

THE UNIVERSITY OF CHICAGO

CROSS-STACK OPTIMIZATION OF QUANTUM TECHNOLOGIES INSPIRED BY
LOW-LEVEL PHYSICAL PRINCIPLES

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE PHYSICAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF COMPUTER SCIENCE

BY
LENNART MAXIMILIAN SEIFERT

CHICAGO, ILLINOIS

JUNE 2025

ABSTRACT

Over the past few decades various quantum technologies have been proposed and it is their potential impact on the world that is making many researchers eagerly pursue their realization. This endeavor comes with new challenges that we do not encounter in classical computers for instance, as a fundamentally different set of principles determined by quantum mechanics governs the behavior of these novel systems. Quantum information is very delicate and prone to distortion by the environment, therefore the control of such devices needs to be carefully designed with the underlying physical properties in mind. Furthermore, in the case of quantum computing, algorithms and compilation techniques can make use of hardware noise characteristics to improve the quality of results. While eventually we would like to build a computing stack where these high-level design choices can be agnostic to such low-level details, this will not be possible in the near future as complete fault-tolerance is a major research goal that the community regards as a long-term mission.

In this thesis we explore how the low-level understanding of a quantum system aids the optimization of different layers of its device stack. In particular, one major contribution lies in studying the design, improvement and implementation of control pulses that make a system undergo desired operations — in the case of quantum computers (our main focus), this corresponds to the realization of gates that compose programs. However, we further explore quantum sensors as another technology that is promising to have practical use. Here control pulses determine the sensing scheme and the sensor sensitivity, which one seeks to maximize. Pulse optimization covers a lower layer in the quantum stack we study first. We then turn to the top layer and discuss work that improves a particular quantum algorithm, the Variational Quantum Eigensolver, by including efficient hardware noise modeling in its execution. Overall, we demonstrate the impact of design choices for quantum technologies that are inspired by the physics governing the underlying systems.

To Mama and Papa,
for all your love and developing my passion for science,
and to my brother Marc,
for making Marx the greatest duo this world has ever seen.

TABLE OF CONTENTS

LIST OF FIGURES	vii
LIST OF TABLES	xiv
ACKNOWLEDGMENTS	xv
1 INTRODUCTION	1
2 THE QUANTUM TECHNOLOGY STACK	7
2.1 Fundamentals of quantum mechanics	8
2.1.1 From classical to quantum	8
2.1.2 Dirac notation	11
2.1.3 Example: Quantum harmonic oscillator	14
2.2 Quantum computing	16
2.2.1 Quantum bits	16
2.2.2 Quantum gates and measurement	19
2.2.3 Quantum circuits	22
2.2.4 Quantum compilation	24
2.2.5 Clifford gates and circuits	25
2.3 Control of quantum hardware	26
2.3.1 Quantum optimal control	30
2.3.2 Example: The transmon processor	32
2.4 The distorting impact of noise	35
2.4.1 Common sources of errors	37
2.4.2 Quantum error mitigation and error correction	38
3 QUANTUM OPTIMAL CONTROL FOR QUDIT COMPUTATION	40
3.1 Generalized qudit gates	40
3.2 Time-efficient qudit gates through Incremental Pulse Re-seeding	41
3.2.1 Quantum optimal control setup	42
3.2.2 Gate duration optimization	47
3.2.3 Qudit pulse results	56
3.2.4 Conclusion	59
3.3 Exploring ququart computation on a transmon using optimal control	61
3.3.1 Hardware specifications	62
3.3.2 One ququart or two qubits	65
3.3.3 Numerical gate decomposition	67
3.3.4 Quantum optimal control	70
3.3.5 Experimental results	77
3.3.6 Conclusion	85

4	QUANTUM OPTIMAL CONTROL FOR QUANTUM SENSING	87
4.1	Sensing hardware	89
4.1.1	Optical lattice interferometry	89
4.1.2	Direct trajectory optimization	93
4.2	Programming an optical lattice interferometer	97
4.2.1	Gate design	98
4.2.2	Conclusion	107
4.3	End-to-end optimization of interferometer sensitivity	110
4.3.1	Sensing accelerations	111
4.3.2	Fisher information optimization	114
4.3.3	Conclusion	120
5	CONTROL OF A NOISE-PROTECTED QUBIT SYSTEM	122
5.1	Mitigating frequency noise on the physical level	123
5.2	Kerr-cat qubits	124
5.2.1	Protection versus controllability	126
5.3	Frequency-noise-insensitive universal control of Kerr-cat qubits	128
5.3.1	$X(\theta)$ rotations	129
5.3.2	Two-qubit gate	131
5.3.3	$Y(\theta)$ rotations	133
5.3.4	$Z(\theta)$ rotations	139
5.3.5	Extension to time-dependent errors	146
5.3.6	Conclusion	149
6	APPLICATION-TAILORED ERROR MITIGATION	151
6.1	Variational quantum algorithms	151
6.1.1	Variational Quantum Eigensolver	152
6.1.2	CAFQA	153
6.2	Clapton: Clifford-assisted problem transformation for VQA error mitigation	153
6.2.1	Clapton proposal	156
6.2.2	Clapton design	159
6.2.3	Methodology	164
6.2.4	Evaluation	168
6.2.5	Related work	176
6.2.6	Conclusion	177
7	CONCLUSION AND FUTURE DIRECTIONS	178
	REFERENCES	182

LIST OF FIGURES

1.1	Simplified depiction of the stack of a quantum computer as an example of a quantum technology. The low levels consist of the quantum hardware and the control system, which can implement physical operations. A quantum error correction (QEC) scheme translates between physical and logical instructions and is able to dynamically correct errors (lightning symbol). QEC has not been demonstrated in full universality yet, thus the dashed boundary lines. On top we find the high-level programs that are compiled to fundamental operations. Without perfect QEC, errors can propagate all the way up and distort the computation result.	2
2.1	Truncated spectrum of the quantum harmonic oscillator. The black line corresponds to the quadratic potential and the wavefunctions ψ_n of the lowest eigenstates are visualized at their respective energies E_n	16
2.2	Energy spectra showing the encoding of qubit states in larger quantum systems. (a) In neutral atoms such as ^{87}Rb , two hyperfine levels can be used to represent the qubit. (b) In superconducting architectures like transmons, the qubit is typically encoded in the two lowest-energy eigenstates. The colored lines show the probability distributions $\psi_n(\phi)^2$ of the eigenstates $ n\rangle$. Section 2.3.2 includes more background on transmons specifically.	17
2.3	A qubit state $ \psi\rangle$ is represented by a point on the Bloch sphere (red). The poles correspond to the computational basis states $ 0\rangle$ and $ 1\rangle$. Single-qubit gates (see Section 2.2.2) correspond to rotations of the Bloch sphere about specific axes $\hat{n}$ (blue).	18
2.4	Conceptual depiction of a quantum circuit. A sequence of unitary gates U_i is applied to the qubit register that is typically initialized in the state $ 0\rangle^{\otimes N}$. At the end the qubits are measured in the Z basis.	23
2.5	Quantum circuit decomposition of the three-qubit Toffoli gate into single- and two-qubit gates.	24
2.6	A Josephson junction (left) is shunted by a large capacitor (left) to form a transmon.	33
3.1	Real part of an exemplary pulse in the rotating frame, with one carrier wave ($N_f = 1$). (a) Six B-splines shape the envelope of the pulse. Grey dots mark the support of each time-local basis function. (b) The final pulse is constructed from a single carrier wave bounded by the B-spline envelope. With $N_f > 1$, the final pulse is the sum of these individual pulses with one carrier wave.	44

3.2	The Incremental Pulse Re-seeding process. (a) The problem is initialized with a random pulse and duration guess T_{start} . (b) The main optimization function, which searches for the highest-fidelity pulse for a given duration using Juqbox, returning an optimized pulse γ and a fidelity $\mathcal{F}(T; \gamma)$. (c) If an optimization succeeds, save the duration and pulse, then decrease T and start from a truncated version of the successful pulse. (d) If an optimization fails, but a successful pulse has been found in the past, decrease step size and work backwards from T_{best} again; if step size has reached granularity, return best pulse. (e) If no successful pulse has been found so far, but fidelity improved, mark this duration as the best so far, extend the duration, and re-seed with the extended previous pulse. (f) If no successful pulse has been found and fidelity decreased, restart the whole process with the highest-fidelity duration and a random guess.	50
3.3	Applying Incremental Pulse Re-seeding (IPR) to find the shortest duration for the H_4 gate. (a) Visualization of attempted durations during the IPR procedure, with the chronological order of steps indicated by the colored numbers. Green (red) arrows represent steps that did (did not) lead to meeting the pulse target fidelity of $\mathcal{F}_{\text{targ}} = 99.9\%$. Starting from $T_{\text{start}} = 70$ ns and a time step of 8 ns, three pulse extension and re-optimization steps are required before the target fidelity is first reached at 94 ns. The duration is then reduced to 78 ns by successfully re-seeding with truncated pulses. After step 6 the time step is halved because re-optimizing at $T = 70$ ns with the truncated pulse does not meet the target fidelity. Repeating this scheme leads to the final duration of $T_{\text{best}} = 76$ ns, highlighted by the star, as after step 10 the time step falls below the granularity of 1 ns. (b) Plot A shows the optimized pulse after IPR step 2. A new seed for the optimization at $T = 94$ ns is generated by extending the previous result in duration (A*). At 94 ns the target fidelity is met, so the optimized pulse is truncated (B*) and used as a seed for the second $T = 84$ ns optimization, leading to the successful result shown in C.	53
3.4	Evolution of state populations when applying the H_4 pulse of shortest duration (76 ns) to a qudit with $d = 4$. Every computational basis state is transformed into superposition of all four basis states with nearly equal population, deviations only arising from the remaining 0.1% infidelity. At all times the populations of the guard states $ 4\rangle$ and $ 5\rangle$ are below $2 \cdot 10^{-3}$ and 10^{-6} , respectively.	54
3.5	Distribution of naive random-guess optimizations of the X_8 gate for individual durations between 180 ns and 230 ns. Each boxplot consists of 20 fidelities from 20 individual Juqbox optimizations, with the blue line indicating the median value, the box indicating the middle 50% of fidelities, and the whiskers indicating the full range of values. (a) The full data, showing a general increase in fidelity as duration increases. (b) Focusing on the area of interest (fidelities close to 99.9%). The IPR scheme outperforms random guessing, finding a minimum duration of $T_{\text{best}} = 195$ ns at 99.9% fidelity, whereas none of the 20 random guesses at that duration reached the same fidelity.	55

3.6	Dependence of single-qudit gate durations from qudit dimension d . The four gates are defined in Equation (3.2). Each data point shows the minimum duration of applying Incremental Pulse Re-seeding (IPR) from ten different starting times. The unexpected discrepancy between $d = 6$ and $d = 7$ for the generalized T_d gate suggests a highly complicated optimization space and shows that shorter pulses for some d values may be possible. For each gate the scaling of duration with dimension d is analyzed with quadratic and linear regressions, where the latter is less accurate but captures the essential behavior in this low-dimensional regime. The fit coefficients are presented in Table 3.3.	57
3.7	Gate durations of two-qudit SWAP operations for qudit dimensions $d = 2, 3, 4$. The horizontal axis corresponds to the dimension of the full Hilbert space d^2 . Due to the low number of data points, regression is not applied, but a near-linear trend is visible.	59
3.8	False-colored optical image of the device. The left transmon (red) is used as the ququart in the experiment. Single-ququart rotations are sent through the input line coupled to the resonator (blue). The SQUID coupler (purple, zoomed-in region) is kept biased to the flux sweet spot throughout the experiment.	63
3.9	Measurement setup.	64
3.10	Two ways of realizing arbitrary quantum gates on a four-level ququart, which is controlled by three modulated carrier waves with resonant frequencies ω_j . (a) Any ququart gate can be decomposed into a series of at most six qubit-like $Y_{j-1,j}(\theta_i)$ rotations interleaved with ququart phase gates. Given a target unitary, the parameters θ_i and ϕ_i can be calculated numerically. (b) Lifting the gate abstraction, quantum optimal control can be used to determine a control signal that operates on all subspaces in parallel and manipulate the entire ququart state. A final virtual $Z_4(\phi)$ gate adjusts the relative phases. The shown pulses together with $\phi = (4.265, 4.590, 0.000)$ (radians) implement a ququart Hadamard H_4 in $T = 350$ ns.	72
3.11	Optimal control pulses for ququart Hadamard H_4 , with final VZ correction (solid lines, $T = 350$ ns) and without final VZ correction (dashed lines, $T = 420$ ns). The vertical gray line marks the duration of the former.	75
3.12	Optimal control pulses for the tensor product $H \otimes H$ without a final VZ correction, $T = 380$ ns. Allowing a correction did not lead to a significantly reduced gate duration.	77
3.13	Readout characterization. (a) Single-shot histogram after phase rotation and (b) heatmap of detection probabilities (in %). Single-ququart basis states are prepared and measured 5000 times.	78
3.14	Process matrices for quantum optimal control implementations of the (a) ququart H_4 and (b) two-qubit $H \otimes H$. The area of each square encodes the magnitude of the corresponding element while the color encodes its phase. The H_4 process matrix is represented in the generalized Pauli basis while $H \otimes H$ is represented in the Pauli basis for two qubits. The exponent/index $n \in \{0, 1, 2, 3\}$ assumes increasing values from left to right or top to bottom for each tick in a label group.	81

