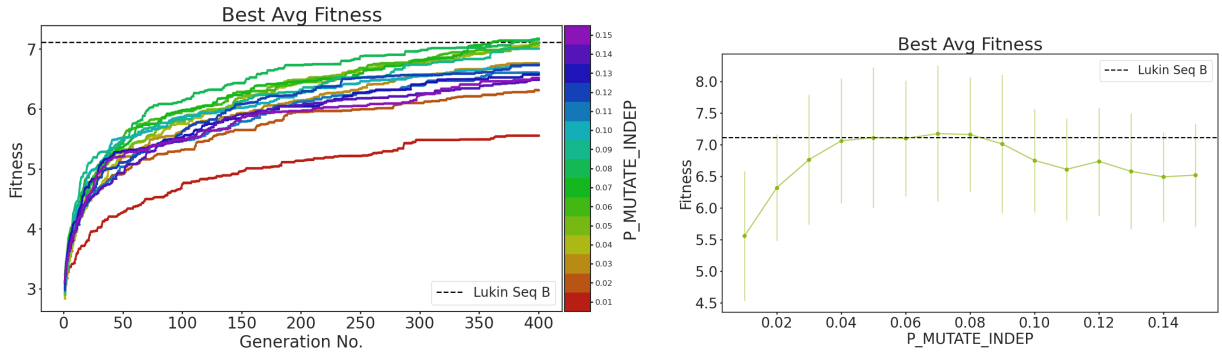


Figure 3.6: Fitness graphs for genetic algorithm executions with individual pulse mutation probability p_M (denoted as P_MUTATE_INDEP) varied between 0.01 and 0.15, each repeated over 50 seeds. **Left:** Each trace shows the best fitness values $\bar{\mathcal{F}}_{best}(n_G)$ averaged over 50 seeds, as a function of generation number n_G. **Right:** Best average fitness values at the last generation (no. 400). Standard deviation across all 50 repetitions are shown as error bars. Black dashed line shows fitness of sequence B under these conditions. Hamiltonian parameters: $N = 3,\ J_{int} = 1\,\text{kHz},\ \tau = 2\mu s,\ \alpha = 10\,\text{kHz}$, +5% pulse rotation error. GA parameters: $N_{\text{pop}} = 101,\ N_G = 400,\ N_T = 3, \mathbf{P}_C = 0.5,\ p_C = 0.1,\ \mathbf{P}_M = 0.5,\ p_M = 0.06$

3.5.3 Deviations from Hamiltonian parameters

As we optimize genetic algorithm parameters, we study the pulse sequences discovered by our searches. For any given pulse sequence, we wish to evaluate its performance (fitness) when properties of the Hamiltonian is changed. Specifically, we study the best pulse sequence from Fig. 3.4 and compare it to DROID-60, where there is a 5% pulse angle error. The colorplots in Fig. 3.7 show the fitness of this pulse sequence as α, J_{int}, τ and pulse angle errors are varied, two at a time. The white dotted line indicates the parameters used for the GA executions.

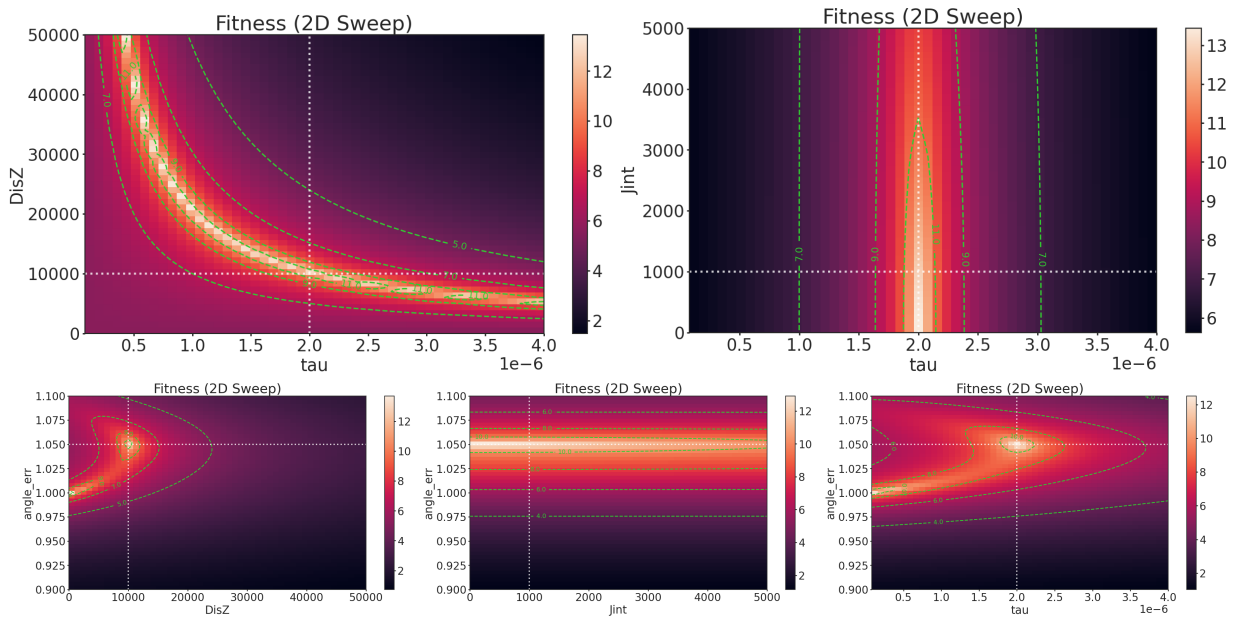


Figure 3.7: For the pulse sequence with the highest fitness from Fig. 3.4 (right), plots show fitness when Hamiltonian parameters are changed in the presence of +5% pulse error. Top left: α and τ are varied. Top right: J_{int} and τ are varied. Bottom left: Rotation angle error and α are varied. Bottom middle: Rotation angle error and J_{int} are varied. Bottom right: Rotation angle error and τ are varied. White dotted line indicates parameters used for GA execution and pulse sequence search. Green dashed lines mark fitness contours for this pulse sequence.

Next we show the fitness of DROID-60 with the Hamiltonian parameters varied around the same range, shown in Fig. 3.8. Performance is relatively homogeneous at $\mathcal{F} \sim 7$ with changes of a few percent. In contrast, the fitness of the GA-derived pulse sequence varies much more in Fig. 3.7. Also, we typically expect the performance of decoupling sequences to improve at smaller values τ, α and J_{int}. This is reflected in DROID-60, but not for the sequence found by the GA. Instead, we see local regions of higher fitness at specific values of α and τ (top left plot in Fig. 3.7). However, we do note improved fitness values as we decrease J_{int} (top right plot in Fig. 3.7), but only if τ is fixed.

To show relative fitness improvement over DROID-60, we subtract the reference fitness values in Fig. 3.8 from those in Fig. 3.7, and plot them below in Fig. 3.9. The GA pulse

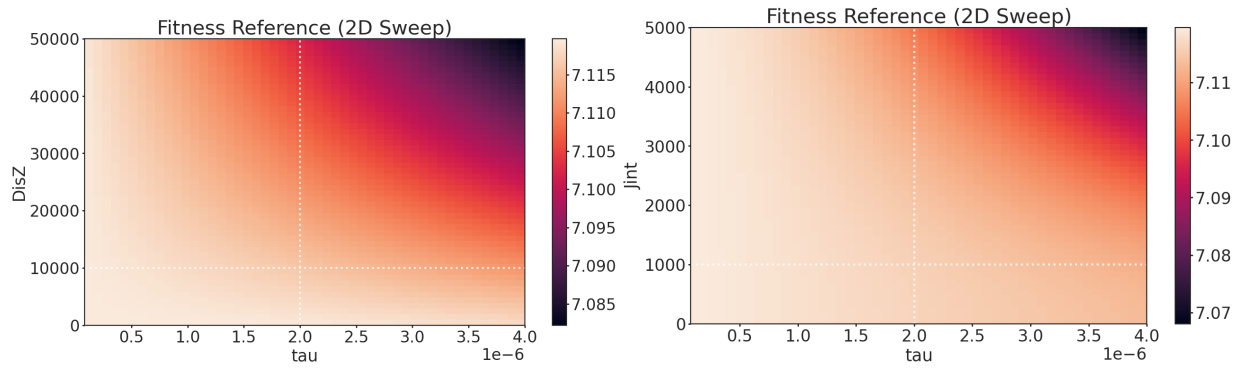


Figure 3.8: Fitness of DROID-60 sequence over a range of Hamiltonian parameters, in the presence of pulse error +5%. Left: α and τ are varied. Right: J_{int} and τ are varied. White dotted line indicates parameters used for GA execution and pulse sequence search

sequence obtains higher fitness values than DROID-60 in regions of parameter space bounded by a yellow dashed line. We also plot fitness improvements when the rotation angle error is changed. We conclude that the discovered pulse sequence only achieves higher fitness values in the presence of these pulse errors, but suffers greatly when they are absent i.e. when angle_err ~ 1. This is made more evident in Fig. 3.10; for fractional pulse angle errors close to 1 (no error), DROID-60 achieves excellent fitness values close to the maximum allowed (30), far superior than the numbers achieved by this pulse sequence.

