

Control and Characterization of the Central Spin System

by

Jiahui Chen

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

Waterloo, Ontario, Canada, 2025

© Jiahui Chen 2025

Examining Committee Membership

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

External Examiner: David Meyer
Professor, University of California San Diego

Supervisor(s): David Cory
Professor, University of Waterloo

Committee Member: Raffi Budakian
Professor, University of Waterloo

Committee Member: Joseph Emerson
Associate Professor, University of Waterloo

Committee Member: Dmitry Pushin
Associate Professor, University of Waterloo

Author's Declaration

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

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

Abstract

Precise, coherent, robust quantum control and characterization of quantum systems play important roles in the development of applications of quantum technologies. In particular, advancing the quality of control requires precise characterization, which, in turn, depends on the quality of control.

In the first part of the thesis, we introduce a general framework for designing efficient, precise, and robust quantum control strategies using effective Hamiltonian engineering. The methods enable designs that are robust to systematic control errors and variations in the Hamiltonian. The efficiency benefit of achieving control at zeroth order in the Magnus expansion is highlighted. Design tools, such as methods that identify the space of achievable effective Hamiltonians at each order from the Magnus expansion, are introduced. Objective functions for engineering arbitrary effective Hamiltonians are provided and can be used by numerical optimizers for control sequence design.

The second part of the thesis explores the characterization of general noise models based on experiments on a central spin system. The noise is probed through stimulated echo experiments, multi-dimensional correlation spectroscopy, and multi-quantum experiments to characterize system/environment correlation and environmental memory effects. Combined with Bayesian inference, these experiments provide quantitative measures of correlation growth, environmental mixing, and deviations from stochastic noise models. Measures that influence the choice of control schemes include non-Gaussianity, non-stationarity, and non-Markovianity. The multi-quantum experiments can also reveal an extended environment and show how the environmental mixing propagates quantum information throughout the environment.

Acknowledgements

I am deeply grateful to my supervisor, David Cory, for his guidance throughout this PhD. His dedication to mentorship has shaped the way I think and work. His high standards pushed me further than I thought possible.

I thank Xinhua Peng and Dmitry Pushin for recommending me to join the Cory group for my PhD. I also thank Dmitry Pushin, Joseph Emerson, and Raffi Budakian for serving on my committee and for their support during committee meetings. I thank David Meyer for agreeing to be my external examiner.

I cannot overstate how fortunate I feel to have met some of the most remarkable people during my time here. I thank Peter Sprenger for his friendship, support, and mentorship. His positivity, passion, and steady presence inspired me every day to take on challenges with courage. I thank Dmitry Akhmetzyanov for his friendship and encouragement. His kindness helped me through some of the hardest days. I also thank Maryam Mirkamali for many engaging discussions and her encouragement. I thank Saba Sadeghi for her support and helpful advice. I am grateful to Christopher Wyenberg for the many interesting conversations in the office, and to Fatemeh Fani Sani for sharing opportunities with me. I thank Mohamad Niknam for teaching me NMR during my Master's studies. I thank George Nicols for kindly helping take care of the magnet, and Sara Clark for her help around the lab. I thank Alexander Cooper-Roy, Carola Purser, and Dusan Sarenac for their helpful advice.

I thank my parents for their unwavering love and unconditional support, which have been the bedrock of my pursuit of this dream. I thank my grandparents for their trust, love, and the lessons they gave me. I'm sorry I wasn't by your side. I hope we'll meet again one day. I thank my other family members for their support and understanding. To my partner, Lucy: thank you for always standing by me. With you by my side, I fear nothing, no matter what comes next.

.

Dedication

To my parents

Table of Contents

List of Figures	xiii
List of Tables	xxii
1 Introduction	1
1.1 Robust quantum control via Hamiltonian engineering	2
1.2 Characterizing noise in open quantum systems using a central spin model . .	9
2 Engineering Precise and Robust Effective Hamiltonians	13
2.1 Introduction	13
2.2 Effective Hamiltonian engineering and toggling frame	17
2.3 Control Setup	18
2.4 Flowchart for effective Hamiltonian engineering	20
2.5 Examples	27
2.5.1 Robust Hadamard	27
2.5.2 Robust Hadamard with finite bandwidth	30
2.5.3 Robust control with nonlinear system	31
2.5.4 Two-qubit gates	35
2.5.5 Simulating a Tunable Quantum Harmonic Oscillator	36
2.5.6 Sensing with a Scalable Spin Amplifier	38
2.6 Conclusion	42

3	Engineering Higher-order Effective Hamiltonians	44
3.1	Engineering $\bar{H}^{(r-1)}$	45
3.2	Engineering $\bar{H}^{(r-1)}$ on qubit networks	46
3.3	Choosing Target Effective Hamiltonians	49
3.4	Examples	50
3.4.1	2nd-order decoupling	51
3.4.2	Engineering nonzero first-order effective Hamiltonian	52
3.5	Conclusion and Discussion	55
4	Efficiently Characterizing Quantum Information Flow, Loss and Recovery in the Central Spin System	56
4.1	Introduction	56
4.2	Theory	58
4.3	Method	64
4.4	Results	66
4.4.1	Measure change of local field	66
4.4.2	Measure change of sensitivity	67
4.4.3	Measure strength of H_e	68
4.5	Discussion	68
4.6	Conclusion	69
5	Characterizing Correlated Noise on a Central Spin Using Multi-Dimensional Experiments	74
5.1	Introduction	74
5.2	Theory	75
5.2.1	The Gauss-Markov Approximation	76
5.2.2	Measuring the joint distributions through multi-dimensional experiments	77
5.3	Method	80

5.4	Results	81
5.4.1	1D Experiment	81
5.4.2	2D Experiment	81
5.5	Projections of 2D experiments	86
5.6	3D Experiment	88
5.7	Discussion	93
5.7.1	1D experiment	93
5.7.2	2D experiment	93
5.7.3	Projections of 2D spectra	93
5.7.4	3D experiment	94
5.8	Conclusion	96
6	Probing the Structure of the Environment through the Central Spin Using Multi-quantum Experiments	97
6.1	Standard multi-quantum experiments	99
6.2	Coherence growth in the presence of environmental dynamics	102
6.3	Refocusing the environmental dynamics	105
6.4	Inferring the number of spins in the environment	105
6.5	Conclusion	107
7	Conclusion and Future Work	112
7.1	Summary	112
7.2	Future work	114
A	Supplemental information: Engineering Nonzero Effective Hamiltonians	116
A.1	Calculating $\mathcal{C}(\mathbf{g}_{\text{pri}}, H_{\text{pert}})$	116
A.2	Finding the Set of Achievable H_{eff}	118
A.3	Generating random points for controllability test	122
A.4	Discretization of $\vec{B}(t)$ for Integration	124

A.5	Calculating $\mathcal{C}$ -integrals for shaped pulse	125
A.5.1	$\mathcal{C}$ -integrals of a square pulse	126
A.5.2	$\mathcal{C}$ -integrals of a shaped pulse	127
A.5.3	Cross $\mathcal{C}$ -integrals	127
A.6	Robustness against Control Errors	128
A.6.1	Robustness in $U_{\text{pri}}(T_{\text{seq}})$	129
A.6.2	Robustness in $U_{\text{pert}}(T_{\text{seq}})$	130
A.6.3	Control Robustness is Always Achievable	132
A.7	Higher-order corrections	133
A.8	Numerical Setup For Sequence Design	134
A.8.1	Objective functions	134
A.9	Optimization algorithms	135
A.10	Supplementary Information on Examples	139
A.10.1	Broadband robust sequences	139
A.10.2	Robust control in nonlinear systems	140
A.10.3	Bloch-Siegert Shift	141
A.10.4	Leakage in multi-qubit systems	144
A.10.5	Examples in spectroscopy	145
A.10.6	Spin System with Two Clusters	147
B	Supplementary: Efficiently characterizing quantum information flow, loss and recovery in the central spin system	149
B.1	Coherence Pathway	149
B.2	Detailed analysis of the stimulated echo	151
B.3	Bayesian analysis for the stimulated echo	154

C	Supplementary: Experimentally Characterizing Fluctuation of Noise in the Central Spin System	158
C.1	Dynamics of the Central Spin System	158
C.2	Properties of the Ornstein–Uhlenbeck Process	161
C.3	Bayesian analysis of time-domain data	163
C.4	Time-domain data and pure-phase spectrum	164
C.5	Bayesian analysis	165
C.5.1	Multi-dimensional experiment	167
D	Supplementary: Probing the structure of the environment through the central spin using multi-quantum experiments	169
D.1	Dynamics without mixing	169
D.2	Pulse Sequences	172
E	Randomized Benchmarking of the Central Spin System	175
E.1	Randomized Benchmarking	175
E.1.1	Benchmarking the Central Spin System	177
F	Discrete Methods for Pulse Engineering	182
F.1	Pulse Design Using Tree Search	182
F.2	Sequence with Finite Gate Set	184
F.3	Genetic Algorithm for Sequence Optimization	186
G	Optimization Algorithms for Pulse Engineering	190
G.1	Nelder-Mead Simplex Method	190
G.2	Conjugate Gradient Method	191
G.3	Quadratic Line Search	191
G.4	Nonlinear Conjugate Gradient	194

H Markov Approximation of the Environmental Dynamics	200
H.1 Non-Markovianity of the Central Spin System	200
H.2 Stochastic Description using the Gauss-Markov Process	202
References	204

List of Figures

1.1	The WAHUA-4 sequence and the toggling-frame Hamiltonian of H_{zz} . Here, $H_{\alpha\alpha} = 3\sigma_{\alpha}\sigma_{\alpha} - \sigma \cdot \sigma$ for $\alpha = x, y$ and z . The total length of the sequence is $T_{\text{seq}} = 6\tau$. Since $H_{xx} + H_{yy} + H_{zz} = 0$, the sequence decouples H_{zz} to zeroth order when the pulse widths are negligible. $\bar{H}^{(1)}$ is also zero since the sequence is symmetric ($H_{\text{tog}}(t) = H_{\text{tog}}(T_{\text{seq}} - t)$).	4
1.2	Control landscapes of a $\pi/2$ x rotation implemented as a simple square $\pi/2$ rotation, a composite $\pi/2$ rotation (BB1) and an optimized robust pulse. The composite rotation maintains a high fidelity over a larger range of Rabi field variations (ϵ) compared with the square pulse. Using numerical optimization, robust pulses can be found that outperform the composite pulse and also incorporate robustness against other errors. Here, the optimized pulse is also robust against variations in energy offset ($\Delta\omega$) of the qubit. . .	7
1.3	Time-dependent amplitude $\omega_1(t)$ and phase $\phi(t)$ of the robust optimized pulse. Wait periods are included at the beginning and the end to allow easy composition between different pulses. In a more complicated implementation, the pulse would be designed for the control system's bandwidth and nonlinearities.	8
1.4	Dynamics of quantum control and evolution under Markovian and non-Markovian noise. Here, the average gate fidelity of a square $\pi/2$ pulse and the concurrence of a Bell state are shown. Insets provide a close-up of the initial decay behavior.	10
1.5	Schematic of a CS model. Here, a spin-1/2 CS is coupled to an external environment of many, interacting spin-1/2 particles.	11

2.1	Illustration of a very simplified control system. The input signals $\{a_k(t)\}$ are designed by the user. The control system outputs a continuous control field $\vec{B}(t)$ acting on the qubit system in the lab frame. The design of $\{a_k(t)\}$ should account for the model of the system so that it captures the distortion of the hardware along with additive noise.	19
2.2	Schematic for finding achievable scaling factors. The set of achievable zeroth-order average Hamiltonians $\mathcal{O}'(\mathbf{g}_{\text{pri}}, H_{\text{pert}})$ is a convex set in $\mathcal{C}(\mathbf{g}_{\text{pri}}, H_{\text{pert}})$. The convex set is approximated by the polygon formed by $\{\vec{v}_i = U_i H_{\text{pert}} U_i^\dagger\}$. The achievable scaling factors for H_{target} can then be found via linear programming.	25
2.3	Hadamard gate (robust to Rabi and offset variations) and the contour and 3D plots of its control landscape. The bottom plots are the control landscapes of a non-robust Hadamard gate with the same T_{seq}	29
2.4	Bloch sphere trajectories of the state $ 0\rangle\langle 0 $ under the robust Hadamard for different random values of ϵ and $\Delta\omega$ drawn from their distributions. Even though the trajectories are different, the final states are all close to $ +\rangle\langle + $	30
2.5	Hadamard gate (robust to W and δ variations) and the contour and 3D plots of its control landscape. The blue and orange curves represent the control parameters and relative amplitudes of $B_x(t)$ and $B_y(t)$ respectively. The bottom plots show the control landscapes of a non-robust sequence.	32
2.6	A simple resonant circuit with $R_s = 50 \text{ } \Omega$, $R = 0.01 \text{ } \Omega$, $C_m = 11 \text{ fF}$, $C_t = 1.479 \text{ pF}$, and $L = 170 \text{ pH}$. The resonance frequency is 10 GHz	33
2.7	The control parameters (blue) and the x and y components of $\vec{B}(t)/\omega_{\text{max}}$ (orange) at $\alpha_L = 0$ of the sequence that is nonrobust (top) and robust (bottom) to variations in α_L	34
2.8	Robustness plot of different sequences against variations of the nonlinear coefficient α_L	35
2.9	Minimal length of sequence for a CNOT gate with different detunings and interaction forms.	35
2.10	Time-dependent control parameters of an optimized sequence that engineers a tunable QHO.	37
2.11	Control landscape (left) and simulation of the evolution (right) of the state $(n = 1\rangle + n = 3\rangle)/\sqrt{2}$ under the sequence shown in Fig. 2.10.	38

2.12	Time-dependent control parameters of the sequence that achieves a scalable spin amplifier through engineering a conditioned decoupling.	40
2.13	Control landscape of the sequence in Fig. 2.12. The parameter $\Delta\omega$ determines the width of the distribution of the coupling strength between the target qubit and the spin amplifier.	40
2.14	Simulated survival probabilities of the states $ \psi_{0/1}\rangle$ under the sequence in Fig. 2.12.	41
3.1	Nonisomorphic graphs associated with different Magnus terms.	48
3.2	Examples of network topologies and the corresponding loop-free graphs contributing to $\bar{H}^{(2)}$. From left to right: spin pairs, nearest-neighbor spin chain, square spin lattice, and all-to-all coupled spin ensemble.	49
3.3	Control parameters of the sequence for 2nd-order decoupling. The horizontal and vertical axes are time (μs) and amplitude.	51
3.4	Simulated signals for different sequences. For reference, we also show the simulated free evolution signal.	52
3.5	Control parameters for engineering a nonzero three-body average Hamiltonian at first order.	54
3.6	Simulated echoes of the sequence for 1st-order average Hamiltonian engineering.	55
4.1	Schematic of a central spin system. The CS is a spin-1/2 particle and interacts with an environment consisting of spin-1/2 particles. The system can provide a complete model of an open quantum system of a qubit.	57
4.2	Measuring multi-time joint distributions of local fields using nD experiments. The 1D experiment yields the distribution of local field when there is no mixing. The 2D experiment yields the two-time joint distribution of local fields before and after a period of mixing, which measures the change of local field. nD experiments ($n \geq 3$) can then produce the n -time joint distributions that measure more subtle properties of the memory of the environment.	59
4.3	The second moments of the spectra of the STE experiments by Fourier transform with respect to either t or Δ give the autocorrelation functions $\langle H'_{cs/e}(\Delta)H'_{cs/e}(0) \rangle$ and $\langle H'_e(t)H'_e(0) \rangle$, allowing efficient measurement of change of the local field and correlation growth.	60

4.4	Schematic of the forced Rayleigh scattering (top) and stimulated echo (bottom) experiments. Like the FRS experiment, the STE experiment consists of three steps: encoding, mixing, and decoding, which correspond to grating, diffusion, and scattering in the FRS experiment. The quantum information initially resides at the CS. By comparing it with the echo after decoding, one can learn about mixing.	61
4.5	Coherence pathway for STE. The empty rectangle represents the effective excitation pulse corresponding to the cross-polarization. By combining experiments with carefully chosen phases, one can achieve the desired coherence pathway shown here which gives the unitary propagator of the STE experiment.	65
4.6	The pulse sequences for the (a) STE and the (b) decoupling experiments. CP and RF refer to cross-polarization and RF irradiation. Saturation with $\pi/2$ pulses is applied to the environment spins prior to the start of each experiment to prevent correlation between experiments. Spin temperature alteration [182] is incorporated in the phase cycling to remove artifacts from the cross-polarization. During t_d , a continuous weak measurement is performed on the CS. A strong continuous wave is applied to the environment during this period to decouple the CS from the environment. Then an integration is taken over t_d to give a complex signal point.	70
4.7	Time-domain data (left) and spectra (right) of the STE experiments for (a) different Δ , (b) different t , and (c) those of the decoupling experiments for different ω_{cs} . Here, only real parts of the data are shown. (a) and (b) are obtained from the STE experiments (Fig. 4.6 (a)), while (c) is from the decoupling experiments (Fig. 4.6 (b)).	71
4.8	Measure of (a) second moments, (b) linewidths and (c) second moments derived from the linewidth of the STE experiments as functions of Δ . The spectra are obtained by Fourier transform with respect to t . The data are obtained from their corresponding posterior predictive distributions. The error bars correspond to the 1/4 and 3/4 quantiles. The solid line represents the best fit using the posterior mean for the optimal model with $L = 2$ (Eq. (4.11)).	72
4.9	Probability histogram plots of the posterior distributions of C_2 , C_4 , C_6 , D_2 , D_4 , and D_6	72

4.10	Measure of (a) linewidths and (b) second moments derived from the linewidths of the STE experiments as functions of t . They correspond to spectra obtained by Fourier transform of S_{STE} with respect to Δ	73
4.11	Linewidths of the spectra measured from the decoupling experiments as a function of the strength ω_{dec} of the RF irradiation.	73
5.1	Time-domain data and corresponding spectrum of a 1D experiment. The solid lines are the best Gaussian fits.	81
5.2	Measured 2D spectra for different Δ . The plots are normalized to the maximal point of the spectrum of $\Delta = 0\mu s$	82
5.3	Simulated 2D spectra for different Δ . The plots are normalized to the maximal point of the spectrum of $\Delta = 0\mu s$	83
5.4	3D views of Measured 2D spectra with various Δ . (a) $\Delta = 0\mu s$. (b) $\Delta = 50\mu s$. (c) $\Delta = 200\mu s$. (d) $\Delta = 1.5s$	84
5.5	MSE of the fits to the bivariate Gaussian distribution for different Δ	84
5.6	Differences of projections of 2D spectra along $\omega_1 = 0$ and $\omega_2 = 0$ for different Δ . The one at $\Delta = 50\mu s$ shows that the dynamics are non-stationary.	85
5.7	MSE of fits to bivariate Gaussian distributed local field of the STE and STAE data as a function of Δ	86
5.8	Absolute mean difference (AMD) of bivariate Gaussian fits of the STE and STAE data.	86
5.9	Projections of 2D spectra along $\omega_1 = \omega_2$ for different Δ . The black solid lines are the best Gaussian fits.	87
5.10	Projections of 2D spectra along $\omega_1 = -\omega_2$ for different Δ . The black solid lines are the best Gaussian fits.	87
5.11	Coefficients of determination (R^2) of Gaussian fits of the projections of the 2D spectra along different directions for different Δ . Larger R^2 corresponds to better fits. A logarithmic scale of Δ is used to better represent the initial behaviors.	88
5.12	Measured autocorrelation function $\langle \Omega_2 \Omega_1 \rangle$ (left) with the best exponential fit (dashed line) and the linewidths of the projections (right) with the best exponential fits (solid lines) and the best Gaussian fit (dashed line). The blue squares and the orange dots are for projections along $\omega_2 = \omega_1$ and $\omega_2 = -\omega_1$, respectively.	89

5.13	3D spectra (left 3) and their sectional views (right 3) for different $\Delta_2 = \Delta_1$ (from left to right: 0, $20\mu s$ and $100\mu s$). From top to bottom: Measured spectra, Markov reconstructions, and best Gaussian fit.	89
5.14	Projections of the 3D spectra along (1, 1, 1). From top to bottom: Measured spectra, Markov reconstructions, and best Gauss-Markov fit. From left to right: $\Delta_1 = \Delta_2 = 0$, $20\mu s$ and $100\mu s$	90
5.15	Residual correlation for different 3D experiments.	91
5.16	3D displays (left half, one contour at 0.5 maximum) and sectional views (right half, contours at 0.25, 0.5 and 0.75 maximum) of measured (top) and simulated (bottom) spectra for $\Delta_2 = 200\mu s/\Delta_1 = 30\mu s$ and $\Delta_1 = 30\mu s/\Delta_2 = 200\mu s$	92
5.17	Posterior distributions of correlations $\langle \Omega(\Delta_1 + \Delta_2)\Omega(0) \rangle$. The posterior mean estimates are $0.31^{+0.04}_{-0.07}$ ($200\mu s/30\mu s$) and $0.322^{+0.028}_{-0.031}$ ($30\mu s/200\mu s$). As a comparison, $\langle \Omega_3\Omega_1 \rangle$ for all experiments are plotted on the left.	92
5.18	Spectra of decoupling experiments. Black lines are the best Gaussian fits. . .	95
6.1	Schematic of learning the environment through a CS. The environment mixing may destroy the CS/environment interaction, reducing learnability of the environment. If refocused, however, the mixing may help reveal a larger environment as it can reach parts not directly observable via $H_{cs/e}$	100
6.2	Time-domain data of standard MQ experiments.	101
6.3	Amplitudes of different coherence orders (from -17 to 18) for different t	101
6.4	Amplitudes of different coherence orders as functions of t	102
6.5	Modified magic echo sequence. a and b are the durations of delays and the spin-locking field, respectively. By varying a and b , different effective Hamiltonians can be achieved under the sequence.	102
6.6	Data of the LE experiments. From left to right: $2a - b = 0, 1, 2, 3, 4, 5, 6 \mu s$. Top to bottom: (before scaling) $t = 96, 192, 288, 384, 480, 576, 672 \mu s$. For large t and $2a - b$ (right bottom corner), the signal-to-noise ratio goes to zero.	103
6.7	Echo intensity as functions of t . The decay becomes faster as the amount of mixing ($2a - b$) increases for large t ($\geq 200\mu s$).	104
6.8	Entropy of coherence amplitudes of LE experiments. Left: Entropy as a function of t for different mixing ($2a - b$). Right: Entropy for fixed t as a function of $2a - b$	105

6.9	Intensity (vertical axes) of different coherence orders k as functions of evolution time t (μs) (horizontal axes) for different amount of mixing. $k = 0$ (blue dots), 1 (orange squares), 2 (green diamonds), 3 (red up triangles) and 4 (purple down triangles).	106
6.10	Intensity (vertical axes) of the same coherence order for different amount of mixing as functions of evolution time $t(\mu s)$. $2a - b = 0 \mu s$ (blue dots), $1\mu s$ (orange squares), $2 \mu s$ (green diamonds), $3 \mu s$ (red up triangles) and $4 \mu s$ (purple down triangles). The crossings of the different lines (e.g., the blue and orange line for $k = 1$) indicate that the mixing aids in growth of coherence for short t	107
6.11	Measured amplitudes of different coherence orders for different t and mixing time Δ . From left to right: $\Delta = 15, 30, 45, 60, 75, 90 \mu s$. Top to bottom: $t = 72, 144, 216, 288 \mu s$. Increased mixing (Δ) introduces more structure to the distributions of coherence orders and spreads it toward higher-order coherence.	108
6.12	Measured entropy of distributions of coherence orders.	108
6.13	Zoomed-in versions of distributions of coherence orders. The bottom row exhibits a non-convex feature (the first and second rows) and longer tail (the last row) of the distributions.	109
6.14	Best fit (orange squares) to the data (blue dots) for $t = 672 \mu s$. The improvement of the fitting of the last one is diminishing and not justified under the present uncertainty.	110
6.15	AIC for different number of spins N . It suggests possible over-fitting at $N = 16$	110
7.1	Schematic for controlling and characterizing open quantum systems. The black solid boxes are subjects related to the control and characterization of open quantum systems in the thesis. The red boxes represent potential directions for future work. Arrows indicate how one element informs or enables the next step in the research.	113
A.1	A network of three clusters. Different clusters are controlled differently (not necessarily independently) and have their own internal Hamiltonians.	118

A.2	Schematic for finding the range of achievable scaling factors. The set of achievable zeroth-order average Hamiltonians is a convex set in $\mathcal{C}(\mathbf{g}_{\text{pri}}, H_{\text{pert}})$. By sampling random points from the set of extreme points $\mathcal{Q}(\mathbf{g}_{\text{pri}}, H_{\text{pert}})$, the convex set can be approximated with a polygon. The achievable scaling factors for a desired H_{target} can then be found using linear programming. . .	120
A.3	Volume of the simulated $\bar{\mathcal{Q}}$ as a function of the number of samples. It saturates after around 500 samples.	124
A.4	Illustration of discretization of output field parameters. The control parameters $\{a_k(t)\}$ are mapped to $\vec{B}(t)$ through the model of the control system. Then the continuous field $\vec{B}(t)$ is discretized into M steps for integration. .	125
A.5	Different types of legalizers for box constraint: Clip, bounce and reflect. The clip legalizer stops the particle when it touches the boundary while the bounce and reflect legalizers allow the particle to continue its movement according to different rules.	136
A.6	Control landscapes of non-robust, 0th-order, 1st-order, 2nd-order and broadband robust sequences. The broadband sequences achieve a much larger robust region than the 2nd-order robust sequence.	139
A.7	A simple resonant circuit with $R_s = 50\Omega$, $R = 0.01\Omega$, $C_m = 11\text{fF}$, $C_t = 1.479\text{pF}$, and $L = 170\text{pH}$. The resonance frequency is 10GHz. We choose $\kappa_i=1\text{V}^{-1}$ and $\kappa_o=2\text{A}^{-1}$	140
A.8	Illustration of x (left) and y (right) components of $\Delta H(t)$ (top) and $\Delta H'(t)$ (bottom), both blue, of Bloch-Siegert shift at $\omega_r = 1.6\text{MHz}$. The red area corresponds to the input parameters. Note that the error $\Delta H'(t)$ grows linearly over time.	143
A.9	Illustration of the Bloch-Siegert error (blue) when discretized using the middle-point approximation at $\omega_r = 1.6\text{MHz}$. It greatly overestimates the intensity of the error.	143
A.10	Control landscapes of robust and non-robust sequences against variations in the resonance frequency ω_r . Both sequences account for the Bloch-Siegert shift.	144
A.11	Regions of achievable effective Hamiltonians. The dots are $(s_{\pm} \cos \theta, s_{\pm} \sin \theta)$ for $\theta = 0, \frac{\pi}{64}, \dots, \frac{63\pi}{64}$. They trace out the convex set for all achievable 0th-order average Hamiltonians.	148
B.1	Example of coherence pathway.	152

B.2	Posterior distributions of derived second moments from the measured linewidths of the spectra. They are well approximated by Gaussian distributions.	156
B.3	Information criterion values for models of different order L for $\langle H'_{cs/e}(\Delta)H'_{cs/e}(0) \rangle$ (left) and $\langle H'_e(t)H'_e(0) \rangle$ (right).	157
C.1	Coherence transfer pathways for 2D and 3D experiments. The solid lines achieve backward evolutions in the corresponding periods while the dashed lines achieve forward evolutions. By combining forward and backward evolutions, one cosine- or sine-modulation in that period can be selected.	166
C.2	Preparation pulse sequence. The phase is denoted in analogy to a $\pi/2$ pulse, indicating that the state after preparation is $(\sigma_x^{\text{CS}} \sin \phi - \sigma_y^{\text{CS}} \cos \phi) \otimes \mathbb{1}^{\otimes N}$. A saturation consisting of 50 $\pi/2$ pulses separated by 1 ms delays is applied to eliminate proton magnetization, ensuring that all experiments start with the same initial state. After a 5 s delay, cross-polarization is used to transfer polarization from proton spins to phosphorus spins. The duration of the spin-locking fields is 1400 μs	167
C.3	Measurement pulse sequence.	167
C.4	Pulse sequences 1D (top), 2D (middle) and 3D (bottom) experiments.	168
D.1	Pulse sequence for multi-quantum experiments. The phase of the magic echo sequence in the middle is changed so that it applies a θ encoding to the environment spins.	173
D.2	Pulse sequence for Loschmidt echo experiments. SME8 [†] is a variations of SME8 where delays in the magic echo sequences are adjusted to introduce mixing.	173
D.3	Pulse sequence for multi-quantum experiments with mixing. $\Delta_+ = 2\Delta_-$ so the environmental dynamics is refocused at the end of Δ_-	174

List of Tables

2.1	Overview of applications of AHT with respective targets, goals, and figures of merit (FoM).	15
2.2	Achievable ranges of scaling factors for different target effective Hamiltonians. They correspond to conditioned decoupling, double/zero quantum dynamics, and double/single quantum dynamics.	39
3.1	Ranks of $\oplus_{G_r} \mathbf{A}(G_r)$ for decoupling the secular dipolar interaction for different network structures (Fig. 3.2).	51
4.1	Posterior mean estimations of the commutation functions (defined in the caption of Fig. 4.9). 95% intervals are used to obtain the error bars.	67
5.1	Different time scales measured by the 1D and 2D experiments.	95
A.1	Conditions for achieving different objectives.	134
A.2	Achievable ranges of scaling factors for different target effective Hamiltonians. Here, for simplicity, only unnormalized target Hamiltonians are shown. The scaling factors are from normalized Hamiltonians to normalized ones. The last target Hamiltonian is an example of arbitrary desired two-body Hamiltonians.	146
A.3	Achievable ranges of scaling factors for different target effective Hamiltonians for a 3-spin spin chain when the interaction is included in the primary Hamiltonian.	146

Chapter 1

Introduction

Quantum technologies rely on the ability to manipulate quantum systems with high precision and reliability. Coherent control (the process of steering quantum dynamics using tailored external fields) is essential for enabling quantum computing, sensing [50], simulation [61, 192], and communication [11]. However, the quality of quantum control is fundamentally limited by noise, system variability, and imperfections in control hardware. These issues become increasingly problematic as systems scale in size and complexity. A major open challenge is to develop control methods that remain effective under such non-ideal conditions and across different physical platforms. This thesis is concerned with engineering coherent control, in particular addressing robustness and precision. We use Average Hamiltonian Theory (AHT) [82, 81] to achieve efficient, precise, and robust quantum control on qubit systems. This is accomplished by characterizing the set of achievable effective Hamiltonians at each order in the Magnus expansion and developing general strategies to engineer these terms toward desired targets. Efficiency is achieved by reaching the target control at the lowest order possible. Robustness is achieved through numerical optimization that minimizes the sensitivity to variations in the system. The methods developed are applicable to all modalities of qubits and to all applications of quantum technologies.

When a quantum system interacts with an external environment, coherent evolution is disrupted by information exchange between the system and its surroundings. The environment can store and return information to the system, leading to non-Markovian effects that require more complex models of quantum dynamics, in particular beyond models based on Lindblad dynamics [121, 71]. Such decoherence can reduce the effectiveness of control sequences by introducing persistent system/environment correlations and limiting the refocusability of quantum dynamics. To explore this, we develop experimental tools based on a central spin (CS) model [49, 5] to characterize these correlations and environ-

mental memory effects. We use stimulated echo [83, 188] methods to measure the time correlation of the noise along with multi-dimensional correlation spectroscopy [58] and multi-quantum experiments [139] to characterize the growth of system/environment correlations and the mixing induced by the environmental dynamics. The insights gained from these experiments lead to more accurate models of decoherence and enable the design of control strategies that are more robust to environmental influence.

1.1 Robust quantum control via Hamiltonian engineering

The first part of the thesis provides a method for designing precise, robust quantum control strategies using Hamiltonian engineering. Here, we illustrate the impact of these control methods by considering a simple case where the total Hamiltonian of a quantum system is divided into a control Hamiltonian $H_c(t)$ and an internal Hamiltonian H_{int} :

$$H_{\text{tot}}(t) = H_c(t) + H_{\text{int}}. \quad (1.1)$$

Quantum control aims to design the time dependence of $H_c(t)$ to achieve a target unitary process over a specified interval of time. The full, general implementation of quantum control is contained in Chapters 2 and 3.

Effective Hamiltonian engineering, based on Average Hamiltonian Theory (AHT), can be used to find $H_c(t)$ sequences that achieve the target unitary evolution of the system through an effective Hamiltonian H_{eff} . By using the Magnus expansion, AHT enables the designing of scalable, robust sequences perturbatively. Specifically, AHT uses the fact that the total unitary of a sequence can be equivalently represented by an effective Hamiltonian H_{eff} and the duration of the sequence T_{seq} :

$$U(T_{\text{seq}}) = \mathcal{T} \exp \left[-i \int_0^{T_{\text{seq}}} H_{\text{tot}}(t) dt \right] = \exp(-i H_{\text{eff}} T_{\text{seq}}), \quad (1.2)$$

where $\mathcal{T}$ is the Dyson time-ordering operator. The Magnus expansion of H_{eff} is a perturbative expansion in $|H_{\text{tot}}| \tau_c$ [81] where τ_c is the correlation time of $H_{\text{tot}}(t)$:

$$H_{\text{eff}} = \bar{H}^{(0)} + \bar{H}^{(1)} + \dots, \quad (1.3)$$

where each $\bar{H}^{(n)}$ is the n th-order average Hamiltonian. For example, the zeroth- and first-order average Hamiltonians are

$$\begin{aligned}\bar{H}^{(0)} &= \frac{1}{T_{\text{seq}}} \int_0^{T_{\text{seq}}} dt_1 H_{\text{tot}}(t_1), \\ \bar{H}^{(1)} &= -\frac{i}{2T_{\text{seq}}} \int_0^{T_{\text{seq}}} dt_1 \int_0^{t_1} dt_2 [H_{\text{tot}}(t_1), H_{\text{tot}}(t_2)].\end{aligned}\tag{1.4}$$

Since each $\bar{H}^{(n)}$ is Hermitian, truncation of the expansion at any order is also Hermitian and gives a valid approximation of H_{eff} . Typically, when the expansion converges, higher-order terms are increasingly smaller quantities.

The toggling frame [129] (defined as the interaction frame of $H_c(t)$) provides a convenient picture to separate the control from the evolution and to study how the internal dynamics is being modulated. Instead of $H_{\text{tot}}(t)$, the Magnus expansion can be defined with respect to the toggling-frame Hamiltonian:

$$H_{\text{tog}}(t) = U_c^\dagger(t) H_{\text{int}} U_c(t),\tag{1.5}$$

where $U_c(t) = \mathcal{T} \exp[-i \int_0^t H_c(t') dt']$. For example, in the WAHUHA-4 sequence [206] (Fig. 1.1), the toggling-frame Hamiltonian of the high-field dipolar interaction $H_{\text{int}} = H_{zz} = 3\sigma_z\sigma_z - \sigma \cdot \sigma$ during each window is shown in Fig. 1.1. Both $\bar{H}^{(0)}$ and $\bar{H}^{(1)}$ are zero when the pulse widths are negligible. In the presence of finite pulse widths, the WAHUHA-4 sequence has a residual error at zeroth-order [81]. This is accounted for in the MREV-8 sequence [127, 163] which combines two WAHUHA-4 sequences with different phases.

To build a general framework for quantum control using AHT, it is useful to consider the following questions:

1. What are the achievable effective Hamiltonians under the Magnus expansion, given the available control Hamiltonian H_c and the internal Hamiltonian H_{int} ? This question defines the controllability of the system, that is, whether the available controls are sufficient to generate a desired target unitary evolution.
2. What is achievable at zeroth order, and can an additional control term render a target Hamiltonian achievable at zeroth order? Achieving the target at the lowest possible order is typically most efficient.
3. How to design sequences that are robust to variations in H_{int} and the control system? Unintended variations in parameters are unavoidable, but control solutions can be found that are largely insensitive to such variations, particularly if they have correlation times longer than Rabi period (correlation time of variations $> \frac{2\pi}{\omega_{\text{Rabi}}}$).

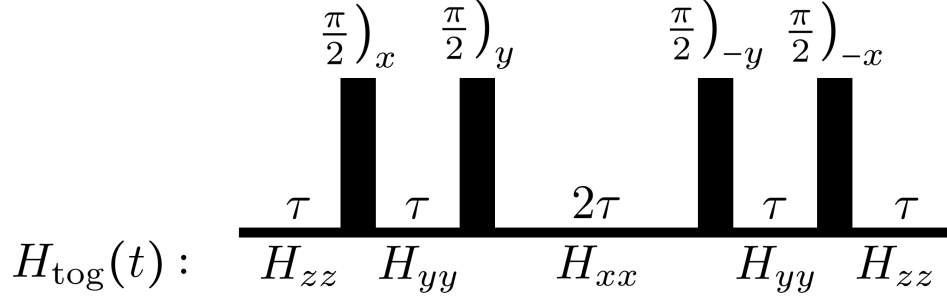


Figure 1.1: The WAHUA-4 sequence and the toggling-frame Hamiltonian of H_{zz} . Here, $H_{\alpha\alpha} = 3\sigma_\alpha\sigma_\alpha - \sigma \cdot \sigma$ for $\alpha = x, y$ and z . The total length of the sequence is $T_{\text{seq}} = 6\tau$. Since $H_{xx} + H_{yy} + H_{zz} = 0$, the sequence decouples H_{zz} to zeroth order when the pulse widths are negligible. $\bar{H}^{(1)}$ is also zero since the sequence is symmetric ($H_{\text{tog}}(t) = H_{\text{tog}}(T_{\text{seq}} - t)$).

4. How to minimize higher-order corrections and cross terms in the Magnus expansion? These terms involve nested commutators of the Hamiltonians and typically scale with T_{seq} . Errors due to these terms can become significant after accumulation. So, minimizing higher-order terms and cross terms is important for achieving high precision in engineered effective Hamiltonian.
5. How to engineer nonzero higher-order terms in the Magnus expansion? This is useful since not all effective Hamiltonians are achievable at zeroth order (e.g., Hamiltonians involving multi-qubit correlations absent in H_c or H_{int} , such as a higher-order grade-raising Hamiltonian $H = \sigma_+\sigma_+\sigma_+ + \sigma_-\sigma_-\sigma_-$).

Studying controllability is essential for verifying the feasibility of the target control given the available resources (e.g., tunable parameters and power bandwidth product) and determining modification of the system (e.g., inclusion of additional controls) when the target is unreachable. This enables efficient allocation of resources including power, bandwidth and duration of the sequence. Additionally, the set of achievable effective Hamiltonians is also the set of possible errors. Understanding controllability helps with error analysis and can be used to design more robust sequences.

Robustness is needed for most implementations. A universal approach, regardless of the modality of the control system and the properties of H_{int} , is useful for making the method more applicable. It also allows knowledge gained with one modality to be transferred to others.

Minimizing higher-order corrections is necessary for achieving precise effective Hamiltonians, even when terms in the Magnus expansion cannot be calculated directly, such

as in large systems or when the Hamiltonian is uncertain. Symmetry-based methods for minimizing higher-order and cross terms remain useful.

We aim to identify symmetry-based conditions that minimize each term in the Magnus expansion, capturing the minimal requirements necessary for suppression. These conditions should reduce to exact Magnus expansion evaluation in the case of perfect knowledge of the Hamiltonians. For instance, the first-order term $\bar{H}^{(1)}$ can be eliminated by symmetrizing the control sequence; however, while such symmetry is sufficient, it is not strictly necessary and may not always be realizable in practice. Indeed, some sequences lacking explicit symmetry [43] still achieve $\bar{H}^{(1)} = 0$. Moreover, we seek conditions that are numerically computable, enabling integration into optimization routines for control sequence design.

Since not all effective Hamiltonians can be realized at zeroth order, a systematic approach for engineering nonzero higher-order terms is used to expand the capabilities of AHT. In particular, higher-order terms are preferred when the target unitary includes more qubits than those in the control or internal Hamiltonians or when the target involves multi-qubit correlations absent in the control and internal Hamiltonians. Although the control and internal Hamiltonians are typically limited to one- and two-body terms, shaped pulses can be used to deliberately amplify higher-order contributions, enabling the creation of effective multi-body interactions and control over multi-spin pathways [203]. For example, in dipole-coupled spin systems subject to collective spin rotations, it is possible to engineer double-quantum Hamiltonians (e.g., the grade-raising Hamiltonian [9], $H = \sigma_+ \sigma_+ + \sigma_- \sigma_-$) at zeroth order. At first order, the effective Hamiltonian includes commutators of toggling-frame components, which can give rise to triple-quantum terms. Additionally, establishing conditions for constructing nonzero symmetry-based solutions are important for designing scalable protocols, such as spin amplifiers [36], that remain invariant to uninteresting parameters in the Hamiltonian, such as variations in coupling strengths across a qubit ensemble.

The time dependence of $H_c(t)$ is controlled via a set of control parameters $\{a_k(t)\}$ (e.g., the time-dependent amplitudes and phases of the control field). $\{a_k(t)\}$ is mapped to a continuous output field $\vec{B}(t)$ acting on the qubit system. The map depends on a set of model parameters $\{\mu_j\}$ (e.g., impedance and nonlinearities). In practice, $\{\mu_j\}$ are subject to variations, including environmental fluctuations. H_{int} also depends on a set of internal parameters $\{\eta_{\text{int}}^i\}$ (e.g., coupling strengths and offsets) that can additionally have variations. Robust quantum control requires the designed total unitary evolution under H_{tot} to be insensitive to these variations. For example, Fig. 1.2 compares the control landscapes (fidelity maps over variations), for a simple square $\pi/2$ rotation, a composite

$\pi/2$ rotation designed for robustness [212],

$$\text{BB1} = R_0(\pi/2)R_{\phi_0}(\pi)R_{3\phi_0}(2\pi)R_{\phi_0}(\pi), \quad (1.6)$$

where $R_\phi(\theta) = \exp[-i(\sigma_x \cos(\phi) + \sigma_y \sin(\phi))\theta/2]$ and $\phi_0 = \arccos(-1/8)$, and an optimized robust pulse. However, solutions such as the composite pulses are typically designed for specific kinds of control imperfections (e.g., Rabi field and offset variations). It is useful to have a general method for achieving robustness against arbitrary parameter variations, independent of the physical implementation. Using numerical optimization (e.g., gradient descent [97]), we can design robust pulses that outperform conventional composite pulses (Figs. 1.2 and 1.3).

To address these questions, we develop a new framework that is suitable for:

1. Fully characterizing the space/set of achievable unitary propagators and effective Hamiltonians in each order in the Magnus expansion (Chapters 2 and 3). This reveals exactly what control targets are achievable and determines the resulting set of errors. This establishes a foundation for Hamiltonian engineering using AHT.
2. Achieving robustness against most systematic uncertainties in the control system (Chapter 2). This enables the design of robust control sequences, including higher-order robustness, independent of the system's physical modality.
3. Combining numerical integration of the Schrödinger equation and perturbative expansions using the Magnus expansion (Chapter 2). By generalizing the toggling frame to the interaction frame of any component of H_{tot} , we can reshape the set of achievable effective Hamiltonians and apply numerical integration or perturbative evaluation selectively. This allows precise control over how the solution depends on specific parameters, such as the control functions $\{a_k(t)\}$.
4. Designing solutions that are parameter-dependent and/or symmetry-based, depending on the knowledge of the Hamiltonians (Chapter 3). First, this generalizes existing well-known symmetry conditions in AHT [33] to enable the identification of necessary and sufficient conditions for achieving any target effective Hamiltonian, zero or nonzero, at each order within the achievable set. Second, it connects with numerical methods for directly evaluating the Magnus expansion (e.g., using the van Loan equations [79]), enabling the design of solutions that interpolate between parameter-dependent and symmetry-based approaches, depending on the available knowledge of the Hamiltonian and specific design requirements.

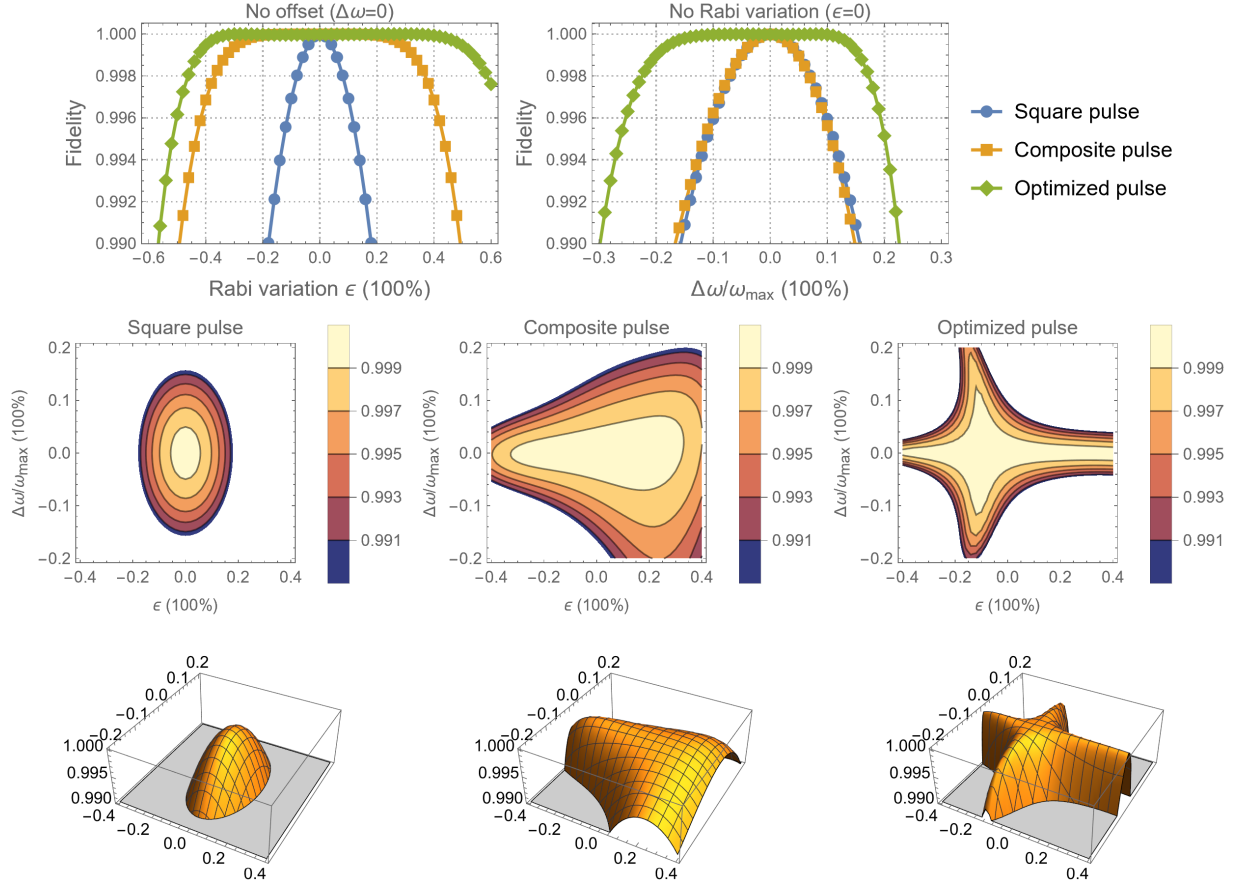


Figure 1.2: Control landscapes of a $\pi/2$ x rotation implemented as a simple square $\pi/2$ rotation, a composite $\pi/2$ rotation (BB1) and an optimized robust pulse. The composite rotation maintains a high fidelity over a larger range of Rabi field variations (ϵ) compared with the square pulse. Using numerical optimization, robust pulses can be found that outperform the composite pulse and also incorporate robustness against other errors. Here, the optimized pulse is also robust against variations in energy offset ($\Delta\omega$) of the qubit.

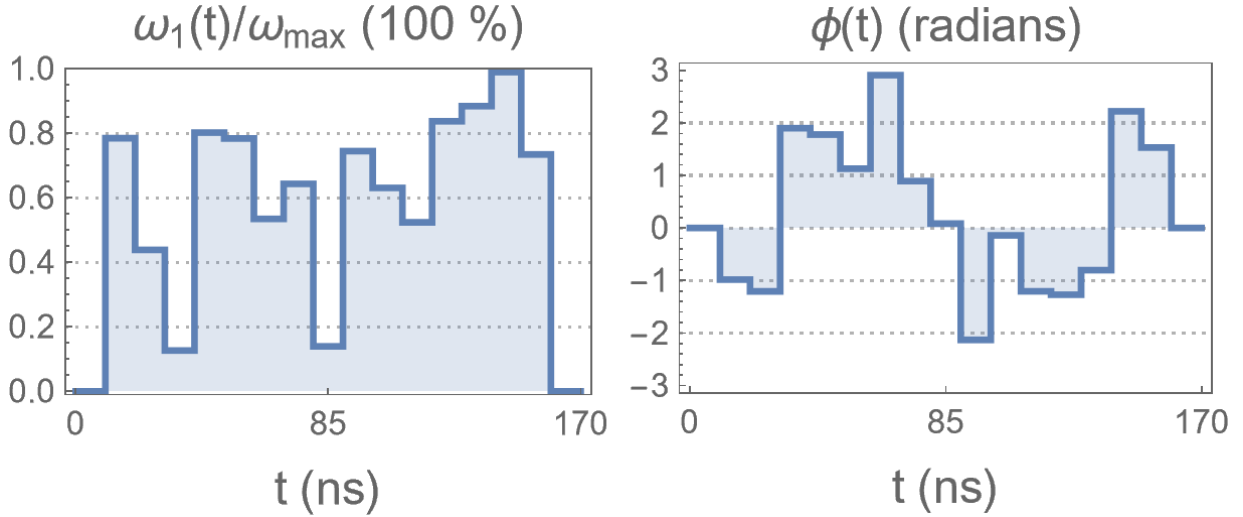


Figure 1.3: Time-dependent amplitude $\omega_1(t)$ and phase $\phi(t)$ of the robust optimized pulse. Wait periods are included at the beginning and the end to allow easy composition between different pulses. In a more complicated implementation, the pulse would be designed for the control system’s bandwidth and nonlinearities.

5. Minimizing higher-order corrections, including cross terms between different components of H_{int} and those between the control error and H_{int} (Chapters 2 and 3). Higher-order robustness in the presence of control imperfections is achieved by minimizing both the higher-order terms and the cross terms between control errors and the internal Hamiltonian. Minimizing cross terms between different components of H_{int} enables higher precision in engineering the target effective Hamiltonian.
6. Providing necessary and sufficient conditions for engineering target average Hamiltonians in each order in the Magnus expansion, easily used by numerical optimizers for sequence design (Chapters 2 and 3). This enables the design of optimized control sequences that accommodate multiple objectives simultaneously, including achieving a target unitary and/or effective Hamiltonian, as well as robustness.

Control efficiency can be achieved by designing nonzero H_{eff} in the lowest-order term in the Magnus expansion whenever possible and minimizing other terms. The proposed method is independent of the modality of the system. The applications span across gate design, spectroscopy, dynamical decoupling, quantum simulation, and sensing. The effectiveness of the method is demonstrated through diverse examples, including robust control under Rabi field and offset inhomogeneities, handling imperfections in both linear and nonlinear

control systems, simulation of a tunable quantum harmonic oscillator, realization of a scalable spin amplifier for sensing, higher-order decoupling of dipole-coupled systems, and the engineering of symmetry-based, scalable three-body interactions. The quality of the designed sequences is evaluated through calculating their control landscape and figure of merit (average gate fidelity over distributions of $\{\mu_j\}$ and $\{\eta_{\text{int}}^i\}$) including a simple model of decoherence of the system (e.g., a depolarizing channel).

1.2 Characterizing noise in open quantum systems using a central spin model

Designing control sequences becomes more complex when the quantum system interacts strongly with an external environment. In these regimes, decoherence is typically non-Markovian, and the environment can retain long-lived memory of the system’s evolution. As a result, standard figures of merit derived from Markovian noise models generally fail to accurately predict the performance of control sequences since it can depend on the system’s prior evolution history. The evolution of the quantum state may have complex, non-exponential structure depending on the memory effect of the environment (Fig. 1.4). In Fig. 1.4, the average gate fidelity of a simple one-qubit square $\pi/2$ rotation is plotted as a function of the pulse duration t . A simple depolarization channel is used to simulate the Markovian noise. The non-Markovian noise is simulated by coupling the qubit to a four-qubit environment with internal dynamics and analyzing the evolution of its reduced density matrix. Under Markovian noise, the fidelity decays exponentially, reflecting the memoryless nature of the environment. In contrast, under non-Markovian noise, the fidelity exhibits non-exponential behavior, including possible revivals and saturation, due to memory effects in the environment. Fig. 1.4 also shows the concurrence of a Bell state, $(|00\rangle + |11\rangle)/\sqrt{2}$, under both types of noise. Similar to the fidelity behavior, non-Markovian noise leads to a non-monotonic decay of entanglement, often featuring revivals and delayed loss compared to the purely exponential decay observed under Markovian noise. The insets of Fig. 1.4 show the Zeno regions [135, 59] under the non-Markovian noise where the environment holds short-term memory and feeds back information. This can be useful for designing quantum control since the system can be stabilized or revived using tailored pulse sequences that exploit memory effects which enable more efficient or robust control than is possible under Markovian noise. Developing experimental tools to characterize system/environment dynamics, particularly to probe system/environment correlations, environmental mixing, information backflow (e.g., revival of correlations), and long-lived memory (e.g., correlations insensitive to mixing), is essential for constructing

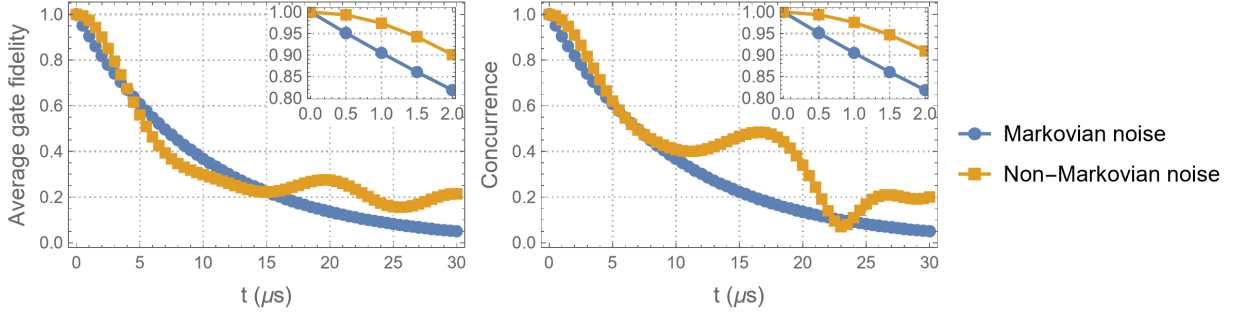


Figure 1.4: Dynamics of quantum control and evolution under Markovian and non-Markovian noise. Here, the average gate fidelity of a square $\pi/2$ pulse and the concurrence of a Bell state are shown. Insets provide a close-up of the initial decay behavior.

more accurate models of decoherence. Such improved models can, in turn, enable the design of more effective quantum control protocols and noise mitigation strategies, including quantum error correction [12, 100].

In the second part of the thesis, methods are described for obtaining a more complete understanding of the environmental dynamics. The goals of characterizing the environment are to:

1. Quantify the system/environment correlations, i.e., classical and quantum correlations between the states of the system and the environment. This is important for designing effective control for non-Markovian noise because these correlations determine how past interactions influence future dynamics.
2. Understand how the local field changes under environmental mixing, i.e., environmental evolution that scrambles its state, which may reduce system/environment correlations. The evolution of the quantum system is driven by the local field. Understanding its changes under mixing is essential to predict, control, or suppress decoherence and unwanted interactions.
3. Quantify non-stochastic processes. Non-stochastic processes including environmental mixing can produce structured, memory-dependent dynamics that are not captured by standard noise models like the Gauss-Markov model, so quantifying them is crucial for accurate modeling and control of realistic open quantum systems.
4. Understand how environmental mixing moves information within the environment, including from a directly connected part to the extended environment. This helps di-

agnose when and where information becomes inaccessible, irreversibly lost, or scrambled in a way that can only be recovered with additional controls (e.g., collective control over the environment), setting fundamental limits on control fidelity and predictability.

Answering the above questions enables a more accurate decoherence model, which will inform improved control design in future work.

A central spin (CS) model (Fig. 1.5) is used to characterize the system/environment dynamics. The Hamiltonian of the CS system is

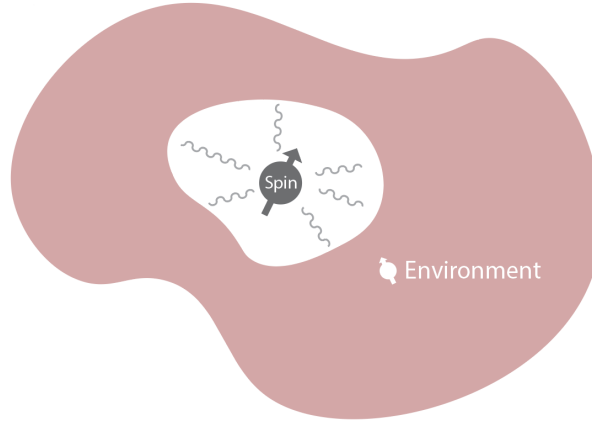


Figure 1.5: Schematic of a CS model. Here, a spin-1/2 CS is coupled to an external environment of many, interacting spin-1/2 particles.

$$H = H_{cs} + H_{cs/e} + H_e, \quad (1.7)$$

where H_{cs} , $H_{cs/e}$, and H_e are the CS, CS/environment interaction and the environment Hamiltonians. The CS system provides a quantum model for a qubit coupled to an environment that can hold quantum information. $H_{cs/e}$ is responsible for the growth of system/environment correlations and the flow of quantum information from the CS to the environment. Evolution under H_e can scramble the CS/environment correlations and potentially move the information to an extended environment that is not directly connected to the CS.

To study the system/environment correlations and the memory effect of the environment, we use three different experimental approaches and use a solid-state NMR system to experimentally test these methods. Stimulated echo (STE) experiments (Chapter 4) are

similar to forced Rayleigh scattering experiments, with encoding, mixing, and decoding processes analogous to grating diffusion and scattering. The STE experiments provide an efficient way to quantify the growth of system/environment correlations and the extent of mixing. The multi-dimensional (MD) correlation spectroscopy (Chapter 5) provides a more complete picture by producing the multi-time joint probability distributions of the noise on the CS. This can be used to quantify non-Gaussianity, non-stationarity, and non-Markovianity of the dynamics of the environment. The multi-quantum (MQ) experiments (Chapter 6) measure the growth of CS/environment correlations via Hamming weights [143]. They enable studying how the environmental mixing moves the quantum information within the environment and are used to measure the size of the environment. The MQ experiments reveal a directly connected part of the environment and an extended part that is not directly connected to the CS. Bayesian analysis is used to draw quantitative conclusions.

Chapter 2

Engineering Precise and Robust Effective Hamiltonians

Engineering effective Hamiltonians is essential for advancing quantum technologies including quantum simulation, sensing, and computing. This chapter presents a general framework for effective Hamiltonian engineering, enabling robust, accurate, and efficient quantum control strategies. We focus on creating target zeroth-order effective Hamiltonians while minimizing higher-order contribution and enhancing robustness against systematic errors. The control design identifies the minimal subspace of the toggling frame Hamiltonian and the full set of achievable, zeroth-order, effective Hamiltonians. Examples are included to illustrate the achievable precision and robustness.

2.1 Introduction

Quantum processors have extraordinary potential to transform sensing, secure communication, and enhance computational efficiency, including in the simulation of quantum physics. To realize this potential, three key engineering tasks must be addressed: (1) efficient and robust coherent control, (2) the design of complex, interconnected quantum processors, and (3) the implementation of quantum error correction. This chapter provides tools for designing coherent control methods to enhance precision and efficiency, using Average Hamiltonian Theory (AHT) as the foundational framework. Introduced by John Waugh over fifty years ago [82], AHT describes coherent averaging initially developed for NMR [206]. It has since been widely applied to decoupling [127, 163, 34, 185, 43], recoupling

[160, 118], spectroscopy [164, 80], sensing [146], quantum simulation [192, 22], multi-body physics [207, 143, 153, 90], and quantum computing [179]. The formalism Waugh developed for the coherent control of coupled spin-1/2 particles naturally extends to the control of coupled qubits, as the control of qubits and spin-1/2 particles are isomorphic.

AHT is based on the principle that the integrated action of a time-varying total Hamiltonian, $H_{\text{tot}}(t)$, over a period $[0, T_{\text{seq}})$, can be represented by a unitary operator or, equivalently, an effective Hamiltonian, H_{eff} , and its duration, T_{seq} . The total Hamiltonian of a quantum system is typically partitioned into a user-defined, time-dependent control Hamiltonian and a time-independent internal Hamiltonian

$$H_{\text{tot}}(t) = H_c(t) + H_{\text{int}}. \quad (2.1)$$

The total unitary propagator is

$$\begin{aligned} U(T_{\text{seq}}) &= \mathcal{T} \exp \left(-i \int_0^{T_{\text{seq}}} H_{\text{tot}}(t) dt \right) \\ &= \exp \left(-i H_{\text{eff}} T_{\text{seq}} \right), \end{aligned} \quad (2.2)$$

where $\mathcal{T}$ is the Dyson time-ordering operator, and H_{eff} is the effective Hamiltonian. When designing an effective Hamiltonian perturbatively, the most efficient solution is often achieved at zeroth order, as higher-order terms typically make increasingly smaller contributions¹. Methods for accounting for higher-order corrections are essential for robust applications of AHT. Table A.1 lists common use cases of AHT, along with their goals and typical figures of merit.

To illustrate the desired outcome, consider a simple and time-independent internal Hamiltonian

$$H_{\text{int}} = \sum_{i \neq j}^N B_{ij} \vec{\sigma}_i \cdot \mathbf{D} \cdot \vec{\sigma}_j + \sum_{i=1}^N \delta_i \vec{n} \cdot \vec{\sigma}_i \quad (2.3)$$

where $\vec{\sigma}_i = (\sigma_x^i, \sigma_y^i, \sigma_z^i)$ and

$$\sigma_\alpha^i = \mathbb{1}^{\otimes(i-1)} \otimes \sigma_\alpha \otimes \mathbb{1}^{\otimes(N-i)}, \quad (2.4)$$

for $\alpha = \{x, y, z\}$, are the Pauli operators acting on the i th qubit, B_{ij} are coupling strengths, $\mathbf{D}$ is a rank-2 tensor determining the form of the qubit interactions, $\vec{n}$ is a unit vector determining the direction, δ_i the strength of the qubit-dependent energy offset, and N is the total number of qubits. Note that the algorithms outlined in the paper do not depend

¹Typically, when designing an effective Hamiltonian using the Magnus expansion, the parameter $|H_{\text{tot}}| \tau_c$ is small, so higher-order terms contribute progressively less [82]. A general convergence condition for the expansion can be found in [16].

Table 2.1: Overview of applications of AHT with respective targets, goals, and figures of merit (FoM).

Use case	Target	Goal	FoM
Decoupling	$H_{\text{eff}} = s \sum_i \delta_i \vec{v}_{\text{eff}} \cdot \sigma_i$	spectroscopy	line narrowing $(\frac{1}{\pi T_2}), sT_{2x}$
Time suspension	$H_{\text{eff}} = 0$	store quantum information	decoherence $(\frac{1}{\pi T_2})$
Simulation	$H_{\text{eff}} = H_{\text{sim}},$ $U(nT_{\text{seq}}) = e^{-iH_{\text{sim}}nT_{\text{seq}}}$	simulating quantum systems	$\ H_{\text{eff}} - H_{\text{sim}}\ $
Sensing	$H_{\text{eff}} = H_{\text{sen}}$	parameter estimation	Fisher information
Gates	$U(T_{\text{seq}}) = U_{\text{target}}$	computation	gate fidelity

on the particular organization of the Hamiltonians. We will show later that more complex internal Hamiltonians are easily accommodated.

In Table A.1, the goal of spectroscopy is to measure select contributions to the internal Hamiltonian while suppressing unwanted terms. A common goal is to measure the linear terms while suppressing the bilinear terms in H_{int} [82]. The scaling factor s is preferably large and the effective direction $\vec{v}_{\text{eff}}$ is typically chosen to be along the quantization axis. Time suspension [43, 196, 219] is an extension of decoupling that aims to suppress all of the terms in H_{int} so, $H_{\text{eff}} = 0$. It is enabling of qubit storage. H_{sim} and H_{sen} are the target effective Hamiltonians for simulation and sensing tasks. In quantum simulation, the goal is to transform H_{int} into H_{sim} , typically to explore the dynamics of the simulated system. Quantum simulation is often performed stroboscopically with $H_{\text{sim}}(T_{\text{seq}})$ being the effective Hamiltonian over a repeated cycle of length T_{seq} and n being the index of the stroboscopic measurement. In sensing, the goal is to engineer a Hamiltonian, H_{sen} , that offers advantages, such as enhanced Fisher information, for measuring target parameters such as the strength of an external field. In gate design, the goal is to implement a target unitary U_{target} efficiently and with robustness against errors. The figure of merit (FoM) is the gate fidelity.

Recent advances have enhanced the application of AHT. For example, J. Choi et al. [40] proposed a symmetry-based method for engineering effective Hamiltonians using control sequences composed of $\pi/2$ pulses and delays, specifically targeting the decoupling of the zeroth-order average Hamiltonian. It enables the design of high-performance sequences for

spectroscopy and quantum simulations, demonstrating superior performance compared to existing methods in the presence of strong disorders. H. Zhou et al. [221] and M. Tyler et al. [193] extended this approach to account for higher-order terms. However, these methods remain limited to control with $\pi/2$ rotations of the qubits, anisotropic Heisenberg model, and lower-order corrections.

Reinforcement learning has been utilized to engineer effective Hamiltonians using $\pi/2$ pulses to decouple secular dipolar interactions [152]. This method highlights the potential of advanced numerical algorithms to significantly enhance the design of sequences for Hamiltonian engineering [62].

The above methods mainly target time suspension applications. For gate design, H. Haas *et al.* [79] utilized the van Loan equation [69] to enable numerical evaluation of the Magnus expansion for control sequences. The authors further integrated their framework with GRAPE [97], facilitating efficient numerical optimization for designing sequences in Hamiltonian engineering. Other numerical methods have also been proposed [158].

Several key challenges remain unresolved:

- **Controllability:** A systematic method to find the achievable set of effective Hamiltonians for arbitrary systems and controls [47] is not well-known.
- **Arbitrary control errors:** There is no general methodology for designing robust sequences that mitigate arbitrary systematic errors in control systems.
- **Higher-order decoupling:** A generic symmetry-based framework for eliminating higher-order correction terms is still lacking.

This paper aims to provide a general framework suitable for designing control sequences in the presence of control imperfections and dispersion of H_{int} . The method includes:

1. Algorithms for characterizing achievable sets of zeroth-order average Hamiltonians for any H_{int} and H_c .
2. Algorithms for finding numerically optimized sequences for engineering zero and nonzero effective Hamiltonians.
3. Method to reduce the effects of most systematic errors in the control system.
4. Method to reduce higher-order corrections from the Magnus expansion.

2.2 Effective Hamiltonian engineering and toggling frame

For robust control, the internal and control Hamiltonians are broken up into isochromats that describe the range of possible parameters. For each isochromat, the control is unitary so an effective Hamiltonian (Eq. 2.2) always exists. For the dynamics averaged over a distribution of isochromats, the overall map is a convex sum of unitary processes and so is a completely positive trace-preserving (CPTP) map. The effective Hamiltonian (H_{eff}) is formally calculated by taking the matrix logarithm of $U(T_{\text{seq}})$. H_{eff} describes the system's effective dynamics over the interval $[0, T_{\text{seq}})$ and is not unique since adding integer multiples of $2\pi/T_{\text{seq}}$ to its eigenvalues yields the same unitary $U(T_{\text{seq}})$. In place of matrix logarithm, AHT often uses a perturbative series approximation to evaluate H_{eff} via the Magnus expansion:

$$H_{\text{eff}}(T_{\text{seq}}) = \sum_{n=0}^{\infty} \bar{H}^{(n)}(T_{\text{seq}}), \quad (2.5)$$

where n is the perturbation order, and $\bar{H}^{(n)}$ is the n th-order average Hamiltonian. The zeroth-order average Hamiltonian is

$$\bar{H}^{(0)}(T_{\text{seq}}) = \frac{1}{T_{\text{seq}}} \int_0^{T_{\text{seq}}} H_{\text{tot}}(t) dt. \quad (2.6)$$

When the total Hamiltonian commutes with itself at all times, $H_{\text{eff}}(T_{\text{seq}}) = \bar{H}^{(0)}(T_{\text{seq}})$. We are interested in the case where we focus on finding a control sequence such that

$$H_{\text{target}} = \bar{H}^{(0)}(T_{\text{seq}}). \quad (2.7)$$

Having a zeroth-order average Hamiltonian implementation of H_{target} is advantageous since when a solution is achievable in the zeroth order, it is typically the most efficient implementation in terms of resources such as the power bandwidth product. If H_{target} is achievable in zeroth order, a good strategy is to find a control sequence achieving Eq. (2.7) and to simultaneously minimize the residual errors in the higher-order terms in the Magnus expansion:

$$\bar{H}^{(1)}(T_{\text{seq}}), \bar{H}^{(2)}(T_{\text{seq}}), \dots \quad (2.8)$$

Of course, not all potential H_{target} are achievable in zeroth order.

The toggling frame is a powerful tool that simplifies finding average Hamiltonians. It bridges numerical integration and perturbative expansion by partitioning $H_{\text{tot}}(t)$ into a primary part and a perturbative part:

$$H_{\text{tot}}(t) = H_{\text{pri}}(t) + H_{\text{pert}}(t). \quad (2.9)$$

The toggling frame is the interaction frame of $H_{\text{pri}}(t)$. A common use of AHT is to take the control Hamiltonian H_c as H_{pri} and the internal Hamiltonian H_{int} as H_{pert} . This restriction, however, is unnecessary, and the freedom to choose any partition of H_{tot} enables a larger set of achievable controls in zeroth order. The toggling frame Hamiltonian of $H_{\text{pert}}(t)$ is

$$H_{\text{tog}}(t) = U_{\text{pri}}^\dagger(t) H_{\text{pert}}(t) U_{\text{pri}}(t), \quad (2.10)$$

where the primary unitary $U_{\text{pri}}(t)$ is

$$U_{\text{pri}}(t) = \mathcal{T} \exp \left(-i \int_0^t H_{\text{pri}}(t') dt' \right). \quad (2.11)$$

The perturbative unitary is

$$\begin{aligned} U_{\text{pert}}(T_{\text{seq}}) &= \mathcal{T} \exp \left(-i \int_0^{T_{\text{seq}}} H_{\text{tog}}(t) dt \right) \\ &= \exp(-i H_{\text{eff}} T_{\text{seq}}). \end{aligned} \quad (2.12)$$

Then, the total propagator is

$$U(T_{\text{seq}}) = U_{\text{pri}}(T_{\text{seq}}) U_{\text{pert}}(T_{\text{seq}}), \quad (2.13)$$

where U_{pri} and U_{pert} can be evaluated through numerical integration and perturbative expansion, respectively.

2.3 Control Setup

The control Hamiltonian is most often implemented as a sequence of piecewise constant intervals of control parameters $\{a_k(t)\}$. For example, it can be

$$\begin{aligned} H_c(t) &= \sum_{i=1}^N \left[\omega_1^i(t) \cos(\phi_i(t)) \sigma_x^i + \omega_1^i(t) \sin(\phi_i(t)) \sigma_y^i + (\omega_0 - \omega_r^i(t)) \sigma_z^i \right], \end{aligned} \quad (2.14)$$

where the time-dependent amplitude ($\omega_1^i(t)$), frequency ($\omega_r^i(t)$) and phase of the control field ($\phi_i(t)$) on the i th qubit are control parameters ($a_k(t)$). $a_k(t)$ is typically set to zero at the beginning and the end of each sequence [62], enabling a simple composition of gates.

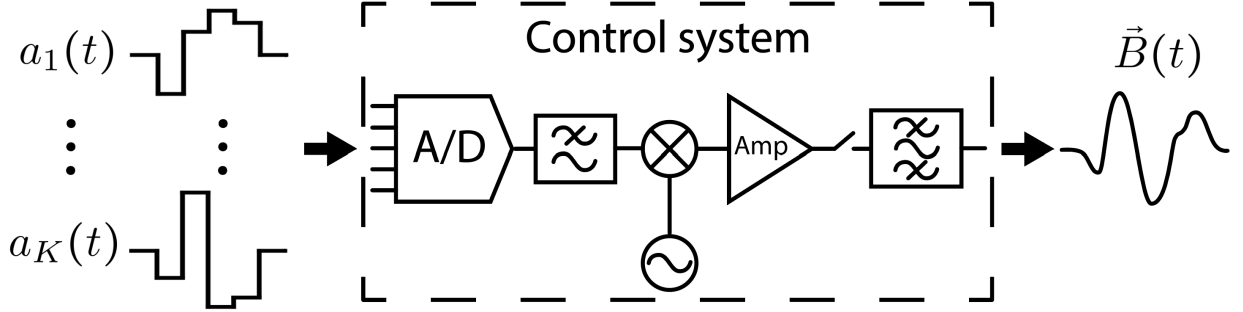


Figure 2.1: Illustration of a very simplified control system. The input signals $\{a_k(t)\}$ are designed by the user. The control system outputs a continuous control field $\vec{B}(t)$ acting on the qubit system in the lab frame. The design of $\{a_k(t)\}$ should account for the model of the system so that it captures the distortion of the hardware along with additive noise.

The control system (Fig. 2.1) maps $\{a_k(t)\}$ to a continuous output field $\vec{B}(t)$ that is conveniently described in the lab frame. This map is generally a non-invertible, smooth, nonlinear function depending on time-independent parameters of the control system, $\{\mu_j\}$, such as reactance, spatial and temporal variations of control fields, nonlinearities from kinetic inductance, etc. Variations in $\{\mu_j\}$ result in a modification of the control Hamiltonian by an error Hamiltonian ΔH . The error Hamiltonian ΔH can be evaluated by applying the Taylor expansion to the map from $\{a_k(t)\}$ to $\vec{B}(t)$ with respect to $\{\mu_j\}$. Robustness to variations in $\{\mu_j\}$ can then be achieved by minimizing the effective Hamiltonian of ΔH . The following methods are capable of finding high-precision control fields that account for $\{\mu_j\}$ and are robust to their variations.

The internal Hamiltonian also depends on a set of parameters, $\{\eta_{\text{int}}^i\}$, which can include terms such as the Zeeman splitting, chemical shift, coupling strengths, etc. Typically, $\{\eta_{\text{int}}^i\}$ are governed by a classical probability distribution with a correlation time that is often longer than the total length of a control cycle (T_{seq}).

Given a sequence $\{a_k(t)\}$, as well as the internal and model parameters ($\{\eta_{\text{int}}^i\}$ and $\{\mu_j\}$), the density matrix of the system obeys the master equation

$$\frac{d}{dt}\rho(t) = -i[H_{\text{tot}}(\{a_k(t)\}; \{\mu_j\}, \{\eta_{\text{int}}^i\}), \rho(t)] + \mathcal{L}_{\text{relaxation}}[\rho(t)] \quad (2.15)$$

where $\mathcal{L}_{\text{relaxation}}$ is the Lindbladian for relaxation. The solution to Eq. (2.15) over the interval $[0, T_{\text{seq}}]$ is

$$\rho(T_{\text{seq}}) = \Lambda(\{a_k(t)\}; \{\mu_j\}, \{\eta_{\text{int}}^i\})(\rho_0), \quad (2.16)$$

where ρ_0 is the initial density matrix and $\Lambda(\{a_k(t)\}; \{\mu_j\}, \{\eta_{\text{int}}^i\})$ is a CPTP map. The density matrix of the system for an expectation-valued measurement is the ensemble average of Eq. (2.16) over the distributions of $\{\mu_j\}$ and $\{\eta_{\text{int}}^i\}$:

$$\bar{\rho}(T_{\text{seq}}) = \bar{\Lambda}(\rho_0), \quad (2.17)$$

where the average CPTP map $\bar{\Lambda}$ is

$$\bar{\Lambda} = \int d\{\mu_j\} d\{\eta_{\text{int}}^i\} p(\{\mu_j\}) p(\{\eta_{\text{int}}^i\}) \Lambda(\{a_k(t)\}; \{\mu_j\}, \{\eta_{\text{int}}^i\}). \quad (2.18)$$

The goal is to design $\{a_k(t)\}$ to achieve the desired FoM with respect to $\bar{\Lambda}$ or $\bar{\rho}(T_{\text{seq}})$ for tasks in Table A.1. For example, if the overall unitary target for $U(T_{\text{seq}})$ is U_0 , then the average gate fidelity of the sequence can be calculated as

$$\mathcal{F}_{\text{ave}}(\{a_k(t)\}, U_0) = \int \langle \psi | U_0^\dagger \bar{\Lambda}(|\psi\rangle\langle\psi|) U_0 | \psi \rangle d\psi. \quad (2.19)$$

When knowing the details of the errors is important (e.g., in implementing quantum error correction, noiseless subsystems, and distinguishing between coherent and stochastic errors), one should report the full average CPTP map $\bar{\Lambda}$.

2.4 Flowchart for effective Hamiltonian engineering

The design method is given as a flowchart which references algorithms that are provided in pseudocode. The flowchart is the same for all applications in Table A.1.

1. Describe the quantum system and controls:
 - (a) Provide a description of H_{int} including distributions of $\{\eta_{\text{int}}^i\}$.
 - (b) Provide a description of $H_c(t)$ and a model of the control system including finite bandwidth, maximal power, and nonlinearities and distributions of the model parameters $\{\mu_j\}$. Specify the error Hamiltonian ΔH .
 - (c) Provide relaxation times when desired.
2. Identify from Table A.1 the use case and its objectives. Partition H_{tot} into H_{pri} and H_{pert} accordingly. Specify the target unitary U_{target} and/or effective Hamiltonian H_{target} and the required FoM.