3.15	Experimental results for (a) ququart RB (C_4 Clifford group, red) and (b) two-qubit RB ($C_2^{\otimes 2}$ Clifford group, red) on a four-level transmon. Using IRB we benchmark (a) H_4 and (b) $H \otimes H$ in the appropriate setting, for both their gate-level decomposition (DEC, yellow) and their quantum optimal control (QOC, blue) implementation. Markers represent the mean survival probabilities and error bars (smaller than the marker size) are the standard deviations of these means.	83
4.1	Band structure of the optical lattice Hamiltonian (see Equation (4.1)) for $\varphi \equiv 0$. The potential $V(x)$ for the specific choice $V_0 = 10\omega_R$ is shown in black, while modulating φ corresponds to shifting the potential to the left or right. The Bloch states $\psi_\nu(x) = \langle x \nu\rangle_B$ are shown at their corresponding energy eigenvalue, clearly defining a valence and a conduction band subspaces, with latter characterized by nearly degenerate pairs of opposite-parity states. A subset of the operations considered in this work are indicated by red arrows (see also Section 4.2.1 and Figure 4.2).	91
4.2	Visualization of a variety of operations that can be realized in the optical lattice interferometer. Neighboring pairs of Bloch states are identified as qubit subspaces and represented by separate Bloch spheres. Upon initialization the system is in $ 0\rangle_B$ in the valence band (Π_1 qubit). SPLIT $_\nu$ maps to a Bloch state $ \nu\rangle_B$ in the conduction band, here shown for $\nu = 4$. Unitary BOOST $_\nu$ / SLOW $_\nu$ operations map between adjacent conduction band qubits, for instance BOOST $_4 : \Pi_4 \rightarrow \hat{\Pi}_6$ and SLOW $_6 : \Pi_6 \rightarrow \Pi_4$. The mapping $\Pi_4 \rightarrow \Pi_1$ corresponds to bringing the atoms to rest (HOLD operation, inverse is RELEASE). Single-qubit operations on the conduction band qubits are visualized; we provide interpretations in terms of atom-cloud trajectories in the main text. The yellow circles in the conduction band Bloch spheres correspond to approximate momentum eigenstates $ \pm\nu k_L\rangle_p$	99
4.3	SPLIT $_3 : 0\rangle_B \rightarrow 3\rangle_B$ operation, $T = 1.88\omega_R^{-1}$, $\mathcal{I} = 3.27 \cdot 10^{-5}$. The top panel shows the state evolution in the Bloch basis while the lower panel shows the corresponding control function.	100
4.4	Realization of logical qubit gates on the Π_4 qubit. Each column shows the state evolution of one basis state, the control function as well as its Fourier transform. The desired frequency cutoff at $\omega_c = 70\omega_R$ is achieved through an additional term in the objective and enables smooth waveforms. The gate properties (duration T , infidelity $\mathcal{I}$) are $(1.50\omega_R^{-1}, 4.05 \cdot 10^{-3})$ for X_4 , $(3.00\omega_R^{-1}, 2.13 \cdot 10^{-3})$ for Z_4 , $(3.50\omega_R^{-1}, 9.39 \cdot 10^{-5})$ for H_4 and $(2.50\omega_R^{-1}, 3.61 \cdot 10^{-4})$ for T_4	102
4.5	Z_6 gate acting on qubit Π_6 , $T = 4.50\omega_R^{-1}$, $\mathcal{I} = 2.90 \cdot 10^{-3}$. It is equivalent to a MIRROR operation for atoms in $ \pm 6k_L\rangle_p$. The top panel shows the state evolution in the Bloch basis while the lower panel shows the corresponding control function.	106
4.6	Control functions for the HOLD operation (left, $\mathcal{I} = 1.37 \cdot 10^{-3}$) and the BOOST $_4$ operation (right, $\mathcal{I} = 9.13 \cdot 10^{-5}$). For both operations the gate time is $T = 1.75\omega_r^{-1}$	108
4.7	Space-time representation of a simple circuit for a lattice-hold interferometer sequence written using the gates proposed in this work.	109

4.8	Optimization of the accelerometer's sensitivity in an end-to-end fashion. The red vertical lines indicate the insertion of free-propagation windows of duration $2.0 \cdot 2\pi\omega_R^{-1} \approx 12.6 \omega_R^{-1}$ in order to mimic the structure of a Mach-Zehnder interferometer. (a) Evolution of the Bloch states throughout the protocol. (b) Evolution of the (classical) Fisher information, normalized to the Fisher information of the idealized Mach-Zehnder interferometer $\mathcal{I}_{\text{ideal}}(a = 0; T)$ shown in Equation (4.48).	120
5.1	(a) In frequency-tunable transmons (inset), uncertainties in the flux $\Phi_{x,\text{err}}$ lead to unknown frequency shifts Δ (orange arrows around green point) when operated away from sweet spots (pink stars). (b) Shifts Δ break the degeneracy of the KC qubit (orange circle), causing phase errors. The gap opening is exponentially suppressed with the cat size α^2 , creating robustness to Δ .	125
5.2	(a) Bloch sphere of the KC qubit showing Wigner functions of the X and Z eigenstates as well as rotations about the axes X (easy, green), Y and Z (hard, red), and an arbitrary axis in the X - Y plane $\hat{n}_\varphi$ (blue). (b) Dependence of the magnitude of the drive matrix element h_φ on the angle φ . As the cat size and protection increase, rotations about all axes in the X - Y plane except for the X axis get suppressed.	126
5.3	Pulse implementation and infidelity robustness for the $X(\pi/2)$ gate. (a) The $X(\theta)$ rotation is native to the KC and is realized through a simple single-photon drive pulse. (b) The infidelity of the pulse (a) when simulated over a range of detuning errors Δ . (c) The average infidelity $\bar{\mathcal{I}}$ of pulses of type (a) for different durations and cat sizes. The infidelity decreases when the cat size is increased.	131
5.4	(a) Pulse scheme for the Y rotation, which includes tuning the two-photon the pump strength $\tilde{\varepsilon}_2$ to lower the cat size and a DRAG term ε_x . (b) The infidelity of the pulse (a) when simulated over a range of detuning errors Δ . (c) The average infidelity $\bar{\mathcal{I}}$ of pulses of type (a) for different durations and cat sizes. The increase in robustness diminishes when increasing the cat size beyond $\alpha^2 \gtrsim 2$.	134
5.5	Optimization results of the $Y(\pi/2)$ gate for different settings, showing average infidelity $\bar{\mathcal{I}}$, drive amplitude $\varepsilon_{y,0}$ and relative cat size $\eta = \alpha^2(T/2)/\alpha^2$ at the most ramped-down point. (a) Gate scheme including the DRAG correction ε_x . This is the main result, which is also included in Figure 5.4. The cat size is barely ramped down to retain the KC protection, however this makes large drive amplitudes necessary, which can become a limiting factor. (b) When constraining the drive amplitude, the cat size has to be ramped increasingly more for larger initial α^2 in order to counteract the suppressed h_y element. Sacrificing the KC protection limits the achievable average infidelity. (c) Without the implementation of a DRAG correction, the coupling to excited states combined with the suppressed h_y element leads to a Stark shift and ultimately to a loss in fidelity, hence the found solution is to ramp the cat size down all the way to zero. No robustness improvement over the bare qubit ($\alpha^2 = 0$) is achieved.	137

5.6	Adiabatic realization of the $Z(-\pi/2)$ gate using the scheme $\textcircled{1} \leftrightarrow \textcircled{2}$. (a) The energy gap derivative $\partial_\delta E_{01}(\delta, \alpha^2)$ as a function of the detuning δ and cat size α^2 . The robust line $\delta_{\text{rob}}(\alpha^2)$ is highlighted by a black dashed line. An example gate trajectory is shown with an orange solid line. The inset (same legend as (d)) shows the robust line in the energy spectrum. (b) Control pulse for the solid trajectory shown in (a), which simultaneously tunes the detuning δ and the two-photon pump strength via $\tilde{\varepsilon}_2$. (c) The infidelity of the pulse (b) when simulated over a range of detuning errors Δ . (d) The average infidelity $\bar{\mathcal{I}}$ of pulses of type (b) for different durations and cat sizes. We note that the $\alpha^2 = 1$ line ends at $T \approx 22 K^{-1}$ because our scheme $\textcircled{1} \leftrightarrow \textcircled{2}$ is not applicable in the case of small cat sizes ($\alpha^2 \lesssim 1$), which is also why $\alpha^2 = 0$ is excluded here. We discuss an alternative scheme $\textcircled{1} \leftrightarrow \textcircled{2'}$ below, which works for all cat sizes but is less suited for dynamical errors $\Delta(t)$	140
5.7	Alternative realization of the $Z(-\pi/2)$ gate using the $\textcircled{1} \leftrightarrow \textcircled{2'}$ scheme. (a) The energy gap derivative $\partial_\delta E_{01}(\delta, \alpha^2)$ as a function of the detuning δ and cat size α^2 . The robust line $\delta_{\text{rob}}(\alpha^2)$ (highlighted by a black dashed line) separates the parameter space into two regions with different signs of the energy gap derivative $\partial_\delta E_{01}(\delta, \alpha^2)$, which enables first-order insensitivity to static errors Δ on averaging over the trajectory. An example trajectory is drawn with an orange solid line. The inset shows the energy spectrum. (b) Control pulse for the solid trajectory shown in (a), which simultaneously tunes the detuning δ and the two-photon pump strength via $\tilde{\varepsilon}_2$. (c) The infidelity of the pulse (b) when simulated over a range of detuning errors Δ . (d) The average infidelity $\bar{\mathcal{I}}$ of pulses of type (b) for different durations and cat sizes.	144
6.1	VQAs combine a classical optimizer which adjusts the parameters in each iteration with a quantum device producing function evaluations by running the corresponding quantum circuits.	152
6.2	The fundamental idea behind Clapton: Transforming a VQA problem structure to match the noise structure on a quantum device. Having identified a good algorithmic state, transforming the VQA problem such that it is mapped to a good hardware state improves the fidelity under noisy evaluation, leading to improved performance of the entire VQA instance.	156
6.3	Key result: Clifford operations enable the combination of classically simulable noise and VQA Hamiltonian transformation, which ultimately gives Clapton its advantage over other initialization schemes in terms of solution quality (lower energy) and modeling accuracy.	157

6.4	The Clapton transformation engine. Given a problem Hamiltonian $\mathcal{H}$, a noise-equipped version of the VQE ansatz (at the zero point) $\tilde{A}(\mathbf{0})$ and a Clifford transformation ansatz $C(\boldsymbol{\gamma})$, Clapton employs genetic algorithms (GAs) to search over transformation parameters $\boldsymbol{\gamma}$ to find a solution $\hat{\boldsymbol{\gamma}}$. It instantiates $\mathbf{s}$ processes $\mathbf{GA}_1, \dots, \mathbf{GA}_s$, starting from random populations $S_1, \dots, S_s$, that search the parameter space independently. Internally, each $\mathbf{GA}$ instance performs $\mathbf{m}$ iterations of transformation optimization, where for different $\boldsymbol{\gamma}$ the Hamiltonian is transformed and the noisy ($\mathcal{L}_{\mathcal{N}}$) as well as noiseless ($\mathcal{L}_0$) energy of the initial point is evaluated. The GAs produce final populations S_i^* , of which Clapton takes the top $\mathbf{k}$ solutions, mixes them and randomly constructs new starting populations S_i , additionally adding new random guesses that are not derived from the previous round. This entire scheme is repeated until the global loss value does not decrease between rounds (convergence).	162
6.5	Comparing the VQE initialization quality of Clapton against the baselines CAFQA and nCAFQA, as measured by achieved energy and relative improvement η , across a variety of benchmarks (color-coded). Results are shown for three noise models (nairobi , toronto , mumbai) and one real quantum device (hanoi). Top rows: Energy values of found solutions for each method on an absolute, normalized scale (see main text). Different symbols refer to the evaluation of each solution in different noise environments. Energies under full model/device evaluation ($\times$) are the most relevant. The dark $\times$ refers to the energy after a couple hundred VQE iterations (final point (f)) when starting from the colored $\times$ (initial point (i)). Bottom rows: Relative improvements of Clapton over CAFQA and nCAFQA, evaluated for the initial $\times$ and the final $\times$. The inset text shows the geometric mean $\bar{\eta}$ across all benchmarks.	169
6.6	Visualizations of the VQE convergence of the ten-qubit XXZ model with (a) $J = 0.25$ and (b) $J = 1.00$, run on two different noise models. The Clapton initialization starts better than the baselines and typically converges better. Initial and final evaluations on real machine hanoi are shown with stars. Hardware experiments yield worse final points in some cases, suggesting a device-model discrepancy that Clapton does not seem to be strongly affected by.	172
6.7	The impact of Clapton when solely sweeping the single-qubit gate error p (two-qubit gate error $10p$), for three different decay times and a selection of benchmarks. Relative improvement is evaluated at the initial VQE point.	173
6.8	The impact of Clapton when solely sweeping the measurement error p , for three different decay times and a selection of benchmarks. Relative improvement is evaluated at the initial VQE point.	174
6.9	Scaling of Clapton optimization time with the number of qubits N for the Ising model ($J = 0.25$). Each instance is run from five initial guesses using $\mathbf{s} = 10$. The inset presents the scaling of the average optimization time per round, which follows an $\mathcal{O}(N^2)$ trend. The baseline CAFQA method shows a considerably faster performance as it only computes noiseless evaluations.	175

LIST OF TABLES

3.1	Important parameters for optimization in Jupyter	45
3.2	Variables used in Incremental Pulse Re-seeding	49
3.3	Linear and quadratic fit coefficients for single-qudit gate durations	57
3.4	Ququart characterization. Relaxation times T_1 and Ramsey dephasing times T_R between neighboring energy levels $ j-1\rangle \leftrightarrow j\rangle$ ($j = 1, 2, 3$) are listed, as well as transition frequencies ω_j and on-chip drive responses r_j to unit power AWG driving.	63
3.5	Numerically obtained decomposition parameters for several gates according to Equation (3.11), together with sequence execution durations T computed using Equation (3.10). All angles are measured in radians. We recognize decimal approximations of parameters as multiples of $\frac{\pi}{2}$ or special analytical expressions which we show in the lower table.	71
3.6	Sets of pre-rotation ($\mathcal{S}$) and post-rotation ($\mathcal{T}$) gates to implement quantum state tomography and quantum process tomography.	79
3.7	Single-ququart gate fidelity characterized by QPT, for decomposed gates (DEC) and optimally controlled gates (QOC). In the QOC implementation of $H \otimes H$ (*) we only considered the case without a final virtual Z (VZ).	81
3.8	Single-ququart gate fidelity characterized by different types of RB experiments, for decomposed gates (DEC) and optimally controlled gates (QOC). The H_4 gate fidelity is characterized using the Clifford group $\mathcal{C}_4$, other gates are characterized using the Clifford group $\mathcal{C}_2^{\otimes 2}$. Base fidelity for RB with $\mathcal{C}_4$ is 96.22(14)%, and with $\mathcal{C}_2^{\otimes 2}$ is 95.84(5)%. We only consider the VZ implementation for the QOC H_4 pulse.	84
3.9	Ququart gate infidelities extracted from simulation. DEC: decomposition, QOC: quantum optimal control.	85

ACKNOWLEDGMENTS

I want to express my utmost gratitude to my advisor Fred Chong, who supported me through this PhD journey with excellent guidance and constant encouragement. He always had faith in me, my capabilities and my work, especially during times when I struggled to see it myself. I am thankful for the opportunities he provided me to grow as a researcher, like the many conferences I was able to attend and present at as well as the numerous fellow researchers he connected me with. Furthermore, our one-on-one meetings would also often include chats not related to work, and I appreciate many bits of advice on life and emotional support when I needed it.

I want to thank my committee members, Bill Fefferman and Liang Jiang, for their time and feedback on my thesis. Their expertise and curiosity created exam situations that were both challenging and enjoyable, with insightful conversations after each presentation.

Special thanks have to be extended to my external collaborators who were instrumental in the success of many projects included in this thesis. Thank you to Ziqian Li, who showed me the challenges when experimenting with quantum hardware, to Michael Perlin and Victor Colussi, my mentors during my internship at Inflection who taught me about atom interferometry and the exciting field that is quantum sensing, and to Connor Hann and Kyungjoo Noh, who introduced me to the world of cat qubits when I interned at the AWS Center for Quantum Computing. I am looking forward to being reunited with their team after graduation.

I am incredibly grateful to have been part of a research lab that was not only highly productive and collaborative, but also felt like family. We worked on many exciting projects together, went on fun conference trips and shared many laughs. I am happy that we all became such good friends and I could not have made it this far without your support, encouragement and kindness. While all of you are very dear to me, I would like to mention a few people in particular. Thank you to Siddharth Dangwal, whose personality and humor

aligns with mine in so many ways, which lead to an abundance of funny and meaningful memories, to Christopher Kang, for constant support with work and life and logical reasoning one cannot argue against, to Jason Chadwick, who I wrote my first paper with and biked through all of Chicago, to Willers Yang, who I shared many adventures with and who would always build me up when I was down, to Jonathan Baker and Andrew Litteken, who mentored me early on and made the start to the PhD so easy, and to Gokul Ravi, whose excitement for Clifford gates can be contagious and is only surpassed by his enthusiasm and motivational skills.