3.5.4 Outlook

In this section, we implement a genetic algorithm implementation to dynamical decouple an interacting spin chain coupled by dipolar interactions, with magnetic disorder and pulse rotation angle errors included. These results are compared to a selected pulse sequence DROID-60, which has been experimentally demonstrated to effectively decouple strongly interacting solid-state spins in diamond while retaining a response to an AC magnetic field.

Initial results suggest that the performance of the GA can readily exceed that of the ref-

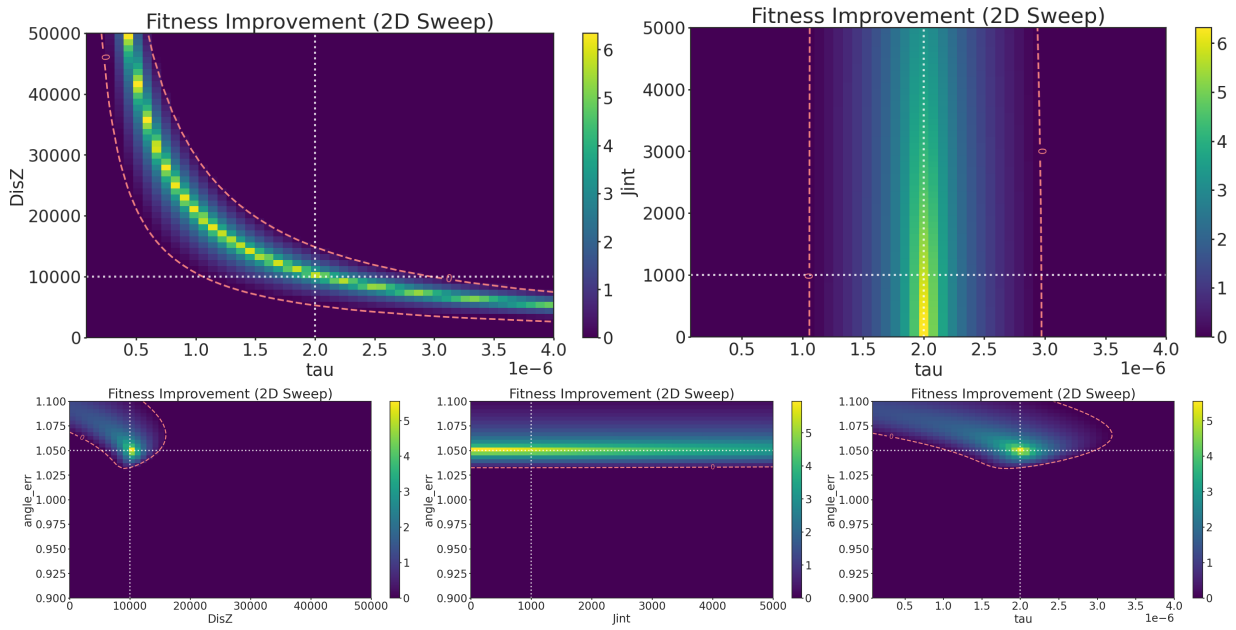


Figure 3.9: Improvement of pulse sequence fitness relative to DROID-60 sequence, over a range of Hamiltonian parameters, in the presence of pulse error +5%. Top left: α and τ are varied. Top right: J_{int} and τ are varied. Bottom left: Rotation angle error and α are varied. Bottom middle: Rotation angle error and J_{int} are varied. Bottom right: Rotation angle error and τ are varied. White dotted line indicates parameters used for GA execution and pulse sequence search. Pink dashed lines mark the fitness contour for DROID-60 (no improvement). Minimum colorscale value is manually capped at 0, which indicates no fitness improvement.

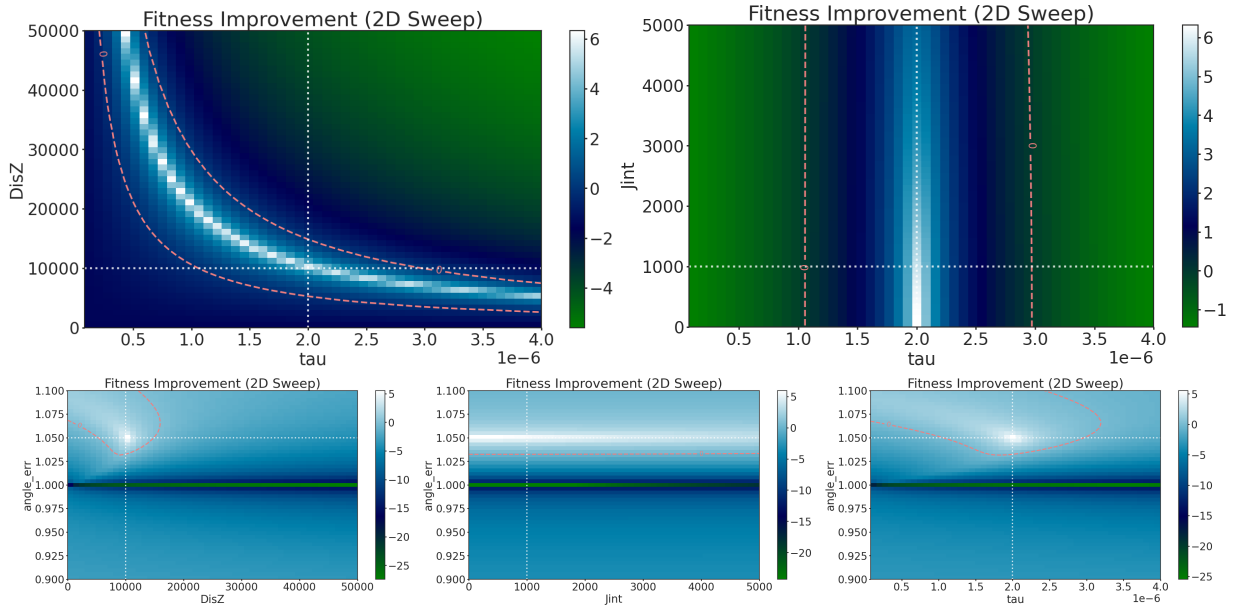


Figure 3.10: Same data as in Fig. 3.9 but with the colorscale edited so the minimum fitness improvement is not capped at 0.

erence pulse sequence. However, when the best-performing pulse sequence is tested on spin systems with adjusted Hamiltonian parameters, immediate reductions in fitness values are observed. These analyses suggest that the GA is over-fitting to the spin environment. Potential future developments include the consideration of varying spin environment parameters to more accurately emulate experimental settings, such as the use of a distribution of Hamiltonian parameters to represent a realistic ensemble of spin chains. In addition, alternate fitness functions may be developed to represent sensing performance (for example, by including a target signal field), enabling the GA to search for pulse sequences optimized for broadband or narrowband sensing.

Chapter 4: TEMPO: A Python Package for Time Evolution of Pulse Sequences in QuTiP

This chapter has been adapted from the preprint titled

***TEMPO: A Python Package for Time Evolution of Pulse Sequences in QuTiP**,*

with authors Jner Tzern Oon, Sonja A. Hakala, George A. Witt, and Ronald Walsworth,

posted on the Arxiv[108].

My contributions include project conception, code structure design, manuscript preparation, and supervision of all project components. I am proud to share a statement of equal contribution with Sonja A. Hakala, who led code development and worked with me on all structural decisions in addition to initial drafts of the manuscript. George Witt developed a suite of unit tests for the package. Sonja A. Hakala and I drafted the initial manuscript, with assistance from George A. Witt, and Ronald Walsworth. The work was performed under the supervision of Ronald L. Walsworth.

4.1 Summary

TEMPO (Time-dependent Evolution of Multiple Pulse Operations) offers accessible and efficient simulations of pulse sequences in Python, using the suite of master equation solvers available in the Quantum Toolbox in Python (QuTiP). It enables straightforward definition of

pulse sequence structures, including any underlying time-dependent Hamiltonians and pulse timing information, and faster simulations of pulse sequence dynamics (compared to naive implementations using QuTiP) while remaining compatible with the existing collection of QuTiP subpackages. Given the ubiquitous use of pulse sequences throughout quantum information/computing sciences, magnetic resonance studies, and quantum metrology, this work has immediate relevance to a wide array of research applications.

4.2 Statement of Need

Pulse sequences typically contain a series of discrete operations (pulses) using radio frequency, microwave, or optical fields. Their application for quantum control has an extensive history across atomic physics ([109]); nuclear magnetic resonance ([110]); solid-state spin systems used in quantum technologies ([9]); and a broad range of other platforms. In recent years, research into quantum technologies has driven the development of advanced software tools for numerical simulations of quantum systems ([111]). In particular, the QuTiP (Quantum Toolbox in Python) framework provides open-source tools for simulations of open quantum systems, and has received prolific use across numerous quantum applications ([112,113]), solvers that are native to QuTiP, TEMPO provides two key advantages for numerical simulations of pulse sequence dynamics.