3. Determine the controllability of $U(T_{\text{seq}})$:
 - (a) Controllability of U_{pri} : Use Algorithm 1 to determine if U_{target} is achievable.
 - (b) Space of achievable $\bar{H}^{(0)}$: Use Algorithm 2 to determine if H_{target} is in the space of achievable zeroth-order average Hamiltonians ($\bar{H}^{(0)}$).
 - (c) Range of achievable scaling factors for H_{target} : Use Algorithm 3 to check if sH_{target} is achievable. If yes, set $H_{\text{target}} \leftarrow sH_{\text{target}}$.
 - (d) If U_{pri} or sH_{target} are not achievable, go to Step 1 to change $H_c(t)$ if additional control is available in the physical system. Sometimes (see below for an explanation of **Step 3** (d)), one can go to Step 2 to choose a different partitioning to enable a larger set of achievable $\bar{H}^{(0)}$.
4. Using a numerical optimizer to find a control sequence that maximizes the FoM targets. The FoM targets are achieved by reaching the desired precision in $U_{\text{pri}}(T_{\text{seq}}) \leftarrow U_{\text{target}}$ and $\bar{H}^{(0)} \leftarrow H_{\text{target}}$ while accounting for the model of the control system, minimizing sensitivity to variations in the $\{\mu_j\}$ and $\{\eta_{\text{int}}^i\}$, and removing higher-order corrections $\bar{H}^{(1)}, \bar{H}^{(2)}, \dots$.
5. Calculate the control landscape, accounting for relaxation and the full model of the control system. Check if the desired FoM and robustness are reached. If yes, return the sequence. Otherwise, if the FoM is low due to relaxation, go to Step 3 (c) to choose a larger s . If the robustness is not sufficient, go to Step 4 to include higher-order robustness.

The following sections expand upon the flowchart step by step.

Step 1. Description of the quantum system and controls

(a) *Description of H_{int} .* The description of H_{int} includes the operators, amplitudes and distributions of $\{\eta_{\text{int}}^i\}$. For example, for H_{int} in Eq. (3.9), one needs to specify $\mathbf{D}$ and $\vec{n}$, distributions of the coupling strengths (B_{ij}) and the offsets (δ_i), usually obtained via spectroscopy.

(b) *Description of $H_c(t)$.* The description of $H_c(t)$ includes the input parameters $\{a_k(t)\}$ and a mapping, given by the model of the control system, that maps $\{a_k(t)\}$ to the output field $\vec{B}(t)$ in the lab frame. The map is generally given as first-order state-space differential equations. For time-invariant linear systems, the map is determined by the transfer function. Once a map from $\{a_k(t)\}$ to $\vec{B}(t)$ is given, $H_c(t)$ is obtained by

moving into the reference frame of the qubits. The distributions of the model parameters $\{\mu_j\}$ (obtained via “tune-up” cycles [67]) can be included. The error Hamiltonian ΔH is obtained by expanding the map from $\{a_k(t)\}$ to $\vec{B}(t)$ with respect to $\{\mu_j\}$.

(c) *Relaxation time.* Provide an estimate of relaxation times. Typically, T_1 and T_2 are included. When available, $T_{1\rho}$ can be included. T_2^* is best included in the distribution of qubit energies. Alternatively, one can use quantum process tomography to determine the full system Lindbladian [208].

Step 2. Define partitioning of H_{tot} and objectives

The partitioning of H_{tot} is for computational convenience, enabling an efficient use of the toggling frame to determine average Hamiltonians. Default partitioning is $H_{\text{pri}} = H_c + H_{\text{int}}$ and H_{pert} includes variations in H_{int} , as in gate design and simulation and sensing when the Lie algebra of the system can be calculated exactly. For decoupling and time suspension, and simulation and sensing with large systems or when one wants to control the rate of the dynamics, H_{int} can be included in H_{pert} and engineered perturbatively.

Specify U_{target} , H_{target} and T_{seq} . Choose target fidelity for the process. Verify that for the relaxation time and T_{seq} the target FoM is achievable.

Step 3. Controllability of U_{pri}

(a) *Controllability of U_{pri} .* The controllability of U_{pri} is found by calculating the minimal Lie algebra containing $H_{\text{pri}}(t)$. For example, the Lie algebra is $\mathfrak{g}_{\text{pri}} = \text{span}_{\text{Lie}} \{\sigma_x, \sigma_y, \sigma_z\}$ for universal control on a single qubit and the set of Hamiltonians is the generators of $\mathfrak{g}_{\text{pri}}$. The corresponding Lie group $e^{\mathfrak{g}_{\text{pri}}} = \{e^g | g \in \mathfrak{g}_{\text{pri}}\}$ gives the set of achievable $U_{\text{pri}}(T_{\text{seq}})$ [47] and states [85]. Given a set of generators $\{H_i\}$, Algorithm 1 calculates the generated Lie algebra [47].

Algorithm 1 Algorithm for calculating the Lie algebra generated by $\{H_i\}$. The function `FINDLIEALGEBRA`($\{H_i\}$) takes a set of generators $\{H_i\}$ and returns an orthonormal basis of the generated Lie algebra.

```

1: function FINDLIEALGEBRA( $\{H_i\}$ )
2:    $\mathcal{A}_0 \leftarrow i\{H_i\}$ 
3:    $v \leftarrow 0$ 
4:   while  $\mathcal{A}_v \neq \emptyset$  and  $|\cup_{i=1}^v \mathcal{A}_i| < 2^{2N} - 1$  do  $\triangleright 2^N$  is the dimension of the Hilbert space
5:      $\mathcal{A}_{v+1} \leftarrow \emptyset$ 
6:     for  $g \in \mathcal{A}_0$  and  $h \in \mathcal{A}_v$  do
7:       if  $[h, g]$  is linearly independent with  $\cup_{i=1}^{v+1} \mathcal{A}_i$  then  $\triangleright$  Linear dependence can
       be checked by calculating the rank of the matrix of the vectorizations of the operators.
8:         Append  $[h, g]$  to  $\mathcal{A}_{v+1}$ 
9:       end if
10:    end for
11:     $v \leftarrow v + 1$ 
12:  end while
13:  Use the Gram–Schmidt process to orthonormalize  $\cup_{i=1}^v \mathcal{A}_i$  and return the result.
14: end function

```

(b) *Minimal Subspace of $H_{tog}(t)$.* The minimal subspace containing the toggling Hamiltonian is also the subspace of achievable zeroth-order average Hamiltonians. It is

$$\mathcal{C}(\mathbf{g}_{\text{pri}}, H_{\text{pert}}) = \text{span}_{\mathbb{R}} \left\{ U_{\text{pri}}^\dagger(t) H_{\text{pert}} U_{\text{pri}}(t) \mid U_{\text{pri}}(t') \in e^{\mathbf{g}_{\text{pri}}} \right\}. \quad (2.20)$$

Since $H_{\text{tog}}(t)$ is determined by the action of $H_{\text{pri}}(t)$ on H_{pert} , the subspace can be calculated by taking nested commutators between $\mathbf{g}_{\text{pri}}$ and H_{pert} :

$$\mathcal{C}(\mathbf{g}_{\text{pri}}, H_{\text{pert}}) = \text{span}_{\mathbb{R}} \left\{ [g_1, \dots [g_L, H_{\text{pert}}] \dots] \mid g_l \in \mathbf{g}_{\text{pri}}, L \geq 0 \right\}. \quad (2.21)$$

Algorithm 2 calculates an orthonormal basis of $\mathcal{C}(\mathbf{g}_{\text{pri}}, H_{\text{pert}})$:

Algorithm 2 Algorithm for calculating an orthonormal basis of $\mathcal{C}(\mathbf{g}_{\text{pri}}, H_{\text{pert}})$. The function $\text{FINDC}(\mathcal{B}(\mathbf{g}_{\text{pri}}), H_{\text{pert}})$ takes as input a basis $\mathcal{B}(\mathbf{g}_{\text{pri}}) = e_1, \dots, e_{|\mathbf{g}_{\text{pri}}|}$ of $\mathbf{g}_{\text{pri}}$ and the operator H_{pert} , and returns an orthonormal basis of $\mathcal{C}(\mathbf{g}_{\text{pri}}, H_{\text{pert}})$.

```

1: function FINDC( $\mathcal{B}(\mathbf{g}_{\text{pri}}), H_{\text{pert}}$ )
2:    $b_1 \leftarrow H_{\text{pert}}$ 
3:    $i_{\text{max}} \leftarrow 1$  ▷ Total number of basis vectors.
4:    $i_{\text{new}} \leftarrow 1$  ▷ Number of new basis vectors.
5:   while  $i_{\text{new}} > 0$  and  $i_{\text{max}} \leq n^2$  do
6:      $t \leftarrow 0$ 
7:     for  $i_{\text{max}} - i_{\text{new}} + 1 < i < i_{\text{max}}$  do
8:       for  $1 < j < |\mathbf{g}_{\text{pri}}|$  do
9:         if  $b_1, \dots, b_{i_{\text{max}}+t}, [e_j, b_i]$  are linearly independent then
10:           $t \leftarrow t + 1$ 
11:           $b_{i_{\text{max}}+t} \leftarrow [e_j, b_i]$ 
12:        end if
13:      end for
14:    end for
15:     $i_{\text{max}} \leftarrow i_{\text{max}} + t$ 
16:     $i_{\text{new}} \leftarrow t$ 
17:  end while
18:  Use the Gram–Schmidt process to orthonormalize  $(b_1, \dots, b_{i_{\text{max}}})$  and return the
    result
19: end function

```

(c) *Achievable range of scaling factors.* The set of achievable $\bar{H}^{(0)}(T_{\text{seq}})$ is

$$\begin{aligned}
& \mathcal{O}'(\mathbf{g}_{\text{pri}}, H_{\text{pert}}) \\
&= \left\{ \frac{1}{T_{\text{seq}}} \int_0^{T_{\text{seq}}} U_{\text{pri}}^\dagger(t) H_{\text{pert}} U_{\text{pri}}(t) dt \right\}.
\end{aligned} \tag{2.22}$$

It is the convex hull of $\{U_i^\dagger H_{\text{pert}} U_i\}$ ($U_i \in e^{\mathbf{g}_{\text{pri}}}$), which can be approximated through random sampling [133]. Therefore, along each $H_{\text{target}} \in \mathcal{C}(\mathbf{g}_{\text{pri}}, H_{\text{pert}})$, the range of achievable scaling factors can be found by solving a linear program associated with the sampled set (Fig. A.2). The procedure is summarized in Algorithm 4.

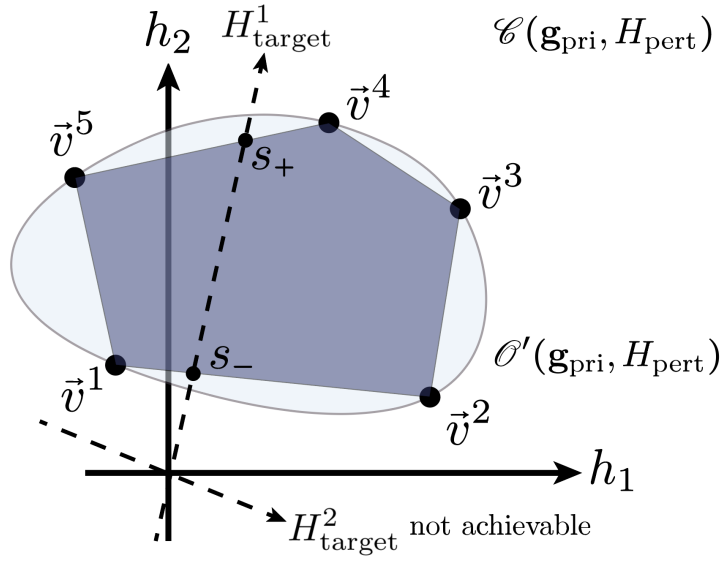


Figure 2.2: Schematic for finding achievable scaling factors. The set of achievable zeroth-order average Hamiltonians $\mathcal{O}'(\mathbf{g}_{\text{pri}}, H_{\text{pert}})$ is a convex set in $\mathcal{C}(\mathbf{g}_{\text{pri}}, H_{\text{pert}})$. The convex set is approximated by the polygon formed by $\{\vec{v}_i = U_i H_{\text{pert}} U_i^\dagger\}$. The achievable scaling factors for H_{target} can then be found via linear programming.

Algorithm 3 (informal version) Algorithm for finding the range of achievable scaling factors for H_{target} .

- 1: Sample random unitaries $U_i \in e^{\mathfrak{g}_{\text{pri}}}$ to form the set $\{U_i^\dagger H_{\text{pert}} U_i\}$
 - 2: Solve the linear program that maximizes (minimizes) the projection of the convex combination of $\{U_i^\dagger H_{\text{pert}} U_i\}$ onto H_{target} ; the solution corresponds to s_+ (s_-)
 - 3: **return** $[s_-, s_+]$ as the range of achievable scaling factors
-

(d) *Modify control or partitioning:* For gate design, simulation and sensing where $H_{\text{pri}} = H_c + H_{\text{int}}$, spectroscopy, and decoupling, if U_{target} or sH_{target} are not achievable, one needs to introduce additional control fields (e.g., RF, microwave, or laser fields) to modify the controllability. For simulation and sensing where H_{pert} includes H_{int} , one can put part of H_{int} in H_{pri} to modify controllability of U_{pri} and $\bar{H}^{(0)}$.

Step 4. Find optimized sequence

Given the relaxation times and desired FoM, determine the maximal T_{seq} of a sequence that allows the target FoM. Specify higher-order corrections including higher-order terms in the Magnus expansion of ΔH and H_{pert} and their cross terms. Use a numerical optimizer, such as gradient descent [97] or simulated annealing [191], to find $\{a_k(t)\}$ that achieves U_{target} and H_{target} within the desired fidelity while accounting for designated error terms and higher-order corrections. A solution should be searched with short T_{seq} first and with an increased T_{seq} if no solution can be found with the current value of T_{seq} so that shorter sequences can be found.

Step 5. Estimate sequence performance

The control landscape is obtained by calculating the fidelity between Λ and target total unitary U_0 for each set of values of $\{\mu_j\}$ and $\{\eta_{\text{int}}^i\}$ drawn from their respective distributions. Simulate the resulting average CPTP map $\bar{\Lambda}$ and/or the average density matrix $\bar{\rho}(T_{\text{seq}})$ of the system (Eq. (2.18)) using, e.g., the Monte Carlo method. Calculate FoM through either the average gate fidelity (Eq. (2.19)) or the state fidelity $\text{Tr}(\bar{\rho}(T_{\text{seq}})\rho_{\text{target}})$. If the FoM is unsatisfactory, check $\bar{\Lambda}$ to identify the error. If the error is due to a depolarization, check if T_{seq} is too long. If not, then the robustness is insufficient, go to Step 4 to include higher-order corrections involving ΔH and variations of H_{int} . If the error is coherent, go to Step 4 to increase target fidelity or include higher-order corrections.

2.5 Examples

We now present examples to demonstrate how to achieve robustness to control errors and variations in H_{int} , characterize controllability, design effective Hamiltonians, and incorporate higher-order corrections. We focus on designing high-precision, robust quantum gates. Then, we will describe the setup for applications in simulation and sensing. An example in spectroscopy is given in the Appendix. More detail will be published in the near future.

2.5.1 Robust Hadamard

We design a Hadamard gate that is robust to variations in Rabi field strength and offset. This sequence is useful for high-fidelity quantum operations in trapped-ion, quantum dot, Rydberg atom, and superconducting systems.

Step 1. System and controls. The internal and control Hamiltonians are

$$\begin{aligned} H_{\text{int}} &= \omega_0 \sigma_z, \\ H_c(t) &= \omega_1(t) \cos(\omega_r t + \phi(t)) \sigma_x, \end{aligned} \tag{2.23}$$

where $\omega_0 = 1$ GHz ($1 \text{ Hz} = 2\pi \text{ rad}\cdot\text{s}^{-1}$) is the resonance frequency of the qubit, ω_r is the transmitter microwave frequency. The control parameters are the amplitude of the applied microwave field ($\omega_1(t)$) and its phase ($\phi(t)$). The maximal value of $\omega_1(t)$ is 20 MHz. The control system has a 500 MHz bandwidth.

It is convenient to move to an interaction frame that rotates at the transmitter frequency ω_r and take the rotating wave approximation [176]. The internal and control Hamiltonians in this frame are

$$\begin{aligned} H_{\text{int}} &= \Delta\omega \sigma_z, \\ H_c(t) &= \omega_1(t) [\cos(\phi(t)) \sigma_x + \sin(\phi(t)) \sigma_y], \end{aligned} \tag{2.24}$$

where $\Delta\omega = \omega_0 - \omega_r$ is the offset.

The gate is robust to variations in the amplitude of the applied microwave field, as the actual amplitude is $(1+\epsilon)\omega_1(t)$ where ϵ is a random variable sampled from $\text{Unif}[-5\%, 5\%]$. Thus, the error Hamiltonian is $\Delta H(t) = \epsilon H_c(t)$. The gate is also robust to variations in $\Delta\omega$, sampled from $\text{Unif}[-2.5 \text{ MHz}, 2.5 \text{ MHz}]$.

The qubit decoherence is modeled as a depolarizing channel with a characteristic time of 100 ms.

Step 2. Objectives and partitioning of H_{tot} . Choose the partitioning $H_{\text{pri}} = H_c(t)$ and $H_{\text{pert}} = H_{\text{int}}$. The objectives are

$$U_{\text{target}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \quad (2.25)$$

and $H_{\text{target}} = 0$. The target FoM is 0.9999.

Step 3. Controllability. The set of generators of $\mathbf{g}_{\text{pri}}$ is $\{\sigma_x, \sigma_y\}$, and their composition via Algorithm 1 provides universal control over the qubit space. Therefore, U_{target} is achievable. Using Algorithm 2, an orthonormal basis of $\mathcal{C}(\mathbf{g}_{\text{pri}}, H_{\text{int}})$ is $\{\sigma_x, \sigma_y, \sigma_z\}/\sqrt{2}$. Choose $H_{\text{target}} = s\sigma_z/\sqrt{2}$. Via Algorithm 3, the range of achievable scaling factors is $[-1, 1]$. Therefore, $H_{\text{target}} = 0$ is achievable in zeroth order.

Step 4. Design robust sequence. Designing for a fidelity of 0.9999 or higher, so given the strength of the depolarizing channel, the maximum gate time T_{seq} is 10 μs . Within this range, the number of intervals can be varied until the target fidelity is achieved. Here, results are shown for a sequence with 35 intervals and $T_{\text{seq}} = 0.35 \mu\text{s}$. The higher-order terms to minimize in the Magnus expansion include the first-order corrections of ΔH and H_{int} , with their cross term. A generalized simulated annealing optimizer [191, 169] is used to search for an optimized sequence. The sequence and its control landscape are shown in Fig. 2.3. For comparison, the control landscapes of a non-robust sequence with the same T_{seq} are also shown.

Step 5. Calculate FoM. To simulate its FoM, assume that $\epsilon \sim \mathcal{N}(0, 0.01^2)$ and that $\Delta \sim \mathcal{N}(0, (0.5\text{MHz})^2)$. The average gate fidelity (Eq. (2.19)) of the resulting CPTP map of the robust and non-robust sequences are estimated to be $0.99990^{+0.00009}_{-0.00004}$ and $0.994^{+0.005}_{-0.003}$. The FoM satisfies the requirement of being at least 0.9999. Fig. 2.4 shows Bloch sphere trajectories of $|0\rangle\langle 0|$ under the sequence for different configurations of ϵ and $\Delta\omega$. The performance of the robust and non-robust sequences can also be compared by calculating the average CPTP maps using Eq. (2.18). The representations under the basis $\{\mathbb{1}, \sigma_x, \sigma_y, \sigma_z\}/\sqrt{2}$ are

$$\begin{aligned} \bar{\Lambda}_{\text{robust}} &= \begin{pmatrix} 1 & & & \\ & 0.0329 & 0.0036 & 0.9994 \\ & -0.0076 & -0.9999 & 0.0039 \\ & 0.9994 & -0.0077 & -0.0329 \end{pmatrix}, \\ \bar{\Lambda}_{\text{non-robust}} &= \begin{pmatrix} 1 & & & \\ & 0.0101 & 0.0067 & 0.9708 \\ & 0.0089 & -0.9865 & 0.0240 \\ & 0.9747 & 0.0336 & -0.0227 \end{pmatrix}. \end{aligned} \quad (2.26)$$

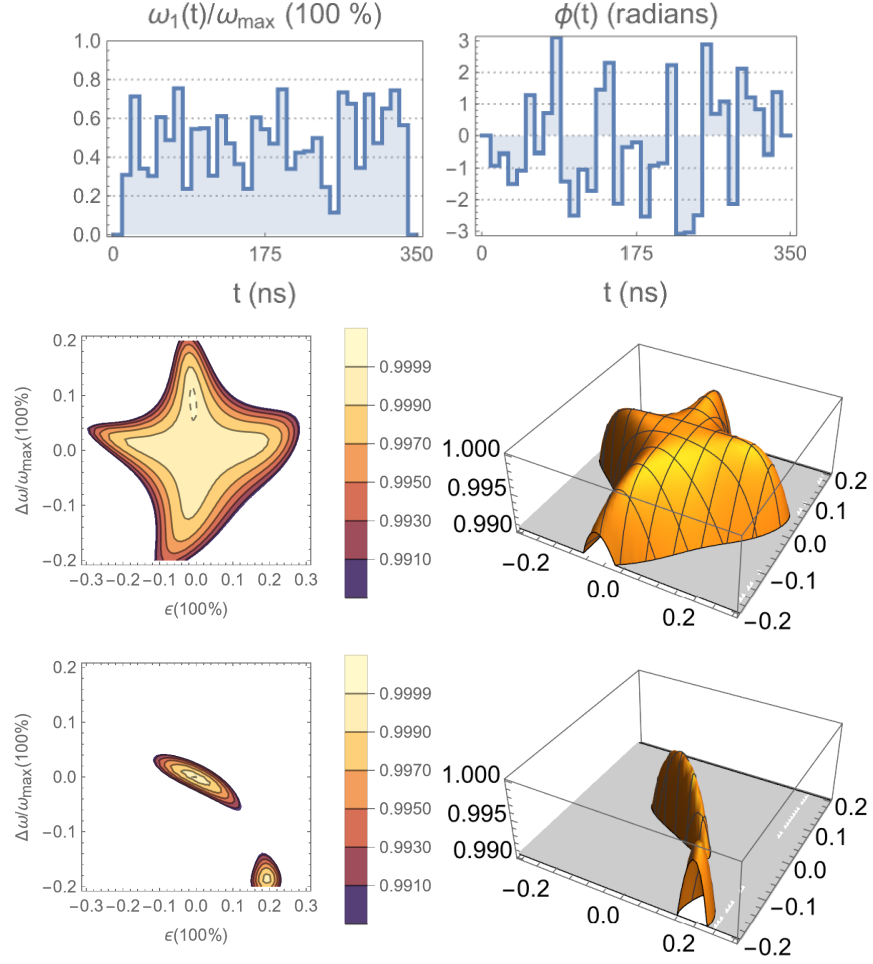


Figure 2.3: Hadamard gate (robust to Rabi and offset variations) and the contour and 3D plots of its control landscape. The bottom plots are the control landscapes of a non-robust Hadamard gate with the same T_{seq} .

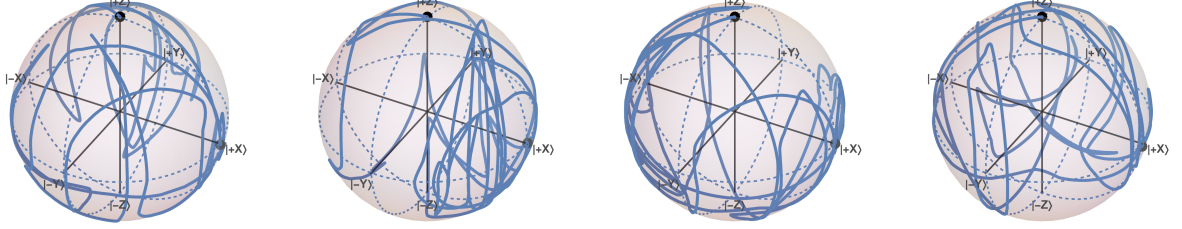


Figure 2.4: Bloch sphere trajectories of the state $|0\rangle\langle 0|$ under the robust Hadamard for different random values of ϵ and $\Delta\omega$ drawn from their distributions. Even though the trajectories are different, the final states are all close to $|+\rangle\langle +|$.

The low FoM of the non-robust sequence is mainly due to its low robustness rather than unitary error. Averaging over parameter distributions causes depolarization of the density matrix, evident in the reduced absolute values of the anti-diagonal elements in the bottom-right block of the CPTP map.

2.5.2 Robust Hadamard with finite bandwidth

Now, consider the same task of designing a Hadamard gate that accounts for a finite bandwidth (80 MHz) in a linear control system. The gate is robust to variations in bandwidth and small anti-symmetric phase transient.

Step 1. System and controls. In the rotating frame ($\omega_0 = \omega_r = 1$ GHz), the internal and control Hamiltonians are

$$\begin{aligned} H_{\text{int}} &= 0, \\ H_c(t) &= B_x(t)\sigma_x + B_y(t)\sigma_y, \end{aligned} \tag{2.27}$$

where $B_x(t)$ and $B_y(t)$ are the time-dependent x and y components of the control field in the rotating frame. The control parameters are $\omega_1(t)$ (maximal value is 20 MHz) and $\phi(t)$ ($H_c(t)$ in Eqs. (2.24)), they are mapped to $\vec{B}(t)$ via a convolution with a kernel Φ

$$\begin{pmatrix} B_x(t) \\ B_y(t) \end{pmatrix} = \Phi \star \begin{pmatrix} \omega_1(t) \cos(\theta(t)) \\ \omega_1(t) \sin(\theta(t)) \end{pmatrix}, \tag{2.28}$$

with

$$\Phi = \begin{pmatrix} \phi^A & \phi^P \\ -\phi^P & \phi^A \end{pmatrix}, \tag{2.29}$$

and

$$\phi^A = \begin{cases} W e^{-Wt} & t > 0 \\ 0 & t \leq 0 \end{cases} \quad \text{and} \quad \phi^P = \begin{cases} \delta (1 - Wt) e^{-Wt} & t > 0 \\ 0 & t \leq 0 \end{cases}. \quad (2.30)$$

Here, the bandwidth W and $\delta = \omega^* - 1$ GHz are the model parameters, and ω^* is the free ringing frequency of the system [130, 7]. W and δ are random variables that follow $W \sim \text{Unif}[70, 90]$ MHz and $\delta \sim \text{Unif}[-20, 20]$ MHz. The leading-order approximation of the error Hamiltonian is

$$\Delta H \approx (W - 80 \text{ MHz}) \frac{\partial H_c}{\partial W}(W = 80 \text{ MHz}) + \delta \frac{\partial H_c}{\partial \delta}(\delta = 0). \quad (2.31)$$

Steps 2 and 3 are the same as example above, except that no H_{target} is needed here.

Step 4. Robustness and optimization. Similar with the previous example, we design for a fidelity of 0.9999 or higher, the maximum gate time T_{seq} is 10 μs . Robustness to ΔH is achieved by minimizing the zeroth-order average Hamiltonians of $\frac{\partial H_c}{\partial W}$ and $\frac{\partial H_c}{\partial \delta}$ simultaneously. Fig. 2.5 shows an optimized sequence ($T_{\text{seq}} = 180$ ns) and its control landscape. As a comparison, the control landscape of a non-robust sequence with the same T_{seq} is also shown.

Step 5. Calculate FoM. Assume that $W \sim \mathcal{N}(80 \text{ MHz}, (2 \text{ MHz})^2)$ ($5\sigma = 10 \text{ MHz}$) and that $\delta \sim \mathcal{N}(0, (4 \text{ MHz})^2)$ ($5\sigma = 20 \text{ MHz}$). The FoM of the robust and non-robust sequences are $0.999995^{+0.000003}_{-0.000002}$ and $0.99994^{+0.00004}_{-0.00003}$.

Note that it is possible to go beyond the perturbative regime and achieve large robust regions by considering alternative partitionings of H_{tot} and enforcing robustness across all of them simultaneously. This is discussed in detail in Appendix A.

2.5.3 Robust control with nonlinear system

Generally, the output field $\vec{B}(t)$ and the input parameter $\{a_k(t)\}$ are related through a state-space differential equation

$$\frac{d\vec{B}}{dt} = h(\vec{B}(t), \{a_k(t)\}, t; \{\mu_j\}), \quad (2.32)$$

where h is a nonlinear function. We can design control sequences robust to variations in $\{\mu_j\}$ given that the equation is not chaotic. As an example, consider the resonant circuit in Fig. A.7. We design a Hadamard gate that is robust to small kinetic inductance. This

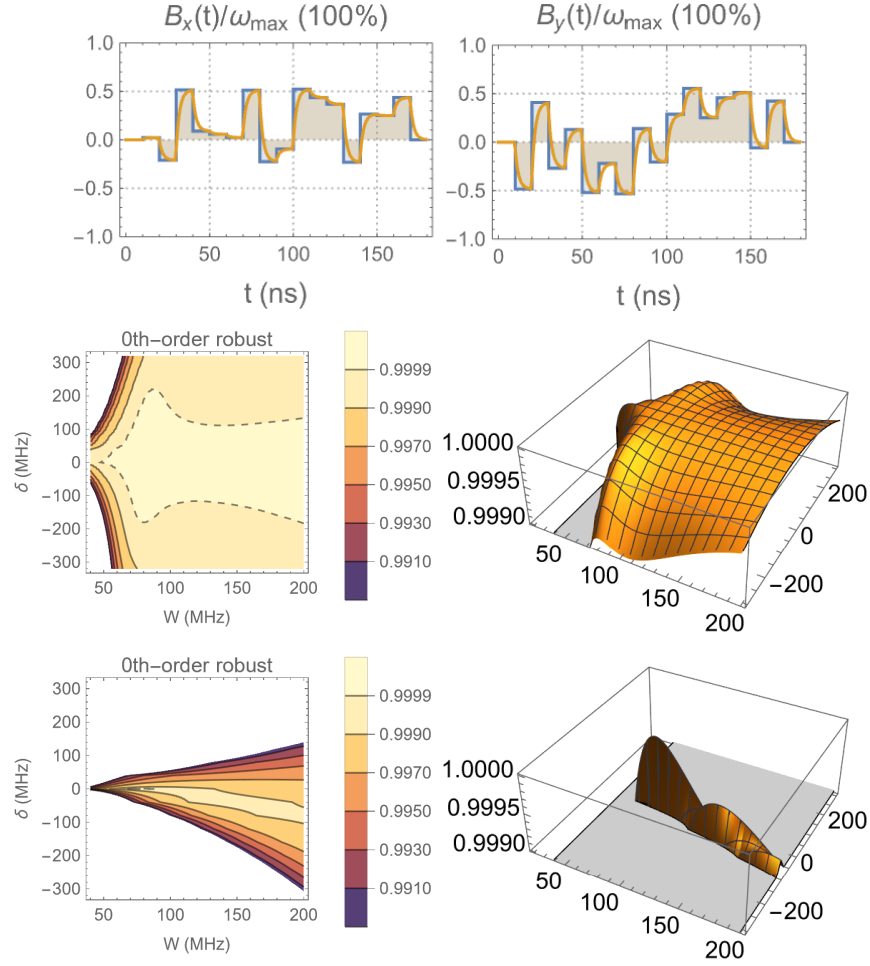


Figure 2.5: Hadamard gate (robust to W and δ variations) and the contour and 3D plots of its control landscape. The blue and orange curves represent the control parameters and relative amplitudes of $B_x(t)$ and $B_y(t)$ respectively. The bottom plots show the control landscapes of a non-robust sequence.

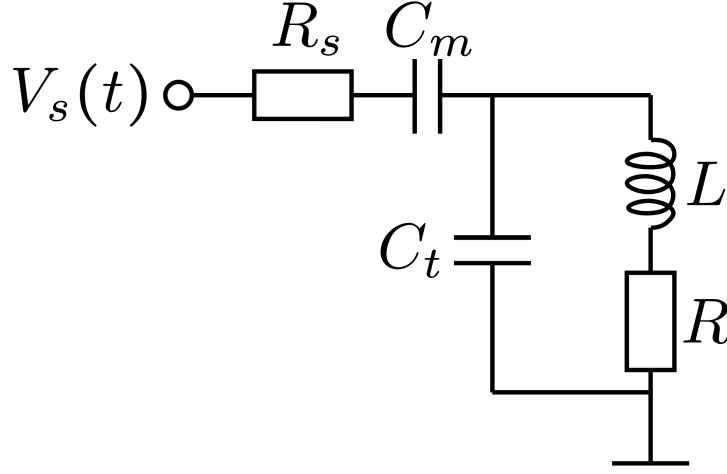


Figure 2.6: A simple resonant circuit with $R_s = 50 \, \Omega$, $R = 0.01 \, \Omega$, $C_m = 11 \, \text{fF}$, $C_t = 1.479 \, \text{pF}$, and $L = 170 \, \text{pH}$. The resonance frequency is 10 GHz.

method can be used to design robust sequences against variations in parameters in general nonlinear control systems.

Step 1. System and controls. Start with the same internal and control Hamiltonians in Eqs. (2.27). The control field $\vec{B}(t)$ can be obtained by solving the following differential equation in the rotating frame [86]

$$\frac{dx}{dt} = A(x)x + \alpha(t)u, \quad (2.33)$$

where

$$x = \begin{pmatrix} \tilde{I}_L \\ \tilde{V}_{C_m} \\ \tilde{V}_{C_t} \end{pmatrix}, \quad A(x) = \begin{pmatrix} -\frac{R}{L} & 0 & \frac{1}{L} \\ 0 & \frac{-1}{R_L C_m} & \frac{1}{R_L C_m} \\ \frac{-1}{C_t} & \frac{-1}{R_L C_t} & \frac{1}{R_L C_t} \end{pmatrix} - i\omega_r \mathbb{1}, \quad u = \begin{pmatrix} 0 \\ \frac{1}{R_L C_m} \\ \frac{1}{R_L C_t} \end{pmatrix}, \quad (2.34)$$

and the control parameters $\{a_k(t)\}$ and the components of the output field $\vec{B}(t)$ are

$$a_1(t) = \kappa_i \text{Re}(\alpha(t)), \quad a_2(t) = \kappa_i \text{Im}(\alpha(t)), \quad B_x(t) = \kappa_o \text{Re}(\tilde{I}_L) \omega_{\max}, \quad B_y(t) = \kappa_o \text{Im}(\tilde{I}_L) \omega_{\max}, \quad (2.35)$$

where $\omega_{\max} = 2 \, \text{MHz}$, $\kappa_i = 1 \, \text{V}^{-1}$ and $\kappa_o = 20 \, \text{A}^{-1}$.

In the presence of kinetic inductance, the inductance L depends on the current:

$$L = L_0(1 + \alpha_L |I_L|^2), \quad (2.36)$$

where α_L is a constant, leading to nonlinear behavior in the circuit [123, 138, 45]. The model parameters are the electrical component parameters in the circuit where $\alpha_L \sim \text{Unif}[-0.001, 0.001]$ A⁻². Note, α_L is typically positive, but we can still design around $\alpha_L = 0$. The solution is robust to α_L that is either positive or negative. The error Hamiltonian ΔH can be calculated by solving the differential equation of $\partial x / \partial \alpha_L$. The general method is provided in the Supplemental material.

Steps 2 and 3 are the same as the previous example.

Step 4. Robustness and optimization. Design for a fidelity of 0.9999 or higher. Fig. 2.7 shows robust and non-robust Hadamard gates. Their control landscapes are shown in Fig. 2.8.

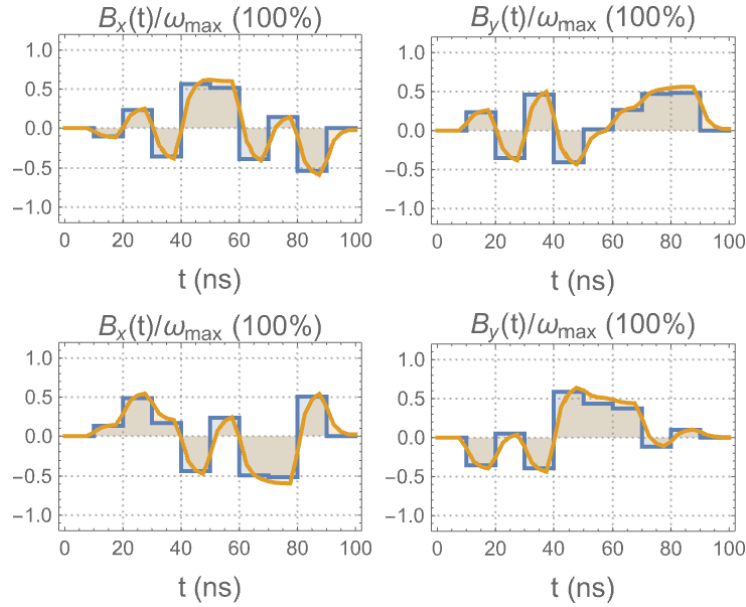


Figure 2.7: The control parameters (blue) and the x and y components of $\vec{B}(t)/\omega_{\max}$ (orange) at $\alpha_L = 0$ of the sequence that is nonrobust (top) and robust (bottom) to variations in α_L .

Step 5. Calculate FoM. To simulate its FoM, assume that α_L obeys a half Gaussian distribution, $\alpha_L \sim \mathcal{HN}(0, (0.0002 \text{ A}^{-2})^2)$ ($5\sigma = 0.001 \text{ A}^{-2}$). The FoM of the robust and non-robust sequences are $0.99991^{0.00008}_{-0.00003}$ and $0.997^{+0.003}_{-0.002}$.

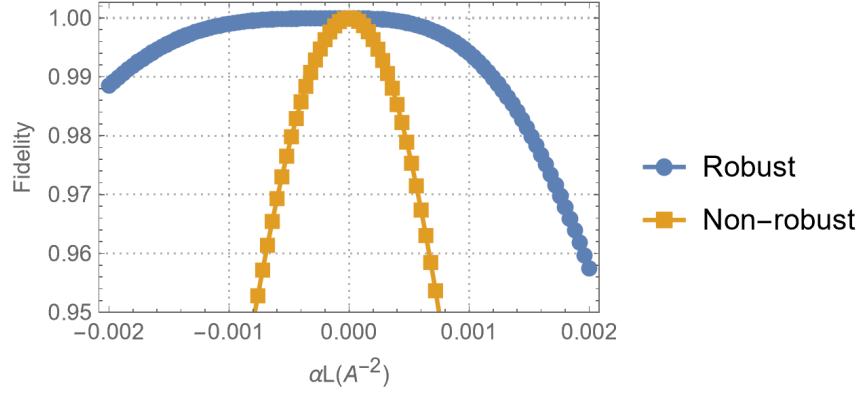


Figure 2.8: Robustness plot of different sequences against variations of the nonlinear coefficient α_L .

2.5.4 Two-qubit gates

The methods for designing robust two-qubit gates are the same as that discussed above. Note that, however, leakage between controls on different qubits can reduce the efficiency of a sequence. Additionally, different forms of qubit interactions can result in different efficiencies. Fig. 2.9 shows the minimal lengths of the sequences we found for a CNOT. The minimal length increases as the detuning Δ between the qubits decreases (leakage increases). It also shows a difference in efficiency due to different forms of the interaction.

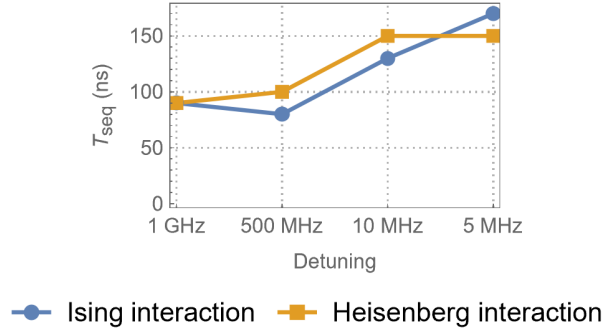


Figure 2.9: Minimal length of sequence for a CNOT gate with different detunings and interaction forms.

2.5.5 Simulating a Tunable Quantum Harmonic Oscillator

We show how to simulate a truncated quantum harmonic oscillator (QHO) using two qubits. The Hamiltonian of a QHO is $H_{\text{QHO}} = \hbar\Omega(\hat{N} + 1/2)$. Using the code [178]

$$\begin{aligned} |n=0\rangle &= |00\rangle, \\ |n=1\rangle &= |01\rangle, \\ |n=2\rangle &= |11\rangle, \\ |n=3\rangle &= |10\rangle, \end{aligned} \tag{2.37}$$

an evolution under H_{QHO} can be simulated by implementing

$$U(T_{\text{seq}}) = \exp[i\{\sigma_z^2(1 + \sigma_z^1/2) - 2\}\Omega T_{\text{seq}}]. \tag{2.38}$$

With universal control over the two qubits, a solution to realize $U(T_{\text{seq}})$ can be easily found using algorithms such as GRAPE [97]. We consider a different approach using Hamiltonian engineering that enables the direct control of the scaling of the simulated dynamics with an input parameter.

Step 1. System and controls. Assume that the bandwidth of the control system is very large and variations in control and model parameters are negligible. In the rotating frame of the qubits, the control and internal Hamiltonians are

$$\begin{aligned} H_c(t) &= \omega_1^1(t) [\cos(\phi_1(t))\sigma_x^1 + \sin(\phi_1(t))\sigma_y^1] + \Delta\omega_1(t)\sigma_z^1 \\ &\quad + \omega_1^2(t) [\cos(\phi_2(t))\sigma_x^2 + \sin(\phi_2(t))\sigma_y^2] + \Delta\omega_2(t)\sigma_z^2, \\ H_{\text{int}} &= J(\sigma_x^1\sigma_x^2 + \sigma_y^1\sigma_y^2 + \sigma_z^1\sigma_z^2). \end{aligned} \tag{2.39}$$

The control parameters are $\omega_1^i(t)$, $\theta_i(t)$ and $\Delta\omega_i$ for $i = 1, 2$. $J = 1.4$ MHz, $\Delta\omega \sim \text{Unif}[-1, 1]$ MHz and $T_1 = 100$ ms.

Step 2. Objectives and partitioning of H_{tot} . Set $\Delta\omega_1(t) \equiv \Delta\omega$ and $\Delta\omega_2(t) \equiv 0$, and define the primary and perturbative Hamiltonians as

$$\begin{aligned} H_{\text{pri}} &= \omega_1^1(t)(\cos(\phi_1(t))\sigma_x^1 + \sin(\phi_1(t))\sigma_y^1) \\ &\quad + \omega_1^2(t)(\cos(\phi_2(t))\sigma_x^2 + \sin(\phi_2(t))\sigma_y^2) \\ &\quad + J(\sigma_x^1\sigma_x^2 + \sigma_y^1\sigma_y^2 + \sigma_z^1\sigma_z^2), \\ H_{\text{pert}} &= \Delta\omega\sigma_z^1. \end{aligned} \tag{2.40}$$

H_{pert} depends on the offset $\Delta\omega$ which can be adjusted at the start of each experiment to modify the rate of the dynamics. The objectives are $U_{\text{pri}} = \mathbb{1}$ and $H_{\text{target}} = s\Delta\omega(\sigma_z^2(1 + \sigma_z^1/2))$.

Step 3. Controllability. Using Algorithm 1, $\mathbf{g}_{\text{pri}}$ generates universal control on the two qubits. Thus, U_{pri} is achievable. Using Algorithm 2, $\mathcal{C}(\mathbf{g}_{\text{pri}}, H_{\text{pert}})$ is fifteen-dimension and includes H_{target} . Using Algorithm 3, we find that the achievable range of s is $[-0.6(2), 0.6(8)]$. Since the scaling of H_{target} can be controlled by adjusting $\Delta\omega$, we design for $s > 0.1$.

Step 4. Robustness and optimization. Design for fidelity 0.9999 or higher. To achieve precision in the engineered effective Hamiltonian, we choose to minimize $\bar{H}^{(1)}$. Note, since H_{int} is included in H_{pri} , $\bar{H}^{(1)}$ cannot be removed by symmetrizing the sequence. Instead, the generic condition (Appendix) for removing $\bar{H}^{(1)}$ is used in the optimization. Fig. 2.10 shows an optimized sequence ($T_{\text{seq}} = 120$ ns), the control landscape is shown in the left of Fig. 2.11. This sequence achieves a scaling factor of 0.221.

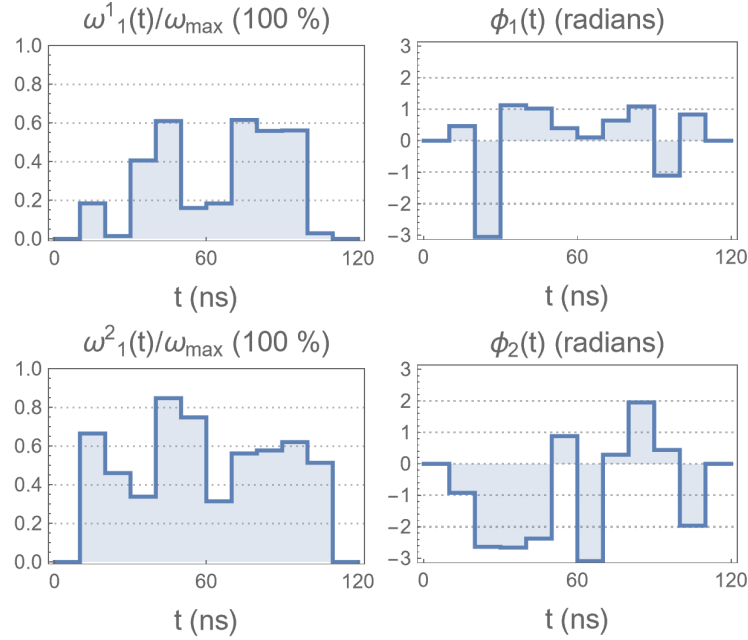


Figure 2.10: Time-dependent control parameters of an optimized sequence that engineers a tunable QHO.

Step 5. Calculate FoM. Assume that $\Delta\omega \sim \mathcal{N}(0, (200 \text{ kHz}^2))$, the FoM of the sequence is $0.999997^{+0.000006}_{-0.000005}$.

As mentioned before, the sequence allows simulation of a tunable QHO whose frequency can be adjusted by directly changing $\Delta\omega_1(1) \equiv \Delta\omega$. The right-hand side of Fig. 2.11 shows the survival probability of $(|n=1\rangle + |n=3\rangle)/\sqrt{2}$ under the sequence with different values of $\Delta\omega$.

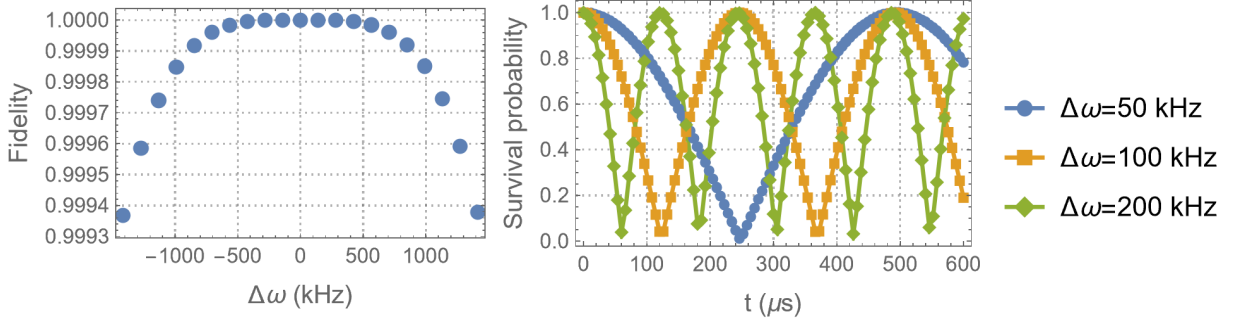


Figure 2.11: Control landscape (left) and simulation of the evolution (right) of the state $(|n=1\rangle + |n=3\rangle)/\sqrt{2}$ under the sequence shown in Fig. 2.10.

2.5.6 Sensing with a Scalable Spin Amplifier

Consider designing a scalable solution for realizing a spin amplifier for measuring the state of a qubit. The quantum system of the target qubit and the spin amplifier is initially prepared in the state $|\phi_T\rangle \otimes |0_1 0_2 \dots 0_N\rangle$. The goal is to design a sequence such that at the end of the sequence we can learn the state of the target qubit $|\phi_T\rangle$ by measuring the collective magnetization M_z of qubits 1 to N in the spin amplifier.

A scalable solution can be obtained by engineering conditioned dynamics of the spin amplifier controlled by the state of the target qubit [36]. This can be achieved using internal interactions of the system and only collective rotations of qubits in the spin amplifier.

We design a conditional decoupling of the internal dynamics of the spin amplifier. Under the applied sequence, the internal interaction is suppressed when $|\phi_T\rangle = |0\rangle$ and active when $|\phi_T\rangle = |1\rangle$. In the experiment, a collective $\pi/2)_x$ pulse first rotates the amplifier polarization to the transverse plane, after which the system evolves under the engineered dynamics, followed by a $-\pi/2)_x$ readout pulse. When the interaction is active ($|\phi_T\rangle = |1\rangle$), the amplifier evolves into an entangled state with an average longitudinal magnetization $M_z \approx 0$ for large N , whereas when it is decoupled ($|\phi_T\rangle = |0\rangle$), M_z remains close to its initial value. Measuring M_z at the end of the sequence thus reveals the state of $|\phi_T\rangle$.

Step 1. System and controls. The internal and control Hamiltonians are

$$\begin{aligned}
 H_c(t) &= C(t) \left(\cos(\phi(t)) \sum_{i=1}^N \sigma_x^i + \sin(\phi(t)) \sum_{i=1}^N \sigma_y^i \right), \\
 H_{\text{int}} &= \sigma_z \otimes \sum_{i=1}^N \omega_i \sigma_z^i + \mathbb{1} \otimes \sum_{i < j}^N \omega_{ij} (2\sigma_z^i \otimes \sigma_z^j - \sigma_x^i \otimes \sigma_x^j - \sigma_y^i \otimes \sigma_y^j).
 \end{aligned} \tag{2.41}$$

The control parameters are $\omega(t)$ ($\omega_{\max} = 20$ MHz) and $\phi(t)$. The internal parameters are the coupling strength: $\omega_i \sim \text{Unif}[441, 459]$ kHz and $\omega_{ij} \sim \text{Unif}[-22.5, 22.5]$ kHz. $T_1 = 100$ ms.

Step 2. Objectives and partitioning of H_{tot} . Define the primary and perturbative Hamiltonians as

$$\begin{aligned} H_{\text{pri}}(t) &= H_c(t) + \sigma_z \otimes \sum_{i=1}^N \omega_0 \sigma_z^i, \\ H_{\text{pert}}^1 &= \sigma_z \otimes \sum_{i=1}^N (\omega_i - \omega_0) \sigma_z^i, \\ H_{\text{pert}}^2 &= \mathbb{1} \otimes \sum_{i < j}^N \omega_{ij} (2\sigma_z^i \otimes \sigma_z^j - \sigma_x^i \otimes \sigma_x^j - \sigma_y^i \otimes \sigma_y^j), \end{aligned} \quad (2.42)$$

where $\omega_0 = 450$ kHz. The objectives are

$$\begin{aligned} U_{\text{pri}} &= \mathbb{1}, \\ H_{\text{target}}^2 &= s\sqrt{2}(|1\rangle\langle 1| \otimes \sum_{i < j}^N \omega_{ij} (2\sigma_z^i \otimes \sigma_z^j - \sigma_x^i \otimes \sigma_x^j - \sigma_y^i \otimes \sigma_y^j)). \end{aligned} \quad (2.43)$$

Step 3. Controllability. To check the controllability, it is sufficient to consider two spins from the spin amplifier. Via Algorithm 1, $\mathbf{g}_{\text{pri}}$ is three-dimensional and $U_{\text{pri}} = \mathbb{1}$ is achievable. Using Algorithm 2, $\mathcal{C}_2(\mathbf{g}_{\text{pri}}, H_{\text{pert}}^2)$ is ten-dimensional and includes H_{target}^2 . Using Algorithm 3, we can find the achievable range of the scaling factor s (Table 2.2). For comparison, Table 2.2 also lists some other achievable H_{target}^2 and their ranges of scaling factors.

Table 2.2: Achievable ranges of scaling factors for different target effective Hamiltonians. They correspond to conditioned decoupling, double/zero quantum dynamics, and double/single quantum dynamics.

Target Hamiltonian	Achievable range
$ 1\rangle\langle 1 \otimes (2\sigma_z \otimes \sigma_z - \sigma_x \otimes \sigma_x - \sigma_y \otimes \sigma_y)$	$[-0.353(5), 0.706(8)]$
$ 0\rangle\langle 0 \otimes (\sigma_x \otimes \sigma_x - \sigma_y \otimes \sigma_y) + 1\rangle\langle 1 \otimes (2\sigma_z \otimes \sigma_z - \sigma_x \otimes \sigma_x - \sigma_y \otimes \sigma_y)$	$[-0.408(2), 0.811(4)]$
$ 0\rangle\langle 0 \otimes (\sigma_x \otimes \sigma_y + \sigma_y \otimes \sigma_x) + 1\rangle\langle 1 \otimes (\sigma_x \otimes \sigma_z + \sigma_z \otimes \sigma_x)$	$[-0.576(4), 0.576(8)]$

Step 4. Robustness and optimization. Design for a fidelity of 0.9999 or higher. Fig. 2.12 shows an optimized sequence ($T_{\text{seq}} = 120$ ns). It achieves a scaling factor of $s = 0.327$. To obtain its control landscape, consider a spin amplifier of 6 qubits with the coupling

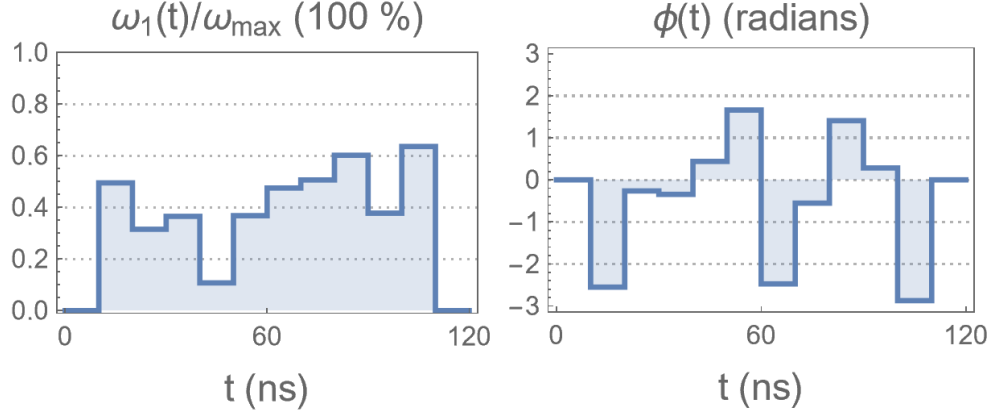


Figure 2.12: Time-dependent control parameters of the sequence that achieves a scalable spin amplifier through engineering a conditioned decoupling.

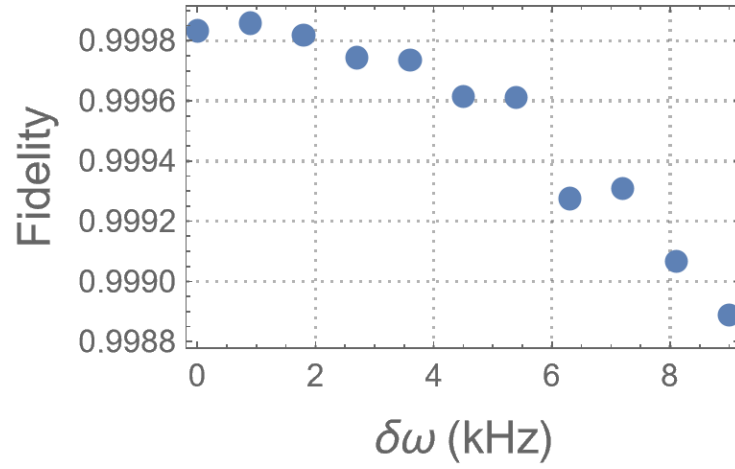


Figure 2.13: Control landscape of the sequence in Fig. 2.12. The parameter $\Delta\omega$ determines the width of the distribution of the coupling strength between the target qubit and the spin amplifier.

strengths generated from $\omega_i \sim \text{Unif}[450 + [-\delta\omega, \delta\omega]]$ kHz and $\omega_{ij} \sim \text{Unif}[-22.5, 22.5]$ kHz. The mean fidelity for different values of $\delta\omega$ is plotted in Fig. 2.13.

Step 5. Calculate FoM. To calculate the FoM of the sequence, use a spin amplifier of 6 qubits with the coupling strengths generated from $\omega_i \sim \text{Unif}[441, 459]$ kHz and $\omega_{ij} \sim \text{Unif}[-22.5, 22.5]$ kHz. The FoM, calculated as the average gate fidelity of the sequence, is $0.9997^{+0.0001}_{-0.0001}$.

To test the sequence, we simulate the same spin amplifier used in the calculation of the FoM. The initial state of the system evolves under the internal dynamics modulated by the designed sequence. Before and after each simulated sequence, a $\pi/2)_x$ and a $-\pi/2)_x$ pulse are applied, respectively. The survival probability of the initial state is then measured at the end of the simulation. Repeating this procedure for different numbers of applied sequences yields the time evolution of the survival probability. The survival probabilities of the states

$$|\psi_{0/1}\rangle = |\phi_T = 0/1\rangle \otimes |00\cdots 0\rangle \quad (2.44)$$

under the sequence are plotted in Fig. 2.14. The survival probability is proportional to M_z and can be used to learn the state $|\phi_T\rangle$. As can be seen from the result, after 20 μs ,

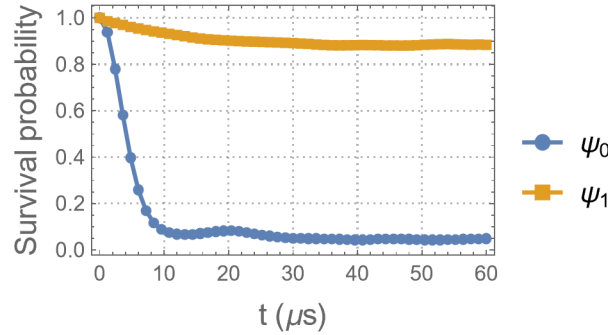


Figure 2.14: Simulated survival probabilities of the states $|\psi_{0/1}\rangle$ under the sequence in Fig. 2.12.

the survival probability of $|\psi_0\rangle$ is above 0.8 while that of $|\psi_1\rangle$ is below 0.1. So the states $|\psi_{0/1}\rangle$ can be distinguished by measuring the magnetization of the spin amplifier, which is proportional to the survival probability.

In principle, one can include robustness by requiring $H_{\text{target}}^1 = 0$ and including higher-order corrections ². In practice, however, we could not find a sequence that fulfills these requirements simultaneously while maintaining a high scaling factor without significantly increasing the length of the sequence. In turn, decreased scaling factor or fidelity of

²To minimize the cross term between H_{pert}^1 and H_{pert}^2 . To achieve this, consider $N = 2$, the total

U_{pri} result in decreased FoM and inferior overall performance in sensing. This is because $|H_{\text{pert}}^1| \ll |H_{\text{pert}}^2|$, thus a larger scaling factor on H_{pert}^2 is much more important. Also, large scaling factors are crucial when engineering nonzero effective Hamiltonians, as higher-order corrections are typically much smaller and less significant than the dominant low-order scaling.

2.6 Conclusion

In conclusion, this work presents a general framework for precise and robust quantum control through effective Hamiltonian engineering which addresses two important issues. First, robustness to variations in parameters of any control system can be achieved by minimizing the effective Hamiltonian of the error Hamiltonian. Second, when designing control through effective Hamiltonians, the method identifies the subspace of the achievable zeroth-order average Hamiltonians and finds the range of scaling factors through linear programming. Controllability can be adjusted by partitioning the total Hamiltonian into a primary Hamiltonian and a perturbative Hamiltonian. Generic conditions for removing higher-order corrections are also introduced. The framework is combined with numerical optimization to design optimized sequences.

There are three natural extensions of the present work:

1. The conditions for removing higher-order corrections can be further refined by considering connectivity of the Hamiltonians. This also identifies achievable effective Hamiltonians in higher orders, enabling engineering higher-order terms.
2. In practice, the controllability of achievable effective Hamiltonians may be limited by the efficiency of the sequences. Large scaling factor or higher-order robustness may require a very long sequence, making the solution less useful. It would be important to understand the upper bound of efficiency and how to achieve it in the presence of limited resources (e.g., power bandwidth product and gate time).

Hamiltonian is

$$H_{\text{tot}}(t) = H_{\text{pri}}(t) + (\omega_1 - \omega_0)\sigma_z \otimes \sigma_z \otimes \mathbb{1} + (\omega_2 - \omega_0)\sigma_z \otimes \mathbb{1} \otimes \sigma_z + H_{\text{pert}}^2. \quad (2.45)$$

Since both $H_{\text{pri}}(t)$ and H_{pert}^2 are symmetric with respect to a permutation of the two qubits in the spin amplifier, we only need to consider $(\omega_1 - \omega_0)\sigma_z \otimes \sigma_z \otimes \mathbb{1}$ for leading-order robustness.

3. It would also be interesting to extend the method for designing controls against decoherence in open quantum systems when the memory effect of the environment is considered.

This work offers a deeper understanding of effective Hamiltonian engineering and represents an important step toward the automated design of high-performance sequences for quantum control in different systems.

Chapter 3

Engineering Higher-order Effective Hamiltonians

Last chapter focuses on engineering the zeroth-order average Hamiltonian while minimizing higher-order terms. However, not all effective Hamiltonians can be engineered in zeroth order. Here, we describe a systematic methodology for achieving the precision, robustness, and complexity required for practical applications, particularly through the engineering of higher-order processes and effective Hamiltonians. We focus on three key use cases:

1. When the zeroth-order solution correctly implements the desired process, higher-order terms must be suppressed.
2. When the target unitary involves more qubits than the internal and control operators, making a zeroth-order solution unavailable.
3. When the target unitary includes qubit correlations absent in the internal and control Hamiltonians.

In cases 2 and 3, the desired control terms emerge only through nested commutators of the control and internal Hamiltonians.

We will show that the topology of the qubit network determines the achievable interactions. We derive the space of feasible higher-order effective Hamiltonians of each order and provide necessary and sufficient conditions for engineering target Hamiltonians using numerically optimized sequences.

The foundation of the method is the space of the achievable zeroth-order average Hamiltonian $\mathcal{C}(\mathbf{g}_{\text{pri}}, H_{\text{pert}})$ introduced in the last chapter. Represented in $\mathcal{C}(\mathbf{g}_{\text{pri}}, H_{\text{pert}})$, the toggling Hamiltonian is

$$|H_{\text{tog}}(t)\rangle\rangle = (c_1(t), c_2(t), \dots, c_{|\mathcal{C}|}(t)). \quad (3.1)$$

Denote the time-ordered integrals of the coefficients $c_i(t)$ ($\mathcal{C}$ -integrals) as

$$\bar{c}_{i_1 \dots i_r} = \int_0^{T_{\text{seq}}} dt_1 \dots \int_0^{t_{r-1}} dt_r c_{i_1}(t_1) \dots c_{i_r}(t_r), \quad (3.2)$$

a generic condition for $\bar{H}^{(r-1)} = 0$ is

$$\bar{c}_{i_1 i_2 \dots i_r} = (-1)^r \bar{c}_{i_r i_{r-1} \dots i_1}, \quad (3.3)$$

for all $i_1, i_2, \dots, i_r$ except when $i_1 = i_2 = \dots = i_r$. As will be shown below, this condition can be refined.

3.1 Engineering $\bar{H}^{(r-1)}$

When the values of $\{\eta_{\text{int}}^i\}$ are exact, $\bar{H}^{(r-1)}$ can be calculated using, for example, the van Loan equations [79]. When $\{\eta_{\text{int}}^i\}$ are not exact, if the variations are small, robust solutions can be found using the method in the last chapter. When the variations are large, we seek symmetry-based solutions that hold for all $\{\eta_{\text{int}}^i\}$. We consider the situation when $H_{\text{pri}}(t)$ is known exactly, and the perturbative Hamiltonian H_{pert} depends on parameters $\{\eta_{\text{pert}}^m\}_{m=1}^M$, which is the subset of $\{\eta_{\text{int}}^i\}$ subjected to variations.

Assuming that H_{pert} is linear in $\{\eta_{\text{pert}}^m\}$, then $\bar{H}^{(r-1)}$ and any achievable target Hamiltonian $\bar{H}_{\text{target}}^{(r-1)}$ are polynomials of $\{\eta_{\text{pert}}^m\}$ of order up to r . Therefore, $\bar{H}^{(r-1)} = \bar{H}_{\text{target}}^{(r-1)}$ for all $\{\eta_{\text{pert}}^m\}$ if and only if

$$\frac{\partial^r \bar{H}^{(r-1)}}{\partial \eta_{\text{pert}}^{m_1} \dots \partial \eta_{\text{pert}}^{m_r}} = \frac{\partial^r \bar{H}_{\text{target}}^{(r-1)}}{\partial \eta_{\text{pert}}^{m_1} \dots \partial \eta_{\text{pert}}^{m_r}} \bigg|_{\bar{\eta}_{\text{pert}} = \bar{0}}, \quad (3.4)$$

for $m_1, m_2, \dots, m_s \in \{1, 2, \dots, M\}$. Given a basis $\{h_1, \dots, h_{|\mathcal{C}|}\}$ of $\mathcal{C}(\mathbf{g}_{\text{pri}}, H_{\text{pert}})$,

$$\begin{aligned} & \bar{H}^{(r-1)} \\ &= \int_0^t dt_1 \dots \int_0^{t_{r-1}} dt_r F_r(H_{\text{tog}}(t_1), H_{\text{tog}}(t_2), \dots, H_{\text{tog}}(t_r)), \\ &= \sum_{i_1, \dots, i_r=1}^{|\mathcal{C}|} F_r(h_{i_1}, h_{i_2}, \dots, h_{i_r}) \bar{c}_{i_1 \dots i_r}. \end{aligned} \quad (3.5)$$

where

$$F_r(h_{i_1}, h_{i_2}, \dots, h_{i_r}) = \sum_{\pi} (-i)^{r-1} (-1)^{d_b} \frac{d_a! d_b!}{r!} \pi(h_{i_1} h_{i_2} \dots h_{i_r}), \quad (3.6)$$

where the sum is taken over all possible permutations π of $h_{i_1}, h_{i_2}, \dots, h_{i_r}$. d_a and d_b are ascents and descents of π . Thus, Eq. (3.4) becomes

$$\sum_{i_1, \dots, i_r=1}^{|\mathcal{C}|} \frac{\partial^r F_r(h_{i_1}, h_{i_2}, \dots, h_{i_r})}{\partial \eta_{\text{pert}}^{m_1} \dots \partial \eta_{\text{pert}}^{m_r}} \bar{c}_{i_1 \dots i_r} = \left. \frac{\partial^r \bar{H}_{\text{target}}^{(r-1)}}{\partial \eta_{\text{pert}}^{m_1} \dots \partial \eta_{\text{pert}}^{m_r}} \right|_{\bar{\eta}_{\text{pert}} = \bar{0}}. \quad (3.7)$$

For decoupling, the right-hand side of the equation is 0. It reduces to Eq. (3.3) if it is required to hold for any $h_{i_1}, h_{i_2}, \dots, h_{i_r}$. The space of achievable $(r-1)$ th-order effective Hamiltonian is isomorphic to that spanned by $\left\{ \left. \frac{\partial^r F_r(h_{i_1}, h_{i_2}, \dots, h_{i_r})}{\partial \eta_{\text{pert}}^{m_1} \dots \partial \eta_{\text{pert}}^{m_r}} \right|_{\bar{\eta}_{\text{pert}} = \bar{0}} \right\}$.

3.2 Engineering $\bar{H}^{(r-1)}$ on qubit networks

As an example application for Eq. (3.7), consider engineering effective Hamiltonians on qubit networks, such as nearest-neighbor chains and square lattices. They can be represented by a graph $G = (V(G), E(G))$ where the set of vertices $V(G)$ and the set of edges $E(G)$ represent the qubits and their connectivity, respectively. For simplicity, we will consider collective control of the qubits

$$\begin{aligned} H_{\text{pri}}(t) &= H_c(t) \\ &= \omega_1(t) \left[\cos(\phi(t)) \sum_{i=1}^N \sigma_x^i + \sin(\phi(t)) \sum_{i=1}^N \sigma_y^i \right] + (\omega_0 - \omega_r(t)) \sum_{i=1}^N \sigma_z^i \end{aligned} \quad (3.8)$$

where $\omega_1(t)$, $\omega_r(t)$ and $\phi(t)$ are the time-dependent amplitude, frequency, and phase of the control field, and ω_0 is the resonance frequency, and an internal Hamiltonian

$$\begin{aligned} H_{\text{pert}} &= H_{\text{int}} = \sum_{i < j}^N B_{ij} \vec{\sigma}_i \cdot \mathbf{D} \cdot \vec{\sigma}_j + \sum_{i=1}^N \delta_i \sigma_z^i \\ &= H_{\text{pert}}^1 + H_{\text{pert}}^2 \end{aligned} \quad (3.9)$$

where $\vec{\sigma}_i = (\sigma_x^i, \sigma_y^i, \sigma_z^i)$ and

$$\sigma_{\alpha}^i = \mathbb{1}^{\otimes(i-1)} \otimes \sigma_{\alpha} \otimes \mathbb{1}^{\otimes(N-i)}, \quad (3.10)$$

for $\alpha = \{x, y, z\}$, are the Pauli operators acting on the i th qubit. $\{B_{ij}\}$ are coupling strengths, $\mathbf{D}$ is a rank-2 tensor determining the form of the qubit interactions, δ_i the strength of the qubit-dependent energy offset, and N is the total number of qubits.

Let $\mathcal{B}(\mathcal{C}_1) = \{h_1^i, \dots, h_{|\mathcal{C}_1|}^i\}$ be an orthonormal basis of $\mathcal{C}_i(\mathbf{g}_{\text{pri}}, H_{\text{pert}}^i)$, for $i = 1, 2$. Define $\mathcal{C} = \mathcal{C}_1(\mathbf{g}_{\text{pri}}, H_{\text{pert}}^1) \oplus \mathcal{C}_2(\mathbf{g}_{\text{pri}}, H_{\text{pert}}^2)$, its basis is

$$\mathcal{B}(\mathcal{C}) = \{h_1, \dots, h_{|\mathcal{C}|}\} = \{h_1^1, \dots, h_{|\mathcal{C}_1|}^1, h_1^2, \dots, h_{|\mathcal{C}_2|}^2\}. \quad (3.11)$$

Denote a labeled graph with r edges as G_r , and define

$$\begin{aligned} & F_r(G_r; i_1, \dots, i_r) \\ &= \sum_{\pi(j_1 \dots j_r)} F_r(h_{i_1}(e_{j_1}), h_{i_2}(e_{j_2}), \dots, h_{i_r}(e_{j_r})) \\ &= \left. \frac{\partial^r F_r(h_{i_1}, h_{i_2}, \dots, h_{i_r})}{\partial \eta_{\text{int}}^{e_{j_1}} \dots \partial \eta_{\text{int}}^{e_{j_r}}} \right|_{\vec{\eta}_{\text{int}} = \vec{0}}, \end{aligned} \quad (3.12)$$

where $e_{j_1}, \dots, e_{j_r} \in E(G_r)$ are different edges of G_r ,

$$\eta_{\text{int}}^{ij} = \begin{cases} B_{ij} & i \neq j \\ \delta_i & i = j \end{cases}, \quad (3.13)$$

and

$$h(e) = \text{Tr}_{\text{rest}}(h) \otimes \mathbb{1}_{\text{rest}}, \quad (3.14)$$

where the partial trace, $\text{Tr}_{\text{rest}}(h)$, is taken over the Hilbert space of all qubits except those incident to $e \in E(G_r)$, and $\mathbb{1}_{\text{rest}}$ represents the identity operator acting on these qubits. $F_r(G_r; i_1, \dots, i_r)$ vanishes if G_r is not connected. Denote

$$H_{\text{target}}^{(r-1)}(G_r) = \left. \frac{\partial^r H_{\text{target}}^{(r-1)}}{\partial \eta_{\text{int}}^{e_{j_1}} \dots \partial \eta_{\text{int}}^{e_{j_r}}} \right|_{\vec{B} = \vec{0}}, \quad (3.15)$$

Eq. (3.7) becomes

$$\sum_{i_1, \dots, i_r=1}^{|\mathcal{C}|} F_r(G_r; i_1, \dots, i_r) \bar{c}_{i_1, \dots, i_r} = H_{\text{target}}^{(r-1)}(G_r) \quad (3.16)$$

for all nonisomorphic G_r that are weakly isomorphic to connected subgraphs of G . Weak isomorphism is a bijection that preserves the vertex set and connectivity but allows differing numbers of parallel edges and loops, unlike isomorphism, which also preserves

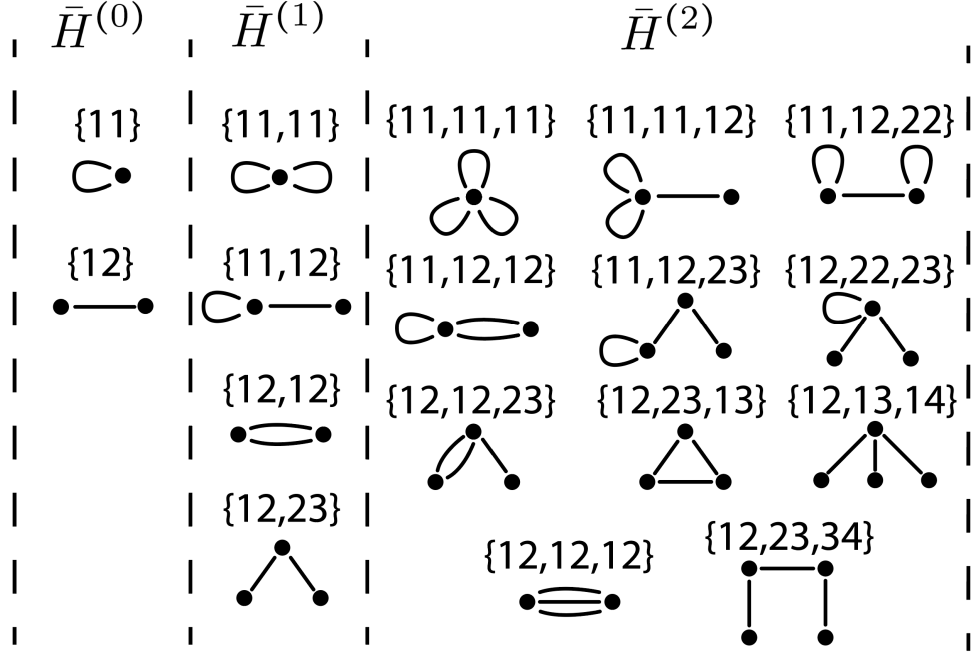


Figure 3.1: Nonisomorphic graphs associated with different Magnus terms.

edge multiplicities. The space of achievable $H_{\text{target}}^{(r-1)}$ is isomorphic to that spanned by $\{F_r(G_r; i_1, \dots, i_r)\}$. An effective Hamiltonian corresponding to G_r can be mapped to the real effective Hamiltonian acting on the total Hilbert space of the network through

$$H_{\text{eff}}^{(r-1)}(N) = \sum_{G'_r \sim G_r} \eta_{\text{int}}(G'_r) H_{\text{eff}}^{(r-1)}(G_r), \quad (3.17)$$

where the sum is taken over all G'_r that are isomorphic to G_r and

$$\eta_{\text{int}}(G'_r) = \sum_{e_1, \dots, e_r \in E(G'_r)} \eta_{\text{int}}^{e_1} \dots \eta_{\text{int}}^{e_r}. \quad (3.18)$$

The total effective Hamiltonian is then $\sum_{G_r} H_{\text{eff}}^{(r-1)}(G_r)$ where the sum is taken over all nonisomorphic graphs with r edges. Fig. 3.1 shows nonisomorphic graphs for different $\bar{H}^{(r-1)}$ for a network with all-to-all interactions. Fig. 3.2 shows nonisomorphic graphs for $\bar{H}^{(2)}$ for qubit systems with different topology. If one wishes to minimize the effect from the offset, then $H_{\text{target}}^{(r-1)}(G_r) = 0$ for all G_r with loops.

$\bar{H}^{(2)}$				
	✓	✓	✓	✓
	✗	✓	✓	✓
	✗	✗	✗	✓
	✗	✗	✓	✓
	✗	✓	✓	✓

Figure 3.2: Examples of network topologies and the corresponding loop-free graphs contributing to $\bar{H}^{(2)}$. From left to right: spin pairs, nearest-neighbor spin chain, square spin lattice, and all-to-all coupled spin ensemble.