Outside the immediate work environment, I am thankful to have made many more friends that made these four years in Chicago so special. Important mentions here are Geneva and Trevor Anderberg, the most energetic and positive couple I know and grew very close with, thank you for many Frisbee games, paddles on the lake, countless barbecues and including me in your wedding party. Thank you to Ashley Uphoff and Adrian Lehmann, my German refuge in the United States who helped me with many troubles and were always encouraging, to Maria Vinokurskaya, who, together with Jason, was always fun to hang out, play games and have deep conversations with, and to Carys Obertone, Ethan Hyland and Charles Greenspon, for holding me accountable at the gym and being the best workout buddies I could have wished for.

I want to extend my gratitude to all my friends back home in Germany, whose support made me dare this step of pursuing a PhD abroad in the first place. I appreciate your friendship throughout all those years, and I want to mention Anna Böhme, Fabian Scheurer, Isaac Justice, Annina Missikewitsch and Martin Grotz in particular.

Thank you to my family, who has always had my back in all decisions and has been believing in me from the day I was born. Thank you, Mama and Papa, for your love and constant encouragement in all stages of my life and teaching me all the values I hold so dear. I could not imagine better role models than you. And finally, thank you to my little brother

Marc, who is as proud of me as I am of him, for the way we support each other and always make the greatest team.

Thank you all for being part of this incredible journey. I am excited about the next chapter and I am looking forward to sharing it with you.

CHAPTER 1

INTRODUCTION

In the last few years, the development of quantum technologies has seen tremendous progress. One large endeavor pursued by scientists around the globe is the design and engineering of a fault-tolerant quantum computer, which is expected to introduce a new revolutionary paradigm of computing. The trajectory of achievements is promising: The number of quantum bits (*qubits*) on hardware and their quality has been steadily increasing across different quantum computing platforms, and researchers have been designing more reliable ways to encode and manipulate the fragile quantum information.

However, further improvements to various parts of the quantum computing stack are needed to reach a point of practical utility for these machines. The simplified version of the stack design is shown in Figure 1.1. Starting at the bottom, these challenges include the fabrication of high-fidelity qubits, which requires a deeper understanding of material properties, a careful choice of physical device parameters (especially in the case of artificial qubits like transmons) and layout, and better shielding from interaction with the environment. On the next level, where we wish to purposefully manipulate the quantum hardware to perform computations, researchers are concerned with optimizing the control system that implements logic gates and performs the readout of quantum information. Work at this layer is still largely platform-specific as different architectures like superconducting qubits, atomic qubits or semiconductor qubits come with their own set of control knobs as well as unique features and limitations. As an example, superconducting transmons [99] are operated with microwave pulses, while neutral atoms [201] are mostly controlled with laser beams.

It is intuitive that both these layers bear many sources of imperfections that can degrade the performance of the quantum computer. Quantum devices are operated at extreme conditions (cryogenic temperature, ultra-high vacuum, etc.) and any distortion typically leads to unwanted corruption of the quantum information. On the control side, the pulses that drive

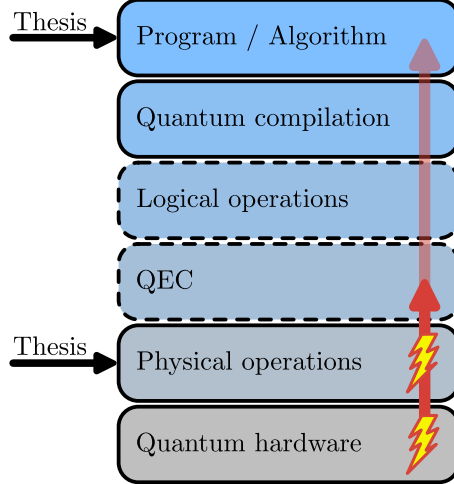


Figure 1.1: Simplified depiction of the stack of a quantum computer as an example of a quantum technology. The low levels consist of the quantum hardware and the control system, which can implement physical operations. A quantum error correction (QEC) scheme translates between physical and logical instructions and is able to dynamically correct errors (lightning symbol). QEC has not been demonstrated in full universality yet, thus the dashed boundary lines. On top we find the high-level programs that are compiled to fundamental operations. Without perfect QEC, errors can propagate all the way up and distort the computation result.

the qubits themselves can have errors stemming from the control electronics or interact with the quantum system in a way that is not intended (e.g., leakage outside the computational space). Each factor can be limiting to the qubit coherence and impose a fundamental barrier on the capability of a device. Generally, these issues are regarded collectively and referred to as *noise* affecting the hardware. A major goal by every experimentalist is to reduce the noise in their system as much as possible.

In an ideal world without noise, the layers that follow this low-level part of the stack can be more and more agnostic to these hardware details. A specific platform offers a set of native gates depending on the available control knobs and further defines a certain qubit topology, both being properties a quantum compiler has to know about. For instance, in a neutral atom architecture, a multi-qubit gate can be achieved with a single, native operation that exploits the so-called *Rydberg blockade* [183], which is not possible on a transmon processor as it lacks this quantum feature and also typically only allows nearest-neighbor connectivity.

Above that, on the program or algorithm level, no knowledge of the underlying quantum system should be necessary, allowing for the universal design of applications that can be executed on any quantum computer.

However, in reality, noise is a serious concern because these low-level distortions lead to errors on the top level, limiting the fidelity of algorithms. While an obvious effort is to reduce the noise at the source, much research has also been dedicated to higher layers in the stack to reduce the impact of errors. For instance, a variety of error mitigation techniques have been proposed, like noise-aware compilation schemes [129] or zero-noise extrapolation [178]. While these methods have proven helpful, in recent years, the community has come to the realization that better hardware engineering and error mitigation will not be sufficient to build machines that can tackle large-scale problems, the error rates will not be sufficiently low enough. This is why the field of quantum error correction (QEC) has been experiencing a large growth in attention and major breakthroughs have been achieved [21]. The idea is to redundantly encode quantum information in a larger quantum system in order to detect and correct errors dynamically. This would define intermediate layers in the stack as shown in Figure 1.1. QEC offers a systematic path to achieve the long-term goal of completely fault-tolerant quantum computing, however, its redundancy overhead is critically determined by the quality and the control of the hardware components, which is why improving the low-level stack is still of utmost importance. Furthermore, until then the high-level algorithm side will still be exposed to hardware noise, therefore tailoring application programs to deal with such distortions constitutes a worthwhile research direction.

Taking inspirations from these observations for quantum computers, this thesis addresses the space of optimizations across the device stacks of different quantum technologies. We demonstrate how to use the knowledge about the physics that governs these quantum technologies to improve their performance and inform architectural design choices at different layers of the stack. Our focus lies on optimizations that can be achieved through theory,

simulations and software. While quantum computing platforms serve as the main example for the major part of this thesis, another promising technology we consider are quantum sensors. The key idea behind quantum sensing is making use of quantum effects to detect physical properties that are not accessible to us with classical tools, for instance, magnetic fields [159] or spatial accelerations [30] at a very fine scale. We describe how the concept of a stack is translated to this technology and how it can benefit from various physics-inspired optimizations as well.

Before we dive into the technical research content of this thesis, we first give a more detailed introduction to the quantum technology stack and the formalism of quantum physics in Chapter 2. We discuss the concepts of wavefunctions and state vectors describing quantum systems as well as their evolution and controlled manipulation, which is relevant in different parts of the thesis. Especially the latter, the control of quantum systems and the room for optimization it provides, represents a key aspect of our research. Over the course of the following three chapters, we show that improvements at this low-level pulse layer enable positive implications on the device stack’s quality for the cases of three different quantum architectures.

First, in Chapter 3, we explore how pulse control enables hardware-specific operations that go beyond what we would intuitively deem necessary to realize standard quantum computation. Specifically, many quantum systems allow access to higher energy levels, enabling the implementation of d -dimensional qudits. In the past qudits have been investigated as architectural components on the gate level with promising results [6, 7]. In order to assess their feasibility in reality, we present hardware-aware studies where we numerically design control pulses using quantum optimal control [199] and further benchmark them in experiment on a quantum device.

Second, we apply the tools from quantum optimal control to a sensing technology based on neutral atoms. We describe in Chapter 4 how we design control sequences that enable a

“programmable sensing” framework, which includes proposing gate primitives and a language similar to quantum computing, thus defining higher-level stack layers specifically for this sensing platform. The specific pulse implementation of a sensing scheme influences the sensitivity it can sense the desired parameters with. Pushing this sensitivity to the limit is a key goal when constructing such a device, which is why we furthermore present an end-to-end pulse optimization scheme that directly boosts the sensor’s sensitivity.

Third, we turn our attention back to quantum computing in Chapter 5 and consider a qubit system with an inherent noise protection. This is particularly appealing in the context of quantum error correction as it could reduce the redundancy overhead that is usually required. In particular, we focus on Kerr-cat qubits which are robust to undesired phase flip errors. However, this gain in error protection comes at the cost of reduced controllability, which poses a significant obstacle for their practical use. In order to address this challenge, we study the specific instance of operating a Kerr-cat qubit encoded in superconducting hardware that is subject to unknown frequency fluctuations. We demonstrate pulse schemes that circumvent this tradeoff, thereby recovering universal computation while maintaining robustness to induced phase flips.

These chapters extensively cover the low-level part of the technology stack. In Chapter 6, we move up the stack to target the algorithm level, where we leverage knowledge of hardware noise characteristics to inform quantum programs and increase their fidelity. This adaption is especially useful in the transitional period from current noisy hardware to fully fault-tolerant machines. In particular, we focus on an algorithm that lends itself to this endeavor in an intuitive way: the Variational Quantum Eigensolver. Its purpose is to find the ground state of a (modeled) quantum system. We investigate how a specific ground state problem can be transformed given the knowledge of hardware noise to yield a more noise-resilient formulation. As a result we propose a pre-processing error mitigation technique that exploits the structure of hardware noise and evaluate its performance in simulation as well

as experiment.

To reiterate, in this thesis we present a variety of physics-inspired optimizations that can be applied to different layers of the quantum technology stack, which is also highlighted in Figure 1.1. We discuss multiple architectures and show how their unique features enable device-specific improvements.

CHAPTER 2

THE QUANTUM TECHNOLOGY STACK

For a *classical* computer, like our phone, laptop or smart TV, the concept of its *stack* is well established. It describes the different layers that define the technology, from the bare hardware (the “silicon”) at the bottom to complex software applications at the top. The space in between is filled by a number of layers that add an increasing amount of abstraction to bridge the gap between top and bottom; these layers include the basic instruction set, compilers and the operating system for instance.

As *quantum* technologies emerge and develop, this concept can be extended to them as well and we already introduced the adaption for a quantum computer with Figure 1.1 in the previous chapter. We find the underlying quantum hardware at the bottom, which is controlled with classical electronics that implements physical operations. The quantum effects coming to life inside such a device are highly delicate and prone to noise, limiting a quantum device’s capability. In order to build a powerful quantum technology able to tackle practical problems, this limitation needs to be overcome. For quantum computing specifically, great efforts are being devoted to the development of *quantum error mitigation* and *quantum error correction* (QEC) schemes. A QEC code defines logical operations as compositions of physical operations, which then form the fundamental set of instructions. At the top, a quantum compiler takes a high-level algorithm and makes it compatible with the underlying layers by translating it to basic instructions.

In this chapter, we go over the different layers of the stack in more detail and introduce physical concepts, terminology and example platforms that are of interest in our works.

2.1 Fundamentals of quantum mechanics

At the heart of all quantum technologies lie the principles of quantum mechanics, which enable their unique features. Therefore, in this section, we start with an introduction to the basic concepts of quantum mechanics that are relevant throughout this thesis. For more in-depth material, we refer to standard textbooks like Refs. [167, 71, 136].

2.1.1 From classical to quantum

In classical mechanics (we refer to Ref. [60] for a detailed introduction), the state of a system is fully determined by a set of *spatial coordinates* $\mathbf{x} = [x_1 \dots x_d]^\top$ together with corresponding *conjugate momenta* $\mathbf{p} = [p_1 \dots p_d]^\top$. Collectively, one refers to the $2d$ -tuple $(\mathbf{x}, \mathbf{p})$ as *canonical coordinates*. While there exist many such coordinate systems that are equivalent in the sense that they can be transformed into each other (keyword: *canonical transformation*), in this thesis it is sufficient to consider the common case that $\mathbf{x}$ represents Cartesian coordinates; also, it simplifies this brief discussion. As an example, the state of a system of two particles in three dimensions could be described by the canonical coordinates $[x_1 \ x_2 \ x_3 \ x_4 \ x_5 \ x_6 \ p_1 \ p_2 \ p_3 \ p_4 \ p_5 \ p_6]^\top$, where the indices 1 – 3 refer to the first particle and 4 – 6 to the second.

The system evolves in time according to Hamilton’s equations of motion:

$$\dot{\mathbf{x}}(t) = \frac{\partial \mathcal{H}(\mathbf{x}(t), \mathbf{p}(t), t)}{\partial \mathbf{p}}, \quad \dot{\mathbf{p}}(t) = -\frac{\partial \mathcal{H}(\mathbf{x}(t), \mathbf{p}(t), t)}{\partial \mathbf{x}}. \quad (2.1)$$

The function $\mathcal{H} : (\mathbf{x}, \mathbf{p}, t) \mapsto \mathcal{H}(\mathbf{x}, \mathbf{p}, t)$ is called the *Hamiltonian* and often takes the form $\mathcal{H} = T + V$, where T is the *kinetic energy* and V the *potential energy* (or simply, *potential*) of the system, thus representing the total energy of the system. For instance, the Hamiltonian of a harmonic oscillator (mass m , eigenfrequency ω_0) subject to an external force $F(t)$ in

one dimension is given by

$$\mathcal{H}(x, p, t) = \frac{p^2}{2m} + \frac{1}{2}m\omega_0^2 x^2 - F(t)x, \quad (2.2)$$

which leads to equations of motion that are familiar from Newtonian mechanics:

$$\dot{x}(t) = \frac{p(t)}{m}, \quad \dot{p}(t) = -m\omega_0^2 x(t) + F(t). \quad (2.3)$$

However, certain phenomena cannot be explained with the laws of classical mechanics, for instance, the behavior of electrons inside atoms. The necessary transition to quantum mechanics is achieved via *canonical quantization*, where the canonical coordinates x_k and p_k are replaced with Hermitian operators (i.e., linear maps) $\hat{x}_k : \mathcal{V} \rightarrow \mathcal{V}$ and $\hat{p}_k : \mathcal{V} \rightarrow \mathcal{V}$ in the Hamiltonian $\mathcal{H}$, giving rise to an Hamiltonian operator $\hat{\mathcal{H}} : \mathcal{V} \rightarrow \mathcal{V}$ describing the energy of the quantum system:

$$\mathcal{H}(\mathbf{x}, \mathbf{p}, t) \rightarrow \hat{\mathcal{H}}(\hat{\mathbf{x}}, \hat{\mathbf{p}}, t). \quad (2.4)$$

These operators act on elements $\psi \in \mathcal{V}$ of a vector space of functions $\mathcal{V} \subset \{f : \mathbb{R}^d \rightarrow \mathbb{C}\}$:

$$\hat{x}_k \psi(\mathbf{x}) = x_k \psi(\mathbf{x}), \quad \hat{p}_k \psi(\mathbf{x}) = -i\hbar \frac{\partial \psi(\mathbf{x})}{\partial x_k}. \quad (2.5)$$

The *reduced Planck constant* $\hbar \approx 1.054 \cdot 10^{-34}$ Js is a fundamental, scale-defining constant in quantum mechanics; however, it is common to work in units where $\hbar = 1$. This has the effect that energy and frequency are measured in the same units. We set the symbol in lightgray in this section and it does not appear in the rest of this thesis.