Ease of Use: By incorporating both the characteristics of a Hamiltonian and its time constraints as necessary properties, pulses are first constructed individually, then collated to form a 'pulse sequence'. Time evolution is performed without the need to manually deactivate each pulse Hamiltonian outside its given time interval. Using pulse 'recipes' in TEMPO, the creation

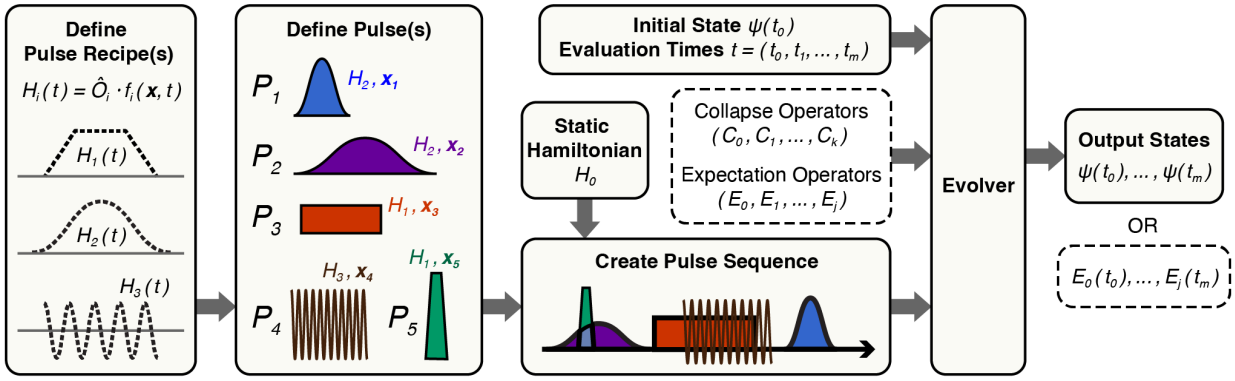


Figure 4.1: Outline of steps in TEMPO used to perform time evolution due to a pulse sequence. First, each pulse recipe contains the functional form of a time-dependent Hamiltonian $H_i(t) = \hat{O}_i \cdot f_i(\mathbf{x}, t)$. Pulse recipe(s) are used to create individual pulse(s) $P_1, \ldots, P_n$, with individual parameter settings $\mathbf{x_1}, \ldots, \mathbf{x_n}$, respectively. The pulses are organized into a sequence along with an optional static Hamiltonian. Next, the pulse sequence is provided to the Evolver, with an initial system state $\psi(t_0)$ and an array of evaluation times $(t_0, t_1, \ldots, t_m)$. Time evolution returns the state $\psi(t_0), \ldots, \psi(t_m)$ at these times; if operators $E_0, \ldots, E_j$ are provided, the operator expectation values $E_0(t_0), \ldots, E_j(t_0)$ are calculated instead.

of pulses with overlapping functional forms is streamlined, with parameters that can be tuned for individual pulses.

Faster Executions of Time Evolution: TEMPO organizes each pulse sequence as a series of time segments, preserving only the pulses that are active within each segment. This avoids overheads incurred by repeated inspections of inactive pulse(s), significantly speeding up evaluation of system evolution.

4.3 Usage

There are five main classes that make up a simulation in TEMPO: `Pulse_Recipe`, `Pulse`, `Hamiltonian`, `Pulse_Sequence`, and `Evolver`. In general, simulations can be easily executed by following the steps outlined below, which are illustrated in Figure 4.1.

1. Create a pulse recipe by defining the functional form of the time-dependent Hamiltonian

$$\hat{H}_i(t) = \hat{O}_i \cdot f_i(\mathbf{x}, t).$$

   The user provides an operator/matrix $\hat{O}$ and scalar function $f(\mathbf{x}, t)$ that depends on input parameters $\mathbf{x}$ and time t.

2. Create individual pulse(s) $P_1, \ldots, P_n$ by providing a pulse recipe, individual parameter settings $\mathbf{x_1}, \ldots, \mathbf{x_n}$, and pulse timing information.

3. Generate a pulse sequence by inputting the pulses $P_1, \ldots, P_n$, and optionally a time-independent (static) Hamiltonian.

4. To evolve the system in time, the pulse sequence is provided to the Evolver class along with the initial system state $\psi(t_0)$ (state vector or density matrix) and an array of time points $t_0, t_1, \ldots, t_m$. The system state is returned at these times. It is also possible to provide collapse operators for evolution using the Linbladian (open) master equation, along with operators for calculation of expectation values at these times.

5. The Evolver returns the system state $\psi(t_0), \ldots, \psi(t_m)$ or operator expectation values at the times provided.

4.4 Efficient Simulation Times

Using the existing master equation solvers available in QuTiP, TEMPO provides significant speedups in simulations of multi-pulse sequences. Without TEMPO, the simulation (wall-clock)

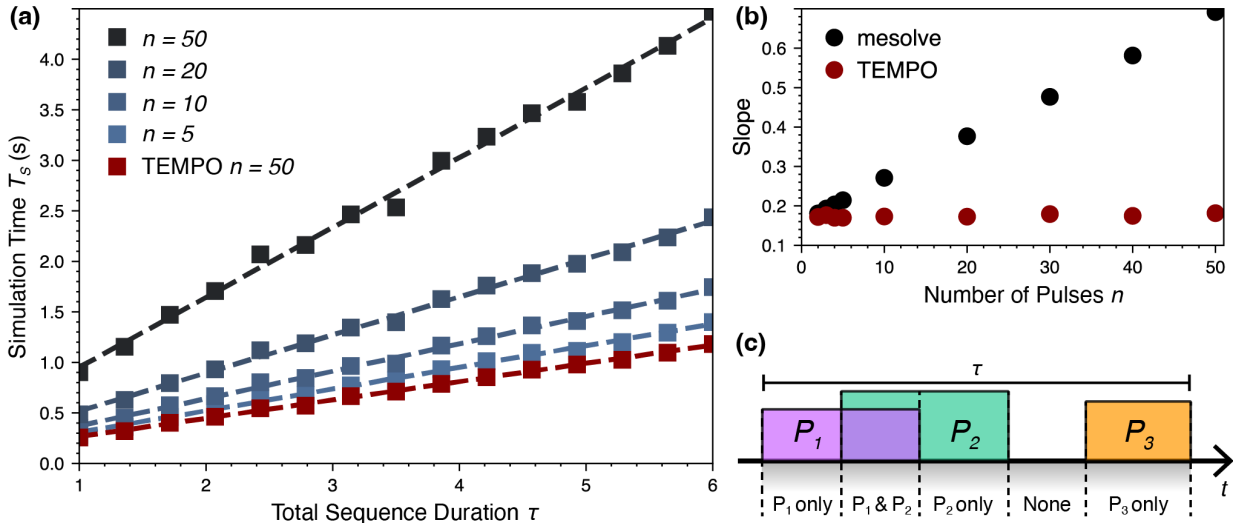


Figure 4.2: (a) Average simulation wall-clock time T_S as a function of pulse sequence duration τ without using TEMPO, for number of pulses $n = 5, 10, 20, 50$, shown in shades of blue. Average simulation times using TEMPO are shown in red at high pulse number count $(n = 50)$. Best-fit curves are displayed as dashed lines, with slope values plotted in (b) for measurements with TEMPO in red and without TEMPO (via QuTiP's mesolve function) in black. (c) Example visualization of pulse sequence segmentation performed by TEMPO, where $n = 3$ pulses result in $2n - 1 = 5$ time segments for solver evaluation. All measurements of T_S are average values over 20 repeated runs.

time T_S generally follow a linear trend with respect to 1) the total duration of a pulse sequence τ, and 2) the number of pulses n, roughly obeying

$$T_S \propto n \cdot \tau.$$

This dependence is highlighted in Figure 4.2(a), which shows measurements of the average simulation time with varying τ. Each set of measurements, shown as points of the same color, is performed for a fixed number of pulses n. Without TEMPO (shades of blue), the fitted slope for each set of measurements increases with n, exhibiting a linear dependence as shown in Figure 4.2(b). In contrast, the use of TEMPO (red) results in simulation times that are nearly independent of the number of pulses n, instead following $T_S \propto \tau$. As a result, the slope of each best-fit line using

TEMPO stays roughly constant with respect to n, as shown in Figure 4.2(b).

This advantage can be understood by considering how the solvers are constructed natively (without TEMPO). At each timestep during the pulse sequence, the solver typically inspects if each pulse is active at this time. This results in n checks at each timestep and thus the rough linear dependence of T_S with respect to n.

Instead, TEMPO first divides the pulse sequence into consecutive time segments for efficient time evolution, illustrated by an example in Figure 4.2(c). By creating segment breaks at the start and end of each pulse, TEMPO preserves only the active pulses in each of the $2n - 1$ segments. The solver is then executed consecutively across each segment, without any redundant checks of inactive pulses. As a result, simulation times are largely independent of n, besides minor overheads from repeated use of the solver.