3.3 Choosing Target Effective Hamiltonians

Since G_r contains at most $r + 1$ vertices, it is sufficient to consider a $(r + 1)$ -qubit system when engineering $\bar{H}^{(r-1)}$. Denote

$$\mathcal{S}(G_r) = \text{span}_{i_1, \dots, i_r=1}^{|\mathcal{C}|} \{F_r(G_r; i_1, \dots, i_r)\}. \quad (3.19)$$

$\mathcal{S}(G_r)$ represents the space of achievable $\bar{H}^{(r-1)}$ for G_r and are disjoint for different G_r (since they have different dependence on $\{\eta_{\text{int}}^{ij}\}$). Thus, the space of achievable $H_{\text{target}}^{(r-1)}$ is isomorphic to $\cup_{G_r} \mathcal{S}(G_r)$. Given an orthonormal basis of $\mathcal{S}(G_r)$

$$\mathcal{B}(\mathcal{S}(G_r)) = \{h_i(G_r)\}, \quad (3.20)$$

define

$$\vec{c}^{(r-1)} = (\bar{c}_{1\dots 11}, \bar{c}_{1\dots 12}, \dots, \bar{c}_{|\mathcal{C}|\dots |\mathcal{C}||\mathcal{C}|})^T, \quad (3.21)$$

the condition for engineering $\bar{H}^{(r-1)}$ (Eq. (3.16)) becomes

$$\mathbf{A}(G_r) \cdot \vec{c}^{(r-1)} = \vec{d}(G_r), \quad (3.22)$$

where

$$\mathbf{A}(G_r) = \begin{pmatrix} \langle\langle h_1(G_r)|F_r(G_r; 1, \dots, 1) \rangle\rangle & \cdots & \langle\langle h_1(G_r)|F_r(G_r; |\mathcal{S}(G_r)|, \dots, |\mathcal{S}(G_r)|) \rangle\rangle \\ \vdots & \ddots & \vdots \\ \langle\langle h_{|\mathcal{S}(G_r)|}(G_r)|F_r(G_r; 1, \dots, 1) \rangle\rangle & \cdots & \langle\langle h_{|\mathcal{S}(G_r)|}(G_r)|F_r(G_r; |\mathcal{S}(G_r)|, \dots, |\mathcal{S}(G_r)|) \rangle\rangle \end{pmatrix}, \quad (3.23)$$

and

$$\vec{d}(G_r) = \left(\langle\langle h_1(G_r)|H_{\text{target}}^{(r-1)}(G_r) \rangle\rangle, \dots, \langle\langle h_{|\mathcal{S}(G_r)|}(G_r)|H_{\text{target}}^{(r-1)}(G_r) \rangle\rangle \right)^T, \quad (3.24)$$

where $\langle\langle \cdot | \cdot \rangle\rangle = \text{Tr}(\cdot \cdot^\dagger)$. The total condition for engineering $\bar{H}^{(r-1)}$ can be represented as

$$\bigoplus_{G_r} \mathbf{A}(G_r) = \bigoplus_{G_r} \vec{d}(G_r), \quad (3.25)$$

where the direct sum is taken over all nonisomorphic graphs with r edges for $\bar{H}^{(r-1)}$.

Assuming that the graphs associated with $\bar{H}^{(r-1)}$ are $G_r^1, G_r^2, \dots, G_r^{n_r}$, target effective Hamiltonians are chosen from $\mathcal{S}(G_r^1), \mathcal{S}(G_r^2), \dots, \mathcal{S}(G_r^{n_r})$ individually. The choices must not contradict one another. The space of achievable $\bar{H}^{(r-1)}$ for G_r^k conditioned on targets $\vec{d}(G_r^1), \dots, \vec{d}(G_r^{k-1})$ is

$$\mathcal{S}(G_r^k | G_r^{k-1}, \vec{d}(G_r^{k-1}); \dots; G_r^1, \vec{d}(G_r^1)) = \text{span} \left\{ \sum_{i_1, \dots, i_r=1}^{|\mathcal{C}|} \bar{c}_{i_1 \dots i_r}^{(r-1)} F_r(G_r^k; i_1, \dots, i_r) \right\}_{\vec{c}^{(r-1)} \in \bigcap_{j=1}^{k-1} \mathcal{N}(G_r^j, \vec{d}(G_r^j))}, \quad (3.26)$$

where $\mathcal{N}(G_r^j, \vec{d}(G_r^j))$ is the solution space of Eq. (3.22).

3.4 Examples

Here, examples are provided to demonstrate how to use Eq. (3.22) to engineer higher-order effective Hamiltonian. A qubit network with all-to-all dipolar interactions under a high external field is considered, with the internal Hamiltonian given by

$$H_{\text{int}} = \sum_{i < j}^N B_{ij} (2\sigma_z^i \otimes \sigma_z^j - \sigma_x^i \otimes \sigma_x^j - \sigma_y^i \otimes \sigma_y^j). \quad (3.27)$$

The control Hamiltonian given by Eq. (3.8). For simplicity, the qubit offsets are not considered here. Thus, when calculating the conditions for engineering $\bar{H}^{(r-1)}$ (Eq. (3.22)), only the graphs without loops are considered. We use on-resonance control ($\omega_r(t) \equiv \omega_0$), so, the control parameters are the time-dependent amplitude and phase ($\omega_1(t)$ and $\phi(t)$) of the control field. The maximal amplitude ω_{max} of the Rabi field is 200 kHz, and the time duration of each step Δt is 1 μs .

3.4.1 2nd-order decoupling

We design an optimized sequence that decouples H_{int} (Eq. (3.27)) up to the 2nd-order. The sequence is shown in Fig. 3.3. The conditions for removing $\bar{H}^{(1)}$ and $\bar{H}^{(2)}$ are found by using Eq. (3.22) where $\vec{d}(G_r) \equiv \vec{0}$. There are respectively two and five graphs for $\bar{H}^{(1)}$ and $\bar{H}^{(2)}$ that need to be considered. Table 3.1 shows the ranks of $\oplus_{G_r} \mathbf{A}(G_r)$ (number of linearly independent constraints) for removing different higher-order terms. As comparison, we also show the ranks for other networks (Fig. 3.2). Spin pairs require least number of constraints to decouple while nearest-neighbor spin chains, square lattice and all-to-all spin ensemble have the same $\oplus_{G_r} \mathbf{A}(G_r)$ up to second order.

Table 3.1: Ranks of $\oplus_{G_r} \mathbf{A}(G_r)$ for decoupling the secular dipolar interaction for different network structures (Fig. 3.2).

	Pairs	NN chain	Square lattice	All-to-all ensemble
$\bar{H}^{(0)} = 0$	5	5	5	5
$\bar{H}^{(1)} = 0$	3	10	10	10
$\bar{H}^{(2)} = 0$	5	29	29	29
$\bar{H}^{(3)} = 0$	3	75	95	95

For simplicity, a symmetric sequence is used to remove odd-order corrections.

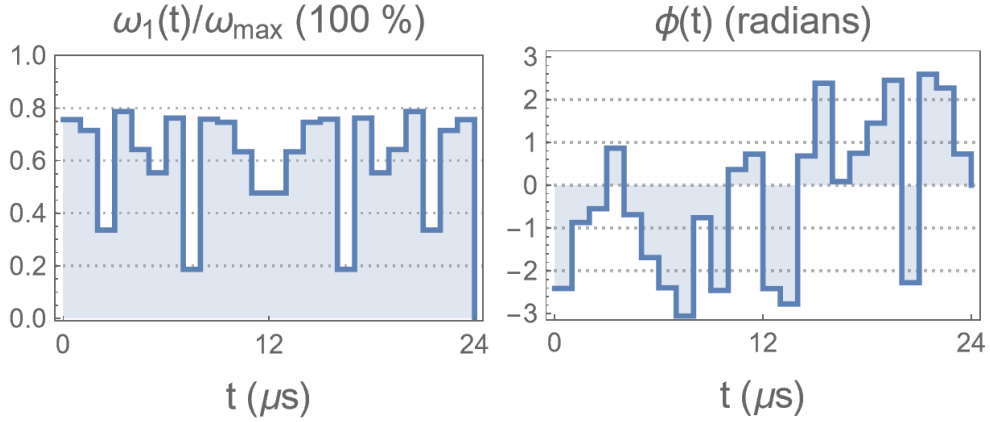


Figure 3.3: Control parameters of the sequence for 2nd-order decoupling. The horizontal and vertical axes are time (μs) and amplitude.

To test the performance, we perform a simulation with an ensemble of networks with

$N = 7$ with uniformly distributed random coupling strengths $19\text{kHz} \leq B_{ij} \leq 35\text{kHz}$. The result is an average of over 1000 realizations of the network. The total length of the sequence T_{seq} is $24\mu\text{s}$. The comparison with a first-order decoupling sequence is shown in Fig. 3.4.

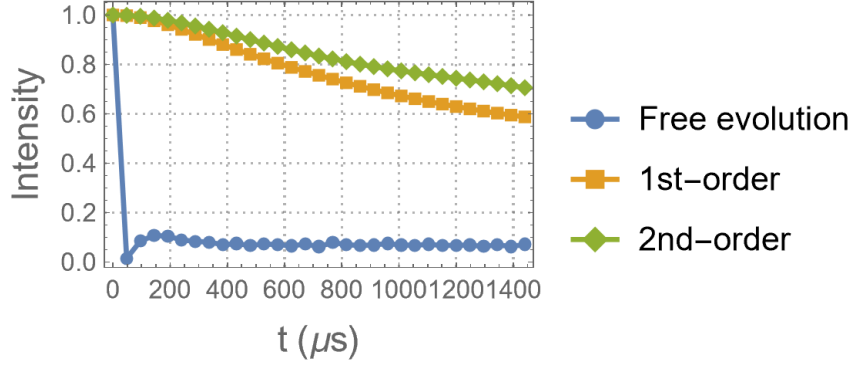


Figure 3.4: Simulated signals for different sequences. For reference, we also show the simulated free evolution signal.

3.4.2 Engineering nonzero first-order effective Hamiltonian

Here, we show how to engineer a three-body interaction in $\bar{H}^{(1)}$. There are two graphically inequivalent graphs $G_1^1 = \{12, 12\}$ and $G_1^2 = \{12, 23\}$ which give a 10-dimensional space for achievable $\bar{H}^{(1)}$:

$$\begin{aligned}
& \{xi + ix, yi + iy, zi + iz\}, \\
& \{yyx - 2yxy + xyy - zzx - 4zxx + xzz, \\
& \quad xxy - 2xyx + yxx - zzy - 4zyz + yzz, \\
& \quad xyz + xzy - yxz - yzx - zxy - zyx, \\
& \quad xxz + xzx + zxx - yyz - yzy - zyy, \\
& \quad -xxz - zxx - yyz - zyy + 2zzz, \\
& \quad zzx + xzz - yyx - xyy - 2yxy, \\
& \quad xxy + yxx + zzy + yzz - yyy, \\
& \quad xxy + yxx - zzy - yzz + 2yxy, \\
& \quad yyx + xyy + zzx + xzz - 2xxx, \\
& \quad xzx + yzy\}.
\end{aligned} \tag{3.28}$$

For simplicity, the shorthand x, y, z, i is used for the Pauli matrices and identity without the normalization constants. Using Eq. (3.22), the target for G_1^1 is

$$\mathbf{A}(G_1^1)\vec{c}^{(1)} = 0. \quad (3.29)$$

The conditional subspace $\mathcal{S}(G_1^2|G_1^1, \vec{d}(G_1^1) = 0)$ (Eq. (3.26)) is spanned by

$$\begin{aligned} &\{xxy + xyx + yxx - 2yyy + zzy + zyz + yzz, \\ &\quad xxz + xzx + zxx - 2zzz + yyz + yzy + zyy, \\ &\quad yyx + yxy + xyx - 3xxx + 2zzx + 2zxx + 2xzz, \\ &\quad -xxy - xyx - yxx + zzy + zyz + yzz, \\ &\quad xxz + xzx + zxx - yyz - yzy - zyy, \\ &\quad -3yyx - 3yxy - 3xyy + xxx + 2zzx + 2zxx + 2xzz, \\ &\quad -xyz - xzy - yxz - yzx - zxy - zyx\}. \end{aligned} \quad (3.30)$$

$\mathbf{A}(G_1^2|G_1^1, \vec{d}(G_1^1) = 0)$ can be calculated using Eq. (3.23). As an example, we choose to engineer the first basis Hamiltonian from $\mathcal{S}(G_1^2|G_1^1, \vec{d}(G_1^1) = 0)$

$$H_{\text{eff}}^{(1)} \propto xxy + xyx + yxx - 2yyy + zzy + zyz + yzz \quad (3.31)$$

which gives the targets

$$\begin{aligned} &\mathbf{A}(G_0^1)\vec{c}^{(0)} = 0, \\ &\mathbf{A}(G_1^1)\vec{c}^{(1)} = 0, \\ &\mathbf{A}(G_1^2|G_1^1, \vec{d}(G_1^1) = 0)\vec{c}^{(1)} = \vec{d}, \end{aligned} \quad (3.32)$$

where $G_0^1 = \{12\}$ and $\vec{d} = (a, 0, 0, 0, 0, 0, 0)$. The corresponding effective Hamiltonian (Eq. (3.17)) is

$$\begin{aligned} H_{\text{eff}}^{(r-1)}(N) &= \sum_{G' \sim G} \omega(G') H_{\text{eff}}^{(r-1)}(G) \\ &= \alpha \sum_{i < j < k}^N (B_{ij}B_{ik} + B_{ij}B_{jk} + B_{ik}B_{jk}) \\ &\quad (x^i x^j y^k + x^i y^j x^k + y^i x^j x^k \\ &\quad - 2y^i y^j y^k + z^i z^j y^k + z^i y^j z^k + y^i z^j z^k) \end{aligned} \quad (3.33)$$

α depends on the normalization constant, a and the normalization of $\mathbf{A}$. In practice, we can obtain α by calculating the overlap between the achieved effective Hamiltonian and

H_{target} . Instead of setting a to a fixed value, we require it to be larger than 0.25. No primary target is set as it can be corrected in post-processing [42]. The sequence found is shown in Fig. 3.5, which achieves

$$\vec{d} = (0.259, -0.007, -0.005, 0.001, -0.002, -0.004, 0.0044) \quad (3.34)$$

An echo is simulated to test the performance of the sequence with the same setup from

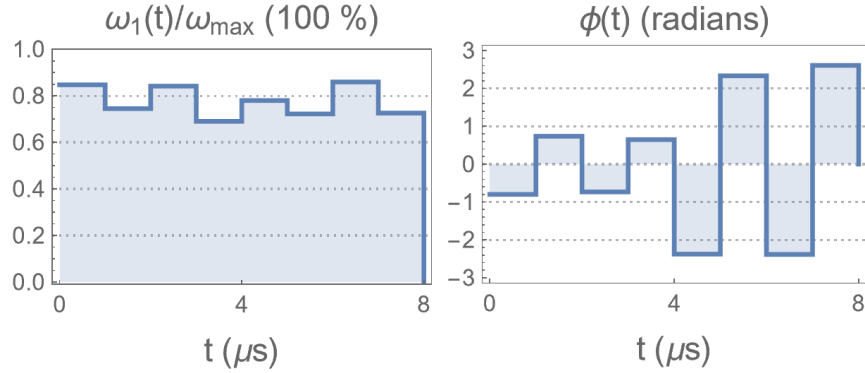


Figure 3.5: Control parameters for engineering a nonzero three-body average Hamiltonian at first order.

the last example. The unitary propagator of the simulation is

$$U_s(k_1, k_2) = \exp\left(ik_1 H_{\text{eff}}^{(1)}(N)t\right) U^{\otimes k_2}, \quad (3.35)$$

where

$$U = \mathcal{T} \exp\left(-i \int_0^t dt' \left[\sum_{k=1}^2 a_k(t') H_c^k + H_{\text{int}} \right]\right). \quad (3.36)$$

If $H_{\text{eff}} \approx H_{\text{eff}}^{(1)}$, an echo should appear at $k_1 = k_2$.

Using

$$\alpha = \frac{\text{Tr}\left[-\frac{\ln(U_c^\dagger U)}{it} H_{\text{eff}}^{(r-1)}(N)\right]}{\text{Tr}\left[H_{\text{eff}}^{(r-1)}(N) H_{\text{eff}}^{(r-1)}(N)\right]}, \quad (3.37)$$

we found that $\alpha = 0.123(5)$. Fig. 3.6 shows the simulated echoes. As can be seen, the echoes appear at $k_1 = k_2$ as expected, indicating that the desired dynamics has been engineered under the sequence. As k_2 increases, the intensities of the echoes decay, which is due to higher-order corrections and residual errors. More robust results can be achieved by considering third- and higher-order corrections.

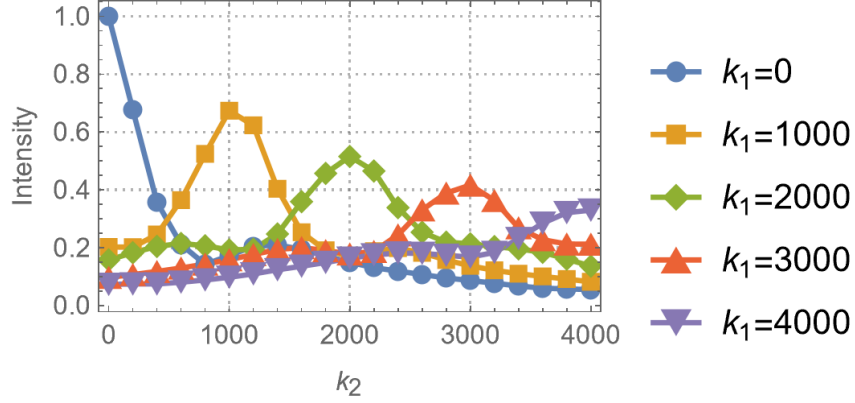


Figure 3.6: Simulated echoes of the sequence for 1st-order average Hamiltonian engineering.

3.5 Conclusion and Discussion

Based on the framework provided in the last chapter, we presented a general criterion (Eq. (3.7)) to engineer higher-order effective Hamiltonians. As an example, we showed how to engineer the internal Hamiltonian of a qubit network, two-body interactions, and offsets, with collective control. The topology of the network determines the space of achievable higher-order effective Hamiltonians and provides computable necessary and sufficient conditions for engineering arbitrary target Hamiltonians (Eq. (3.22)). Explicit examples show how to design optimized sequences for higher-order decoupling and a target three-body interaction in a qubit network with all-to-all interactions.

The method is symmetry-based and does not depend on the distributions of the internal parameters $\{\eta_{\text{int}}^i\}$. The statistical properties of the distributions, however, can complement the symmetry-based method and yield more efficient approximate solutions. The result will be published elsewhere.

The method provides a complete solution for engineering higher-order effective Hamiltonians in arbitrary systems, enabling the design of more precise and robust solutions, and realizing complex Hamiltonians, such as multi-body interactions, not available in the zeroth order.

Chapter 4

Efficiently Characterizing Quantum Information Flow, Loss and Recovery in the Central Spin System

¹ Understanding the flow, loss, and recovery of information between a system and its environment is essential for advancing quantum technologies. The central spin system serves as a useful model for a single qubit, offering valuable insights into how quantum systems can be manipulated and protected from decoherence. This work uses the stimulated echo experiment to track information flow between the central spin and its environment, providing a direct measure of the sensitivity of system/environment correlations to environmental dynamics. The extent of mixing and the growth of correlations are quantified through autocorrelation functions of the noise and environmental dynamics, which also enable the estimation of nested commutators between the system/environment and environmental Hamiltonians. Complementary decoupling experiments offer a straightforward measure of the strength of the system Hamiltonians. The approach is demonstrated experimentally on a spin system.

4.1 Introduction

Advances in quantum technologies are transforming fields such as communication [68], sensing [50], and physics simulation [66, 13], but decoherence remains an obstacle. In

¹This Chapter is largely based on [38].

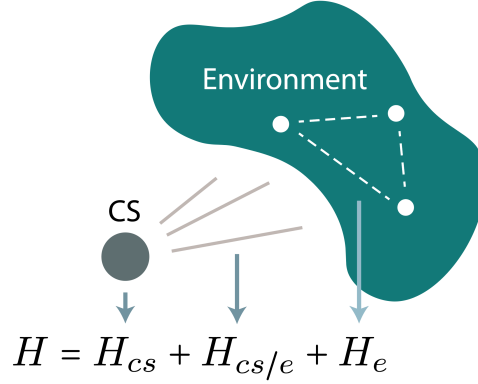


Figure 4.1: Schematic of a central spin system. The CS is a spin-1/2 particle and interacts with an environment consisting of spin-1/2 particles. The system can provide a complete model of an open quantum system of a qubit.

particular, the memory of the environment plays a key role in determining the optimal strategies for control [66, 79, 39, 218] and error correction [189], which is not captured by Markovian approximations such as Lindbladians [121]. Here, we present the theory and experimental implementation of a central spin (CS) model (Fig. 4.1) [49, 5] of qubit decoherence that characterizes the memory effects in the environment. The CS is a spin-1/2 interacting with an environment that can hold quantum information. The Hamiltonian of the system is

$$H = H_{cs} + H_{cs/e} + H_e, \quad (4.1)$$

where H_{cs} and H_e are the Hamiltonians of the CS and environment, respectively, and $H_{cs/e}$ describes the coupling of the CS and the environment. Note that there is no dissipation in this system; the total dynamics are unitary and thus perfectly preserve quantum information. The essence of this model is that an initially separable CS becomes correlated to the environment through the CS/environment coupling [143]. The coherence time of the CS, to first order, is governed by the rate at which these correlations form. The correlations are refocusable unless disrupted by the environment Hamiltonian, leading to a loss of refocusability with a characteristic time, T , determined by $[H_{cs/e}, H_e]$. As higher correlations build, the quantum information shared among larger correlated states becomes more susceptible to H_e . The loss of refocusability has an initial Zeno region [89], marked by a Gaussian signature in the initial decay of the autocorrelation function, that represents the short-term memory of the environment. This region, where the quantum information is well preserved, is important to the optimal design of quantum gates. On the contrary, the environment may hold part of the quantum information and have a long-term memory,

which can be much harder to detect. This long-term memory can lead to bias in noise on the CS and be harmful to error correction.

There are various ways to characterize the noise from the environment. For example, dynamical decoupling using filter functions can measure the power spectral density [35, 150, 15, 3]. This has also been measured by observing the relaxation rate in the rotating frame ($T_{1\rho}$) [216]. These methods typically rely on a stochastic description of the noise and take the Gauss-Markov approximation of the environmental dynamics. There have been generalizations to measuring polyspectra of non-Gaussian noise [183, 145]. However, these still rely on the truncation of higher-order terms and a classical description of the noise. An experimentally viable method to fully capture the quantum nature of the noise remains elusive.

Previously, M. Niknam, *et al.* [143, 144] introduced an experimental method for directly detecting the growth of multi-spin correlations in a CS system and examining decoherence time. They defined a metric for quantifying information flow and entropy distribution in these systems.

This work provides an efficient experimental method, using the stimulation echo (STE) [83, 188] and decoupling experiments, to track the short- and long-term memory of the environment and to measure changes of the local field and correlation growth. No approximation is taken so that the complete quantum nature of the system is preserved. The memory effect and information flow from the CS to the environment are quantified through autocorrelation functions, which are measured without any truncation. Additionally, nested commutators between $H_{cs/e}$ and H_e are estimated using Bayesian analysis [91, 64], which characterize the growth of memory on different levels.

4.2 Theory

Since our interest is to characterize the buildup of the CS/environment correlations, the initial state of a fiducial state of the CS $|0\rangle\langle 0|$ and a fully mixed environment is considered. $H_{cs/e}$ leads to the correlation of the CS and the environment which is well described by a local field model [19]. The local field is given by the eigenvalues of $H_{cs/e}$ and depends on the environment's state, which determines the structure of the spectrum of the CS. The growth of correlation under the local field has an initial Gaussian signature and is refocusable by a simple echoing of the CS.

Environmental mixing can alter the state of the environment. For instance, if H_e takes

the form of a "flip-flop" Hamiltonian $\sigma_+ \otimes \sigma_- + \sigma_- \otimes \sigma_+$, then

$$|01\rangle \xrightarrow{\sigma_+ \otimes \sigma_- + \sigma_- \otimes \sigma_+} |10\rangle. \quad (4.2)$$

Such changes in the environment's state result in a corresponding change in the local field seen by the CS.

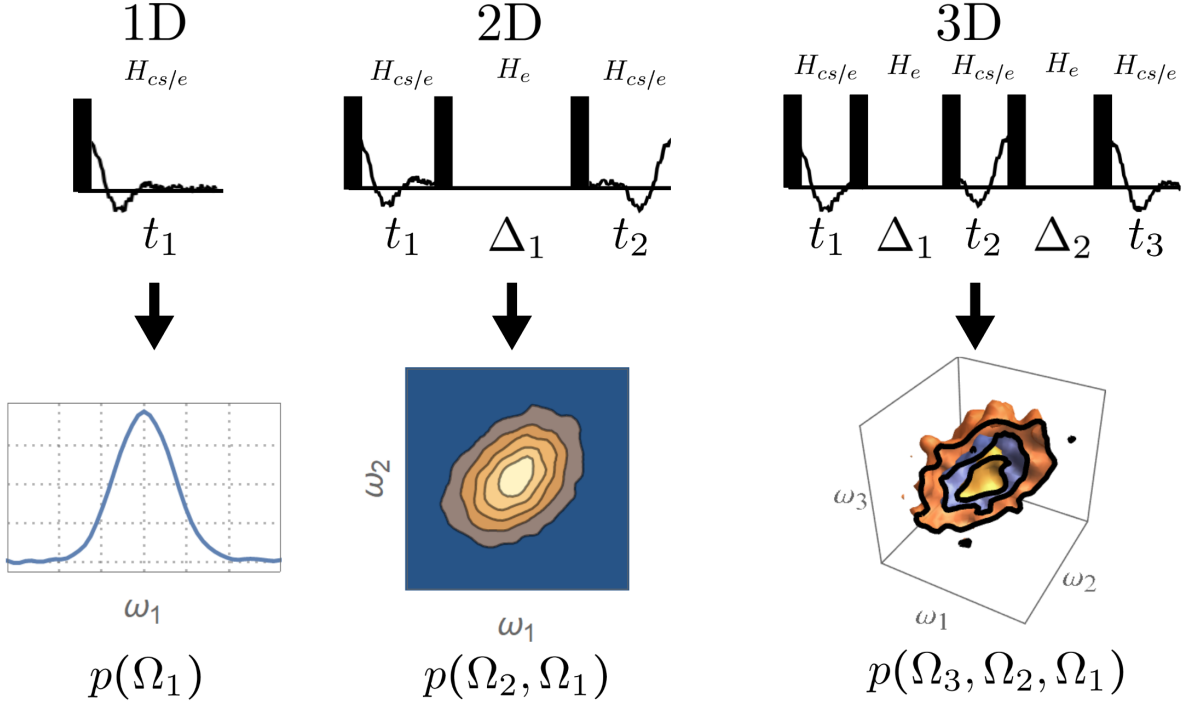


Figure 4.2: Measuring multi-time joint distributions of local fields using n D experiments. The 1D experiment yields the distribution of local field when there is no mixing. The 2D experiment yields the two-time joint distribution of local fields before and after a period of mixing, which measures the change of local field. n D experiments ($n \geq 3$) can then produce the n -time joint distributions that measure more subtle properties of the memory of the environment.

The effect of the change of the local field can be described by joint distributions $p(\Omega_n, \dots, \Omega_1)$ where Ω_i is a random variable representing the local field at time t_i (Fig. 4.2). Δ 's are the mixing time. The joint distributions can be measured by an n -dimensional

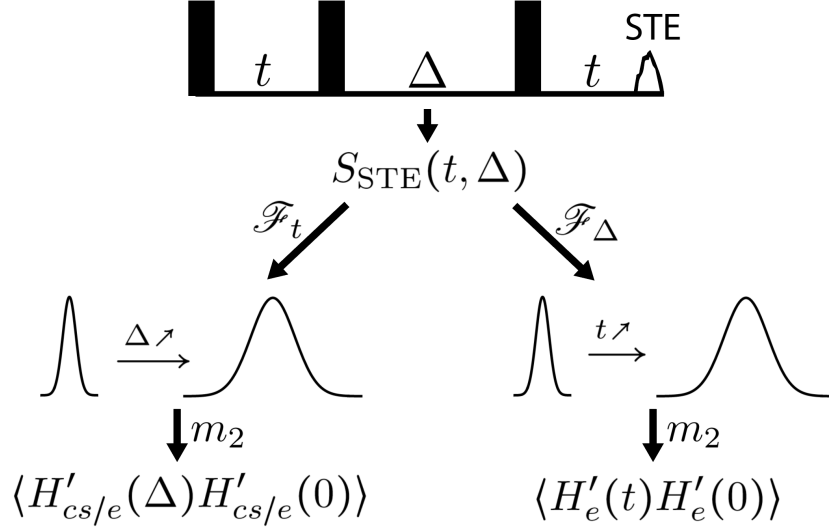


Figure 4.3: The second moments of the spectra of the STE experiments by Fourier transform with respect to either t or Δ give the autocorrelation functions $\langle H'_{cs/e}(\Delta)H'_{cs/e}(0) \rangle$ and $\langle H'_e(t)H'_e(0) \rangle$, allowing efficient measurement of change of the local field and correlation growth.

(n D) experiment which corresponds to the following unitary propagator

$$U_{nD} = e^{-iH_{cs/e}t_n} e^{-iH_e\Delta_{n-1}} e^{-iH_{cs/e}t_{n-1}} \dots e^{-iH_e\Delta_1} e^{-iH_{cs/e}t_1}. \quad (4.3)$$

It separates the effect of the CS/environment interaction and the environmental dynamics so that the full dynamics under $H_{cs/e}$ or H_e during each time period can be sampled by continuously varying t_i or Δ_i . A joint distribution $p(\Omega_n, \dots, \Omega_1)$ can then be obtained by an n D Fourier transform with respect to t_i . In the quantum picture, the variation in the local field corresponds to $H'_{cs/e}(\Delta) = e^{iH_e\Delta} H_{cs/e} e^{-iH_e\Delta}$ which represents $H_{cs/e}$ in the interaction frame of H_e . The n D experiment effectively compares $H'_{s/e}(\Delta)$ at different times and measures its time dependence caused by the environmental dynamics H_e .

The n D experiments provide a more complete picture of the growing correlations, but we do not necessarily need that level of detail. A more efficient approach is to measure the memory of the environment. If the interest is in the multi-time correlation functions

$$\text{Corr}(\Omega_n, \dots, \Omega_1) = \langle H'_{cs/e}(\Delta_{n-1} + \dots + \Delta_1) \dots H'_{cs/e}(0) \rangle, \quad (4.4)$$

a more efficient approach is to measure the memory of the environment. Here, it is sufficient to implement reduced versions of the n D experiments that are typical of lower dimensions

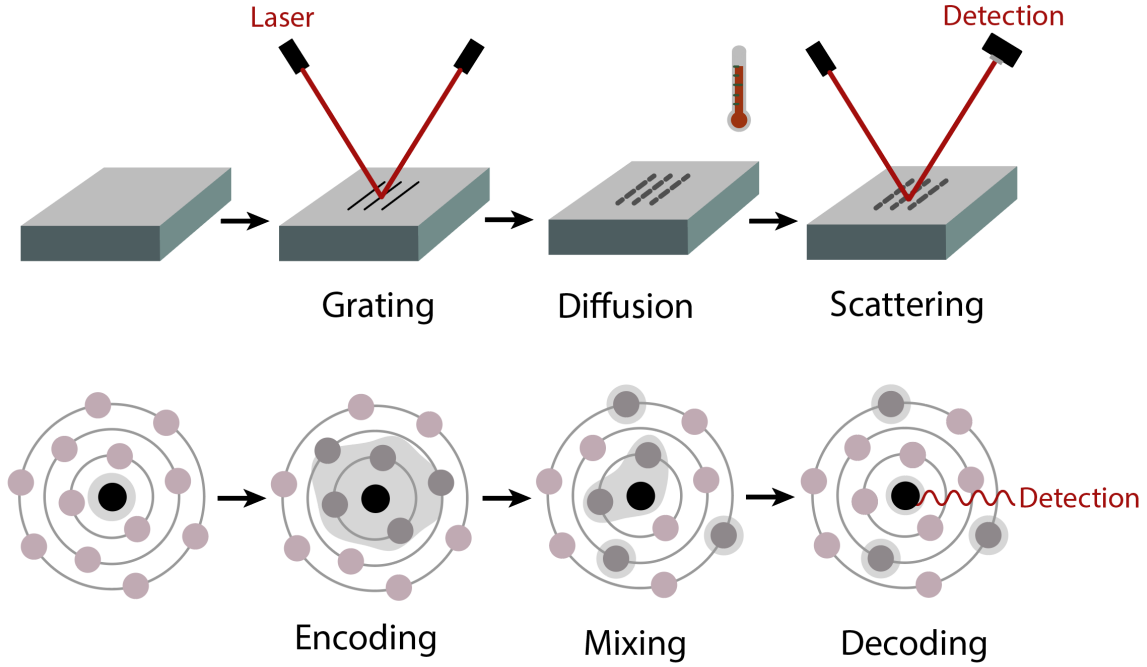


Figure 4.4: Schematic of the forced Rayleigh scattering (top) and stimulated echo (bottom) experiments. Like the FRS experiment, the STE experiment consists of three steps: encoding, mixing, and decoding, which correspond to grating, diffusion, and scattering in the FRS experiment. The quantum information initially resides at the CS. By comparing it with the echo after decoding, one can learn about mixing.

[170]. According to the central slice theorem, the spectra of the reduced experiments correspond to projections of the n D spectra, which can extract useful information from the spectra.

Here, we show how to measure the autocorrelation function $\langle H'_{cs/e}(\Delta)H'_{cs/e}(0) \rangle$ through the STE experiment. The experiment is a reduced version of the 2D experiment where t_1 and t_2 are set equal and are incremented together. To better understand how the STE experiment works, we can compare the change of the local field to diffusion in materials. Indeed, environmental dynamics has been modeled by spin diffusion [18, 220], inspired by thermal diffusion. A classic experimental method for measuring thermal diffusion in materials is forced Rayleigh scattering (FRS) [156]. The basic procedure is summarized below:

1. **Creation of spatial grating:** A structured spatial grating is created within the

material by intersecting two coherent laser beams.

2. **Thermal diffusion and grating decay:** The material is then allowed to undergo thermal diffusion which causes the grating to blur.
3. **Monitoring diffusion via scattered light:** A monitoring laser beam (with the same frequency as the initial grating-forming lasers) is applied to the material. The intensity of the scattered light from this probe beam is measured, which indicates the state of the grating. The decay in the scattered intensity corresponds to the extent of diffusion, reflecting the material's diffusion properties.

The STE experiment (Fig. 4.4) consists of three similar steps:

1. **Creation of CS/environment correlation:** The CS interacts with the environment via $H_{cs/e}$, forming a correlation with it.
2. **Environment mixing and correlation decay:** The CS undergoes a mixing period under H_e , which perturbs the system/environment correlation.
3. **Measuring mixing through echo:** The CS evolves under $-H_{cs/e}$ to form an echo, and the echo intensity is measured to reflect the extent of mixing and reveal properties of the environment.

Here, the correlation plays a similar role to that of the grating, which is used to track mixing in the environment. Then the system is let to evolve backward under the same Hamiltonian $H_{cs/e}$, similar to the monitoring laser beam, to give a returned signal which is then used to measure the extent of mixing. By separating the dynamics under $H_{cs/e}$ and H_e , it allows us to control the extent of dynamics under $H_{cs/e}$ and H_e and achieve a better estimation of parameters in the system.

To understand what and how information can be extracted from the STE experiment, consider its unitary propagator

$$U_{\text{STE}} = \exp(iH_{cs/e}t) \exp(-iH_e\Delta) \exp(-iH_{cs/e}t), \quad (4.5)$$

the signal of the experiment is

$$S_{\text{STE}}(t, \Delta) = \text{Tr}(U_{\text{STE}}\rho(0)U_{\text{STE}}^\dagger\rho_{\text{ref}}), \quad (4.6)$$

where $\rho(0)$ is the density matrix of the initial state of the CS system and ρ_{ref} is the observable corresponding to a measurement on the CS.

Intuitively, the autocorrelation function $\langle H'_{cs/e}(\Delta)H'_{cs/e}(0) \rangle$ is directly related to the time dependence of S_{STE} on Δ : the more mixing there is, the less correlated $H'_{cs/e}(\Delta)$ and $H'_{cs/e}(0)$ are and the weaker the echo is. Indeed, $\langle H'_{cs/e}(\Delta)H'_{cs/e}(0) \rangle$ can be calculated from the second moment of the spectrum obtained by a Fourier transform of $S_{\text{STE}}(t, \Delta)$ with respect to t . It measures the extent of mixing over any Δ . Fundamentally, the change of local field arises because $[H_{cs/e}, H_e] \neq 0$, and this information is encoded in the Δ -dependence of $\langle H'_{cs/e}(\Delta)H'_{cs/e}(0) \rangle$. In fact, the autocorrelation function can be expanded with respect to Δ as

$$\langle H'_{cs/e}(\Delta)H'_{cs/e}(0) \rangle = -\frac{\Delta^2}{2!} \langle [H_e, [H_e, H_{cs/e}]]H_{cs/e} \rangle + \frac{\Delta^4}{4!} \langle [H_e, [H_e, [H_e, [H_e, H_{cs/e}]]]]H_{cs/e} \rangle + \dots \quad (4.7)$$

Only even-order terms appear because S_{STE} and $\langle H'_{cs/e}(\Delta)H'_{cs/e}(0) \rangle$ are even functions of Δ ; changing the sign of Δ does not change S_{STE} . Therefore, the non-commutativity of $H_{cs/e}$ and H_e can be measured through estimating the functions $\langle [H_e, [H_e, H_{cs/e}]]H_{cs/e} \rangle, \dots$, providing an absolute measure of mixing that is independent of Δ .

The STE experiment also provides a measure of information flow between the CS and the environment. Note that the decay rate of S_{STE} with respect to Δ for fixed t depends on the amount of correlation between the CS and the environment: the more correlation there is, the more sensitive the system is to H_e and the faster the decay is. Therefore, the decay rate with respect to Δ can be used to monitor the amount of correlation between the CS and the environment. This can also be measured through the spectra obtained by a Fourier transform of $S_{\text{STE}}(t, \Delta)$ with respect to Δ . The second moment of the spectra is directly related to this sensitivity and gives the autocorrelation function $\langle H'_e(t)H'_e(0) \rangle$, where $H'_e(t) = \exp(iH_{cs/e}t)H_e \exp(-iH_{cs/e}t)$ represents H_e in the interaction frame of $H_{cs/e}$. $H'_e(t)$ can be interpreted as the CS's perception of the environment. As correlations build, the CS's view of the environment shifts, causing $\langle H'_e(t)H'_e(0) \rangle$ to decay.

The time dependence of $\langle H'_e(t)H'_e(0) \rangle$ can also be used to estimate non-commutativity of $H_{cs/e}$ and H_e . Particularly, we can measure terms such as $\langle [H_{cs/e}, [H_{cs/e}, H_e]]H_e \rangle$ and higher-order nested commutators. These terms provide an absolute measure of the rate of information flow, independent of t .

Since the STE measures correlation and mixing through the commutation relation between H_e and $H_{cs/e}$, it captures only the components that lead to a nonvanishing $[H_e, H_{cs/e}]$. There can always be a component with a vanishing commutator that remains unmeasured. For $\langle H'_{cs/e}(\Delta)H'_{cs/e}(0) \rangle$, this unmeasured component corresponds to the part of the environment that does not disrupt the correlation. Conversely, for $\langle H'_e(t)H'_e(0) \rangle$, it corresponds to the portion of the correlation that remains unaffected by the environmental dynamics.

A direct measurement of $H_{cs/e}$ and H_e is essential as it establishes a baseline for the intrinsic dynamics of the environment. $H_{cs/e}$ can be determined by analyzing the evolution of the CS under $H_{cs/e}$ while decoupling H_e as has been achieved in [143]; a simple spectrum then characterizes the CS/environment coupling. Here, we are mostly interested in measuring $\|H_e\|$. For this purpose, the decoupling experiment is employed to assess these dynamics. The experiment explores how the loss of refocusability of the echo caused by H_e changes in the presence of continuous RF irradiation on the environment. The unitary propagator of the experiment is

$$U_{\text{dec}} = \exp(-i(-H_{cs/e} + H_e + \omega_{\text{dec}}J_x)t) \exp(-i(H_{cs/e} + H_e + \omega_{\text{dec}}J_x)t), \quad (4.8)$$

where J_x is the x component of the total angular momentum operator of the environment and ω_{dec} is the frequency of the RF irradiation. The signal of the experiment is

$$S_{\text{dec}}(t) = \text{Tr}(U_{\text{dec}}\rho(0)U_{\text{dec}}^\dagger\rho_{\text{ref}}). \quad (4.9)$$

When ω_{dec} is small, H_e disrupts the CS/environment correlation, leading to fast decay of S_{dec} with respect to t . When $\omega_{\text{dec}} > \|H_e\|$, the RF irradiation becomes dominant, effectively suppressing H_e and decoupling the CS from the environment [81]; a rapid slowing in the decay rate can be observed. Therefore, the frequency ω_{dec} at which the transition in the decay rate happens provides a quantitative measure of $\|H_e\|$.

4.3 Method

A spin system is used to experimentally demonstrate the method. The experiments are implemented with a high-field (7T) solid-state NMR system at room temperature using a Bruker spectrometer. A powder sample of triphenylphosphine is used to simulate the CS system. The phosphorus spins and the proton spins serve as the CS and environment spins respectively. The total Hamiltonian of the system is

$$H = \delta_{\text{CS}}\sigma_z^{\text{CS}} + \sigma_z^{\text{CS}} \otimes \sum_{i=1}^N \omega_i \sigma_z^i + \sum_{i=1}^N \delta_i \sigma_z^i + \sum_{i < j}^N \omega_{ij} (2\sigma_z^i \otimes \sigma_z^j - \sigma_x^i \otimes \sigma_x^j - \sigma_y^i \otimes \sigma_y^j), \quad (4.10)$$

where N is the number of hydrogen spins in the environment, δ_{CS} and δ_i are chemical shift anisotropy of the phosphorus and hydrogen spins, and ω_i and ω_{ij} are coupling strengths of P-H hetero- and H-H homo-nuclear dipolar interactions.

The pulse sequences of the STE and decoupling experiments are shown in Fig. 4.6. Correct phase cycling of the $\pi/2$ pulses in the STE experiment is chosen to select the

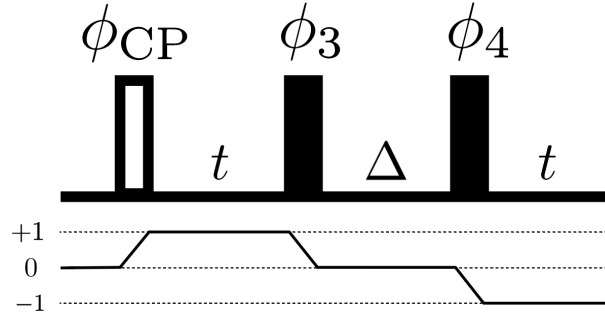


Figure 4.5: Coherence pathway for STE. The empty rectangle represents the effective excitation pulse corresponding to the cross-polarization. By combining experiments with carefully chosen phases, one can achieve the desired coherence pathway shown here which gives the unitary propagator of the STE experiment.

desired coherence pathway [20, 177] (Fig. 4.5). Additionally, EXORCYCLE [95] is applied to the π pulse in the decoupling experiments. Chemical shift anisotropy of phosphorus spins is refocused by the pulse sequences. That of the proton spins ($\sim 100\text{Hz}$) is insignificant and ignored throughout the experiments. The Hamiltonians $H_{cs/e}$ and H_e are then given by the P-H heteronuclear dipolar interactions and the H-H homonuclear dipolar interactions respectively. The sample is doped with a relaxation agent (chromium(iii) acetylacetonate) to reduce its T_1 . The time scale of the experiments ($< 10\text{ms}$) is much smaller than the T_1 of the phosphorus spins (55s) and the hydrogen spins (1s). Cross-polarization [84, 154] is used to transfer polarization from the proton spins to the phosphorus spins to increase the signal-to-noise ratio.

During the encode and decode periods in the STE experiments, the symmetric magic-echo train 8 (SME8) [27] is used to turn off the environmental dynamics H_e while keeping the system/environment interaction $H_{cs/e}$ with a scaling factor of 0.342 and along the same direction. This, together with the coherence pathway selection, achieves the desired unitary propagator Eq. (4.5).

Bayesian inference is used to estimate the autocorrelation functions and non-commutativity of $H_{cs/e}$ and H_e using their corresponding posterior distributions. In addition, a Bayesian model selection [205] is used to evaluate the goodness of different levels of models in order to estimate the commutation relations.

4.4 Results

4.4.1 Measure change of local field

STE experiments (Fig. 4.6 (a)) are implemented on the CS system for every combination of $t = 48, 96, \dots, 1200 \mu s$ and $\Delta = 0, 7.5, \dots, 150 \mu s$ to collect the signals $S_{\text{STE}}(t, \Delta)$ (Eq. (4.6)). A Fourier transform on S_{STE} with respect to t yields the spectra $\tilde{S}_{\text{STE}}(\omega, \Delta)$. The time-domain data S_{STE} and the corresponding spectra $\tilde{S}_{\text{STE}}$ are shown in Fig. 4.7 (a). As expected, the decay of S_{STE} with respect to t for a given Δ becomes faster as Δ increases, indicating more loss of correlation. This is also shown as a broadening of the spectra.

The second moments of the spectra are calculated to obtain the autocorrelation function $\langle H'_{cs/e}(\Delta) H'_{cs/e}(0) \rangle$. To reduce the uncertainty, the calculation is restricted to the region of the peaks ($[-3.6, 3.6]$ kHz). Note that the frequencies used here are all angular frequencies. The second moments are then evaluated using a Bayesian approach to incorporate the uncertainty (Fig. 4.8 (a)). As seen in the result, the second moment increases (corresponding to decay of the autocorrelation function) as Δ increases, indicating an increase of change of local field. However, due to imperfect refocusing, the signature of leading-order (second-order) dependence on Δ cannot be observed from the result.

To better explore the features predicted by Eq. (4.7), the linewidths of the spectra are measured and plotted (Fig. 4.8 (b)). The linewidths are measured with less uncertainty and show a clear Gaussian signature for small Δ . Then the second moments derived from the measured linewidths by taking a Gaussian approximation ($m_2 \approx \text{FWHM}^2/2.355^2$) are calculated and plotted in Fig. 4.8 (c). The derived second moments also show a Gaussian signature as predicted by Eq. (4.7) and has the same overall timescale as the measured second moments. However, after the initial Gaussian decay, there is a transition (at $\sim 80 \mu s$) to a slower decay, suggesting two components of the environment that correspond to different mixing rates.

In order to measure the non-commutativity $\langle [H_e, [H_e, H_{cs/e}]] H_{cs/e} \rangle, \dots$, Bayesian model learning (a combination of model selection and parameter estimation) is applied to the derived second moments according to the model given by Eq. (4.7). To better capture the decay of the autocorrelation function, it is remodeled as a mixture of Gaussians:

$$\langle H'_{cs/e}(\Delta) H'_{cs/e}(0) \rangle = \sum_{l=1}^L p_l \exp(-\Delta^2/\lambda_l), \quad (4.11)$$

which is equivalent to Eq. (4.7) as $L \rightarrow \infty$. The procedure for estimating the parameters is as follows: First, the most probable level of the model L is selected according to information

criteria calculated by maximizing the likelihood functions; then, a Bayesian analysis is applied to the selected model to obtain the posterior distributions of the parameters p_l, λ_l ; finally, the posterior distributions of the functions $\langle [H_e, [H_e, H_{cs/e}]] H_{cs/e} \rangle, \dots$ are obtained from the aforementioned posterior distributions to give the desired estimations.

The selected model is of $L = 2$, confirming the conjecture that there are likely two different components of the environment. The obtained posterior distributions of the first three terms in $\langle [H_e, [H_e, H_{cs/e}]] H_{cs/e} \rangle, \dots$ are plotted in Fig. 4.9 (top row). The corresponding estimations are summarized in Table 4.1. The estimations give an absolute measure of the effect of mixing seen by the CS. Reasonably, higher-order nested commutators are estimated with increased uncertainty. This can be improved by collecting more data for large Δ since the higher-order effects are more important at the tail.

Table 4.1: Posterior mean estimations of the commutation functions (defined in the caption of Fig. 4.9). 95% intervals are used to obtain the error bars.

C_2	C_4	C_6
$2.02^{+0.25}_{-0.22} \times 10^{-4} (\text{MHz}^2)$	$-3.1^{+0.7}_{-0.9} \times 10^{-7} (\text{MHz}^4)$	$8.4^{+5}_{-2.9} \times 10^{-10} (\text{MHz}^6)$
D_2	D_4	D_6
$0.9^{+0.7}_{-0.35} (\text{kHz}^2)$	$-5.3^{+3.5}_{-10} \times 10^{-6} (\text{kHz}^4)$	$0.6^{+2}_{-0.5} \times 10^{-10} (\text{kHz}^6)$

4.4.2 Measure change of sensitivity

A similar analysis is applied to the STE experiment data to quantify information flow. Fig. 4.7 (b) presents the time-domain data and spectra where the roles of Δ and t are exchanged. The data show an increased decay rate (spectral broadening) for larger t , reflecting enhanced sensitivity to environmental dynamics due to stronger system-environment correlations.

The linewidths of the spectra are distorted by the Gibbs ringing due to truncation of the time-domain data. Therefore, instead of relying on direct measurement of the linewidths, the decay rates are estimated by fitting the time-domain data to an exponential function. Then the decay rate is used to estimate the corresponding second moments (Fig. 4.10). The optimal model, determined by Bayesian model selection, has $L = 1$. Then the commutation relations are estimated by fitting the data to the model. The posterior distributions are shown in Fig. 4.9 (bottom row), with estimates provided in Table 4.1. Due to the high uncertainty of the data at the tail ($t > 700\mu\text{s}$), estimates of higher-order commutation relations are significantly worse.

4.4.3 Measure strength of H_e

To measure the absolute strength of the mixing $\|H_e\|$, the decoupling experiments (Fig. 4.6 (b)) are implemented on the CS system for $t = 48, 96, \dots, 1200 \mu s$ and different strengths of the RF irradiation. The time-domain data and the corresponding spectra are plotted in Fig. 4.7 (c). The linewidths of the spectra are calculated and plotted in Fig. 4.11 as a function of ω_{dec} . The result shows that the decay slows as the ω_{dec} increases, implying that a larger portion of the mixing is suppressed. The Gaussian signature at small ω_{dec} corresponds to the regime where ω_{dec} becomes larger than $\|H_{cs/e}\|$ and the information flow is being suppressed. It can be seen that the linewidth starts to saturate at around $\omega_{\text{dec}} = 30 \text{ kHz}$, indicating the condition that $\omega_{\text{dec}} \gg \|H_e\|$ is satisfied for the majority of the environment. Indeed, the strongest interaction between environment spins, calculated according to the crystal structure of the molecule [106], is around 30 kHz.

4.5 Discussion

The measured autocorrelation function $\langle H'_{cs/e}(\Delta) H'_{cs/e}(0) \rangle$ reveals an overall characteristic time $T \sim 90 \mu s$ for environment mixing (Fig. 4.8 (a)). However, its structure deviates from an exponential form, indicating memory effects in the environment. Initially, a Zeno regime $\Delta \sim [0, 30) \mu s$ is observed, corresponding to an initial resistance to environmental perturbation due to the unitary dynamics of the environment. The duration of the period is approximately equal to $1/\|H_e\|$ as measured by the decoupling experiment. Then, this period is followed by a rapid change in the local field, characterized by a Gaussian signature, suggesting that information is being rapidly scrambled in the environment. The environment size is sufficiently large to ensure that this scrambling as seen by the CS appears uncorrelated. Around $\Delta \sim 70 \mu s$, a transition to a slower change in the local field occurs. Model selection further confirms the presence of two Gaussian components: one fast and one slow. This indicates the existence of a long-lived state within the system/environment correlation that is less sensitive to environmental dynamics. The rapid mixing corresponds to dynamics in a finite-sized part of the environment. Once mixing in this region saturates, information flow extends to a larger part of the environment, leading to additional information loss. However, this extended part contributes more slowly, likely due to weaker connectivity with the primary environment. This observation is further confirmed through the estimation of non-commutativity, suggesting a structured environment with both long and short memory. A detailed understanding of this memory structure will undoubtedly be crucial for comprehending the environment's dynamics and for designing improved control and error correction protocols for the system.

Mathematically, the information flow in a many-body system can be quantified by the information flux [53, 52]. For information flow between the CS and the environment, the information flux is closely related to the correlation measure given by the Rènyi entropies [144]. The autocorrelation function $\langle H'_e(t)H'_e(0) \rangle$ and the nested commutation relations $D_2, D_4, \dots$ aim to provide an efficient measure of this correlation. As the results show, the sensitivity to H_e , as measured by the second moments of the spectra with respect to Δ in STE experiments reveals a Zeno regime around $t \sim [0, 250\mu s)$. Similarly, the non-exponential dynamics of the autocorrelation function shows that the system/environment interaction is highly non-Markovian. According to Niknam et al. [143], first-order coherence peaks around $t \sim 200\mu s$ (adjusting for scaling differences between MREV8 and SME8 sequences), suggesting that rapid increase of sensitivity to perturbation mainly arises from higher-order coherence terms. Their results show that coherence growth saturates at $t \sim 590\mu s$. However, in our measurements, sensitivity continues to increase beyond this time. This may imply that highly sensitive, multi-spin correlations continue to develop, even if this does not directly enhance coherence, as most multi-spin correlations are still of lower-order coherence. Thus, STE experiments may be capable of revealing the growth of higher-order correlations.

4.6 Conclusion

In conclusion, an efficient experimental approach combining stimulated echo (STE) and decoupling experiments provides quantitative measures of information flow and mixing in a central spin system. The STE experiment distinguishes the effects of system/environment interaction from environment mixing, allowing focused study on either information flow or mixing. The second moments of STE spectra yield autocorrelation functions that capture changes in local fields and sensitivity, offering detailed quantitative insights into the landscape of information flow and mixing. Through Bayesian analysis, the STE experiment directly measures the non-commutativity between $H_{cs/e}$ and H_e , providing absolute measures of the rates of correlation growth and mixing across different levels. The decoupling experiment, in turn, offers a straightforward and effective method to directly probe the strength of H_e . Together, the STE and decoupling experiments deliver a comprehensive metric of environment mixing. The results reveal a complex structure in the environmental dynamics, with both short-term and long-term memory effects in the system, highlighting the rich dynamics of the central spin system and the importance of thoroughly understanding them. While the method is demonstrated using a spin system, it is applicable to any qubit-environment interacting system.

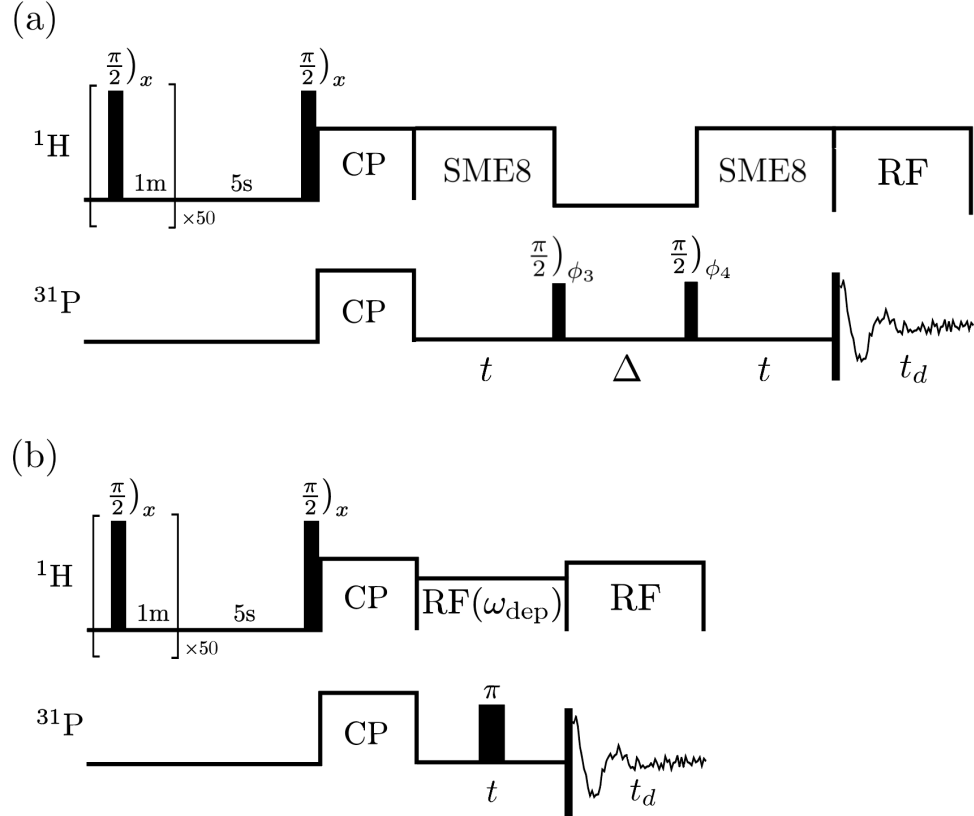


Figure 4.6: The pulse sequences for the (a) STE and the (b) decoupling experiments. CP and RF refer to cross-polarization and RF irradiation. Saturation with $\pi/2$ pulses is applied to the environment spins prior to the start of each experiment to prevent correlation between experiments. Spin temperature alteration [182] is incorporated in the phase cycling to remove artifacts from the cross-polarization. During t_d , a continuous weak measurement is performed on the CS. A strong continuous wave is applied to the environment during this period to decouple the CS from the environment. Then an integration is taken over t_d to give a complex signal point.

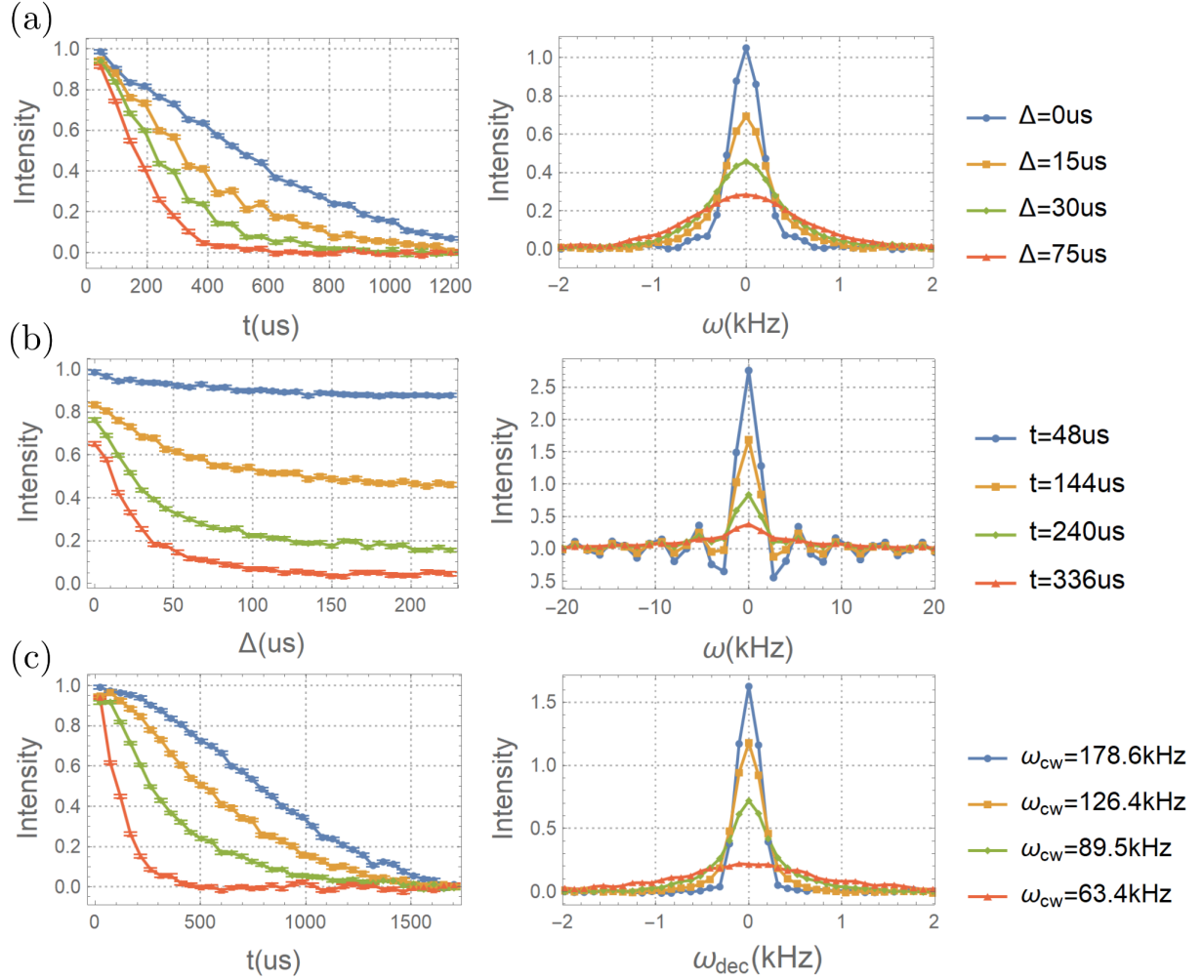


Figure 4.7: Time-domain data (left) and spectra (right) of the STE experiments for (a) different Δ , (b) different t , and (c) those of the decoupling experiments for different ω_{cs} . Here, only real parts of the data are shown. (a) and (b) are obtained from the STE experiments (Fig. 4.6 (a)), while (c) is from the decoupling experiments (Fig. 4.6 (b)).

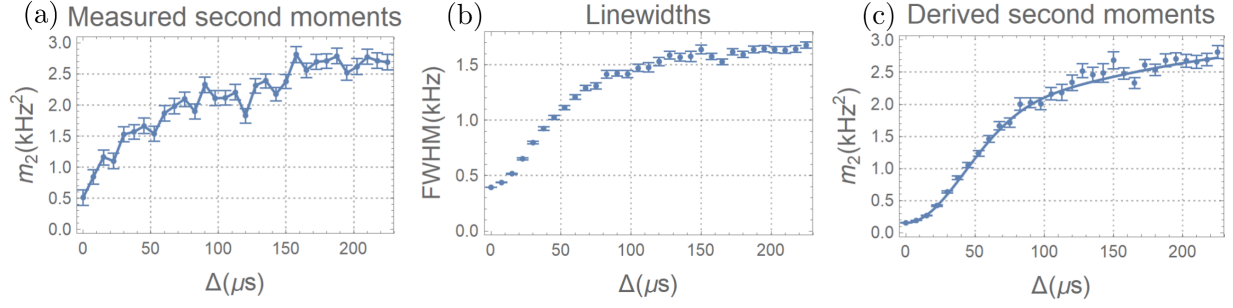


Figure 4.8: Measure of (a) second moments, (b) linewidths and (c) second moments derived from the linewidth of the STE experiments as functions of Δ . The spectra are obtained by Fourier transform with respect to t . The data are obtained from their corresponding posterior predictive distributions. The error bars correspond to the 1/4 and 3/4 quantiles. The solid line represents the best fit using the posterior mean for the optimal model with $L = 2$ (Eq. (4.11)).

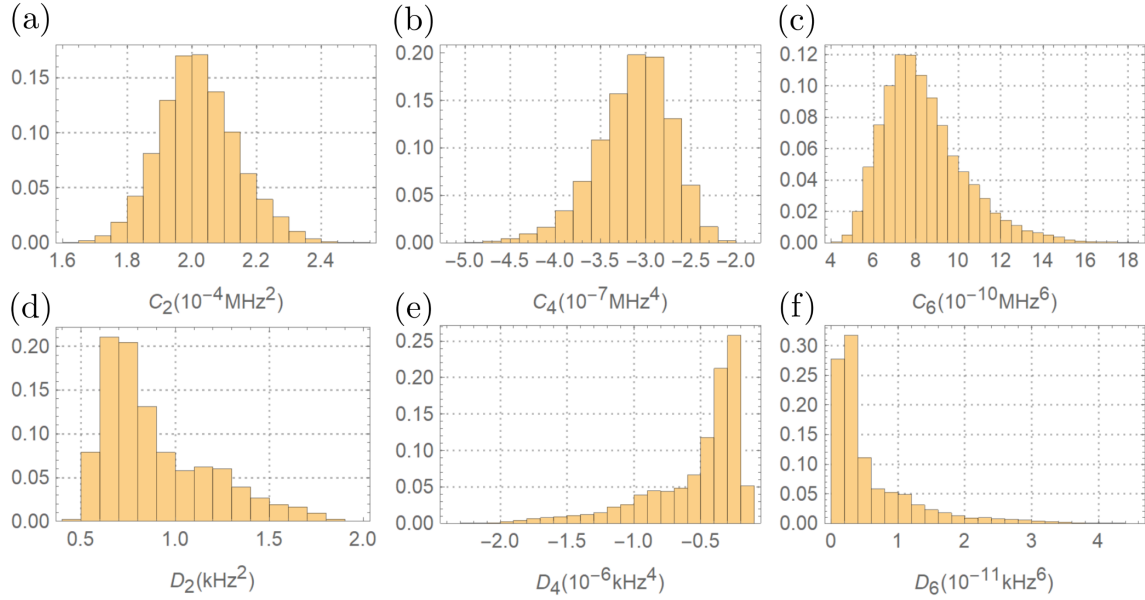


Figure 4.9: Probability histogram plots of the posterior distributions of (a) $C_2 = \langle [H_e, [H_e, H_{cs/e}]] H_{cs/e} \rangle$, (b) $C_4 = \langle [H_e, [H_e, [H_e, [H_e, H_{cs/e}]]]] H_{cs/e} \rangle$, (c) $C_6 = \langle [H_e, [H_e, [H_e, [H_e, [H_e, [H_e, H_{cs/e}]]]]]] H_{cs/e} \rangle$, (d) $D_2 = \langle [H_{cs/e}, [H_{cs/e}, H_e]] H_e \rangle$, (e) $D_4 = \langle [H_{cs/e}, [H_{cs/e}, [H_{cs/e}, [H_{cs/e}, H_e]]]] H_e \rangle$ and (f) $D_6 = \langle [H_{cs/e}, [H_{cs/e}, [H_{cs/e}, [H_{cs/e}, [H_{cs/e}, [H_{cs/e}, H_e]]]]]] H_e \rangle$.

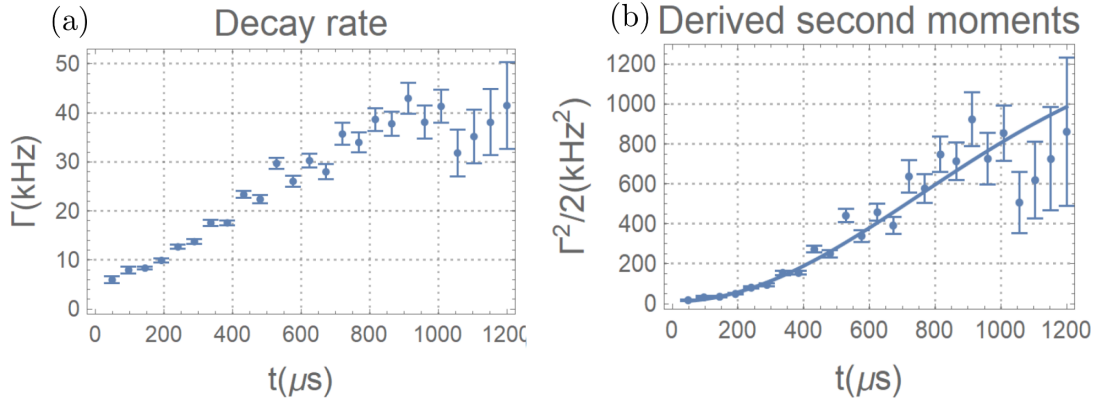


Figure 4.10: Measure of (a) linewidths and (b) second moments derived from the linewidths of the STE experiments as functions of t . They correspond to spectra obtained by Fourier transform of S_{STE} with respect to Δ .

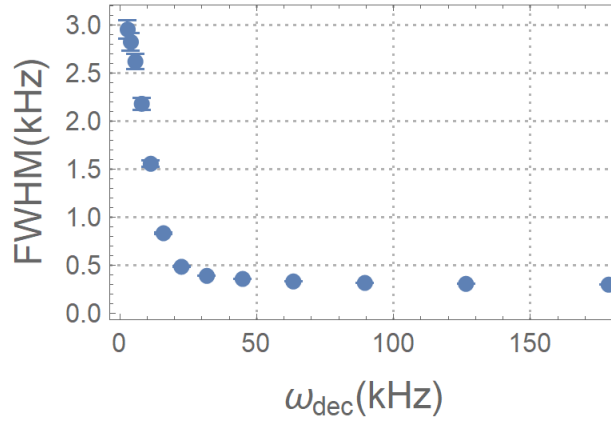


Figure 4.11: Linewidths of the spectra measured from the decoupling experiments as a function of the strength ω_{dec} of the RF irradiation.

Chapter 5

Characterizing Correlated Noise on a Central Spin Using Multi-Dimensional Experiments

Here, we investigate how central spin dynamics, combined with correlation spectroscopy, can characterize information flow in a qubit system. Our method estimates multi-time noise distributions and employs Bayesian analysis to quantify environmental memory. This reveals persistent memory and correlated fluctuations, allowing us to assess non-Gaussianity, non-stationarity, and non-Markovian features of the noise.

5.1 Introduction

Understanding a qubit's noise is crucial for developing simpler, more effective control methods in engineered quantum devices. The properties of the noise determine the performance of noise reduction strategies. Examples include decoupling [37, 131, 196, 194], which depends on noise having a long correlation time, and quantum error correction [174, 12, 100], whose efficiency depends on the symmetry of the system/environment interaction. Passive noise reduction strategies such as noiseless subsystems [101, 63] and decoherence-free subspaces [120] depend on the noise being collective and respecting certain symmetries.

Here, we present an experimental method to characterize the properties of the noise as seen by a qubit, including higher-order memory effects. The qubit can be modeled as a

central spin (CS) [49, 5] coupled to an environment that holds quantum information. We are interested in three questions:

1. What is the distribution of the frequency of the noise?
2. What is the lifetime of the CS/environment correlations?
3. What are the environmental memory effects, e.g., correlation in environmental fluctuations and long-lived states?

It is not our goal to reduce the measured dynamics to a Markovian model; rather, to learn properties of the environmental memory enables quantum efficiencies in noise rejection.

There have been numerous studies on characterizing noise in open quantum systems. R. Kubo [111] and P. Anderson [199] modeled the noise as a Gauss-Markov process which successfully predicted the line shapes in NMR and ESR experiments. Measurements have been designed to extract effective Lindblad operators [182, 126], and to measure the noise spectra of T_2 and $T_{1\rho}$ [35, 216]. These have then been reduced to an effective Markovian model that captures the loss of information in the qubit. Decoupling has been used to experimentally measure the power spectral density of the noise based on the Gauss-Markov model [3, 151]. This method has been generalized to measure higher-order polyspectra of non-Gaussian noise [145]. Recently, M. Niknam et al [143, 144] have experimentally measured the growth of CS/environment correlations. J. Chen et al [38] used stimulated echo and decoupling experiments to quantify the environmental dynamics and correlation growth in the CS system.

Here, we explore the additional information that is available from recording higher-order correlations of the noise. Correlation spectroscopy [58, 92, 171, 170] works as a simple scattering experiment. In the most basic implementation, it consists of two encoding periods separated by a variable storage time. Its higher-dimensional versions [171, 75] include multiple encoding periods and produce spectra that correspond to the multi-time joint probability distributions of the noise. Combining statistical inference with correlation spectroscopy provides a quantitative measure of memory effects including non-stationary, non-Gaussian, and non-Markovian effects.

5.2 Theory

Recall that the Hamiltonian of the CS system is

$$H = H_{cs/e} + H_e, \tag{5.1}$$

where $H_{cs/e}$ and H_e are the Hamiltonians of the CS/environment interaction and environmental mixing. We assume that the CS and the environment are initially uncorrelated with an initial state $\rho(0) = \rho_{cs}(0) \otimes \rho_e(0)$. An evolution under $H_{cs/e}$ creates correlation between the CS and the environment, which is well described by a local field model [19]. The local field depends on the eigenvalues of $H_{cs/e}$ and the environment's state which can be altered by environmental mixing. For instance, if H_e takes the form of a "flip-flop" Hamiltonian $\sigma_+ \otimes \sigma_- + \sigma_- \otimes \sigma_+$, then

$$|01\rangle \xrightarrow{\sigma_+ \otimes \sigma_- + \sigma_- \otimes \sigma_+} |10\rangle. \quad (5.2)$$

Such change in the environment's state results in a corresponding change in the local field seen by the CS, which can be described by multi-time joint distributions of the frequency of the noise $p(\Omega_n, \dots, \Omega_1)$, where Ω_i is a random variable representing the local field at the i th time instance (Fig. 4.2).

5.2.1 The Gauss-Markov Approximation

We are interested in characterizing the memory effect of the environment. Therefore, it is important to understand the assumptions made by the memoryless model, the Gauss-Markov model, and when the approximation fails.

When the noise results from contributions from a large number of weakly correlated sources, the spectrum of the local field is approximately Gaussian due to the central limit theorem. Similarly, the dynamics of environment fluctuations can be approximately Gaussian if the eigenstructure of H_e is complex and the dynamics involve averaging over different realizations. Thus, the joint distribution $p(\Omega_n, \Omega_{n-1}, \dots, \Omega_1)$ is approximately a multi-variate Gaussian distribution.

The Markov approximation claims that

$$p(\Omega_n | \Omega_{n-1}, \dots, \Omega_1) = p(\Omega_n | \Omega_{n-1}). \quad (5.3)$$

Namely, the environmental dynamics is memoryless and the present state of the noise provides all the information needed to predict the future state. For example, consider an environment of one qubit with an initial state $a|0\rangle + b|1\rangle$. $|0\rangle$ and $|1\rangle$ are eigenstates of $H_{cs/e}$ with eigenvalues $+1$ and -1 . An evolution under H_e evolves the state to $a'|0\rangle + b'|1\rangle$. The distribution of the local field changes from $\{(|a|^2, +1), (|b|^2, -1)\}$ to $\{(|a'|^2, +1), (|b'|^2, -1)\}$. In general, the new distribution depends on the values of a and b . However, if the quantum coherence $ab^*|0\rangle\langle 1|$ and $a^*b|1\rangle\langle 0|$ decay fast, the new distribution of the local field depends

only on $|a|^2$ and $|b|^2$ approximately. The process of the change of the local field can be approximately treated as a classical random process that obeys the Markov property Eq. (5.3).

The noise is stationary, if its statistical properties, e.g., mean, variance, and autocorrelation, do not change over time. A necessary and sufficient condition is that the joint distribution $p(\Omega_n, \dots, \Omega_1)$ is invariant under time translations, which requires $e^{-i\Delta H_e} \rho_e(0) e^{i\Delta H_e} = \rho_e(0)$ for all Δ . For example, if the environment is in the fully mixed state, the distribution of the local field is given by the full set of eigenvalues of $H_{cs/e}$, and

$$p(\Omega_n) \equiv p(\Omega_1), \quad (5.4)$$

for all n , which satisfies the stationary property.

The above approximations are captured by the Gauss-Markov model that treats the environmental dynamics as an Ornstein-Uhlenbeck process [108] (a stationary Gauss-Markov random process). In addition to the properties mentioned above, the process has an exponential decaying autocorrelation

$$\langle \Omega_2 \Omega_1 \rangle = e^{-\theta \Delta} \langle \Omega_1^2 \rangle, \quad (5.5)$$

and vanishing cumulants of orders higher than two.

The memory effects, due to correlated fluctuations, residual correlation in the environment, or fluctuations of the state of the environment, will show up as violations of Gaussian, Markovian, or stationary properties, respectively.

5.2.2 Measuring the joint distributions through multi-dimensional experiments

The joint distributions can be measured by an n -dimensional (n D) experiment, which has the following unitary propagator

$$U_{nD} = e^{-iH_{cs/e}t_n} e^{-iH_e\Delta_{n-1}} e^{-iH_{cs/e}t_{n-1}} \dots e^{-iH_e\Delta_1} e^{-iH_{cs/e}t_1}. \quad (5.6)$$

This separates the effect of the CS/environment interaction and the environmental dynamics so that the full dynamics under $H_{cs/e}$ or H_e during each period can be sampled by continuously varying t_i or Δ_i . A joint distribution $p(\Omega_n, \dots, \Omega_1)$ can then be obtained by an n D Fourier transform with respect to t_i . The n D experiment effectively compares $H'_{s/e}(\Delta)$ at different times and measures its time dependence. For example, the 1D spectrum gives

$$\tilde{S}_{1D}(\omega_1) = p(\Omega_1), \quad (5.7)$$

which gives the distribution of the frequencies of the noise at $\Delta = 0$. In the presence of small T_1 and T_2 , the spectrum is a convolution between $p(\omega_1, 0)$ and a Lorentzian $L(\omega)$. Here, we are interested in the case where $p(\Omega_1)$ is much broader than $L(\omega)$. The 1D experiments correspond to the growth of CS/environment correlation in the absence of the environmental dynamics.

A 2D spectrum is

$$\tilde{S}_{2D}(\omega_2, \omega_1) = p(\Omega_2, \Omega_1) = p(\Omega_2|\Omega_1)p(\Omega_1). \quad (5.8)$$

Compared to the 1D spectra, the additional information lies in the transition probability $p(\Omega_2|\Omega_1)$ which measures the change of the local field. All second-order statistical properties of the noise, such as the joint moments $\langle \Omega_2^{k_2} \Omega_1^{k_1} \rangle$ and the covariance $\text{Cov}(\Omega_2, \Omega_1)$ can be calculated using $\tilde{S}_{2D}(\omega_2, \omega_1)$.

The joint distributions $p(\Omega_2, \Omega_1)$ can be used to quantify systematic discrepancy from standard models. For instance, we can use goodness of fit, such as the mean square error (MSE) and the coefficient of determination (R^2), to a 2-variate Gaussian distribution to measure non-Gaussianity.

We can also use parameter estimation to quantify discrepancy. For example, consider measuring the non-stationarity of the noise. A stationary noise obeys

$$p(\Omega_2) \equiv p(\Omega_1), \quad (5.9)$$

for all Δ , which implies that the means of Ω_2 and Ω_1 are the same. Thus, when fitted to a 2-variate normal distribution, the absolute difference of the means (ADM) $|\mu_2 - \mu_1|$ provides a measure of non-stationarity of the noise. Note that a nonzero measure implies non-stationarity, but the opposite is not true.