The complex-valued function ψ is called the quantum system's *wavefunction* and it fully describes its state, similar to the canonical coordinates in the classical case. For example, in the case of N electrons inside an atom, the wavefunction ψ with $d = 3N$ describes their global state in three dimensions. While $\psi(\mathbf{x})$ itself has no clear interpretation, its squared absolute

value $|\psi(\mathbf{x})|^2$ can be understood as a probability density, i.e., $|\psi(\mathbf{x})|^2 d^d \mathbf{x}$ corresponds to the probability of *observing* the system in a small volume $d^d \mathbf{x}$ around $\mathbf{x}$. Therefore, we require a physically meaningful wavefunction to be normalized:

$$\int |\psi(\mathbf{x})|^2 d^d \mathbf{x} = \|\psi\|_{\mathcal{L}^2}^2 = 1. \quad (2.6)$$

This means that wavefunctions are elements of the Hilbert space $\mathcal{L}^2(\mathbb{R}^d, \mathbb{C})$, which is the vector space of square-integrable functions (i.e., functions with finite norm) that map from $\mathbb{R}^d$ (here the coordinates $\mathbf{x}$) to $\mathbb{C}$. The space of all wavefunctions (including unnormalized ones) is thus a subspace $\mathcal{V} \subset \mathcal{L}^2(\mathbb{R}^d, \mathbb{C})$ and furthermore inherits the standard inner product from $\mathcal{L}^2(\mathbb{R}^d, \mathbb{C})$,

$$\langle \cdot | \cdot \rangle : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{C}, \quad \langle \phi | \psi \rangle = \int \phi^*(\mathbf{x}) \psi(\mathbf{x}) d^d \mathbf{x}, \quad (2.7)$$

making $\mathcal{V}$ a Hilbert space itself.

The evolution of the quantum state in time leads to a smoothly parameterized family of wavefunctions $\psi(t) \in \mathcal{V}$, where the dynamics is determined by the famous *time-dependent Schrödinger equation* (TDSE)

$$i\hbar \dot{\psi}(\mathbf{x}, t) = \hat{\mathcal{H}}(\hat{\mathbf{x}}, \hat{\mathbf{p}}, t) \psi(\mathbf{x}, t). \quad (2.8)$$

Applied to our example of a simple harmonic oscillator in Equation (2.2), the TDSE governing a driven quantum harmonic oscillator reads

$$i\hbar \dot{\psi}(x, t) = \left(\frac{\hat{p}^2}{2m} + \frac{1}{2} m \omega_0^2 \hat{x}^2 - F(t) \hat{x} \right) \psi(x, t) = \left(-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + \frac{1}{2} m \omega_0^2 x^2 - F(t)x \right) \psi(x, t). \quad (2.9)$$

This is a partial linear differential equation of first order in time and second order in space, which is the common case as the kinetic energy is typically quadratic in $\mathbf{p}$.

2.1.2 Dirac notation

We transition to a more general formulation of quantum mechanics using a notation that was first introduced by Paul Dirac in 1939 [47] and is now ubiquitously used by quantum scientists. It enables the description of quantum effects that cannot be modeled with wave-functions, for instance, the spin of an electron.

The language of quantum mechanics is linear algebra. The *quantum state* of a system is represented by an element (called *ket*) $|\psi\rangle \in \mathcal{V}$ of a complex Hilbert space $\mathcal{V}$ that is equipped with an inner product $\langle \cdot | \cdot \rangle : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{C}$, and can therefore be expanded in an orthonormal basis $\{|\psi_k\rangle\}_k$ ($\langle \psi_k | \psi_l \rangle = \delta_{kl}$),

$$|\psi\rangle = \sum_k \alpha_k |\psi_k\rangle, \quad (2.10)$$

with complex coefficients $\alpha_k = \langle \psi_k | \psi \rangle \in \mathbb{C}$. The coefficient α_k determine the probability $p_k = |\alpha_k|^2$ of observing the system in the state $|\psi_k\rangle$, which implies the normalization constraint $\sum_k |\alpha_k|^2 = 1$. Due to the vector space structure, and as Equation (2.10) shows, a quantum system can exist in a *superposition* of basis states. This is one of the unique properties of quantum mechanics that are essential to the functionality of a quantum technology, which becomes clearer in the next section.

Another crucial feature is *entanglement*, which can emerge between multiple subsystems with Hilbert spaces $\mathcal{V}_j$ of a composite quantum system with Hilbert space $\mathcal{V} = \bigotimes_j \mathcal{V}_j$. If a state $|\psi\rangle \in \mathcal{V}$ can be factorized, i.e., written in the form $|\psi\rangle = \bigotimes_j |\psi^j\rangle$ with $|\psi^j\rangle \in \mathcal{V}_j$, it is called a *product state*. Otherwise, the state is called *entangled*, which means that properties of the subsystems cannot be described independently of each other. For example, the state

$$|\psi\rangle = \frac{|\psi_1^1\rangle \otimes |\psi_1^2\rangle + |\psi_2^1\rangle \otimes |\psi_2^2\rangle}{\sqrt{2}} \equiv \frac{|\psi_1^1 \psi_1^2\rangle + |\psi_2^1 \psi_2^2\rangle}{\sqrt{2}}, \quad (2.11)$$

where $\{|\psi_k^j\rangle\}_k$ denotes an orthonormal basis of $\mathcal{V}_j$, is entangled. Here we also introduced the shorthand notation for product states, e.g., $|\psi_1^1 \psi_1^2\rangle \equiv |\psi_1^1\rangle \otimes |\psi_1^2\rangle$. Upon measurement,

we either find both subsystems in their first basis states $|\psi_1^1\psi_1^2\rangle$ or their second basis states $|\psi_2^1\psi_2^2\rangle$, with a 50% probability each, thus the individual observations are correlated. Entanglement is a form of correlation that has no classical counterpart and cannot be explained by local properties of the subsystems (keywords: local hidden variable model, Bell inequality [10]), making it the key ingredient that quantum technologies aim to exploit.

Mathematically, the quantum state $|\psi\rangle$ can often be thought of as a (possibly infinitely dimensional) column vector of complex numbers. Its dual (called *bra*) $\langle\psi|$ is identified by the adjoint, $\langle\psi| \cong |\psi\rangle^\dagger$, which then corresponds to the complex-conjugated and transposed row vector. In that case the inner product $\langle\cdot|\cdot\rangle$ is given by the familiar standard inner product of two complex vectors, which is captured in the notation: $\langle\phi|\psi\rangle = \langle\phi| |\psi\rangle$ for $\phi, \psi \in \mathcal{V}$.

Information about a quantum system is extracted by measuring an *observable* O , which is an Hermitian operator ($O = O^\dagger$) with real eigenvalues o_k and corresponding eigenstates $|o_k\rangle$, so $O|o_k\rangle = o_k|o_k\rangle$. We say that we “measure O ” or “measure in the O basis”. Upon measurement, the outcome o_k is obtained with probability $p(o_k) = |\langle o_k|\psi\rangle|^2$ and the system collapses to the eigenstate $|o_k\rangle$, which may lead to the loss of entanglement correlations to other systems. The expectation value of the observable is then given by $\langle O \rangle = \langle\psi|O|\psi\rangle$. As an example, we can draw the connections to wavefunctions: Let $\{|\mathbf{x}\rangle\}_{\mathbf{x}}$ be the set of eigenstates of the d -dimensional position operator $\hat{\mathbf{x}}$, defining the *position basis*, with eigenvalues $\{\mathbf{x}\}_{\mathbf{x}}$. The projection of the quantum state onto this basis corresponds to the wavefunction $\psi(\mathbf{x}) = \langle\mathbf{x}|\psi\rangle$ and relates to the probability density¹ $|\psi(\mathbf{x})|^2$ of the outcome $\mathbf{x}$, as discussed before.

To reiterate, the probability of any physical observation is determined by the squared absolute value $|\langle o_k|\psi\rangle|^2$ of its corresponding state component $\langle o_k|\psi\rangle$. This has the direct consequence that we are unable to differentiate between a state $|\psi\rangle$ and a state $|\psi'\rangle = e^{i\phi}|\psi\rangle$ with $\phi \in [0, 2\pi)$, therefore, the *global phase* of a quantum state is unobservable. Thus, we can

1. Due to the uncountably infinite dimension of the position basis $\{|\mathbf{x}\rangle\}$, the notion of a probability density is required.

associate with $|\psi\rangle$ a family of equivalent states $\left\{e^{i\phi}|\psi\rangle \mid \phi \in [0, 2\pi)\right\}$, but work in derivations with a particular representative $|\psi\rangle$ with a fixed global phase without loss of generality. This freedom is in stark contrast to the relevance of *relative phases* between different basis states in a superposition, for instance,

$$\frac{|\psi_1\rangle + |\psi_2\rangle - |\psi_3\rangle}{\sqrt{3}} \quad \text{versus} \quad \frac{|\psi_1\rangle - i|\psi_2\rangle + |\psi_3\rangle}{\sqrt{3}}, \quad (2.12)$$

because they determine the constructive or destructive interference of basis states as the system evolves.

A particularly interesting observable is the Hamiltonian $\hat{\mathcal{H}}$, which describes the system's energy. Its eigenvalue problem is known as the *time-independent Schrödinger equation* (TISE)

$$\hat{\mathcal{H}}|\psi_k\rangle = E_k|\psi_k\rangle, \quad (2.13)$$

which defines the *eigenenergies* E_k and eigenstates $|\psi_k\rangle$ of the system. It is often useful to choose the energy eigenbasis to express arbitrary states (see Equation (2.10)). The Hamiltonian determines the evolution of the system's state $|\psi(t)\rangle$ in time, which is governed by the *time-dependent Schrödinger equation* (TDSE)

$$i|\dot{\psi}(t)\rangle = \hat{\mathcal{H}}(t)|\psi(t)\rangle, \quad (2.14)$$

which generalizes Equation (2.8) to arbitrary quantum systems. We included the possibility that the Hamiltonian can be time-dependent, which is a fundamental requirement for building a quantum device because the system needs to be controllable. The formal solution of the TDSE is given by

$$|\psi(t)\rangle = \underbrace{\mathcal{T} \exp\left(-i \int_0^t \hat{\mathcal{H}}(t') dt'\right)}_{\equiv \hat{U}(t)} |\psi(t_0)\rangle = \hat{U}(t) |\psi(0)\rangle. \quad (2.15)$$

Here, $\mathcal{T}$ denotes the time-ordering operator, which is required because the Hamiltonian generally does not commute with itself at different times, i.e., $[\hat{\mathcal{H}}(t), \hat{\mathcal{H}}(t')] \neq 0$. Effectively, the dynamics is described by the unitary *time evolution operator* $\hat{U}(t) = \left(\hat{U}(t)^\dagger\right)^{-1}$ that evolves any input state at time t_0 to its corresponding output state at time t . By purposefully engineering the Hamiltonian $\hat{\mathcal{H}}(t)$, different such unitary transformations on the quantum system can be realized, enabling the practical use of a quantum technology as we discuss in Section 2.2.

2.1.3 Example: Quantum harmonic oscillator

To conclude the introduction of basic concepts, we briefly describe the algebraic solution of the one-dimensional quantum harmonic oscillator (QHO) as it is a fundamental model in quantum mechanics and we encounter it multiple times throughout this thesis.

We recall the Hamiltonian from Equation (2.9), where we consider the case of no driving ($F(t) = 0$) for now:

$$\hat{\mathcal{H}} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega_0^2\hat{x}^2 \quad (2.16)$$

By introducing the new operators

$$\begin{aligned} \hat{a} &= \sqrt{\frac{m\omega_0}{2}}\hat{x} + i\frac{1}{\sqrt{2m\omega_0}}\hat{p}, & \hat{a}^\dagger &= \sqrt{\frac{m\omega_0}{2}}\hat{x} - i\frac{1}{\sqrt{2m\omega_0}}\hat{p} \\ \longrightarrow \hat{x} &= \frac{1}{\sqrt{2m\omega_0}}(\hat{a} + \hat{a}^\dagger), & \hat{p} &= -i\sqrt{\frac{m\omega_0}{2}}(\hat{a} - \hat{a}^\dagger) \end{aligned} \quad (2.17)$$

the Hamiltonian can be rewritten as (using the commutation relation $[\hat{x}, \hat{p}] = i$)

$$\hat{\mathcal{H}} = \omega_0\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right) = \omega_0\left(\hat{n} + \frac{1}{2}\right), \quad (2.18)$$

The solution of the TISE (2.13) of the QHO is determined by the *number operator* $\hat{n} = \hat{a}^\dagger\hat{a}$,

which can be shown to fulfill

$$\hat{n} |n\rangle = n |n\rangle \quad (2.19)$$

with eigenvalues $n \in \mathbb{N}$. Therefore, the energies $E_n = \omega_0 \left(n + \frac{1}{2}\right)$ of the QHO are equally spaced with gap size ω_0 .

The operators $\hat{a}$ and $\hat{a}^\dagger$ are called *annihilation* and *creation operator*, respectively, as they lower or raise the energy by one quanta:

$$\hat{a} |n\rangle = \sqrt{n} |n-1\rangle, \quad \hat{a}^\dagger |n\rangle = \sqrt{n+1} |n+1\rangle. \quad (2.20)$$

They enable moving up and down the ladder of eigenstates, thus they are also referred to as *ladder operators*. Starting from the ground state $|0\rangle$ with property $\hat{a} |0\rangle = 0$, the spectrum is constructed by repeated application of $\hat{a}^\dagger$. This way the corresponding wavefunctions $\psi_n(x) = \langle x | n \rangle$ can be determined:

$$\begin{aligned} \left(\sqrt{\frac{m\omega_0}{2}} x + \frac{1}{\sqrt{2m\omega_0}} \frac{\partial}{\partial x} \right) \psi_0(x) = 0 &\longrightarrow \psi_0(x) = \left(\frac{m\omega_0}{\pi} \right)^{\frac{1}{4}} e^{-\frac{m\omega_0 x^2}{2}} \\ \psi_{n+1}(x) = \frac{1}{\sqrt{n+1}} \left(\sqrt{\frac{m\omega_0}{2}} x - \frac{1}{\sqrt{2m\omega_0}} \frac{\partial}{\partial x} \right) \psi_n(x). \end{aligned} \quad (2.21)$$

The prefactor in Equation (2.21) takes care of the normalization requirement. We visualize the first few wavefunctions of the QHO in Figure 2.1.

Adding back the driving term $F(t)$ leads to the time-dependent Hamiltonian

$$\hat{\mathcal{H}}(t) = \omega_0 \left(\hat{n} + \frac{1}{2} \right) - \frac{F(t)}{\sqrt{2m\omega_0}} (\hat{a} + \hat{a}^\dagger) \quad (2.22)$$

in terms of ladder operators, which shows the possibility to add or remove excitations to the system using the force input $F(t)$.