Conclusion

This dissertation presents four studies that span experiment, simulation, theory, and algorithmic learning.

Chapter 1 characterized electron Ramsey envelope modulation (EREEM) in ^{15}NV diamond magnetometry, revealing how magnetic field misalignments produce unexpected envelope modulation effects that degrade sensor performance. Even modest misalignments significantly reduce the contrast at envelope modulation nodes, substantially degrading sensor performance when bias fields are not perfectly aligned with NV axes. The study showed that these modulations exhibit characteristic beat frequencies and amplitudes that depend on both the bias field magnitude and its orientation relative to the NV quantization axis. Through experimental validation and an intuitive vector model, we showed that this effect stems from electron-nuclear hyperfine coupling in transverse magnetic fields—a phenomenon well-studied in electron spin echo measurements but previously uncharacterized in Ramsey protocols. The analysis revealed systematic differences between experimental measurements and previous literature values for hyperfine parameters, and showed that double-quantum protocols can dramatically suppress these envelope modulations, explained by nearly antiparallel electronic effective fields that result in negligible modulation amplitudes.

Chapter 2 examined the breakdown of Average Hamiltonian Theory in quantum sensing

applications, revealing that this foundational framework fails in the operating regimes most relevant for solid-state quantum sensors. The work established that the conditions guaranteeing Magnus expansion convergence conflict with common operating conditions for magnetometry with solid-state spins, where energy scales and sequence durations routinely exceed convergence thresholds. In this regime, the Magnus series produces unreliable predictions for sensor performance, unless higher order contributions are suppressed. We developed exact methods for calculating sensor coupling factors and discovered that specific symmetry structures allow the Magnus approximation to remain accurate well beyond its guaranteed radius of convergence. In particular, rapid echo structures maintain AHT accuracy outside the convergence radius. These findings provide exact methods that are valid regardless of Magnus expansion convergence and establish new principles for pulse sequence design that operate beyond Average Hamiltonian Theory limitations.

Chapter 3 introduced Ensemble Cluster Sampling (ECS) to address the fundamental challenge of applying quantum control algorithms to experimental ensembles with parameter heterogeneities. We demonstrated that reinforcement learning and beam search algorithms, when trained on fixed parameter systems, produce pulse sequences that overfit to specific conditions and exhibit poor generalization performance across parameter distributions encountered in real experiments. The ECS framework was shown to overcome this limitation by representing experimental ensembles as collections of smaller clusters, each sampling from distributions of Hamiltonian parameters that capture realistic variations in dipolar coupling strengths, magnetic disorder, and control field imperfections. The study applied both reinforcement learning and beam search algorithms to the ECS framework, finding that training these algorithms on heterogeneous environments rather than single-instance systems yields substantial improvements in decoupling

fidelity when pulse sequences are evaluated on ensemble distributions. The results showed that ensemble training produces sequence patterns that are inherently robust to parameter variations, representing a fundamentally different class of control protocols optimized for high average fidelity across parameter spaces rather than peak performance under specific Hamiltonian parameter conditions.

Chapter 4 presented TEMPO, a Python package addressing computational barriers that limit the scope and scale of pulse sequence optimization studies. The software overcomes the scaling limitations of existing simulation approaches by organizing each pulse sequence as a series of time segments, preserving only the pulses that are active within each segment and avoiding redundant checks of inactive pulses. Without this approach, simulation times scale unfavorably with both the number of pulses and sequence duration, creating computational bottlenecks that restrict pulse sequence complexity. In contrast, TEMPO achieves simulation times that are nearly independent of the number of pulses, scaling primarily with sequence duration and providing substantial speedups for multi-pulse sequences. The package maintains compatibility with existing QuTiP workflows while incorporating structures familiar to experimentalists working with pulse generator hardware.

The work reveals the interconnected nature of theoretical, computational, and experimental challenges in quantum sensing. The exact methods developed for beyond-AHT analysis could be extended to higher-dimensional systems and more complex noise environments, potentially revealing new classes of robust pulse sequences. The ECS framework provides a platform for exploring quantum control in increasingly realistic systems with applications extending to quantum error correction and many-body quantum simulation. The computational insights from TEMPO could inform next-generation quantum control software that seamlessly integrates theoretical de-

sign with experimental implementation.

Several patterns emerge from these studies. First, the consistent discovery of unexpected experimental phenomena—envelope modulations that degrade sensitivity, theoretical frameworks that break down in relevant operating regimes, optimization algorithms that fail to generalize, and computational bottlenecks that limit practical applications—demonstrates that discrepancies between idealized models and experimental reality cannot be ignored. Second, the success of approaches that explicitly model experimental complexity, whether through exact theoretical methods or heterogeneous ensemble sampling, addresses the mismatch between ideal theoretical predictions and realistic experimental outcomes.

This work demonstrates that discrepancies between idealized models and experimental reality represent not obstacles but sources of scientific insight and technological innovation. Systematically characterizing and addressing these discrepancies help transition quantum sensing from proof-of-concept laboratory demonstrations to robust technologies capable of impactful applications. The path forward requires continued development of theoretical frameworks and computational tools that prioritize robustness, reliability, and scalability rather than treating them as secondary considerations.

Bibliography

[1] Pai Peng, Xiaoyang Huang, Chao Yin, Linta Joseph, Chandrasekhar Ramanathan, and Paola Cappellaro. Deep reinforcement learning for quantum hamiltonian engineering. *Physical Review Applied*, 18(2):024033, 2022. doi: 10.1103/physrevapplied.18.024033.

[2] Jonathan P. Dowling and Gerard J. Milburn. Quantum technology: the second quantum revolution. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 361(1809):1655–1674, June 2003. ISSN 1471-2962. doi: 10.1098/rsta.2003.1227. URL http://dx.doi.org/10.1098/rsta.2003.1227.

[3] John Preskill. Quantum computing in the nisq era and beyond. *Quantum*, 2:79, August 2018. ISSN 2521-327X. doi: 10.22331/q-2018-08-06-79. URL http://dx.doi.org/10.22331/q-2018-08-06-79.

[4] C. L. Degen, F. Reinhard, and P. Cappellaro. Quantum sensing. *Reviews of Modern Physics*, 89(3):035002, 2017. doi: 10.1103/revmodphys.89.035002.

[5] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone. Quantum-enhanced measurements: Beating the standard quantum limit. *Science*, 306(5700):1330–1336, November 2004. ISSN 1095-9203. doi: 10.1126/science.1104149. URL http://dx.doi.org/10.1126/science.1104149.

[6] Luca Pezzè, Augusto Smerzi, Markus K. Oberthaler, Roman Schmied, and Philipp Treutlein. Quantum metrology with nonclassical states of atomic ensembles. *Reviews of Modern Physics*, 90(3), September 2018. ISSN 1539-0756. doi: 10.1103/revmodphys.90.035005. URL http://dx.doi.org/10.1103/RevModPhys.90.035005.

[7] Antonio Acín, Immanuel Bloch, Harry Buhrman, Tommaso Calarco, Christopher Eichler, Jens Eisert, Daniel Esteve, Nicolas Gisin, Steffen J Glaser, Fedor Jelezko, Stefan Kuhr, Maciej Lewenstein, Max F Riedel, Piet O Schmidt, Rob Thew, Andreas Wallraff, Ian Walmsley, and Frank K Wilhelm. The quantum technologies roadmap: a european community view. *New Journal of Physics*, 20(8):080201, August 2018. ISSN 1367-2630. doi: 10.1088/1367-2630/aad1ea. URL http://dx.doi.org/10.1088/1367-2630/aad1ea.

[8] Edlyn V. Levine, Matthew J. Turner, Pauli Kehayias, Connor A. Hart, Nicholas Langellier, Raisa Trubko, David R. Glenn, Roger R. Fu, and Ronald L. Walsworth. Principles and techniques of the quantum diamond microscope. *Nanophotonics*, 8(11):1945–1973, sep 2019. doi: 10.1515/nanoph-2019-0209. URL https://doi.org/10.1515%2Fnanoph-2019-0209.

[9] John F. Barry, Jennifer M. Schloss, Erik Bauch, Matthew J. Turner, Connor A. Hart, Linh M. Pham, and Ronald L. Walsworth. Sensitivity optimization for NV-diamond magnetometry. *Rev. Mod. Phys.*, 92:015004, Mar 2020. doi: 10.1103/RevModPhys.92.015004. URL https://link.aps.org/doi/10.1103/RevModPhys.92.015004.

[10] Nabeel Aslam, Hengyun Zhou, Elana K. Urbach, Matthew J. Turner, Ronald L. Walsworth, Mikhail D. Lukin, and Hongkun Park. Quantum sensors for biomedical applications. *Nature Reviews Physics*, 5(3):157–169, February 2023. ISSN 2522-5820. doi: 10.1038/s42254-023-00558-3. URL http://dx.doi.org/10.1038/s42254-023-00558-3.