Many properties of $p(\Omega_2, \Omega_1)$ can be obtained from projections of the 2D spectra, which can be obtained using reduced 2D experiments. For example, consider the stimulated echo (STE) and stimulated anti-echo (STAE) experiments, which correspond to projections of the 2D spectra along $\omega_2 = \omega_1$ and $\omega_2 = -\omega_1$ according to the central slice theorem. Their signals are

$$\begin{aligned} S_{\text{STE}}(t) &= \int d\Omega_2 d\Omega_1 e^{-i(\Omega_2 - \Omega_1)t} p(\Omega_2, \Omega_1), \\ S_{\text{STAE}}(t) &= \int d\Omega_2 d\Omega_1 e^{-i(\Omega_2 + \Omega_1)t} p(\Omega_2, \Omega_1). \end{aligned} \quad (5.10)$$

$S_{\text{STE}}(t)$ comes from the frequencies that remain the same ($\Omega_2 = \Omega_1$), while $S_{\text{STAE}}(t)$ comes from frequencies whose signs are reversed ($\Omega_2 = -\Omega_1$). For short mixing time

Δ , the frequency has a high probability to remain the same, thus $S_{\text{STE}}(t) > S_{\text{STAE}}(t)$. As Δ increases, $S_{\text{STE}}(t)$ decreases while $S_{\text{STAE}}(t)$ increases. When all correlation is lost, $S_{\text{STAE}}(t) = S_{\text{STE}}(t)$, since the system forgets the initial value of the local field Ω_1 . Additionally, the reduced experiments can be used for the statistical analysis discussed above, including quantifying non-Gaussianity and non-stationarity, since statistical inference typically requires only a subset of the data from a 2D experiment.

The 3D experiments give the 3-time joint distributions

$$\tilde{S}_{3\text{D}}(\omega_3, \omega_2, \omega_1) = p(\Omega_3, \Omega_2, \Omega_1) = p(\Omega_3|\Omega_2, \Omega_1)p(\Omega_2|\Omega_1)p(\Omega_1). \quad (5.11)$$

Compared to the 2D experiments, the additional information is in $p(\Omega_3|\Omega_2, \Omega_1)$, which measures the memory of the environmental dynamics. One way to detect the memory effect is to compare the 3D spectra with their Markovian counterparts. Define the Markov reconstruction of a 3D spectrum as

$$\tilde{S}_{3\text{D}}^M = \frac{p(\Omega_3, \Omega_2)p(\Omega_2, \Omega_1)}{p(\Omega_2)}, \quad (5.12)$$

which should be equal to $\tilde{S}_{3\text{D}}$ if the noise were Markovian: $p(\Omega_3|\Omega_2, \Omega_1) = p(\Omega_3|\Omega_2)$. Therefore, systematic discrepancy between $\tilde{S}_{3\text{D}}$ and $\tilde{S}_{3\text{D}}^M$ implies non-Markovianity. Motivated by this, define the residual correlation

$$R_\lambda^2 = (\langle \Omega_1 \Omega_3 \rangle - \langle \Omega_1 \Omega_2 \rangle \langle \Omega_2 \Omega_3 \rangle)^2, \quad (5.13)$$

which measures the difference between the direct correlation of Ω_1 and Ω_3 and that through Ω_2 . It can be evaluated using the 3D spectra and provides a quantitative measure of non-Markovianity.

The two environment time parameters Δ_1 and Δ_2 in 3D experiments provide an alternative approach to measure non-Markovianity by measuring the divisibility of the dynamics. A Markovian process is divisible, namely, the total action of a map stays the same wherever it is intercepted. Thus, the correlation $\langle \Omega_3 \Omega_1 \rangle$ should only depend on the total time of environment evolution $\Delta_1 + \Delta_2$ and not on the values of Δ_1 or Δ_2 . Thus, we can fix $\Delta_1 + \Delta_2$ and measure how $\langle \Omega_3 \Omega_1 \rangle$ changes as a function of Δ_1 to probe non-Markovianity.

Although the above discussion adopts a classical probabilistic description of the noise, the n D experiments do not rely on this and have a fully quantum interpretation where the time dependence of the local field is replaced by that of $H_{cs/e}$ in the interaction frame of H_e (Appendix).

5.3 Method

The same spin system used in the last Chapter is used to experimentally demonstrate the method. Recall that the total Hamiltonian of the system is

$$H = \delta_{\text{CS}}\sigma_z^{\text{CS}} + \sigma_z^{\text{CS}} \otimes \sum_{i=1}^N \omega_i \sigma_z^i + \sum_{i=1}^N \delta_i \sigma_z^i + \sum_{i<j}^N \omega_{ij} (2\sigma_z^i \otimes \sigma_z^j - \sigma_x^i \otimes \sigma_x^j - \sigma_y^i \otimes \sigma_y^j), \quad (5.14)$$

where N is the number of hydrogen spins in the environment; δ_{CS} and δ_i are the chemical shift anisotropy of the phosphorus and hydrogen spins, respectively; and ω_i and ω_{ij} are the coupling strengths of the P-H hetero- and H-H homo-nuclear dipolar interactions, respectively.

The pulse sequences of 1D, 2D and 3D experiments are included in the Appendix. Chemical shift anisotropy of the phosphorus spins is refocused by the pulse sequences. That of the proton spins (~ 100 Hz) is insignificant and ignored throughout the experiments. The Hamiltonians $H_{cs/e}$ and H_e are then given by the P-H heteronuclear dipolar interactions and the H-H homonuclear dipolar interactions, respectively. The sample is doped with a relaxation agent (chromium(iii) acetylacetonate) to reduce its T_1 . The time scale of evolutions under $H_{cs/e}$ (< 10 ms) is much smaller than the T_1 of the phosphorus spins (55 s) and the hydrogen spins (1 s). For mixing times ~ 1 s, the relaxation of the proton spin serves as an additional mechanism that can remove residual CS/environment correlation.

For the 2D experiments, during encode and decode periods, the symmetric magic-echo train 8 (SME8) [27] is used to turn off the environmental dynamics H_e while keeping the system/environment interaction $H_{cs/e}$ with a scaling factor of 0.342 and along the same direction. This, together with the coherence pathway selection, achieves the desired unitary propagator Eq. (5.6). 3D and higher-dimensional experiments are obtained by combining 2D experiments [76]. The coherence pathway selection also achieves amplitude-modulated time-domain signals, used to obtain pure-phase spectra by the States method [181].

Bayesian analysis is used to fit the time-domain data to that given by Gaussian distributed local field. The resulting predictive posterior distributions are used to estimate functions of the model, which are used to provide quantitative measures to non-Gaussianity, non-stationarity, and non-Markovianity.

5.4 Results

5.4.1 1D Experiment

The time-domain data and spectrum of the 1D experiment are shown in Fig. 5.1. The data are fitted to a Gaussian shape (black line in Fig. 5.1).

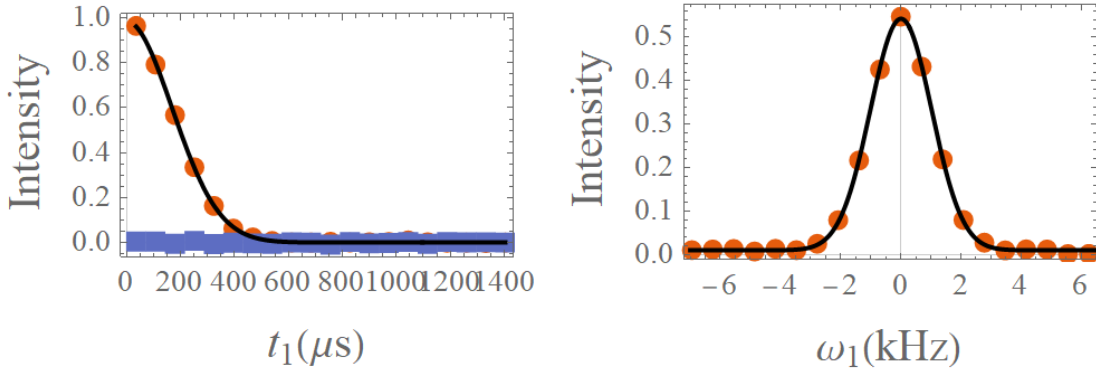


Figure 5.1: Time-domain data and corresponding spectrum of a 1D experiment. The solid lines are the best Gaussian fits.

5.4.2 2D Experiment

2D experiments of different mixing times Δ are implemented. The spectra are shown in Figs. 5.2 and 5.4, which show a broadening along $\omega_1 = -\omega_2$ as Δ increases.

The simulated 2D spectra using the Gauss-Markov model are shown in Fig. 5.3. There is a difference in shapes between the measured (Fig. 5.2) and simulated (Fig. 5.3) spectra for intermediate values of Δ . For example, for $\Delta = 50 \mu$ s, the simulated spectrum has a more rectangular shape.

To quantify the deviation of the data from the model, the MSE of the fit is calculated (Fig. 5.5). It shows a small deviation initially, which increases as Δ increases to 50μ s and decreases as Δ further increases. There is a plateau of the deviation for 100μ s $< \Delta < 1500 \mu$ s. The deviations further increase as Δ goes beyond 0.5 s.

To see if the process is stationary, projections of the spectra onto $\omega_1 = 0$ and $\omega_2 = 0$ are calculated. Their differences for different Δ are shown in Fig. 5.6. The nonzero difference at $\Delta = 50 \mu$ s suggests potential non-stationarity.

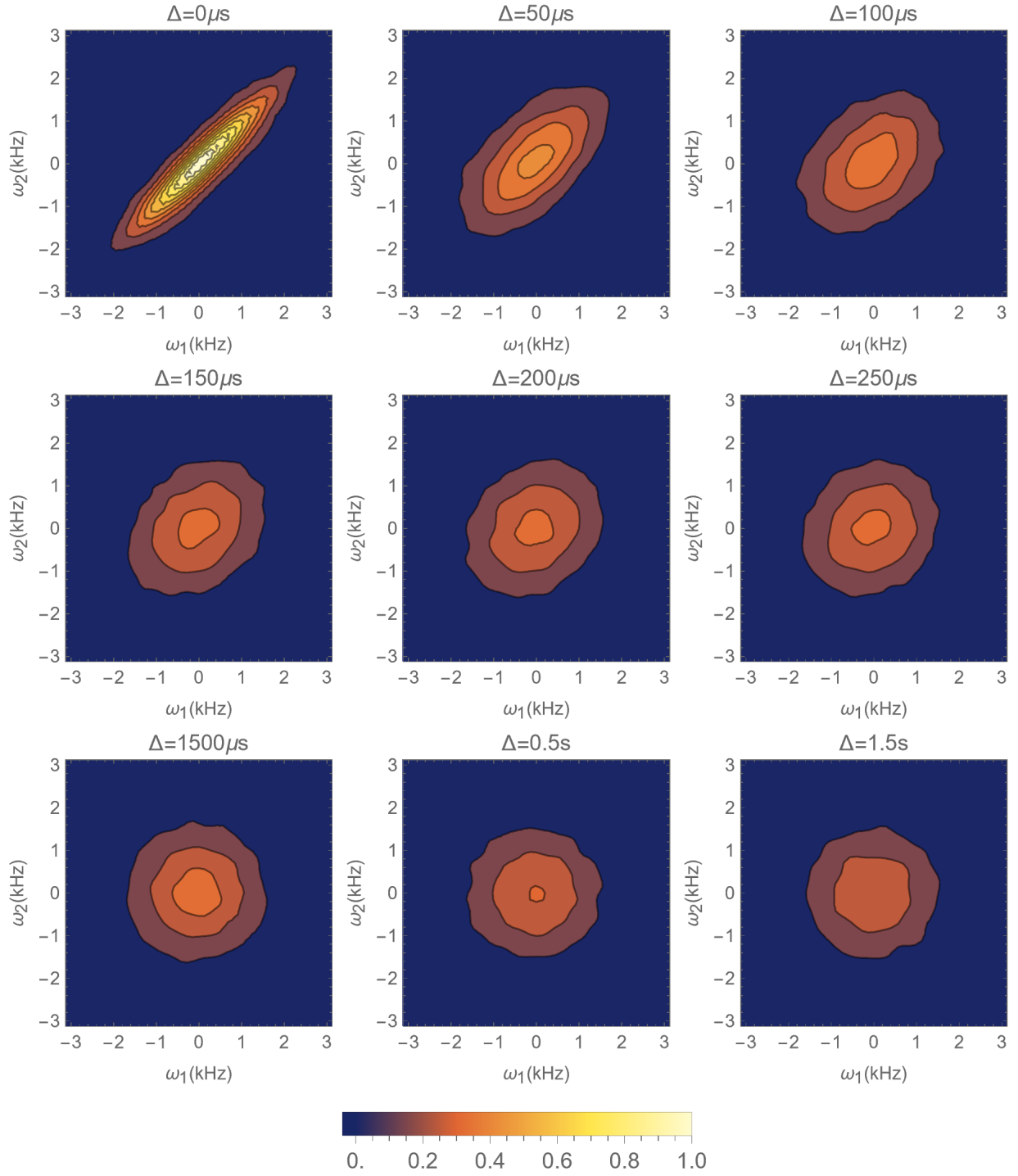


Figure 5.2: Measured 2D spectra for different Δ . The plots are normalized to the maximal point of the spectrum of $\Delta = 0 \mu s$.

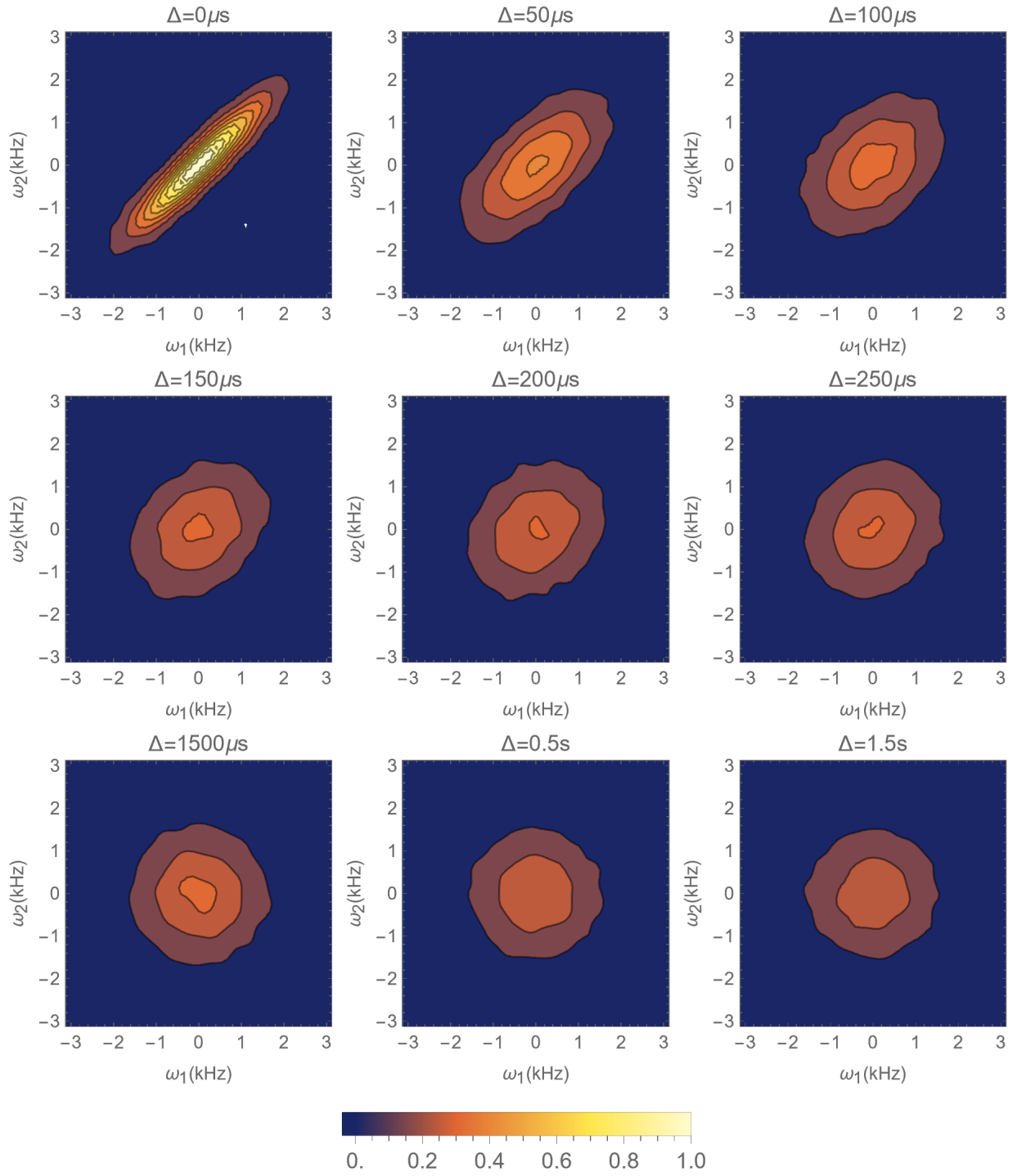


Figure 5.3: Simulated 2D spectra for different Δ . The plots are normalized to the maximal point of the spectrum of $\Delta = 0\mu s$.

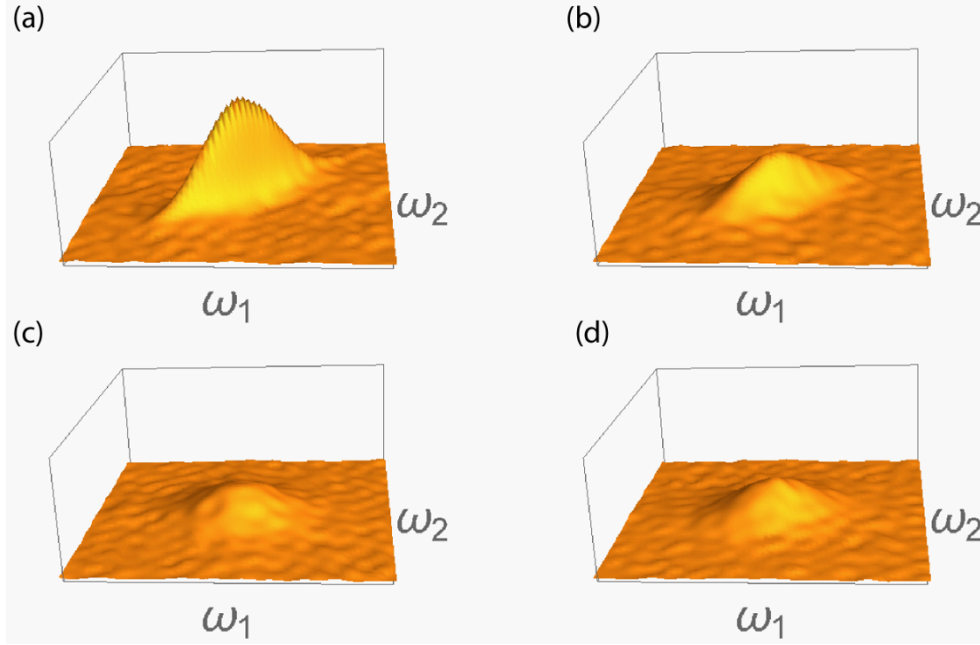


Figure 5.4: 3D views of Measured 2D spectra with various Δ . (a) $\Delta = 0 \mu s$. (b) $\Delta = 50 \mu s$. (c) $\Delta = 200 \mu s$. (d) $\Delta = 1.5 s$.

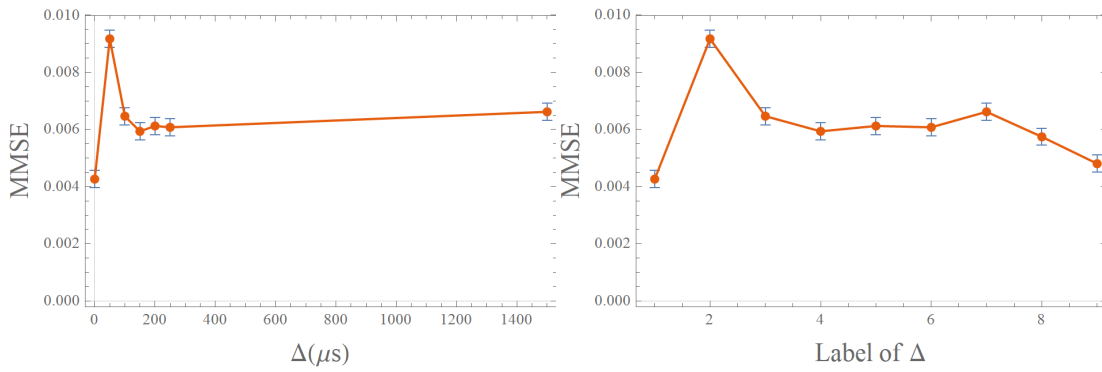


Figure 5.5: MSE of the fits to the bivariate Gaussian distribution for different Δ .

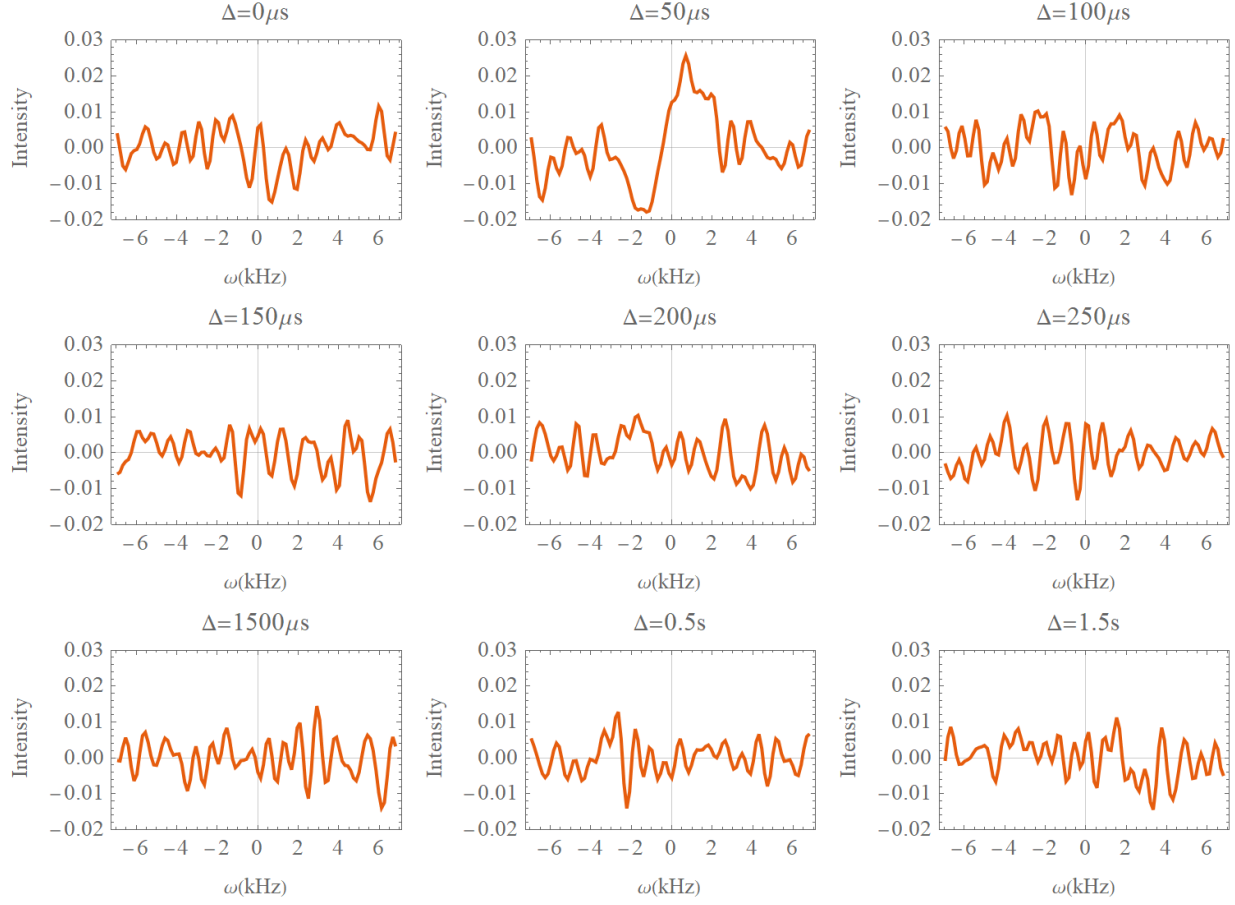


Figure 5.6: Differences of projections of 2D spectra along $\omega_1 = 0$ and $\omega_2 = 0$ for different Δ . The one at $\Delta = 50 \mu s$ shows that the dynamics are non-stationary.

5.5 Projections of 2D experiments

Projections of the 2D spectra along $\omega_2 = \omega_1$ and $\omega_2 = -\omega_1$ are measured through STE ($t_2 = -t_1$) and STAE ($t_2 = t_1$) experiments, respectively. The deviations of the data from Gauss-Markov measured by MSE are shown in Fig. 5.7. The deviation starts small, increases to a maximum at $\Delta \sim 30\mu s$, decreases as Δ further increases, and saturates at $\Delta \sim 150\mu s$. To measure stationarity, the AMD of the fit to the Gauss-Markov model is

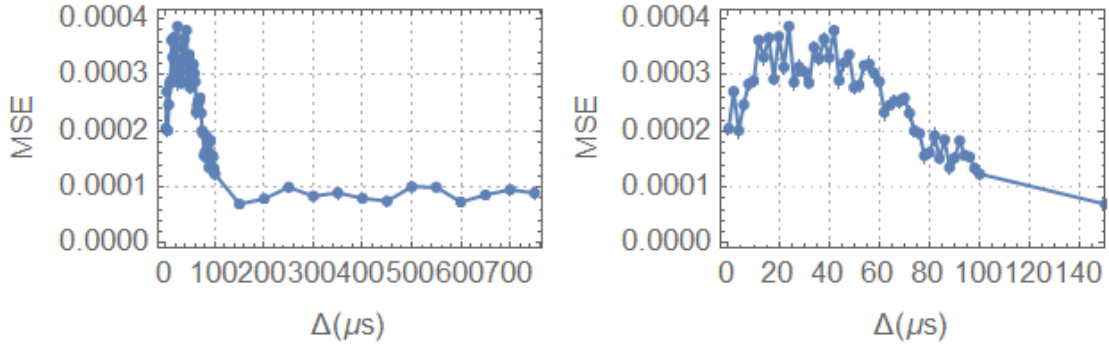


Figure 5.7: MSE of fits to bivariate Gaussian distributed local field of the STE and STAE data as a function of Δ .

calculated and plotted in Fig. 5.8.

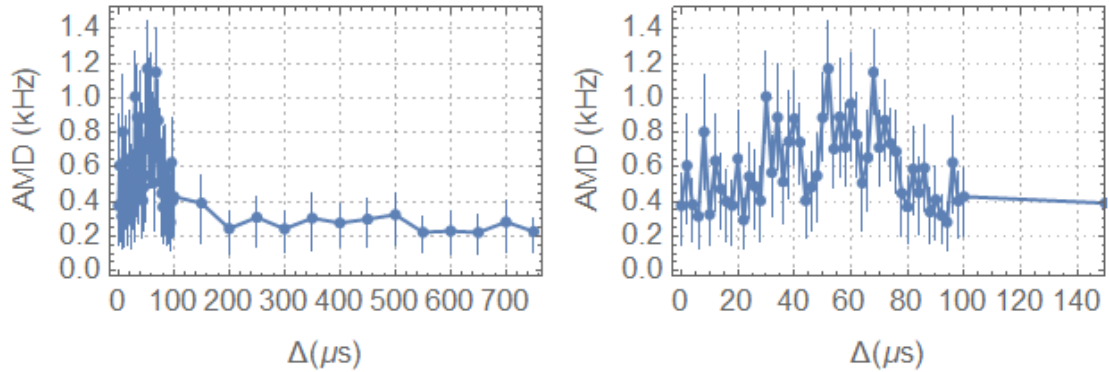


Figure 5.8: Absolute mean difference (AMD) of bivariate Gaussian fits of the STE and STAE data.

Examples of the spectra are shown in Figs. 5.9 and 5.10. To compare their contri-

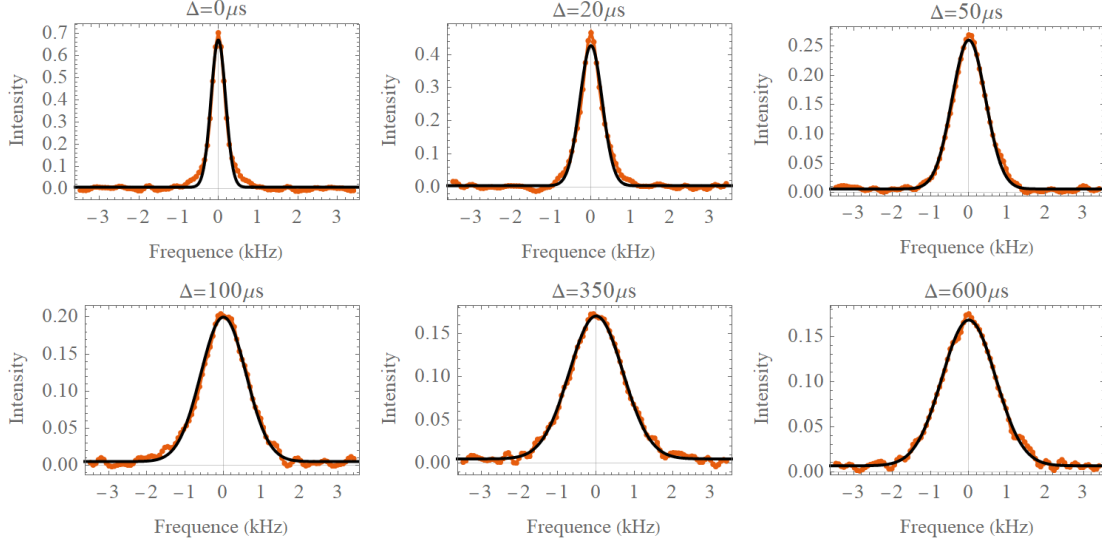


Figure 5.9: Projections of 2D spectra along $\omega_1 = \omega_2$ for different Δ . The black solid lines are the best Gaussian fits.

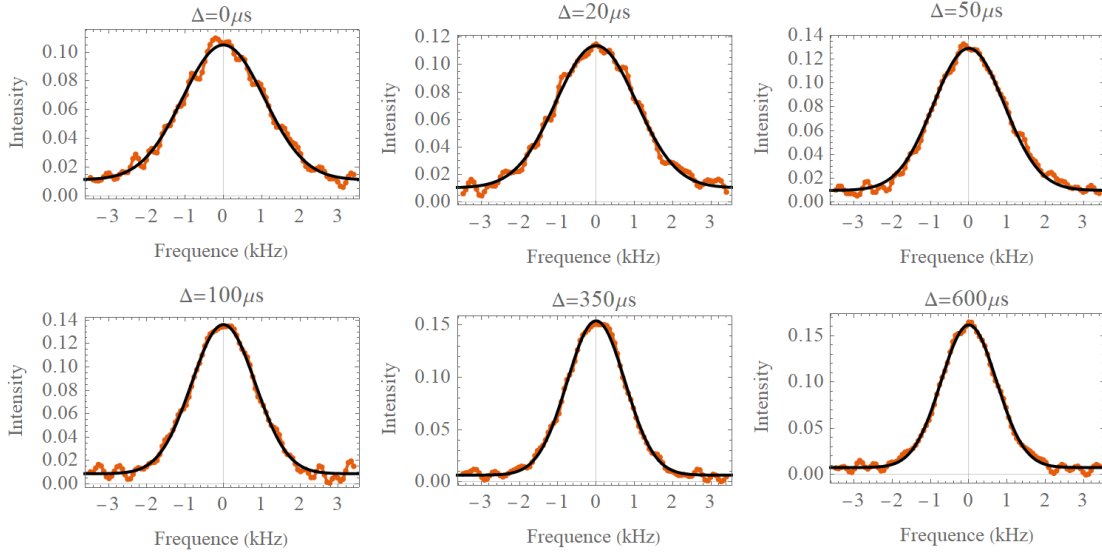


Figure 5.10: Projections of 2D spectra along $\omega_1 = -\omega_2$ for different Δ . The black solid lines are the best Gaussian fits.

butions to the deviations, the spectra are fitted to Gaussian shapes individually, and the corresponding coefficients of determination R^2 are shown in Fig. 5.11. The deviations from

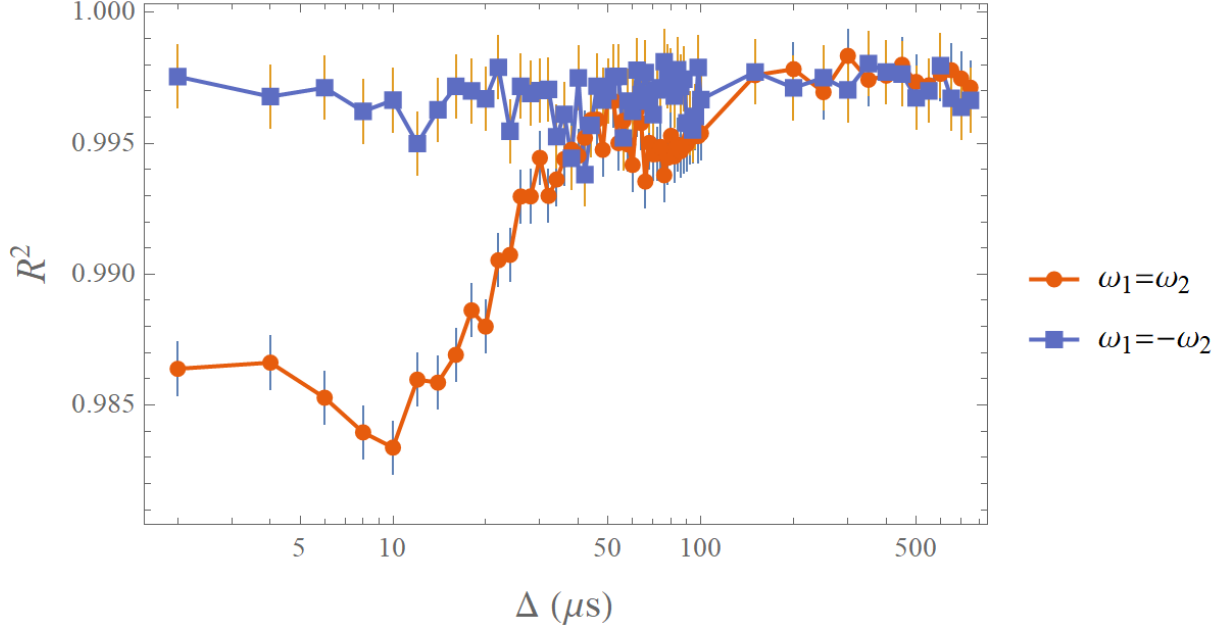


Figure 5.11: Coefficients of determination (R^2) of Gaussian fits of the projections of the 2D spectra along different directions for different Δ . Larger R^2 corresponds to better fits. A logarithmic scale of Δ is used to better represent the initial behaviors.

the STE experiments show a similar trend as the overall deviations, while those from the STAE experiments stay small, suggesting the former is the main source of the deviations.

The autocorrelation function $\langle \Omega_2 \Omega_1 \rangle$ calculated from the simulated echo and anti-echo are shown in Fig. 5.12. As a comparison, the full widths at half maximum of the spectra of the simulated echo and anti-echo experiments are also plotted (Fig. 5.12 right).

5.6 3D Experiment

3D experiments are implemented to explore the memory effect of the environmental dynamics. Fig. 5.13 shows the measured spectra ($\Delta_1 = \Delta_2$). In comparison, the Markov reconstructions of the spectra are plotted beneath them. The measured spectra exhibit a more structured surface than their Markov reconstructions, indicating a possible memory

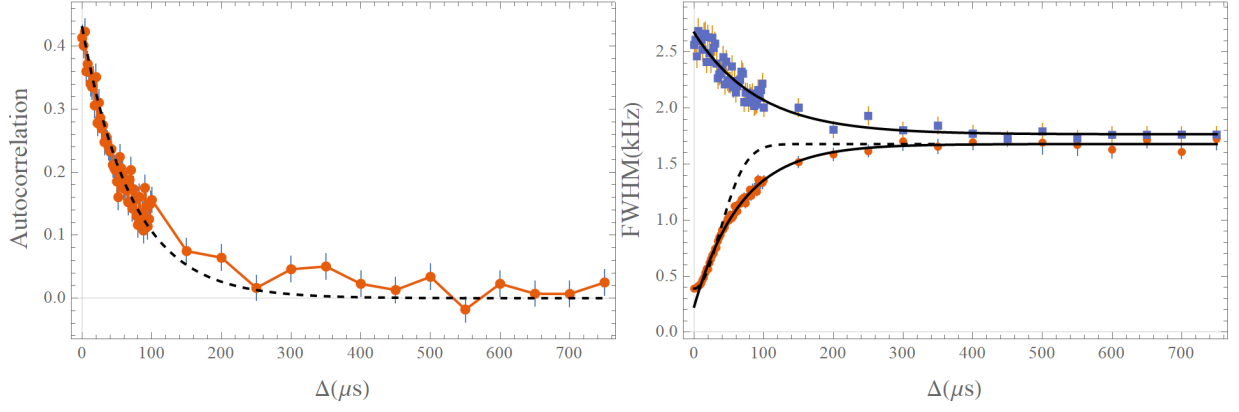


Figure 5.12: Measured autocorrelation function $\langle \Omega_2 \Omega_1 \rangle$ (left) with the best exponential fit (dashed line) and the linewidths of the projections (right) with the best exponential fits (solid lines) and the best Gaussian fit (dashed line). The blue squares and the orange dots are for projections along $\omega_2 = \omega_1$ and $\omega_2 = -\omega_1$, respectively.

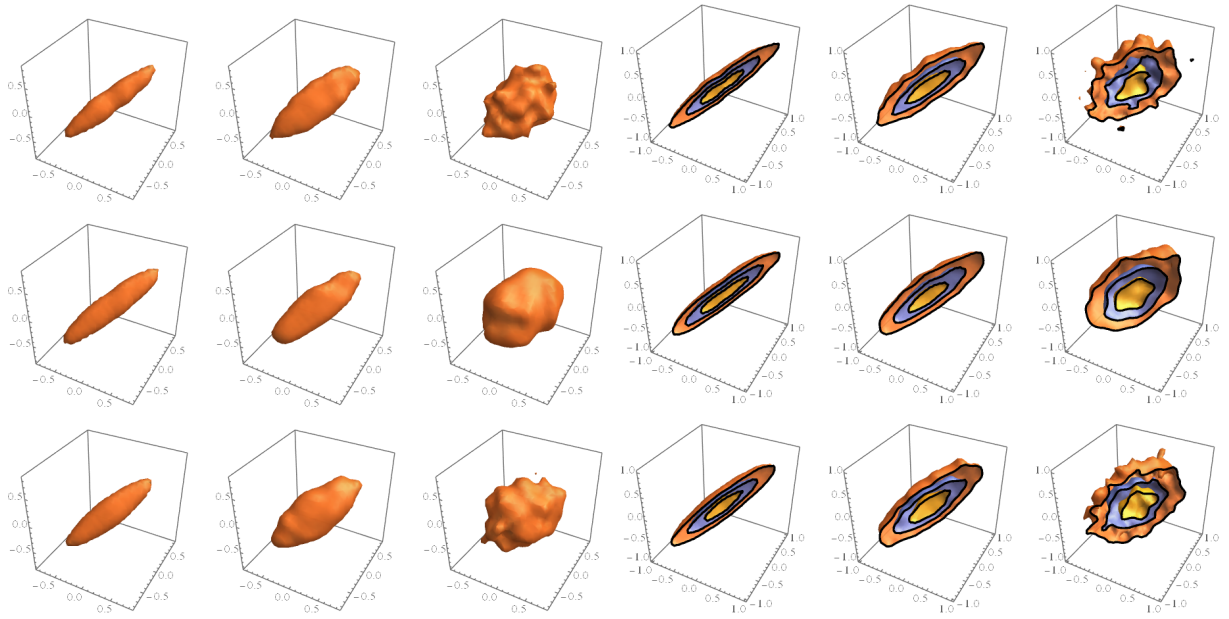


Figure 5.13: 3D spectra (left 3) and their sectional views (right 3) for different $\Delta_2 = \Delta_1$ (from left to right: 0, $20\mu\text{s}$ and $100\mu\text{s}$). From top to bottom: Measured spectra, Markov reconstructions, and best Gaussian fit.

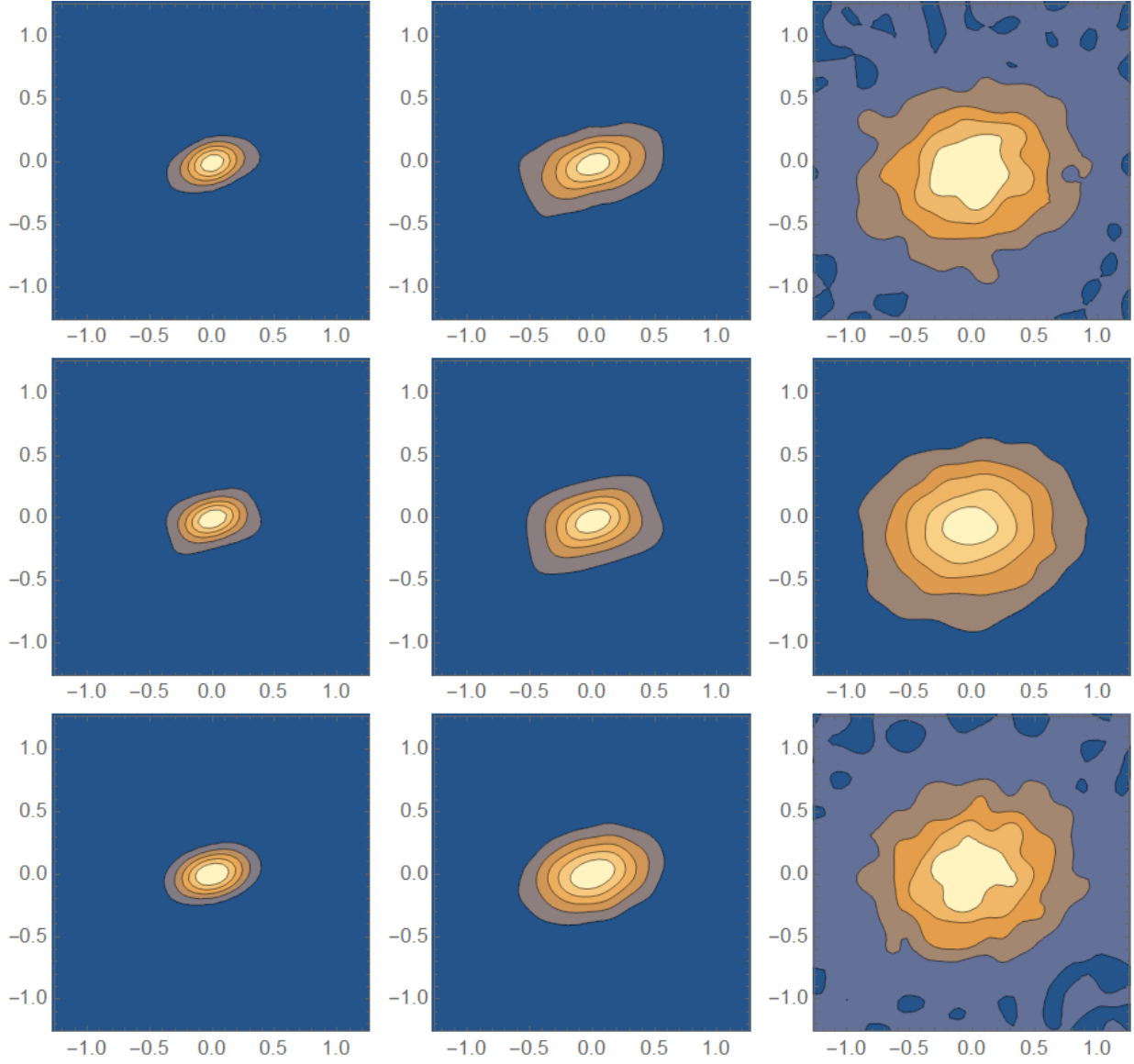


Figure 5.14: Projections of the 3D spectra along $(1,1,1)$. From top to bottom: Measured spectra, Markov reconstructions, and best Gauss-Markov fit. From left to right: $\Delta_1 = \Delta_2 = 0, 20\mu s$ and $100\mu s$.

effect. To check whether the features arise from the measurement noise, the data are fitted to the trivariate normal distributions (Fig. 5.13 bottom) and used to calculate the residual correlation (Fig. 5.15). The result shows a similar result with that observed from the 2D

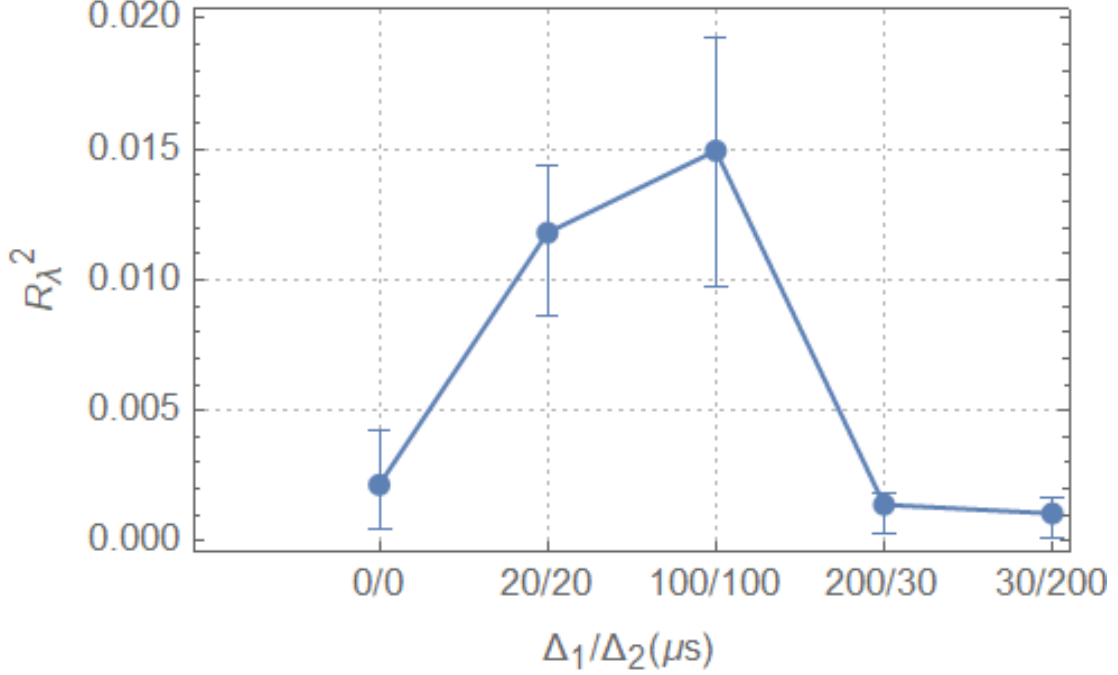


Figure 5.15: Residual correlation for different 3D experiments.

experiments: the residual correlation starts low and increases as the storage times increase. However, the memory persists longer ($> \Delta_1 + \Delta_2 = 200\mu s$) than previously measured in the 2D experiment ($\Delta < 150\mu s$).

To further explore the memory effect, two 3D experiments are performed, where $\Delta_1/\Delta_2 = 200\mu s/30\mu s$ and $30\mu s/200\mu s$. The spectra are shown in Fig. 5.16. Their corresponding residual correlations are calculated (Fig. 5.15). The posterior distributions of the total correlation $\langle \Omega_3 \Omega_1 \rangle$ are shown in Fig. 5.17. The first experiment ($200\mu s/30\mu s$) shows less correlation.

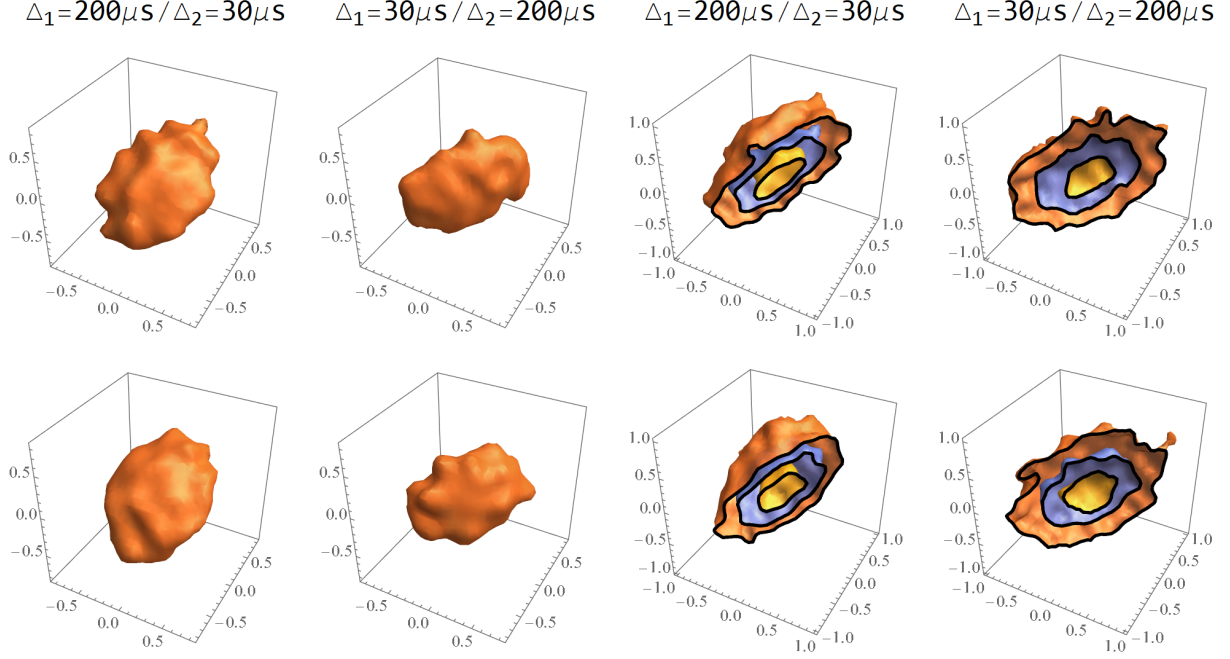


Figure 5.16: 3D displays (left half, one contour at 0.5 maximum) and sectional views (right half, contours at 0.25, 0.5 and 0.75 maximum) of measured (top) and simulated (bottom) spectra for $\Delta_2 = 200\mu s / \Delta_1 = 30\mu s$ and $\Delta_1 = 30\mu s / \Delta_2 = 200\mu s$.

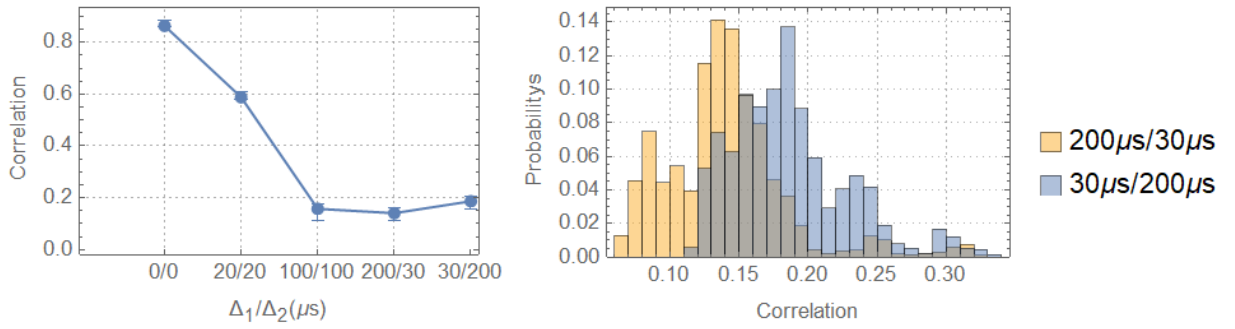


Figure 5.17: Posterior distributions of correlations $\langle \Omega(\Delta_1 + \Delta_2) \Omega(0) \rangle$. The posterior mean estimates are $0.31^{+0.04}_{-0.07}$ ($200\mu s / 30\mu s$) and $0.322^{+0.028}_{-0.031}$ ($30\mu s / 200\mu s$). As a comparison, $\langle \Omega_3 \Omega_1 \rangle$ for all experiments are plotted on the left.

5.7 Discussion

5.7.1 1D experiment

The 1D experiment yields the distribution of the local field in the absence of the environmental dynamics, which determines the growth of CS/environment correlation and the coherence time of the CS. The distribution is well approximated by a Gaussian shape due to the complexity of $H_{cs/e}$ and averaging over many realizations. The memory effect of the environment is not captured by the 1D experiment.

5.7.2 2D experiment

The 2D experiments measure the change of the local field. The spectrum at $\Delta = 0\mu s$ centers at $\omega_2 = \omega_1$, suggesting no change of the local field in the absence of H_e . Ω_2 and Ω_1 are highly correlated

$$p(\Omega_2, \Omega_1) \sim \delta(\Omega_2 - \Omega_1)p(\Omega_1). \quad (5.15)$$

This correlation decreases as Δ increases, shown as the broadening of the spectra along $\omega_2 = -\omega_1$. The MSE of the fit to the bivariate Gaussian distribution measures non-Gaussianity of the environmental dynamics as a function of Δ (Fig. 5.5). The MSE is small at $\Delta = 0\mu s$, consistent with the result from the 1D experiment (in the absence of H_e , the dynamics under $H_{cs/e}$ is well captured by a Gaussian distribution). The MSE increases as Δ increases, indicating increased non-Gaussianity in the dynamics, due to correlated fluctuations in the environmental dynamics. The decrease of the MSE as Δ further increases suggests a transition to a fully mixed regime where the correlation between environment fluctuations is lost. However, a long-lived residual correlation persists up to $\Delta = 1500\mu s$, as shown by the decrease of the MSE at $\Delta = 0.5s$ and $1.5s$, which involves relaxation in the environment, suggesting a constant of motion in the CS/environment correlation to the unitary dynamics of the environment.

The difference between the projections of the spectra onto $\omega_1 = 0$ and $\omega_2 = 0$ at $\Delta = 50\mu s$ suggests potential non-stationarity (Fig. 5.6).

5.7.3 Projections of 2D spectra

The projections of the 2D spectra along $\omega_2 = \omega_1$ and $\omega_2 = -\omega_1$, measured by the stimulated echo and anti-echo experiments, provide complementary information to the 2D experiments.

The MSE of the fit to the bivariate Gaussian distribution (Fig. 5.5), equivalent to that measured in the 2D experiments (Fig. 5.7), but with more points, shows a similar trend, confirming the non-Gaussianity at intermediate values of Δ . It reaches the maximum at $\Delta \sim 30\mu s$, around $1/|H_e|$ ($|H_e| \sim 30$ kHz), and falls to a minimum at $\Delta \sim 150\mu s$, suggesting a transition from a Gaussian regime to a non-Gaussian regime then back to the Gaussian regime.

The AMD of the fit provides a measure of non-stationarity of the noise (Fig. 5.8). At small Δ , non-stationarity is small due to small environmental dynamics. However, the nonzero AMD in $20\mu s < \Delta < 80\mu s$ indicates non-stationarity, consistent with Fig. 5.6.

The coefficient of determination measured for Gaussian fits suggests that the non-Gaussianity mainly results from the STE, which is consistent since the initial environmental dynamics are mainly captured by the STE. At small Δ (e.g., $\Delta = 0\mu s$ and $20\mu s$), there is a pedestal of the spectra of the STE, not captured by the Gaussian fit, around the baseline. This is possibly due to a part of the environment with high frequency that is not decoupled by the SME8 sequence. To test this hypothesis, compare the spectra with those obtained from the decoupling experiments (Fig. 5.18), which also shows the same pedestal around baseline at $\omega_{\text{dec}} \sim |H_e|$. The pedestal disappears as ω_{dec} increases above 60 kHz, suggesting a part of the environment, with a frequency higher than $|H_e|$, is suppressed.

The linewidths of the STE and STAE become closer to each other as Δ increases, suggesting enhanced mixing. However, they do not saturate to the same value, indicating residual correlation, as observed from the 2D experiments (Fig. 5.5). Although no obvious deviation from exponential decay is observed from the calculated autocorrelation function due to high uncertainty, the linewidths of the STE spectra show a Zeno regime marked by a Gaussian signature in $0\mu s < \Delta < 20\mu s$, before transitioning to an exponential decay.

Table 5.1 summarizes different time scales observed so far in the system.

5.7.4 3D experiment

The residual correlations in the 3D experiments show that the memory effect persists up to $\Delta_1 + \Delta_2 = 200\mu s$ which is as long as the mixing time. This is likely due to the interception by the evolution under $H_{cs/e}$ between the first and second mixing periods, which reduces the total action of the mixing. This is further confirmed by the 3D experiments where $\Delta_1 + \Delta_2$ is fixed. The experiment where the mixing is intercepted earlier ($\Delta_1 = 30\mu s$) shows less residual correlation and more total correlation between Ω_3 and Ω_1 than the experiment where $\Delta_1 = 200\mu s$. The interception at the early stage of environment mixing disrupts the accumulation of higher-order effects, thus reducing its overall effectiveness.

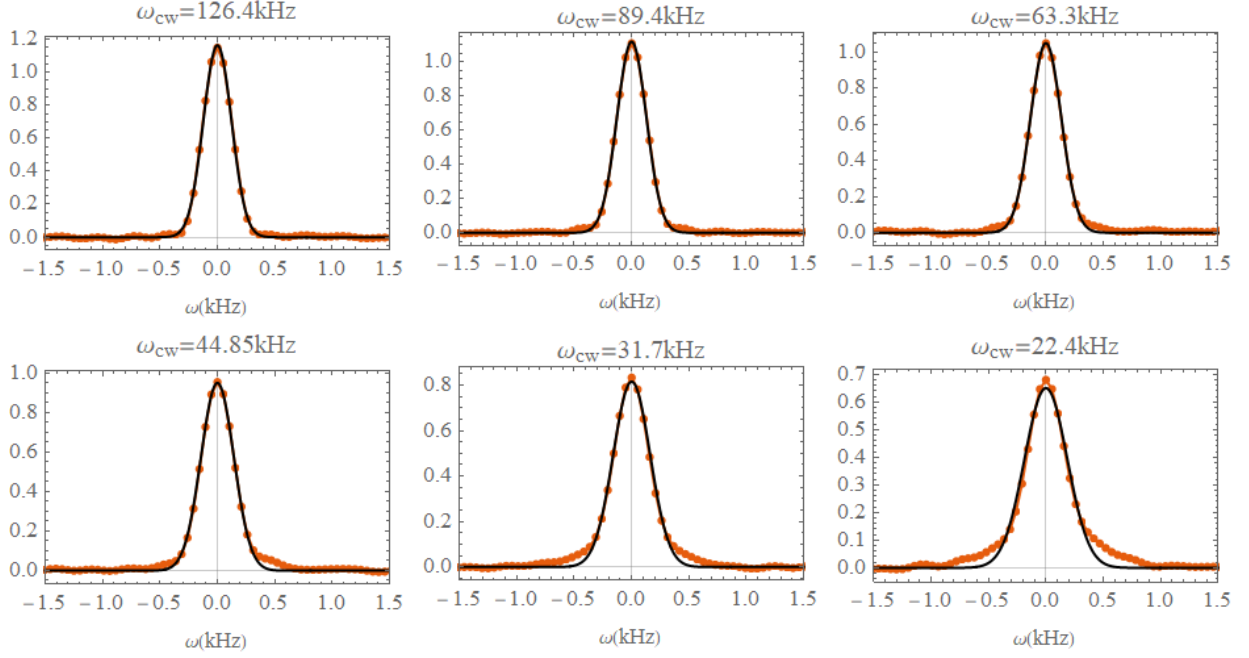


Figure 5.18: Spectra of decoupling experiments. Black lines are the best Gaussian fits.

Table 5.1: Different time scales measured by the 1D and 2D experiments.

Name	Time scale
CS coherence time	$t \sim 300\mu s$
Mixing time	$\Delta \sim 200\mu s$
Non-Gaussian	$\Delta \sim [0, 150]\mu s$
Quantum Zeno	$\Delta \sim [0, 20]\mu s$
Gaussian-exponential transition	$\Delta \sim 50\mu s$
Non-stationary	$\Delta \sim [0, 100]\mu s$
Constant of motion	$\Delta > 750\mu s$

Such interception is less effective at a later stage, however, as the mixing has reached saturation.

5.8 Conclusion

We presented a general experimental method for comprehensively characterizing highly correlated quantum noise in a central spin system using multi-dimensional experiments. The experiments measure the multi-time joint distributions of the noise, represented by the multi-dimensional spectra, which contain all the statistical properties of the noise. Reduced multi-dimensional experiments, which produce projections of the multi-dimensional spectra, provide an efficient way to probe the noise and can be used in statistical inference to provide a quantitative measure of the memory effect of the environment.

Using the multi-dimensional experiments, we experimentally characterized the noise in a central spin system. The one-, two-, and three-dimensional experiments measure the distribution of the local field, change of local field by environmental dynamics, and memory effect of the environmental dynamics, respectively. Through quantitative measures of non-Gaussianity and non-stationarity, the two-dimensional experiments and their reduced versions, the stimulated echo and anti-echo experiments, reveal a non-Gaussian period ($0\mu s < \Delta < 150\mu s$) due to correlated environmental fluctuations and a non-stationary period ($0\mu s < \Delta < 100\mu s$) due to return of information from the environment to the central spin. Residual system/environment correlation, which persists beyond the mixing time ($200\mu s < \Delta < 1500\mu s$) has been revealed, suggesting a component insensitive to the environmental dynamics. The 3D experiments provide a quantitative measure of non-Markovianity of the environmental dynamics through residual correlation and non-divisibility, revealing persistent memory effect above the mixing time of $200\mu s$. Intercepting the environmental dynamics by information flow between the central spin and the environment reduces the effectiveness of higher-order effects of the environmental dynamics, leading to reduced total action and prolonged system/environment correlation.

The quantification of different aspects of the memory effect, including non-Gaussianity, non-stationarity, and non-Markovianity, provides a benchmark for noise modeling, a threshold for error correction and control protocols, and an effective sensor for learning dynamics of many-body systems.

Chapter 6

Probing the Structure of the Environment through the Central Spin Using Multi-quantum Experiments

Understanding CS/environment interaction and information flow is essential for advancing quantum technologies. The environmental dynamics can scramble the CS/environment correlation and hold this information in subspaces that are not directly available to the CS. As shown in Chapter 4, stimulated echo experiments provide a simple way to study information flow between the CS and the environment. In Chapter 5, the multi-dimensional correlation spectroscopy provides a more complete picture of the memory effect due to the environmental dynamics.

Although it is possible to measure all relevant information of the noise on the CS without knowing the detailed structure of the environment, a better understanding of how information is stored in the environment can enable better prediction of non-Markovian behaviors in the noise [155, 65, 35]. The additional information can be measured by the multi-quantum (MQ) experiments which were first introduced to study multi-body correlations in NMR systems [204, 9, 140]. The basic idea is that different components of the density matrix have different sensitivity to external perturbation. By encoding the qubit system with collective external rotations, the MQ experiments can resolve these components. To see how this can be applied to study dynamics of the CS system, consider an initial state where the quantum information starts at the CS: $\rho(0) = \sigma_x \otimes \mathbb{1}^{\otimes N} / 2^{N+1}$.

Under the CS/environment interaction

$$H_{cs/e} = \sigma_z \otimes \sum_{i=1}^N \omega_i \sigma_z^i, \quad (6.1)$$

the state evolves to

$$\begin{aligned} \rho(t) = \frac{1}{2^{N+1}} & \left(\sigma_x \otimes \mathbb{1}^N \prod_{i=1}^N \cos \omega_i t + \sum_{i=1}^N \sigma_y \otimes \sigma_z^i \sin \omega_i t \prod_{j \neq i}^N \cos \omega_j t + \dots \right. \\ & \left. + \sigma_{[N]} \otimes \sigma_z^1 \otimes \dots \otimes \sigma_z^N \prod_{i=1}^N \sin \omega_i t \right). \end{aligned} \quad (6.2)$$

The multi-body terms represent the CS/environment correlation. The highest number of operators that can be involved is equal to the number of spins in the environment N . $\sigma_{[N]} = \sigma_x$ if N is even and $\sigma_{[N]} = \sigma_y$ if N is odd. In the absence of environmental dynamics, $\rho(t)$ can be returned back to $\rho(0)$ with a simple echoing.

The MQ experiments use collective x rotations on the environment to encode the CS/environment correlation so that the echo carries information from the encoding. Under the basis $\{\mathbb{1}, \sigma_{+x}, \sigma_{-x}, \sigma_x\}^{\otimes N}$, operators with the same coherence order ($n_+ - n_-$, where n_+ and n_- are number of σ_+^x and σ_-^x) span an invariant subspace of the x rotations with the same eigenvalue ($e^{i(n_+ - n_-)\theta}$ where θ is the rotation angle). Thus, the x rotations encode projections of $\rho(t)$ onto the subspaces with their respective eigenvalues. Given N , the number of invariant subspaces is $2N + 1$. So, if we perform $2N + 1$ experiments with different θ , we can solve the $2N + 1$ linear equations and obtain the projection of $\rho(t)$ onto each subspace. Conventionally, θ is incremented from 0 to $4\pi N/(2N + 1)$ in step of $2\pi/(2N + 1)$ and a Fourier transform is used to obtain the projections. This can be compared with permutationally invariant state tomography [190] where the density matrix is partially measured with respect to the symmetry imposed by the x rotations. The population of the quantum information in these subspaces depends on the connectivity between the CS and the environment and the structure of the environment. Note, this additional information cannot be obtained in the multi-dimensional experiments; it is the added control, the x collective rotations on the environment, that enables us to do so.

Since the number of invariant subspaces is directly related to the number of spins in the environment, the MQ experiments effectively measure the size of the environment. M. Niknam, et al used this method to measure the growth of the correlation in the CS system [143] and used entropies to quantify the amount of CS/environment correlation [144].

In this chapter, we build on these results and show how MQ experiments can also be used to study the environmental dynamics (Fig. 6.1). We study two variations. In the

first variation, we study how information leakage can affect the performance of the MQ experiments. [143] also studied sensitivity of CS/system correlation to environmental mixing. Here, our approach is related to the Loschmidt echo (LE) experiments [90, 72, 167] where the goal is to study how an ideally perfect echo is affected by errors during the implementation of the experiment. The focus here is the refocusability and learnability of the environment by the CS, which is important for implementing robust quantum operations and sensing. MQ experiments provide powerful tools for studying this sensitivity, as they can track the sensitivity of components of each coherence order.

In the second variation of MQ experiments, we design experiments where a period of environmental mixing is added after the CS has established correlation with the environment. Then after the encoding, we refocus both the mixing and CS/environment interaction. This allows us to encode the dynamics of the environment and show how environmental mixing induces flow of information between subspaces of different coherence orders. This provides us with a more detailed view on how quantum information is scrambled by the mixing. As will be shown with the experimental results, the mixing creates a structured population on the subspaces and connects the CS to a larger, extended environment. This environment can store quantum information not directly seen by the CS and is responsible for long memory.

We also show how Bayesian inference can be combined with the MQ experiments to gain improved sensitivity of the measurement of the size of the environment. Taking the standard MQ experiments as an example, we show that the CS can develop correlation with up to fifteen spins from the environment which is more than what can be measured from counting the coherence orders.

6.1 Standard multi-quantum experiments

First, we implement the standard MQ experiments for measuring growth of CS/environment correlation [143]. The procedure of the experiments is as follows:

$$\rho(0) \xrightarrow{\exp(-iH_s/e t)} \rho(t) \xrightarrow{\exp(-iJ_x\theta/2)} \rho(t, \theta) \xrightarrow{\exp(iH_s/e t)} \text{measure}, \quad (6.3)$$

where θ is systematically incremented from 0 to $2\pi * 31/32$. The MQ experiments for different values of t are performed on the CS system using the same solid-state NMR system introduced in Chapter 4. Fig. 6.2 shows the time-domain data of the experiments. As can be seen here, the intensity of the echo becomes more sensitive to θ as t increases, suggesting a growth of CS/environment correlation. Fig. 6.3 shows measured amplitudes

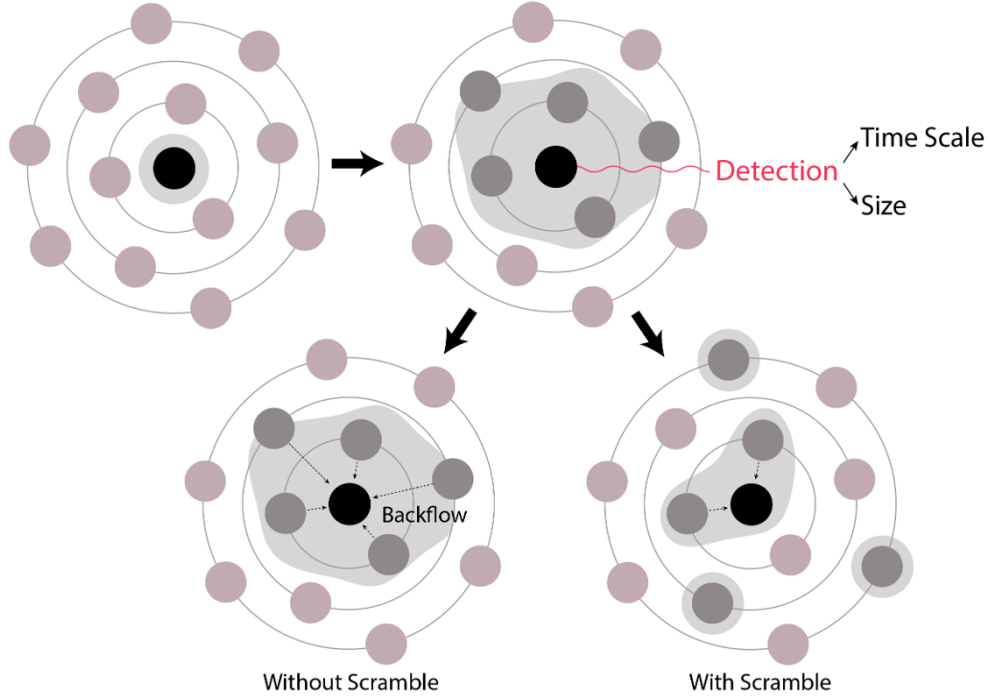


Figure 6.1: Schematic of learning the environment through a CS. The environment mixing may destroy the CS/environment interaction, reducing learnability of the environment. If refocused, however, the mixing may help reveal a larger environment as it can reach parts not directly observable via $H_{cs/e}$.

of different coherence orders for different t . To see the time evolution of different coherence order, Fig. 6.4 shows the time-dependence of different coherence orders. The picture shows a detailed time dependence of the correlation build-up between the CS and the environment determined by $H_{s/e}$. A thorough study of this effect can be found in [143, 144]. (Note that they had used a different sequence with different scaling factors and were looking at total amplitudes of absolute coherence orders.) The result shows that the CS sees a larger portion of the environment as the evolution time, t , increases. This is manifested in the growth of higher-order coherence and the broadening of the coherence distribution.

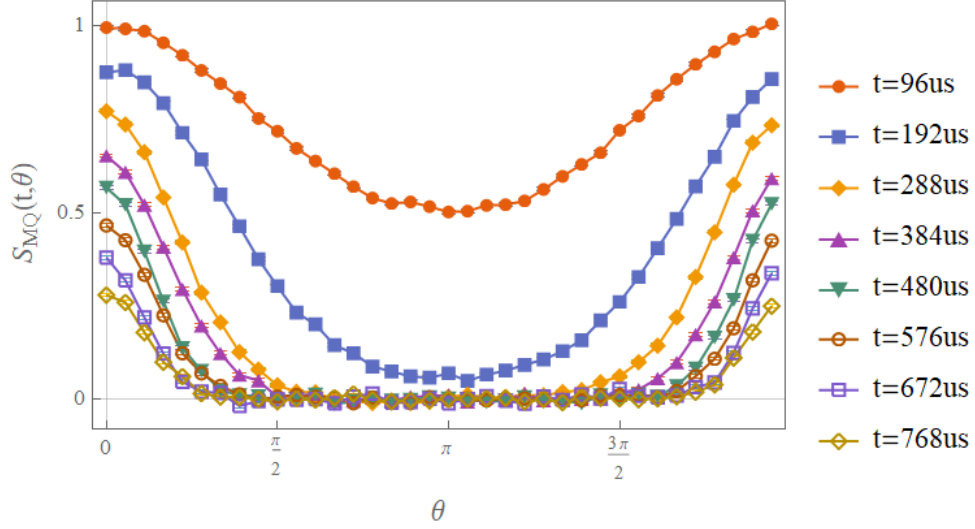


Figure 6.2: Time-domain data of standard MQ experiments.

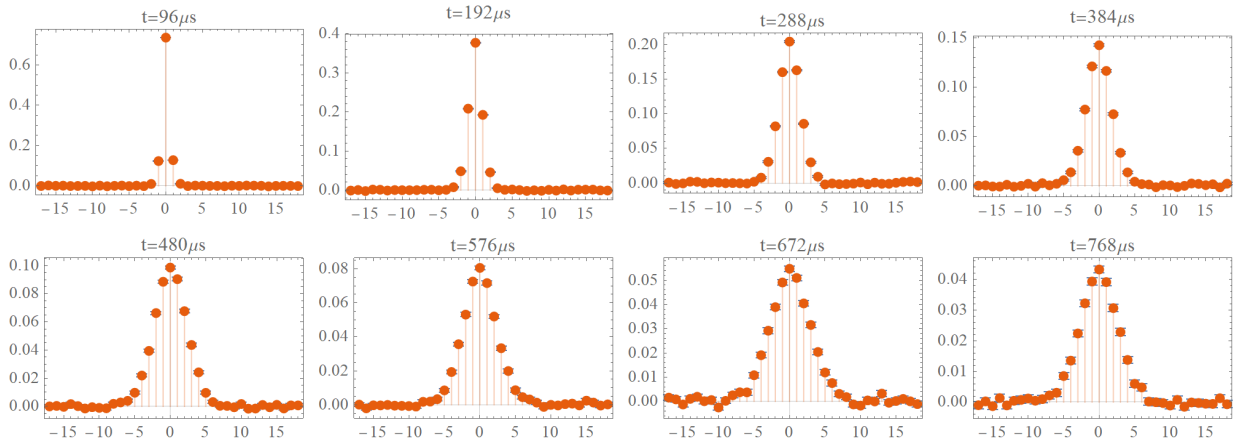


Figure 6.3: Amplitudes of different coherence orders (from -17 to 18) for different t .

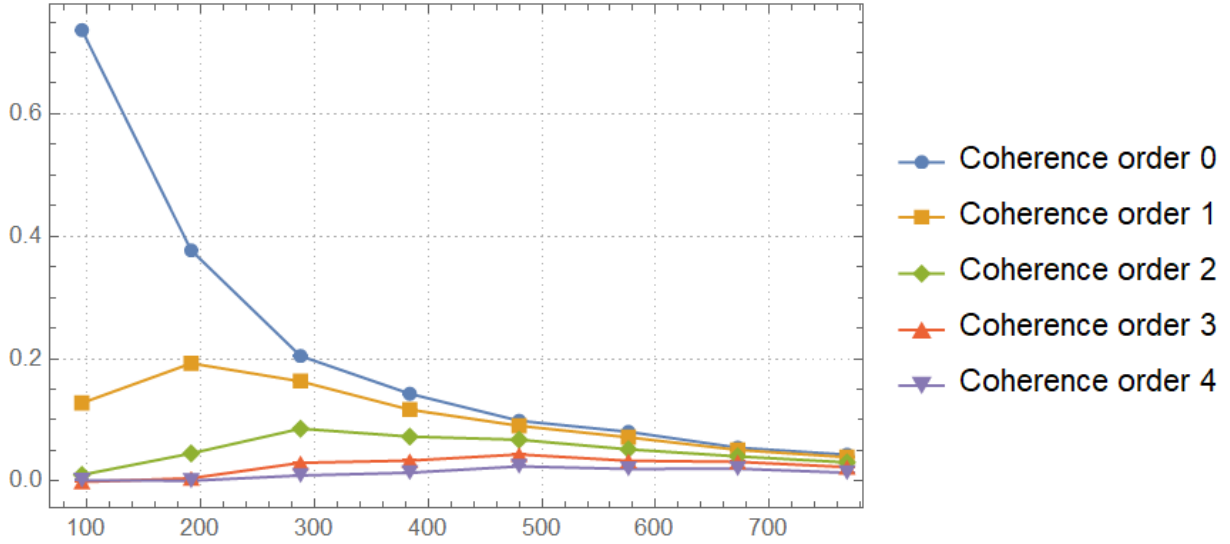


Figure 6.4: Amplitudes of different coherence orders as functions of t .

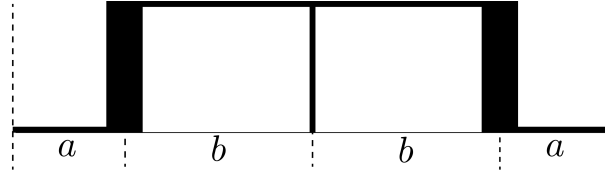


Figure 6.5: Modified magic echo sequence. a and b are the durations of delays and the spin-locking field, respectively. By varying a and b , different effective Hamiltonians can be achieved under the sequence.