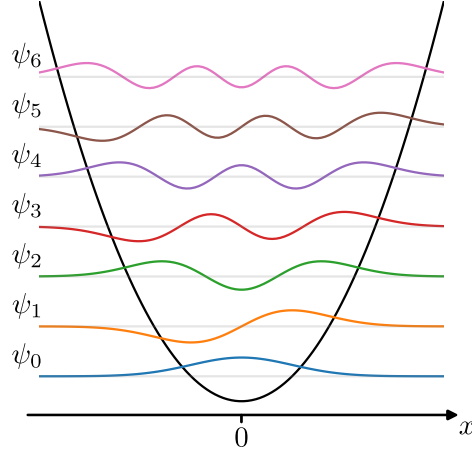


Figure 2.1: Truncated spectrum of the quantum harmonic oscillator. The black line corresponds to the quadratic potential and the wavefunctions ψ_n of the lowest eigenstates are visualized at their respective energies E_n .

2.2 Quantum computing

In this section we discuss the example of a quantum computer in order to give intuition for how the unique features of quantum mechanics can be leveraged to design novel technologies. Most projects in this thesis are centered around this specific technology.

Going forward, we drop the hat on operators (e.g., $\hat{\mathcal{H}} \rightarrow \mathcal{H}$), which is commonly done in research works and it will be clear from context which symbols denote operators.

2.2.1 Quantum bits

The idea behind quantum computing is to encode information in the quantum system and purposefully manipulate it, using superposition and entanglement, to enable computing schemes that a classical computer does not have access to. While the fundamental unit of computation in the classical case is the bit, which can take the value of *either* 0 or 1, the quantum computer operates on quantum bits, called *qubits*, with basis states denoted by

$$|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \quad (2.23)$$

This is also often referred to as the *computational basis*. In hardware, the qubit is typically encoded in two energy eigenstates of the underlying quantum system, $|0\rangle \equiv |E_{k_1}\rangle$ and $|1\rangle \equiv |E_{k_2}\rangle$ for $k_1 \neq k_2$. Figure 2.2 visualizes this for the popular platforms based on neutral atoms and superconducting transmons (see also Section 2.3.2). However, other encoding schemes exist and may offer specific advantages — we refer to Chapter 5 for an example. The specific hardware implementation defines the lowest layer of the quantum computing stack shown in Figure 1.1.

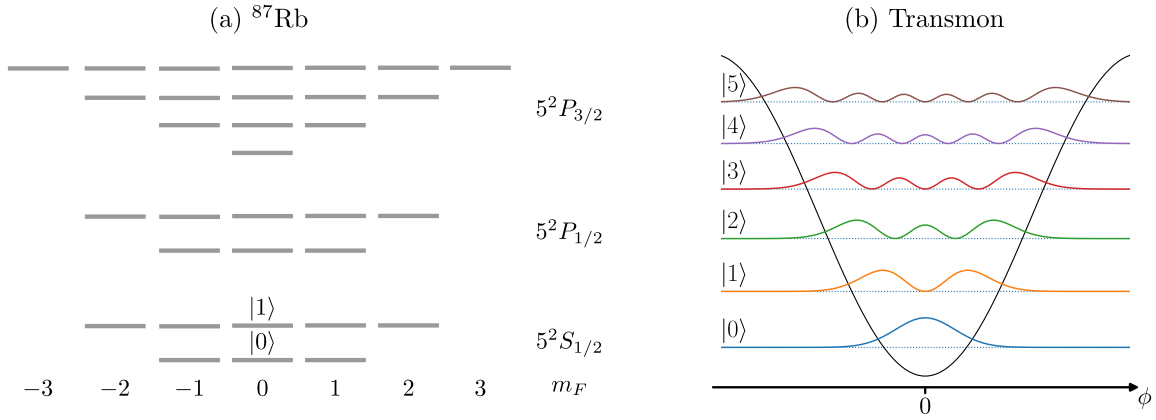


Figure 2.2: Energy spectra showing the encoding of qubit states in larger quantum systems. (a) In neutral atoms such as ^{87}Rb , two hyperfine levels can be used to represent the qubit. (b) In superconducting architectures like transmons, the qubit is typically encoded in the two lowest-energy eigenstates. The colored lines show the probability distributions $\psi_n(\phi)^2$ of the eigenstates $|n\rangle$. Section 2.3.2 includes more background on transmons specifically.

Being a quantum object, a qubit may exist in a superposition of the basis states (Equation (2.10)), so its general state is typically written as

$$|\psi\rangle = \alpha |0\rangle + \beta |1\rangle \in \mathbb{C}^2. \quad (2.24)$$

with complex amplitudes $\alpha, \beta \in \mathbb{C}$, hence the Hilbert space is $\mathcal{V} = \mathbb{C}^2$. We know the meaning of these coefficients from the general discussion in the previous section: The probabilities of observing the qubit in the states $|0\rangle$ and $|1\rangle$ are given by $|\alpha|^2$ and $|\beta|^2$, respectively, and therefore we have the normalization condition $|\alpha|^2 + |\beta|^2 = 1$. Fixing the global phase, an

alternative representation of an arbitrary qubit state is thus given by

$$|\psi\rangle = \cos\left(\frac{\theta}{2}\right) |0\rangle + e^{i\phi} \sin\left(\frac{\theta}{2}\right) |1\rangle, \quad (2.25)$$

with spherical angles $\theta \in [0, \pi]$ and $\phi \in [0, 2\pi)$. Therefore, a common representation of the state space for a single qubit is the *Bloch sphere*, shown in Figure 2.3, which is a unit sphere in $\mathbb{R}^3$ with the poles corresponding to the computational basis states $|0\rangle$ and $|1\rangle$.

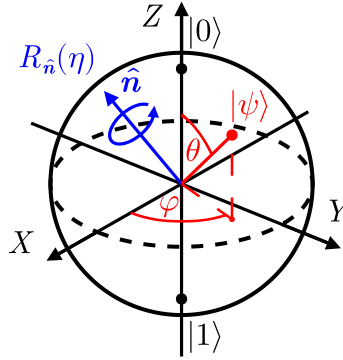


Figure 2.3: A qubit state $|\psi\rangle$ is represented by a point on the Bloch sphere (red). The poles correspond to the computational basis states $|0\rangle$ and $|1\rangle$. Single-qubit gates (see Section 2.2.2) correspond to rotations of the Bloch sphere about specific axes $\hat{n}$ (blue).

A quantum computer requires a large number of qubits to perform powerful computations. In an N -qubit system the computational basis is given by the product basis $\{|0\rangle, |1\rangle\}^{\otimes N}$, thus each basis state corresponds to one of the 2^N classical bitstrings. For instance, in the case of two qubits, the four basis states are $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$. While the state of a classical device is described by exactly one bitstring at a time, the quantum computer can exist in a superposition of all 2^N basis states simultaneously. This exponential factor is the foundation for the potential advantage of quantum applications over their classical counterparts. However, at the same time it makes the simulation of quantum computing on classical machines exponentially hard as it necessitates the representation of elements of $\mathbb{C}^{2^N}$, which is very costly in memory. We expand on this discussion throughout the next few subsections.

2.2.2 Quantum gates and measurement

In order to perform computation, we wish to implement *quantum gates* U that alter the state of the qubits:

$$|\psi\rangle \rightarrow |\psi'\rangle = U|\psi\rangle. \quad (2.26)$$

Since the qubits' dynamics is governed by the TDSE (2.14), any such desired transformation U has to correspond to a unitary operator. Specifically, the underlying Hamiltonian $\mathcal{H}(t)$ has to be tuned during the time interval $[0, T]$ in such a way that the time evolution operator $U(T)$ equals the target unitary, $U(T) \cong U$. Due to the global phase freedom on quantum states, this equality is to be understood as “up to global phase factor” as well and thus denoted with “ $\cong$ ”. T denotes the gate duration, which depends on the specific hardware implementation and how the control of $\mathcal{H}(t)$ is realized. We dive more into the topic of quantum control in Section 2.3 to set the stage for the majority of the thesis, which discusses different control projects (see Chapters 3, 4 and 5). With respect to the device stack (see Figure 1.1), the implementation of gates through physical control signals constitutes the second layer from the bottom.

A quantum gate U acting on N qubits can be represented by a $2^N \times 2^N$ unitary matrix when expressed in a particular basis, which is typically the computational basis. A fundamental set of single-qubit gates that are ubiquitous in the field of quantum information are the *Pauli operators* $\mathcal{P} = \{I, X, Y, Z\}$, defined as

$$I \equiv \sigma_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad X \equiv \sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad Y \equiv \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad Z \equiv \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \quad (2.27)$$

The identity gate I leaves the qubit state unaltered and corresponds to an idling operation. The X and Y gate both flip the computational basis states $|0\rangle \leftrightarrow |1\rangle$ but influence the relative phase of superpositions differently. The Z gate is diagonal in the computational basis (hence the terms “computational basis” and “ Z basis” are often used interchangeably)

and only imposes a relative phase flip: $|0\rangle \rightarrow |0\rangle$, but $|1\rangle \rightarrow -|1\rangle$.

The Pauli operators are unitary but also Hermitian, and in fact they form a basis for the vector space of all 2×2 Hermitian matrices, therefore they can generate any possible single-qubit gate, which we parameterize as (up to global phase)

$$R_{\hat{\mathbf{n}}}(\eta) \equiv e^{i\frac{\eta}{2}\hat{\mathbf{n}}\boldsymbol{\sigma}} = \cos\left(\frac{\eta}{2}\right)I + i\sin\left(\frac{\eta}{2}\right)\hat{\mathbf{n}}\boldsymbol{\sigma}. \quad (2.28)$$

Here $\hat{\mathbf{n}} \in \mathbb{R}^3$ denotes a unit vector ($\|\hat{\mathbf{n}}\| = 1$) and $\boldsymbol{\sigma} = [\sigma_1 \ \sigma_2 \ \sigma_3]^\top$ is the vector of nontrivial Pauli operators. The gate $R_{\hat{\mathbf{n}}}(\eta)$ corresponds to a rotation about the axis $\hat{\mathbf{n}}$ by an angle η in the Bloch sphere, which is highlighted in Figure 2.3, and thus they are also referred to as *rotation gates*. We use a special notation for rotations about the principal axes X , Y and Z :

$$X(\eta) \equiv R_{\hat{\mathbf{x}}}(\eta), \quad Y(\eta) \equiv R_{\hat{\mathbf{y}}}(\eta), \quad Z(\eta) \equiv R_{\hat{\mathbf{z}}}(\eta). \quad (2.29)$$

As a brief observation we point out that the nontrivial Pauli gates can be viewed as rotations by an angle π , for instance, $X \cong X(\pi)$.

Some frequently encountered single-qubit gates are given special symbols, such as the phase gates

$$S = Z^{\frac{1}{2}} = \begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix} \cong Z\left(\frac{\pi}{2}\right), \quad T = Z^{\frac{1}{4}} = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{4}} \end{bmatrix} \cong Z\left(\frac{\pi}{4}\right), \quad (2.30)$$

and the *Hadamard gate*

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \cong R_{\frac{\hat{\mathbf{x}}+\hat{\mathbf{z}}}{\sqrt{2}}}(\pi) \quad \longrightarrow \quad H = \begin{cases} |0\rangle \leftrightarrow (|0\rangle + |1\rangle)/\sqrt{2}, \\ |1\rangle \leftrightarrow (|0\rangle - |1\rangle)/\sqrt{2}, \end{cases} \quad (2.31)$$

which maps between the computational basis states and equal-weight superpositions.

In order to create entanglement between qubits, we require multi-qubit gates. The sim-

plest kind are two-qubit gates, with two common examples being

$$\text{CNOT} = \text{CX} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad \text{CPHASE} = \text{CZ} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}. \quad (2.32)$$

The CNOT (“controlled NOT”) gate flips the state of the *target qubit* q_1 conditioned on the state of the *control qubit* q_0 , i.e., $\text{CNOT} |q_1\rangle |q_2\rangle = |q_1\rangle |q_2 \oplus q_1\rangle$ for $q_1, q_2 \in \{0, 1\}$. In other words, it applies a conditional X gate on the target qubit, hence the equivalent notation CX (“controlled X ”). Similarly, the CPHASE (“controlled PHASE”) gate applies a conditional Z gate on the target qubit ($\rightarrow$ CZ, “controlled Z ”), which corresponds to flipping the phase of the $|11\rangle$ basis state and leaving the other basis states unaltered. Another crucial two-qubit gate is the SWAP gate, $\text{SWAP} |q_1 q_2\rangle = |q_2 q_1\rangle$, which cannot create entanglement but is very useful to move quantum information around between qubits.

While higher-dimensional multi-qubit gates can be defined (such as the Toffoli gate, see Figure 2.5), any such gate can be decomposed into a sequence of single- and two-qubit gates. Specifically, the access to arbitrary single-qubit operations together with one entangling gate is sufficient for this task. Moreover, the former can be approximately realized (to arbitrary precision) using sequences of gates from a special finite single-qubit gate set, according to the Solovay-Kitaev theorem [42]. Such gate sets are called *universal* and a common universal single-qubit gate set is given by $\{H, T\}$. Therefore, universal quantum computing on any number of qubits can be achieved with $\{H, T, \text{CNOT}\}$ for instance.

As mentioned before, the Pauli operators are Hermitian, therefore they also constitute qubit observables. When measuring one of the nontrivial operators X , Y or Z , the state collapses to one of the eigenstates

$$X : |\pm\rangle = \frac{|0\rangle \pm |1\rangle}{\sqrt{2}}, \quad Y : |\pm i\rangle = \frac{|0\rangle \pm i |1\rangle}{\sqrt{2}}, \quad Z : |0/1\rangle, \quad (2.33)$$

depending on the probabilistic measurement outcome ± 1 (eigenvalue). “Measuring” I corresponds to no measurement as all qubit states are eigenstates of I with eigenvalue $+1$.

Typically, the native measurement basis supported on a quantum computer is the Z basis. In order to measure in a different basis, specific gates have to be applied to the qubit state prior to measurement, effectively rotating the measurement basis. For instance, the Hadamard gate H , recalling its definition in Equation (2.31), can transform the Z basis measurement into an X basis measurement.

The simultaneous measurement of Pauli operators on N qubits is described by a *Pauli string*

$$P = \bigotimes_{j=1}^N P^j \equiv P^1 P^2 \dots P^N \in \mathcal{P}_N, \quad (2.34)$$

which is an element of $\mathcal{P}_N = \mathcal{P}^{\otimes N}$. Here P^k denotes the Pauli operator acting on the k th qubit. Often I components are not written explicitly, so the four-qubit Pauli string $IXIZ$ would be abbreviated as $X^2 Z^4$ for instance. The possible measurement outcomes of a Pauli string are ± 1 as well, and are equal to the product of the individual qubit measurement results.

The set of all N -qubit Pauli strings forms a basis for the vector space of Hermitian operators on N qubits, thus any N -qubit observable O can be expressed as a linear combination

$$O = \sum_k c_k P_k \quad (2.35)$$

with real coefficients $c_k \in \mathbb{R}$. Most importantly, any N -qubit Hamiltonian $\mathcal{H}$ can be represented this way, which becomes particularly relevant in Chapter 6.