[11] Jner Tzern Oon, Jiashen Tang, Connor A. Hart, Kevin S. Olsson, Matthew J. Turner, Jennifer M. Schloss, and Ronald L. Walsworth. Ramsey envelope modulation in nv diamond magnetometry. *Physical Review B*, 106(5), August 2022. ISSN 2469-9969. doi: 10.1103/physrevb.106.054110. URL http://dx.doi.org/10.1103/PhysRevB.106.054110.

[12] Norman F. Ramsey. A molecular beam resonance method with separated oscillating fields. *Phys. Rev.*, 78:695–699, Jun 1950. doi: 10.1103/PhysRev.78.695. URL https://link.aps.org/doi/10.1103/PhysRev.78.695.

[13] L Rondin, J-P Tetienne, T Hingant, J-F Roch, P Maletinsky, and V Jacques. Magnetometry with nitrogen-vacancy defects in diamond. *Reports on Progress in Physics*, 77(5):056503, may 2014. doi: 10.1088/0034-4885/77/5/056503. URL https://doi.org/10.1088/0034-4885/77/5/056503.

[14] Priyadharshini Balasubramanian, Christian Osterkamp, Yu Chen, Xiuliang Chen, Tokuyuki Teraji, E. Wu, Boris Naydenov, and Fedor Jelezko. dc magnetometry with engineered nitrogen-vacancy spin ensembles in diamond. *Nano Lett.*, 19(9):6681–6686, 2019. doi: 10.1021/acs.nanolett.9b02993.

[15] Connor A. Hart, Jennifer M. Schloss, Matthew J. Turner, Patrick J. Scheidegger, Erik Bauch, and Ronald L. Walsworth. N-V–diamond magnetic microscopy using a double quantum 4-ramsey protocol. *Phys. Rev. Appl.*, 15:044020, Apr 2021. doi: 10.1103/PhysRevApplied.15.044020. URL https://link.aps.org/doi/10.1103/PhysRevApplied.15.044020.

[16] I. Lovchinsky, A. O. Sushkov, E. Urbach, N. P. de Leon, S. Choi, K. De Greve, R. Evans, R. Gertner, E. Bersin, C. Müller, L. McGuinness, F. Jelezko, R. L. Walsworth, H. Park, and M. D. Lukin. Nuclear magnetic resonance detection and spectroscopy of single proteins using quantum logic. *Science*, 351(6275):836–841, 2016. doi: 10.1126/science.aad8022. URL https://www.science.org/doi/abs/10.1126/science.aad8022.

[17] Nithya Arunkumar, Kevin S. Olsson, Jner Tzern Oon, Connor Hart, Dominik B. Bucher, David Glenn, Mikhail D. Lukin, Hongkun Park, Donhee Ham, and Ronald L. Walsworth. Quantum logic enhanced sensing in solid-state spin ensembles, 2022. URL https://arxiv.org/abs/2203.12501.

[18] D. R. Glenn, R. R. Fu, P. Kehayias, D. Le Sage, E. A. Lima, B. P. Weiss, and R. L. Walsworth. Micrometer-scale magnetic imaging of geological samples using a quantum diamond microscope. *Geochemistry, Geophysics, Geosystems*, 18(8):3254–3267, 2017. doi: https://doi.org/10.1002/2017GC006946. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017GC006946.

[19] Jennifer M. Schloss, John F. Barry, Matthew J. Turner, and Ronald L. Walsworth. Simultaneous broadband vector magnetometry using solid-state spins. *Phys. Rev. Appl.*, 10:034044, Sep 2018. doi: 10.1103/PhysRevApplied.10.034044. URL https://link.aps.org/doi/10.1103/PhysRevApplied.10.034044.

[20] Matthew J. Turner, Nicholas Langellier, Rachel Bainbridge, Dan Walters, Srujan Meesala, Thomas M. Babinec, Pauli Kehayias, Amir Yacoby, Evelyn Hu, Marko Lončar, Ronald L. Walsworth, and Edlyn V. Levine. Magnetic field fingerprinting of integrated-circuit activity with a quantum diamond microscope. *Phys. Rev. Appl.*, 14:014097, Jul 2020. doi: 10.1103/PhysRevApplied.14.014097. URL https://link.aps.org/doi/10.1103/PhysRevApplied.14.014097.

[21] K. Jensen, N. Leefer, A. Jarmola, Y. Dumeige, V. M. Acosta, P. Kehayias, B. Patton, and D. Budker. Cavity-enhanced room-temperature magnetometry using absorption by nitrogen-vacancy centers in diamond. *Phys. Rev. Lett.*, 112:160802, Apr 2014. doi: 10.1103/PhysRevLett.112.160802. URL https://link.aps.org/doi/10.1103/PhysRevLett.112.160802.

[22] John F. Barry, Matthew J. Turner, Jennifer M. Schloss, David R. Glenn, Yuyu Song, Mikhail D. Lukin, Hongkun Park, and Ronald L. Walsworth. Optical magnetic detection of single-neuron action potentials using quantum defects in diamond. *Proc. Natl. Acad. Sci. U.S.A.*, 113(49):14133–14138, nov 2016. doi: 10.1073/pnas.1601513113. URL https://doi.org/10.1073/pnas.1601513113.

[23] Till Lenz, Georgios Chatzidrosos, Zhiyuan Wang, Lykourgos Bougas, Yannick Dumeige, Arne Wickenbrock, Nico Kerber, Jakub Zázvorka, Fabian Kammerbauer, Mathias Kläui, Zeeshawn Kazi, Kai-Mei C. Fu, Kohei M. Itoh, Hideyuki Watanabe, and Dmitry Budker. Imaging topological spin structures using light-polarization and magnetic microscopy. *Phys. Rev. Applied*, 15:024040, Feb 2021. doi: 10.1103/PhysRevApplied.15.024040. URL https://link.aps.org/doi/10.1103/PhysRevApplied.15.024040.

[24] Scott T. Alsid, Jennifer M. Schloss, Matthew H. Steinecker, John F. Barry, Andrew C. Maccabe, Guoqing Wang, Paola Cappellaro, and Danielle A. Braje. A solid-state microwave magnetometer with picotesla-level sensitivity, 2022. URL https://arxiv.org/abs/2206.15440.

[25] L. G. Rowan, E. L. Hahn, and W. B. Mims. Electron-spin-echo envelope modulation. *Phys. Rev.*, 137(1A):A61–A71, January 1965. doi: 10.1103/physrev.137.a61. URL https://doi.org/10.1103/physrev.137.a61.

[26] Torsten Gaebel, Michael Domhan, Iulian Popa, Christoffer Wittmann, Philipp Neumann, Fedor Jelezko, James R. Rabeau, Nikolas Stavrias, Andrew D. Greentree, Steven Prawer, Jan Meijer, Jason Twamley, Philip R. Hemmer, and Jörg Wrachtrup. Room-temperature coherent coupling of single spins in diamond. *Nat. Phys.*, 2(6):408–413, may 2006. doi: 10.1038/nphys318. URL https://doi.org/10.1038%2Fnphys318.

[27] L. Childress, M. V. Gurudev Dutt, J. M. Taylor, A. S. Zibrov, F. Jelezko, J. Wrachtrup, P. R. Hemmer, and M. D. Lukin. Coherent dynamics of coupled electron and nuclear spin qubits in diamond. *Science*, 314(5797):281–285, October 2006. doi: 10.1126/science.1131871. URL https://doi.org/10.1126/science.1131871.

[28] Friedemann Reinhard, Fazhan Shi, Nan Zhao, Florian Rempp, Boris Naydenov, Jan Meijer, Liam T. Hall, Lloyd Hollenberg, Jiangfeng Du, Ren-Bao Liu, and Jörg Wrachtrup. Tuning a spin bath through the quantum-classical transition. *Phys. Rev. Lett.*, 108:200402, May 2012. doi: 10.1103/PhysRevLett.108.200402. URL https://link.aps.org/doi/10.1103/PhysRevLett.108.200402.

[29] Kejie Fang, Victor M. Acosta, Charles Santori, Zhihong Huang, Kohei M. Itoh, Hideyuki Watanabe, Shinichi Shikata, and Raymond G. Beausoleil. High-sensitivity magnetometry based on quantum beats in diamond nitrogen-vacancy centers. *Phys. Rev. Lett.*, 110: 130802, Mar 2013. doi: 10.1103/PhysRevLett.110.130802. URL https://link.aps.org/doi/10.1103/PhysRevLett.110.130802.

[30] H. J. Mamin, M. H. Sherwood, M. Kim, C. T. Rettner, K. Ohno, D. D. Awschalom, and D. Rugar. Multipulse double-quantum magnetometry with near-surface nitrogen-vacancy centers. *Phys. Rev. Lett.*, 113:030803, Jul 2014. doi: 10.1103/PhysRevLett.113.030803. URL https://link.aps.org/doi/10.1103/PhysRevLett.113.030803.

[31] Erik Bauch, Connor A. Hart, Jennifer M. Schloss, Matthew J. Turner, John F. Barry, Pauli Kehayias, Swati Singh, and Ronald L. Walsworth. Ultralong dephasing times in solid-state spin ensembles via quantum control. *Phys. Rev. X*, 8:031025, Jul 2018. doi: 10.1103/PhysRevX.8.031025. URL https://link.aps.org/doi/10.1103/PhysRevX.8.031025.