6.2 Coherence growth in the presence of environmental dynamics

The procedure of MQ experiments in the presence of environmental perturbation, the LE experiments, is

$$\rho(0) \xrightarrow{\exp(-i(s_{cs/e}H_{cs/e}+s_eH_e)t)} \rho(t) \xrightarrow{\exp(-iJ_x\theta/2)} \rho(t,\theta) \xrightarrow{\exp(-i(-s_{cs/e}H_{cs/e}+s_eH_e)t)} \text{measure.} \quad (6.4)$$

The added environmental dynamics, multiplied by a scaling factor s_e , results in an imperfect inversion of the dynamics and loss of quantum information.

Mixing is introduced by changing the durations of the RF radiation (b) in the magic

echo sequences (Fig. 6.5). This results in the scaling factors

$$s_{cs/e} \approx \frac{a}{a+b}, \quad s_e \approx \frac{2a-b}{2(a+b)}. \quad (6.5)$$

The exact scaling factors are obtained by calculating the average Hamiltonian of $H_{cs/e}$ and H_e . We keep $a+b$ fixed so that the decoherence is the same for different experiments and vary $2a-b$. Note that the encoding time t needs to be scaled according to $s_{cs/e}$. Experiments of different $2a-b$ are implemented. Fig. 6.6 shows measured amplitudes of coherence orders. The total amplitudes of the echoes are shown in Fig. 6.2. When no

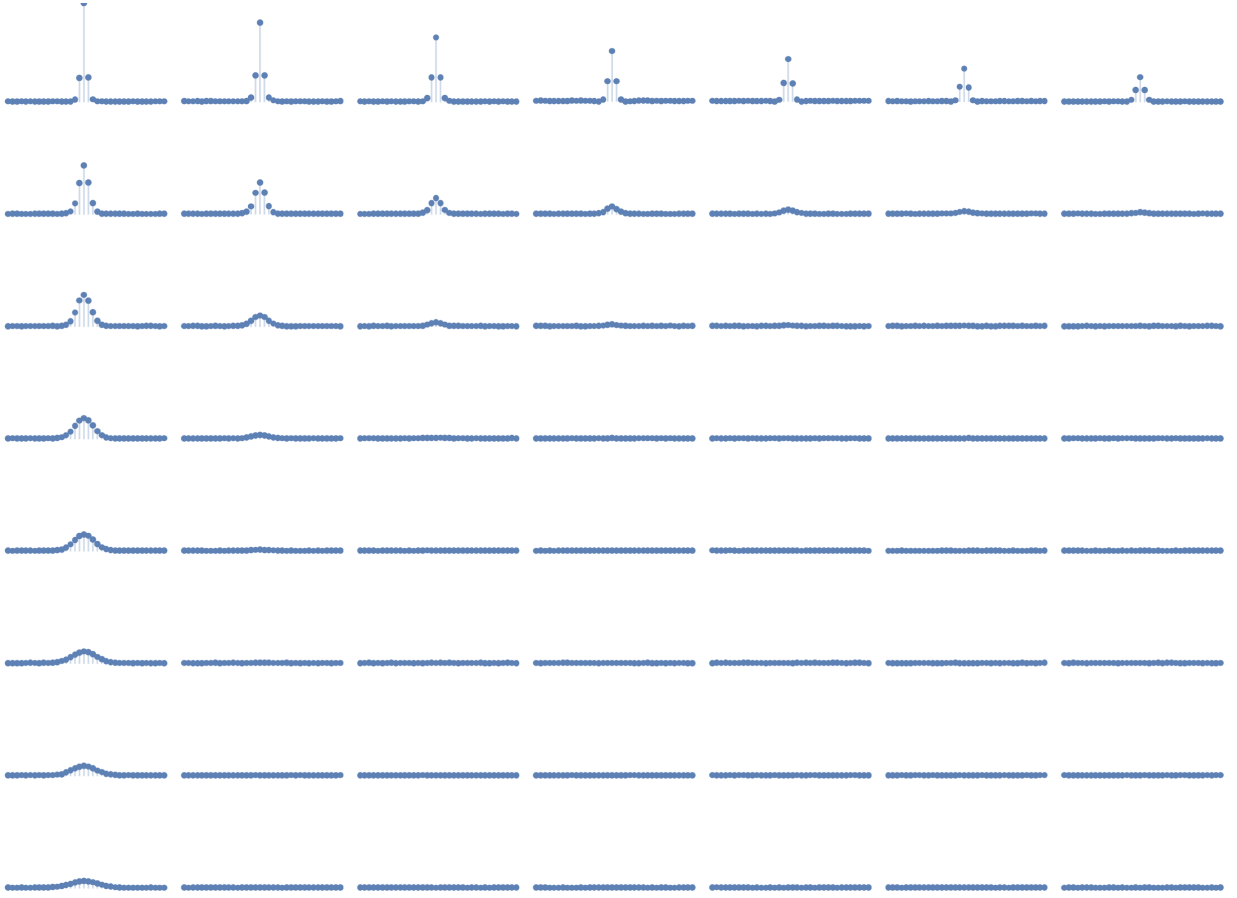


Figure 6.6: Data of the LE experiments. From left to right: $2a-b = 0, 1, 2, 3, 4, 5, 6 \mu s$. Top to bottom: (before scaling) $t = 96, 192, 288, 384, 480, 576, 672 \mu s$. For large t and $2a-b$ (right bottom corner), the signal-to-noise ratio goes to zero.

environmental dynamics is present, the correlation grows monotonically as the evolution

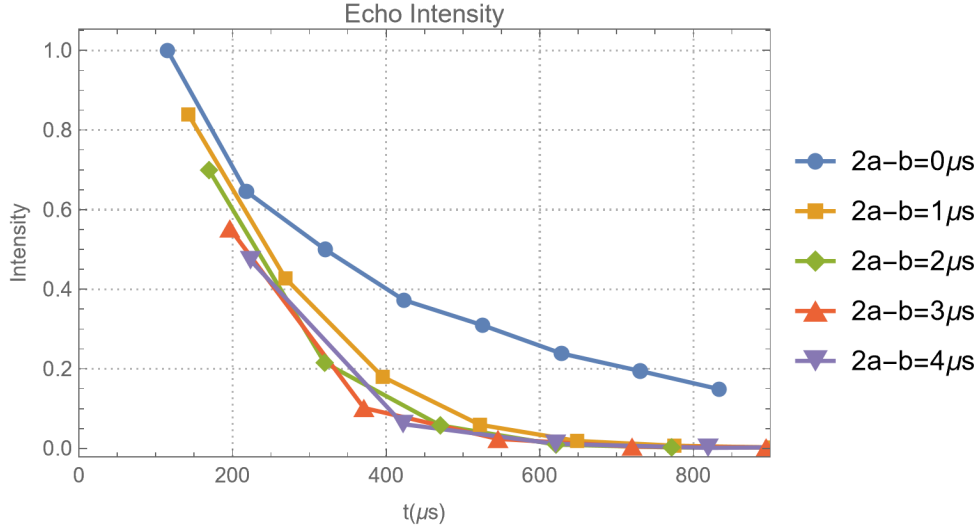


Figure 6.7: Echo intensity as functions of t . The decay becomes faster as the amount of mixing ($2a - b$) increases for large t ($\geq 200\mu s$).

time t increases. In the presence of the environmental dynamics, however, there is a limit to how much the CS can learn about the environment: For small t more correlation is built up as t increases. The correlation growth saturates earlier as more environmental dynamics is included. When t increases even more, the CS/environment correlation decreases and eventually becomes completely destroyed. To quantify the amount of correlation between the CS and the environment, we calculate the Shannon entropy of the distribution of the amplitudes (Fig. 6.8). The entropy provides a measure of effective size and learnability of the environment via the CS. When the environmental dynamics is nonzero, the entropy reaches a maximal value at intermediate t . However, the crossings between different lines suggest that the entropy can increase with increased environmental dynamics (Fig. 6.8 right). There could be a transient period when the environmental dynamics can aid growth of CS/environment correlation. Additionally, for short t and large s_e (e.g. the last four plots of the first row in Fig. 6.6), the distributions show non-convex patterns (small wiggles around the baseline), which was not observed in the absence of mixing.

Figs. 6.9 and 6.10 show the growth of coherence orders after accounting for the scaling factors ($s_{cs/e}$) of $H_{cs/e}$. In Fig. 6.10, there are multiple places where different lines cross. This means that, for the same correlation time t , nonzero mixing can indeed lead to more growth of higher-order coherence terms ($t < 200\mu s$).

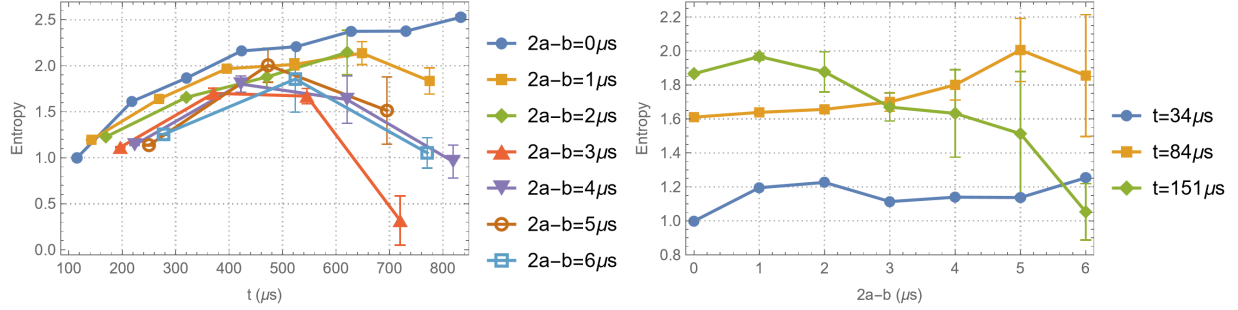


Figure 6.8: Entropy of coherence amplitudes of LE experiments. Left: Entropy as a function of t for different mixing ($2a-b$). Right: Entropy for fixed t as a function of $2a-b$.

6.3 Refocusing the environmental dynamics

To investigate how the environmental mixing affects quantum information in subspaces of different coherence order, consider the following experiments:

$$\rho(0) \xrightarrow{\exp(-i(H_{cs/e}t + H_e\Delta))} \rho(t, \Delta) \xrightarrow{\exp(-iJ_x\theta/2)} \rho(t, \Delta, \theta) \xrightarrow{\exp(i(H_{cs/e}t + H_e\Delta))} \text{measure}. \quad (6.6)$$

By refocusing the environmental dynamics, there is no loss of correlation.

The measured amplitudes of coherence orders are shown in Fig. 6.11 and their corresponding Shannon entropy is shown in Fig. 6.12. The mixing spreads the coherence over a larger space, resulting in a broader distribution of coherence orders and increased entropy. Due to higher-order errors in the sequence, the signal eventually decays to zero. The growth of coherence under mixing follows a different pattern compared with that under $H_{s/e}$. Notably, a non-convex shape emerges for large Δ (Fig. 6.13) since the mixing introduces a double-quantum term, which is more effective at populating higher-order coherences.

Compared with the results from the last section, the experiments suggest that mixing can aid in correlation growth by redistributing the correlations. In addition, Fig. 6.13 shows that the largest nonzero coherence order at $\Delta = 90\mu s$ is larger than that at $\Delta = 15\mu s$, indicating that the mixing has created correlation with spins that are initially not directly coupled to the CS.

6.4 Inferring the number of spins in the environment

The MQ experiments provide a way to measure the effective size of the environment seen by the CS. However, higher-order coherence can have weak intensity and can be a poor

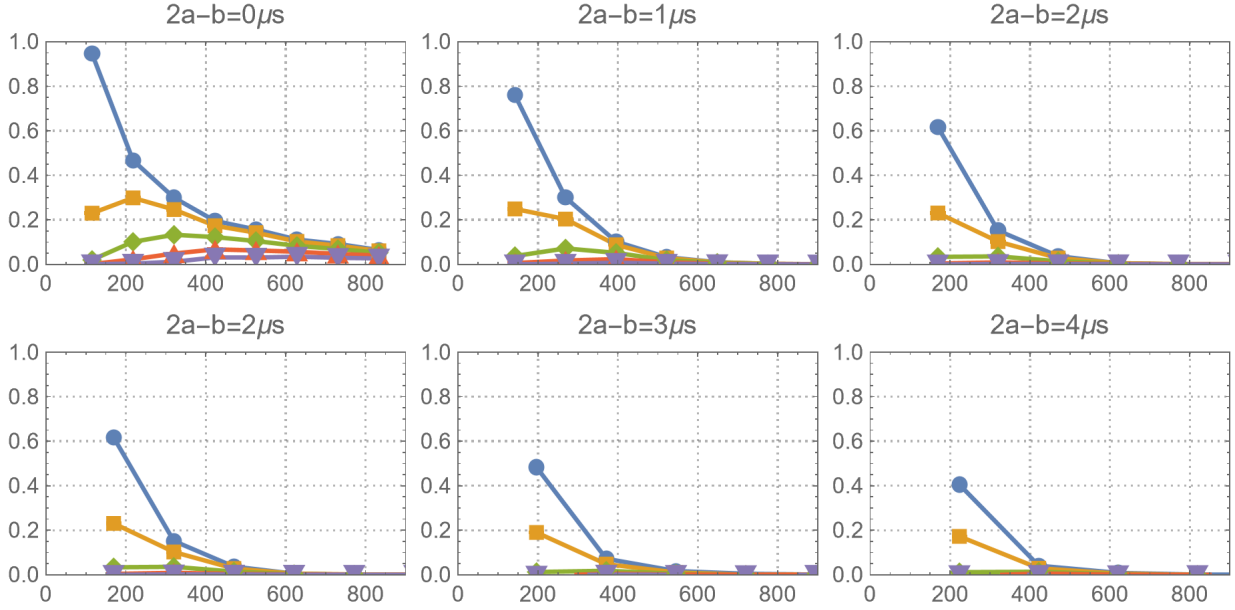


Figure 6.9: Intensity (vertical axes) of different coherence orders k as functions of evolution time t (μs) (horizontal axes) for different amount of mixing. $k = 0$ (blue dots), 1 (orange squares), 2 (green diamonds), 3 (red up triangles) and 4 (purple down triangles).

measure of N in the presence of measurement noise. Here, we show how to use Bayesian inference to estimate N . Instead of relying solely on the highest visible coherence order, this approach considers all information from the measurement including the shape of the distribution.

As an example, we infer the number of spins from the environment using the standard MQ experiment for $t = 672 \mu s$. We use the maximum likelihood estimation and a model selection with the Akaike information criterion (AIC). Figure 6.14 shows the best fits for different values of N , based on the measured distribution of coherence orders at $t = 672 \mu s$, with the corresponding AIC values plotted in Figure 6.15. The results suggest that, although the highest nonzero measured coherence order is around seven, the CS sees as many as fifteen spins. The result suggests that in order to replicate the observed data, a model with at least fifteen spins is required. Recall that each molecule of the sample (triphenylphosphine) has fifteen environment spins. Note that, however, this does not mean that all spins from the same molecule are correlated to the CS. Since the sample is in powder form, MQ experiments measure an average over molecules with different orientations. This provides more information than any single molecule, offering a more

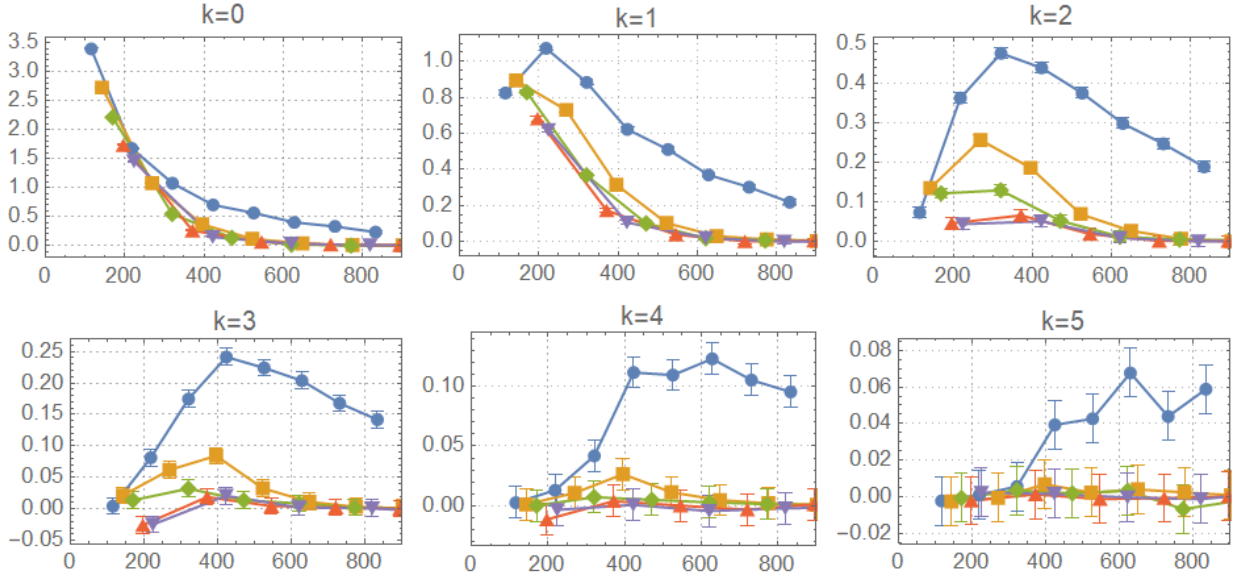


Figure 6.10: Intensity (vertical axes) of the same coherence order for different amount of mixing as functions of evolution time $t(\mu s)$. $2a - b = 0 \mu s$ (blue dots), $1 \mu s$ (orange squares), $2 \mu s$ (green diamonds), $3 \mu s$ (red up triangles) and $4 \mu s$ (purple down triangles). The crossings of the different lines (e.g., the blue and orange line for $k = 1$) indicate that the mixing aids in growth of coherence for short t .

complete view of the environment. However, some spins may remain uncorrelated with the CS if their interaction aligns with the magic angle.

6.5 Conclusion

In summary, this chapter demonstrates how multi-quantum experiments and their variations can be used to measure system/environment correlation growth in the central spin system and characterize the size of the environment. When acting as a perturbation, the environmental mixing can reduce the learnability of the environment via the CS. However, there is a transient period where mixing redistributes correlations between the CS and the environment, facilitating the build-up of higher-order coherence. This effect is more pronounced in the experiments when mixing is refocused. The results show that mixing helps the CS correlate with a larger portion of the environment that is not directly connected via $H_{cs/e}$. Bayesian inference is used to estimate the number of spins detectable by the CS.

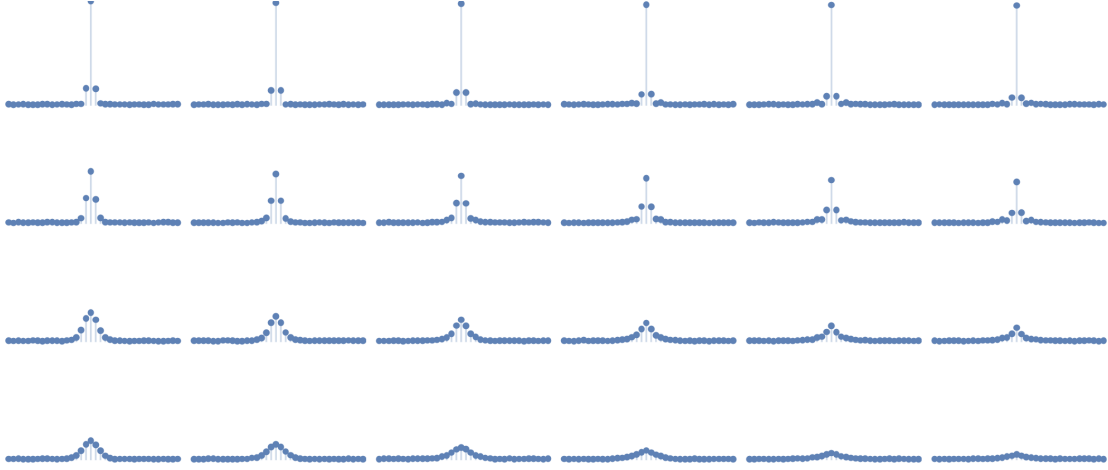


Figure 6.11: Measured amplitudes of different coherence orders for different t and mixing time Δ . From left to right: $\Delta = 15, 30, 45, 60, 75, 90 \mu s$. Top to bottom: $t = 72, 144, 216, 288 \mu s$. Increased mixing (Δ) introduces more structure to the distributions of coherence orders and spreads it toward higher-order coherence.

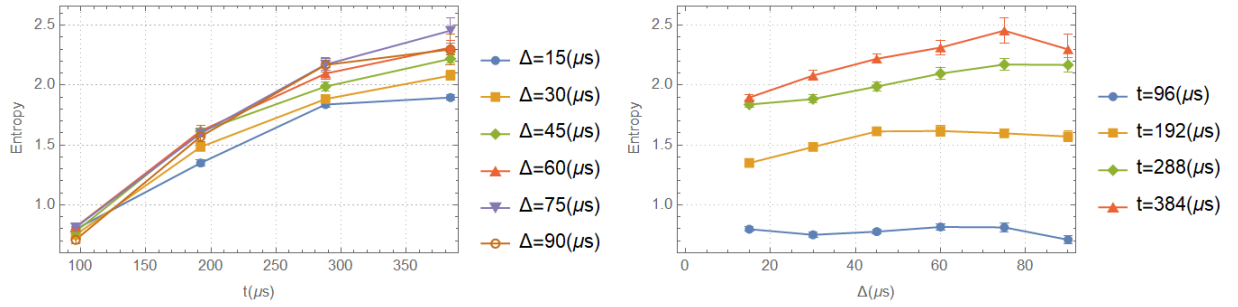


Figure 6.12: Measured entropy of distributions of coherence orders.

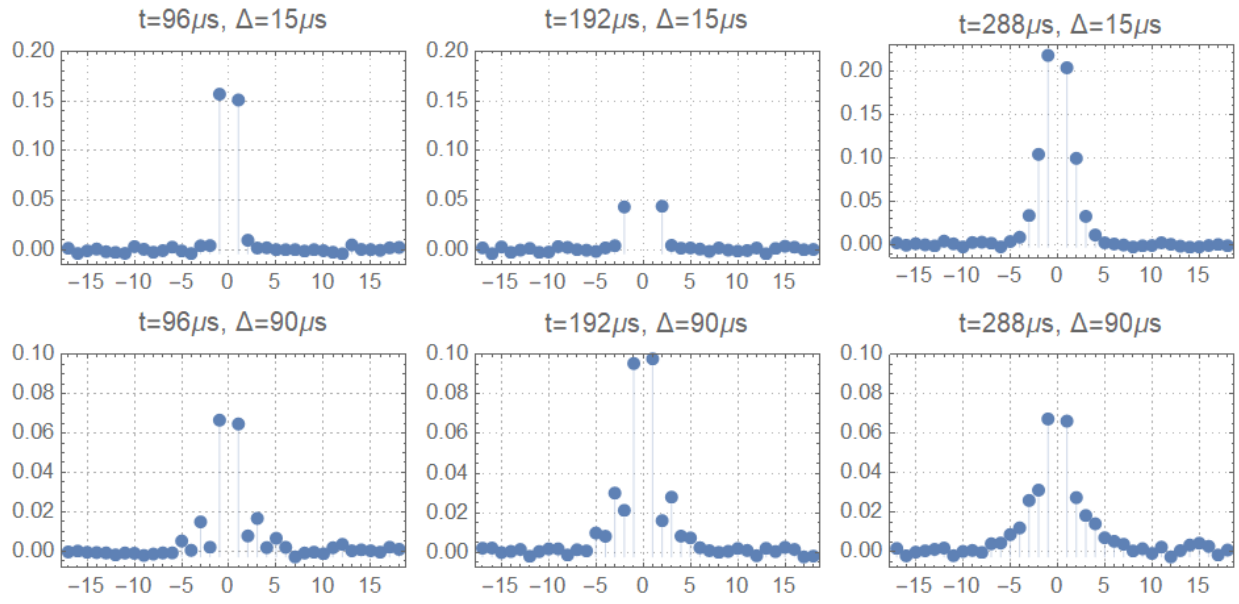


Figure 6.13: Zoomed-in versions of distributions of coherence orders. The bottom row exhibits a non-convex feature (the first and second rows) and longer tail (the last row) of the distributions.

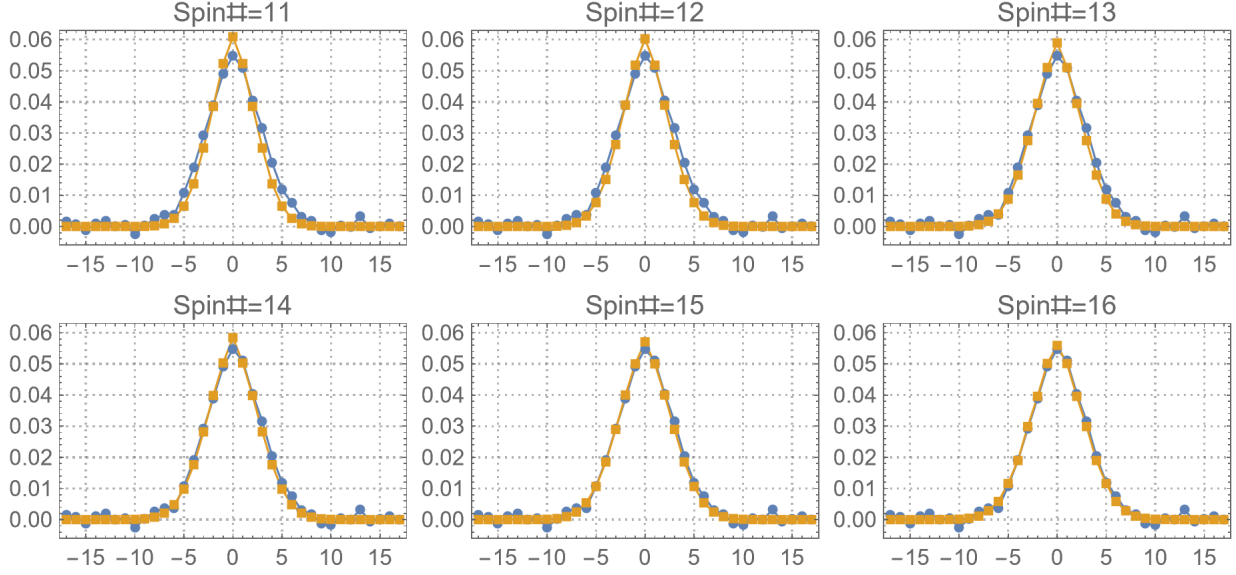


Figure 6.14: Best fit (orange squares) to the data (blue dots) for $t = 672 \mu s$. The improvement of the fitting of the last one is diminishing and not justified under the present uncertainty.

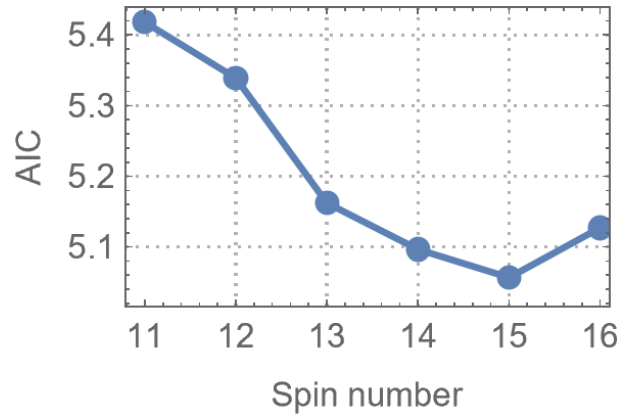


Figure 6.15: AIC for different number of spins N . It suggests possible over-fitting at $N = 16$.

The memory effects of the noise due to the environmental dynamics depend on the structure of the environment. The results here reveal an extended environment that can hold long memory and a transient period where mixing can aid in CS/environment correlation. The method allows us to map out information flows due to mixing between different subspaces of the environment. Understanding these flows can help us understand the origin of memory effects and design better noise reduction strategies.

Chapter 7

Conclusion and Future Work

7.1 Summary

This thesis provides tools to aid the development of coherence control methods for quantum information processing. The ultimate outcomes are a set of measurements for characterizing the quantum system and its environment, a general framework for designing robust and precise quantum control in the presence of uncertainties, and methods for validating system performance. Fig. 7.1 attempts to summarize the results from the thesis as a chart, showing tasks and interconnections. It is color-coded to indicate those tasks that are fully implemented in the thesis (black) and those that remain for future work (red boxes). Overall, integrating an environment with memory into pulse finding is the key future step.

In the first part (Chapters 2 and 3), we established a general framework for designing numerically optimized, precise, and robust quantum control sequences using AHT for a system with the Hamiltonian

$$H_{\text{tot}}(t) = H_c(\{a_k(t)\}; \{\mu_j\}) + H_{\text{int}}(\{\eta_{\text{int}}^i\}). \quad (7.1)$$

In Fig. 7.1, a control sequence, represented as a set of time-dependent control parameters $(\{a_k(t)\})$, is mapped via a control system to a continuous output field $\vec{B}(t)$ acting on the quantum system (arrows A and B in Fig. 7.1). The control field achieves a desired unitary target and/or an effective Hamiltonian target. Algorithms were introduced to characterize the set/space of achievable unitaries and effective Hamiltonian at each order from the Magnus expansion, including how to minimize higher-order corrections and cross terms.

The method incorporates robustness against variations in the control system (arrow C) and the internal Hamiltonian (arrow D), and is independent of the modality of the control system. All tasks are fully reduced to pseudocodes/flowcharts, enabling all researchers to implement them.

Chapters 4, 5, and 6 used the stimulated echo (STE), multi-dimensional (MD), and multi-quantum (MQ) experiments to study the system environment correlation and environmental memory effects using solid-state NMR experiments in a central spin (CS) system ($H = H_{cs/e} + H_e$). The STE and MD experiments measure CS/environment correlation and environmental mixing (arrows E). The STE experiments provide efficient measures of growth of CS/environment correlations and environmental mixing by measuring the auto-correlation functions and commutators between $H_{cs/e}$ and H_e . The MD experiments give a more complete investigation of the environmental dynamics by producing the multi-time joint probability distributions of the local field. The experiments quantify non-Gaussianity, non-stationarity, and non-Markovianity of the environmental dynamics. The MQ experiments explore how the environmental mixing moves information within the environment and reveal an extended environment that can hold long-lived memory (arrow F).

7.2 Future work

An important future step is to use the information learned from these experiments to provide a better model of decoherence (arrows G and H) which can aid the design of better control (arrow I).

The experiment methods developed in the thesis can be extended to characterize noise in a two-central-spin (2CS) system. STE and MD experiments can be used to characterize CS/environment correlation growth, environmental mixing, and memory effects of the environment with respect to each CS (arrows J and K). There could be information flow between the CSs through the environment, resulting in a correlation between the joint distributions of the noise. Bayesian hierarchical modeling may be used to capture such correlation. The information can potentially improve the design of two-qubit controls. The environment is divided into different parts that have different connectivity with the CSs. The difference in connectivity can lead to different abilities of scrambling system/environment correlations and holding memory. This difference can affect performance in two-qubit gates on the 2CS system. MQ experiments can be extended to the 2CS system to track the coherence orders of the correlations between each CS and the environment and potentially be used to study the structure of the environment (arrow K).

The 2CS system can also be used to explore environmental influence on different quantum states and measure the symmetry of the environment. This information may aid in the design of symmetry-dependent noise reduction strategies, e.g., quantum error correction [12, 100], decoherence-free subspaces [120], and noiseless subsystems [101, 63] (arrow L).

The memory effects of noise in an open quantum system can affect the reliability of randomized benchmarking [102, 73], which is a standard tool, originally designed for Markovian noise, for characterizing the performance of a gate set. It is important to understand its robustness against memory effects and how it can be used to measure correlated noise (arrow M).

Simulation of quantum operations is important for understanding fault-tolerant properties of quantum circuits. Markovian noise can be efficiently approximated using Monte Carlo simulation [126]. If the noise is non-Markovian, however, ancillary qubits may be needed to store the memory of the process. It would be useful to see how to faithfully represent a non-Markovian noise and design efficient simulations (arrow N).

Appendix A

Supplemental information: Engineering Nonzero Effective Hamiltonians

A.1 Calculating $\mathcal{C}(\mathbf{g}_{\text{pri}}, H_{\text{pert}})$

Here, a theorem that enables the calculation of $\mathcal{C}(\mathbf{g}_{\text{pri}}, H_{\text{pert}})$ using Algorithm 2 in the main text is proved. Next, we discuss the representation for calculating $\mathcal{C}(\mathbf{g}_{\text{pri}}, H_{\text{pert}})$ and how to engineer multiple H_{pert}^w in parallel.

Define the adjoint map $\text{ad}_A(B) = [A, B]$ for $\forall$ (anti-)Hermitian operators A, B acting on the total Hilbert space $\mathcal{H}$.

Theorem A.1.1. *For a Hamiltonian H_{pert} , define*

$$\mathcal{C}(\mathbf{g}_{\text{pri}}, H_{\text{pert}}) = \text{span}_{\mathbb{R}} \{[g_1, \dots [g_L, H_{\text{pert}}] \dots] | g_l \in \mathbf{g}_{\text{pri}}, L \geq 0\}, \quad (\text{A.1})$$

and

$$\mathcal{O}(\mathbf{g}_{\text{pri}}, H_{\text{pert}}) = \text{span}_{\mathbb{R}} \{U_c^\dagger(t) H_{\text{pert}} U_{\text{pri}}(t) dt | U_{\text{pri}}(t') \in e^{\mathbf{g}_{\text{pri}}}\}. \quad (\text{A.2})$$

Then $\mathcal{C}(\mathbf{g}_{\text{pri}}, H_{\text{pert}}) = \mathcal{O}(\mathbf{g}_{\text{pri}}, H_{\text{pert}})$.

Proof. Consider piecewise-continuous $U(t)$. For all $U(t) \in e^{\mathbf{g}_{\text{pri}}}$, there exists $g_t \in \mathbf{g}_{\text{pri}}$, such that $U(t) = e^{g_t}$. Therefore, using $e^{-g_t} H_{\text{pert}} e^{g_t} = e^{-\text{ad}_{g_t}}(H_{\text{pert}}) \in \mathcal{C}(\mathbf{g}_{\text{pri}}, H_{\text{pert}})$. Therefore, $\mathcal{O}(\mathbf{g}_{\text{pri}}, H_{\text{pert}}) \subset \mathcal{C}(\mathbf{g}_{\text{pri}}, H_{\text{pert}})$.

Since $\mathcal{H}$ is finite-dimensional, there exists $m < |\mathcal{H}|^2 - 1$ such that $\{H_{\text{pert}}, \text{ad}_g(H_{\text{pert}}), \dots, \text{ad}_g^{m+1}(H_{\text{pert}})\}$ is linearly dependent $\forall g \in \mathfrak{g}_{\text{pri}}$. Assume that $\{H_{\text{pert}}, \text{ad}_g(H_{\text{pert}}), \dots, \text{ad}_g^m(H_{\text{pert}})\}$ is linearly independent, $e^{gt_i} H_{\text{pert}} e^{-gt_i}$ is in

$$\text{span}_{\mathbb{R}}\{H_{\text{pert}}, \text{ad}_g(H_{\text{pert}}), \dots, \text{ad}_g^m(H_{\text{pert}})\}.$$

Therefore, we can choose t_i for $i = 1, \dots, m$ such that

$$[g, H_{\text{pert}}] \in \text{span}_{\mathbb{R}}\{e^{gt_i} H_{\text{pert}} e^{-gt_i} | i = 1, \dots, m\}, \quad (\text{A.3})$$

$\forall g_1, g_2 \in \mathfrak{g}_{\text{pri}}$, we have

$$\begin{aligned} [g_2, [g_1, H_{\text{pert}}]] &= \sum_{j=1}^s a_j^2 e^{g_2 t'_j} [g_1, H_{\text{pert}}] e^{-g_2 t'_j} \\ &= \sum_{j=1}^s a_j^2 e^{g_2 t'_j} \left(\sum_{i=1}^m a_i^1 e^{g_1 t_i} H_{\text{pert}} e^{-g_1 t_i} \right) e^{-g_2 t'_j} \\ &= \sum_{j=1}^s \sum_{i=1}^m a_j^2 a_i^1 e^{g_{ij}} H_{\text{pert}} e^{-g_{ij}}, \end{aligned} \quad (\text{A.4})$$

for some $g_{ij} \in \mathfrak{g}_{\text{pri}}$. The last equality is due to the Baker–Campbell–Hausdorff theorem. Therefore, by induction, $[g_1, \dots, [g_L, H_{\text{pert}}]] \in \mathcal{O}(\mathfrak{g}_{\text{pri}}, H_{\text{pert}}) \forall g_1, \dots, g_L \in \mathfrak{g}_{\text{pri}}$. Thus, $\mathcal{C}(\mathfrak{g}_{\text{pri}}, H_{\text{pert}}) \subset \mathcal{O}(\mathfrak{g}_{\text{pri}}, H_{\text{pert}})$.

Combining the result with the first part of the proof, we have $\mathcal{C}(\mathfrak{g}_{\text{pri}}, H_{\text{pert}}) = \mathcal{O}(\mathfrak{g}_{\text{pri}}, H_{\text{pert}})$. $\square$

For time-dependent $H_{\text{pert}}(t)$, the space is defined as $\mathcal{C}(\mathfrak{g}_{\text{pri}}, \mathfrak{g}_{\text{pert}}) = [\mathfrak{g}_{\text{pri}}, \mathfrak{g}_{\text{pert}}]$, where $\mathfrak{g}_{\text{pert}}$ is the minimal Lie algebra containing $H_{\text{pert}}(t)$.

It is convenient to choose a representation ϕ of $\mathfrak{g}_{\text{pri}}$ and H_{pert} that is restricted to a subspace of $\mathcal{H}$ to calculate $\mathcal{C}(\mathfrak{g}_{\text{pri}}, H_{\text{pert}})$. ϕ needs to preserve the structure of $\mathcal{C}(\mathfrak{g}_{\text{pri}}, H_{\text{pert}})$, namely

$$[\phi(g), \phi(h_1)] = \phi(h_2) \iff [g, h_1] = h_2, \quad (\text{A.5})$$

for all $g \in \mathfrak{g}_{\text{pri}}$ and $h_1, h_2 \in \mathcal{C}(\mathfrak{g}_{\text{pri}}, H_{\text{pert}})$. The representation is bijective and unique up to isomorphism.

When a finite-dimensional representation does not exist, sometimes we can decompose H_{pert} into components and define a finite-dimensional representation for each component. An example is a qubit network with different clusters that are controlled independently (Fig. A.1, ref. Sec. A.10.6).

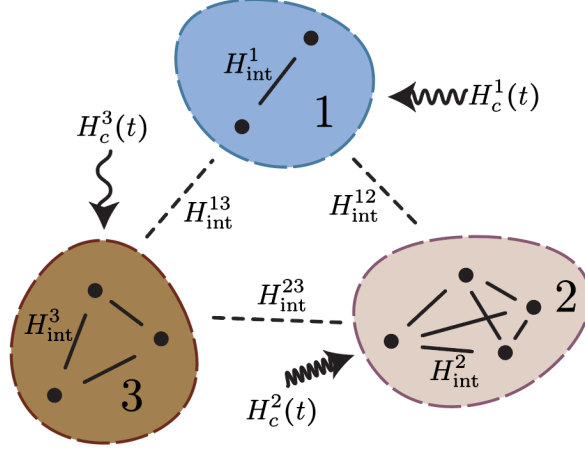


Figure A.1: A network of three clusters. Different clusters are controlled differently (not necessarily independently) and have their own internal Hamiltonians.

A.2 Finding the Set of Achievable H_{eff}

In this section, we first prove that the set of achievable $\bar{H}^{(0)}$ is the convex hull of $\{U_i^\dagger H_{\text{pert}} U_i\}$, $U_i \in e^{\mathbf{g}_{\text{pri}}}$. Then, the formal version of Algorithm 3 in the main text is given. Finally, the result is generalized to characterizing the range of achievable scaling factors when multiple H_{pert}^w are being engineered.

Theorem A.2.1. *Define*

$$\mathcal{O}'(\mathbf{g}_{\text{pri}}, H_{\text{pert}}) = \left\{ \frac{1}{T_{\text{seq}}} \int_0^{T_{\text{seq}}} dt |U^\dagger(t) H_{\text{pert}} U(t)\rangle\right\rangle \left| U(t) \in e^{\mathbf{g}_{\text{pri}}} \right\}, \quad (\text{A.6})$$

$$\mathcal{Q}(\mathbf{g}_{\text{pri}}, H_{\text{pert}}) = \left\{ |U^\dagger H_{\text{pert}} U\rangle\right\rangle | U \in e^{\mathbf{g}_{\text{pri}}} \right\}. \quad (\text{A.7})$$

Denote the convex hull of $\mathcal{Q}(\mathbf{g}_{\text{pri}}, H_{\text{pert}})$ as $\bar{\mathcal{Q}}(\mathbf{g}_{\text{pri}}, H_{\text{pert}})$. Assume that any $U \in e^{\mathbf{g}_{\text{pri}}}$ can be implemented in a time t_p much shorter than $1/|H_{\text{pert}}|$, then $\mathcal{O}'(\mathbf{g}_{\text{pri}}, H_{\text{pert}}) = \bar{\mathcal{Q}}(\mathbf{g}_{\text{pri}}, H_{\text{pert}})$.

Proof. An element from $\mathcal{Q}(\mathbf{g}_{\text{pri}}, |H_{\text{pert}}\rangle\right\rangle)$ takes the form $\sum_{i=1}^M p_i |U_i^\dagger H_{\text{pert}} U_i\rangle\right\rangle$ for a positive

integer M and probabilities $\{p_i\}$. Let

$$g(t) = \begin{cases} \frac{g_i}{t_p} & t \in [0, t_p) \\ 0 & t \in [t_p, p_i T_{\text{seq}} - t_p) \\ -\frac{g_i}{t_p} & t \in [p_i T_{\text{seq}} - t_p, p_i t] \end{cases}, \quad (\text{A.8})$$

for a fixed set of $\{p_i\}$. Let $U(t) = \mathcal{T} \exp \left[\int_0^t dt' g(t') \right]$. Then

$$\lim_{t_p \rightarrow 0^+} \frac{1}{T_{\text{seq}}} \int_0^{T_{\text{seq}}} |U^\dagger(t) H_{\text{pert}} U(t)\rangle\rangle dt = \sum_{i=1}^M p_i |e^{-g_i} H_{\text{pert}} e^{g_i}\rangle\rangle. \quad (\text{A.9})$$

Therefore, $\bar{\mathcal{Q}}(\mathbf{g}_{\text{pri}}, H_{\text{pert}}) \subset \mathcal{O}'(\mathbf{g}_{\text{pri}}, H_{\text{pert}})$. For all $U(t) \in e^{\mathbf{g}_{\text{pri}}}$,

$$\begin{aligned} & \frac{1}{T_{\text{seq}}} \int_0^{T_{\text{seq}}} |U^\dagger(t) H_{\text{pert}} U(t)\rangle\rangle dt \\ &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \left| U^\dagger \left(\frac{i T_{\text{seq}}}{N} \right) H_{\text{pert}} U \left(\frac{i T_{\text{seq}}}{N} \right) \right\rangle\rangle \in \bar{\mathcal{Q}}(\mathbf{g}_{\text{pri}}, H_{\text{pert}}). \end{aligned} \quad (\text{A.10})$$

The limit is valid if $t_p \ll 1/|H_{\text{pert}}|$. Thus, $\mathcal{O}'(\mathbf{g}_{\text{pri}}, H_{\text{pert}}) \subset \bar{\mathcal{Q}}(\mathbf{g}_{\text{pri}}, H_{\text{pert}})$. Therefore, $\mathcal{O}'(\mathbf{g}_{\text{pri}}, H_{\text{pert}}) = \bar{\mathcal{Q}}(\mathbf{g}_{\text{pri}}, H_{\text{pert}})$. $\square$

If one can implement

$$\mathcal{T} \exp \left(-i \int_0^t H_{\text{tot}}(t') dt' \right) = \exp(-i H_{\text{pert}} t) U_t, \quad (\text{A.11})$$

for all $t > 0$ and some $U_t \in e^{\mathbf{g}_{\text{pri}}}$, e.g., $H_{\text{pri}}(t) = H_c(t)$ and $H_{\text{pert}} = H_{\text{int}}$, $H_{\text{pri}}(t) = H_c(t) + p H_{\text{int}}$ and $H_{\text{pert}} = (1-p) H_{\text{int}}$, and $H_{\text{pri}}(t) = H_c(t)$ and $H_{\text{pert}}(t) = \epsilon H_c(t)$, the condition $t_p \ll 1/|H_{\text{pert}}|$ is unnecessary. Finite t_p , however, may limit the efficiency of a sequence.

For any $H_{\text{target}} \in \mathcal{C}$, if there exists a sequence $\{a_k(t)\}$ so that $\bar{H}^{(0)}(T_{\text{seq}}) = s H_{\text{target}}$ for a real number s (or equivalently, $s H_{\text{target}} \in \bar{\mathcal{Q}}(\mathbf{g}_{\text{pri}}, H_{\text{pert}})$), then s is an achievable scaling factor for H_{target} . $\bar{\mathcal{Q}}(\mathbf{g}_{\text{pri}}, H_{\text{pert}})$ can be approximated using random unitaries $\{u_j\}$:

$$\bar{\mathcal{Q}}(\mathbf{g}_{\text{pri}}, H_{\text{pert}}) \approx \left\{ \sum_{j=1}^J x_j |u_j^\dagger H_{\text{pert}} u_j\rangle\rangle \left| \sum_{j=1}^J x_j = 1, x_j \geq 0, 1 \leq j \leq J \right. \right\}. \quad (\text{A.12})$$

Choose a matrix T such that $T|H_{\text{target}}\rangle\rangle = (1, 0, 0, \dots, 0)$ and define a set of vertices $\{\tilde{v}^j = T|u_j^\dagger H_{\text{pert}} u_j\rangle\rangle\}$, the range of achievable scaling factors, $[s_-, s_+]$, can be found by

solving the following linear programs

$$\begin{aligned}
\text{LP}_{\pm}: \quad & \text{Find a vector } \vec{x} \\
& \text{that maximizes } s_{\pm} = \vec{c}_{\pm} \cdot \vec{x} \text{ with } \vec{c}_{\pm} = (\pm 1, 0, \dots, 0) \\
& \text{subject to } \sum_i x_i = 1, -1 \leq \sum_i v_1^i x_i \leq 1 \\
& \text{and } \sum_i v_k^i x_i = 0, k = 2, 3, \dots \\
& \text{and } \vec{x} \geq 0
\end{aligned} \tag{A.13}$$

H_{target} is not achievable if the constraint is infeasible (Fig. A.2). The procedure is summarized as the following algorithm.

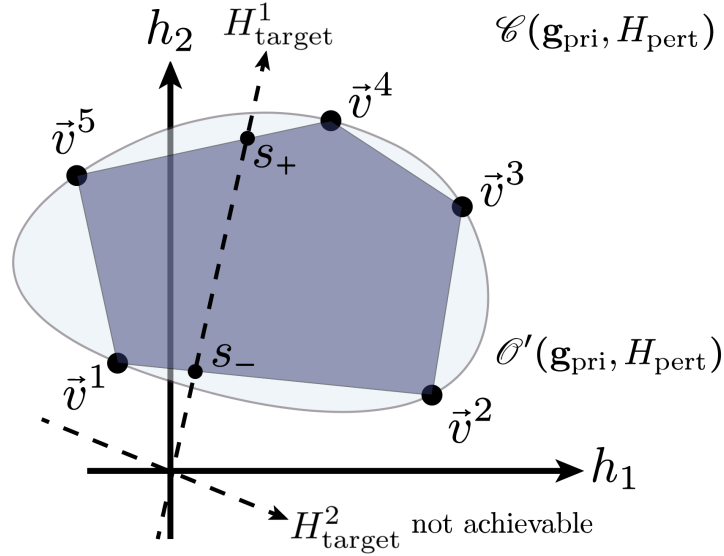


Figure A.2: Schematic for finding the range of achievable scaling factors. The set of achievable zeroth-order average Hamiltonians is a convex set in $\mathcal{C}(\mathbf{g}_{\text{pri}}, H_{\text{pert}})$. By sampling random points from the set of extreme points $\mathcal{Q}(\mathbf{g}_{\text{pri}}, H_{\text{pert}})$, the convex set can be approximated with a polygon. The achievable scaling factors for a desired H_{target} can then be found using linear programming.

Algorithm 4 Algorithm for finding achievable scaling factors. The function **FINDSCALE** takes a basis of the primary Lie algebra $\mathcal{B}(\mathbf{g}_{\text{pri}})$, a perturbative Hamiltonian H_{pert} and a target effective Hamiltonian H_{target} and returns either False when H_{target} is not achievable or the range of achievable scaling factors $[s_-, s_+]$.

```

1: function FINDSCALE( $\mathbf{g}_{\text{pri}}, H_{\text{pert}}$ )
2:   Calculate  $\mathcal{B}(\mathcal{C})$ 
3:   find  $T$  such that  $T|H_{\text{target}}\rangle\rangle = (1, 0, \dots, 0)$ 
4:   for  $j = 1, \dots, J$  do
5:     Sample  $u_j \in e^{\mathbf{g}_{\text{pri}}}$ 
6:      $\tilde{v}^j \leftarrow T|u_j^\dagger H_{\text{pert}} u_j\rangle\rangle$ 
7:   end for
8:   Solve the linear programs  $\text{LP}_+$  and  $\text{LP}_-$ 
9:   if no solution exists then
10:    return False
11:  else
12:    return  $[s_-, s_+]$ 
13:  end if
14: end function

```

When

$$H_{\text{pert}} = \sum_{w=1}^W H_{\text{pert}}^w, \quad (\text{A.14})$$

the condition for parallel effective Hamiltonian engineering is

$$\bigoplus_{m=1}^W |\bar{H}^{(0)w}\rangle\rangle_w = \bigoplus_{w=1}^W |H_{\text{target}}^w\rangle\rangle_w \equiv |H_{\text{target}}\rangle\rangle_{\text{total}}, \quad (\text{A.15})$$

where $|\cdot\rangle\rangle_w$ is the representation in $\mathcal{C}_w(\mathbf{g}_{\text{pri}}, H_{\text{pert}}^w)$. We have

Theorem A.2.2. *Define*

$$\begin{aligned} \mathcal{O}'(\mathbf{g}_{\text{pri}}, \{H_{\text{pert}}^w\}) &= \left\{ \frac{1}{T_{\text{seq}}} \int_0^{T_{\text{seq}}} dt \bigoplus_{w=1}^W |U_{\text{pri}}^\dagger(t) H_{\text{pert}}^w U_{\text{pri}}(t)\rangle\rangle_w \left| U_{\text{pri}}(t) \in e^{\mathbf{g}_{\text{pri}}} \right. \right\}, \\ \mathcal{Q}(\mathbf{g}_{\text{pri}}, \{H_{\text{pert}}^w\}) &= \left\{ \bigoplus_{w=1}^W |U_{\text{pri}}^\dagger H_{\text{pert}}^w U_{\text{pri}}\rangle\rangle_w \left| U_{\text{pri}} \in e^{\mathbf{g}_{\text{pri}}} \right. \right\}. \end{aligned} \quad (\text{A.16})$$

Then $\mathcal{O}'(\mathbf{g}_{\text{pri}}, \{H_{\text{pert}}^w\}) = \bar{\mathcal{Q}}(\mathbf{g}_{\text{pri}}, \{H_{\text{pert}}^w\})$.

Find T such that $T|H_{\text{target}}\rangle_{\text{total}} \propto (1, 0, \dots, 0)$, the vertices of the polygon are

$$\vec{v}^j = T \bigoplus_{w=1}^W |u_j^\dagger H_{\text{pert}}^w u_j\rangle_w. \quad (\text{A.17})$$

Corollary A.2.1. $\mathcal{O}'(\mathbf{g}_{\text{pri}}, \{H_{\text{pert}}^w\}_{w=1}^W | H_{\text{target}}^1 = 0) = \mathcal{O}'(\mathbf{g}_{\text{pri}}, \{H_{\text{pert}}^w\}_{w=2}^W).$

Namely, decoupling of any part of H_{pert} does not affect controllability of other parts.

A.3 Generating random points for controllability test

Here, we present algorithms for sampling random unitaries from $e^{\mathbf{g}_{\text{pri}}}$. For standard groups such as the unitary groups, Haar random samples can be generated using the QR decomposition [133].

Algorithm 5 Algorithm for generating random unitaries based on QR decomposition.

- 1: **function** RANDOMUQR(n)
 - 2: generate two random real $n \times n$ matrices r_1 and r_2 whose elements are sampled from $\mathcal{N}(0, 1)$
 - 3: $z \leftarrow (r_1 + ir_2)/\sqrt{2}$
 - 4: use QR decomposition to decompose z into $z = QR$ where Q is unitary and R is upper-triangle
 - 5: take signs of the diagonal of R and form a diagonal matrix Λ
 - 6: **return** $Q\Lambda$
 - 7: **end function**
-

More generally, we generate random unitaries from $e^{\mathbf{g}_{\text{pri}}}$ using a random walk algorithm [116].

Algorithm 6 Algorithm for generating random unitaries based on random walk. Fix step length ϵ_{step} . For the second function, instead of using thin, we can sample $N > n_{\text{sample}}$ then sample n_{sample} uniformly at random from N . They both destroy correlations between samples. This approach is more effective but less efficient and less simple compared to thinning.

```

1: function RANDOMULIE( $\mathcal{B}(\mathbf{g}_{\text{pri}})$ )
2:    $g \leftarrow 0$ 
3:   for  $h \in \mathcal{B}(\mathbf{g}_{\text{pri}})$  do
4:     sample  $p$  from  $\mathcal{N}(0, 1)$ 
5:      $g \leftarrow g + ph$ 
6:   end for
7:   return  $e^g$ 
8: end function
9: function RANDOMWALKU( $\mathcal{B}(\mathbf{g}_{\text{pri}}), n_{\text{burn}}, n_{\text{thin}}, n_{\text{sample}}$ )
10:   $x \leftarrow \mathbb{1}$ 
11:  for  $1 \leq i_{\text{burn}} \leq n_{\text{burn}}$  do
12:     $x \leftarrow \text{RANDOMULIE}(B(\mathbf{g}_{\text{pri}})) \cdot x$ 
13:  end for
14:   $l_{\text{sample}} \leftarrow \{\}$ 
15:  for  $1 \leq i_{\text{sample}} \leq n_{\text{sample}}$  do
16:    for  $1 \leq i_{\text{thin}} \leq n_{\text{thin}}$  do
17:       $x \leftarrow \text{RANDOMULIE}(B(\mathbf{g}_{\text{pri}})) \cdot x$ 
18:    end for
19:    append  $x$  to  $l_{\text{sample}}$ 
20:  end for
21:  return  $l_{\text{sample}}$ 
22: end function

```

The convergence of the algorithm can be checked using, e.g., the coupling method. Alternatively, one can check the volume of the simulated $\mathcal{Q}$. Fig. A.3 shows the volume of the simulated $\mathcal{Q}$ of the secular dipoler interaction under collective control of the qubits (Eqs. (A.97)). The volume is calculated using the quick hull algorithm [8, 197]. We note that the simulated range of achievable scaling factors usually converges more slowly and should be checked independently.

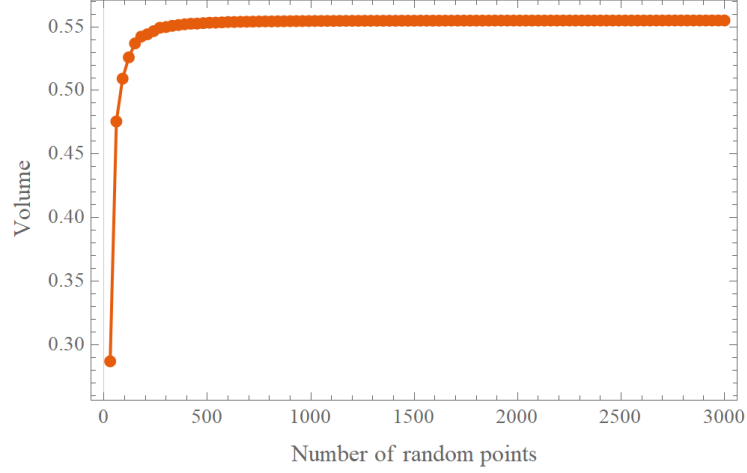


Figure A.3: Volume of the simulated $\bar{\mathcal{Q}}$ as a function of the number of samples. It saturates after around 500 samples.

A.4 Discretization of $\vec{B}(t)$ for Integration

The control parameters $a_k(t)$ are typically piecewise-constant:

$$a_{k,n} = a_k[(n - 1/2)dt], \quad (\text{A.18})$$

for $n = 1, \dots, N$. Here, an equal step length dt is assumed for the discretization for simplicity. Denote components of $\vec{B}(t)$ in rotation frames of the qubits by $\{b_k(t)\}$, they can be discretized into M steps (Fig. A.4) and the value at the m th step $b_{k,m}$ is approximated by the value of $b_k(t)$ at the middle of the step

$$b_{l,m} = b_k[(m - 1/2)\delta t], \quad (\text{A.19})$$

so that $Ndt = M\delta t = T_{\text{seq}}$. M should be large enough in order to accurately represent $\vec{B}(t)$. The approximation (Eq. (A.19)) is reasonable when the time-dependence of $b_l(t)$ is much slower compared to δ . When the assumption does not hold, a better approximation is given by

$$b_{l,m} = \frac{1}{\delta t} \int_{(m-1)\delta t}^{m\delta t} b_l(t) dt. \quad (\text{A.20})$$

The discretization enables straightforward calculation of the $\mathcal{C}$ -integrals (defined below) by first computing them for each step and then composing them using a chain rule.

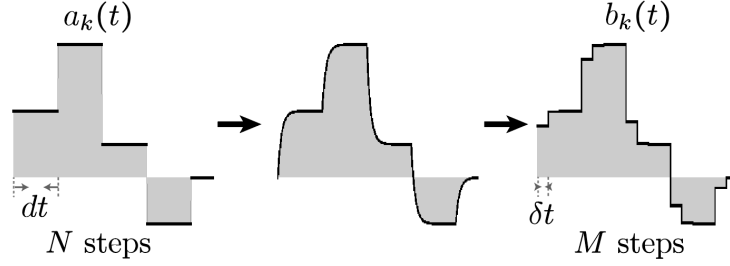


Figure A.4: Illustration of discretization of output field parameters. The control parameters $\{a_k(t)\}$ are mapped to $\vec{B}(t)$ through the model of the control system. Then the continuous field $\vec{B}(t)$ is discretized into M steps for integration.

A.5 Calculating $\mathcal{C}$ -integrals for shaped pulse

In this section, we show how to calculate the time-ordered integrals ($\mathcal{C}$ -integrals)

$$\bar{c}_{i_1 \dots i_r}(T_{\text{seq}}) = \int_0^{T_{\text{seq}}} dt_1 \dots \int_0^{t_{r-1}} dt_r c_{i_1}(t_1) \dots c_{i_r}(t_r). \quad (\text{A.21})$$

The $\mathcal{C}$ -integrals are used for achieving robustness and accounting for higher-order corrections. First, we need to represent the toggling-frame Hamiltonian and the primary unitaries in $\mathcal{C}(\mathbf{g}_{\text{pri}}, H_{\text{pert}})$. Given an orthonormal basis of $\mathcal{C}(\mathbf{g}_{\text{pri}}, H_{\text{pert}})$

$$\mathcal{B}(\mathcal{C}) = \{h_1, \dots, h_{|\mathcal{C}|}\}, \quad (\text{A.22})$$

where $|\mathcal{C}|$ is the dimension of $\mathcal{C}(\mathbf{g}_{\text{pri}}, H_{\text{pert}})$, for a Hamiltonian H , a unitary propagator U and a generator g of the unitary $U = e^g$ acting on $\mathcal{C}$, their representations are denoted by $|H\rangle\rangle$, $\mathcal{D}(U)$ and $\mathcal{D}(\text{ad}_g)$ respectively. They obey

$$|UHU^\dagger\rangle\rangle = \mathcal{D}(U)|H\rangle\rangle, \quad \mathcal{D}(e^g) = e^{\mathcal{D}(\text{ad}_g)}, \quad (\text{A.23})$$

and can be calculated as

$$\begin{aligned} |H\rangle\rangle &= \begin{pmatrix} \text{Tr}(Hh_1^\dagger) \\ \vdots \\ \text{Tr}(Hh_{|\mathcal{C}|}^\dagger) \end{pmatrix}, \quad \mathcal{D}(U) = \begin{pmatrix} \langle\langle h_1 | U h_1 U^\dagger \rangle\rangle & \dots & \langle\langle h_1 | U h_{|\mathcal{C}|} U^\dagger \rangle\rangle \\ \vdots & \ddots & \vdots \\ \langle\langle h_{|\mathcal{C}|} | U h_1 U^\dagger \rangle\rangle & \dots & \langle\langle h_{|\mathcal{C}|} | U h_{|\mathcal{C}|} U^\dagger \rangle\rangle \end{pmatrix}, \\ \mathcal{D}(\text{ad}_g) &= \begin{pmatrix} \langle\langle h_1 | [g, h_1] \rangle\rangle & \dots & \langle\langle h_1 | [g, h_{|\mathcal{C}|}] \rangle\rangle \\ \vdots & \ddots & \vdots \\ \langle\langle h_{|\mathcal{C}|} | [g, h_1] \rangle\rangle & \dots & \langle\langle h_{|\mathcal{C}|} | [g, h_{|\mathcal{C}|}] \rangle\rangle \end{pmatrix}, \end{aligned} \quad (\text{A.24})$$

where $\langle\langle \cdot | \cdot \rangle\rangle = \text{Tr}(\cdot \cdot^\dagger)$.

Define

$$\vec{c}^{(r-1)} = (\bar{c}_{11\dots 1}, \bar{c}_{11\dots 2}, \dots, \bar{c}_{|\mathcal{C}||\mathcal{C}|\dots|\mathcal{C}|}), \quad (\text{A.25})$$

then

$$\vec{c}^{(r-1)} = \int_0^{T_{\text{seq}}} dt_1 \dots \int_0^{t_{r-1}} dt_r \mathcal{D}(U_{\text{pri}}^\dagger)(t_1) \otimes \dots \otimes \mathcal{D}(U_{\text{pri}}^\dagger)(t_r) |H_{\text{pert}}(t_1)\rangle\rangle \otimes \dots \otimes |H_{\text{pert}}(t_r)\rangle\rangle. \quad (\text{A.26})$$

The discretization of $\{b_k(t)\}$ is chosen to be fine enough so that during each step both $H_{\text{pri}}(t)$ and $H_{\text{pert}}(t)$ can be approximately viewed as constant. Denote their discretized versions corresponding to $\vec{b}$ by

$$\begin{aligned} \vec{H}_{\text{pri}} &= (H_{\text{pri}}^1, H_{\text{pri}}^2, \dots, H_{\text{pri}}^Q), \\ \vec{H}_{\text{pert}} &= (H_{\text{pert}}^1, H_{\text{pert}}^2, \dots, H_{\text{pert}}^Q). \end{aligned} \quad (\text{A.27})$$

And the durations of the steps are given by

$$\Delta \vec{t} = (\Delta t_1, \Delta t_2, \dots, \Delta t_Q). \quad (\text{A.28})$$

A recipe for calculating the $\mathcal{C}$ -integrals for any sequence requires two ingredients: (1) Their calculation for a square pulse. (2) A chain rule of the $\mathcal{C}$ -integrals that allows to calculate those of a sequence given the $\mathcal{C}$ -integrals of the components of the sequence.

A.5.1 $\mathcal{C}$ -integrals of a square pulse

During the q th step, $H_{\text{pri}}(t) = H_{\text{pri}}^q$ and $H_{\text{pert}}(t) = H_{\text{pert}}^q$, so the $\mathcal{C}$ -integrals become

$$\vec{c}^{(r-1)} = \int_0^{T_{\text{seq}}} dt_1 \dots \int_0^{t_{r-1}} dt_r \exp(-i\mathcal{D}(\text{ad}_{H_{\text{pri}}^q})t_1) \otimes \dots \otimes \exp(-i\mathcal{D}(\text{ad}_{H_{\text{pri}}^q})t_r) |H_{\text{pert}}^q\rangle\rangle^{\otimes r}. \quad (\text{A.29})$$

Let $\mathcal{D}(\text{ad}_{H_{\text{pri}}^q}) = TDT^\dagger$ where D is diagonal and T is a unitary matrix, the integral becomes

$$\vec{c}^{(r-1)} = T^{\otimes r} \left(I(\lambda_1, \lambda_1, \dots, \lambda_1), I(\lambda_1, \lambda_1, \dots, \lambda_2), \dots, I(\lambda_{|\mathcal{C}|}, \lambda_{|\mathcal{C}|}, \dots, \lambda_{|\mathcal{C}|}) \right)^T T^{\dagger \otimes r} |H_{\text{pert}}^q\rangle\rangle^{\otimes r}. \quad (\text{A.30})$$

where $\lambda_1, \dots, \lambda_{|\mathcal{C}|}$ are eigenvalues of ad_{H_c} and

$$I(\lambda_{i_1}, \dots, \lambda_{i_r}) = \int_0^{T_{\text{seq}}} \dots \int_0^{t_{r-1}} dt_r e^{-i\lambda_{i_1}t_1} \dots e^{-i\lambda_{i_r}t_r}. \quad (\text{A.31})$$

The diagonalization of ad_{H_c} can be found using numerical Lanczos method.

A.5.2 $\mathcal{C}$ -integrals of a shaped pulse

Denote the $\mathcal{C}$ -integral of the q th step by $\vec{c}^{(r-1)}(q)$. Define

$$\vec{c}_{\text{tog}}^{(r-1)}(q) = [\mathcal{D}(U_1^\dagger) \cdots \mathcal{D}(U_{q-1}^\dagger)]^{\otimes r} \cdot \vec{c}^{(r-1)}(q), \quad (\text{A.32})$$

where U_q is the primary unitary of the q th step. Then

$$\vec{c}^{(r-1)} = \sum_{r_1 + \cdots + r_v = r} \sum_{Q \geq q_1 > \cdots > q_v \geq 1} \vec{c}_{\text{tog}}^{(r_1-1)}(q_1) \otimes \vec{c}_{\text{tog}}^{(r_2-1)}(q_2) \otimes \cdots \otimes \vec{c}_{\text{tog}}^{(r_v-1)}(q_v). \quad (\text{A.33})$$

For $r = 1, 2$ and 3 , we have

$$\begin{aligned} \vec{c}^{(0)} &= \sum_{q=1}^Q \vec{c}_{\text{tog}}^{(0)}(q), \\ \vec{c}^{(1)} &= \sum_{q=1}^Q \vec{c}_{\text{tog}}^{(1)}(q) + \sum_{q_1=2}^Q \sum_{q_2=1}^{q_1-1} \vec{c}_{\text{tog}}^{(0)}(q_1) \otimes \vec{c}_{\text{tog}}^{(0)}(q_2), \\ \vec{c}^{(2)} &= \sum_{q=1}^Q \vec{c}_{\text{tog}}^{(2)}(q) + \sum_{q_1=2}^Q \sum_{q_2=1}^{q_1-1} \left[\vec{c}_{\text{tog}}^{(0)}(q_1) \otimes \vec{c}_{\text{tog}}^{(1)}(q_2) + \vec{c}_{\text{tog}}^{(1)}(q_1) \otimes \vec{c}_{\text{tog}}^{(0)}(q_2) \right] \\ &\quad + \sum_{q_1=3}^Q \sum_{q_2=2}^{q_1-1} \sum_{q_3=1}^{q_2-1} \vec{c}_{\text{tog}}^{(0)}(q_1) \otimes \vec{c}_{\text{tog}}^{(0)}(q_2) \otimes \vec{c}_{\text{tog}}^{(0)}(q_3). \end{aligned} \quad (\text{A.34})$$

A.5.3 Cross $\mathcal{C}$ -integrals

Eq. (A.33) can be generalized to cross $\mathcal{C}$ -integrals that are defined on different spaces. In particular,

$$\begin{aligned} \vec{c}_{\text{cross}}^{(1)} &= \int_0^{T_{\text{seq}}} dt_1 \int_0^{t_1} dt_2 \mathcal{D}_{\text{pert}}(U_{\text{pri}}^\dagger(t_1)) \otimes \mathcal{D}_{\text{err}}(U_{\text{pri}}^\dagger(t_2)) |H_{\text{pert}}\rangle \otimes |A\rangle \\ &= \sum_{q_1=1}^Q \vec{c}_{\text{cross}}^{(1)}(q_1) + \sum_{q_1=2}^Q \sum_{q_2=1}^{q_1-1} \vec{c}_{\text{pert}}^{(0)}(q_1) \otimes \vec{c}_{\text{err}}^{(0)}(q_2), \end{aligned} \quad (\text{A.35})$$

where A is an error Hamiltonian (e.g., $\partial_j \Delta H$ and $\partial_{j_1} \partial_{j_2} \Delta H$). This is used in calculation of left-hand side of Eq. (A.61).

A.6 Robustness against Control Errors

This section develops the general method for achieving control robustness. The conditions for achieving robustness will be expressed in terms of the $\mathcal{C}$ -integrals.

When the model parameters $\vec{\mu} = \{\mu_j\}$ are subject to variations $\vec{\epsilon}$, the total Hamiltonian is

$$H_{\text{tot}}(t) = H_c(t, \vec{\mu}) + \Delta H(t, \vec{\epsilon}) + H_{\text{int}}, \quad (\text{A.36})$$

where

$$\Delta H(t, \vec{\epsilon}) = H_c(t, \vec{\mu} + \vec{\epsilon}) - H_c(t, \vec{\mu}), \quad (\text{A.37})$$

and $\Delta H(t, \vec{0}) = 0$. The total unitary propagator becomes

$$U(T_{\text{seq}}, \vec{\epsilon}) = U_{\text{pri}}(T_{\text{seq}}) \mathcal{T} \exp \left[-i \int_0^{T_{\text{seq}}} (\Delta H_{\text{tog}}(t, \vec{\epsilon}) + H_{\text{tog}}(t)) dt \right], \quad (\text{A.38})$$

where

$$\begin{aligned} \Delta H_{\text{tog}}(t, \vec{\epsilon}) &= U_{\text{pri}}^\dagger(t) \Delta H(t, \vec{\epsilon}) U_{\text{pri}}(t), \\ H_{\text{tog}}(t) &= U_{\text{pri}}^\dagger(t) H_{\text{pert}} U_{\text{pri}}(t). \end{aligned} \quad (\text{A.39})$$

Applying toggling transformation under ΔH_{tog} to the second factor, we have

$$U(T_{\text{seq}}, \vec{\epsilon}) = U_{\text{pri}}(T_{\text{seq}}) U_{\text{err}}(T_{\text{seq}}, \vec{\epsilon}) \mathcal{T} \exp \left(-i \int_0^{T_{\text{seq}}} \tilde{H}_{\text{tog}}(t, \vec{\epsilon}) dt \right), \quad (\text{A.40})$$

where

$$U_{\text{err}}(T_{\text{seq}}, \vec{\epsilon}) = \mathcal{T} \exp \left[-i \int_0^{T_{\text{seq}}} \Delta H_{\text{tog}}(t, \vec{\epsilon}) dt \right] \quad (\text{A.41})$$

and

$$\tilde{H}_{\text{tog}}(t, \vec{\epsilon}) = U_{\text{pri}}^\dagger(t) U_{\text{err}}^\dagger(t, \vec{\epsilon}) H_{\text{int}} U_{\text{err}}(t, \vec{\epsilon}) U_{\text{pri}}(t). \quad (\text{A.42})$$

Define the imperfect primary and perturbative unitary propagators as

$$\begin{aligned} \tilde{U}_{\text{pri}}(t, \vec{\epsilon}) &= U_{\text{pri}}(t) U_{\text{err}}(t, \vec{\epsilon}), \\ \tilde{U}_{\text{pert}}(t, \vec{\epsilon}) &= \mathcal{T} \exp \left(-i \int_0^{T_{\text{seq}}} \tilde{H}_{\text{tog}}(t, \vec{\epsilon}) dt \right), \end{aligned} \quad (\text{A.43})$$

we have

$$U(T_{\text{seq}}, \vec{\epsilon}) = \tilde{U}_{\text{pri}}(T_{\text{seq}}, \vec{\epsilon}) \tilde{U}_{\text{pert}}(T_{\text{seq}}, \vec{\epsilon}). \quad (\text{A.44})$$

We want to achieve robustness for both the primary and perturbative unitary propagators independently

$$\tilde{U}_{\text{pri}}(T_{\text{seq}}, \vec{\epsilon}) = U_{\text{target}} + O(\vec{\epsilon}^{n+1}) \quad (\text{A.45})$$

and

$$\tilde{U}_{\text{pert}}(t, \vec{\epsilon}) = \exp(-iH_{\text{eff}}T_{\text{seq}}) + O(\vec{\epsilon}^{n+1}). \quad (\text{A.46})$$

Together, they achieve the n th-order robustness in $U(T_{\text{seq}})$:

$$U(T_{\text{seq}}, \vec{\epsilon}) = U_{\text{target}} \exp(-iH_{\text{eff}}T_{\text{seq}}) + O(\vec{\epsilon}^{n+1}). \quad (\text{A.47})$$

Next, we will discuss how to achieve Eqs. (A.45) and (A.46) individually.