2.2.3 Quantum circuits

With the fundamental building blocks at the lowest two levels of the stack available (i.e., physical qubits and gates), we can consider the top layer of the stack: quantum programs.

In analogy to classical programs, a quantum program or algorithm consists of a sequence of quantum gates that is applied to an N -qubit register, followed by measurement of the register to read out the result of the computation. Programs are commonly represented by *quantum circuits*, which show the succession of operations acting on the qubits. Figure 2.4 visualizes the structure of a quantum circuit.

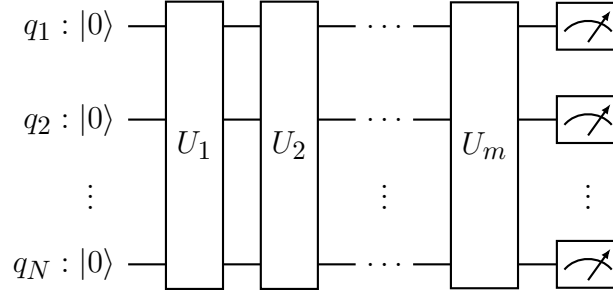


Figure 2.4: Conceptual depiction of a quantum circuit. A sequence of unitary gates U_i is applied to the qubit register that is typically initialized in the state $|0\rangle^{\otimes N}$. At the end the qubits are measured in the Z basis.

The qubits are typically initialized to the $|0\rangle$ state and each wire represents the state evolution of a qubit. Time flows from left to right, so the application of gate layers U_1 through U_m corresponds to the circuit unitary

$$U = U_m \dots U_2 U_1. \quad (2.36)$$

Each layer U_i might act only on a subset of qubits. The total number of layers m is called the *circuit depth*, while the number of qubits N is also referred to as the *circuit width*. The meter symbol at the end of the circuit denotes the measurement in the Z basis.

A quantum circuit may include complex multi-qubit operations which can be challenging to realize on hardware directly. We know from above that these gates can be decomposed into sequences of fundamental basis gates from a universal gate set, therefore, if the quantum computer supports such a gate set, this challenge can be addressed. As an example, we show in Figure 2.5 the decomposition of the Toffoli gate [136], which is also called the CCNOT

gate as it flips the state of the target qubit given that two control qubits are in the state $|1\rangle$.

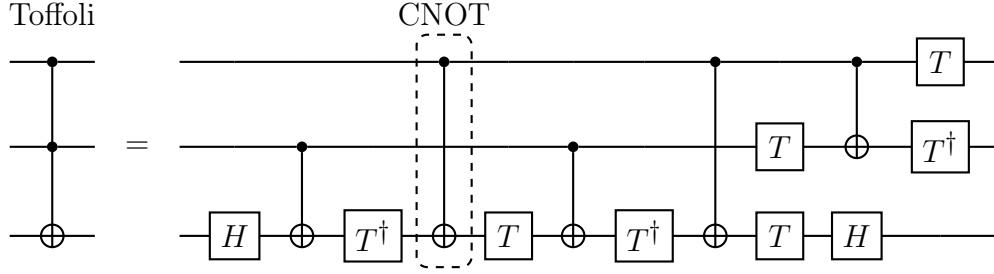


Figure 2.5: Quantum circuit decomposition of the three-qubit Toffoli gate into single- and two-qubit gates.

2.2.4 Quantum compilation

Given a high-level quantum circuit, it usually needs to be modified to meet the requirements of the underlying quantum hardware. Devices in the noisy intermediate-scale quantum (NISQ) era as well as future fault-tolerant machines impose constraints such as limited qubit connectivity or finite basic instruction sets. For instance, in superconducting hardware we usually find physical qubits laid out in a grid (or even sparser) structure, whereas on neutral atom platforms qubits can interact over longer distances.

The translation process that transforms the input circuit to an equivalent circuit respecting all the constraints is called *quantum compilation*. Key components of compilation pipelines include rewriting the circuit in terms of the device’s native gate set, mapping program qubits to hardware qubits, routing qubits to compensate for limited connectivity, and scheduling the instructions in time. We refer to Ref. [40] for a more detailed description of quantum compilation and insights on how optimized compilation can boost the success probability of quantum programs.

2.2.5 Clifford gates and circuits

While generally the complexity of classically simulating quantum circuits scales exponentially with the size of the qubit register N due to the exponential scaling of the Hilbert space dimension $d = 2^N$, there are certain gate families that allow for classically efficient simulation because of their special properties.

One such family is given by the *Clifford group* $\mathcal{C}_N$ on N qubits. It includes all quantum gates C that normalize the corresponding Pauli group $\mathcal{P}_N$

$$CPC^\dagger \in \pm\mathcal{P}_N \quad \forall P \in \mathcal{P}_N. \quad (2.37)$$

The expression $CPC^\dagger$ is called the *conjugation* of the Pauli operator P with the Clifford operator C and by definition yields a Pauli operator again. However, it can introduce a sign flip, indicated with $\pm$ here. We denote the conjugation of an input Pauli P to an output Pauli P' by $P \rightarrow \pm P'$. The Clifford group includes the set of Pauli gates, $\mathcal{P}_N \subset \mathcal{C}_N$.

Just like any other quantum gate, multi-qubit Clifford gates can be decomposed into single- and two-qubit Clifford operations. Some single-qubit Clifford gates and their conjugation rules include:

$$H : X \leftrightarrow Z, \quad S : X \rightarrow Y \rightarrow -X \rightarrow -Y \rightarrow X. \quad (2.38)$$

The CNOT gate $CX_{c \rightarrow t}$ (with control on qubit c and target on qubit t) is a commonly encountered two-qubit Clifford gate:

$$CX_{c \rightarrow t} = \begin{cases} X^c \rightarrow X^c X^t, & X^t \rightarrow X^t, \\ Z^c \rightarrow Z^c, & Z^t \rightarrow Z^c Z^t. \end{cases} \quad (2.39)$$

It is important to note how it correlates certain Pauli terms with each other. This observation

is a crucial ingredient for the work we present in Chapter 6.

Clifford circuits are quantum circuits that only include gates from the Clifford group. They can be efficiently simulated because it suffices to represent the action of each gate by its specific conjugation rules, thus an exponentially-scaling matrix representation is not required. This enables polynomially-scaling simulation [64] of these circuits, which is implemented in libraries like Stim [58].

Due to this nature, Clifford gates are not sufficient for universal quantum computation as they can only explore a limited subset of Hilbert space. Clifford gates on N qubits can map between *stabilizer states* $|\psi_{\mathcal{S}}\rangle$, which are defined as the mutual $+1$ eigenstates of their respective N commuting *stabilizers* $\mathcal{S} = \{S_i\}_{i=1}^N \subset \mathcal{P}_N$:

$$S_i |\psi_{\mathcal{S}}\rangle = |\psi_{\mathcal{S}}\rangle \quad \forall S_i \in \mathcal{S}. \quad (2.40)$$

For the initial state $|0\rangle^{\otimes N}$ we have $\mathcal{S} = \{Z^i\}_{i=1}^N$ for instance. The action of a Clifford gate C yields

$$C |\psi_{\mathcal{S}}\rangle = C S_i |\psi_{\mathcal{S}}\rangle = (C S_i C^\dagger) C |\psi_{\mathcal{S}}\rangle \quad \forall S_i \in \mathcal{S}, \quad (2.41)$$

thus the output state is described by the conjugated stabilizer group $C\mathcal{S}C^\dagger = \{C S_i C^\dagger\}_{i=1}^N$.

Universality can be recovered by adding a non-Clifford gate to the instruction set. A popular choice is the T gate, leading to the “Clifford+ T ” set that is of high interest in the context of arbitrary gate decomposition and fault-tolerant quantum computing [158].

2.3 Control of quantum hardware

We now turn to the task of purposefully manipulating a quantum system in order to realize desired gates U , which is at the heart of the second layer of the stack (see Figure 1.1). The

system needs to be governed by a tunable Hamiltonian $\mathcal{H}(t)$, which can be decomposed as

$$\mathcal{H}(t) = \mathcal{H}_0 + \mathcal{H}_c(t) = \mathcal{H}_0 + \sum_j f_j(t) \mathcal{H}_j. \quad (2.42)$$

Here $\mathcal{H}_0$ denotes the system's internal *drift Hamiltonian* and $\mathcal{H}_c(t)$ the *control Hamiltonian*, which can be expressed as classical control functions f_j coupling to *drive Hamiltonians* $\mathcal{H}_j$. The goal is to choose the control pulses f_j in such a way that the system evolution (see Equation (2.15)) realizes the desired transformation U_{targ} after a time T :

$$U(T) = \mathcal{T} \exp \left(-i \int_0^T \mathcal{H}(t) dt \right) \stackrel{!}{=} U_{\text{targ}}. \quad (2.43)$$

As an example, we look at the simple case of a two-level qubit. In the so-called *lab frame*, a possible Hamiltonian is given by

$$\mathcal{H}(t) = -\frac{\omega_{01}}{2} Z + \Omega(t) \cos(\omega t + \varphi) X. \quad (2.44)$$

The drift term sets the energy values of the computational basis states, i.e., $E_0 = -\omega_{01}/2$ and $E_1 = \omega_{01}/2$ are the energies of $|0\rangle$ and $|1\rangle$, respectively. Thus, the qubit's energy gap equals $E_{01} = E_1 - E_0 = \omega_{01}$. The classical control is given by a cosine signal $f(t) = \Omega(t) \cos(\omega t + \varphi)$ with drive phase φ and tunable amplitude $\Omega(t)$. Typically, the timescale set by the qubit and the drive is much shorter than the timescale of the gate, i.e., $\omega_{01}^{-1}, \omega^{-1} \ll T$. Moving to a time-dependent *rotating frame* defined by a unitary basis transformation $U_{\text{rot}}(t)$,

$$\begin{aligned} |\psi_{\text{rot}}(t)\rangle &= U_{\text{rot}}^\dagger(t) |\psi(t)\rangle, \\ \mathcal{H}_{\text{rot}}(t) &= U_{\text{rot}}^\dagger(t) \mathcal{H}(t) U_{\text{rot}}(t) - i U_{\text{rot}}^\dagger(t) \dot{U}_{\text{rot}}(t) \end{aligned} \quad (2.45)$$

allows focusing on the slow dynamics of the gate determined by $\Omega(t)$. In our example, we choose a rotating frame $U_{\text{rot}}(t) = e^{i\frac{\omega t}{2} Z}$ defined by the drive frequency ω , which leads to the

transformed Hamiltonian

$$\mathcal{H}_{\text{rot}}(t) = -\frac{\delta}{2}Z + \frac{\Omega(t)}{2}((\cos(\varphi) + \cos(2\omega t + \varphi))X - (\sin(\varphi) + \sin(2\omega t + \varphi))Y) \quad (2.46)$$

$$\rightarrow -\frac{\delta}{2}Z + \frac{\Omega(t)}{2}(\cos(\varphi)X - \sin(\varphi)Y). \quad (2.47)$$

Here $\delta = \omega_{01} - \omega$ denotes the *detuning* between the qubit frequency and the drive frequency. The crucial step happens in Equation (2.47), where the *rotating wave approximation* (RWA) is applied: The fast-oscillating terms ($\sim 2\omega$) are neglected as they can be approximated to average to zero on the relevant timescale of the gate, assuming weak driving, i.e., $\Omega(t) \ll \omega_{01}, \omega$. The RWA is widespread in the study of the control of quantum systems and computation is typically performed in such a rotating frame.

When driving the qubit *on resonance*, i.e., $\omega = \omega_{01}$, the RWA Hamiltonian simplifies to

$$\mathcal{H}_{\text{rot}}(t) = \frac{\Omega(t)}{2}(\cos(\varphi)X - \sin(\varphi)Y), \quad (2.48)$$

leading to the time evolution

$$U(T) = \exp\left(-i\frac{\int_0^T \Omega(t) dt}{2}(\cos(\varphi)X - \sin(\varphi)Y)\right) = R_{\hat{\mathbf{n}}}(\eta), \quad (2.49)$$

which realizes a rotation about the axis $\hat{\mathbf{n}} = (\cos(\varphi), -\sin(\varphi), 0)$ by an angle $\eta = \int_0^T \Omega(t) dt$ in the Bloch sphere (see Equation (2.28)). Therefore, driving a qubit on resonance enables the direct implementation of arbitrary rotations about axes in the X - Y plane of the Bloch sphere, depending on the drive phase φ . In this setup the Z axis is not directly accessible, but physical Z rotations could be achieved by combining X and Y rotations. However, this is not necessary as Z rotations can be implemented virtually in zero time by advancing the drive phase of subsequent operations in a quantum circuit [119]. We directly exploit this property for the gate optimizations we consider in Section 3.3.

If the drive frequency is strongly detuned from the qubit frequency, i.e., $|\delta| \gg |\Omega(t)|$, then no population transfer between the computational states is possible. This can be illustrated with the example of a constant X drive with $\Omega(t) = \Omega$ and $\varphi = 0$:

$$\mathcal{H}_{\text{rot}} = -\frac{\delta}{2}Z + \frac{\Omega}{2}X \quad \longrightarrow \quad U(T) = R_{\hat{\mathbf{n}}}(\eta) \quad \text{with} \quad \hat{\mathbf{n}} = \frac{1}{\sqrt{\Omega^2 + \delta^2}} \begin{bmatrix} \Omega \\ 0 \\ -\delta \end{bmatrix} \xrightarrow{|\delta| \gg |\Omega|} -\hat{\mathbf{z}}. \quad (2.50)$$

Therefore, a qubit needs to be driven on resonance to effectively move the state around the Bloch sphere. This means that a system of two levels can only be addressed when the drive frequency matches (or is close to) its energy gap. When the qubit is encoded in a bigger Hilbert space, it is thus important to engineer the system in such a way that this individual addressability is assured, i.e., the qubit's energy gap is sufficiently different from transition energies to neighboring eigenstates. Examples include superconducting platforms, which we discuss in more detail in Section 2.3.2.

While the control of an isolated qubit is rather simple, the control problem becomes more complex as more energy levels get involved. Generalizing from the idea of the encoded qubit, we are often interested in controlling only a subspace $\tilde{\mathcal{V}} \subset \mathcal{V}$ of a system's full Hilbert space $\mathcal{V}$, i.e., implement transformations on a subset of states. We refer to $\tilde{\mathcal{V}}$ as the *computational subspace* with dimension $d = \dim \tilde{\mathcal{V}}$. The challenge becomes to design control pulses that implement the desired transformations on $\tilde{\mathcal{V}}$ without causing *leakage* outside the computational subspace, i.e., populating states outside $\tilde{\mathcal{V}}$, which relates back to the idea of individual addressability from the previous paragraph.