[32] Marcus W. Doherty, Neil B. Manson, Paul Delaney, Fedor Jelezko, Jörg Wrachtrup, and Lloyd C.L. Hollenberg. The nitrogen-vacancy colour centre in diamond. *Phys. Rep.*, 528(1):1–45, 2013. ISSN 0370-1573. doi: https://doi.org/10.1016/j.physrep.2013.02.001. URL https://www.sciencedirect.com/science/article/pii/S0370157313000562. The nitrogen-vacancy colour centre in diamond.

[33] V. M. Acosta, E. Bauch, M. P. Ledbetter, A. Waxman, L.-S. Bouchard, and D. Budker. Temperature dependence of the nitrogen-vacancy magnetic resonance in diamond. *Phys. Rev. Lett.*, 104:070801, Feb 2010. doi: 10.1103/PhysRevLett.104.070801. URL https://link.aps.org/doi/10.1103/PhysRevLett.104.070801.

[34] S. Felton, A. M. Edmonds, M. E. Newton, P. M. Martineau, D. Fisher, D. J. Twitchen, and J. M. Baker. Hyperfine interaction in the ground state of the negatively charged nitrogen vacancy center in diamond. *Phys. Rev. B*, 79:075203, Feb 2009. doi: 10.1103/PhysRevB.79.075203. URL https://link.aps.org/doi/10.1103/PhysRevB.79.075203.

[35] Bryan Myers. *Quantum decoherence of near-surface nitrogen-vacancy centers in diamond and implications for nanoscale imaging*. PhD thesis, University of California, Santa Barbara, 2016.

[36] Additional details are included in the supplemental material.

[37] Benjamin Smeltzer, Jean McIntyre, and Lilian Childress. Robust control of individual nuclear spins in diamond. *Phys. Rev. A*, 80:050302, Nov 2009. doi: 10.1103/PhysRevA.80.050302. URL https://link.aps.org/doi/10.1103/PhysRevA.80.050302.

[38] M. Steiner, P. Neumann, J. Beck, F. Jelezko, and J. Wrachtrup. Universal enhancement of the optical readout fidelity of single electron spins at nitrogen-vacancy centers in diamond. *Phys. Rev. B*, 81:035205, Jan 2010. doi: 10.1103/PhysRevB.81.035205. URL https://link.aps.org/doi/10.1103/PhysRevB.81.035205.

[39] A. Dréau, M. Lesik, L. Rondin, P. Spinicelli, O. Arcizet, J.-F. Roch, and V. Jacques. Avoiding power broadening in optically detected magnetic resonance of single nv defects for enhanced dc magnetic field sensitivity. *Phys. Rev. B*, 84:195204, Nov 2011. doi: 10.1103/PhysRevB.84.195204. URL https://link.aps.org/doi/10.1103/PhysRevB.84.195204.

[40] Thomas J. DiCiccio and Bradley Efron. Bootstrap confidence intervals. *Stat. Sci.*, 11(3): 189 – 228, 1996. doi: 10.1214/ss/1032280214. URL https://doi.org/10.1214/ss/1032280214.

[41] J.R. Johansson, P.D. Nation, and Franco Nori. QuTiP: An open-source python framework for the dynamics of open quantum systems. *Comput. Phys. Commun.*, 183(8):1760–1772, August 2012. doi: 10.1016/j.cpc.2012.02.021. URL https://doi.org/10.1016/j.cpc.2012.02.021.

[42] J.R. Johansson, P.D. Nation, and Franco Nori. QuTiP 2: A python framework for the dynamics of open quantum systems. *Comput. Phys. Commun.*, 184(4):1234–1240, April 2013. doi: 10.1016/j.cpc.2012.11.019. URL https://doi.org/10.1016/j.cpc.2012.11.019.

[43] Mo Chen, Masashi Hirose, and Paola Cappellaro. Measurement of transverse hyperfine interaction by forbidden transitions. *Phys. Rev. B*, 92:020101, Jul 2015. doi: 10.1103/PhysRevB.92.020101. URL https://link.aps.org/doi/10.1103/PhysRevB.92.020101.

[44] S Sangtawesin, C A McLellan, B A Myers, A C Bleszynski Jayich, D D Awschalom, and J R Petta. Hyperfine-enhanced gyromagnetic ratio of a nuclear spin in diamond. *New Journal of Physics*, 18(8):083016, aug 2016. doi: 10.1088/1367-2630/18/8/083016. URL https://doi.org/10.1088/1367-2630/18/8/083016.

[45] J. M. Taylor, P. Cappellaro, L. Childress, L. Jiang, D. Budker, P. R. Hemmer, A. Yacoby, R. Walsworth, and M. D. Lukin. High-sensitivity diamond magnetometer with nanoscale resolution. *Nature Physics*, 4(10):810–816, sep 2008. doi: 10.1038/nphys1075. URL https://doi.org/10.1038%2Fnphys1075.

[46] Jner Tzern Oon, Sebastian C. Carrasco, Connor A. Hart, George Witt, Vladimir S. Malinovsky, and Ronald Walsworth. Beyond average hamiltonian theory for quantum sensing, 2024. URL https://arxiv.org/abs/2410.04296.

[47] Wilhelm Magnus. On the exponential solution of differential equations for a linear operator. *Communications on Pure and Applied Mathematics*, 7(4):649–673, November 1954. ISSN 1097-0312. doi: 10.1002/cpa.3160070404. URL http://dx.doi.org/10.1002/cpa.3160070404.

[48] J A Oteo and J Ros. The magnus expansion for classical hamiltonian systems. *Journal of Physics A: Mathematical and General*, 24(24):5751–5762, December 1991. ISSN 1361-6447. doi: 10.1088/0305-4470/24/24/011. URL http://dx.doi.org/10.1088/0305-4470/24/24/011.

[49] Alain Connes and Dirk Kreimer. Renormalization in quantum field theory and the riemann-hilbert problem i: The hopf algebra structure of graphs and the main theorem. *Communications in Mathematical Physics*, 210(1):249, 2000. doi: 10.1007/s002200050779.

[50] D W Robinson. Multiple coulomb excitations of deformed nuclei. *Helvetica Physica Acta (Switzerland)*, 36, 1 1963. URL https://www.osti.gov/biblio/4703875.

[51] Jörg Eichler. Magnus approximation for k-shell ionization by heavy-ion impact. *Physical Review A*, 15(5):1856–1862, May 1977. ISSN 0556-2791. doi: 10.1103/physreva.15.1856. URL http://dx.doi.org/10.1103/PhysRevA.15.1856.

[52] S. Blanes, F. Casas, J.A. Oteo, and J. Ros. The magnus expansion and some of its applications. *Physics Reports*, 470(5–6):151–238, January 2009. ISSN 0370-1573. doi: 10.1016/j.physrep.2008.11.001. URL http://dx.doi.org/10.1016/j.physrep.2008.11.001.

[53] Fernando Casas. Sufficient conditions for the convergence of the magnus expansion. *Journal of Physics A: Mathematical and Theoretical*, 40(50):15001–15017, November 2007. ISSN 1751-8121. doi: 10.1088/1751-8113/40/50/006. URL http://dx.doi.org/10.1088/1751-8113/40/50/006.

[54] Andreas Brinkmann. Introduction to average Hamiltonian theory. i. basics. *Concepts Magnetic Resonance*, 45A(6), 2016. doi: 10.1002/cmr.a.21414.

[55] W.A.B Evans. On some applications of the magnus expansion in nuclear magnetic resonance. *Annals of Physics*, 48(1):72–93, June 1968. ISSN 0003-4916. doi: 10.1016/0003-4916(68)90270-4. URL http://dx.doi.org/10.1016/0003-4916(68)90270-4.

[56] U. Haeberlen and J. S. Waugh. Coherent averaging effects in magnetic resonance. *Physical Review*, 175(2):453–467, November 1968. ISSN 0031-899X. doi: 10.1103/physrev.175.453. URL http://dx.doi.org/10.1103/PhysRev.175.453.

[57] Mattias Edén and Malcolm H. Levitt. Pulse sequence symmetries in the nuclear magnetic resonance of spinning solids: Application to heteronuclear decoupling. *The Journal of Chemical Physics*, 111(4):1511–1519, July 1999. ISSN 1089-7690. doi: 10.1063/1.479410. URL http://dx.doi.org/10.1063/1.479410.

[58] M. Goldman, P. J. Grandinetti, A. Llor, Z. Olejniczak, J. R. Sachleben, and J. W. Zwanziger. Theoretical aspects of higher-order truncations in solid-state nuclear magnetic resonance. *The Journal of Chemical Physics*, 97(12):8947–8960, December 1992. ISSN 1089-7690. doi: 10.1063/1.463321. URL http://dx.doi.org/10.1063/1.463321.

[59] Eugène S. Mananga and Thibault Charpentier. Introduction of the Floquet-Magnus expansion in solid-state nuclear magnetic resonance spectroscopy. *The Journal of Chemical Physics*, 135(4), 2011. doi: 10.1063/1.3610943.