A.6.1 Robustness in $U_{\text{pri}}(T_{\text{seq}})$

Represent the error unitary (Eq. (A.41)) by an effective Hamiltonian H_{err}

$$U_{\text{err}}(T_{\text{seq}}, \vec{\epsilon}) = \exp(-iH_{\text{err}}T_{\text{seq}}), \quad (\text{A.48})$$

where H_{err} can be calculated using the Magnus expansion $H_{\text{err}} = \sum_{i=0}^{\infty} \bar{H}_{\text{err}}^{(i)}$

$$\begin{aligned} \bar{H}_{\text{err}}^{(0)} &= \frac{1}{T_{\text{seq}}} \int_0^{T_{\text{seq}}} \Delta H_{\text{tog}}(t_1) dt_1, \\ \bar{H}_{\text{err}}^{(1)} &= -\frac{i}{2T_{\text{seq}}} \int_0^{T_{\text{seq}}} dt_1 \int_0^{t_1} dt_2 [\Delta H_{\text{tog}}(t_1), \Delta H_{\text{tog}}(t_2)] \\ &\dots \end{aligned} \quad (\text{A.49})$$

Denote

$$\partial_{j_1} \dots \partial_{j_i} \Delta H_{\text{tog}} = \frac{\partial^i}{\partial \mu_{j_1} \dots \partial \mu_{j_i}} \Delta H_{\text{tog}} \Big|_{\vec{\epsilon}=\vec{0}}. \quad (\text{A.50})$$

The robustness is achieved by requiring

$$\begin{aligned} \frac{1}{T_{\text{seq}}} \int_0^{T_{\text{seq}}} \partial_{j_1} \Delta H_{\text{tog}}(t) dt &= 0, \\ \frac{1}{T_{\text{seq}}} \int_0^{T_{\text{seq}}} \partial_{j_1} \partial_{j_2} \Delta H_{\text{tog}}(t) dt &= 0, \\ -\frac{i}{2T_{\text{seq}}} \int_0^{T_{\text{seq}}} dt_1 \int_0^{t_1} dt_2 [\partial_{j_1} \Delta H_{\text{tog}}(t_1), \partial_{j_2} \Delta H_{\text{tog}}(t_2)] &= 0, \\ &\dots \end{aligned} \quad (\text{A.51})$$

In $\mathcal{C}_{\text{err}}(\mathbf{g}_{\text{pri}}, \Delta H)$, the time-ordered $\mathcal{C}$ -integrals corresponding to the conditions (A.51) are

$$\begin{aligned} & \vec{c}^{(0)}(\partial_{j_1} \Delta H(t, \vec{0})), \\ & \vec{c}^{(0)}(\partial_{j_1} \partial_{j_2} \Delta H(t, \vec{0})), \\ & \vec{c}^{(1)}(\{\partial_{j_1} \Delta H(t, \vec{0}), \partial_{j_2} \Delta H(t, \vec{0})\}), \\ & \dots, \end{aligned} \tag{A.52}$$

where the $\mathcal{C}$ -integrals are defined as

$$\begin{aligned} & \vec{c}^{(r-1)}(\{A_1(t), \dots, A_r(t)\}) \\ & = \int_0^{T_{\text{seq}}} dt_1 \dots \int_0^{t_{r-1}} dt_r \mathcal{D}(U_{\text{pri}}^\dagger)(t_1) \otimes \dots \otimes \mathcal{D}(U_{\text{pri}}^\dagger)(t_r) |A_1(t_1)\rangle \otimes \dots \otimes |A_r(t_r)\rangle. \end{aligned} \tag{A.53}$$

The corresponding robustness conditions are

$$\begin{aligned} & |\vec{c}^{(0)}(\partial_{j_1} \Delta H(t, \vec{0}))| = 0, \\ & |\vec{c}^{(0)}(\partial_{j_1} \partial_{j_2} \Delta H(t, \vec{0}))| = 0, \\ & |\bar{c}_{i_1 i_2}(\{\partial_{j_1} \Delta H(t, \vec{0}), \partial_{j_2} \Delta H(t, \vec{0})\}) + \bar{c}_{i_2 i_1}(\{\partial_{j_1} \Delta H(t, \vec{0}), \partial_{j_2} \Delta H(t, \vec{0})\})| = 0. \end{aligned} \tag{A.54}$$

A.6.2 Robustness in $U_{\text{pert}}(T_{\text{seq}})$

Representing the imperfect toggling-frame Hamiltonian $\tilde{H}_{\text{tog}}(t, \vec{\epsilon})$ (Eq. (A.42)) in $\mathcal{C}_{\text{pert}}(\mathbf{g}_{\text{pri}}, H_{\text{pert}})$, we have

$$|\tilde{H}_{\text{tog}}(t)\rangle_{\text{pert}} = \mathcal{D}_{\text{pert}}(U_{\text{err}}^\dagger(t, \vec{\epsilon})) |H_{\text{tog}}(t)\rangle_{\text{pert}} \tag{A.55}$$

using the Dyson series, we have

$$\mathcal{D}_{\text{pert}}(U_{\text{err}}^\dagger(t, \vec{\epsilon})) = \mathbb{1} + i \int_0^t \mathcal{D}_{\text{pert}}(\text{ad}_{\Delta H_{\text{tog}}}(t', \vec{\epsilon})) dt' - \int_0^t dt_1 \int_0^{t_1} dt_2 \mathcal{D}_{\text{pert}}(\text{ad}_{\Delta H_{\text{tog}}}(t_1, \vec{\epsilon}) \text{ad}_{\Delta H_{\text{tog}}}(t_2, \vec{\epsilon})) + \dots. \tag{A.56}$$

Expand ΔH to the second order of $\vec{\epsilon}$, we have

$$\begin{aligned} \mathcal{D}_{\text{pert}}(U_{\text{err}}^\dagger(t, \vec{\epsilon})) &= \mathbb{1} + i \int_0^t D_{\text{pert}} \left[\vec{\epsilon} \cdot \vec{\partial}(\text{ad}_{\Delta H_{\text{tog}}})(t_1, \vec{0}) + \frac{(\vec{\epsilon} \cdot \vec{\partial})^2}{2}(\text{ad}_{\Delta H_{\text{tog}}})(t_1, \vec{0}) \right] dt_1 \\ &\quad - \int_0^t \int_0^{t_1} D_{\text{pert}} [\vec{\epsilon} \cdot \vec{\partial}(\text{ad}_{\Delta H_{\text{tog}}})(t_1, \vec{0}) \vec{\epsilon} \cdot \vec{\partial}(\text{ad}_{\Delta H_{\text{tog}}})(t_2, \vec{0})] dt_1 dt_2 + O(\vec{\epsilon}^3). \end{aligned} \tag{A.57}$$

Substitute it into the $\vec{c}^{(0)}$ and $\vec{c}^{(1)}$, we have

$$\begin{aligned}
\vec{c}^{(0)} &= \int_0^{T_{\text{seq}}} dt \mathcal{D}_{\text{pert}}(U_{\text{err}}^\dagger(t, \vec{\epsilon})) |H_{\text{tog}}(t)\rangle\rangle_{\text{pert}} \\
&= \int_0^{T_{\text{seq}}} dt |H_{\text{tog}}(t)\rangle\rangle_{\text{pert}} \\
&\quad + i \int_0^{T_{\text{seq}}} dt \int_0^t \mathcal{D}_{\text{pert}} \left[\vec{\epsilon} \cdot \vec{\partial}(\text{ad}_{\Delta H_{\text{tog}}})(t_1, \vec{0}) + \frac{(\vec{\epsilon} \cdot \vec{\partial})^2}{2}(\text{ad}_{\Delta H_{\text{tog}}})(t_1, \vec{0}) \right] dt_1 |H_{\text{tog}}(t)\rangle\rangle_{\text{pert}} \\
&\quad - \int_0^{T_{\text{seq}}} dt \int_0^t \int_0^{t_1} \mathcal{D}_{\text{pert}} [\vec{\epsilon} \cdot \vec{\partial}(\text{ad}_{\Delta H_{\text{tog}}})(t_1, \vec{0}) \vec{\epsilon} \cdot \vec{\partial}(\text{ad}_{\Delta H_{\text{tog}}})(t_2, \vec{0})] dt_1 dt_2 |H_{\text{tog}}(t)\rangle\rangle_{\text{pert}} \\
&\quad + \dots.
\end{aligned} \tag{A.58}$$

and

$$\begin{aligned}
\vec{c}^{(1)} &= \int_0^{T_{\text{seq}}} dt \int_0^t dt' \mathcal{D}_{\text{pert}}(U_{\text{err}}^\dagger(t, \vec{\epsilon})) \otimes \mathcal{D}_{\text{pert}}(U_{\text{err}}^\dagger(t', \vec{\epsilon})) |H_{\text{tog}}(t)\rangle\rangle_{\text{pert}} \otimes |H_{\text{tog}}(t')\rangle\rangle_{\text{pert}} \\
&= \int_0^{T_{\text{seq}}} dt \int_0^t dt' |H_{\text{tog}}(t)\rangle\rangle_{\text{pert}} \otimes |H_{\text{tog}}(t')\rangle\rangle_{\text{pert}} \\
&\quad + i \int_0^{T_{\text{seq}}} dt \int_0^t \mathcal{D}_{\text{pert}} \left[\vec{\epsilon} \cdot \vec{\partial}(\text{ad}_{\Delta H_{\text{tog}}})(t_1, \vec{0}) + \frac{(\vec{\epsilon} \cdot \vec{\partial})^2}{2}(\text{ad}_{\Delta H_{\text{tog}}})(t_1, \vec{0}) \right] dt_1 |H_{\text{tog}}(t)\rangle\rangle_{\text{pert}} \otimes \\
&\quad \int_0^t dt' |H_{\text{tog}}(t')\rangle\rangle_{\text{pert}} \\
&\quad + i \int_0^{T_{\text{seq}}} dt |H_{\text{tog}}(t)\rangle\rangle_{\text{pert}} \otimes \\
&\quad \int_0^t dt' \int_0^{t'} \mathcal{D}_{\text{pert}} \left[\vec{\epsilon} \cdot \vec{\partial}(\text{ad}_{\Delta H_{\text{tog}}})(t_1, \vec{0}) + \frac{(\vec{\epsilon} \cdot \vec{\partial})^2}{2}(\text{ad}_{\Delta H_{\text{tog}}})(t_1, \vec{0}) \right] dt_1 |H_{\text{tog}}(t')\rangle\rangle_{\text{pert}} \\
&\quad + \dots.
\end{aligned} \tag{A.59}$$

The leading term of the error (in $\vec{c}^{(0)}$) is removed by requiring

$$i \int_0^{T_{\text{seq}}} dt \int_0^t \mathcal{D}_{\text{pert}} [\partial_j(\text{ad}_{\Delta H_{\text{tog}}})(t_1, \vec{0})] dt_1 |H_{\text{tog}}(t)\rangle\rangle_{\text{pert}} = 0, \tag{A.60}$$

for all $j = 1, \dots, J$. Since $\partial_j \Delta H_{\text{tog}}(t) = \sum_{l=1}^{|\mathcal{C}_{\text{err}}|} c_l^{\text{err}}(t; \partial_j \Delta H, \mathcal{C}_{\text{err}}) h_{\text{err}}^l$ and $H_{\text{tog}}(t) = \sum_{s=1}^{|\mathcal{C}_{\text{pert}}|} c_s^{\text{pert}}(t) h_{\text{pert}}^s$ where $\{h_{\text{err}}^l\}$ and $\{h_{\text{pert}}^s\}$ are the bases of their corresponding spaces, the above condition becomes

$$\left\| \sum_{l=1}^{|\mathcal{C}_{\text{err}}|} \sum_{s=1}^{|\mathcal{C}_{\text{pert}}|} \int_0^{T_{\text{seq}}} dt_1 \int_0^{t_1} dt_2 c_s^{\text{pert}}(t_1) c_l^{\text{err}}(t_2; \partial_j \Delta H) [i h_{\text{err}}^l, h_{\text{pert}}^s] \right\|_2 = 0. \tag{A.61}$$

A.6.3 Control Robustness is Always Achievable

In the presence of an error Hamiltonian $\Delta H(t, \mu + \epsilon)$, assume the following: (1) there exists an implementation of the identity operator that is robust to ϵ to at least zeroth order (e.g., a wait period where $a_k(t) \equiv 0$); (2) ΔH continuously depends on $a_k(t)$ and μ ; and (3) the correlation time of ΔH is finite so that we can compose gates without interference, then the following result holds.

Theorem A.6.1. *Zeroth-order robustness to ϵ is always achievable to any systematic error Hamiltonian ΔH .*

Proof. The total Hamiltonian is

$$H_{\text{tot}}(t) = H_c(t) + \epsilon \frac{\partial}{\partial \epsilon} \Delta H(t, 0) + O(\epsilon^2). \quad (\text{A.62})$$

The goal is to design a unitary U_{target} so that

$$\begin{aligned} U(T_{\text{seq}}, 0) &= U_{\text{target}}, \\ U(T_{\text{seq}}, \epsilon) &= U_{\text{target}} + O(\epsilon^2). \end{aligned} \quad (\text{A.63})$$

Suppose that we have found a sequence $U_0(T_{\text{seq}}, \epsilon)$ that achieves the first condition in Eqs. (A.63). In general, we have

$$U_0(T_{\text{seq}}, \epsilon) = U_{\text{target}} e^{-i\epsilon\sigma_0\theta} + O(\epsilon^2). \quad (\text{A.64})$$

If we could implement a sequence U^* so that

$$U^*(T_{\text{seq}}, \epsilon) = e^{i\epsilon\sigma_0\theta} + O(\epsilon^2), \quad (\text{A.65})$$

then $U_0 U^*$ achieves a robust implementation of U_{target} in zeroth order.

To this end, note that we can always achieve an identity operation U_1 with no control error (a wait period, for example) and an identity U_2 with some error.

Since we can continuously change U_1 to U_2 , there exists $U_r = \exp(-ir\epsilon\sigma') + O(\epsilon^2)$, $\forall r > 0$ and some σ' . Therefore, if we choose r small enough, we can achieve $U_r^N \approx \exp(-is\epsilon\sigma') + O(\epsilon^2)$ to arbitrary accuracy, $\forall s > 0$ and some positive integer N .

Due to universal control, we can rotate σ' to any direction. Thus, we can implement $U_\alpha(s) = \exp(-i(s\epsilon\sigma_\alpha + \epsilon\sigma_{r_\alpha})) + O(\epsilon^2)$, for $\alpha = \pm x, \pm y$ and $\pm z$. Here, σ_{r_α} corresponds to the error introduced by the imperfect rotation. Since we can choose any $s > 0$ and σ_{r_α} does not depend on s , therefore, $\forall \sigma_0$, there exist $s_\alpha > 0$ so that $\prod_\alpha U_\alpha(s_\alpha) = e^{i\epsilon\sigma_0\theta} + O(\epsilon^2)$. $\square$

The result is a generalization of the result from [30]. All the errors considered in this work (Rabi field variation, variations in linear or nonlinear models) satisfy the assumptions and therefore can be corrected to at least zeroth order. Additionally, requiring control robustness does not affect the controllability of U_{pri} or U_{pert} , so when characterizing the achievable set of H_{target} , we do not need to consider ΔH .

A.7 Higher-order corrections

Having structured the control and internal Hamiltonian such that the desired effective Hamiltonian is achievable in zeroth order, the remaining task is to engineer the average Hamiltonian of higher orders to be under a threshold, which needs to be accomplished for the entire distributions of Hamiltonian parameters. This is necessary for achieving high precision when the target control is achieved through the effective Hamiltonian.

Although, the frequency of modulation is typically much larger than $\|H_{\text{pert}}\|$, making $\bar{H}^{(n)} \sim (H_{\text{pert}}\tau_c)^{n+1}/T_{\text{seq}}$ a small quantity of order $n+1$, where τ_c is the characteristic time of the modulation [82, 128], modulation of controls can lead to higher-order terms being larger than low-order terms.

It is well-known that odd-order correction terms can be eliminated by symmetrizing the control sequence [127]. This is widely used for decoupling and is straightforward to implement for simulations. When H_{pri} contains H_{int} , however, a sequence cannot be symmetrized in general, since the sign of H_{int} cannot be reversed. Here, we provide computable conditions in terms of the $\mathcal{C}$ -integrals.

The higher-order corrections terms in the Magnus expansion are

$$\begin{aligned}\bar{H}^{(1)} &= -\frac{i}{2T_{\text{seq}}} \int_0^{T_{\text{seq}}} dt_1 \int_0^{t_1} dt_2 [H_{\text{tog}}(t_1), H_{\text{tog}}(t_2)], \\ \bar{H}^{(2)} &= -\frac{1}{6T_{\text{seq}}} \int_0^{T_{\text{seq}}} dt_1 \int_0^{t_1} dt_2 \int_0^{t_2} dt_3 ([H_{\text{tog}}(t_1), [H_{\text{tog}}(t_2), H_{\text{tog}}(t_3)]] \\ &\quad + [H_{\text{tog}}(t_3), [H_{\text{tog}}(t_2), H_{\text{tog}}(t_1)]]) , \\ &\dots\end{aligned}\tag{A.66}$$

Denote

$$\begin{aligned}F_2(A_1, A_2) &= [A_1, A_2], \\ F_3(A_1, A_2, A_3) &= [A_1, [A_2, A_3]] + [A_3, [A_2, A_1]], \\ &\dots,\end{aligned}\tag{A.67}$$

then $F_r(A_1, \dots, A_r) = (-1)^{r-1} F_r(A_r, \dots, A_1)$. Represent $H_{\text{tog}}(t)$ in $\mathcal{C}(\mathbf{g}_{\text{pri}}, H_{\text{pert}})$, we have

$$F_r(H_{\text{tog}}(t_1), \dots, H_{\text{tog}}(t_r)) = \sum_{i_1, \dots, i_r=1}^{|\mathcal{C}|} c_{i_1}(t_1) \cdots c_{i_r}(t_r) F_r(h_{i_1}, \dots, h_{i_r}). \quad (\text{A.68})$$

Therefore, since the $(r-1)$ th correction term is

$$\bar{H}^{(r-1)} \propto \int_0^{T_{\text{seq}}} dt_1 \cdots \int_0^{t_{r-1}} dt_r F_r(H_{\text{tog}}(t_1), \dots, H_{\text{tog}}(t_r)), \quad (\text{A.69})$$

the condition for $\bar{H}^{(r-1)} \equiv 0$ for all $h_1, \dots, h_r$ is given by

$$|\bar{c}_{i_1 \dots i_r}(T_{\text{seq}}) - (-1)^{r-1} \bar{c}_{i_r \dots i_1}(T_{\text{seq}})| = 0, \quad (\text{A.70})$$

for all $i_1, \dots, i_r$ excluding the situations where $i_1 = \dots = i_r$.

A.8 Numerical Setup For Sequence Design

A.8.1 Objective functions

A sequence is designed to achieve a target primary unitary, engineer a target zeroth-order average Hamiltonian while accounting for higher-order corrections, and meet robustness requirements. Table A.1 summarizes these requirements. The criteria of the objectives

Table A.1: Conditions for achieving different objectives.

Objective name	Subject	Target	Criterion
Primary unitary	$U_{\text{pri}}(T_{\text{seq}})$	U_{target}	$1 - \frac{\text{Abs}[\text{Tr}(U_{\text{pri}} U_{\text{target}}^\dagger)]}{\text{Tr}(U_{\text{target}} U_{\text{target}}^\dagger)} = 0$
First-order primary robustness	$\partial_j \Delta H(t, \vec{0})$	0	Eq. 1 of (A.54)
Second-order primary robustness	$\partial_{j_1} \partial_{j_2} \Delta H(t, \vec{0})$	0	Eq. 2 of (A.54)
First-order cross primary robustness	$\partial_{j_1} \Delta H(t, \vec{0}), \partial_{j_2} \Delta H(t, \vec{0})$	0	Eq. 3 of (A.54)
Zero-order average Hamiltonian	$\bar{H}^{(0)}(T_{\text{seq}})$	H_{target}	$ \bar{c}^{(0)} - H_{\text{target}}\rangle\rangle = 0$
$(r-1)$ th-order correction ($r > 1$)	$\bar{H}^{(r-1)}(T_{\text{seq}})$	0	Eq. (A.70)
First-order correction	$\partial_j \Delta H(t), H_{\text{tog}}(t)$	0	Eq. (A.61)

take the form

$$f(\text{sequence}, \text{objective}) = 0 \quad (\text{A.71})$$

where f is a positive function of the sequence and the objective. Therefore, the total cost function (the objective function is just 1-the cost function) can be written as

$$f_{\text{tot}}(\{a_k(t)\}) = \sum_i w_i f_i, \quad (\text{A.72})$$

where f_i is the function value of f of the i th objective and w_i are positive weights. One can adjust the weight to reflect the difference in importance of different objectives. All the objectives are met when $f_{\text{tot}}(\{a_k(t)\}) = 0$. Therefore, the task of designing a sequence is an optimization (minimization of the cost function) problem whose goal is to find $\{a_k(t)\}$ that minimizes $f_{\text{tot}}(\{a_k(t)\})$.

A.9 Optimization algorithms

A generalized simulated annealing for continuous optimization is used for sequence design in this work. Simulated annealing is a stochastic optimizer that drew inspiration from the physical process of annealing in metallurgy [98]. The generalized simulated annealing [191] combines classical simulated annealing and fast simulated annealing and provides tunable hyperparameters that control the behavior of the algorithms. It has been proven very helpful in finding global optimum of complex multi-dimensional functions where methods such as the gradient descent may be trapped easily in local optimum and has been applied in quantum learning [44] and optimization [54].

The algorithm is typically designed for unconstrained optimization. To adapt it to pulse engineering where the power of control is constrained, a legalizer is used to restrain the particle within the feasible region. For box constraint $[-1, 1]^n$, define a bouncing function

$$B(x) = (-1)^{\lfloor \frac{x+1}{2} \rfloor} \left(x - 2 \left\lfloor \frac{x+1}{2} \right\rfloor \right), \quad (\text{A.73})$$

then the box constraint (bounce) can be applied by defining a function

$$F(\vec{x}) = f(B(\vec{x})). \quad (\text{A.74})$$

Applying the optimization algorithm to F instead of f is equivalent to letting the particle bouncing off the wall which allows it to lie within the constraint without sacrificing exploration of the energy landscape. Other strategies for designing a legalizer are depicted in Fig. A.5. Alternatively, a constraint can be added as a penalty function [32].

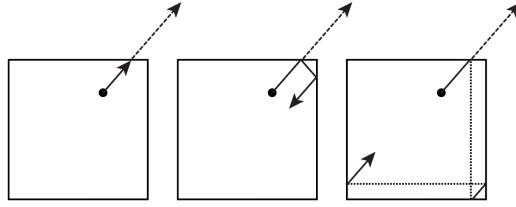


Figure A.5: Different types of legalizers for box constraint: Clip, bounce and reflect. The clip legalizer stops the particle when it touches the boundary while the bounce and reflect legalizers allow the particle to continue its movement according to different rules.

Below is the pseudocode of the algorithm. A D-dimensional Tsallis random number generator introduced in [169] is used. For best efficiency, the best solution should always be stored [56].

Algorithm 7 Generalized simulated annealing. For a fixed $q_a \in \mathbb{R}$, a fixed $1 < q_v < 3$, an initial temperature T_0 , a maximal number of iterations $t_{\max}$, the function GSA takes an D -dimensional Energy function E , an initial guess $\vec{x}_1$ and the target energy value E_{target} as input and return a solution $\vec{x}_{\text{best}}$ such that either $E(\vec{x}_{\text{best}}) < E_{\text{target}}$ or that is the best solution found within maximal number of iterations.

```

1: function TSALLISRNG( $T$ )
2:    $p \leftarrow \frac{3-q_v}{2(q_v-1)}$ 
3:    $s \leftarrow \frac{\sqrt{2(q_v-1)}}{(T)^{1/(3-q_v)}}$ 
4:   Generate iid  $x_i \sim \mathcal{N}(0, 1)$  for  $i = 1, 2, \dots, D$ 
5:    $\vec{x} \leftarrow (x_1, x_2, \dots, x_D)$ 
6:   Generate a gamma random number  $u \sim \gamma(1, p)$ 
7:   return  $\frac{\vec{x}}{s\sqrt{u}}$ 
8: end function
9: function TEMPERATURE( $t$ ) ▷ Cooling schedule
10:  return  $\frac{(q_v-1)^2-1}{(q_v-1)^{t+1}-1} T_0$ 
11: end function
12: function COMPAREENERGY( $E_1, E_2, T$ )
13:   $r \leftarrow$  random real number from  $[0, 1]$ 
14:   $a \leftarrow 1 + \frac{(E_2-E_1)(q_a-1)}{T}$ 
15:   $b \leftarrow \left[ 1 + \frac{(E_2-E_1)(q_a-1)}{T} \right]^{-\frac{1}{q_a-1}}$ 
16:  if  $E_2 < E_1$  then
17:    return True
18:  else
19:    if  $q_a > 1$  then
20:      return Boole( $r \leq b$ )
21:    else if  $q_a < 1$  then
22:      return Boole( $a \geq 0$  and  $r \leq b$ )
23:    else
24:      return Boole( $r \leq \exp[(E_1 - E_2)/T]$ )
25:    end if
26:  end if
27: end function

```

Algorithm 8 Generalized simulated annealing (continued).

```
28: function GSA( $E, \vec{x}_1$ )
29:    $E_1 \leftarrow E(\vec{x}_1)$ 
30:    $E_{\text{best}} \leftarrow E_1$ 
31:    $\vec{x}_{\text{best}} \leftarrow \vec{x}_1$ 
32:    $t \leftarrow 1$ 
33:   while  $t < t_{\text{max}}$  and  $E_{\text{best}} > E_{\text{target}}$  do
34:      $T \leftarrow \text{TEMPERATURE}(t)$ 
35:      $\vec{x}_2 \leftarrow \text{LEGALIZE}(\vec{x}_1 + \text{TSALLISRNG}(T))$ 
36:      $E_2 \leftarrow E(\vec{x}_2)$ 
37:     if  $E_2 < E_{\text{best}}$  then
38:        $E_{\text{best}} \leftarrow E_2$ 
39:        $\vec{x}_{\text{best}} \leftarrow \vec{x}_2$ 
40:     end if
41:     if  $\text{COMPAREENERGY}(E_1, E_2, T)$  then
42:        $\vec{x}_1 \leftarrow \vec{x}_2$ 
43:        $E_1 \leftarrow E_2$ 
44:     end if
45:      $t \leftarrow t + 1$ 
46:   end while
47:   return  $\vec{x}_{\text{best}}$ 
48: end function
```

Parallel searches [161] that divide the search into multiple stages are helpful in exploring the most promising regions and are used for multi-objective optimization in sequence design. The performance of the algorithm is largely determined by the hyperparameters: The initial temperature T_0 , the cooling rate parameter q_v , the exploration rate parameter q_a and the maximal number of iterations t_{max} . A general strategy is to choose a high initial temperature (e.g., two times the highest acceptable hill of $E(\vec{x})$), a small q_v and a large q_a to ensure enough exploration, and a large i_{max} to ensure convergence. To find optimized hyperparameters, test runs are implemented on complex problem with a fixed small i_{max} , then a Nelder-Mead optimizer [142, 113] is used to adjust T_0 , q_v and q_a to optimize the average performance [149].

A.10 Supplementary Information on Examples

A.10.1 Broadband robust sequences

Robustness beyond the perturbative limitation can be achieved by choosing proper partitionings of H_{tot} . As an example, we design a broadband robust sequence against variations in offset. For simplicity, Rabi field variations are assumed to be negligible. Choose three different partitionings:

$$H_{\text{pri}}^i = H_c(t) + \omega_i H_{\text{int}}, \quad H_{\text{pert}}^i = (\Delta\omega - \omega_i) H_{\text{int}}, \quad (\text{A.75})$$

for $i = 1, 2$ and 3 where $\omega_i/\omega_{\text{max}} = -0.2, 0$ and 0.2 . The goal is to design $\{a_k(t)\}$ so that U_{pri}^i achieves a Hadamard gate and $\bar{H}_i^{(0)} = 0$ for $i = 1, 2$ and 3 simultaneously. Fig. A.6 shows

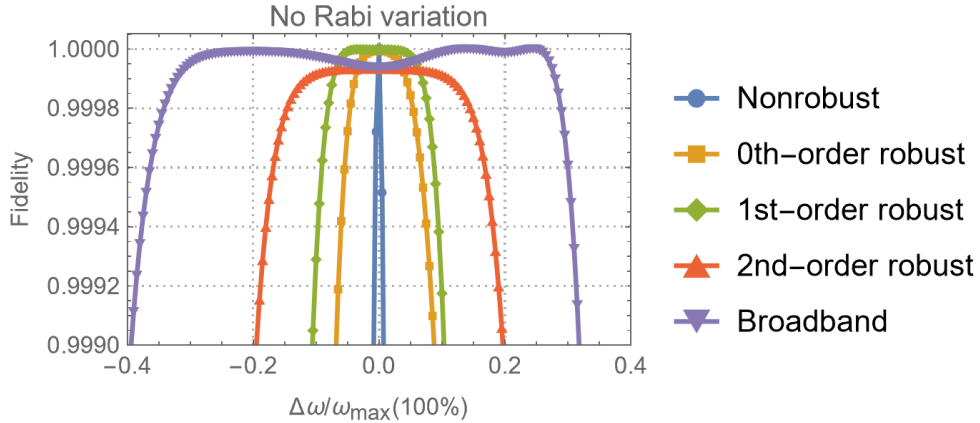


Figure A.6: Control landscapes of non-robust, 0th-order, 1st-order, 2nd-order and broadband robust sequences. The broadband sequences achieve a much larger robust region than the 2nd-order robust sequence.

the control landscapes of a broadband robust sequence, together with the landscapes of other sequences in comparison. Note that the broadband robust sequence is 190 ns, shorter than the 2nd-order robust sequence (300 ns). Thus, by choosing different partitioning of H_{tot} and combining different robust regions, we can potentially design robust sequences across a large region that is also more efficient.

A.10.2 Robust control in nonlinear systems

For general nonlinear control systems, the output field $\vec{B}(t)$ and the input parameter $\vec{a}(t)$ are related through a nonlinear differential equation

$$\frac{d\vec{B}}{dt} = h(\vec{B}, \vec{a}, t; \vec{\mu}_0), \quad (\text{A.76})$$

where h is a general nonlinear function. Although nonlinear effects may exhibit chaotic behaviors, in the presence of small nonlinearity, the system may have a non-chaotic regime where the output field $\vec{B}(t)$ can be calculated exactly.

A variation $\vec{\epsilon}$ in the model parameter $\vec{\mu}$ results in a different output field $\vec{B}(t, \vec{\mu}_0 + \vec{\epsilon}) = \vec{B}(t, \vec{\mu}_0) + \sum_i \epsilon_i \frac{\partial \vec{B}}{\partial \mu_i}(t, \vec{\mu}_0) + \dots$. The differential equation becomes

$$\frac{d}{dt} \left(\vec{B} + \sum_i \epsilon_i \frac{\partial \vec{B}}{\partial \mu_i} + \dots \right) = h(\vec{B}(\vec{\mu}_0), \vec{a}, t; \vec{\mu}_0) + \sum_i \epsilon_i \mathbf{J}(h, \vec{B})|_{\vec{\mu}=\vec{\mu}_0} \frac{\partial \vec{B}}{\partial \epsilon_i}(\vec{\mu}_0) + \dots, \quad (\text{A.77})$$

where $\mathbf{J}$ is the Jacobian matrix. By equating terms of the same order in ϵ_i , we can obtain coupled differential equations that can be solved for $\frac{\partial \vec{B}}{\partial \mu_i}, \dots$, enabling evaluation of the error Hamiltonian ΔH . As an example, consider the resonant circuit in Fig. A.7. In a

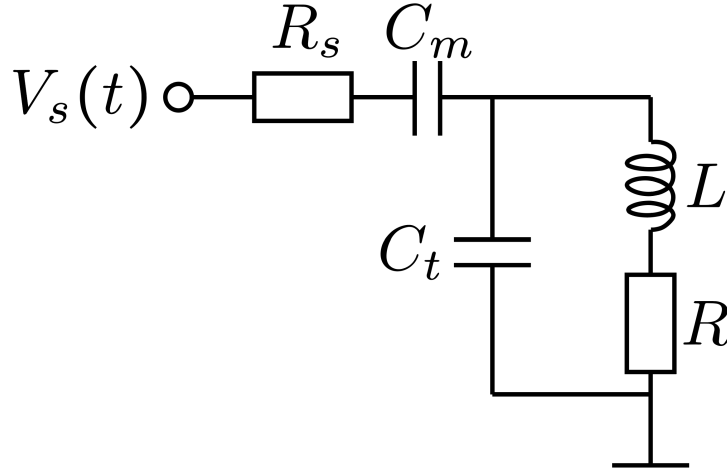


Figure A.7: A simple resonant circuit with $R_s = 50\Omega$, $R = 0.01\Omega$, $C_m = 11\text{fF}$, $C_t = 1.479\text{pF}$, and $L = 170\text{pH}$. The resonance frequency is 10GHz. We choose $\kappa_i = 1\text{V}^{-1}$ and $\kappa_o = 2\text{A}^{-1}$.

rotating frame of frequency ω_r , it obeys the following differential equation [86]

$$\frac{dx}{dt} = A(x)x + \alpha(t)u, \quad (\text{A.78})$$

where

$$x = \begin{pmatrix} \tilde{I}_L \\ \tilde{V}_{C_m} \\ \tilde{V}_{C_t} \end{pmatrix}, \quad A(x) = \begin{pmatrix} -\frac{R}{L} & 0 & \frac{1}{L} \\ 0 & \frac{-1}{R_L C_m} & \frac{1}{R_L C_m} \\ \frac{-1}{C_t} & \frac{1}{R_L C_t} & \frac{1}{R_L C_t} \end{pmatrix} - i\omega_r \mathbb{1}, \quad u = \begin{pmatrix} 0 \\ \frac{1}{R_L C_m} \\ \frac{1}{R_L C_t} \end{pmatrix}, \quad (\text{A.79})$$

and

$$a_1(t) = \kappa_i \text{Re}(\alpha(t)), \quad a_2(t) = \kappa_i \text{Im}(\alpha(t)), \quad b_1(t) = \kappa_o \text{Re}(\tilde{I}_L), \quad b_2(t) = \kappa_o \text{Im}(\tilde{I}_L), \quad (\text{A.80})$$

where κ_i and κ_o are normalization constants that depend on the geometry of the control system.

In the presence of kinetic inductance, the value of L depends on the current:

$$L = L_0(1 + \alpha_L |I_L|^2), \quad (\text{A.81})$$

where α_L is a constant. We want to design for $\alpha_L = 0$ and to achieve robustness for small α_L . Using Eq. (A.77), the equations for $x(\alpha_L = 0)$ and $\frac{\partial x}{\partial \alpha_L}(\alpha_L = 0)$ are

$$\frac{d}{dt} \begin{pmatrix} x \\ \frac{\partial x}{\partial \alpha_L} \end{pmatrix} = \begin{pmatrix} A(x) & 0 \\ \frac{\partial A}{\partial \alpha_L}(x) & A(x) \end{pmatrix} \begin{pmatrix} x \\ \frac{\partial x}{\partial \alpha_L} \end{pmatrix} + \alpha(t) \begin{pmatrix} u \\ \frac{\partial u}{\partial \alpha_L} \end{pmatrix} \Big|_{\alpha_L=0}, \quad (\text{A.82})$$

where

$$\frac{\partial A}{\partial \alpha_L}(x, \alpha_L = 0) = \begin{pmatrix} \frac{R|\tilde{I}_L|^2}{L_0} & 0 & -\frac{|\tilde{I}_L|^2}{L_0} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \frac{\partial u}{\partial \alpha_L} = 0. \quad (\text{A.83})$$

The distortion of the input pulse and the corresponding error Hamiltonian are thus calculated by solving the system of differential equations.

A.10.3 Bloch-Siegert Shift

The Bloch-Siegert shift of spectra [17] is due to distortion of the control field by a counter-rotating component in the rotating frame. To see how this arises, start with the total Hamiltonian in the lab frame is

$$H_{\text{lab}}(t) = A(t) \cos \left(\int_0^t \omega_r(t') dt' + \phi(t) \right) \sigma_x + \omega_0 \sigma_z / 2. \quad (\text{A.84})$$

Move into the rotating frame (interaction frame) of $H_r(t) = \omega_r(t)\sigma_z/2$, define

$$\theta(t) = \int_0^t \omega_r(t')dt', \quad U_r(t) = \exp(-i\theta(t)\sigma_z/2), \quad \delta(t) = \omega_0 - \omega_r(t). \quad (\text{A.85})$$

The total Hamiltonian in the rotating frame is

$$\begin{aligned} H_{\text{tot}}(t) &= A(t) \cos(\theta(t) + \phi(t)) U_r^\dagger(t) \sigma_x U_r(t) + \delta(t) \sigma_z/2 \\ &= A(t) \cos(\theta(t) + \phi(t)) (\cos(\theta(t)) \sigma_x - \sin(\theta(t)) \sigma_y) + \delta(t) \sigma_z/2 \\ &= A(t) (\cos(\phi(t)) \sigma_x + \sin(\phi(t)) \sigma_y)/2 + \delta(t) \sigma_z/2 + \Delta H, \end{aligned} \quad (\text{A.86})$$

where

$$\Delta H(t) = A(t) [\cos(2\omega_r t + \phi) \sigma_x - \sin(2\omega_r t + \phi) \sigma_y]/2. \quad (\text{A.87})$$

The rotating frame approximation [176] ignores the effect of ΔH . Our goal is to account for this effect and design sequences that are robust to variations in ω_r . For simplicity, consider a constant $\omega_r = \omega_0$, expand ΔH around ω_r , we have

$$\Delta H(\omega_r + \epsilon) = \Delta H(\omega_r) + \epsilon \Delta H'(\omega_r) + O(\epsilon^2), \quad (\text{A.88})$$

where

$$\Delta H'(\omega_r) = -t A(t) [\sin(2\omega_r t + \phi) \sigma_x + \cos(2\omega_r t + \phi) \sigma_y]. \quad (\text{A.89})$$

Fig. A.8 shows $\Delta H(t)$ (top) and $\Delta H'(t)$ (bottom) for a given $H_c(t)$. To design sequences that account for the effect of ΔH , $\Delta H(\omega_r)$ is included in the primary Hamiltonian $H_{\text{pri}}(t)$. The robustness to variations in ω_r is achieved to the first order by decoupling $\Delta H'(\omega_r)$.

Due to the strong time dependence of $\Delta H(t)$ and $\Delta H'(t)$, the middle-point approximation (Eq. (A.19)) in discretization can lead to significant misrepresentation (Fig. A.9). Therefore, the zeroth-order average (Eq. A.20) should be used instead. This leads to the following approximation of the errors terms at the m th step:

$$\begin{aligned} \Delta H[m] &= \frac{A(t)}{4\omega_r \delta t} \{ [\sin(2m\omega_r t + \phi) - \sin(2(m-1)\omega_r t + \phi)] \sigma_x \\ &\quad + [\cos(2m\omega_r t + \phi) - \cos(2(m-1)\omega_r t + \phi)] \sigma_y \}, \\ \Delta H'[m] &= -\frac{A(t)}{4\omega_r^2 \delta t} \{ [2(m-1)\delta t \omega_r \cos(2(m-1)\omega_r t + \phi) - 2m\delta t \omega_r \cos(2m\omega_r t + \phi) \\ &\quad - \sin(2(m-1)\omega_r t + \phi) + \sin(2m\omega_r t + \phi)] \sigma_x \\ &\quad + [-2(m-1)\delta t \omega_r \sin(2(m-1)\omega_r t + \phi) - 2m\delta t \omega_r \sin(2m\omega_r t + \phi) \\ &\quad - \cos(2(m-1)\omega_r t + \phi) + \cos(2m\omega_r t + \phi)] \sigma_y \}. \end{aligned} \quad (\text{A.90})$$

Fig. A.10 shows the control landscape of robust and non-robust sequences against variations in ω_r .

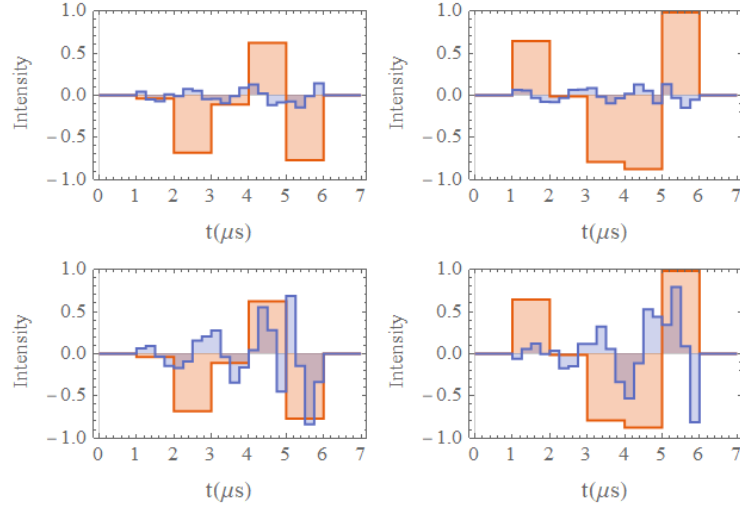


Figure A.8: Illustration of x (left) and y (right) components of $\Delta H(t)$ (top) and $\Delta H'(t)$ (bottom), both blue, of Bloch-Siegert shift at $\omega_r = 1.6\text{MHz}$. The red area corresponds to the input parameters. Note that the error $\Delta H'(t)$ grows linearly over time.

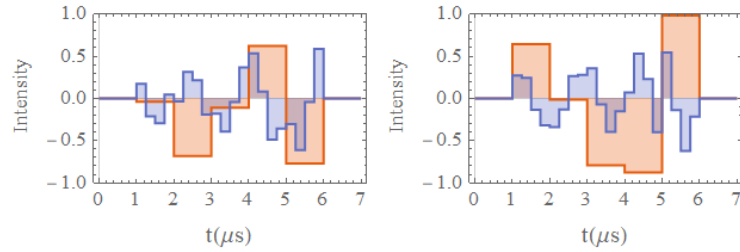


Figure A.9: Illustration of the Bloch-Siegert error (blue) when discretized using the middle-point approximation at $\omega_r = 1.6\text{MHz}$. It greatly overestimates the intensity of the error.

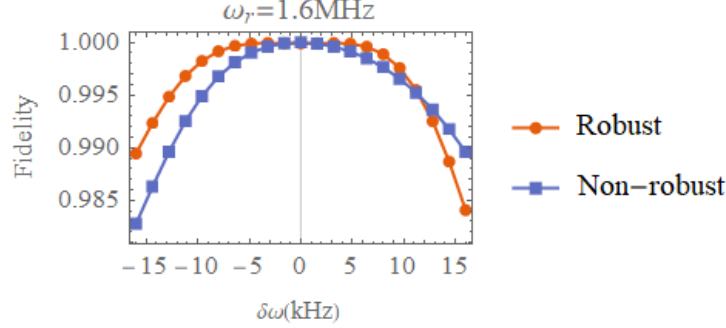


Figure A.10: Control landscapes of robust and non-robust sequences against variations in the resonance frequency ω_r . Both sequences account for the Bloch-Siegert shift.

A.10.4 Leakage in multi-qubit systems

The total Hamiltonian of a two-qubit system is

$$\begin{aligned} H_{\text{tot}}(t) &= H_c(t) + H_{\text{int}}, \\ H_{\text{int}} &= \omega_0^1 \sigma_z^1 + \omega_0^2 \sigma_z^2 + J H_0, \\ H_c(t) &= 2[C_1(t) \cos(\omega_0^1 t + \theta_1(t)) + C_2(t) \cos(\omega_0^2 t + \theta_2(t))](\sigma_x^1 + \sigma_x^2). \end{aligned} \quad (\text{A.91})$$

where H_0 is the interaction Hamiltonian between the two qubits, J is the interaction strength. In the rotating frame of $\omega_0^1 \sigma_z^1 + \omega_0^2 \sigma_z^2$, the control Hamiltonian becomes

$$H_c(t) = H_c^1(t) + H_c^2(t) + H_1'(t) + H_2'(t) + H_{12}'(t), \quad (\text{A.92})$$

where

$$H_c^i(t) = C_i(t)[\cos(\theta_i(t))\sigma_x^i + \sin(\theta_i(t))\sigma_y^i] \quad (\text{A.93})$$

are individual controls on the i th qubit,

$$H_i'(t) = C_i(t)[\cos(2\omega_0^i t + \theta_i(t))\sigma_x^i - \sin(2\omega_0^i t + \theta_i(t))\sigma_y^i] \quad (\text{A.94})$$

are the Bloch-Siegert shifts of the i th qubit, and

$$\begin{aligned} H_{12}'(t) &= 2[C_1(t) \cos(\omega_0^1 t + \theta_1(t))(\cos(\omega_0^2 t)\sigma_x^2 - \sin(\omega_0^2 t)\sigma_y^2) \\ &\quad + C_2(t) \cos(\omega_0^2 t + \theta_2(t))(\cos(\omega_0^1 t)\sigma_x^1 - \sin(\omega_0^1 t)\sigma_y^1)] \\ &= C_1 \cos(\Delta\omega t - \theta_1)\sigma_x^2 + C_1 \cos(\Omega t + \theta_1)\sigma_x^2 + C_2 \cos(\Delta\omega t + \theta_2)\sigma_x^1 + C_2 \cos(\Omega t + \theta_2)\sigma_x^1 - \\ &\quad C_1 \sin(\Delta\omega t - \theta_1)\sigma_y^2 - C_1 \sin(\Omega t + \theta_1)\sigma_y^2 + C_2 \sin(\Delta\omega t + \theta_2)\sigma_y^1 - C_2 \sin(\Omega t + \theta_2)\sigma_y^1 \end{aligned} \quad (\text{A.95})$$

is responsible for the leakage, where $\Delta\omega = \omega_0^2 - \omega_0^1$ and $\Omega = \omega_0^2 + \omega_0^1$. Typically, $\Omega \gg \omega_0^i$, so we can omit terms that depend on Ω . The leakage Hamiltonian becomes

$$H'_{12}(t) \approx C_1 \cos(\Delta\omega t - \theta_1) \sigma_x^2 + C_2 \cos(\Delta\omega t + \theta_2) \sigma_x^1 - C_1 \sin(\Delta\omega t - \theta_1) \sigma_y^2 + C_2 \sin(\Delta\omega t + \theta_2) \sigma_y^1. \quad (\text{A.96})$$

When $\Delta\omega$ is comparable to the Rabi frequencies, the leakage must be accounted for. In addition, large leakage may limit the efficiency of the sequences as shown in the main text.

A.10.5 Examples in spectroscopy

We look at examples for finding the set of achievable $\bar{H}^{(0)}$. Consider an internal Hamiltonian given by dipolar interactions in high field and collective control of the qubits

$$\begin{aligned} H_{\text{int}} &= \sum_{i < j}^N \omega_{ij} (2\sigma_z^i \sigma_z^j - \sigma_x^i \sigma_x^j - \sigma_y^i \sigma_y^j), \\ H_c &= \omega_1(t) \left[\cos(\phi(t)) \sum_{i=1}^N \sigma_x^i + \sin(\phi(t)) \sum_{i=1}^N \sigma_y^i \right] \end{aligned} \quad (\text{A.97})$$

Set $H_{\text{pri}} = H_c$ and $H_{\text{pert}} = H_{\text{int}}$, the space $\mathcal{C}(\mathbf{g}_{\text{pri}}, H_{\text{pert}})$ is spanned by

$$\left\{ \frac{2\sigma_z \otimes \sigma_z - \sigma_x \otimes \sigma_x - \sigma_y \otimes \sigma_y}{2\sqrt{6}}, \frac{\sigma_y \otimes \sigma_z + \sigma_z \otimes \sigma_y}{2\sqrt{2}}, \right. \\ \left. \frac{\sigma_x \otimes \sigma_z + \sigma_z \otimes \sigma_x}{2\sqrt{2}}, \frac{\sigma_y \otimes \sigma_y - \sigma_x \otimes \sigma_x}{2\sqrt{2}}, -\frac{\sigma_x \otimes \sigma_y + \sigma_y \otimes \sigma_x}{2\sqrt{2}} \right\}. \quad (\text{A.98})$$

Table A.2 compares different target zeroth-order average Hamiltonians and their achievable ranges of scaling factors. Alternatively, the Hamiltonian can be partitioned into

$$H_{\text{pri}} = H_c(t) + p H_{\text{int}}, \quad H_{\text{pert}} = (1 - p) H_{\text{int}}. \quad (\text{A.99})$$

The partitioning results in a larger primary Lie algebra, enhancing controllability of both U_{pri} and U_{pert} . For example, for a spin chain with nearest-neighbor interactions, the dimensions of $\mathcal{C}(\mathbf{g}_{\text{pri}}, H_{\text{pert}})$ are 8, 35, 96, and 542, compared to the dimensions of the total Lie algebra being 8, 38, 96, and 542, for $N=2, 3, 4$, and 5. For the case with 3 spins, Table A.3 shows some of the achievable zeroth-order average Hamiltonians.

Table A.2: Achievable ranges of scaling factors for different target effective Hamiltonians. Here, for simplicity, only unnormalized target Hamiltonians are shown. The scaling factors are from normalized Hamiltonians to normalized ones. The last target Hamiltonian is an example of arbitrary desired two-body Hamiltonians.

Target Hamiltonian	Achievable range
$\sigma_+ \otimes \sigma_- + \sigma_- \otimes \sigma_+$	Unachievable
$\sigma_z \otimes \sigma_z$	Unachievable
$\vec{\sigma} \cdot \vec{\sigma}$	Unachievable
$2\sigma_z \otimes \sigma_z - (\sigma_x \otimes \sigma_x + \sigma_y \otimes \sigma_y)$	$[-0.5, 0.99(8)]$
$\sigma_+ \otimes \sigma_+ + \sigma_- \otimes \sigma_-$	$[-0.99(2), 0.99(1)]$
$\sigma_x \otimes \sigma_y + \sigma_y \otimes \sigma_x$	$[-0.99(5), 0.99(4)]$
$2(\sigma_x \otimes \sigma_x + \sigma_y \otimes \sigma_y) + \sqrt{3}(\sigma_y \otimes \sigma_z + \sigma_z \otimes \sigma_y) - 4\sigma_z \otimes \sigma_z$	$[-0.55(2), 0.99(4)]$

Table A.3: Achievable ranges of scaling factors for different target effective Hamiltonians for a 3-spin spin chain when the interaction is included in the primary Hamiltonian.

Target Hamiltonian	Achievable range
$\sigma_+ \otimes \sigma_- + \sigma_- \otimes \sigma_+$	$[-1, 1]$
$\sigma_z \otimes \sigma_z$	$[-1, 0.99(2)]$
$\vec{\sigma} \cdot \vec{\sigma}$	$[-1, 1]$
$\sigma_x \otimes \sigma_y \otimes \sigma_z + \sigma_z \otimes \sigma_y \otimes \sigma_x$	$[-1, 1]$

A.10.6 Spin System with Two Clusters

As a first example to demonstrate how to engineer effective Hamiltonians in an ensemble system, consider parallel engineering of a qubit network divided into two clusters (Fig. A.1). The perturbative Hamiltonian is equal to the internal Hamiltonian

$$\begin{aligned} H_{\text{pert}} = H_{\text{int}} &= \sum_{i < j}^{N_1} A_{ij}^1 + \sum_{i < j}^{N_2} A_{ij}^2 + \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} A_{ij}^{12} \\ &= H_{\text{pert}}^1 + H_{\text{pert}}^2 + H_{\text{pert}}^3. \end{aligned} \quad (\text{A.100})$$

Consider collective control of qubits in individual cluster

$$\begin{aligned} H_c(t) &= \omega_1^1(t) \left(\cos(\phi_1(t)) \sum_{i=1}^{N_1} \sigma_x^{1,i} + \sin(\phi_1(t)) \sum_{i=1}^{N_1} \sigma_y^{1,i} \right) \\ &\quad + \omega_1^2(t) \left(\cos(\phi_2(t)) \sum_{i=1}^{N_2} \sigma_x^{2,i} + \sin(\phi_2(t)) \sum_{i=1}^{N_2} \sigma_y^{2,i} \right). \end{aligned} \quad (\text{A.101})$$

For example, in a dipole-coupled system in high field, the perturbative Hamiltonians are

$$\begin{aligned} H_{\text{pert}}^1(N_1) &= \sum_{i < j}^{N_1} \omega_{ij}^1 \left(2\sigma_z^{1,i} \otimes \sigma_z^{1,j} - \sigma_x^{1,i} \otimes \sigma_x^{1,j} - \sigma_y^{1,i} \otimes \sigma_y^{1,j} \right), \\ H_{\text{pert}}^2(N_2) &= \sum_{i < j}^{N_2} \omega_{ij}^2 \left(2\sigma_z^{2,i} \otimes \sigma_z^{2,j} - \sigma_x^{2,i} \otimes \sigma_x^{2,j} - \sigma_y^{2,i} \otimes \sigma_y^{2,j} \right), \\ H_{\text{pert}}^3(N_1, N_2) &= \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \omega_{ij}^{12} \sigma_z^{1,i} \otimes \sigma_z^{2,j}. \end{aligned} \quad (\text{A.102})$$

Consider decoupling H_{pert}^1 while engineering H_{pert}^2 into a double-quantum interaction

$$H_{\text{pert}}^2 \propto \sum_{ij}^{N_2} \omega_{ij}^{12} \left(\sigma_y^{2,i} \otimes \sigma_y^{2,j} - \sigma_x^{2,i} \otimes \sigma_x^{2,j} \right) \quad (\text{A.103})$$

and keeping H_{int}^{12} . $\mathcal{C}(\mathbf{g}_{\text{pri}}, H_{\text{int}}^{12})$ has a dimension of 9 and is spanned by

$$\left\{ \frac{\sigma_z}{\sqrt{2}}, \frac{\sigma_x}{\sqrt{2}}, \frac{\sigma_y}{\sqrt{2}} \right\} \otimes \left\{ \frac{\sigma_z}{\sqrt{2}}, \frac{\sigma_x}{\sqrt{2}}, \frac{\sigma_y}{\sqrt{2}} \right\}. \quad (\text{A.104})$$

In order to find the set of the achievable zeroth-order average Hamiltonians, define the target effective Hamiltonians as

$$\begin{aligned} |H_{\text{target}}^1\rangle\rangle &= (0, 0, 0, 0, 0)^T, \\ |H_{\text{target}}^2\rangle\rangle &= s(0, 0, 0, \sin \beta, 0)^T, \\ |H_{\text{target}}^3\rangle\rangle &= s(\cos \beta, 0, 0, 0, 0, 0, 0, 0, 0)^T, \end{aligned} \quad (\text{A.105})$$

with a scaling factor s . For each value of $\beta \in [0, 2\pi)$, the range of achievable scaling factors can be found using Algorithm 3. This traces out the region of all achievable effective Hamiltonians (Fig. A.11 left). As a comparison, the region for $H_{\text{eff}}^2 \propto H_{\text{int}}^2$ is shown on the right side of Fig. A.11.

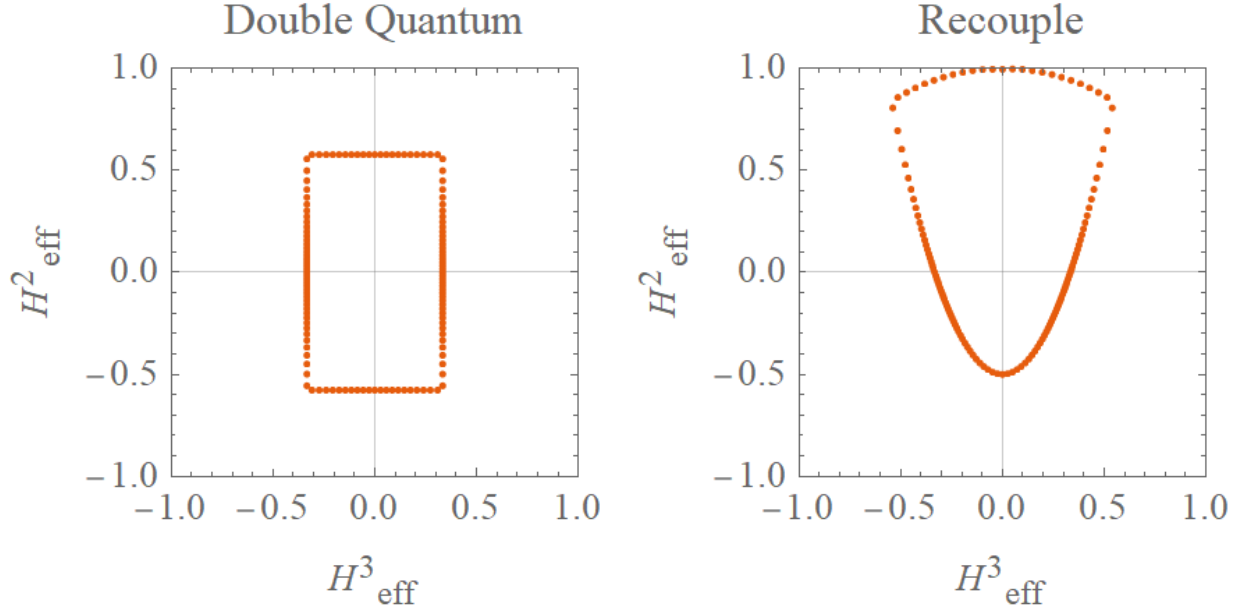


Figure A.11: Regions of achievable effective Hamiltonians. The dots are $(s_{\pm} \cos \theta, s_{\pm} \sin \theta)$ for $\theta = 0, \frac{\pi}{64}, \dots, \frac{63\pi}{64}$. They trace out the convex set for all achievable 0th-order average Hamiltonians.

Appendix B

Supplementary: Efficiently characterizing quantum information flow, loss and recovery in the central spin system

B.1 Coherence Pathway

All experiments in the thesis use the technique of coherence transfer pathways selection [58], thus a brief review is presented here.

Terms such as $|t\rangle\langle u|$ of a density matrix in high-field NMR is said to have coherence order $p_{tu} = M_t - M_u$, where M_t is the magnetic quantum number of the state $|t\rangle$. Therefore, each transition is associated with two coherences $|t\rangle\langle u|$ and $|u\rangle\langle t|$. The secular Hamiltonian always preserves the coherence order. It is the RF pulses that may induce a change in coherence order by transferring $|t\rangle\langle u| \rightarrow |u\rangle\langle t|$. In order to use pathways to describe a pulse experiment, we write the density matrix into terms of different coherence orders

$$\rho(t) = \sum_p \rho^p(t). \quad (\text{B.1})$$

For a system of K spins $1/2$, $-K \leq p \leq K$. The observable by convention is always $J_+ = J_x + iJ_y$. Therefore, only the terms proportional to J_- , which has coherence order $p = -1$, contributes to the signal.

The p -quantum coherence term is an eigenstate of z rotations

$$\exp(-iJ_z\phi)\rho^p\exp(iJ_z\phi) = \rho^p\exp(-ip\phi). \quad (\text{B.2})$$

This property allows us to separate different orders. A propagator U_i , a pulse or a sequence, may cause transfer of coherences

$$U_i\rho^p(t_i^-)U_i^\dagger = \sum_{p'}\rho^{p'}(t_i^+). \quad (\text{B.3})$$

The phase-shifted propagator $U_i(\phi_i) = \exp(-iJ_z\phi_i)U_i(0)\exp(iJ_z\phi_i)$ introduces an additional phase,

$$U_i(\phi_i)\rho^p(t_i^-)U_i^\dagger(\phi_i) = \sum_{p'}\rho^{p'}(t_i^+)\exp(-i\Delta p_i\phi_i), \quad (\text{B.4})$$

where $\Delta p_i = p'(t_i^+) - p(t_i^-)$ corresponds to change in coherence order under the propagator U_i . The phase change will be carried over to the final observable. The total signal is then given by

$$s(\phi_i, t) = \sum_{\Delta p_i} s(\Delta p_i, \phi_i = 0, t) \exp(-i\Delta p_i\phi_i). \quad (\text{B.5})$$

To restrict the coherence transfer, we can do N_i experiments with systematic increment of $\phi_i = k_i 2\pi/N_i$ with $k_i = 0, 1, \dots, N_i - 1$. A discrete Fourier transform gives

$$s(\Delta p_i, t) = \frac{1}{N_i} \sum_{k_i=0}^{N_i-1} s(\phi_i, t) \exp(i\Delta p_i\phi_i). \quad (\text{B.6})$$

Notably, this procedure retains other pathways

$$\Delta p_i^{(\text{selected})} = \Delta p_i^{(\text{desired})} \pm nN_i, \quad n = 0, 1, 2, \dots \quad (\text{B.7})$$

This is reminiscent of aliasing in Fourier transform. It is helpful to write down all possible changes in coherence order, and emphasize the desired change by bold typeface, and indicate Δp_i that must be blocked by setting them in brackets. For example the following notation

$$\Delta p_2 = \dots, -3, -\mathbf{2}, (-1), (0), 1, 2, 3, \dots \quad (\text{B.8})$$

means that we want to select $\Delta p_i = -2$ while rejecting -1 and 0. Other changes are immaterial in this case, even though they may happen. The selection is achieved by shifting the receiver phase by $\phi_{\text{ref}} = -\Delta p_i\phi_i$. Then by selecting a proper N_i , we can block other changes as needed. Here, the minimal value of N_i is 3. Another example of selection is the Exorcycle that selects the pathway $p = 0 \rightarrow +1 \rightarrow -1$ in an echo experiment.

In many cases, one needs to select several consecutive coherence changes

$$\Delta \mathbf{p} = \{\Delta p_1, \dots, \Delta p_n\}. \quad (\text{B.9})$$

Since all pathways start with $p = 0$ and end with $p = -1$, we have

$$\sum_{i=1}^n \Delta p_i = -1. \quad (\text{B.10})$$

If $n - 1$ values of Δp_i are specified by $n - 1$ independent phase cycles, $\Delta \mathbf{p}$ is defined unambiguously. However, to eliminate systematic errors, we often apply independent phase cycles to select Δp_i for all n steps. Thus, one can choose N_i for each Δp_i and cycle the phases independently for $\phi_1 = k_1 2\pi/N_1, \dots, \phi_n = k_n 2\pi/N_n$. The number of experiments is $N = N_1 N_2 \dots N_n$. Because of the sum rule given by Eq. (B.10), it is possible to skip the phase cycling for one of the propagators. To select the desired pathway, the signals need to be combined according to an n -dimensional discrete Fourier transform

$$s(\Delta \mathbf{p}, t) = \frac{1}{N} \sum_{k_1=0}^{N_1-1} \dots \sum_{k_n=0}^{N_n-1} s(\phi, t) \exp(i \Delta \mathbf{p} \cdot \phi). \quad (\text{B.11})$$

This means that we need to set the receiver phase by

$$\phi_{\text{ref}} = -\Delta \mathbf{p} \cdot \phi = -\sum_i \Delta p_i \phi_i. \quad (\text{B.12})$$

We use a diagram such as Fig. C.1 to represent a coherence pathway and its corresponding phase cycling.

B.2 Detailed analysis of the stimulated echo

The following Hamiltonians of the system are considered:

$$H_{cs/e} = \sigma_z \otimes A/2, \quad H_e = \mathbb{1} \otimes B, \quad (\text{B.13})$$

where A and B are operators acting on the environment. The CS/environment interaction is axially symmetric about the z direction. The initial state of the system is $\rho(0) = \sigma_x \otimes \mathbb{1}^{\otimes N}$, where N is the number of spins in the environment. The identity part of the density matrix is excluded, as it does not contribute to the signal. This state represents a scenario where the quantum information initially resides at the CS.

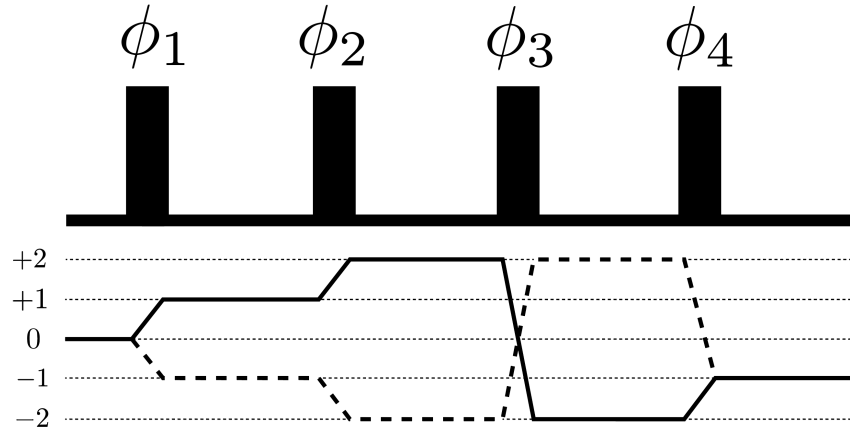


Figure B.1: Coherence pathway. On top is a sequence and the corresponding phases. The black solid squares are propagators that induce changes of coherence orders. The reference phase is not shown here. In the simplest case, these are RF pulses. The possible coherence orders are -2, -1, 0, +1 and +2, indicated by the dotted lines. The solid line shows the desired coherence pathway $p = 0 \rightarrow +1 \rightarrow +2 \rightarrow -2 \rightarrow -1$. It always starts with 0 coherence order and ends with -1 coherence order. The dashed line is a “mirror image” pathway that is important for obtaining pure phase spectrum in multi-dimensional experiments.

The unitary propagator of the experiment in the interaction frame of H_e is

$$U_{\text{STE}} = \exp(iH'_{cs/e}(\Delta)t) \exp(-iH'_{cs/e}(0)t). \quad (\text{B.14})$$

Therefore, the signal is

$$S_{\text{STE}} = \text{Tr}(\exp(-iH'_{cs/e}(0)t)\rho(0) \exp(iH'_{cs/e}(0)t) \exp(-iH'_{cs/e}(\Delta)t)\rho_{\text{ref}}^\dagger \exp(iH'_{cs/e}(\Delta)t)), \quad (\text{B.15})$$

where $\rho_{\text{ref}} = \sigma_- \otimes \mathbb{1}^{\otimes N}/2^{N+1}$ is the observable corresponding to a measurement on the CS. Substitute in $H'_{cs/e}(\Delta) = \sigma_z \otimes A'(\Delta)/2$ where $A'(\Delta) = \exp(iB\Delta)A \exp(-iB\Delta)$ and expand with respect to t , we have

$$\begin{aligned} \exp(-iH'_{cs/e}(0)t)\rho(0) \exp(iH'_{cs/e}(0)t) &= \sigma_x \otimes \cos(A'(0)t) + \sigma_y \otimes \sin(A'(0)t), \\ \exp(-iH'_{cs/e}(\Delta)t)\rho_{\text{ref}} \exp(iH'_{cs/e}(\Delta)t) &= \sigma_+ \otimes \exp(-iA'(\Delta)t)/2^{N+1}. \end{aligned} \quad (\text{B.16})$$

Thus, the signal becomes

$$\begin{aligned} S_{\text{STE}} &= \text{Tr}(\cos(A'(0)t) \exp(-iA'(\Delta)t) + i \sin(A'(0)t) \exp(-iA'(\Delta)t)) / 2^N \\ &= \text{Tr}(\exp(iA'(0)t) \exp(-iA'(\Delta)t)) / 2^N. \end{aligned} \quad (\text{B.17})$$

Note that S_{STE} is invariant under the reversal of t ; thus, it is convenient to write

$$S_{\text{STE}} = \langle \exp(iA'(\Delta)t) \exp(-iA'(0)t) \rangle, \quad (\text{B.18})$$

which ensures consistency with the experiment's time order. The normalization constant is absorbed in the average $\langle \cdot \rangle = \text{Tr}(\cdot)/2^N$ so that the signal is normalized to 1 at $\Delta = 0$.

The spectrum $\tilde{S}_{\text{STE}}$ is defined as the Fourier transform of S_{STE} with respect to t :

$$\tilde{S}_{\text{STE}}(\omega) = \frac{1}{\sqrt{2\pi}} \int e^{i\omega t} S_{\text{STE}}(t) dt. \quad (\text{B.19})$$

Since

$$\frac{1}{\sqrt{2\pi}} \int \omega^n e^{i\omega t} = (-i)^n \sqrt{2\pi} \delta^{(n)}(t), \quad (\text{B.20})$$

and

$$\int f(t) \delta^{(n)}(t) dt = \frac{d^n f}{dt^n}(0), \quad (\text{B.21})$$

we have

$$m_k(\tilde{S}_{\text{STE}}) \equiv \int d\omega \omega^k \tilde{S}_{\text{STE}} = \sqrt{2\pi} (-i)^k \frac{d^k S_{\text{STE}}}{dt^k}(0). \quad (\text{B.22})$$

Thus,

$$\begin{aligned} m_2(\tilde{S}_{\text{STE}}) &= \langle A'(\Delta)A'(\Delta) - A'(\Delta)A'(0) - A'(0)A'(\Delta) - A'(0)A'(0) \rangle \\ &= 2\langle A^2 \rangle - 2\langle A'(\Delta)A'(0) \rangle. \end{aligned} \quad (\text{B.23})$$

Thus, the second moment gives a direct measurement to the autocorrelation function $\langle A'(\Delta)A'(0) \rangle$ which measures the change of local field for a given Δ .

Expand the autocorrelation function with respect to Δ :

$$\langle A'(\Delta)A'(0) \rangle = \langle A^2 \rangle + i\Delta\langle [B, A]A \rangle - \frac{\Delta^2}{2}\langle [B, [B, A]]A \rangle + \dots \quad (\text{B.24})$$

It is easy to check that all the odd-order terms vanish, which can also be confirmed by noting that S_{STE} is an even function of Δ . Therefore, the expansion simplifies to

$$\langle A'(\Delta)A'(0) \rangle = \langle A^2 \rangle - \frac{\Delta^2}{2}\langle [B, [B, A]]A \rangle + \frac{\Delta^4}{4!}\langle [B, [B, [B, [B, A]]]]A \rangle + \dots \quad (\text{B.25})$$

The non-commutativity between A and B (between $H_{cs/e}$ and H_e) can be obtained by analyzing the time dependence of the second moment of $\tilde{S}_{\text{STE}}$.

Note that the signal of the experiment can also be written as

$$\begin{aligned} S_{\text{STE}} &= \langle \exp(-iA'(\Delta)t) \exp(iA'(0)t) \rangle \\ &= \langle \exp(iB\Delta) \exp(-iAt) \exp(-iB\Delta) \exp(iAt) \rangle \\ &= \langle \exp(iB'(t)\Delta) \exp(-iB'(0)\Delta) \rangle, \end{aligned} \quad (\text{B.26})$$

where $B'(t) = \exp(iAt)B \exp(-iAt)$. Denote by $\tilde{S}'_{\text{STE}}$ the spectrum obtained by taking Fourier transform of S_{STE} with respect to Δ , the second moment of the spectrum is

$$m_2(\tilde{S}'_{\text{STE}}) = 2\langle B^2 \rangle - 2\langle B'(t)B'(0) \rangle. \quad (\text{B.27})$$

The auto-correlation function can be similarly expanded as

$$\langle B'(t)B'(0) \rangle = \langle B^2 \rangle - \frac{t^2}{2}\langle [A, [A, B]]B \rangle + \frac{t^4}{4!}\langle [A, [A, [A, [A, B]]]]B \rangle + \dots \quad (\text{B.28})$$

B.3 Bayesian analysis for the stimulated echo

Bayesian analysis is used to extract information from the experiment data and the essential steps of the procedure are described here. Each point x_i in the time-domain data takes the form

$$x_i = y_i + \epsilon_i, \quad (\text{B.29})$$

where y_i is the corresponding true value and ϵ_i is a zero mean white noise $\sim \mathcal{N}(0, \sigma^2)$. The variance σ^2 can be evaluated accurately in advance, so it is not included in the parameters to be estimated. Therefore, the likelihood function $p(x_i|y_i)$ for each x_i is $\mathcal{N}(y_i, \sigma^2)$. If we choose the improper prior $p(x_i) = 1$, then the likelihood function is also the posterior distribution $p(y_i|x_i)$.

The posterior predictive distributions of functions f of y_i can be simulated with samples from $p(\{y_i\}|\{x_i\})$, such as the second moments derived from linewidths or decay rates. Once the posterior predictive distributions are obtained, an estimate is given by taking their means (Bayes estimator).

According to Eq. (B.23), the L th-order model M_L of the second moment is given by Eq. (4.11):

$$M_L(\Delta, \vec{a}) = C + \sum_{l=1}^L p_l \exp(-\Delta^2/\lambda_l), \quad (\text{B.30})$$

where $\vec{a} = (C, p_1, \lambda_1, \dots, p_L, \lambda_L)^T$ are the parameters to be estimated. The constant C absorbs the initial linewidth of the spectra and possible constant of motion so that the focus is on the time-dependence of the model. The likelihood function $p(\{x_i\}|\vec{a})$ can be connected to the second moments of different Δ through

$$p(\{x_i\}|\vec{a}) = \int d\{m_2\} p(\{x_i\}|\{m_2\}) p(\{m_2\}|\vec{a}), \quad (\text{B.31})$$

where $\{m_2\}$ is the collection of second moments for different Δ . Then, $p(\{m_2\}|\vec{a})$ is given by the model M_L :

$$p(\{m_2\}|\vec{a}) = \prod_{\Delta} \delta(m_2(\Delta) - M_L(\Delta, \vec{a})). \quad (\text{B.32})$$

the density $p(\{x_i\}|\{m_2\})$ can be evaluated using the Bayes' rule:

$$p(\{x_i\}|\{m_2\}) = \frac{p(\{m_2\}|\{x_i\})p(\{x_i\})}{p(\{m_2\})}. \quad (\text{B.33})$$

Here, $p(\{m_2\}|\{x_i\})$ are the posterior predictive distributions of m_2 , $p(\{x_i\})$ is a constant independent of $\vec{a}$. Choose the improper noninformative prior $p(\{m_2\}) = 1$, we have $p(\{x_i\}|\{m_2\}) = p(\{m_2\}|\{x_i\})$. Substituting this into the integral, the likelihood function becomes

$$p(\{x_i\}|\vec{a}) = \prod_{\Delta} p(m_2(\Delta) = M_L(\Delta, \vec{a})|\{x_i\}). \quad (\text{B.34})$$

We use the Bayesian information criterion (BIC) [172, 165] to evaluate the goodness of each M_L which is an approximation of the Bayes factor [94] widely used in null hypothesis

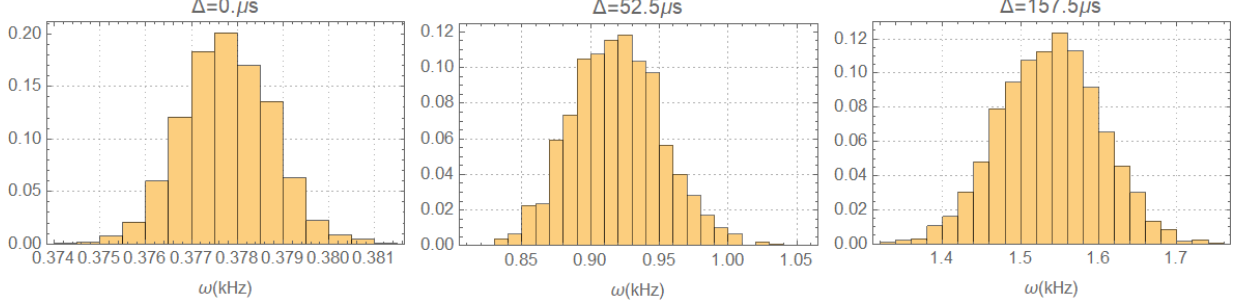


Figure B.2: Posterior distributions of derived second moments from the measured linewidths of the spectra. They are well approximated by Gaussian distributions.

testing. It is defined as

$$\text{BIC} = n_{\text{par}} \ln(n_{\text{data}}) - 2 \ln(\hat{\mathcal{L}}), \quad (\text{B.35})$$

where n_{par} is the number of parameters in the model, n_{data} is the number of data points and $\hat{\mathcal{L}}$ is the maximum likelihood. The first term serves as an Occam's razor that penalizes the complexity of the model while the second term evaluates how well the model fits the data. As a comparison, the Akaike information criterion (AIC) [2] is also calculated, which is defined as

$$\text{AIC} = 2n_{\text{par}} - 2 \ln(\hat{\mathcal{L}}). \quad (\text{B.36})$$

AIC is more forgiving to model's complexity and tries to find the model that best fits the data without overfitting, whereas BIC is trying to find the most plausible underlying model given the available evidence, which is more suitable for our need here.

Instead of using the raw data $\{x_i\}$, we calculate the likelihood function $p(\vec{a}|\{m_2\})$. For convenience of optimization, $p(m_2|\{x_i\})$ are approximated by Gaussian distributions with the same mean and variance which is justified by simulated shapes of the $p(m_2|\{x_i\})$ (Fig. B.2). A generalized simulated annealing optimizer [191, 169] is employed to find the maximum likelihood for each model. Parameter ranges are selected to ensure they are wide enough to encompass any reasonable parameter values. The corresponding information criterion values, based on the maximum likelihood identified in the optimization, are plotted in Fig. B.3. The result shows that both BIC and AIC favor the model with $L = 2$ for $\langle H'_{cs/e}(\Delta)H'_{cs/e}(0) \rangle$ and $L = 1$ for $\langle H'_e(t)H'_e(0) \rangle$.

The posterior distribution of the parameters is given by

$$p(\vec{a}|\{m_2\}) \propto p(\{m_2\}|\vec{a})p(\vec{a}), \quad (\text{B.37})$$

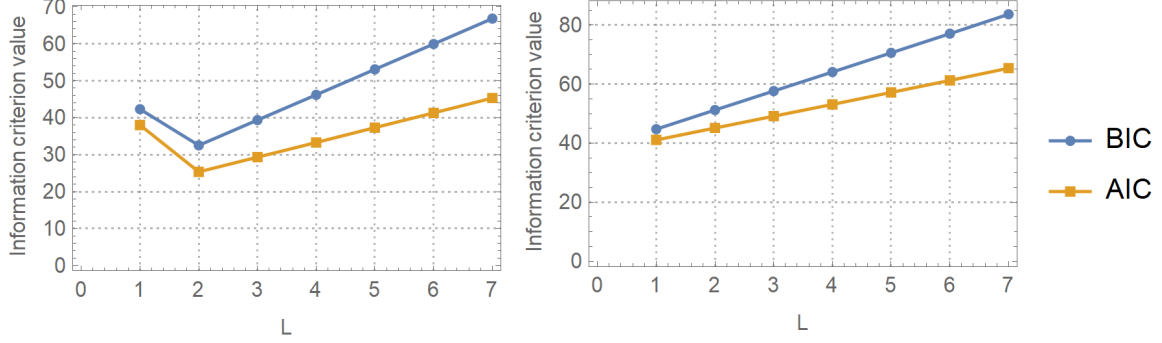


Figure B.3: Information criterion values for models of different order L for $\langle H'_{cs/e}(\Delta)H'_{cs/e}(0) \rangle$ (left) and $\langle H'_e(t)H'_e(0) \rangle$ (right).

and an adaptive Metropolis-Hastings sampler [78] is used to simulate the distributions. The prior distributions for the parameters in $\vec{a}$ are chosen as independent logit-normal distributions with $\sigma = 1$ over the same ranges used in the optimization. The ranges are large enough so that the priors are noninformative. Then the parameters are transform to $(-\infty, +\infty)$ through a logit function to achieve better performance with the sampler.

Once the posterior distributions of the parameters $C, p_1, \lambda_1, \dots, p_L, \lambda_L$ are obtained for the desired model ($L = 2$). Then, the non-commutativity is measured using the corresponding posterior distributions through the following relation:

$$C_k \text{ or } D_k = -\frac{1}{2} \frac{1}{p_1 + \dots + p_L} \frac{k!}{(k/2)!} \sum_{l=1}^L (-1)^{k/2} \frac{p_l}{\lambda_l^{k/2}}, \quad (\text{B.38})$$

where $C_k = \langle [B, [B, A]]A \rangle, \dots$ and $D_k = \langle [A, [A, B]]B \rangle, \dots$, for $k = 2, 4, \dots$ being the level of the nested commutator, are the measured commutation relations. Note that the result is normalized with respect to $\langle A^2 \rangle$ and $\langle B^2 \rangle$ so that $C_0 = D_0 = 1$.

Appendix C

Supplementary: Experimentally Characterizing Fluctuation of Noise in the Central Spin System

C.1 Dynamics of the Central Spin System

Theorem C.1.1. *The signal on the central spin $S(t) = \text{Tr}(\rho(0)\sigma_+ \otimes \mathbb{1}^{\otimes N})$ is*

$$S(t) = \left\langle \mathcal{T} \exp \left(i \int_0^t ac_{A'(t')} dt' / 2 \right) \right\rangle = \exp(K(t)), \quad (\text{C.1})$$

where

$$K(t) = \frac{i}{2} \int_0^t dt_1 \langle ac_{A'(t_1)} \rangle_c - \frac{1}{4} \int_0^t dt_1 \int_0^{t_1} dt_2 \langle ac_{A'(t_1)} ac_{A'(t_2)} \rangle_c + \dots, \quad (\text{C.2})$$

where the anti-commutator is defined as $ac_A(B) = \{A, B\}$ and $\langle \cdot \rangle_c$ is the generalized cumulant [109] defined with respect to $\langle \cdot \rangle = \text{Tr}(\cdot \rho_e(0))$.