The quality of a quantum gate implementation $U(T; \mathbf{f}) : \mathcal{V} \rightarrow \mathcal{V}$, where we explicitly indicate the dependence on the controls $\mathbf{f} = [f_j]_j$, is quantified by its *fidelity* $\mathcal{F}$ or *infidelity* $\mathcal{I}$ with respect to the desired target gate $U_{\text{targ}} : \tilde{\mathcal{V}} \rightarrow \tilde{\mathcal{V}}$, which are defined as

$$\mathcal{F}(T; \mathbf{f}) = \frac{1}{d^2} \left| \text{Tr} \left[U_{\text{targ}}^\dagger P_{\tilde{\mathcal{V}}} U(T; \mathbf{f}) P_{\tilde{\mathcal{V}}}^\dagger \right] \right|^2 \in [0, 1], \quad \mathcal{I}(T; \mathbf{f}) = 1 - \mathcal{F}(T; \mathbf{f}) \in [0, 1]. \quad (2.51)$$

Here $P_{\tilde{\mathcal{V}}} : \mathcal{V} \rightarrow \tilde{\mathcal{V}}$ denotes the projector onto the computational subspace. The fidelity equals one (the infidelity equals zero) if the target gate U_{targ} is realized perfectly up to global phase. Therefore, when optimizing quantum systems, experimentalists are concerned with improving their gate calibrations (i.e., tuning $\mathbf{f}$) in order to push the gate fidelities as close to unity as possible. However, there exist many sources of distortion that influence a quantum system and can limit the gate performance. We devote Section 2.4 to this issue.

2.3.1 Quantum optimal control

Finding analytical control solutions, like we did for the qubit in the previous section, can become challenging as the complexity of the quantum system increases. To address this issue, various numerical optimization frameworks have been proposed in order to find hardware-specific control pulses that realize desired transformations to high accuracy [141, 75, 105, 146, 22, 182, 59]. This is referred to as *quantum optimal control* (QOC), and it represents a major part of this thesis. Specifically, Chapters 3 and 4 deal with applications of quantum optimal control to two different quantum technologies.

While different frameworks vary in their specific implementation, they generally follow the same high-level approach of minimizing a cost function $\mathcal{L}$, which is usually primarily given by the gate infidelity $\mathcal{I}$ (2.51). To this end, the control functions $\mathbf{f}(\boldsymbol{\gamma}) = [f_j(\boldsymbol{\gamma}_j)]_j$ are expressed in terms of optimizable parameters $\boldsymbol{\gamma} = [\boldsymbol{\gamma}_j]_j$, which in the simplest case could correspond to the control values at discrete times $t_l \in [0, T]$:

$$f_j(t_l; \boldsymbol{\gamma}_j) = \gamma_{jl}. \quad (2.52)$$

This is the scenario we encounter in Chapter 4. Alternatively, $\boldsymbol{\gamma}$ can serve as coefficients in a basis expansion

$$f_j(t; \boldsymbol{\gamma}_j) = \sum_s \gamma_{js} B_s(t) \quad (2.53)$$

with basis functions $B_s(t)$, and we apply this approach in Chapter 3.

Given a parameterization of the controls, the infidelity

$$\mathcal{I}(T; \gamma) \equiv \mathcal{I}(T; \mathbf{f}(\gamma)) \quad (2.54)$$

of the resulting gate $U(T; \gamma)$, which is consistent with the dynamics imposed by the TDSE, with respect to a target gate U_{targ} can be estimated via Equation (2.51). The simplest optimization problem is then given by

$$\hat{\gamma} = \arg \min_{\gamma} \mathcal{L}(\gamma) \quad (2.55)$$

with $\mathcal{L}(\gamma) = \mathcal{I}(T; \gamma)$, which is typically solved using gradient-based methods. However, more terms can be added to the cost function $\mathcal{L}$ in order to encourage solutions with certain properties. For instance, we might want to penalize large drive amplitudes to respect power constraints imposed by the hardware or discourage leakage outside the computational subspace during the gate evolution.

The variation of QOC that we have introduced so far is called *open-loop* control. The optimization is based on the model of the quantum system (i.e., the Hamiltonian $\mathcal{H}(t)$), which is assumed to be perfect. However, in practice, the real quantum system deviates from the model due to unwanted distortions (see Section 2.4) or unmodeled effects, inevitably leading to a reduced gate fidelity when implementing the QOC pulses in hardware. While this mismatch can be mitigated by improving the characterization of the hardware to reduce the model error, an alternative approach is given by including the experiment in the optimization routine, thereby *closing the loop*. In this *closed-loop* setting, the parameterized pulses are applied to the hardware and the fidelity determined experimentally (e.g., via tomography or randomized benchmarking, see Section 3.3.5 for instance), which then feeds back into the optimization to tune the controls. Therefore, the model may be imperfect, and moreover

there have been pulse optimization schemes proposed that are completely model-free [9, 198]

Closed-loop optimal control is more advantageous to produce high-fidelity gates that work well on a specific hardware, however, the optimization scheme is quite time-costly and data-intensive, as it requires many experimental runs per iteration to estimate the value of different observables in order to determine the infidelity. In this thesis we focus on open-loop control as we are mostly concerned with theoretical studies of quantum systems.

2.3.2 Example: The transmon processor

We conclude this section by going over one prominent approach to construct quantum computing hardware based on superconducting circuits and discussing their control. We specifically consider the popular transmon architecture introduced in Ref. [96], which is nowadays being researched by many academic groups as well as industry teams like IBM and Google. We sketch the derivation of the model Hamiltonian (which we shall recognize in later chapters), where we follow and summarize the description in Ref. [99].

In its simplest form, the transmon consists of a Josephson junction (JJ) that is shunted by a large capacitor, as shown in Figure 2.6, giving rise to the system Hamiltonian

$$\mathcal{H} = 4E_C n^2 - E_J \cos(\phi). \quad (2.56)$$

Here E_C denotes the capacitive energy per electron, thus $4E_C$ corresponds to the capacitive energy per Cooper pair, and the operator n counts the number of excess Cooper pairs on the superconducting island. E_J describes the JJ's energy and the operator ϕ is the superconducting phase difference across the JJ. For a proper introduction to superconductivity and the Josephson effect, we refer to Ref. [94].

The operators ϕ and n are the conjugate operators for the superconducting circuit just like x and p are the conjugate operators for a moving particle, which lets us draw parallels between

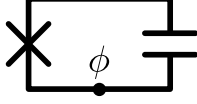


Figure 2.6: A Josephson junction (left) is shunted by a large capacitor (left) to form a transmon.

the different physical systems by identifying $\phi \leftrightarrow x$ and $n \leftrightarrow p$. Thus the term $-E_J \cos(\phi)$ can be interpreted as a potential. Numerically diagonalizing the system leads to the spectrum visualized in Figure 2.2(b). In the following we discuss the common approximate description of the transmon.

In the transmon regime $E_J \gg E_C$, we may expect the eigenstates with the lowest energy to be well localized at the bottom of the potential, therefore we can expand the $\cos(\phi)$ term up to fourth order and obtain (up to constant terms)

$$\mathcal{H} = 4E_C n^2 + E_J \left(\frac{1}{2} \phi^2 - \frac{1}{24} \phi^4 \right). \quad (2.57)$$

Introducing ladder operators $\phi \propto (a + a^\dagger)$ and $n \propto -i(a - a^\dagger)$ as we did for the QHO in Equation (2.17), ultimately yields

$$\mathcal{H} = \omega_0 a^\dagger a - \frac{K}{2} a^{\dagger 2} a^2 = \left(\omega_0 + \frac{K}{2} \right) a^\dagger a - \frac{K}{2} (a^\dagger a)^2. \quad (2.58)$$

Here all terms that have a different number of creation and annihilation operators have been neglected (keyword: rotating wave approximation) and we used $[a, a^\dagger] = 1$. Recalling the discussion on the QHO in Section 2.1.3, we note that the resulting Hamiltonian is diagonal in the basis of the number operator $n = a^\dagger a$. However, the energies are now given by

$$E_n = \left(\omega_0 + \frac{K}{2} \right) n - \frac{K}{2} n^2, \quad n \in \mathbb{N}, \quad (2.59)$$

and most importantly, the gaps between neighboring states are not equal anymore, $E_{n+1} -$

$E_n = \omega_0 - nK$. Therefore, the transmon can be regarded as an *anharmonic oscillator*, with base frequency $\omega_0 \approx \sqrt{8E_C E_J}$ and *anharmonicity* (also called *Kerr nonlinearity*) $K = -E_C$. Typical transmon parameters are in the range $\omega_0/(2\pi) \sim 3 - 6$ GHz and $K/(2\pi) \sim 100 - 300$ MHz.

In a transmon architecture, the qubit is typically encoded in the lowest two eigenstates $|0\rangle$ and $|1\rangle$. The qubit frequency is given by $E_{01} = \omega_0$ and is different from transition energies to higher excited states. This is the reason for using the JJ as an inductive element: The $\cos(\phi)$ potential creates the nonlinear ladder of eigenstates, enabling individual addressing of the qubit transition and thereby control of the qubit subspace. If a linear inductive element had been chosen, creating an LC circuit, the energy levels would have been equally spaced (like in the harmonic oscillator as shown in Figure 2.1) and driving the qubit states on resonance would lead to on-resonance driving of unwanted transitions as well, causing leakage.

Driving the transmon is achieved by coupling to the charge ($\rightarrow n$) or phase ($\rightarrow \phi$) degree of freedom. For instance, the former is realized by capacitively coupling the transmon to a waveform generator, which adds the control Hamiltonian

$$\mathcal{H}_c(t) = f(t)(a + a^\dagger) \quad (2.60)$$

to the system. Here we see the connection to driving a mechanical harmonic oscillator that we showed in Equation (2.22). Given the magnitude of the qubit frequency ω_0 , typical drive frequencies lie in the microwave regime ($\sim$ GHz).

In order to construct a practical quantum computing platform, different qubits have to be able to interact with each other. This is achieved by engineering a coupling like

$$\mathcal{H}_{\text{int}}(t) = g(t)(a_1^\dagger a_2 + a_2^\dagger a_1) \quad (2.61)$$

for instance, also called *beamsplitter interaction*, where excitations are exchanged between

the two transmons labeled 1 and 2. Depending on the architecture, the coupling might be static $g(t) = g$ (IBM's former approach) or tunable $g(t)$ (Google's and also IBM's current approach).

2.4 The distorting impact of noise

Current quantum computers are not perfect but instead prone to errors, caused by unwanted interactions with the environment and distortions to the control of the quantum system. This current state of technology defines the *noisy intermediate-scale quantum* (NISQ) era, where the fidelity of quantum programs is limited and therefore prevents the advantage of quantum computing over classical computing.

Due to the impact of noise and the uncertainty it introduces, a quantum system cannot be described by a so-called *pure* state $|\psi\rangle$ anymore in the general case. Instead, we require the more general concept of the *density operator*

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| \quad (2.62)$$

to describe possibly *mixed states*. ρ is a positive semi-definite ($\rho \succcurlyeq 0$), Hermitian ($\rho = \rho^\dagger$) operator that represents the classical mixtures of pure states $|\psi_i\rangle$ with probabilities p_i . This implies $\text{Tr}[\rho] = 1$. For the case of a qubit, pure states lie on the surface of the Bloch sphere, see Figure 2.3, while mixed states make up the inside.

The quantum mechanical concepts we discussed for pure states can be generalized to mixed states. For instance, given an observable O , the probability of observing a specific outcome o_k and the expectation value of the observable correspond to

$$p(o_k) = \langle o_k | \rho | o_k \rangle \quad \text{and} \quad \langle O \rangle = \text{Tr}[\rho O], \quad (2.63)$$

respectively. While the evolution of a pure state is unitary, $|\psi'\rangle = U|\psi\rangle$, the dynamics of mixed states is described by more general *quantum channels*

$$\rho' = \mathcal{E}(\rho) = \sum_{jk} \chi_{jk} A_j \rho A_k^\dagger. \quad (2.64)$$

Here $\{A_j\}_j$ denotes an operator basis and χ is called the *process matrix*, which parameterizes the quantum channel. In an experimental setting, the process matrix (and thereby the model of the quantum channel) can be reconstructed from measurement data in a scheme called *quantum process tomography*. We employ this method in Section 3.3.5.

A particularly common noise channel is the *depolarizing channel*

$$\mathcal{E}_{\text{dep}}(\rho) = p \frac{I}{d} + (1 - p)\rho, \quad (2.65)$$

where the state ρ gets transformed into the fully mixed state I/d , which has lost all quantum information, with chance p . For an N -qubit system, this can be rewritten as

$$\mathcal{E}_{\text{dep}}(\rho) = (1 - p')\rho + \sum_{j=1}^{4^N-1} \frac{p'}{4^N-1} P_j \rho P_j, \quad (2.66)$$

where the sum goes over all $4^N - 1$ nontrivial Pauli strings $P_j \neq I^{\otimes N}$. Therefore, with chance $p'/(4^N - 1)$, one of these strings is applied to the system (corresponding to an error), while with chance $1 - p'$ the state remains unaltered. The connection to Equation (2.65) is given by $p' = (4^N - 1)p/4^N$. For the single-qubit case specifically, we have

$$\mathcal{E}_{\text{dep}}(\rho) = (1 - p')\rho + \frac{p'}{3}(X\rho X + Y\rho Y + Z\rho Z). \quad (2.67)$$

Therefore, with chance $p'/3$, the qubit experiences either a bit-flip error (X), a phase-flip error (Z), or a combination of both ($Y = iXZ$).

2.4.1 Common sources of errors

While there exist many error mechanisms that vary across different platforms, here we briefly mention three prominent classes of errors.

Thermal relaxation

A quantum system has a natural tendency to decay to a low-energy configuration. In many cases, $|0\rangle$ is encoded in a state with lower energy than $|1\rangle$, thus a qubit in the arbitrary superposition $\alpha|0\rangle + \beta|1\rangle$ progressively decays down to $|0\rangle$ as it evolves over time. The probability of having decayed from $|1\rangle$ to $|0\rangle$ after a time t is given by e^{-t/T_1} , where T_1 is the characteristic decay time of a qubit describing its thermal relaxation.

Gate errors

The quantum gates U that we wish to execute are noisy themselves, stemming from undesired couplings and imperfect control. A good way to model this effect (on average) is given by the noiseless application of U followed by a depolarizing noise channel $\mathcal{E}_{\text{dep}}$, which is realized by the stochastic application of Pauli operators as shown in Equation (2.66). In particular, gate error rates gained from experimental methods like randomized benchmarking correspond to depolarizing error probabilities [113].

Measurement errors

When a qubit is measured, it collapses from a superposition state to one of the basis states. For instance, when measuring in the Z basis, $\alpha|0\rangle + \beta|1\rangle$ collapses to $|0\rangle$ or $|1\rangle$. However, the qubit might suffer errors at the time of measurement due to imperfect measurement control or misclassification of the result, i.e., $|0\rangle$ could be reported as $|1\rangle$, and vice versa.

2.4.2 Quantum error mitigation and error correction

While experimentalists work to construct quantum hardware that experiences lower error rates, there has been much interest in developing other techniques to push the limits of a quantum computer's performance. For instance, multiple proposals of *quantum error mitigation* methods have emerged over the past decade (some examples include Refs. [37, 178, 187, 39]) that try to suppress errors, cancel them out in a clever way or post-process computation results in order to boost the signal of the correct answer. In Chapter 6, we design a pre-processing error mitigation technique specifically designed for a certain class of algorithms, and we refer to Section 6.2.5 for more background on error mitigation work.

Instead of only mitigating errors, being able to correct for errors dynamically during the execution of a program would be even more powerful and enable the implementation of large-scale algorithms. The field of *quantum error correction* (QEC) addresses this challenge and represents a highly active area of research because the quantum community has been realizing that QEC is a vital ingredient towards achieving fault-tolerance. The key idea of QEC is to encode the quantum information of a qubit redundantly in a larger Hilbert space; for instance, in the Hilbert space of multiple physical qubits. Research on QEC explores the design of error correction codes, trying to find good tradeoffs between resource overhead and the ability to correct as many errors as possible.