[60] Davide Candoli, Ilija K. Nikolov, Lucas Z. Brito, Stephen Carr, Samuele Sanna, and Vesna F. Mitrović. Pulsee: A software for the quantum simulation of an extensive set of magnetic resonance observables. *Computer Physics Communications*, 284:108598, March 2023. ISSN 0010-4655. doi: 10.1016/j.cpc.2022.108598. URL http://dx.doi.org/10.1016/j.cpc.2022.108598.

[61] Ozgur Sahin, Erica de Leon Sanchez, Sophie Conti, Amala Akkiraju, Paul Reshetikhin, Emanuel Druga, Aakriti Aggarwal, Benjamin Gilbert, Sunil Bhave, and Ashok Ajoy. High field magnetometry with hyperpolarized nuclear spins. *Nat Commun*, 13(1):5486, 2022. doi: 10.1038/s41467-022-32907-8.

[62] Soonwon Choi, Norman Y. Yao, and Mikhail D. Lukin. Dynamical engineering of interactions in qudit ensembles. *Physical Review letters*, 119(18):183603, 2017. doi: 10.1103/physrevlett.119.183603.

[63] Joonhee Choi, Hengyun Zhou, Helena S. Knowles, Renate Landig, Soonwon Choi, and Mikhail D. Lukin. Robust dynamic hamiltonian engineering of many-body spin systems. *Physical Review X*, 10(3):031002, 2020. doi: 10.1103/physrevx.10.031002.

[64] Michael F O'Keeffe, Lior Horesh, John F Barry, Danielle A Braje, and Isaac L Chuang. Hamiltonian engineering with constrained optimization for quantum sensing and control. *New Journal of Physics*, 21(2):023015, February 2019. ISSN 1367-2630. doi: 10.1088/1367-2630/ab00be. URL http://dx.doi.org/10.1088/1367-2630/ab00be.

[65] Hengyun Zhou, Joonhee Choi, Soonwon Choi, Renate Landig, Alexander M. Douglas, Junichi Isoya, Fedor Jelezko, Shinobu Onoda, Hitoshi Sumiya, Paola Cappellaro, Helena S. Knowles, Hongkun Park, and Mikhail D. Lukin. Quantum metrology with strongly interacting spin systems. *Physical Review X*, 10(3):031003, 2020. doi: 10.1103/physrevx.10.031003.

[66] Nithya Arunkumar, Kevin S. Olsson, Jner Tzern Oon, Connor A. Hart, Dominik B. Bucher, David R. Glenn, Mikhail D. Lukin, Hongkun Park, Donhee Ham, and Ronald L. Walsworth. Quantum logic enhanced sensing in solid-state spin ensembles. *Physical Review letters*, 131(10):100801, 2023. doi: 10.1103/physrevlett.131.100801.

[67] J. S. Waugh, L. M. Huber, and U. Haeberlen. Approach to High-Resolution nmr in Solids. *Physical Review letters*, 20(5):180, 1968. doi: 10.1103/physrevlett.20.180.

[68] Hengyun Zhou, Haoyang Gao, Nathaniel T. Leitao, Oksana Makarova, Iris Cong, Alexander M. Douglas, Leigh S. Martin, and Mikhail D. Lukin. Robust hamiltonian engineering for interacting qudit systems. *Physical Review X*, 14(3):031017, 2024. doi: 10.1103/physrevx.14.031017.

[69] Richard Jozsa. Fidelity for mixed quantum states. *Journal of Modern Optics*, 41(12): 2315–2323, December 1994. ISSN 1362-3044. doi: 10.1080/09500349414552171. URL http://dx.doi.org/10.1080/09500349414552171.

[70] S Blanes, F Casas, J A Oteo, and J Ros. Magnus and fer expansions for matrix differential equations: the convergence problem. *Journal of Physics A: Mathematical and General*, 31(1):259–268, January 1998. ISSN 1361-6447. doi: 10.1088/0305-4470/31/1/023. URL http://dx.doi.org/10.1088/0305-4470/31/1/023.

[71] Per Christian Moan and England Ew. Efficient approximation of sturm-liouville problems using lie-group methods. *Technical Report 1998/NA11*, 04 1999.

[72] Ana Arnal, Fernando Casas, and Cristina Chiralt. A general formula for the magnus expansion in terms of iterated integrals of right-nested commutators. *Journal of Physics Communications*, 2(3):035024, March 2018. ISSN 2399-6528. doi: 10.1088/2399-6528/aab291. URL http://dx.doi.org/10.1088/2399-6528/aab291.

[73] Nicholas J. Higham. *Functions of Matrices: Theory and Computation*. Society for Industrial and Applied Mathematics, January 2008. ISBN 9780898717778. doi: 10.1137/1.9780898717778. URL http://dx.doi.org/10.1137/1.9780898717778.

[74] Leigh S. Martin, Hengyun Zhou, Nathaniel T. Leitao, Nishad Maskara, Oksana Makarova, Haoyang Gao, Qian-Ze Zhu, Mincheol Park, Matthew Tyler, Hongkun Park, Soonwon Choi, and Mikhail D. Lukin. Controlling local thermalization dynamics in a floquet-engineered dipolar ensemble. *Physical Review letters*, 130(21):210403, 2023. doi: 10.1103/physrevlett.130.210403.

[75] Mi Lei, Rikuto Fukumori, Chun-Ju Wu, Edwin Barnes, Sophia Economou, Joonhee Choi, and Andrei Faraon. Quantum thermalization and Floquet engineering in a spin ensemble

with a
clock transition. *arXiv:2408.00252*, 2024. URL https://doi.org/10.48550/arXiv.2408.00252.

[76] Matthew Tyler, Hengyun Zhou, Leigh S. Martin, Nathaniel Leitao, and Mikhail D. Lukin. Higher-order methods for Hamiltonian engineering pulse sequence design. *Physical Review A*, 108(6):062602, 2023. doi: 10.1103/physreva.108.062602.

[77] Łukasz Cywiński, Roman M. Lutchyn, Cody P. Nave, and S. Das Sarma. How to enhance dephasing time in superconducting qubits. *Physical Review B*, 77(17):174509, 2008. doi: 10.1103/physrevb.77.174509.

[78] Götz S Uhrig. Exact results on dynamical decoupling by π pulses in quantum information processes. *New Journal of Physics*, 10(8):083024, August 2008. ISSN 1367-2630. doi: 10.1088/1367-2630/10/8/083024. URL http://dx.doi.org/10.1088/1367-2630/10/8/083024.

[79] Michael J. Biercuk, Hermann Uys, Aaron P. VanDevender, Nobuyasu Shiga, Wayne M. Itano, and John J. Bollinger. Optimized dynamical decoupling in a model quantum memory. *Nature*, 458(7241):996–1000, April 2009. ISSN 1476-4687. doi: 10.1038/nature07951. URL http://dx.doi.org/10.1038/nature07951.

[80] F. Poggiali, P. Cappellaro, and N. Fabbri. Optimal control for one-qubit quantum sensing. *Physical Review X*, 8(2):021059, 2018. doi: 10.1103/physrevx.8.021059.

[81] Hengyun Zhou, Leigh S. Martin, Matthew Tyler, Oksana Makarova, Nathaniel Leitao, Hongkun Park, and Mikhail D. Lukin. Robust higher-order hamiltonian engineering for quantum sensing with strongly interacting systems. *Physical Review letters*, 131(22):220803, 2023. doi: 10.1103/physrevlett.131.220803.

[82] Andrew D. Ludlow, Martin M. Boyd, Jun Ye, E. Peik, and P. O. Schmidt. Optical atomic clocks. *Reviews of Modern Physics*, 87(2):637–701, jun 2015. doi: 10.1103/revmodphys.87.637. URL https://doi.org/10.1103%2Frevmodphys.87.637.

[83] Adam L. Shaw, Ran Finkelstein, Richard Bing-Shiun Tsai, Pascal Scholl, Tai Hyun Yoon, Joonhee Choi, and Manuel Endres. Multi-ensemble metrology by programming local rotations with atom movements. *Nature Physics*, 20(2):195–201, January 2024. ISSN 1745-2481. doi: 10.1038/s41567-023-02323-w. URL http://dx.doi.org/10.1038/s41567-023-02323-w.

[84] Ivan H. Deutsch. Harnessing the power of the second quantum revolution. *PRX Quantum*, 1(2), November 2020. ISSN 2691-3399. doi: 10.1103/prxquantum.1.020101. URL http://dx.doi.org/10.1103/PRXQuantum.1.020101.

[85] Pablo M. Poggi, Nathan K. Lysne, Kevin W. Kuper, Ivan H. Deutsch, and Poul S. Jessen. Quantifying the sensitivity to errors in analog quantum simulation. *PRX Quantum*, 1(2), November 2020. ISSN 2691-3399. doi: 10.1103/prxquantum.1.020308. URL http://dx.doi.org/10.1103/PRXQuantum.1.020308.