Proof. We work with the interaction frame Hamiltonian of the system:

$$H'(t) = \sigma_z \otimes A'(t)/2, \quad (\text{C.3})$$

where $A'(t) = \exp(iBt)A\exp(-iBt)$. The unitary propagator is

$$\begin{aligned}
U(t) &= \mathcal{T}_{\leftarrow} \exp \left[-i \int_0^t dt_1 \sigma_z \otimes A'(t_1)/2 \right] \\
&= \mathbb{1} + \left(\frac{-i}{2} \right) \sigma_z \otimes \int_0^t dt_1 A'(t_1) + \frac{1}{2!} \left(\frac{-i}{2} \right)^2 \mathbb{1} \otimes \int_0^t dt_1 \int_0^{t_1} dt_2 A'(t_1) A'(t_2) + \dots \quad (\text{C.4}) \\
&= \mathbb{1} \otimes \mathcal{T}_{\leftarrow} \cos \left[\int_0^t dt_1 A'(t_1)/2 \right] - i \sigma_z \otimes \mathcal{T}_{\leftarrow} \sin \left[\int_0^t dt_1 A'(t_1)/2 \right].
\end{aligned}$$

Here,

$$\mathcal{T}_{\leftarrow} \cos \left[\int_0^t dt_1 A'(t_1)/2 \right] \pm i \mathcal{T}_{\leftarrow} \sin \left[\int_0^t dt_1 A'(t_1)/2 \right] = \mathcal{T}_{\leftarrow} \exp \left[\pm i \int_0^t dt_1 A'(t_1)/2 \right]. \quad (\text{C.5})$$

Denote

$$\begin{aligned}
C_{\leftarrow} &= \mathcal{T}_{\leftarrow} \cos \left[\int_0^t dt_1 A'(t_1)/2 \right], \\
S_{\leftarrow} &= \mathcal{T}_{\leftarrow} \sin \left[\int_0^t dt_1 A'(t_1)/2 \right].
\end{aligned} \quad (\text{C.6})$$

The signal is (with $\sigma_+ = (\sigma_x + i\sigma_y)/2$)

$$\begin{aligned}
S(t) &= \text{Tr} \left[\sigma_+ \otimes \mathbb{1}_e U(t) \rho(0) U^\dagger(t) \right] \\
&= \text{Tr} \left[\sigma_+ \otimes \mathbb{1}_e (\mathbb{1} \otimes C_{\leftarrow} - i \sigma_z \otimes S_{\leftarrow}) \rho_s(0) \otimes \rho_e(0) (\mathbb{1} \otimes C_{\rightarrow} + i \sigma_z \otimes S_{\rightarrow}) \right].
\end{aligned} \quad (\text{C.7})$$

Let

$$\rho_s(0) = \frac{\mathbb{1}}{2} + n_x \sigma_x + n_y \sigma_y + n_z \sigma_z. \quad (\text{C.8})$$

Only $n_x \sigma_x$ and $n_y \sigma_y$ contribute to the signal. Therefore, the signal becomes

$$S(t) = \text{Tr} \left[\sigma_+ \otimes \mathbb{1}_e (\mathbb{1} \otimes C_{\leftarrow} - i \sigma_z \otimes S_{\leftarrow}) (n_x \sigma_x + n_y \sigma_y) \otimes \rho_e(0) (\mathbb{1} \otimes C_{\rightarrow} + i \sigma_z \otimes S_{\rightarrow}) \right]. \quad (\text{C.9})$$

Since

$$\begin{aligned}
&(\mathbb{1} \otimes C_{\leftarrow} - i \sigma_z \otimes S_{\leftarrow}) (n_x \sigma_x + n_y \sigma_y) \otimes \rho_e(0) (\mathbb{1} \otimes C_{\rightarrow} + i \sigma_z \otimes S_{\rightarrow}) \\
&= (n_x \sigma_x + n_y \sigma_y) \otimes (C_{\leftarrow} \rho_e(0) C_{\rightarrow} - S_{\leftarrow} \rho_e(0) S_{\rightarrow}) + (n_x \sigma_y - n_y \sigma_x) \otimes (C_{\leftarrow} \rho_e(0) S_{\rightarrow} + S_{\leftarrow} \rho_e(0) C_{\rightarrow}),
\end{aligned} \quad (\text{C.10})$$

we have

$$\begin{aligned}
S(t) &= \text{Tr} \{ (n_x + in_y) [C_{\leftarrow} \rho_e(0) C_{\rightarrow} - S_{\leftarrow} \rho_e(0) S_{\rightarrow}] + (-n_y + n_x i) [C_{\leftarrow} \rho_e(0) S_{\rightarrow} + S_{\leftarrow} \rho_e(0) C_{\rightarrow}] \} \\
&= \text{Tr} \left\{ (n_x + in_y) \frac{1}{2} [U_{\leftarrow}^+ \rho_e(0) U_{\rightarrow}^+ + U_{\leftarrow}^- \rho_e(0) U_{\rightarrow}^-] + (-n_y + n_x i) \frac{1}{2i} [U_{\leftarrow}^+ \rho_e(0) U_{\rightarrow}^+ - U_{\leftarrow}^- \rho_e(0) U_{\rightarrow}^-] \right\} \\
&= S(0) \text{Tr} [U_{\leftarrow}^+ \rho_e(0) U_{\rightarrow}^+] \\
&= S(0) \left\langle \mathcal{T} \exp \left(i \int_0^t \text{ac}_{A'(t')} dt' / 2 \right) \right\rangle,
\end{aligned} \tag{C.11}$$

where $\langle \cdot \rangle = \text{Tr}(\cdot \rho_e(0))$, and

$$U_{\leftarrow/\rightarrow}^{\pm} = \mathcal{T}_{\leftarrow/\rightarrow} \exp \left[\pm i \int_0^t dt_1 A'(t_1) / 2 \right]. \tag{C.12}$$

Use the generalized cumulant expansion, we have

$$\left\langle \mathcal{T} \exp \left(i \int_0^t \text{ac}_{A'(t')} dt' / 2 \right) \right\rangle = \exp(K(t)), \tag{C.13}$$

where

$$\begin{aligned}
K(t) &= \left\langle \mathcal{T} \exp \left(i \int_0^t \text{ac}_{A'(t')} dt' / 2 \right) - 1 \right\rangle_c \\
&= \frac{i}{2} \int_0^t dt_1 \langle \text{ac}_{A'(t_1)} \rangle_c - \frac{1}{4} \int_0^t dt_1 \int_0^{t_1} dt_2 \langle \text{ac}_{A'(t_1)} \text{ac}_{A'(t_2)} \rangle_c + \dots
\end{aligned} \tag{C.14}$$

□

The cumulants can be calculated from the multi-time correlation functions.

$$\langle \text{ac}_{A_n} \dots \text{ac}_{A_2} \text{ac}_{A_1} \rangle = 2 \langle \{A_n, \{ \dots, \{A_2, A_1\} \} \} \rangle. \tag{C.15}$$

The unitary propagator of an n D experiment is

$$U_{\text{nD}} = \exp(-iA'_n t_n) \dots \exp(-iA'_1 t_1). \tag{C.16}$$

Thus, the signal of an n D experiment is

$$S_{\text{nD}}(t_1, \dots, t_n; \Delta_1, \dots, \Delta_{n-1}) = \langle \exp(i \text{ac}_{A'_n} t_n / 2) \dots \exp(i \text{ac}_{A'_2} t_2 / 2) \exp(i A'_1 t_1) \rangle. \tag{C.17}$$

The Fourier transform of the exponential of an operator can be formally written as

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} dt e^{i\omega t} \exp(-iAt) = \sqrt{2\pi} \delta(\omega - A), \quad (\text{C.18})$$

where A can be diagonalized as $A = TDT^\dagger$ and $\delta(\omega - A) = T\delta(\omega - D)T^\dagger$ for any Hermitian operator A . Then the n D spectrum is given by

$$\tilde{S}_{nD} = (2\pi)^{n/2} \left\langle \delta(\omega_n + \text{ac}_{A'_n}/2) \cdots \delta(\omega_2 + \text{ac}_{A'_2}/2) \delta(\omega_1 + A'_1) \right\rangle. \quad (\text{C.19})$$

If $[A'(\Delta), A'(0)] = 0$, we have

$$\tilde{S}_{nD} = p(\Omega'(t_n), \Omega'(t_{n-1}), \dots, \Omega'(t_1)). \quad (\text{C.20})$$

However, $\tilde{S}_{nD}$ generally does not have a probabilistic interpretation and higher-order correlations depend on non-commutativity of A' at different times.

Projections of $\tilde{S}_{nD}$ can be obtained from reduced n D experiments. The projections of $\tilde{S}_{2D}$ along $\omega_1 = \omega_2$ and $\omega_1 = -\omega_2$ correspond to spectra of the stimulated echo and anti-echo experiments. The second moments of their spectra give

$$\begin{aligned} m_2(\tilde{S}_{2D}^1) &= \langle A_2'^2 + 2A_2'A_1' + A_1'^2 \rangle, \\ m_2(\tilde{S}_{2D}^{-1}) &= \langle A_2'^2 - 2A_2'A_1' + A_1'^2 \rangle, \end{aligned} \quad (\text{C.21})$$

which can be used to estimate the autocorrelation function.

C.2 Properties of the Ornstein–Uhlenbeck Process

The statistical properties of the Ornstein-Uhlenbeck process have been used to simulate the predicted data. Here, we briefly show how to calculate them. The distribution at any time t is

$$p(x, t) = \sqrt{\frac{\theta}{2\pi D}} \exp\left(-\frac{\theta x^2}{2D}\right). \quad (\text{C.22})$$

The transition probability is

$$p(x_2, t_2 | x_1, t_1) = \sqrt{\frac{\theta}{2\pi D[1 - e^{-2\theta(t_2-t_1)}]}} \exp\left\{-\frac{\theta}{2D} \frac{[x_2 - x_1 e^{-\theta(t_2-t_1)}]^2}{1 - e^{-2\theta(t_2-t_1)}}\right\}. \quad (\text{C.23})$$

Since

$$\begin{aligned} & \int_{-\infty}^{+\infty} dx_1 x_1^n \exp[-a_1(x_2 - \eta_1 x_1)^2] \exp(-\tilde{a}_0 x_1^2) \\ &= \exp\left(-\frac{a_1 \tilde{a}_0}{\tilde{a}_0 + a_1 \eta_1^2} x_2^2\right) (\tilde{a}_0 + a_1 \eta_1^2)^{-\frac{n+2}{2}} \begin{cases} n a_1 \eta_1 \Gamma\left(\frac{n}{2}\right) x_2 \sum_{k=0}^{(n-1)/2} c_k(n) \left(\frac{a_1^2 \eta_1^2 x_2^2}{\tilde{a}_0 + a_1 \eta_1^2}\right)^k & n \text{ is odd} \\ \sqrt{\tilde{a}_0 + a_1 \eta_1^2} \Gamma\left(\frac{n+1}{2}\right) \sum_{k=0}^{n/2} c_k(n) \left(\frac{a_1^2 \eta_1^2 x_2^2}{\tilde{a}_0 + a_1 \eta_1^2}\right)^k & n \text{ is even} \end{cases}, \end{aligned} \quad (\text{C.24})$$

where

$$c_k(n) = \frac{1}{k!} \left. \frac{\partial^k G(x, n)}{\partial x^k} \right|_{x=0}, \quad (\text{C.25})$$

with

$$G(x, n) = e^{-x} \begin{cases} {}_1F_1\left(1 + \frac{n}{2}; \frac{3}{2}; x\right) & n \text{ is odd} \\ {}_1F_1\left(\frac{1+n}{2}; \frac{1}{2}; x\right) & n \text{ is even} \end{cases}. \quad (\text{C.26})$$

The correlation function is

$$\begin{aligned} & \langle X(t_n) \cdots X(t_1) \rangle \\ &= \frac{\sqrt{a_{n-1} \cdots a_1 \tilde{a}_0}}{\pi^{n/2}} \int dx_n \cdots \int dx_1 x_n e^{-a_{n-1}(x_n - \eta_{n-1} x_{n-1})^2} \cdots x_2 e^{-a_1(x_2 - \eta_1 x_1)^2} x_1 e^{-\tilde{a}_0 x_1^2}. \end{aligned} \quad (\text{C.27})$$

Define

$$\mathcal{L}_k = \int dx_k e^{-a_k(x_{k+1} - \eta_k x_k)^2} x_k, \quad (\text{C.28})$$

then the integral in the correlation function becomes $\int dx_n x_n \mathcal{L}_{n-1} \cdots \mathcal{L}_1 e^{-\tilde{a}_0 x_1^2}$. Define $I_0 = e^{-\tilde{a}_0 x_1^2}$ and $I_{k+1} = \mathcal{L}_{k+1} I_k$.

$$I_k = e^{-\tilde{a}_k x_{k+1}^2} \begin{cases} \sum_{j=0}^{(k-1)/2} C_j(k) x_{k+1}^{2j+1} & k \text{ is odd} \\ \sum_{j=0}^{k/2} C_j(k) x_{k+1}^{2j} & k \text{ is even} \end{cases}, \quad (\text{C.29})$$

where

$$\tilde{a}_k = \frac{a_k \tilde{a}_{k-1}}{\tilde{a}_{k-1} + a_k \eta_k^2}. \quad (\text{C.30})$$

Using the recurrence relation Eq. (C.24), we have

$$C_k(n) = \left(\frac{a_n^2 \eta_n^2}{\tilde{a}_{n-1} + a_n \eta_n^2} \right)^k \begin{cases} \sum_{l=k}^{(n-1)/2} C_l(n-1) (\tilde{a}_{n-1} + a_n \eta_n^2)^{-\frac{2l+3}{2}} (2l+1) a_n \eta_n \Gamma\left(\frac{2l+1}{2}\right) c_k(2l+1) & n \text{ is odd} \\ \sum_{l=\max\{k-1, 0\}}^{n/2-1} C_l(n-1) (\tilde{a}_{n-1} + a_n \eta_n^2)^{-\frac{2l+3}{2}} \Gamma\left(\frac{2l+3}{2}\right) c_k(2l+2) & n \text{ is even} \end{cases}. \quad (\text{C.31})$$

Finally, the correlation function is given by

$$\begin{aligned} \langle X(t_n) \cdots X(t_1) \rangle &= \frac{\sqrt{a_{n-1} \cdots a_1 \tilde{a}_0}}{\pi^{n/2}} \int dx_n x_n I_{n-1} \\ &= \begin{cases} 0 & n \text{ is odd} \\ \frac{\sqrt{a_{n-1} \cdots a_1 \tilde{a}_0}}{\pi^{n/2}} \sum_{j=0}^{(n-2)/2} C_j(n-1) \tilde{a}_{n-1}^{-\frac{2j+3}{2}} \Gamma\left(\frac{2j+3}{2}\right) & n \text{ is even} \end{cases}. \end{aligned} \quad (\text{C.32})$$

Let $\tilde{a}_0 = \frac{\theta}{2D}$, $\eta_k = \exp(-\theta\Delta_k)$ and $a_k = \frac{\theta}{2D(1-\eta_k^2)}$ with $\Delta_k = t_{k+1} - t_k$, we show below examples of the correlation functions

$$\begin{aligned} \langle X(t_2)X(t_1) \rangle &= \frac{D}{\theta} e^{-\theta\Delta_1}, \\ \langle X(t_4)X(t_3)X(t_2)X(t_1) \rangle &= \frac{D^2}{\theta^2} e^{-\theta(\Delta_1+2\Delta_2+\Delta_3)} (2 + e^{2\theta\Delta_2}), \\ \langle X(t_6)X(t_5)X(t_4)X(t_3)X(t_2)X(t_1) \rangle \\ &= \frac{D^3}{\theta^3} e^{-\theta(\Delta_1+2\Delta_2+3\Delta_3+2\Delta_4+\Delta_5)} [6 + e^{2\theta\Delta_3} (2 + e^{2\theta\Delta_2}) (2 + e^{2\theta\Delta_4})]. \end{aligned} \quad (\text{C.33})$$

Alternatively, one can use the Isserlis' theorem to derive the results, since the only nonzero cumulant is of second order.

C.3 Bayesian analysis of time-domain data

The time-domain data of the 2D and 3D experiments are fitted to that predicted by the Gauss-Markov model to provide quantitative measures of memory effect. According to the Gauss-Markov model, the local field obeys a multi-variate Gaussian distribution:

$$p(\Omega_n, \dots, \Omega_1) = p(\Omega_n | \Omega_{n-1}) \cdots p(\Omega_2 | \Omega_1) p(\Omega_1), \quad (\text{C.34})$$

where

$$\begin{aligned} p(\Omega_1) &= \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{\Omega_1^2}{2\sigma^2}\right), \\ p(\Omega_{i+1} | \Omega_i) &= \frac{1}{\sigma\sqrt{2\pi(1-\lambda_{i,i+1}^2)}} \exp\left\{-\frac{1}{2\sigma^2} \frac{(\Omega_{i+1} - \Omega_i \lambda_{i,i+1})^2}{1-\lambda_{i,i+1}^2}\right\}, \end{aligned} \quad (\text{C.35})$$

with the correlation being an exponential function of the time $\lambda_{i,i+1} = \exp(-\theta\Delta_i)$. Given the distribution of the local field, the time-domain data of an n D experiment is

$$S_{nD} = \int p(\omega_n, \dots, \omega_1) e^{i\omega_n t_n} \dots e^{i\omega_1 t_1} d\omega_n \dots d\omega_1. \quad (\text{C.36})$$

If $p(\omega_n, \dots, \omega_1)$ is a Gaussian distribution with a covariance matrix Σ and mean $\vec{\mu} = (\mu_1, \dots, \mu_n)$, we have

$$S_{nD}^G = \exp\left(-i\vec{t} \cdot \vec{\mu} - \frac{1}{2}\vec{t} \cdot \Sigma \cdot \vec{t}\right). \quad (\text{C.37})$$

The Gauss-Markov model requires $\vec{\mu} = \vec{0}$ and

$$\Sigma = \sigma^2 \begin{pmatrix} 1 & \lambda_{12} & \dots & \lambda_{1n} \\ \lambda_{12} & 1 & & \\ \vdots & & \ddots & \\ \lambda_{1n} & & & 1 \end{pmatrix}, \quad (\text{C.38})$$

where $\lambda_{ij} = \exp[-\theta(\Delta_i + \dots + \Delta_{j-1})]$. For 2D and 3D experiments, these are

$$\begin{aligned} S_{2D}^{\text{MG}} &= \exp\left[-\frac{\sigma^2}{2} (t_1^2 + t_2^2 + 2t_1 t_2 \lambda_{12})\right], \\ S_{3D}^{\text{MG}} &= \exp\left[-\frac{\sigma^2}{2} (t_1^2 + t_2^2 + t_3^2 + 2t_1 t_2 \lambda_{12} + 2t_1 t_3 \lambda_{13} + 2t_2 t_3 \lambda_{23})\right]. \end{aligned} \quad (\text{C.39})$$

These models are used to provide quantitative measure to non-Gaussianity and non-Markovianity.

To measure non-stationarity in the 2D data, the model is allowed to have different means: $\mu_1 \neq \mu_2$, the signal becomes

$$S'_{2D} = \exp\left[-i(\mu_1 t_1 + \mu_2 t_2) - \frac{\sigma^2}{2} (t_1^2 + t_2^2 + 2t_1 t_2 \lambda_{12})\right]. \quad (\text{C.40})$$

C.4 Time-domain data and pure-phase spectrum

The States' method [181, 96] is used to obtain pure-phase spectra. Define

$$\begin{aligned} S_c(t_1, t_2) &= \frac{1}{2} [S_{2D}(t_1, t_2) + S_{2D}(-t_1, t_2)], \\ S_s(t_1, t_2) &= \frac{1}{2i} [S_{2D}(t_1, t_2) - S_{2D}(-t_1, t_2)], \end{aligned} \quad (\text{C.41})$$

the pure absorption spectrum is given by

$$\tilde{S}_{2D}^{\text{States}} = \text{Re}(\mathcal{F}_1 \{ \text{Re}[\mathcal{F}_2(S_c)] \}) - \text{Im}(\mathcal{F}_1 \{ \text{Re}[\mathcal{F}_2(S_s)] \}) \quad (\text{C.42})$$

where $\tilde{S}'_{c/s} = \mathcal{F}_1 \{ \text{Re}[\mathcal{F}_2(S_{c/s})] \}$. For 3D experiments, define

$$\begin{aligned} S_{cc}(t_1, t_2, t_3) &= \frac{1}{4} [S_{3D}(t_1, t_2, t_3) + S_{3D}(t_1, -t_2, t_3) + S_{3D}(-t_1, t_2, t_3) + S_{3D}(-t_1, -t_2, t_3)], \\ S_{cs}(t_1, t_2, t_3) &= \frac{1}{4i} [S_{3D}(t_1, t_2, t_3) - S_{3D}(t_1, -t_2, t_3) + S_{3D}(-t_1, t_2, t_3) - S_{3D}(-t_1, -t_2, t_3)], \\ S_{sc}(t_1, t_2, t_3) &= \frac{1}{4i} [S_{3D}(t_1, t_2, t_3) + S_{3D}(t_1, -t_2, t_3) - S_{3D}(-t_1, t_2, t_3) - S_{3D}(-t_1, -t_2, t_3)], \\ S_{ss}(t_1, t_2, t_3) &= \frac{1}{4} [-S_{3D}(t_1, t_2, t_3) + S_{3D}(t_1, -t_2, t_3) + S_{3D}(-t_1, t_2, t_3) - S_{3D}(-t_1, -t_2, t_3)], \end{aligned} \quad (\text{C.43})$$

the pure-phase spectrum is

$$\begin{aligned} \tilde{S}_{3D}^{\text{States}} &= \text{Re} \left[\mathcal{F}_1 \left(\text{Re} \left(\mathcal{F}_2 \left\{ \text{Re}[\mathcal{F}_3(\tilde{S}_{cc})] \right\} \right) \right) - \text{Im} \left(\mathcal{F}_2 \left\{ \text{Re}[\mathcal{F}_3(\tilde{S}_{cs})] \right\} \right) \right] \\ &\quad - \text{Im} \left[\mathcal{F}_1 \left(\text{Re} \left(\mathcal{F}_2 \left\{ \text{Re}[\mathcal{F}_3(\tilde{S}_{sc})] \right\} \right) \right) - \text{Im} \left(\mathcal{F}_2 \left\{ \text{Re}[\mathcal{F}_3(\tilde{S}_{ss})] \right\} \right) \right]. \end{aligned} \quad (\text{C.44})$$

Fig. C.1 shows the coherence transfer pathways [20] for the 2D and 3D experiments. Initial scaling is applied to the time-domain data prior to Fourier transform to avoid ridges [148, 171].

C.5 Bayesian analysis

Bayesian analysis is used to extract information from the experiment data, and the essential steps of the procedure are described here. Each point x_i in the time-domain data takes the form

$$x_i = y_i + \epsilon_i, \quad (\text{C.45})$$

where y_i is the corresponding true value and ϵ_i is a zero mean white noise $\sim \mathcal{N}(0, \sigma_{\text{ref}}^2)$.

We use Bayesian posterior model checking [64, 132] to provide measures to environmental memory effect. The test quantities are, mean square error (MSE) for quantifying non-Gaussianity,

$$\text{MSE} = \int d\vec{x}^{\text{rep}} p(\vec{x}^{\text{rep}} | \vec{x}) \overline{(\vec{x}^{\text{rep}} - \vec{x})^2}, \quad (\text{C.46})$$

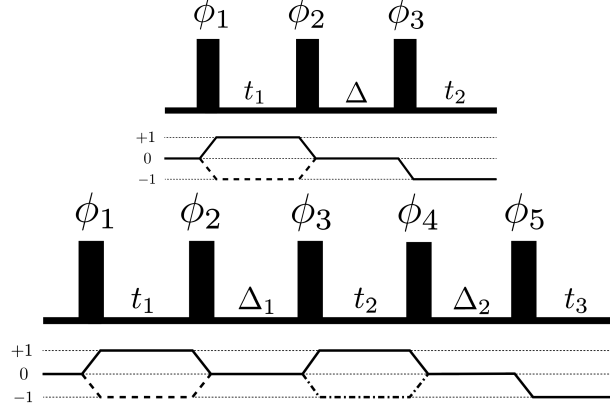


Figure C.1: Coherence transfer pathways for 2D and 3D experiments. The solid lines achieve backward evolutions in the corresponding periods while the dashed lines achieve forward evolutions. By combining forward and backward evolutions, one cosine- or sine-modulation in that period can be selected.

absolute mean difference (AMD) for quantifying non-stationarity,

$$\text{AMD} = \int d\mu_1 d\mu_2 p(\mu_1, \mu_2 | \vec{x}) |\mu_1 - \mu_2| \quad (\text{C.47})$$

and the residual correlation

$$R_\lambda^2 = \int d\lambda_{12} \lambda_{13} \lambda_{23} p(\lambda_{12}, \lambda_{13}, \lambda_{23} | \vec{x}) (\lambda_{13} - \lambda_{12} \lambda_{23})^2 \quad (\text{C.48})$$

and the correlation

$$\mathbb{E}[\lambda_{13} | \vec{x}] = \int d\lambda_{13} p(\lambda_{13} | \vec{x}) \lambda_{13} \quad (\text{C.49})$$

to quantify non-Markovianity. To obtain the posterior distributions, the 2D time-domain data (Eq. (C.41)) are fitted to $Ah(t_1 + t_2)S_{2D}^{\text{MG}}$ and $Ah(t_1 + t_2)S'_{2D}$. A is a normalization constant and $h(t) = \exp(-t/T_2)$ is the T_2 decay. The 3D time-domain data (Eqs. (C.43)) are fitted to $Ah(t_1 + t_2 + t_3)S_{3D}^{\text{MG}}$. The prior distributions for the parameters in the models are chosen as independent logit-normal distributions with $\sigma = 1$ over ranges that are large enough so that the priors are noninformative. Then, the parameters are transformed to $(-\infty, +\infty)$ through a logit function and the posterior distributions are simulated with a Metropolis-Hastings sampler.

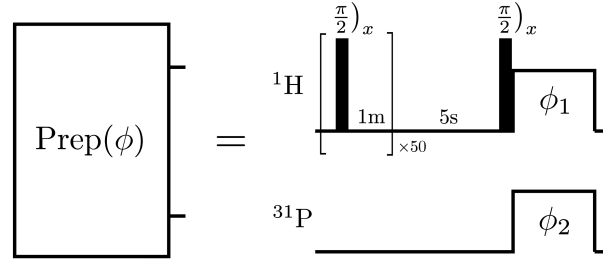


Figure C.2: Preparation pulse sequence. The phase is denoted in analogy to a $\pi/2$ pulse, indicating that the state after preparation is $(\sigma_x^{\text{CS}} \sin \phi - \sigma_y^{\text{CS}} \cos \phi) \otimes \mathbb{1}^{\otimes N}$. A saturation consisting of 50 $\pi/2$ pulses separated by 1 ms delays is applied to eliminate proton magnetization, ensuring that all experiments start with the same initial state. After a 5 s delay, cross-polarization is used to transfer polarization from proton spins to phosphorus spins. The duration of the spin-locking fields is 1400 μs .

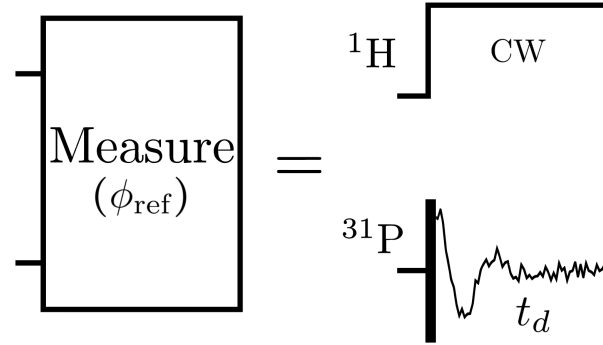


Figure C.3: Measurement pulse sequence.

C.5.1 Multi-dimensional experiment

For convenience, we refer to Figs. C.2 and C.3 for the preparation and measurement of the experiments. The procedure is the same as that used in Chapter 4.

The pulse sequences for 1D, 2D and 3D experiments are shown in Fig. C.4. The preparation consists of a saturation period on the environment spins followed by a cross polarization and the measurement consists of a continuous-wave decoupling on the environment and the continuous sampling on the CS (ref. the design of the STE experiments in chapter 4). During each t_k , the SME8 sequence is used to decouple homonuclear interaction between environment spins while retaining heteronuclear interactions between CS and environment spins. In order to remove chemical shift during each t_k , a π pulse is applied

on the CS during each t_k . A magic-echo sequence is inserted in the SME8 sequence on top of the π pulse to recouple the heteronuclear interaction.

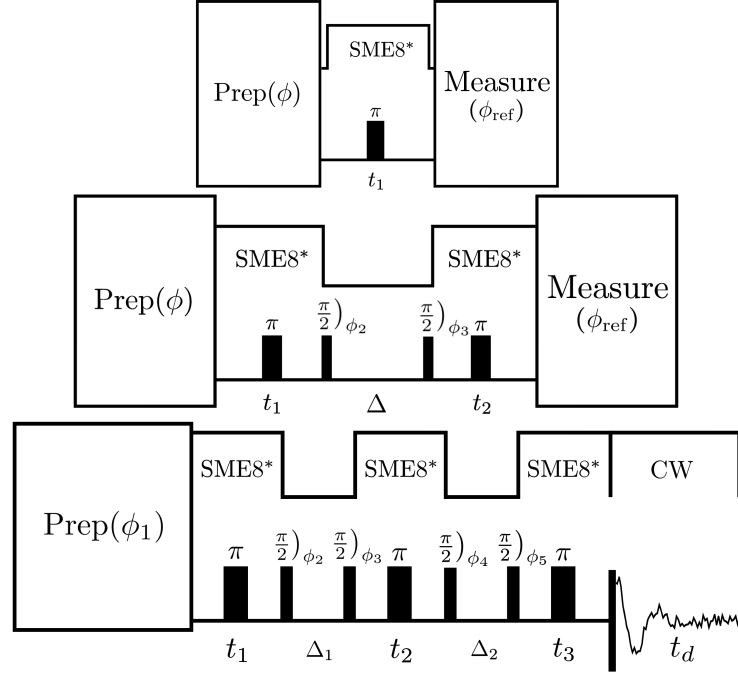


Figure C.4: Pulse sequences 1D (top), 2D (middle) and 3D (bottom) experiments.

Appendix D

Supplementary: Probing the structure of the environment through the central spin using multi-quantum experiments

This appendix presents the general theory of the standard MQ experiments and the pulse sequences for the experiments in the main text.

D.1 Dynamics without mixing

In the absence of environmental dynamics, the initial state, $\rho(0) = \sigma_x \otimes \mathbb{1}^N / 2^{N+1}$, evolves under

$$H_{\text{int}} = \frac{1}{2} \sigma_z^{\text{CS}} \otimes \sum_{i=1}^N \omega_i \sigma_z^i, \quad (\text{D.1})$$

into

$$\begin{aligned} \rho(t) &= \frac{1}{2^{N+1}} \left\{ \sigma_x \otimes \mathbb{1}^N \prod_{i=1}^N \cos \omega_i t + \sum_{i=1}^N \sigma_y \otimes \sigma_z^i \sin \omega_i t \prod_{j \neq i}^N \cos \omega_j t + \dots \right. \\ &\quad \left. + \sigma_{[N]} \otimes \sigma_z^1 \otimes \dots \otimes \sigma_z^N \prod_{i=1}^N \sin \omega_i t \right\} \\ &= \sum_{n=0}^N \frac{A_n}{2^{N+1}} \sigma_{[n]} \otimes \rho_n, \end{aligned} \quad (\text{D.2})$$

for operators $\rho_n \in \text{span}\{\mathbb{1}, \sigma_z\}^{\otimes N}$ with $\text{Tr}(\rho_n \rho_n^\dagger) = 2^N$ where n is the number of σ_z in ρ where

$$\sigma_{[n]} = \sigma_x \left[(-1)^{-\frac{n}{2}} + (-1)^{\frac{n}{2}} \right] + \sigma_y \left[(-1)^{-\frac{n-1}{2}} - (-1)^{\frac{n+1}{2}} \right]. \quad (\text{D.3})$$

Alternatively, A_n^2 can be calculated through

$$A_n^2 = \frac{1}{n!} \left. \frac{d^n r}{dx^n} \right|_{x=0}. \quad (\text{D.4})$$

where

$$r(x) = \prod_{n=1}^N (x \sin^2 \omega_n t + \cos^2 \omega_n t). \quad (\text{D.5})$$

Thus,

$$\begin{aligned} A_0^2 &= \prod_{n=1}^N \cos^2 \omega_n t, \\ A_1^2 &= \sum_{i=1}^N \sin^2 \omega_i t \prod_{j \neq i}^N \cos^2 \omega_j t, \\ &\dots, \\ A_N^2 &= \prod_{n=1}^N \sin^2 \omega_n t. \end{aligned} \quad (\text{D.6})$$

Define

$$\sigma_\pm^x = \frac{\sigma_z \mp i\sigma_y}{2}. \quad (\text{D.7})$$

We have

$$\exp(-i\sigma_x \phi/2) \sigma_\pm^x \exp(i\sigma_x \phi/2) = \sigma_\pm^x \exp(-ip\phi). \quad (\text{D.8})$$

Under the basis $\{\mathbb{1}, \sigma_+^x, \sigma_-^x\}^{\otimes N}$ for the environment,

$$\rho(t) = \frac{1}{2^{N+1}} \left\{ \sigma_x \otimes \mathbb{1}^N \prod_{i=1}^N \cos \omega_i t + \dots + \sigma_{[N]} \otimes (\sigma_+^x + \sigma_-^x)^{\otimes N} \prod_{i=1}^N \sin \omega_i t \right\}. \quad (\text{D.9})$$

In total, there are C_n^p terms of coherence order $n - 2p$ in ρ_n for $p = -n, \dots, n$. We encode them with collective x rotations. The signal of a encoded echo is

$$S(\theta) = \text{Tr} [U^\dagger R_x(\theta) U \rho(0) U^\dagger R_x^\dagger(\theta) U \rho^\dagger(0)] = \text{Tr} [R_x(\theta) \rho(t) R_x^\dagger(\theta) \rho^\dagger(0)], \quad (\text{D.10})$$

where $R_x(\theta)$ is a collective x rotation on the environment. Choose $M \geq N$ and define $A_n^2 = 0$ for $n > N$, the signal is

$$\begin{aligned} S(\theta) &= \sum_{n=0}^M \sum_{p=0}^n A_n^2 \frac{C_n^p}{2^n} \exp[-i(n-2p)\theta] \\ &= \sum_{n=-M}^M A^{(n)^2} \exp(-in\theta). \end{aligned} \quad (\text{D.11})$$

Here, $A^{(n)^2}$ is the total amplitude of terms of coherence order n , which satisfy $\sum_{n=-M}^M A^{(n)^2} = 1$. Let $\theta = 2\pi j/(2M+1)$, $j = -M, -M+1, \dots, M$. A discrete Fourier transform yields

$$\sum_{j=-M}^M \exp(-ik\theta) S(\theta) = A^{(k)^2} = \sum_{n=0}^M \sum_{p=0}^n A_n^2 \frac{C_n^p}{2^n} \delta_{k,n-2p}, \quad (\text{D.12})$$

namely,

$$\begin{aligned} A^{(2k)^2} &= \sum_{j=0}^{\lfloor \frac{M-2k}{2} \rfloor} A_{2k+2j}^2 \frac{C_{2k+2j}^{2j}}{2^{2k+2j}} = A_{2k}^2 \frac{C_{2k}^0}{2^{2k}} + A_{2k+2}^2 \frac{C_{2k+2}^2}{2^{2k+2}} + \dots, \\ A^{(2k+1)^2} &= \sum_{j=0}^{\lfloor \frac{M-2k-1}{2} \rfloor} A_{2k+2j+1}^2 \frac{C_{2k+2j+1}^{2j}}{2^{2k+2j+1}} = A_{2k+1}^2 \frac{C_{2k+1}^0}{2^{2k+1}} + A_{2k+3}^2 \frac{C_{2k+3}^2}{2^{2k+3}} + \dots, \end{aligned} \quad (\text{D.13})$$

for $k = 0, \dots, \lfloor \frac{M}{2} \rfloor$.

Since $A^{(k)^2} = A^{(-k)^2}$, we only need to consider $k \geq 0$. Eq. (D.12) defines a linear map from $\{A_n^2\}$ to $\{A^{(k)^2}\}$

$$\left(A^{(0)^2}, \dots, A^{(M)^2} \right)^T = \mathbf{A} \left(A_0^2, \dots, A_M^2 \right)^T, \quad (\text{D.14})$$

with

$$\mathbf{A}_{ij} = \begin{cases} 0 & j < i \text{ or } (j-i) \bmod 2 = 1 \\ \frac{C_{j-1}^{(j-i)/2}}{2^{j-1}} & d \end{cases} = \begin{pmatrix} \ddots & & \\ & \frac{C_{M-2}^0}{2^{M-2}} & 0 & \frac{C_M^1}{2^M} \\ & & \frac{C_{M-1}^0}{2^{M-1}} & 0 \\ & & & \frac{C_M^0}{2^M} \end{pmatrix}. \quad (\text{D.15})$$

This allows us to calculate $\{A_m^2\}$ or $\{A^{(k)^2}\}$ when the other is known.

For example, if

$$\begin{aligned}
\rho(t) &= \frac{1}{2^4} \left[A_0 \sigma_x \otimes \mathbb{1}^{\otimes 3} + A_1 \sigma_y \otimes (\sigma_z \otimes \mathbb{1} \otimes \mathbb{1} + \mathbb{1} \otimes \mathbb{1} \otimes \sigma_z) / \sqrt{2} \right. \\
&\quad \left. - A_2 \sigma_x \otimes \sigma_z \otimes \mathbb{1} \otimes \sigma_z - A_3 \sigma_y \otimes \sigma_z \otimes \sigma_z \otimes \sigma_z \right] \\
&= \frac{1}{2^4} \left[A_0 \sigma_x \otimes \mathbb{1}^{\otimes 3} + A_1 \sigma_y \otimes (\sigma_+^x \otimes \mathbb{1} \otimes \mathbb{1} + \sigma_-^x \otimes \mathbb{1} \otimes \mathbb{1} + \mathbb{1} \otimes \mathbb{1} \otimes \sigma_+^x + \mathbb{1} \otimes \mathbb{1} \otimes \sigma_-^x) / \sqrt{2} \right. \\
&\quad - A_2 \sigma_x \otimes (\sigma_+^x \otimes \mathbb{1} \otimes \sigma_+^x + \sigma_+^x \otimes \mathbb{1} \otimes \sigma_-^x + \sigma_-^x \otimes \mathbb{1} \otimes \sigma_+^x + \sigma_-^x \otimes \mathbb{1} \otimes \sigma_-^x) \\
&\quad \left. - A_3 \sigma_y \otimes (\sigma_+^x \otimes \sigma_+^x \otimes \sigma_+^x + \sigma_+^x \otimes \sigma_+^x \otimes \sigma_-^x + \sigma_+^x \otimes \sigma_-^x \otimes \sigma_+^x + \sigma_-^x \otimes \sigma_+^x \otimes \sigma_+^x \right. \\
&\quad \left. + \sigma_-^x \otimes \sigma_-^x \otimes \sigma_+^x + \sigma_-^x \otimes \sigma_+^x \otimes \sigma_-^x + \sigma_-^x \otimes \sigma_-^x \otimes \sigma_+^x + \sigma_-^x \otimes \sigma_-^x \otimes \sigma_-^x) \right]. \tag{D.16}
\end{aligned}$$

According to Eq. (D.12), the coherence amplitudes are

$$\begin{aligned}
A^{(0)^2} &= A_0^2 C_0^0 + A_2^2 \frac{C_2^1}{2^2} = A_0^2 + \frac{A_2^2}{2}, \\
A^{(\pm 1)^2} &= A_1^2 \frac{C_1^0}{2^1} + A_3^2 \frac{C_3^1}{2^3} = \frac{A_1^2}{2} + \frac{3A_3^2}{8}, \\
A^{(\pm 2)^2} &= A_2^2 \frac{C_2^2}{2^2} = \frac{A_2^2}{4}, \\
A^{(\pm 3)^2} &= A_3^2 \frac{C_3^0}{2^3} = \frac{A_3^2}{8}.
\end{aligned} \tag{D.17}$$

The transformation matrix and its inverse are

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{3}{8} \\ 0 & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 0 & \frac{1}{8} \end{pmatrix} \quad \text{and} \quad \mathbf{A}^{-1} = \begin{pmatrix} 1 & 0 & -2 & 0 \\ 0 & 2 & 0 & -6 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 8 \end{pmatrix}. \tag{D.18}$$

Thus,

$$\begin{aligned}
\left(A^{(0)^2}, A^{(1)^2}, A^{(2)^2}, A^{(3)^2} \right)^T &= \mathbf{A} \left(A_0^2, A_1^2, A_2^2, A_3^2 \right)^T \\
&= \left(A_0^2 + A_2^2/2, A_1^2/2 + 3A_3^2/8, A_2^2/4, A_3^2/8 \right)^T, \\
\left(A_0^2, A_1^2, A_2^2, A_3^2 \right)^T &= \mathbf{A}^{-1} \left(A^{(0)^2}, A^{(1)^2}, A^{(2)^2}, A^{(3)^2} \right)^T \\
&= \left(A^{(0)^2} - 2A^{(2)^2}, 2A^{(1)^2} - 6A^{(3)^2}, 4A^{(2)^2}, 8A^{(3)^2} \right)^T.
\end{aligned} \tag{D.19}$$

D.2 Pulse Sequences

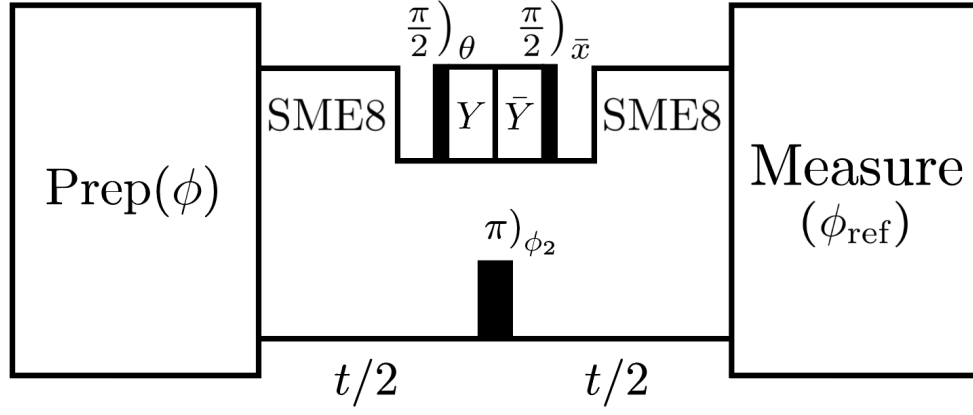


Figure D.1: Pulse sequence for multi-quantum experiments. The phase of the magic echo sequence in the middle is changed so that it applies a θ encoding to the environment spins.

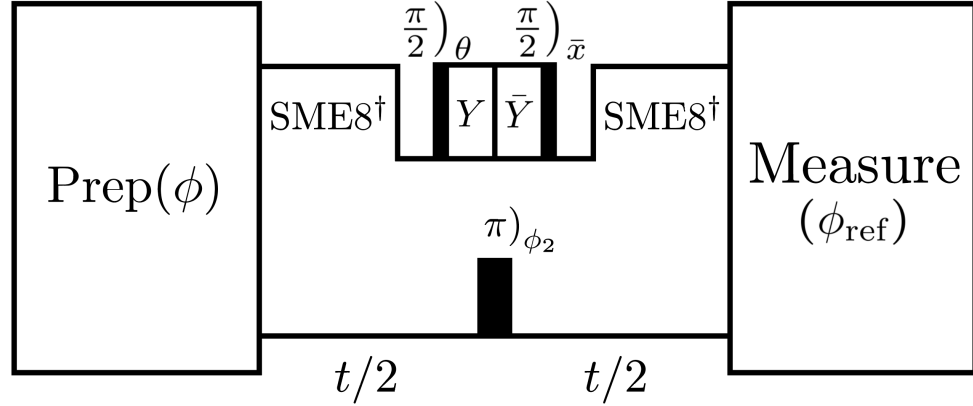


Figure D.2: Pulse sequence for Loschmidt echo experiments. $\text{SME8}^\dagger$ is a variations of SME8 where delays in the magic echo sequences are adjusted to introduce mixing.

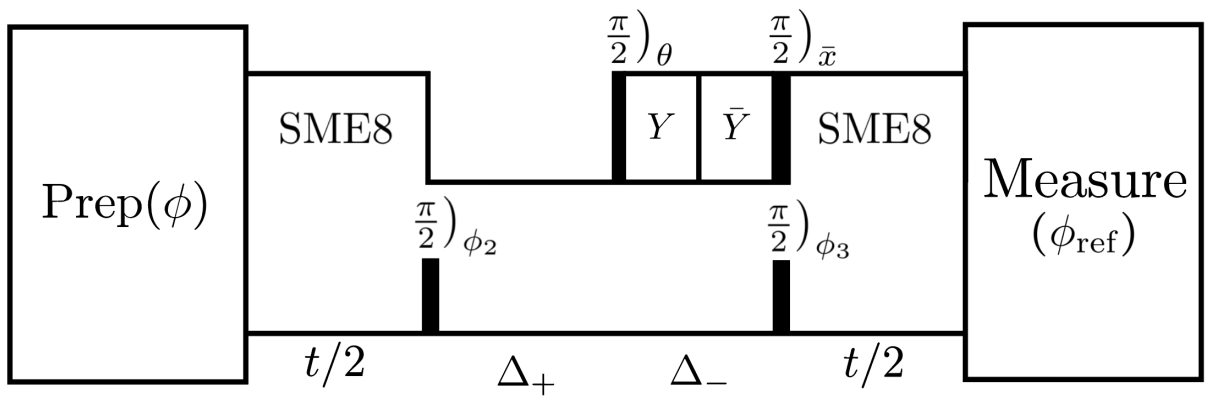


Figure D.3: Pulse sequence for multi-quantum experiments with mixing. $\Delta_+ = 2\Delta_-$ so the environmental dynamics is refocused at the end of Δ_- .

Appendix E

Randomized Benchmarking of the Central Spin System

This appendix presents some general results on randomized benchmarking on the CS system.

E.1 Randomized Benchmarking

Randomized benchmarking (RB) [102, 124, 125] are scalable and robust methods that characterize the noise of a set of gates. The standard model predicts an exponential decay under Markovian noise. Therefore, deviation from an exponential decay can indicate memory effect.

First, we briefly review the basic model of standard RB. For a unital matrix $\hat{E}$ acting on $\mathfrak{u}(n)$, its average gate fidelity is [57]

$$\tilde{F}(E) = \int \langle\langle \phi | \hat{E} | \phi \rangle\rangle d\phi = \int_{U \in \text{SU}(n)} \langle\langle \phi | \hat{U} \circ \hat{E} \circ \hat{U}^\dagger | \phi \rangle\rangle d\tau_U = \langle\langle \phi | \hat{E}_{\text{twirl}} | \phi \rangle\rangle, \quad (\text{E.1})$$

with the unitary twirl defined as

$$\hat{E}_{\text{twirl}} = \int_{U \in \text{SU}(n)} \hat{U} \circ \hat{E} \circ \hat{U}^\dagger d\tau_U. \quad (\text{E.2})$$

$\hat{U}$ is the conjugate representation of U on $\mathfrak{u}(n)$. $d\tau_U$ is an invariant measure such that $d\tau_U = d\tau_{VU}$ for $\forall V \in \text{U}(n)$. We have $[\hat{E}_{\text{twirl}}, \hat{U}] = 0 \ \forall U \in \text{U}(n)$. According to Schur's

Lemma, $\hat{E}_{\text{twirl}}$ is a multiple of the identity in each irreducible subspace: $\{\mathbb{1}\}$ and $\text{su}(n)$ [77]. Namely, $\hat{E}_{\text{twirl}} = \text{Diag}(1, \lambda, \dots, \lambda)$ where $(n^2 - 1)\lambda + 1 = \text{Tr}(\hat{E})$. Therefore, for a density matrix ρ , we have

$$\hat{E}_{\text{twirl}}^{\circ k}(\rho) = \frac{\mathbb{1}_n}{n} + \lambda^k \left(\rho - \frac{\mathbb{1}_n}{n} \right). \quad (\text{E.3})$$

The survival probability is an exponential function of k

$$\bar{S}_k = \text{Tr}[\rho \hat{E}_{\text{twirl}}^{\circ k}(\rho)^\dagger] = \frac{1}{n} + \lambda^k \left(1 - \frac{1}{n} \right). \quad (\text{E.4})$$

The average gate fidelity is

$$\tilde{F}(\hat{E}) = S_1 = \lambda + \frac{1 - \lambda}{n}. \quad (\text{E.5})$$

For $\rho_0 = (\mathbb{1} + \epsilon \sigma_x)/2$, the signal after $E_{\text{twirl}}^{\circ k}$ is

$$S_{\text{RB}}(k) = \text{Tr}[\sigma_x \hat{E}_{\text{twirl}}^{\circ k}(\rho_0)^\dagger] = \epsilon \lambda^k. \quad (\text{E.6})$$

The twirl of E can be realized using the Clifford group $\mathcal{C}$:

$$\hat{E}_{\text{twirl}} = \int_{U \in \text{SU}(n)} \hat{U} \circ \hat{E} \circ \hat{U}^\dagger d\tau_U = \frac{1}{|\mathcal{C}|} \sum_{i=1}^{|\mathcal{C}|} \hat{C}_i \circ \hat{E} \circ \hat{C}_i^\dagger. \quad (\text{E.7})$$

Hence, by the rearrangement theorem

$$\begin{aligned} \hat{E}_{\text{twirl}}^{\circ k} &= \frac{1}{|\mathcal{C}|^k} \left(\sum_{i=1}^{|\mathcal{C}|} \hat{C}_i \circ \hat{E} \circ \hat{C}_i^\dagger \right)^k \\ &= \sum_{i_1, \dots, i_k} \hat{C}_{i_{k+1}} \circ \hat{E} \circ \hat{C}_{i_k} \circ \hat{E} \circ \dots \circ \hat{C}_{i_1} \\ &= \mathbb{E}_{\vec{C} \in \mathcal{C}^{\times k}} (\hat{C}_{i_{k+1}} \circ \hat{E} \circ \hat{C}_{i_k} \circ \hat{E} \circ \dots \circ \hat{C}_{i_1}), \end{aligned} \quad (\text{E.8})$$

where $C_{i_{k+1}} = (C_{i_k} \dots C_{i_1})^\dagger$ and $\vec{C} = (C_{i_k}, \dots, C_{i_1})$. In randomized benchmarking (RB), one can estimate

$$\mathbb{E}_{\vec{C} \in \mathcal{C}^{\times k}} (\hat{C}_{i_{k+1}} \circ \hat{E} \circ \hat{C}_{i_k} \circ \hat{E} \circ \dots \circ \hat{C}_{i_1}) \quad (\text{E.9})$$

by picking uniformly at random from the Clifford group for m times for each C_{i_j} [124, 125]. RB is free of state preparation and measurement error. It is a more efficient and reliable way for estimating the average gate fidelity compared to state tomography [214]. It is also a useful tool for diagnosing noise in experiments [147, 200].

E.1.1 Benchmarking the Central Spin System

We want to use an interleaved RB to characterize the dynamics of the CS system. We assume that the system starts with $\rho(0) = \sigma_x \otimes \mathbb{1}^{\otimes n}/2^{n+1}$. Consider an experiment that measures

$$S_{\text{RB}}(t, k) = \mathbb{E}_{\tilde{C} \in \mathcal{C}^{\times k}} \left(\langle\langle \sigma_+ \otimes \mathbb{1}^n | \hat{C}_{i_{k+1}} \otimes \hat{\mathbb{1}}^{\otimes n} \circ \hat{E}(t) \circ \dots \circ \hat{E}(t) \circ \hat{C}_{i_1} \otimes \hat{\mathbb{1}}^{\otimes n} | \rho(0) \rangle\rangle \right). \quad (\text{E.10})$$

Here, $\mathcal{C}$ is the 1-qubit Clifford group and $E(t) = \exp(-iHt)$ where $H = H_{cs/e} + H_e$ and

$$H_{cs/e} = \sigma_z \otimes A/2, \quad H_e = \mathbb{1} \otimes B. \quad (\text{E.11})$$

In the interaction frame of H_e , we have

$$S_{\text{RB}}(t, k) = \mathbb{E}_{\tilde{C} \in \mathcal{C}^{\times k}} \left(\langle\langle \sigma_+ \otimes \mathbb{1}^n | \hat{C}_{i_{k+1}} \otimes \hat{\mathbb{1}}^{\otimes n} \circ \hat{E}'_k \circ \dots \circ \hat{E}'_1 \circ \hat{C}_{i_1} \otimes \hat{\mathbb{1}}^{\otimes n} | \rho(0) \rangle\rangle \right), \quad (\text{E.12})$$

where

$$E_k = \mathbb{1} \otimes M^{1-k} \mathcal{T} \exp \left(-i \int_0^t \sigma_z \otimes A'(t')/2 dt' \right) \mathbb{1} \otimes M^{k-1}, \quad (\text{E.13})$$

with $M = \exp(-iBt)$.

For an operator $\hat{\Lambda}$ acting on $\mathfrak{u}(2^{n+1})$, define

$$\hat{\Lambda}_{1\text{-twirl}} = \int_{U \in \text{SU}(2)} \hat{U} \otimes \hat{\mathbb{1}}^{\otimes n} \circ \hat{\Lambda} \circ \hat{U}^\dagger \otimes \hat{\mathbb{1}}^{\otimes n} d\tau_U. \quad (\text{E.14})$$

$\hat{\Lambda}_{1\text{-twirl}}$ is in general a block-diagonal matrix

$$\hat{\Lambda}_{1\text{-twirl}} = \begin{pmatrix} \Lambda_1 & & & \\ & \Lambda_2 & & \\ & & \Lambda_2 & \\ & & & \Lambda_2 \end{pmatrix}, \quad (\text{E.15})$$

with Λ_1 and Λ_2 being $4^n \times 4^n$ blocks and $\Lambda_1 + 3\Lambda_2 = \text{Tr}_1(\hat{\Lambda})$. In addition, if we denote the diagonal $4^n \times 4^n$ blocks of $\hat{\Lambda}$ as Λ_{ii} for $i = 1, 2, 3$ and 4, then

$$\begin{aligned} \Lambda_1 &= \Lambda_{11}, \\ \Lambda_2 &= (\Lambda_{11} + \Lambda_{22} + \Lambda_{33})/3. \end{aligned} \quad (\text{E.16})$$

In order to calculate the corresponding Λ_1 and Λ_2 of $E(t)$ defined in Eq. (E.15), we only need to know its action in the invariant subspaces.

Recall that

$$\begin{aligned} U(t) &= \mathcal{T}_\leftarrow \exp \left[-i \int_0^t dt_1 \sigma_z \otimes A'(t_1)/2 \right] \\ &= \mathbb{1} \otimes C_\leftarrow - i \sigma_z \otimes S_\leftarrow, \end{aligned} \quad (\text{E.17})$$

with

$$\begin{aligned} C_\leftarrow &= \mathcal{T}_\leftarrow \cos \left[\int_0^t dt_1 A'(t_1)/2 \right] = \frac{1}{2} (U_\leftarrow^+ + U_\leftarrow^-), \\ S_\leftarrow &= \mathcal{T}_\leftarrow \sin \left[\int_0^t dt_1 A'(t_1)/2 \right] = \frac{1}{2i} (U_\leftarrow^+ - U_\leftarrow^-). \end{aligned} \quad (\text{E.18})$$

We first calculate Λ_1 . Since

$$\begin{aligned} &\text{Tr} (\mathbb{1} \otimes \sigma (\mathbb{1} \otimes C_\leftarrow - i \sigma_\alpha \otimes S_\leftarrow) \mathbb{1} \otimes \rho (\mathbb{1} \otimes C_\rightarrow + i \sigma_\alpha \otimes S_\rightarrow)) \\ &= 2 \text{Tr} (\sigma (C_\leftarrow \rho C_\rightarrow + S_\leftarrow \rho S_\rightarrow)) \\ &= \text{Tr} (\sigma (U_\leftarrow^+ \rho U_\rightarrow^- + U_\leftarrow^- \rho U_\rightarrow^+)). \end{aligned} \quad (\text{E.19})$$

Since

$$\begin{aligned} U_\leftarrow^+ \rho U_\rightarrow^- + U_\leftarrow^- \rho U_\rightarrow^+ &= \left(\mathcal{T} \exp \left[i \int_0^t \text{ad}_{A'(t')} dt'/2 \right] + \mathcal{T} \exp \left[-i \int_0^t \text{ad}_{A'(t')} dt'/2 \right] \right) (\rho) \\ &= 2 \mathcal{T} \cos \left[\int_0^t \text{ad}_{A'(t')} dt'/2 \right] (\rho), \end{aligned} \quad (\text{E.20})$$

we have

$$\Lambda_1 = \mathcal{T} \cos \left[\int_0^t \text{ad}_{A'(t')} dt'/2 \right]. \quad (\text{E.21})$$

Using the fact that

$$\begin{aligned} \sigma_\alpha \sigma_\alpha &= \mathbb{1}, \\ \sigma_\alpha \sigma_\beta &= i \epsilon_{\alpha\beta\gamma} \sigma_\gamma, \end{aligned} \quad (\text{E.22})$$

for $\alpha, \beta, \gamma = x, y, z$ and $\alpha \neq \beta$, with similar calculation, we have

$$\begin{aligned} \text{Tr} (\sigma_\alpha \otimes \sigma U(t) \sigma_\alpha \otimes \rho U^\dagger(t)) &= 2 \text{Tr} \left(\sigma \mathcal{T} \cos \left[\int_0^t \text{ad}_{A'(t')} dt'/2 \right] (\rho) \right), \\ \text{Tr} (\sigma_\alpha \otimes \sigma U(t) \sigma_\beta \otimes \rho U^\dagger(t)) &= 2 \text{Tr} \left(\sigma \mathcal{T} \cos \left[\int_0^t \text{ad}_{A'(t')} dt'/2 \right] (\rho) \right). \end{aligned} \quad (\text{E.23})$$

Thus,

$$\Lambda_2 = \frac{1}{3} \left(\mathcal{T} \cos \left[\int_0^t \text{ad}_{A'(t')} dt'/2 \right] + 2 \mathcal{T} \cos \left[\int_0^t \text{ad}_{A'(t')} dt'/2 \right] \right). \quad (\text{E.24})$$

Let

$$\exp(i\bar{A}_k t/2) \equiv M^{1-k} \mathcal{T} \exp \left[-i \int_0^t A'(t') dt' \right] M^{k-1}, \quad (\text{E.25})$$

then

$$\hat{E}_{k,1\text{-twirl}} = \begin{pmatrix} \Lambda'_1(k) & & & \\ & \Lambda'_2(k) & & \\ & & \Lambda'_2(k) & \\ & & & \Lambda'_2(k) \end{pmatrix}, \quad (\text{E.26})$$

with

$$\begin{aligned} \Lambda'_1(k) &= \cos(\text{ad}_{\bar{A}_k} t/2) \\ \Lambda'_2(k) &= \frac{1}{3} [\cos(\text{ad}_{\bar{A}_k} t/2) + 2 \cos(\text{ac}_{\bar{A}_k} t/2)]. \end{aligned} \quad (\text{E.27})$$

Thus, the RB signal is

$$S_{\text{RB}}(t, k) = \frac{1}{2^n} \langle\langle \mathbb{1}^{\otimes n} | \bigcirc_{j=1}^k \hat{\Lambda}'_2(j) | \mathbb{1}^{\otimes n} \rangle\rangle. \quad (\text{E.28})$$

Below we consider three different scenarios. When $B = 0$, denote the RB signal by $S_{\text{RB}}^*(t, k)$, we have

$$\begin{aligned} S_{\text{RB}}^*(t, k) &= \frac{1}{2^{n+1}} \langle\langle \sigma_+ \otimes \mathbb{1}^{\otimes n} | \hat{E}_{1\text{-twirl}}^{\circ k}(t) | \sigma_- \otimes \mathbb{1}^{\otimes n} \rangle\rangle \\ &= \frac{1}{2^n} \langle\langle \mathbb{1}^{\otimes n} | \Lambda_2^{\circ k} | \mathbb{1}^{\otimes n} \rangle\rangle \\ &= \frac{1}{2^n} \left\langle\left\langle \mathbb{1}^{\otimes n} \left| \left[\frac{\cos(\text{ad}_A t/2) + 2 \cos(\text{ac}_A t/2)}{3} \right]^{\circ k} \right| \mathbb{1}^{\otimes n} \right\rangle\right\rangle. \end{aligned} \quad (\text{E.29})$$

Let the eigenbasis of A be $\{|1\rangle, \dots, |2^n\rangle\}$, $A|j\rangle = \Omega_j|j\rangle$ we have $\text{ac}_A(|j\rangle\langle j|) = 2\Omega_j|j\rangle\langle j|$ and $\text{ad}_A(|j\rangle\langle j|) = 0$. Express $|\mathbb{1}\rangle$ in terms of $|j\rangle\langle j|$, we get

$$S_{\text{RB}}^*(t, k) = \frac{1}{2^n} \sum_{j=1}^{2^n} \left[\frac{1 + 2 \cos(\Omega_j t)}{3} \right]^k. \quad (\text{E.30})$$

When $\|t^2[B, [B, A]]\| \ll 1$, we can make the following truncation

$$\hat{\Lambda}'_2(j+1) \approx \hat{\Lambda}'_2(j) + t [\text{ad}_{iB}, \hat{\Lambda}'_2(j)] + \frac{t^2}{2} [\text{ad}_{iB}, [\text{ad}_{iB}, \hat{\Lambda}'_2(j)]]. \quad (\text{E.31})$$

We can show that

$$\begin{aligned}\langle\langle \mathbb{1}^{\otimes n} | [\text{ad}_{iB}, \hat{\Lambda}_2] \circ \hat{\Lambda}_2 | \mathbb{1}^{\otimes n} \rangle\rangle &= 0, \\ \langle\langle \mathbb{1}^{\otimes n} | [\text{ad}_{iB}, [\text{ad}_{iB}, \hat{\Lambda}_2]] \circ \hat{\Lambda}_2 | \mathbb{1}^{\otimes n} \rangle\rangle &< 0.\end{aligned}\tag{E.32}$$

The first identity comes from the fact that $\text{ad}_{iB}(\mathbb{1}^{\otimes n}) = 0$ and

$$\langle\langle \mathbb{1}^{\otimes n} | \hat{\Lambda}_2 \circ \text{ad}_{iB} \circ \hat{\Lambda}_2 | \mathbb{1}^{\otimes n} \rangle\rangle = \langle\langle \hat{\Lambda}_2(\mathbb{1}^{\otimes n}) | \text{ad}_{iB} | \hat{\Lambda}_2(\mathbb{1}^{\otimes n}) \rangle\rangle = 0.\tag{E.33}$$

To obtain the second inequality, we have

$$\langle\langle \mathbb{1}^{\otimes n} | [\text{ad}_{iB}, [\text{ad}_{iB}, \hat{\Lambda}_2]] \circ \hat{\Lambda}_2 | \mathbb{1}^{\otimes n} \rangle\rangle = \langle\langle \hat{\Lambda}_2(\mathbb{1}^{\otimes n}) | \text{ad}_{iB}^2 | \hat{\Lambda}_2(\mathbb{1}^{\otimes n}) \rangle\rangle < 0.\tag{E.34}$$

Since ad_{iB}^2 is negative semi-definite, the above quantity must be no larger than 0. It is equal to 0 if and only if $[A, B] = 0$. Therefore,

$$\begin{aligned}S_{\text{RB}}(t, 2) &= \frac{1}{2^n} \langle\langle \mathbb{1}^{\otimes n} | \hat{\Lambda}_2'(2) \circ \hat{\Lambda}_2(1) | \mathbb{1}^{\otimes n} \rangle\rangle \\ &\approx S_{\text{RB}}^*(t, 2) + \frac{t^2}{2^{n+1}} \langle\langle \mathbb{1}^{\otimes n} | [\text{ad}_{iB}, [\text{ad}_{iB}, \hat{\Lambda}_2]] | \mathbb{1}^{\otimes n} \rangle\rangle \leq S_{\text{RB}}^*(t, 2).\end{aligned}\tag{E.35}$$

The last equality holds if and only if $[A, B] = 0$.

When mixing is very strong $\|B\| \gg \|A\|$, the effective Hamiltonian is given by

$$H_{SM} = \sigma_z \otimes \bar{A}/2,\tag{E.36}$$

where $\bar{A}$ is given by the following theorem.

Theorem E.1.1. *Let A, B be $n \times n$ Hermitian matrices and T be a unitary matrix that diagonalizes B :*

$$TBT^\dagger = \begin{pmatrix} \lambda_1 \mathbb{1}_{\mu_1} & & \\ & \ddots & \\ & & \lambda_m \mathbb{1}_{\mu_m} \end{pmatrix},\tag{E.37}$$

where λ_i are distinct eigenvalues of B and μ_i are the corresponding degeneracies. Let $A' = TAT^\dagger = \{A'_{ij}\}_{i,j=1}^m$, where A'_{ij} are $\mu_i \times \mu_j$ blocks. Denote $\bar{A} = T^\dagger (\oplus_{i=1}^m A_{ii}) T$ and $U_0(t, x) = \exp[-ixBt]$. Then,

$$\lim_{x \rightarrow \infty} \mathcal{T} \exp \left[-i \int_0^t U_0^\dagger(t', x) A U_0(t', x) dt' \right] = \exp(-i\bar{A}t).\tag{E.38}$$

Denote the eigenvalue vectors of A and $\bar{A}$ by Ω and $\bar{\Omega}$, respectively.

When mixing is strong $|B| \gg 1/t$, $\bar{A}_j \equiv \bar{A}$. $\bar{A}$ can be calculated using Theorem E.1.1. The survival probability is

$$S_{\text{RB}}(t, k) = \frac{1}{2^n} \sum_{j=1}^{2^n} \bar{p}_j^k, \quad (\text{E.39})$$

with

$$\bar{p}_j = \frac{2 \cos(\bar{\Omega}_j t) + 1}{3}. \quad (\text{E.40})$$

$\bar{\Omega}_j$ are eigenvalues of $\bar{A}$. According to Schur majorization theorem, Ω_j majorizes $\bar{\Omega}_j$

$$\sum_{i=1}^k \Omega_j^\downarrow \geq \sum_{i=1}^k \bar{\Omega}_j^\downarrow, \quad (\text{E.41})$$

for $k = 1, \dots, n$. The corresponding local fields are more closely centered at 0, resulting in a more exponential and slower decay. Additionally, there exists a doubly stochastic matrix S , such that $\bar{\Omega} = S\Omega$. The matrix S determines the shape of the decay of the RB signal.

1. In the absence of the environmental mixing, the decay of the RB signal is non-exponential and determined by the distribution of the local field (Eq. (E.30)).
2. The shape of the decay is robust to small environmental mixing due to the lack of first-order dependence (Eq. (E.35)).
3. Small mixing leads to faster initial decay (Eq. (E.35)).
4. Strong mixing leads to a slower and more exponential decay (Eq. (E.39)).

The increase of decay rate due to small mixing is due to enhanced energy transfer between the CS and the environment when the strength of mixing approaches the frequency of the Clifford gates are being applied (cf. cross polarization and dynamical decoupling). We expect the slowdown to saturate at $\|jt^2[B, [B, A]]\| \sim 1$. Strong mixing removes CS/environment correlation and leads to a faster and more exponential decay. Any residual deviation from an exponential decay is due to residual correlation and can indicate constant of motion to the environmental mixing.

Appendix F

Discrete Methods for Pulse Engineering

Here, we present the general framework for designing sequences for effective Hamiltonians engineering using gates from a finite set.

F.1 Pulse Design Using Tree Search

The chain rule for $\mathcal{C}$ -integrals (Appendix A) allows constructing a control sequence from a finite set of components (e.g., square $\pi/2$ rotations or Clifford gates). Because the $\mathcal{C}$ -integrals of a composite sequence follow from those of its parts, we only need to calculate the $\mathcal{C}$ -integrals for each component in the finite set. Any sequence made from these components can then be derived using the chain rule. Then, the $\mathcal{C}$ -integrals fully determine the robustness and Hamiltonian-engineering properties of the sequence.

When designing sequences, we must account for constraints like limited total power and finite bandwidth. A convenient way to accommodate these constraints is to structure the solution space as a tree, where the tree policy (how new children are generated given the current node) can then be easily determined by these constraints. Then the design of a sequence is viewed as a tree search problem [21].

We propose to use Monte Carlo tree search which has achieved super-human performance in games [175, 195] and applied to quantum circuit design [202]. Here, we provide the pseudocode (Algorithm 9) for the upper confidence bound (UCT) search [103, 104, 31].

Algorithm 9 The UCT algorithm.

```
1: function UCTSEARCH( $s_0$ )
2:   create root node  $v_0$  with state  $s_0$ 
3:   while  $v_0$  is non-terminal do
4:     while within computational budget do
5:        $v_l \leftarrow \text{TREEPOLICY}(v_0)$ 
6:        $\text{DEFAULTPOLICY}(s(v_l))$ 
7:        $\text{BACKUP}(v_l, \Delta)$ 
8:     end while
9:      $v_0 \leftarrow \text{BESTCHILD}(v_0, 0)$ 
10:  end while
11:  return the state of  $v_0$ 
12: end function
13: function TREEPOLICY( $v$ )
14:   while  $v$  is nonterminal do
15:     if  $v$  not fully expanded then
16:       return  $\text{EXPAND}(v)$ 
17:     else
18:        $v \leftarrow \text{BESTCHILD}(v, C_p)$ 
19:     end if
20:   end while
21:   return  $v$ 
22: end function
23: function EXPAND( $v$ )
24:   choose  $v'$  from untried children of  $v$ 
25:   return  $v'$ 
26: end function
27: function BESTCHILD( $v, c$ )
28:    $\arg \max_{v' \in \text{children of } v} \frac{Q(v')}{N(v')} + c\sqrt{\frac{2 \ln N(v)}{N(v' )}}$ 
29: end function
30: function DEFAULTPOLICY( $s$ )
31:   while  $s$  is non-terminal do
32:      $s \leftarrow$  random child of  $s$ 
33:   end while
34:   return reward for state  $s$ 
35: end function
36: function BACKUP( $(v, \Delta)$ )
37:   while  $v$  is not null do
38:      $N(v) \leftarrow N(v) + 1$ 
39:      $Q(v) \leftarrow Q(v) + \Delta(v)$ 
40:      $v \leftarrow$  parent of  $v$ 
41:   end while
42: end function
```

F.2 Sequence with Finite Gate Set

For pulse sequences with delays and short pulses (e.g., solid-echo based sequences), it is important to first ensure that the errors during the windows are minimized. Then, one can choose the gates that induce the right transition of the toggling-frame Hamiltonian between two adjacent windows and minimize the total gate-dependent error. Assume that the unitary propagators of the candidate gates form a finite group $\mathcal{S}$, then the toggling-frame Hamiltonian during the p th window is chosen from a finite set with W elements:

$$\mathcal{T} = \{U^\dagger H_{\text{pert}} U | U \in \mathcal{S}\} = \{H_{\text{tog}}(1), \dots, H_{\text{tog}}(W)\}. \quad (\text{F.1})$$

Then for a sequence with P windows, Assume that we can apply any gate from $\mathcal{S}$ in front of each window, then a possible sequence of the toggling-frame Hamiltonian $H_{\text{tog}}(w)$ is $(\{H_{\text{tog}}(w_1), t_1\}, \dots, \{H_{\text{tog}}(w_P), t_P\})$, where $H_{\text{tog}}(w_p)$ is the toggling-frame Hamiltonian during the p th window and t_p is the length of the window. For simplicity, consider the case where $t_1 = \dots = t_P$. we will call the sequence a toggling-frame sequence and denote it with the labels $w_1, \dots, w_P$:

$$s = (w_1 \cdots w_P)_{\text{tog}}, \quad (\text{F.2})$$

where $1 \leq w_1, \dots, w_P \leq W$. The goal is to design s so that the accumulated errors during the windows are minimized. These errors, similar to those of any sequence, are determined by the $\mathcal{C}$ -integrals of s . Since $\mathcal{T}$ is typically restricted to a subspace of $\mathcal{C}(\mathbf{g}_{\text{pri}}, H_{\text{pert}})$, we can also use representation restricted to the subspace to calculate the $\mathcal{C}$ -integrals. An algorithm for finding the desired toggling-frame sequence is given at the end of the appendix.

Using the conditions found in the last section, we first try to find toggling sequences that decouple the interactions up to 4th order. We will design sequences for spin ensembles as examples since they are the hardest. We will take the dipolar interaction as an example.

Take as example engineering the secular dipolar interaction under high field with collective control, $\mathcal{C}(\mathbf{g}_{\text{pri}}, H_{\text{pert}})$ is five-dimensional. But for sequences with $\pi/2$ x and y rotations, $\mathcal{T}$ is restricted to a two-dimensional subspace spanned by

$$\frac{1}{2} \left\{ \frac{2\sigma_z \otimes \sigma_z - \sigma_x \otimes \sigma_x - \sigma_y \otimes \sigma_y}{\sqrt{6}}, \frac{\sigma_x \otimes \sigma_x - \sigma_y \otimes \sigma_y}{\sqrt{2}} \right\}. \quad (\text{F.3})$$

$\mathcal{T}$ has three elements:

$$|H_1\rangle\rangle = \left(-\frac{1}{2}, \frac{\sqrt{3}}{2}\right)^T, |H_2\rangle\rangle = \left(-\frac{1}{2}, -\frac{\sqrt{3}}{2}\right)^T, |H_3\rangle\rangle = (1, 0)^T. \quad (\text{F.4})$$

The necessary and sufficient conditions (using the method from chapter 3) for $\bar{H}^{(r-1)}(T_{\text{seq}}) = 0$ are

$$\begin{aligned}
\bar{H}^{(0)} = 0 : \quad & \bar{c}_1 = 0 \\
& \bar{c}_2 = 0 \\
\bar{H}^{(1)} = 0 : \quad & \bar{c}_{12} - \bar{c}_{21} = 0 \\
\bar{H}^{(2)} = 0 : \quad & \bar{c}_{112} - \frac{1}{2}\bar{c}_{121} + \bar{c}_{211} = 0 \\
& \bar{c}_{122} - \frac{1}{2}b\bar{c}_{212} + \bar{c}_{221} = 0 \\
\bar{H}^{(3)} = 0 : \quad & \bar{c}_{1121} - \bar{c}_{1211} = 0 \\
& \bar{c}_{1122} - \bar{c}_{1212} + \bar{c}_{2121} - \bar{c}_{2211} = 0 \\
& \bar{c}_{2122} - \bar{c}_{2212} = 0 \\
\bar{H}^{(4)} = 0 : \quad & \bar{c}_{11112} + \bar{c}_{11121} - \frac{1}{4}\bar{c}_{11211} + \bar{c}_{12111} + \bar{c}_{21111} = 0 \\
& \bar{c}_{11122} - \frac{1}{2}\bar{c}_{12112} + \bar{c}_{12121} + \bar{c}_{21112} - \frac{1}{2}\bar{c}_{21121} + \bar{c}_{22111} = 0 \\
& \bar{c}_{11212} + \bar{c}_{11221} - \frac{1}{2}\bar{c}_{12112} + \bar{c}_{12211} - \frac{1}{2}\bar{c}_{21121} + \bar{c}_{21211} = 0 \\
& \bar{c}_{11222} - \frac{1}{2}\bar{c}_{12212} + \bar{c}_{12221} + \bar{c}_{21212} - \frac{1}{2}\bar{c}_{21221} + \bar{c}_{22211} = 0 \\
& \bar{c}_{12122} - \frac{1}{2}\bar{c}_{12212} + \bar{c}_{21122} - \frac{1}{2}\bar{c}_{21221} + \bar{c}_{22112} + \bar{c}_{22121} = 0 \\
& \bar{c}_{12222} + \bar{c}_{21222} - \frac{1}{4}\bar{c}_{22122} + \bar{c}_{22212} + \bar{c}_{22221} = 0 \\
& \dots
\end{aligned} \tag{F.5}$$

Below are some of the toggling-frame sequences that achieve $\bar{H}^{(r-1)}(T_{\text{seq}}) = 0$ up to $r = 5$:

$$\begin{aligned}
& (133331221211213223) \\
& (131333221232121333112121222133313122) \\
& (121132232322332123111322333221323211321331131212113322211311231323322311) \\
& \dots
\end{aligned} \tag{F.6}$$

Alternatively, we can also design composite toggling-frame sequences using a set of

toggling-frame sequences. For example, let

$$\begin{aligned} \{s_1 = (123), s_2 = (132), s_3 = (213), \\ s_4 = (231), s_5 = (312), s_6 = (321)\}. \end{aligned} \quad (\text{F.7})$$

Then a composite toggling-frame sequence can be built using these ingredients. For example,

$$[s_1 s_6 s_2] = [(123)(321)(132)] = (123321132). \quad (\text{F.8})$$

This allows us to build sequences using subcycles that account for lower-order errors (the zeroth-order error in this case) to enhance convergence of the sequence. Below are some composite toggling-frame sequences that achieve $\bar{H}^{(r-1)}(T_{\text{seq}}) = 0$ up to $r = 5$:

$$\begin{aligned} & [s_3 s_4 s_4 s_1 s_2 s_2] \\ & [s_1 s_1 s_2 s_2 s_4 s_5 s_3 s_6 s_4 s_6 s_5 s_3] \\ & [s_1 s_3 s_6 s_6 s_6 s_3 s_4 s_3 s_5 s_4 s_4 s_6 s_4 s_5 s_2 s_6 s_2 s_5 s_1 s_4 s_1 s_5 s_2 s_3] \\ & \dots \end{aligned} \quad (\text{F.9})$$

Once a desired toggling-frame sequence is found, we can choose gates from $\mathcal{S}$ to insert between windows to make the correct transition of H_{tog} and minimize total gate-dependent errors using Algorithm 9.

F.3 Genetic Algorithm for Sequence Optimization

In principle, we can also use Algorithm 9 to design toggling-frame sequences. But we found that the genetic algorithms often perform better. Hence, we briefly describe the method here.

Genetic algorithms make relatively few assumptions, and are therefore applicable to many problems [88, 210]. The success of a genetic algorithm depends heavily on encoding and choice of genetic operators [136]. To choose a good encoding, we first note that the Hamiltonian-engineering properties of a toggling-frame sequence are invariant under rotations and inversion. A good encoding that exploits this property is the path representation which has been applied to the traveling salesman problem [115]. We also incorporate the zeroth-order decoupling condition into our encoding to reduce the size of sample space. We choose an order-based crossover operator [48] since it depends on the relative order of the gates. For mutation operator, we use a repeated exchange mutation operator [6, 4]. It connects every two states in the sample space, therefore, ensures accessibility of every possible

solution. In our experience, a high mutation probability is necessary to prevent premature convergence and maintain population diversity. A steady-state modification may also be applied to mitigate the problem [184]. Below, we give a simple genetic algorithm and the corresponding genetic operators for path encoding (for designing toggling-frame sequences) and N -character encoding (for designing composite toggling-frame sequences).