[86] Christiane P. Koch, Ugo Boscain, Tommaso Calarco, Gunther Dirr, Stefan Filipp, Steffen J. Glaser, Ronnie Kosloff, Simone Montangero, Thomas Schulte-Herbrüggen, Dominique Sugny, and Frank K. Wilhelm. Quantum optimal control in quantum technologies. strategic report on current status, visions and goals for research in europe. *EPJ Quantum Technology*, 9(1), July 2022. ISSN 2196-0763. doi: 10.1140/epjqt/s40507-022-00138-x. URL http://dx.doi.org/10.1140/epjqt/s40507-022-00138-x.

[87] Lorenza Viola and Seth Lloyd. Dynamical suppression of decoherence in two-state quantum systems. *Physical Review A*, 58(4):2733–2744, October 1998. ISSN 1094-1622. doi: 10.1103/physreva.58.2733. URL http://dx.doi.org/10.1103/PhysRevA.58.2733.

[88] Lorenza Viola, Emanuel Knill, and Seth Lloyd. Dynamical decoupling of open quantum systems. *Physical Review Letters*, 82(12):2417–2421, March 1999. ISSN 1079-7114. doi: 10.1103/physrevlett.82.2417. URL http://dx.doi.org/10.1103/PhysRevLett.82.2417.

[89] Götz S. Uhrig. Keeping a quantum bit alive by optimized¡mml:math xmlns:mml="http://www.w3.org/1998/math/mathml" display="inline"¿¡mml:mi¿π¡/mml:mi¿¡/mml:math¿-pulse sequences. *Physical Review Letters*, 98(10), March 2007. ISSN 1079-7114. doi: 10.1103/physrevlett.98.100504. URL http://dx.doi.org/10.1103/PhysRevLett.98.100504.

[90] Marin Bukov, Alexandre G. R. Day, Dries Sels, Phillip Weinberg, Anatoli Polkovnikov, and Pankaj Mehta. Reinforcement learning in different phases of quantum control. *Physical Review X*, 8(3), September 2018. ISSN 2160-3308. doi: 10.1103/physrevx.8.031086. URL http://dx.doi.org/10.1103/PhysRevX.8.031086.

[91] G. Kucsko, S. Choi, J. Choi, P. C. Maurer, H. Zhou, R. Landig, H. Sumiya, S. Onoda, J. Isoya, F. Jelezko, E. Demler, N. Y. Yao, and M. D. Lukin. Critical thermalization of a disordered dipolar spin system in diamond. *Physical Review Letters*, 121(2), July 2018. doi: 10.1103/physrevlett.121.023601. URL https://doi.org/10.1103/physrevlett.121.023601.

[92] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *CoRR*, abs/1707.06347, 2017. URL http://arxiv.org/abs/1707.06347.

[93] Terry L. Rankin. Heuristics: Intelligent search strategies for computer problem solving. judea pearl. addison-wesley publishing. 1984. 382 pp. *AI Magazine*, 7(1):87–89, 1986. doi: https://doi.org/10.1609/aimag.v7i1.534. URL https://onlinelibrary.wiley.com/doi/abs/10.1609/aimag.v7i1.534.

[94] J.-S. Li and N. Khaneja. Ensemble control of bloch equations. *IEEE Transactions on Automatic Control*, 54(3):528–536, March 2009. ISSN 1558-2523. doi: 10.1109/tac.2009.2012983. URL http://dx.doi.org/10.1109/TAC.2009.2012983.

[95] Erik Bauch, Swati Singh, Junghyun Lee, Connor A. Hart, Jennifer M. Schloss, Matthew J. Turner, John F. Barry, Linh M. Pham, Nir Bar-Gill, Susanne F. Yelin, and Ronald L. Walsworth. Decoherence of ensembles of nitrogen-vacancy centers in diamond. *Physical Review B*, 102(13), October 2020. ISSN 2469-9969. doi: 10.1103/physrevb.102.134210. URL http://dx.doi.org/10.1103/PhysRevB.102.134210.

[96] Dmitry Budker and Michael Romalis. Optical magnetometry. *Nature Physics*, 3(4): 227–234, apr 2007. doi: 10.1038/nphys566. URL https://doi.org/10.1038%2Fnphys566.

[97] Manuel Vogel, Wolfgang Quint, and Wilfried Nörtershäuser. Trapped ion oscillation frequencies as sensors for spectroscopy. *Sensors*, 10(3):2169–2187, March 2010. ISSN 1424-8220. doi: 10.3390/s100302169. URL http://dx.doi.org/10.3390/s100302169.

[98] Kenneth R. Brown, John Chiaverini, Jeremy M. Sage, and Hartmut Häffner. Materials challenges for trapped-ion quantum computers. *Nature Reviews Materials*, 6(10):892–905, March 2021. ISSN 2058-8437. doi: 10.1038/s41578-021-00292-1. URL http://dx.doi.org/10.1038/s41578-021-00292-1.

[99] Kevin A. Gilmore, Matthew Affolter, Robert J. Lewis-Swan, Diego Barberena, Elena Jordan, Ana Maria Rey, and John J. Bollinger. Quantum-enhanced sensing of displacements and electric fields with two-dimensional trapped-ion crystals. *Science*, 373 (6555):673–678, August 2021. ISSN 1095-9203. doi: 10.1126/science.abi5226. URL http://dx.doi.org/10.1126/science.abi5226.

[100] Joonhyuk Kwon, William J. Setzer, Michael Gehl, Nicholas Karl, Jay Van Der Wall, Ryan Law, Matthew G. Blain, Daniel Stick, and Hayden J. McGuinness. Multi-site integrated optical addressing of trapped ions. *Nature Communications*, 15(1), May 2024. ISSN 2041-1723. doi: 10.1038/s41467-024-47882-5. URL http://dx.doi.org/10.1038/s41467-024-47882-5.

[101] A. Abragam and L. C. Hebel. The principles of nuclear magnetism. *American Journal of Physics*, 29(12):860–861, December 1961. ISSN 1943-2909. doi: 10.1119/1.1937646. URL http://dx.doi.org/10.1119/1.1937646.

[102] S. Kullback and R. A. Leibler. On information and sufficiency. *The Annals of Mathematical Statistics*, 22(1):79–86, March 1951. ISSN 0003-4851. doi: 10.1214/aoms/1177729694. URL http://dx.doi.org/10.1214/aoms/1177729694.

[103] Georg Pólya. Über den zentralen grenzwertsatz der wahrscheinlichkeitsrechnung und das momentenproblem. *Mathematische Zeitschrift*, 8(3–4):171–181, September 1920. ISSN 1432-1823. doi: 10.1007/bf01206525. URL http://dx.doi.org/10.1007/BF01206525.

[104] Hans Fischer. *A History of the Central Limit Theorem: From Classical to Modern Probability Theory*. Springer New York, 2011. ISBN 9780387878577.

doi: 10.1007/978-0-387-87857-7. URL http://dx.doi.org/10.1007/978-0-387-87857-7.

[105] R.S. Sutton and A.G. Barto. Reinforcement learning: An introduction. *IEEE Transactions on Neural Networks*, 9(5):1054–1054, September 1998. ISSN 1941-0093. doi: 10.1109/tnn.1998.712192. URL http://dx.doi.org/10.1109/TNN.1998.712192.

[106] Stuart Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach*. Prentice Hall Press, USA, 3rd edition, 2009. ISBN 0136042597.

[107] J. Carr. An introduction to genetic algorithms. 2014.

[108] Jner Tzern Oon, Sonja A. Hakala, George A. Witt, and Ronald Walsworth. Tempo: A python package for time evolution of pulse sequences in qutip, 2025. URL https://arxiv.org/abs/2502.11389.

[109] N.V. Vitanov, M. Fleischhauer, B.W. Shore, and K. Bergmann. Coherent manipulation of atoms molecules by sequential laser pulses, 2001. ISSN 1049-250X. URL http://dx.doi.org/10.1016/S1049-250X(01)80063-X.

[110] Malcolm H Levitt. Spin dynamics, May 2013.

[111] Mark Fingerhuth, Tomáš Babej, and Peter Wittek. Open source software in quantum computing. *PLOS ONE*, 13(12):e0208561, December 2018. ISSN 1932-6203. doi: 10.1371/journal.pone.0208561. URL http://dx.doi.org/10.1371/journal.pone.0208561.

[112] J.R. Johansson, P.D. Nation, and Franco Nori. Qutip: An open-source python framework for the dynamics of open quantum systems. *Computer Physics Communications*, 183 (8):1760–1772, August 2012. ISSN 0010-4655. doi: 10.1016/j.cpc.2012.02.021. URL http://dx.doi.org/10.1016/j.cpc.2012.02.021.

[113] J.R. Johansson, P.D. Nation, and Franco Nori. Qutip 2: A python framework for the dynamics of open quantum systems. *Computer Physics Communications*, 184(4): 1234–1240, April 2013. ISSN 0010-4655. doi: 10.1016/j.cpc.2012.11.019. URL http://dx.doi.org/10.1016/j.cpc.2012.11.019.