Algorithm 10 A simple genetic algorithm. For a population size W , a maximal iteration number $i_{\max}$, a crossover probability p_c and a mutation probability p_m , the function **GENETICALGORITHM**(f, f_{target}) takes a fitness function f and a target function value f_{target} as input and returns a solution $(x_1 x_2 \dots x_P)$ such that either $f(x_1 x_2 \dots x_P) \geq f_{\text{target}}$ or that is the best solution found within the maximal number of iterations $i_{\max}$.

```

1: function GENETICALGORITHM( $f, \Delta$ )
2:   for  $w = 1, 2, \dots, W$  do                                      $\triangleright$  Initial population
3:      $\vec{s}_w \leftarrow$  Random sequence of length  $P$ 
4:      $f_w \leftarrow f(\vec{s}_w)$ 
5:   end for
6:    $f_{\text{best}} \leftarrow \max \{f_1, f_2, \dots, f_W\}$                   $\triangleright$  Store the best solution found so far
7:    $\vec{s}_{\text{best}} \leftarrow \vec{s}_w$  that gives  $f_{\text{best}}$ 
8:    $i \leftarrow 1$ 
9:   while  $f_{\text{best}} < f_{\text{target}}$  and  $i < i_{\max}$  do
10:    CROSSOVER(SELECT( $\{f_w\}, \{\vec{s}_w\}$ )) until  $W$  new strings  $\vec{s}'_w$  are generated
11:     $\{\vec{s}_w\} \leftarrow \{\text{MUTATE}(\vec{s}'_w)\}$                         $\triangleright$  Replace the old generation with the offsprings
12:     $\{f_w\} \leftarrow \{f(\vec{s}_w)\}$ 
13:     $f_{\text{best}} \leftarrow \max \{f_{\text{best}}, f_1, f_2, \dots, f_W\}$ 
14:     $\vec{s}_{\text{best}} \leftarrow \vec{s}_w$  that gives the new  $f_{\text{best}}$ 
15:     $i \leftarrow i + 1$ 
16:   end while
17:   return  $\vec{s}_{\text{best}}$ 
18: end function

```

Algorithm 11 Genetic operators for path-encoded algorithms.

```
1: function SELECT( $\{f_w\}, \{\vec{s}_w\}$ ) ▷ Roulette wheel selection
2:    $\bar{f} \leftarrow \sum_{w=1}^W f_w$ 
3:   Sample two  $\vec{s}_w$  randomly according to the distribution  $\{f_w/\bar{f}\}$  and return the
   result
4: end function
5: function CROSSOVER1( $\vec{s}_1, \vec{s}_2$ )
6:   Let  $\vec{s}_{1/2} = (x_1^{1/2} x_2^{1/2} \dots x_P^{1/2})$ 
7:    $r \leftarrow$  random real number from  $[0, 1]$ 
8:   if  $r \leq p_c$  then
9:     Choose a random  $p_1$  and  $p_2$  with  $p_2 > p_1$  from  $\{1, 2, \dots, P\}$ 
10:     $\vec{s}_{1/2}^m \leftarrow (x_{p_1}^{1/2} \dots x_{p_2}^{1/2})$ 
11:     $\vec{s}_{1/2}^o \leftarrow (\vec{s}_{1/2}$  with the first  $p_2 - p_1$  elements that are identical with elements in
     $\vec{s}_{1/2}^m$  removed)
12:     $\vec{s}_{1/2}^n \leftarrow \text{RotateRight}(\text{Join}(\vec{s}_{1/2}^m, \vec{s}_{1/2}^o), \text{by } p_1 - 1)$ 
13:    return  $\vec{s}_1^n$  and  $\vec{s}_2^n$ 
14:   else
15:     return  $\vec{s}_1$  and  $\vec{s}_2$ 
16:   end if
17: end function
18: function MUTATE1( $\vec{s}$ )
19:    $r \leftarrow$  random real number from  $[0, 1]$ 
20:   if  $r \leq p_m$  then
21:     Generate a random integer  $m$  with probability  $p^{m-1}(1-p)$ 
22:     Choose with replacement  $m$  random pairs of elements from  $\vec{s}$  and swap their
    positions
23:   end if
24:   return the new  $\vec{s}$ 
25: end function
```

Algorithm 12 Genetic operators for N -character encoded genetic algorithms.

```

1: function CROSSOVER2( $\vec{s}_1, \vec{s}_2$ ) ▷ 1-point crossover
2:   Let  $\vec{s}_{1/2} = (x_1^{1/2}, x_2^{1/2}, \dots, x_n^{1/2})$ 
3:    $r \leftarrow$  random real number from  $[0, 1]$ 
4:   if  $r \leq p_c$  then
5:     Choose a random  $p'$  from  $\{2, 3, \dots, P\}$ 
6:     return  $(x_1^1, \dots, x_{p'-1}^1, x_{p'}^2, \dots, x_P^2)$  and  $(x_1^2, \dots, x_{p'-1}^2, x_{p'}^1, \dots, x_P^1)$ 
7:   else
8:     return  $\vec{s}_1$  and  $\vec{s}_2$ 
9:   end if
10: end function
11: function MUTATE2( $\vec{s}$ ) ▷ Replacement mutation
12:   Let  $\vec{s} = (x_1, x_2, \dots, x_n)$ 
13:   for  $i = 1, 2, \dots, n$  do
14:      $r \leftarrow$  random real number from  $[0, 1]$ 
15:     if  $r \leq p_m$  then
16:        $x'_i \leftarrow \text{RandomInteger}[1, N]$ 
17:     else
18:        $x'_i \leftarrow x_i$ 
19:     end if
20:   end for
21:   return  $(x'_1, x'_2, \dots, x'_n)$ 
22: end function

```

Appendix G

Optimization Algorithms for Pulse Engineering

All algorithms presented in this thesis were manually implemented in Mathematica to allow for greater customization in pulse engineering. This appendix provides a brief overview of two additional local numerical optimization algorithms utilized in the pulse design process: the Nelder-Mead simplex method and the nonlinear conjugate gradient method.

G.1 Nelder-Mead Simplex Method

The Nelder-Mead simplex method is a direct search local optimizer that does not use the derivative of a function [142]. It is useful compared with gradient-based method when the function is non-differentiable or noisy, or one cares more about a better solution rather than the absolute optimum [105]. It can be used to tune the hyperparameters of a stochastic algorithm [149]. Here, we follow a modern description of the original algorithm [113].

Denote a simplex with its unordered vertices and the corresponding function values as

$$\Delta = \{\{\vec{x}_1, f_1\}, \{\vec{x}_2, f_2\}, \dots, \{\vec{x}_{n+1}, f_{n+1}\}\}. \quad (\text{G.1})$$

Let

$$M = (\vec{x}_1 - \vec{x}_{n+1}, \vec{x}_2 - \vec{x}_{n+1}, \dots, \vec{x}_n - \vec{x}_{n+1}), \quad (\text{G.2})$$

the n -dimensional volume of Δ is

$$\text{vol}(\Delta) = \frac{|\det(M)|}{n!}. \quad (\text{G.3})$$

Δ is nondegenerate if $\text{vol}(\Delta) > 0$. The diameter of Δ is

$$\text{diam}(\Delta) = \max_{i \neq j} \|\vec{x}_i - \vec{x}_j\|. \quad (\text{G.4})$$

We can terminate the algorithm when either the volume or the diameter of the simplex falls below a threshold.

One way for finding an initial simplex as suggested in [142] is to choose a random first point $\vec{x}_1$, then generate the rest by taking $\vec{x}_{i+1} = h_i \vec{e}_i$ for $i = 1, \dots, n$ where h_i are scaling factors and $\vec{e}_i$ are unit vectors of the i th coordinate.

The pseudocode of the Nelder-Mead simplex method is provided in Algorithm 13.

G.2 Conjugate Gradient Method

Gradient-based optimization algorithms are widely used in optimal quantum control. Generally, such an algorithm requires a search direction generator and a line search strategy. The search direction generator determines the search direction using the gradient, while the line search strategy selects the step length.

Here, we briefly describe the method (nonlinear conjugate gradient with quadratic line search) used in this thesis.

G.3 Quadratic Line Search

Here, we define a 3-point bracket that is used in the quadratic line search. A 3-point bracket is $\lambda_1 < \lambda_2 < \lambda_3$, such that $f(\lambda_2) \leq \min\{f(\lambda_1), f(\lambda_3)\}$. Consider $\min f(\lambda)$ over $\lambda \geq 0$. To find a 3-point bracket, generate the sequence

$$\begin{aligned} \lambda_0 &= 0, \\ \lambda_r &= \lambda_{r-1} + 2^{r-1} \Delta, \end{aligned} \quad (\text{G.5})$$

for $r = 1, 2, \dots$ and a small positive number Δ . Find k such that $f(\lambda_0) > \dots > f(\lambda_k)$ and $f(\lambda_{k+1}) > f(\lambda_k)$, the 3-point bracket is then given by

$$\begin{cases} (\lambda_{k-1}, \lambda_k, (\lambda_k + \lambda_{k+1})/2) & f(\lambda_k) < f((\lambda_k + \lambda_{k+1})/2) \\ (\lambda_k, (\lambda_k + \lambda_{k+1})/2, \lambda_{k+1}) & f(\lambda_k) \geq f((\lambda_k + \lambda_{k+1})/2) \end{cases}. \quad (\text{G.6})$$

Algorithm 13 Nelder-Mead simplex method. For fixed ρ, χ, γ and σ . For fixed $i_{\max}$ and ϵ . For bounded constraint, we can legalize the point by projecting it to the boundary when it falls out of the legal region.

```

1: function NM( $f, \Delta$ )
2:   Order  $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_{n+1} \in \mathbb{R}^N$  such that  $f(\vec{x}_1) \leq f(\vec{x}_2) \dots \leq f(\vec{x}_{n+1})$ .
3:    $\bar{x} \leftarrow \sum_{i=1}^n \vec{x}_i / n$ 
4:    $\vec{x}_r \leftarrow \bar{x} + \rho (\bar{x} - \vec{x}_{n+1})$  ▷ Reflect
5:    $f_r \leftarrow f(\vec{x}_r)$ 
6:   if  $f_1 \leq f_r < f_n$  then
7:     return  $\{\{\vec{x}_1, f_1\}, \{\vec{x}_2, f_2\}, \dots, \{\vec{x}_n, f_n\}, \{\vec{x}_r, f_r\}\}$ 
8:   else if  $f_r < f_1$  then
9:      $\vec{x}_e = \bar{x} + \chi (\vec{x}_r - \bar{x})$  ▷ Expand
10:     $f_e \leftarrow f(\vec{x}_e)$ 
11:    if  $f_e < f_r$  then
12:      return  $\{\{\vec{x}_1, f_1\}, \{\vec{x}_2, f_2\}, \dots, \{\vec{x}_n, f_n\}, \{\vec{x}_e, f_e\}\}$ 
13:    else
14:      return  $\{\{\vec{x}_1, f_1\}, \{\vec{x}_2, f_2\}, \dots, \{\vec{x}_n, f_n\}, \{\vec{x}_r, f_r\}\}$ 
15:    end if
16:  else if  $f_n \leq f_r \leq f_{n+1}$  then
17:     $\vec{x}_c \leftarrow \bar{x} + \gamma (\vec{x}_r - \bar{x})$  ▷ Outside contract
18:     $f_c \leftarrow f(\vec{x}_c)$ 
19:    if  $f_c \leq f_r$  then
20:      return  $\{\{\vec{x}_1, f_1\}, \{\vec{x}_2, f_2\}, \dots, \{\vec{x}_n, f_n\}, \{\vec{x}_c, f_c\}\}$ 
21:    end if
22:  else
23:     $\vec{x}_{cc} \leftarrow \bar{x} - \gamma (\bar{x} - \vec{x}_{n+1})$  ▷ Inside contract
24:     $f_{cc} \leftarrow f(\vec{x}_{cc})$ 
25:    if  $f_{cc} < f_{n+1}$  then
26:      return  $\{\{\vec{x}_1, f_1\}, \{\vec{x}_2, f_2\}, \dots, \{\vec{x}_n, f_n\}, \{\vec{x}_{cc}, f_{cc}\}\}$ 
27:    end if
28:  end if
29:  for  $i = 2, \dots, n+1$  do ▷ Shrink
30:     $\vec{v}_i \leftarrow \vec{x}_1 + \sigma (\vec{x}_i - \vec{x}_1)$ 
31:     $f_i \leftarrow f(\vec{v}_i)$ 
32:  end for
33:  return  $\{\{\vec{x}_1, f_1\}, \{\vec{v}_2, f_2\}, \dots, \{\vec{v}_n, f_n\}, \{\vec{v}_{n+1}, f_{n+1}\}\}$ 
34: end function

```

Algorithm 14 Nelder-Mead simplex method (continued).

```

35: function FINDMINNM( $f, \Delta_0, f_{\text{target}}$ )
36:    $i \leftarrow 1$ 
37:   while The diameter of  $\Delta_i$  is greater than  $\epsilon$  and  $i \leq i_{\text{max}}$  do
38:      $\Delta_i \leftarrow \text{NM}(\Delta_{i-1})$ 
39:     return the vertex of  $\Delta_i$  with minimal function value if it is smaller or equal to
        $f_{\text{target}}$ 
40:      $i \leftarrow i + 1$ 
41:   end while
42:   return the vertex of  $\Delta_i$  with minimal function value
43: end function

```

A pseudocode is provided in Algorithm 15.

Let $(\lambda_1, \lambda_2, \lambda_3)$ be the current 3-point bracket. Fit a quadratic approximation $Q(\lambda) = a\lambda^2 + b\lambda + c$ to $f(\lambda)$ using function values at $\lambda_1, \lambda_2, \lambda_3$. The minimum of $Q(\lambda)$ is given by

$$\hat{\lambda} = \frac{(\lambda_2^2 - \lambda_3^2)f(\lambda_1) + (\lambda_3^2 - \lambda_1^2)f(\lambda_2) + (\lambda_1^2 - \lambda_2^2)f(\lambda_3)}{2[(\lambda_2 - \lambda_3)f(\lambda_1) + (\lambda_3 - \lambda_1)f(\lambda_2) + (\lambda_1 - \lambda_2)f(\lambda_3)]} \quad (\text{G.7})$$

If $\hat{\lambda} \neq \lambda_2$, the new bracket is given by

$$\begin{cases} (\lambda_1, \min\{\lambda_2, \hat{\lambda}\}, \max\{\lambda_2, \hat{\lambda}\}) & f(\min\{\lambda_2, \hat{\lambda}\}) < f(\max\{\lambda_2, \hat{\lambda}\}) \\ (\min\{\lambda_2, \hat{\lambda}\}, \max\{\lambda_2, \hat{\lambda}\}, \lambda_3) & f(\min\{\lambda_2, \hat{\lambda}\}) \geq f(\max\{\lambda_2, \hat{\lambda}\}) \end{cases} \quad (\text{G.8})$$

If $\hat{\lambda} = \lambda_2$, return λ_2 if $\lambda_3 - \lambda_1$ is smaller than a positive tolerance ϵ or let

$$\hat{\lambda} = \begin{cases} \lambda_2 + \epsilon/2 & \lambda_2 - \lambda_1 \leq \lambda_3 - \lambda_2 \\ \lambda_2 - \epsilon/2 & \lambda_2 - \lambda_1 > \lambda_3 - \lambda_2 \end{cases} \quad (\text{G.9})$$

Compute $f(\hat{\lambda})$ and go to the above step again. Return $\min_f\{\lambda_2, \hat{\lambda}\}$ if $\max\{f(\lambda_1), f(\lambda_3)\} - f(\lambda_2)$ or $\lambda_3 - \lambda_1$ becomes smaller than a tolerance. A pseudocode is provided in Algorithm 16.

For discrete optimization, choose Δ to be the smallest step allowed. Legalize each λ_i . Terminate the algorithm if $\lambda_3 - \lambda_1 \leq \Delta$.

G.4 Nonlinear Conjugate Gradient

The nonlinear conjugate gradient (NCG) method is used to find a local minimum of a function. Compared with the gradient descent method, NCG replaces the steepest descent direction with a conjugate direction. Given a function $f(\vec{x}_n)$ where $\vec{x}_n = (x_n^1, x_n^2, \dots, x_n^m)$ is a m -dimensional real vector, the steepest descent direction at $\vec{x}_n$ is given by

$$\vec{g}_n = -\nabla f(\vec{x}_n), \quad (\text{G.10})$$

where $\nabla f(\vec{x}_n)$ is the gradient of f at $\vec{x}_n$. The conjugate direction is

$$\vec{s}_n = \vec{g}_n + \beta_n \vec{s}_{n-1}. \quad (\text{G.11})$$

β_n can be calculated using the Polak-Ribière formula

$$\beta_n^{PR} = \frac{\vec{g}_n \cdot (\vec{g}_n - \vec{g}_{n-1})}{\vec{g}_{n-1} \cdot \vec{g}_{n-1}}. \quad (\text{G.12})$$

Convergence of the Polak-Ribière method is guaranteed by choosing

$$\beta_n = \max\{\beta_n^{PR}, 0\}. \quad (\text{G.13})$$

This is equivalent to restarting the NCG algorithm when β_n^{PR} is negative. An introduction to conjugate gradient method can be found in [173]. The gradient descent method is obtained by setting $\beta_n \equiv 0$. A legalizer is used to constrain each component of $\vec{x}_n$ in $[-1, 1]$. Algorithm 17 gives the pseudocode of NCG.

Algorithm 15 Equi-distance three-point bracket. Given an initial step length Δ , a maximal number of iterations and an error tolerance, $\text{THREEPNTBRACKET}(f)$ takes a single-variable function $f(x)$ and returns a three-point bracket as described in Sec. G.3 and their function values.

```

1: function THREEPNTBRACKET( $f$ )
2:    $f_0 \leftarrow f(0)$ 
3:    $\lambda_1 \leftarrow \Delta$ 
4:    $f_1 \leftarrow f(\lambda_1)$ 
5:   while  $f_1 > f_0$  do
6:      $\Delta \leftarrow \Delta/10$ 
7:     if  $\Delta > \epsilon$  then
8:        $\lambda_1 \leftarrow \Delta$ 
9:        $f_1 \leftarrow f(\lambda_1)$ 
10:    else
11:      return  $((0, 0, 0), (f_0, f_0, f_0))$ 
12:    end if
13:  end while
14:   $i \leftarrow 1$ 
15:  while  $i \leq i_{\max}$  and  $f_i < f_{i-1}$  do
16:     $\lambda_{i+1} \leftarrow \lambda_i + 2^i \Delta$ 
17:     $f_{i+1} \leftarrow f(\lambda_{i+1})$ 
18:     $i \leftarrow i + 1$ 
19:  end while
20:   $f_{\text{mid}} \leftarrow f((\lambda_i + \lambda_{i-1})/2)$ 
21:  if  $f_{i-1} < f_{\text{mid}}$  then
22:    return  $((\lambda_{i-2}, \lambda_{i-1}, (\lambda_i + \lambda_{i-1})/2), (f_{i-2}, f_{i-1}, f_{\text{mid}}))$ 
23:  else
24:    return  $((\lambda_{i-1}, (\lambda_i + \lambda_{i-1})/2, \lambda_i), (f_{i-1}, f_{\text{mid}}, f_i))$ 
25:  end if
26: end function

```

Algorithm 16 Quadratic fit line search. Given a maximal number of iterations, an error tolerance ϵ and a minimal increase of function value ϵ_f , QLS(f) takes a single-variable function f and returns x such that x is a local minimum of f .

```

1: function QMIN( $\vec{\lambda}, \vec{f}$ )
2:   return  $\frac{(\lambda_2^2 - \lambda_3^2)f_1 + (\lambda_3^2 - \lambda_1^2)f_2 + (\lambda_1^2 - \lambda_2^2)f_3}{2[(\lambda_2 - \lambda_3)f_1 + (\lambda_3 - \lambda_1)f_2 + (\lambda_1 - \lambda_2)f_3]}$  ▷ See Eq. (G.7).
3: end function
4: function QLS( $f$ )
5:   Define  $\vec{\lambda} = (\lambda_1, \lambda_2, \lambda_3)$  and  $\vec{f} = (f_1, f_2, f_3)$ 
6:    $(\vec{\lambda}, \vec{f}) \leftarrow \text{THREEPNTBRACKET}(f)$ 
7:   if  $\vec{\lambda} = (0, 0, 0)$  then
8:     return 0
9:   end if
10:   $\hat{\lambda} \leftarrow \text{QMIN}(\vec{\lambda}, \vec{f})$ 
11:  while  $i \leq i_{\max}$  and  $\lambda_3 - \lambda_1 > \epsilon$  and  $\max\{f_1, f_3\} - f_2 > \epsilon_f$  do
12:     $\hat{f} \leftarrow f(\hat{\lambda})$ 
13:    if  $\hat{\lambda} > \lambda_2$  then
14:      if  $\hat{f} > f_2$  then
15:         $\vec{\lambda} \leftarrow (\lambda_1, \lambda_2, \hat{\lambda})$ 
16:         $f_3 \leftarrow \hat{f}$ 
17:      else
18:         $\vec{\lambda} \leftarrow (\lambda_2, \hat{\lambda}, \lambda_3)$ 
19:         $f_1 \leftarrow f_2$ 
20:         $f_2 \leftarrow \hat{f}$ 
21:      end if
22:       $\hat{\lambda} \leftarrow \text{QMIN}(\vec{\lambda}, \vec{f})$ 
23:    else if  $\hat{\lambda} < \lambda_2$  then
24:      if  $\hat{f} > f_2$  then
25:         $\vec{\lambda} \leftarrow (\hat{\lambda}, \lambda_2, \lambda_3)$ 
26:         $f_1 \leftarrow \hat{f}$ 
27:      else

```

Algorithm 16 Quadratic fit line search (continued).

```
28:       $\vec{\lambda} \leftarrow (\lambda_1, \hat{\lambda}, \lambda_2)$ 
29:       $f_3 \leftarrow f_2$ 
30:       $f_2 \leftarrow \hat{f}$ 
31:      end if
32:       $\hat{\lambda} \leftarrow \text{QMIN}(\vec{\lambda}, \vec{f})$ 
33:      else
34:      if  $\lambda_2 - \lambda_1 < \lambda_3 - \lambda_2$  then
35:       $\hat{\lambda} \leftarrow \lambda_2 + \epsilon/2$ 
36:      else
37:       $\hat{\lambda} \leftarrow \lambda_2 - \epsilon/2$ 
38:      end if
39:      end if
40:      end while
41:      return  $\lambda_2$ 
42: end function
```

Algorithm 17 Nonlinear conjugate gradient method. Given a maximal iteration number $i_{\max}$, a minimal increase of target function ϵ_f and an error tolerance ϵ , the function $\text{NCG}(f, \vec{x}_0, f_{\text{target}})$ takes a multivariable function f , an initial seed $\vec{x}_0$ and a target function value f_{target} to find $\vec{x}$ such that $f(\vec{x}) \leq f_{\text{target}}$. The algorithm terminates when the maximal number of iterations is reached or $|\vec{g}| \leq \epsilon|\vec{g}_0|$. The CG restarts every $k_{\max}$ iterations. The function $\text{LEGALIZE}(\vec{x}_n, \vec{s}_n)$ legalizes $\vec{s}_n$ with respect to $\vec{x}_n$.

```

1: function  $\text{LEGALIZE}(\vec{x}_n, \vec{s}_n)$ 
2:   Define  $\vec{s} = (s^1, s^2, \dots, s^m)$ 
3:   for  $i \in \{1, 2, \dots, m\}$  do
4:     if  $|x_n^i| = 1$  and  $s_n^i/x_n^i > 0$  then
5:        $s^i \leftarrow 0$ 
6:     else
7:        $s^i \leftarrow s_n^i$ 
8:     end if
9:   end for
10:  return  $\vec{s}$ 
11: end function
12: function  $\text{NCG}(f, \vec{x}_0, f_{\text{target}})$ 
13:   $\vec{x} \leftarrow \vec{x}_0$ 
14:   $\vec{g}_{\text{old}} = -\nabla f(\vec{x})$ 
15:   $\vec{s} \leftarrow \vec{g}_{\text{old}}$ 
16:   $\delta_{\text{new}} \leftarrow |\vec{g}_{\text{old}}|^2$ 
17:   $\delta_0 \leftarrow \delta_{\text{new}}$ 
18:   $i \leftarrow 0$ 
19:   $f_{\text{old}} \leftarrow f(\vec{x})$ 
20:  while  $i < i_{\max}$  and  $\delta_{\text{new}} > \epsilon^2 \delta_0$  and  $f_{\text{old}} > f_{\text{target}}$  do
21:     $\vec{s} \leftarrow \text{LEGALIZE}(\vec{x}, \vec{s})$  ▷ Legalize  $\vec{s}$  with respect to  $\vec{x}$ .
22:    Define  $h(x) = f(\vec{x} + x\vec{s})$ 
23:     $\alpha \leftarrow \text{QLS}(h)$  ▷ Calculate the step length using quadratic fit line search.
24:     $\vec{x} \leftarrow \vec{x} + \alpha\vec{s}$ 
25:     $f_{\text{new}} \leftarrow f(\vec{x})$ 
26:    if  $f_{\text{old}} - f_{\text{new}} \leq \epsilon_f$  then
27:      return  $\vec{x}$ 
28:    end if
29:     $f_{\text{old}} \leftarrow f_{\text{new}}$ 
30:     $\delta_{\text{old}} \leftarrow \delta_{\text{new}}$ 
31:     $\vec{g}_{\text{new}} \leftarrow -\nabla f(\vec{x})$ 
32:     $\delta_{\text{mid}} \leftarrow \vec{g}_{\text{new}} \cdot \vec{g}_{\text{old}}$ 

```

Algorithm 17 Nonlinear conjugate gradient method (continued).

```

33:       $\delta_{\text{new}} \leftarrow \vec{g}_{\text{new}} \cdot \vec{g}_{\text{new}}$ 
34:       $\beta \leftarrow \frac{\delta_{\text{new}} - \delta_{\text{mid}}}{\delta_{\text{old}}}$  ▷ See Eq. (G.12).
35:       $i \leftarrow i + 1$ 
36:       $k \leftarrow k + 1$ 
37:      if  $k \geq k_{\text{max}}$  or  $\beta \leq 0$  then
38:           $\vec{s} \leftarrow \vec{g}_{\text{new}}$  ▷ Restart CG when  $k \geq k_{\text{max}}$  or  $\beta \leq 0$ .
39:           $k \leftarrow 0$ 
40:      else
41:           $\vec{s} \leftarrow \vec{g}_{\text{new}} + \beta \vec{s}$ 
42:      end if
43:       $\vec{g}_{\text{old}} \leftarrow \vec{g}_{\text{new}}$ 
44:  end while
45:  return  $\vec{x}$ 
46: end function

```

Appendix H

Markov Approximation of the Environmental Dynamics

This appendix first derives a master equation of the dynamics of the CS system under a Markovian assumption of the environmental dynamics. Then we show how to derive the Gauss-Markov model of the environmental dynamics.

H.1 Non-Markovianity of the Central Spin System

Here, an exact definition of the non-Markovianity of the environmental dynamics is defined using the projector technique [141, 222, 129].

Consider the Lie algebra of controls on the CS and $H_{cs/e} = \sigma_z \otimes A/2$, denoted as $\mathfrak{g} = \text{span}_{\text{Lie}}\{iH_{cs/e}, \mathfrak{u}(2) \otimes \mathbb{1}^{\otimes N}\}$. The orbit of $\rho(0)$ under $e\mathfrak{g}$ is

$$\mathcal{O}_s = \{U\rho(0)U^\dagger | U \in e\mathfrak{g}\}. \quad (\text{H.1})$$

Any state from the closure $[\mathcal{O}_s]$ can evolve back to $\rho(0)$ under some $U \in e\mathfrak{g}$ with arbitrary accuracy. Therefore, any state from $[\mathcal{O}_s]$ preserves coherence perfectly with respect to $\mathfrak{g}$. The dynamics of the environment lead to the loss of coherence by propagating a state away from $[\mathcal{O}_s]$. For example, consider $\rho(0) = \sigma_x \otimes \mathbb{1}^{\otimes N}$; under the evolution governed by $H_{cs/e}$, the state evolves as

$$\rho(t) = \sigma_x \otimes \cos At + \sigma_y \otimes \sin At. \quad (\text{H.2})$$

The evolution of $\rho(t)$ under the environmental dynamics $H_e = \mathbb{1} \otimes B$ is governed by the time dependence of $A'(t) = e^{iBt} A e^{-iBt}$. In the Liouville representation, defining $P = |A\rangle\langle A|$, $Q = 1 - P$, $L = i\text{ad}_B$, and $|A'_{\text{rel}}\rangle\rangle = P|A'\rangle\rangle$, the Nakajima-Zwanzig equation is given by

$$\frac{d}{dt}|A'_{\text{rel}}\rangle\rangle = PL|A'_{\text{rel}}\rangle\rangle + \int_0^t dt' K(t')|A'_{\text{rel}}(t-t')\rangle\rangle, \quad (\text{H.3})$$

where $K(t)$ is the memory kernel:

$$K(t) = PL e^{QLt} QLP. \quad (\text{H.4})$$

Since $PLP = i\text{Tr}(A[B, A]) = 0$, the memory kernel reduces to

$$K(t) = PL e^{Lt} LP. \quad (\text{H.5})$$

Substituting Eq. (H.5) into Eq. (H.3) and left-multiplying both sides by $\langle\langle A|$, results in

$$\frac{d}{dt}\mathcal{C}(t) = - \int_0^t dt' \langle\langle A|\text{ad}_B e^{i\text{ad}_B t'} \text{ad}_B|A\rangle\rangle \mathcal{C}(t-t'), \quad (\text{H.6})$$

where $PLP = 0$ is used again to eliminate the first term. Here, $\mathcal{C}(t) = \langle\langle A|A'(t)\rangle\rangle$ represents the coherence. The Markovian master equation is obtained by setting $t - t' = t$ in the integrand and extending the upper limit to infinity:

$$\frac{d}{dt}\mathcal{C}^*(t) = - \int_0^\infty dt' \langle\langle A|\text{ad}_B e^{i\text{ad}_B t'} \text{ad}_B|A\rangle\rangle \mathcal{C}^*(t). \quad (\text{H.7})$$

The difference between the observed coherence and $\mathcal{C}^*$ indicates non-Markovian behavior in the environment. Specifically, a characteristic of a Markovian environment is an exponentially decaying autocorrelation function:

$$\langle A'(t)A'(0) \rangle = e^{-t/\tau_c} \langle A^2 \rangle, \quad (\text{H.8})$$

where τ_c is the correlation time, defined as

$$\tau_c = \int_0^\infty \langle A'(t')A'(0) \rangle dt' / \langle A^2 \rangle. \quad (\text{H.9})$$

Note that, the derivation of Eq. (H.7) takes two steps; if only the time convolution is removed without extending the upper limit of the integral to infinity, the equation remains time-local. Although the dynamics of $\mathcal{C}(t)$ can be non-exponential in this scenario, non-Markovianity is not guaranteed. A sufficient condition for non-Markovianity is the non-divisibility of the dynamics resulting from the time convolution.

H.2 Stochastic Description using the Gauss-Markov Process

Here, a derivation of the Gauss-Markov model is provided to help better understand the assumptions. In the local field picture, the state of the CS evolves under $H_{s/e}$ as

$$\rho_s(t) = \sum_k e^{-i\sigma_z \Omega_k t} \rho_s(0) e^{i\sigma_z \Omega_k t}. \quad (\text{H.10})$$

When the number of spins in the environment N is large, the local fields form a continuous spectrum Ω . For each path of $\Omega(t)$, we have

$$\frac{d}{dt} \rho_{cs} = -i[\Omega(t)\sigma_z, \rho_{cs}], \quad (\text{H.11})$$

where $\Omega(t)$ is a stochastic process. The ensemble average of the $\rho_{cs}(t)$ can be represented as a path integral

$$\bar{\rho}_s(t) = \int \mathcal{D}[\Omega(\tau)] P[\Omega(\tau)] e^{-i\sigma_z \Omega t} \rho_s(0) e^{i\sigma_z \Omega t}. \quad (\text{H.12})$$

Consider $H_{cs/e} = \sigma_z \otimes \sum_{k=1}^N \omega_k \sigma_z^k$. Under the computational basis, a typical state of the environment takes the form of $|001011\dots\rangle$. The corresponding local field is $\sum_{i=1}^m \omega_i - \sum_{j=m+1}^N \omega_j$ where m is the number of spin-up spins. We further assume that ω_i are i.i.d. random variables. A change of local field associated with a permutation of two spins is $\delta_{ij} = 2(\omega_i - \omega_j)$ which is also a random variable. When averaged over a large ensemble, the change of local field is approximated by $\bar{\delta}\zeta'$ with each spin swap, where $\bar{\delta} = \mathbb{E}(|\delta_{ij}|)$ and $\zeta' = \pm 1$. Assume that $\Omega_{\max} - \Omega_{\min} = M\bar{\delta}$, we have

$$\Pr(\zeta' = \pm 1) = \frac{1}{2} \mp \frac{\Omega_t}{M\bar{\delta}}, \quad (\text{H.13})$$

where Ω_t is the local field at time t . It has also been assumed that the mean of the distribution of local field is 0, which is valid when N is large. Therefore, after time dt , the change of local field is

$$d\Omega_t = \bar{\delta} \sum_{i=1}^{\lfloor \gamma dt \rfloor} \zeta'_i = -2\gamma \frac{\Omega_t}{M} dt + \bar{\delta} S_{\lfloor \gamma dt \rfloor}, \quad (\text{H.14})$$

where $S_k = \sum_{i=1}^k \zeta_k$ is a 1-dimensional random walk, $\Pr(\zeta_k = \pm 1) = 1/2$ and γ is the number of spin swaps in unit time. Define the rescaled time and local field as $t' = \frac{\gamma t}{M}$ and $\Omega'_{t'} = \frac{\Omega_t}{\bar{\delta}\sqrt{M}}$, we have

$$d\Omega'_{t'} = -2\Omega'_{t'} dt' + dW_{t'}, \quad (\text{H.15})$$

where the Donsker's theorem has been used to yield the Wiener process as a scaling limit of the random walk: $W_{t'} = \lim_{M \rightarrow \infty} S_{[Mdt']}/\sqrt{M}$. Eq. (H.15) is the Itô stochastic differential equation of an Ornstein-Uhlenbeck process. According to Itô's lemma, $dW_{t'} = \sqrt{\frac{dt'}{dt}} dW_t$. Define $\sigma = \bar{\delta}\sqrt{M}$ and $\gamma' = \gamma/M$, Eq. (H.15) can be written as

$$d\Omega_t = -2\gamma'\Omega_t dt + \sigma\sqrt{\gamma'} dW_t. \quad (\text{H.16})$$

γ' is the normalized spin swapping rate, $\sigma^2/4$ is the variance of the local field distribution at time t . The corresponding Fokker-Planck equation is

$$\frac{\partial P}{\partial t} = -\Gamma P(\Omega, t), \quad (\text{H.17})$$

where $P(\Omega, t)$ is the probability to find the environment in the state $|\Omega\rangle$ and $\Gamma = \frac{\partial}{\partial \Omega} \left(2\gamma'\Omega + \frac{\sigma^2\gamma'}{2} \frac{\partial}{\partial \Omega} \right)$. Similar processes have been considered in Kubo's theory of the stochastic Liouville equation [111, 186, 187].

References

- [1] Anatole Abragam. *The principles of nuclear magnetism*. Oxford university press, 1961.
- [2] Hirotugu Akaike. A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6):716–723, 1974.
- [3] Gonzalo A Álvarez and Dieter Suter. Measuring the spectrum of colored noise by dynamical decoupling. *Physical review letters*, 107(23):230501, 2011.
- [4] Balamurali Krishna Ambati, Jayakrishna Ambati, and Mazen Moein Mokhtar. Heuristic combinatorial optimization by simulated darwinian evolution: a polynomial time algorithm for the traveling salesman problem. *Biological Cybernetics*, 65(1):31–35, 1991.
- [5] PW Anderson. Spectral diffusion, phonons, and paramagnetic spin-lattice relaxation. *Physical Review*, 114(4):1002, 1959.
- [6] Wolfgang Banzhaf. The “molecular” traveling salesman. *Biological Cybernetics*, 64(1):7–14, 1990.
- [7] Thomas M Barbara, Joel F Martin, and Jon G Wurl. Phase transients in nmr probe circuits. *Journal of Magnetic Resonance (1969)*, 93(3):497–508, 1991.
- [8] C Bradford Barber, David P Dobkin, and Hannu Huhdanpaa. The quickhull algorithm for convex hulls. *ACM Transactions on Mathematical Software (TOMS)*, 22(4):469–483, 1996.
- [9] Jean Baum, Michael Munowitz, Allen N Garroway, and Alex Pines. Multiple-quantum dynamics in solid state nmr. *The Journal of chemical physics*, 83(5):2015–2025, 1985.

- [10] Yuval Baum, Mirko Amico, Sean Howell, Michael Hush, Maggie Liuzzi, Pranav Mundada, Thomas Merkh, Andre RR Carvalho, and Michael J Biercuk. Experimental deep reinforcement learning for error-robust gate-set design on a superconducting quantum computer. *PRX Quantum*, 2(4):040324, 2021.
- [11] Charles H Bennett and Gilles Brassard. Quantum cryptography: Public key distribution and coin tossing. *Theoretical computer science*, 560:7–11, 2014.
- [12] Charles H Bennett, David P DiVincenzo, John A Smolin, and William K Wootters. Mixed-state entanglement and quantum error correction. *Physical Review A*, 54(5):3824, 1996.
- [13] Ji Bian, Pengfei Lu, Teng Liu, Hao Wu, Xinxin Rao, Kunxu Wang, Qifeng Lao, Yang Liu, Feng Zhu, and Le Luo. Quantum simulation of a general anti-pt-symmetric hamiltonian with a trapped ion qubit. *Fundamental Research*, 3(6):904–908, 2023.
- [14] 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, 2009.
- [15] MJ Biercuk, AC Doherty, and H Uys. Dynamical decoupling sequence construction as a filter-design problem. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 44(15):154002, 2011.
- [16] Sergio Blanes, Fernando Casas, Jose-Angel Oteo, and José Ros. The magnus expansion and some of its applications. *Physics reports*, 470(5-6):151–238, 2009.
- [17] F Bloch and A Siegert. Magnetic resonance for nonrotating fields. *Physical Review*, 57(6):522, 1940.
- [18] Nicolaas Bloembergen. On the interaction of nuclear spins in a crystalline lattice. *Physica*, 15(3-4):386–426, 1949.
- [19] Nicolaas Bloembergen, Edward Mills Purcell, and Robert V Pound. Relaxation effects in nuclear magnetic resonance absorption. *Physical review*, 73(7):679, 1948.
- [20] Geoffrey Bodenhausen, Herbert Kogler, and RR Ernst. Selection of coherence-transfer pathways in nmr pulse experiments. *Journal of Magnetic Resonance (1969)*, 58(3):370–388, 1984.
- [21] John Adrian Bondy, Uppaluri Siva Ramachandra Murty, et al. *Graph theory with applications*, volume 290. Macmillan London, 1976.

- [22] Adam D Bookatz, Pawel Wocjan, and Lorenza Viola. Hamiltonian quantum simulation with bounded-strength controls. *New Journal of Physics*, 16(4):045021, 2014.
- [23] Troy W Borneman and David G Cory. Bandwidth-limited control and ringdown suppression in high-q resonators. *Journal of Magnetic Resonance*, 225:120–129, 2012.
- [24] Troy W Borneman, Christopher E Granade, and David G Cory. Parallel information transfer in a multinode quantum information processor. *Physical Review Letters*, 108(14):140502, 2012.
- [25] Troy W Borneman, Martin D Hürlimann, and David G Cory. Application of optimal control to cpmg refocusing pulse design. *Journal of Magnetic Resonance*, 207(2):220–233, 2010.
- [26] Nicolas Boulant, Timothy F Havel, Marco A Pravia, and David G Cory. Robust method for estimating the lindblad operators of a dissipative quantum process from measurements of the density operator at multiple time points. *Physical Review A*, 67(4):042322, 2003.
- [27] GS Boutis, P Cappellaro, Hyungjoon Cho, C Ramanathan, and DG Cory. Pulse error compensating symmetric magic-echo trains. *Journal of Magnetic Resonance*, 161(2):132–137, 2003.
- [28] Heinz-Peter Breuer, Elsi-Mari Laine, Jyrki Piilo, and Bassano Vacchini. Colloquium: Non-markovian dynamics in open quantum systems. *Reviews of Modern Physics*, 88(2):021002, 2016.
- [29] Heinz-Peter Breuer and Francesco Petruccione. *The theory of open quantum systems*. Oxford University Press, USA, 2002.
- [30] Kenneth R Brown, Aram W Harrow, and Isaac L Chuang. Arbitrarily accurate composite pulse sequences. *Physical Review A—Atomic, Molecular, and Optical Physics*, 70(5):052318, 2004.
- [31] Cameron B Browne, Edward Powley, Daniel Whitehouse, Simon M Lucas, Peter I Cowling, Philipp Rohlfshagen, Stephen Tavener, Diego Perez, Spyridon Samothrakis, and Simon Colton. A survey of monte carlo tree search methods. *IEEE Transactions on Computational Intelligence and AI in games*, 4(1):1–43, 2012.
- [32] Arthur Earl Bryson. *Applied optimal control: optimization, estimation and control*. Routledge, 2018.

- [33] Douglas P Burum. Magnus expansion generator. *Physical Review B*, 24(7):3684, 1981.
- [34] DP Burum and WK Rhim. Analysis of multiple pulse nmr in solids. iii. *The Journal of Chemical Physics*, 71(2):944–956, 1979.
- [35] Jonas Bylander, Simon Gustavsson, Fei Yan, Fumiki Yoshihara, Khalil Harrabi, George Fitch, David G Cory, Yasunobu Nakamura, Jaw-Shen Tsai, and William D Oliver. Noise spectroscopy through dynamical decoupling with a superconducting flux qubit. *Nature Physics*, 7(7):565–570, 2011.
- [36] Paola Cappellaro, Joseph Emerson, Nicolas Boulant, Chandrasekhar Ramanathan, Seth Lloyd, and David G Cory. Entanglement assisted metrology. *Physical review letters*, 94(2):020502, 2005.
- [37] Herman Y Carr and Edward M Purcell. Effects of diffusion on free precession in nuclear magnetic resonance experiments. *Physical review*, 94(3):630, 1954.
- [38] Jiahui Chen, Mohamad Niknam, and David Cory. Efficiently characterizing the quantum information flow, loss, and recovery in the central spin system. *Entropy*, 26(12):1077, 2024.
- [39] Xin-Yu Chen, Na-Na Zhang, Wan-Ting He, Xiang-Yu Kong, Ming-Jie Tao, Fu-Guo Deng, Qing Ai, and Gui-Lu Long. Global correlation and local information flows in controllable non-markovian open quantum dynamics. *npj Quantum Information*, 8(1):22, 2022.
- [40] 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.
- [41] Michael Colarossi, Nathaniel Trask, David P Schmidt, and Mark J Bergander. Multidimensional modeling of condensing two-phase ejector flow. *International journal of refrigeration*, 35(2):290–299, 2012.
- [42] DG Cory, AN Garroway, and JB Miller. Improved resolution of multiple-pulse proton images by oversampling. *Journal of Magnetic Resonance (1969)*, 87(1):202–207, 1990.
- [43] DG Cory, JB Miller, and AN Garroway. Time-suspension multiple-pulse sequences: Applications to solid-state imaging. *Journal of Magnetic Resonance (1969)*, 90(1):205–213, 1990.

- [44] Kristian S Cujia, Konstantin Herb, Jonathan Zopes, John M Abendroth, and Christian L Degen. Parallel detection and spatial mapping of large nuclear spin clusters. *Nature Communications*, 13(1):1260, 2022.
- [45] Thomas Dahm and DJ Scalapino. Theory of intermodulation in a superconducting microstrip resonator. *Journal of applied physics*, 81(4):2002–2009, 1997.
- [46] Yu-Hong Dai and Yaxiang Yuan. A nonlinear conjugate gradient method with a strong global convergence property. *SIAM Journal on optimization*, 10(1):177–182, 1999.
- [47] Domenico d’Alessandro. *Introduction to quantum control and dynamics*. Chapman and hall/CRC, 2021.
- [48] Lawrence Davis et al. Applying adaptive algorithms to epistatic domains. In *IJCAI*, volume 85, pages 162–164. Citeseer, 1985.
- [49] PG De Gennes. Sur la relaxation nucleaire dans les cristaux ioniques. *Journal of Physics and Chemistry of Solids*, 7(4):345–350, 1958.
- [50] Christian L Degen, Friedemann Reinhard, and Paola Cappellaro. Quantum sensing. *Reviews of modern physics*, 89(3):035002, 2017.
- [51] Anatoly E Dementyev, David G Cory, and Chandrasekhar Ramanathan. High-field overhauser dynamic nuclear polarization in silicon below the metal–insulator transition. *The Journal of chemical physics*, 134(15), 2011.
- [52] C Di Franco, Mauro Paternostro, and MS Kim. Nested entangled states for distributed quantum channels. *Physical Review A—Atomic, Molecular, and Optical Physics*, 77(2):020303, 2008.
- [53] C Di Franco, Mauro Paternostro, GM Palma, and MS Kim. Information-flux approach to multiple-spin dynamics. *Physical Review A—Atomic, Molecular, and Optical Physics*, 76(4):042316, 2007.
- [54] Clemens Daska, Kilian Ender, Glen Bigan Mbeng, Andreas Kruckenhauser, Wolfgang Lechner, and Rick van Bijnen. Quantum optimization via four-body rydberg gates. *Physical Review Letters*, 128(12):120503, 2022.
- [55] Jiangfeng Du, Hui Li, Xiaodong Xu, Mingjun Shi, Jihui Wu, Xianyi Zhou, and Rongdian Han. Experimental implementation of the quantum random-walk algorithm. *Physical Review A*, 67(4):042316, 2003.

- [56] Richard W Eglese. Simulated annealing: a tool for operational research. *European journal of operational research*, 46(3):271–281, 1990.
- [57] Joseph Emerson, Robert Alicki, and Karol Życzkowski. Scalable noise estimation with random unitary operators. *Journal of Optics B: Quantum and Semiclassical Optics*, 7(10):S347, 2005.
- [58] Richard R Ernst, Geoffrey Bodenhausen, and Alexander Wokaun. *Principles of nuclear magnetic resonance in one and two dimensions*. Oxford university press, 1990.
- [59] Paolo Facchi, DA Lidar, and Saverio Pascazio. Unification of dynamical decoupling and the quantum zeno effect. *Physical Review A—Atomic, Molecular, and Optical Physics*, 69(3):032314, 2004.
- [60] Edward Farhi and Sam Gutmann. Quantum computation and decision trees. *Physical Review A*, 58(2):915, 1998.
- [61] Richard P Feynman. Simulating physics with computers. In *Feynman and computation*, pages 133–153. cRc Press, 2018.
- [62] Evan M Fortunato, Marco A Pravia, Nicolas Boulant, Grum Teklemariam, Timothy F Havel, and David G Cory. Design of strongly modulating pulses to implement precise effective hamiltonians for quantum information processing. *The Journal of chemical physics*, 116(17):7599–7606, 2002.
- [63] Evan M Fortunato, Lorenza Viola, Marco A Pravia, Emanuel Knill, Raymond Laflamme, Timothy F Havel, and David G Cory. Exploring noiseless subsystems via nuclear magnetic resonance. *Physical Review A*, 67(6):062303, 2003.
- [64] Andrew Gelman, John B Carlin, Hal S Stern, and Donald B Rubin. *Bayesian data analysis*. Chapman and Hall/CRC, 1995.
- [65] Antonio A Gentile, Brian Flynn, Sebastian Knauer, Nathan Wiebe, Stefano Paesani, Christopher E Granade, John G Rarity, Raffaele Santagati, and Anthony Laing. Learning models of quantum systems from experiments. *Nature Physics*, 17(7):837–843, 2021.
- [66] Iulia M Georgescu, Sahel Ashhab, and Franco Nori. Quantum simulation. *Reviews of Modern Physics*, 86(1):153–185, 2014.

- [67] Bernard C Gerstein and Cecil R Dybowski. Transient techniques in nmr of solids: an introduction to theory and practice. 1985.
- [68] Nicolas Gisin and Rob Thew. Quantum communication. *Nature photonics*, 1(3):165–171, 2007.
- [69] Gene H Golub and Charles F Van Loan. *Matrix computations*. JHU press, 2013.
- [70] Thomas Gorin, Tomaž Prosen, Thomas H Seligman, and Marko Žnidarič. Dynamics of loschmidt echoes and fidelity decay. *Physics Reports*, 435(2-5):33–156, 2006.
- [71] Vittorio Gorini, Andrzej Kossakowski, and Ennackal Chandu George Sudarshan. Completely positive dynamical semigroups of n-level systems. *Journal of Mathematical Physics*, 17(5):821–825, 1976.
- [72] Arseni Goussev, Rodolfo A Jalabert, Horacio M Pastawski, and Diego Wisniacki. Loschmidt echo. *arXiv preprint arXiv:1206.6348*, 2012.
- [73] Christopher Granade, Christopher Ferrie, and David G Cory. Accelerated randomized benchmarking. *New Journal of Physics*, 17(1):013042, 2015.
- [74] Christopher E Granade, Christopher Ferrie, Nathan Wiebe, and David G Cory. Robust online hamiltonian learning. *New Journal of Physics*, 14(10):103013, 2012.
- [75] Christian Griesinger, OW Sørensen, and RR Ernst. A practical approach to three-dimensional nmr spectroscopy. *Journal of Magnetic Resonance (1969)*, 73(3):574–579, 1987.
- [76] Christian Griesinger, OW Sorensen, and RR Ernst. Three-dimensional fourier spectroscopy. application to high-resolution nmr. *Journal of Magnetic Resonance (1969)*, 84(1):14–63, 1989.
- [77] David Gross, Koenraad Audenaert, and Jens Eisert. Evenly distributed unitaries: On the structure of unitary designs. *Journal of mathematical physics*, 48(5), 2007.
- [78] Heikki Haario, Eero Saksman, and Johanna Tamminen. An adaptive metropolis algorithm. *Bernoulli*, 7(2):223–242, 2001.
- [79] Holger Haas, Daniel Puzzuoli, Feihao Zhang, and David G Cory. Engineering effective hamiltonians. *New Journal of Physics*, 21(10):103011, 2019.

- [80] Holger Haas, Sahand Tabatabaei, William Rose, Pardis Sahafi, Michèle Piscitelli, Andrew Jordan, Pritam Priyadarsi, Namanish Singh, Ben Yager, Philip J Poole, et al. Nuclear magnetic resonance diffraction with subangstrom precision. *Proceedings of the National Academy of Sciences*, 119(40):e2209213119, 2022.
- [81] Ulrich Haeberlen. *High Resolution NMR in solids selective averaging: supplement 1 advances in magnetic resonance*, volume 1. Elsevier, 2012.
- [82] Ulrich Haeberlen and John S Waugh. Coherent averaging effects in magnetic resonance. *Physical Review*, 175(2):453, 1968.
- [83] Erwin L Hahn. Spin echoes. *Physical review*, 80(4):580, 1950.
- [84] SR Hartmann and EL Hahn. Nuclear double resonance in the rotating frame. *Physical Review*, 128(5):2042, 1962.
- [85] Ian N Hincks, David G Cory, and Chandrasekhar Ramanathan. Equivalent hamiltonians for state to state transfer in the case of partial quantum control. *arXiv preprint arXiv:1111.0944*, 2011.
- [86] IN Hincks, CE Granade, Troy W Borneman, and David G Cory. Controlling quantum devices with nonlinear hardware. *Physical Review Applied*, 4(2):024012, 2015.
- [87] Jonathan S Hodges, Jamie C Yang, Chandrasekhar Ramanathan, and David G Cory. Universal control of nuclear spins via anisotropic hyperfine interactions. *Physical Review A*, 78(1):010303, 2008.
- [88] John H Holland. *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. MIT press, 1992.
- [89] Wayne M Itano, Daniel J Heinzen, John J Bollinger, and David J Wineland. Quantum zeno effect. *Physical Review A*, 41(5):2295, 1990.
- [90] Rodolfo A Jalabert and Horacio M Pastawski. Environment-independent decoherence rate in classically chaotic systems. *Physical review letters*, 86(12):2490, 2001.
- [91] Edwin T Jaynes. *Probability theory: The logic of science*. Cambridge university press, 2003.
- [92] JMBH Jeener, BH Meier, P Bachmann, and Richard R Ernst. Investigation of exchange processes by two-dimensional nmr spectroscopy. *The Journal of chemical physics*, 71(11):4546–4553, 1979.

- [93] Richard Jozsa. Quantum algorithms and the fourier transform. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 454(1969):323–337, 1998.
- [94] Robert E Kass and Adrian E Raftery. Bayes factors. *Journal of the american statistical association*, 90(430):773–795, 1995.
- [95] James Keeler. *Understanding NMR spectroscopy*. John Wiley & Sons, 2010.
- [96] James Keeler and David Neuhaus. Comparison and evaluation of methods for two-dimensional nmr spectra with absorption-mode lineshapes. *Journal of Magnetic Resonance (1969)*, 63(3):454–472, 1985.
- [97] Navin Khaneja, Timo Reiss, Cindie Kehlet, Thomas Schulte-Herbrüggen, and Steffen J Glaser. Optimal control of coupled spin dynamics: design of nmr pulse sequences by gradient ascent algorithms. *Journal of magnetic resonance*, 172(2):296–305, 2005.
- [98] Scott Kirkpatrick, C Daniel Gelatt Jr, and Mario P Vecchi. Optimization by simulated annealing. *science*, 220(4598):671–680, 1983.
- [99] S Klarsfeld and JA Oteo. Recursive generation of higher-order terms in the magnus expansion. *Physical Review A*, 39(7):3270, 1989.
- [100] Emanuel Knill and Raymond Laflamme. Theory of quantum error-correcting codes. *Physical Review A*, 55(2):900, 1997.
- [101] Emanuel Knill, Raymond Laflamme, and Lorenza Viola. Theory of quantum error correction for general noise. *Physical Review Letters*, 84(11):2525, 2000.
- [102] Emanuel Knill, Dietrich Leibfried, Rolf Reichle, Joe Britton, R Brad Blakestad, John D Jost, Chris Langer, Roei Ozeri, Signe Seidelin, and David J Wineland. Randomized benchmarking of quantum gates. *Physical Review A—Atomic, Molecular, and Optical Physics*, 77(1):012307, 2008.
- [103] Levente Kocsis and Csaba Szepesvári. Bandit based monte-carlo planning. In *European conference on machine learning*, pages 282–293. Springer, 2006.
- [104] Levente Kocsis, Csaba Szepesvári, and Jan Willemson. Improved monte-carlo search. *Univ. Tartu, Estonia, Tech. Rep*, 1:1–22, 2006.

- [105] Tamara G Kolda, Robert Michael Lewis, and Virginia Torczon. Optimization by direct search: New perspectives on some classical and modern methods. *SIAM review*, 45(3):385–482, 2003.
- [106] Huub Kooijman, Anthony L Spek, KJC Van Bommel, Willem Verboom, and David N Reinhoudt. A triclinic modification of triphenylphosphine. *Acta Crystallographica Section C: Crystal Structure Communications*, 54(11):1695–1698, 1998.
- [107] Philip Krantz, Morten Kjaergaard, Fei Yan, Terry P Orlando, Simon Gustavsson, and William D Oliver. A quantum engineer’s guide to superconducting qubits. *Applied physics reviews*, 6(2), 2019.
- [108] Rep Kubo. The fluctuation-dissipation theorem. *Reports on progress in physics*, 29(1):255, 1966.
- [109] Ryogo Kubo. Generalized cumulant expansion method. *Journal of the Physical Society of Japan*, 17(7):1100–1120, 1962.
- [110] Ryogo Kubo. Stochastic liouville equations. *Journal of Mathematical Physics*, 4(2):174–183, 1963.
- [111] Ryogo Kubo. A stochastic theory of line shape. *Advances in chemical physics*, 15:101–127, 1969.
- [112] Henning Labuhn, Daniel Barredo, Sylvain Ravets, Sylvain De Léséleuc, Tommaso Macrì, Thierry Lahaye, and Antoine Browaeys. Tunable two-dimensional arrays of single rydberg atoms for realizing quantum ising models. *Nature*, 534(7609):667–670, 2016.
- [113] Jeffrey C Lagarias, James A Reeds, Margaret H Wright, and Paul E Wright. Convergence properties of the nelder–mead simplex method in low dimensions. *SIAM Journal on optimization*, 9(1):112–147, 1998.
- [114] Abdelghani Laraoui, Florian Dolde, Christian Burk, Friedemann Reinhard, Jörg Wrachtrup, and Carlos A Meriles. High-resolution correlation spectroscopy of ^{13}C spins near a nitrogen-vacancy centre in diamond. *Nature communications*, 4(1):1651, 2013.
- [115] Pedro Larranaga, Cindy M. H. Kuijpers, Roberto H. Murga, Inaki Inza, and Sejla Dizdarevic. Genetic algorithms for the travelling salesman problem: A review of representations and operators. *Artificial intelligence review*, 13:129–170, 1999.

- [116] Gregory F Lawler and Vlada Limic. *Random walk: a modern introduction*, volume 123. Cambridge University Press, 2010.
- [117] Malcolm H Levitt. Composite pulses. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 18(2):61–122, 1986.
- [118] MH Levitt, DP Raleigh, F Creuzet, and RG Griffin. Theory and simulations of homonuclear spin pair systems in rotating solids. *The Journal of chemical physics*, 92(11):6347–6364, 1990.
- [119] Jun Li, Xiaodong Yang, Xinhua Peng, and Chang-Pu Sun. Hybrid quantum-classical approach to quantum optimal control. *Physical review letters*, 118(15):150503, 2017.
- [120] Daniel A Lidar, Isaac L Chuang, and K Birgitta Whaley. Decoherence-free subspaces for quantum computation. *Physical Review Letters*, 81(12):2594, 1998.
- [121] Goran Lindblad. On the generators of quantum dynamical semigroups. *Communications in mathematical physics*, 48:119–130, 1976.
- [122] Guang Hao Low, Theodore J Yoder, and Isaac L Chuang. Optimal arbitrarily accurate composite pulse sequences. *Physical Review A*, 89(2):022341, 2014.
- [123] Stephen A Maas. *Nonlinear microwave and RF circuits*. Artech house, 2003.
- [124] Easwar Magesan, Jay M Gambetta, and Joseph Emerson. Scalable and robust randomized benchmarking of quantum processes. *Physical review letters*, 106(18):180504, 2011.
- [125] Easwar Magesan, Jay M Gambetta, and Joseph Emerson. Characterizing quantum gates via randomized benchmarking. *Physical Review A*, 85(4):042311, 2012.
- [126] Easwar Magesan, Daniel Puzzuoli, Christopher E Granade, and David G Cory. Modeling quantum noise for efficient testing of fault-tolerant circuits. *Physical Review A—Atomic, Molecular, and Optical Physics*, 87(1):012324, 2013.
- [127] P Mansfield. Symmetrized pulse sequences in high resolution nmr in solids. *Journal of Physics C: Solid State Physics*, 4(11):1444, 1971.
- [128] M Matti Maricq. Application of average hamiltonian theory to the nmr of solids. *Physical Review B*, 25(11):6622, 1982.

- [129] Michael Mehring. *Principles of high resolution NMR in solids*. Springer Science & Business Media, 2012.
- [130] Michael Mehring and John S Waugh. Phase transients in pulsed nmr spectrometers. *Review of Scientific Instruments*, 43(4):649–653, 1972.
- [131] Saul Meiboom and David Gill. Modified spin-echo method for measuring nuclear relaxation times. *Review of scientific instruments*, 29(8):688–691, 1958.
- [132] Xiao-Li Meng. Posterior predictive p -values. *The annals of statistics*, 22(3):1142–1160, 1994.
- [133] Francesco Mezzadri. How to generate random matrices from the classical compact groups. *arXiv preprint math-ph/0609050*, 2006.
- [134] Maryam Sadat Mirkamali and David G Cory. Mesoscopic spin systems as quantum entanglers. *Physical Review A*, 101(3):032320, 2020.
- [135] Baidyanath Misra and EC George Sudarshan. The zeno’s paradox in quantum theory. *Journal of Mathematical Physics*, 18(4):756–763, 1977.
- [136] Melanie Mitchell. *An introduction to genetic algorithms*. MIT press, 1998.
- [137] Per Christian Moan. *On backward error analysis and Nekhoroshev stability in the numerical analysis of conservative systems of ODEs*. PhD thesis, University of Cambridge, 2002.
- [138] HR Mohebbi, OWB Benningshof, IAJ Taminiau, GX Miao, and DG Cory. Composite arrays of superconducting microstrip line resonators. *Journal of Applied Physics*, 115(9), 2014.
- [139] M Munowitz, A Pines, et al. Principles and applications of multiple-quantum nmr. *Adv. Chem. Phys*, 66:1–152, 1987.
- [140] Michael Munowitz, Alexander Pines, and Michael Mehring. Multiple-quantum dynamics in nmr: A directed walk through liouville space. *The Journal of chemical physics*, 86(6):3172–3182, 1987.
- [141] Sadao Nakajima. On quantum theory of transport phenomena: Steady diffusion. *Progress of Theoretical Physics*, 20(6):948–959, 1958.

- [142] John A Nelder and Roger Mead. A simplex method for function minimization. *The computer journal*, 7(4):308–313, 1965.
- [143] Mohamad Niknam, Lea F Santos, and David G Cory. Sensitivity of quantum information to environment perturbations measured with a nonlocal out-of-time-order correlation function. *Physical Review Research*, 2(1):013200, 2020.
- [144] Mohamad Niknam, Lea F Santos, and David G Cory. Experimental detection of the correlation rényi entropy in the central spin model. *Physical Review Letters*, 127(8):080401, 2021.
- [145] Leigh M Norris, Gerardo A Paz-Silva, and Lorenza Viola. Qubit noise spectroscopy for non-gaussian dephasing environments. *Physical review letters*, 116(15):150503, 2016.
- [146] 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, 2019.
- [147] PJJ O’Malley, J Kelly, R Barends, B Campbell, Y Chen, Z Chen, B Chiaro, A Dunsworth, AG Fowler, I-C Hoi, et al. Qubit metrology of ultralow phase noise using randomized benchmarking. *Physical Review Applied*, 3(4):044009, 2015.
- [148] Gottfried Otting, Hans Widmer, Gerhard Wagner, and Kurt Wüthrich. Origin of t 1 and t 2 ridges in 2d nmr spectra and procedures for suppression. *NMR Struct. Biol*, pages 389–395, 1995.
- [149] Moon-Won Park and Yeong-Dae Kim. A systematic procedure for setting parameters in simulated annealing algorithms. *Computers & Operations Research*, 25(3):207–217, 1998.
- [150] Gerardo A Paz-Silva, Leigh M Norris, and Lorenza Viola. Multiqubit spectroscopy of gaussian quantum noise. *Physical Review A*, 95(2):022121, 2017.
- [151] Gerardo A Paz-Silva and Lorenza Viola. General transfer-function approach to noise filtering in open-loop quantum control. *Physical review letters*, 113(25):250501, 2014.
- [152] 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.

- [153] Pai Peng, Chao Yin, Xiaoyang Huang, Chandrasekhar Ramanathan, and Paola Cappellaro. Floquet prethermalization in dipolar spin chains. *Nature Physics*, 17(4):444–447, 2021.
- [154] A Pines, MG Gibby, and JS Waugh. proton-enhanced nuclear induction spectroscopy: a method for high-resolution nmr of dilute spins in solids. Technical report, Massachusetts Inst. of Tech., Cambridge, 1972.
- [155] F Poggiali, P Cappellaro, and N Fabbri. Optimal control for one-qubit quantum sensing. *Physical Review X*, 8(2):021059, 2018.
- [156] DW Pohl, SE Schwarz, and V Irniger. Forced rayleigh scattering. *Physical Review Letters*, 31(1):32, 1973.
- [157] NV Prokof’Ev and PCE Stamp. Theory of the spin bath. *Reports on Progress in Physics*, 63(4):669, 2000.
- [158] Daniel Puzzuoli, Sophia Fuhui Lin, Moein Malekakhlagh, Emily Pritchett, Benjamin Rosand, and Christopher J Wood. Algorithms for perturbative analysis and simulation of quantum dynamics. *Journal of Computational Physics*, 489:112262, 2023.
- [159] Gregory Quiroz and Daniel A Lidar. Optimized dynamical decoupling via genetic algorithms. *Physical Review A*, 88(5):052306, 2013.
- [160] DP Raleigh, MH Levitt, and Robert G Griffin. Rotational resonance in solid state nmr. *Chemical Physics Letters*, 146(1-2):71–76, 1988.
- [161] D Janaki Ram, TH Sreenivas, and K Ganapathy Subramaniam. Parallel simulated annealing algorithms. *Journal of parallel and distributed computing*, 37(2):207–212, 1996.
- [162] Chandrasekhar Ramanathan, Hyungjoon Cho, Paola Cappellaro, Gregory S Boutis, and David G Cory. Encoding multiple quantum coherences in non-commuting bases. *Chemical physics letters*, 369(3-4):311–317, 2003.
- [163] W-K Rhim, DD Elleman, and RW Vaughan. Analysis of multiple pulse nmr in solids. *The Journal of Chemical Physics*, 59(7):3740–3749, 1973.
- [164] William Rose, Holger Haas, Angela Q Chen, Nari Jeon, Lincoln J Lauhon, David G Cory, and Raffi Budakian. High-resolution nanoscale solid-state nuclear magnetic resonance spectroscopy. *Physical Review X*, 8(1):011030, 2018.

- [165] Jeffrey N Rouder, Paul L Speckman, Dongchu Sun, Richard D Morey, and Geoffrey Iverson. Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic bulletin & review*, 16:225–237, 2009.
- [166] Elodie Salager, Jean-Nicolas Dumez, Lyndon Emsley, and Malcolm H Levitt. A scaling factor theorem for homonuclear dipolar decoupling in solid-state nmr spectroscopy. *Journal of Magnetic Resonance*, 212(1):11–16, 2011.
- [167] CM Sánchez, PR Levstein, RH Acosta, and AK Chattah. Nmr loschmidt echoes as quantifiers of decoherence in interacting spin systems. *Physical Review A—Atomic, Molecular, and Optical Physics*, 80(1):012328, 2009.
- [168] Jack C Saywell, Max S Carey, Philip S Light, Stuart S Szigeti, Alistair R Milne, Karandeep S Gill, Matthew L Goh, Viktor S Perunicic, Nathaniel M Wilson, Calum D Macrae, et al. Enhancing the sensitivity of atom-interferometric inertial sensors using robust control. *Nature Communications*, 14(1):7626, 2023.
- [169] Thomas Schanze. An exact d-dimensional tsallis random number generator for generalized simulated annealing. *Computer physics communications*, 175(11-12):708–712, 2006.
- [170] K Schmidt-Rohr and HW Spiess. Nature of nonexponential loss of correlation above the glass transition investigated by multidimensional nmr. *Physical Review Letters*, 66(23):3020, 1991.
- [171] Klaus Schmidt-Rohr and Hans Wolfgang Spiess. *Multidimensional solid-state NMR and polymers*. Elsevier, 2012.
- [172] Gideon Schwarz. Estimating the dimension of a model. *The annals of statistics*, pages 461–464, 1978.
- [173] Jonathan Richard Shewchuk et al. An introduction to the conjugate gradient method without the agonizing pain, 1994.
- [174] Peter W Shor. Scheme for reducing decoherence in quantum computer memory. *Physical review A*, 52(4):R2493, 1995.
- [175] David Silver, Aja Huang, Chris J Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, et al. Mastering the game of go with deep neural networks and tree search. *nature*, 529(7587):484–489, 2016.

- [176] Charles P Slichter. *Principles of magnetic resonance*, volume 1. Springer Science & Business Media, 2013.
- [177] Aaron Sodickson and David G Cory. A generalized k-space formalism for treating the spatial aspects of a variety of nmr experiments. *Progress in nuclear magnetic resonance spectroscopy*, 33(2):77–108, 1998.
- [178] S Somaroo, CH Tseng, TF Havel, Raymond Laflamme, and David G Cory. Quantum simulations on a quantum computer. *Physical review letters*, 82(26):5381, 1999.
- [179] Alexandre M Souza, Gonzalo A Alvarez, and Dieter Suter. Robust dynamical decoupling for quantum computing and quantum memory. *Physical review letters*, 106(24):240501, 2011.
- [180] Andrew Stasiuk, Pai Peng, Garrett Heller, and Paola Cappellaro. Frame change technique for phase transient cancellation. *Journal of Magnetic Resonance*, 362:107688, 2024.
- [181] DJ States, RA Haberkorn, and DJ Ruben. A two-dimensional nuclear overhauser experiment with pure absorption phase in four quadrants. *Journal of Magnetic Resonance (1969)*, 48(2):286–292, 1982.
- [182] EO Stejskal and Jacob Schaefer. Removal of artifacts from cross-polarization nmr experiments. *Journal of Magnetic Resonance (1969)*, 18(3):560–563, 1975.
- [183] Youngkyu Sung, Félix Beaudoin, Leigh M Norris, Fei Yan, David K Kim, Jack Y Qiu, Uwe von Lüpke, Jonilyn L Yoder, Terry P Orlando, Simon Gustavsson, et al. Non-gaussian noise spectroscopy with a superconducting qubit sensor. *Nature communications*, 10(1):3715, 2019.
- [184] Gilbert Syswerda. A study of reproduction in generational and steady-state genetic algorithms. In *Foundations of genetic algorithms*, volume 1, pages 94–101. Elsevier, 1991.
- [185] K Takegoshi and CA McDowell. A “magic echo” pulse sequence for the high-resolution nmr spectra of abundant spins in solids. *Chemical physics letters*, 116(2-3):100–104, 1985.
- [186] Yoshitaka Tanimura. Stochastic liouville, langevin, fokker–planck, and master equation approaches to quantum dissipative systems. *Journal of the Physical Society of Japan*, 75(8):082001, 2006.

- [187] Yoshitaka Tanimura and Ryogo Kubo. Time evolution of a quantum system in contact with a nearly gaussian-markoffian noise bath. *Journal of the Physical Society of Japan*, 58(1):101–114, 1989.
- [188] John E Tanner. Use of the stimulated echo in nmr diffusion studies. *The Journal of Chemical Physics*, 52(5):2523–2526, 1970.
- [189] Barbara M Terhal. Quantum error correction for quantum memories. *Reviews of Modern Physics*, 87(2):307–346, 2015.
- [190] Géza Tóth, Witłef Wieczorek, David Gross, Roland Krischek, Christian Schwemmer, and Harald Weinfurter. Permutationally invariant quantum tomography. *Physical review letters*, 105(25):250403, 2010.
- [191] Constantino Tsallis and Daniel A Stariolo. Generalized simulated annealing. *Physica A: Statistical Mechanics and its Applications*, 233(1-2):395–406, 1996.
- [192] CH Tseng, S Somaroo, Y Sharf, Emanuel Knill, Raymond Laflamme, Timothy F Havel, and David G Cory. Quantum simulation of a three-body-interaction hamiltonian on an nmr quantum computer. *Physical Review A*, 61(1):012302, 1999.
- [193] Matthew Tyler, Hengyun Zhou, Leigh S Martin, Nathaniel Leitaó, and Mikhail D Lukin. Higher-order methods for hamiltonian engineering pulse sequence design. *Physical Review A*, 108(6):062602, 2023.
- [194] Götz S Uhrig. Keeping a quantum bit alive by optimized π -pulse sequences. *Physical Review Letters*, 98(10):100504, 2007.
- [195] Oriol Vinyals, Igor Babuschkin, Wojciech M Czarnecki, Michaël Mathieu, Andrew Dudzik, Junyoung Chung, David H Choi, Richard Powell, Timo Ewalds, Petko Georgiev, et al. Grandmaster level in starcraft ii using multi-agent reinforcement learning. *Nature*, 575(7782):350–354, 2019.
- [196] Lorenza Viola, Emanuel Knill, and Seth Lloyd. Dynamical decoupling of open quantum systems. *Physical Review Letters*, 82(12):2417, 1999.
- [197] Pauli Virtanen, Ralf Gommers, Travis E Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, et al. Scipy 1.0: fundamental algorithms for scientific computing in python. *Nature methods*, 17(3):261–272, 2020.

- [198] Quoc L Vuong, Jean-François Berret, Jérôme Fresnais, Yves Gossuin, and Olivier Sandre. A universal scaling law to predict the efficiency of magnetic nanoparticles as mri t2-contrast agents. *Advanced healthcare materials*, 1(4):502–512, 2012.
- [199] P W. Anderson. A mathematical model for the narrowing of spectral lines by exchange or motion. *Journal of the Physical Society of Japan*, 9(3):316–339, 1954.
- [200] Joel J Wallman, Marie Barnhill, and Joseph Emerson. Robust characterization of loss rates. *Physical Review Letters*, 115(6):060501, 2015.
- [201] Jianwei Wang, Stefano Paesani, Raffaele Santagati, Sebastian Knauer, Antonio A Gentile, Nathan Wiebe, Maurangelo Petruzzella, Jeremy L O’Brien, John G Rarity, Anthony Laing, et al. Experimental quantum hamiltonian learning. *Nature Physics*, 13(6):551–555, 2017.
- [202] Peiyong Wang, Muhammad Usman, Udaya Parampalli, Lloyd CL Hollenberg, and Casey R Myers. Automated quantum circuit design with nested monte carlo tree search. *IEEE Transactions on Quantum Engineering*, 2023.
- [203] Warren S Warren. Effects of arbitrary laser or nmr pulse shapes on population inversion and coherence. *The Journal of chemical physics*, 81(12):5437–5448, 1984.
- [204] WS Warren, DP Weitekamp, and A Pines. Theory of selective excitation of multiple-quantum transitions. *The Journal of Chemical Physics*, 73(5):2084–2099, 1980.
- [205] Larry Wasserman. Bayesian model selection and model averaging. *Journal of mathematical psychology*, 44(1):92–107, 2000.
- [206] John S Waugh, Lee M Huber, and Ulrich Haeberlen. Approach to high-resolution nmr in solids. *Physical Review Letters*, 20(5):180, 1968.
- [207] Ken Xuan Wei, Chandrasekhar Ramanathan, and Paola Cappellaro. Exploring localization in nuclear spin chains. *Physical review letters*, 120(7):070501, 2018.
- [208] Yaakov S Weinstein, Timothy F Havel, Joseph Emerson, Nicolas Boulant, Marcos Saraceno, Seth Lloyd, and David G Cory. Quantum process tomography of the quantum fourier transform. *The Journal of chemical physics*, 121(13):6117–6133, 2004.
- [209] Yaakov S Weinstein, MA Pravia, EM Fortunato, Seth Lloyd, and David G Cory. Implementation of the quantum fourier transform. *Physical review letters*, 86(9):1889, 2001.

- [210] Darrell Whitley. A genetic algorithm tutorial. *Statistics and computing*, 4:65–85, 1994.
- [211] Nathan Wiebe, Christopher Granade, Christopher Ferrie, and David G Cory. Hamiltonian learning and certification using quantum resources. *Physical review letters*, 112(19):190501, 2014.
- [212] Stephen Wimperis. Broadband, narrowband, and passband composite pulses for use in advanced nmr experiments. *Journal of Magnetic Resonance, Series A*, 109(2):221–231, 1994.
- [213] Christopher J Wood, Troy W Borneman, and David G Cory. Cavity cooling of an ensemble spin system. *Physical review letters*, 112(5):050501, 2014.
- [214] William K Wootters and Brian D Fields. Optimal state-determination by mutually unbiased measurements. *Annals of Physics*, 191(2):363–381, 1989.
- [215] Yilun Xu, Gang Huang, Jan Balewski, Ravi Naik, Alexis Morvan, Bradley Mitchell, Kasra Nowrouzi, David I Santiago, and Irfan Siddiqi. Qubic: An open-source fpga-based control and measurement system for superconducting quantum information processors. *IEEE Transactions on Quantum Engineering*, 2:1–11, 2021.
- [216] Fei Yan, Simon Gustavsson, Jonas Bylander, Xiaoyue Jin, Fumiki Yoshihara, David G Cory, Yasunobu Nakamura, Terry P Orlando, and William D Oliver. Rotating-frame relaxation as a noise spectrum analyser of a superconducting qubit undergoing driven evolution. *Nature communications*, 4(1):2337, 2013.
- [217] Emil A Yuzbashyan, Boris L Altshuler, Vadim B Kuznetsov, and Victor Z Enolskii. Solution for the dynamics of the bcs and central spin problems. *Journal of Physics A: Mathematical and General*, 38(36):7831, 2005.
- [218] Feihao Zhang, Jian Xing, Xiaoxiao Hu, Xinyu Pan, and Guilu Long. Coupling-selective quantum optimal control in weak-coupling nv-13 c system. *AAPPS Bulletin*, 33(1):2, 2023.
- [219] Wenxian Zhang, NP Konstantinidis, VV Dobrovitski, BN Harmon, Lea F Santos, and Lorenza Viola. Long-time electron spin storage via dynamical suppression of hyperfine-induced decoherence in a quantum dot. *Physical Review B—Condensed Matter and Materials Physics*, 77(12):125336, 2008.

- [220] Wurong Zhang and DG Cory. First direct measurement of the spin diffusion rate in a homogenous solid. *Physical review letters*, 80(6):1324, 1998.
- [221] Hengyun Zhou, Leigh S Martin, Matthew Tyler, Oksana Makarova, Nathaniel Leita, 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.
- [222] Robert Zwanzig. Ensemble method in the theory of irreversibility. *The Journal of Chemical Physics*, 33(5):1338–1341, 1960.