As an example, in *stabilizer codes*, the manifold of a *logical* qubit encoded in N physical qubits is determined by a set $\mathcal{S} = \{S_i\}_{i=1}^{N-1}$ of $N - 1$ commuting Pauli stabilizers, that constrain the global quantum state to the mutual $+1$ eigenspace:

$$S_i |\psi\rangle = |\psi\rangle \quad \forall S_i \in \mathcal{S}. \quad (2.68)$$

This is similar to Equation (2.40), however, here we have one stabilizer less, therefore leaving a two-dimensional subspace that describes the logical qubit. Measuring the stabilizers

$\mathcal{S}$ allows detecting errors on the physical qubits (whenever the outcome is not $+1$) and corrections can be applied. A highly popular code is the *surface code* [51], where the qubits are laid out on a 2D grid and the stabilizers are given by $XXXX$ and $ZZZZ$ products of physical qubits that surround the unit cells. Great research efforts target the questions of how to translate error signatures to required corrections (*decoding*) and how to implement operations on logical qubits given a specific code.

A quantum error correction code therefore defines an intermediate layer in the quantum computing stack as shown in Figure 1.1 because it translates physical qubits and gates to logical qubits and operations that the high-level program interacts with. For more background on the vast field of QEC, we refer to the introductory guide in Ref. [157] and the draft textbook in Ref. [66].

CHAPTER 3

QUANTUM OPTIMAL CONTROL FOR QUDIT COMPUTATION

Current efforts to build quantum computers mostly focus on the fabrication and high-fidelity control of two-level qubits. However, many proposed implementations like superconducting transmons or neutral atoms provide a much larger Hilbert space with more energy levels present, as is visualized in Figure 2.2. Their influence is typically aimed to be suppressed through careful device engineering (e.g., by creating an anharmonicity in the transmon, see Section 2.3.2), but these states would be available for computation as well.

Therefore, *qudits*, the d -level generalization of qubits with state vector

$$|\psi\rangle = \alpha_0 |0\rangle + \alpha_1 |1\rangle + \cdots + \alpha_{d-1} |d-1\rangle = \sum_{k=0}^{d-1} \alpha_k |k\rangle, \quad (3.1)$$

have gained a lot of interest as they provide an exponential increase in Hilbert space dimension (d^N versus 2^N for N units), allowing for the adaption and simplification of a variety of algorithms [193, 137, 45, 16, 132]. Extensive research has been conducted for the *qutrit* case $d = 3$: While theoretical studies have shown benefits for quantum compilation [6, 7] and improved schemes for quantum error correction [86, 107, 130, 116], experimental implementations of qutrits have been demonstrated on several architectures [14, 112, 125, 62, 160, 112, 106].

3.1 Generalized qudit gates

The qudit extends the qubit abstraction to d levels as shown in Equation (3.1). This motivates the generalization of common qubit gates (see Section 2.2.2) to the qudit case, and in

the following we specifically consider the gates ($k = 0, \dots, d-1$) [65, 208]

$$\begin{aligned} X_d |k\rangle &= |k+1 \bmod d\rangle, & X_d^s : |0\rangle &\leftrightarrow |d-1\rangle, \\ H_d |k\rangle &= \frac{1}{\sqrt{d}} \sum_{j=0}^{d-1} \omega_d^{kj} |j\rangle, & Z_d |k\rangle &= \omega_d^k |k\rangle. \end{aligned} \tag{3.2}$$

Here $\omega_d = e^{2\pi i/d}$ denotes the primitive d th root of unity. X_d maps a computational basis state to the next higher one, and maps $|d-1\rangle$ back to $|0\rangle$. As an alternative generalization, X_d^s swaps the amplitudes of the lowest and the highest state. The generalized Hadamard gate H_d creates an equal superposition with different relative phases depending on the input state. In analogy to the qubit case, we also consider a generalized $T_d = Z_d^{1/4}$ gate as the fourth root of the Z_d gate. It preserves populations of the basis states but introduces phase differences.

Analogously, two-qudit gates can be generalized as well. For example, the CNOT operation can be generalized to the SUM gate [65] or other useful operations depending on use case [7, 208]. In the next section we specifically focus on the generalized qudit SWAP gate, the straightforward extension to the qubit case, which fully swaps the amplitudes of two qudits.

3.2 Time-efficient qudit gates through Incremental Pulse

Re-seeding

Harnessing qudits could allow current or near-future devices to solve larger, more useful problems sooner than with only qubits. However, qudit gate durations theoretically scale in $\mathcal{O}(d^{2N})$ time [111], which appears to limit the usefulness of higher-dimensional qudits as the benefit of a larger computational space might come at the cost of impractically long gate durations.

In the following study, we address this concern by directly optimizing qudit gates of shortest duration using the quantum optimal control (QOC) tool Juqbox [141]. We refer to Section 2.3.1 for an introduction to QOC. Like many other QOC frameworks, Juqbox is designed to find the most accurate control pulse for a fixed gate and duration. This means that finding the high-accuracy pulse of shortest duration for a given gate requires many independent optimizations at different durations. We present the Incremental Pulse Re-seeding (IPR) scheme, a method for repeatedly changing the pulse duration and using previous results as new initial guesses for the optimizer. IPR provides improved shortest-duration results and decreases the dependence on random initial guesses, which is increasingly important for larger Hilbert space dimension. Using IPR, we then study the dependence of durations of representative gates on the Hilbert space dimension for the exemplary case of a transmon platform.

3.2.1 Quantum optimal control setup

In the following we specify the QOC problem for this theoretical study. We consider a system based on superconducting hardware that consists of two weakly coupled, anharmonic transmons (see Section 2.3.2) with full Hamiltonian

$$\mathcal{H}(t) = \sum_{j=1}^2 \left[\omega_j a_j^\dagger a_j - \frac{K_j}{2} a_j^\dagger a_j^\dagger a_j a_j + f_j(t) (a_j + a_j^\dagger) \right] + g (a_1^\dagger a_2 + a_2^\dagger a_1). \quad (3.3)$$

The indices denote the different subsystems. In order to numerically model qudits, we typically truncate the ladder operators a_j and $a_j^\dagger$ at level $d+2$. Including two additional guard states allows us to penalize population leakage into higher energy states, while we have found that additional guard states beyond two provide marginal fidelity improvements at higher computational cost. We choose realistic physical parameters inspired by Ref. [169]: The 0-1 transition frequencies of the transmons are $\omega_1/(2\pi) = 4.914$ GHz and $\omega_2/(2\pi) = 5.114$ GHz,

and both transmons have the same anharmonicity $K_1/(2\pi) = K_2/(2\pi) = 330$ MHz. They are effectively coupled with $g/(2\pi) = 3.8$ MHz. While our specific results depend explicitly on these parameters, our methods apply to systems with different parameters as well. Control of our model system is possible through microwave drives f_j .

Similar to the discussion of the idealized qubit control in Section 2.3, we move to a rotating frame defined by $U_{\text{rot}}(t) = \exp\left(-i\omega_r\left(a_1^\dagger a_1 + a_2^\dagger a_2\right)t\right)$, yielding the transformed Hamiltonian

$$\begin{aligned} \tilde{\mathcal{H}}(t) = \sum_{j=1}^2 \left[(\omega_j - \omega_r) a_j^\dagger a_j - \frac{K_j}{2} a_j^\dagger a_j^\dagger a_j a_j + \frac{I_j(t)}{2} (a_j + a_j^\dagger) + i \frac{Q_j(t)}{2} (a_j - a_j^\dagger) \right] \\ + g(a_1^\dagger a_2 + a_2^\dagger a_1). \end{aligned} \quad (3.4)$$

The quadrature signals I_j and Q_j in the rotating frame are then related to the lab frame controls through $f_j(t) = \text{Re}[(I_j(t) + iQ_j(t))e^{i\omega_r t}]$ [141].

For the case that we want to describe only a single transmon, we disregard the transmon at index 2 and restrict the model to a single anharmonic oscillator with one lab frame control field $f_1(t)$.

Pulse parameterization

Juqbox [141] is an open-source software package designed to solve quantum optimal control problems in closed systems under the rotating wave approximation. The control functions $I_j(t)$ and $Q_j(t)$ are given by a sum of N_f carrier waves with fixed angular frequencies ω_{jm} , where each carrier wave has an amplitude envelope $B_s(t)$ that consists of N_s time-local

B-spline basis functions:

$$\begin{aligned}
I_j(t, \gamma) &= \sum_{m=1}^{N_f} \sum_{s=1}^{N_s} \text{Re} \left[\gamma_{jms} e^{i\omega_{jm}t} \right] B_s(t) \\
Q_j(t, \gamma) &= \sum_{m=1}^{N_f} \sum_{s=1}^{N_s} \text{Im} \left[\gamma_{jms} e^{i\omega_{jm}t} \right] B_s(t).
\end{aligned} \tag{3.5}$$

Juqbox finds optimal pulses by adjusting the $2N_f N_s$ real coefficients $\gamma_{jms} = \gamma_{jms}^{\text{Re}} + i\gamma_{jms}^{\text{Im}}$ for the quantum devices indexed by j . The collection of these design variables is denoted by γ . The benefit of this approach is that the B-spline parameterization drastically reduces the dimensionality of the optimal control problem, especially for long-duration pulses, and allows carrier wave frequencies to be chosen to specifically address the target state evolutions while avoiding undesired transitions. Figure 3.1 shows an example pulse defined by a single carrier wave and six B-splines.

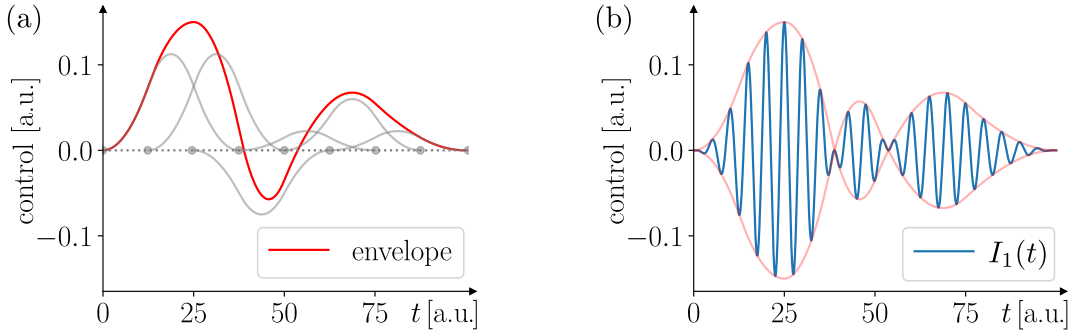


Figure 3.1: Real part of an exemplary pulse in the rotating frame, with one carrier wave ($N_f = 1$). (a) Six B-splines shape the envelope of the pulse. Grey dots mark the support of each time-local basis function. (b) The final pulse is constructed from a single carrier wave bounded by the B-spline envelope. With $N_f > 1$, the final pulse is the sum of these individual pulses with one carrier wave.

The user of this tool must specify a number of physical and numerical parameters for the optimizer, the most significant of which are summarized in Table 3.1 and discussed in the following.

After specifying the required parameters, Juqbox applies gradient-based methods to min-

Table 3.1: Important parameters for optimization in Jupyter

Symbol	Description
T	Total pulse duration
ω_r	Rotating frame frequency
$\tilde{\mathcal{H}}_0$	Drift Hamiltonian in the rotating frame
$\tilde{\mathcal{H}}_c$	Control Hamiltonian in the rotating frame
ω_{jm}	Carrier frequencies in the rotating frame
N_s	Number of B-spline basis functions per carrier wave
U_{targ}	Target unitary gate
γ_{max}	Bound on magnitude of B-splines
<code>error_threshold</code>	Goal infidelity
<code>max_iter</code>	Maximum optimization iterations

imize an objective function consisting mainly of the infidelity $\mathcal{I}(T; \gamma)$ (2.54). For a given target U_{targ} , certain values of duration T will be too small for the optimizer to find a pulse of sufficiently low infidelity, in which case a larger T is needed. In the remainder of Section 3.2, the “gate duration” for a gate U_{targ} refers to the shortest T for which the optimizer converges to our target infidelity.

Choosing optimization parameters

We observe that for high-dimensional quantum optimal control problems it is generally difficult to specify a set of optimization parameters that guarantees finding high-fidelity pulses. Parameters such as N_s , the number of B-splines per pulse, must be large enough to give the optimizer freedom, while not so large as to introduce unnecessary local minima in the highly non-convex search space. We now describe and justify our choosing of relevant optimization parameters.

The carrier frequency and B-spline envelope parameterization of the control pulse has two benefits: It reduces the dimensionality of the pulse, and additionally allows us to easily target the desired state transitions in the qudits by using carrier frequencies directly corresponding

to the qudit transition energies. In the lab frame, the $N_f = d - 1$ carrier frequencies

$$\omega_{jm}^{\text{lab}} = \omega_j - mK_j, \quad m = 0, \dots, d - 2 \quad (3.6)$$

correspond to the resonant frequencies of qudit j . For instance, $\omega_{1,2}^{\text{lab}}$ is the frequency of the $|2\rangle \rightarrow |3\rangle$ transition for qudit 1. For a two-qudit gate, each control pulse includes the complete set of $2(d - 1)$ carrier frequencies in order to make use of the cross-resonance effect [155, 194], where driving one transmon at a resonant frequency of the other transmon triggers a state transition in the latter. We choose not to correct these frequencies to account for the weak coupling between the transmons, since the small corrections fall inside the widths of the respective Fourier peaks of pulses with finite duration. From experience, we have found that considering the complete set of $(d - 1)^2$ resonant frequencies significantly increases the complexity of the problem without providing a noticeable improvement.

Given the fixed set of $2(d - 1)$ lab frame carrier frequencies ω_{jm}^{lab} , we choose the average frequency as the rotating frame frequency $\omega_r = (\omega_{\text{max}}^{\text{lab}} + \omega_{\text{min}}^{\text{lab}})/2$, which is equal for both qudits (the coupling term is then constant in time, as shown in Equation (3.4)). This minimizes the magnitude of the carrier frequencies in the rotating frame, allowing the optimizer to choose larger discrete time steps and decreasing overall optimization run time. As an example we consider two qudits: The lab frame carrier frequencies are $\omega_{jm}^{\text{lab}}/(2\pi) \in \{4.914 \text{ GHz}, 4.584 \text{ GHz}, 5.114 \text{ GHz}, 4.784 \text{ GHz}\}$, where the first two correspond to the two resonant frequencies of transmon 1 and the latter two to the resonant frequencies of transmon 2. The rotating frame frequency is $\omega_r/(2\pi) = (5.114 + 4.584)/2 \text{ GHz} = 4.849 \text{ GHz}$. Thus, the carrier frequencies in the rotating frame are given by $\omega_{jm}/2\pi = (\omega_{jm}^{\text{lab}} - \omega_r)/(2\pi) \in \{65 \text{ MHz}, -265 \text{ MHz}, 265 \text{ MHz}, -65 \text{ MHz}\}$.

The number of B-spline basis functions N_s determines how fast the pulse envelope can vary over time (see Figure 3.1). While a larger number may enable shorter-duration pulses, this will increase the dimensionality of the problem, as well as introduce additional oscilla-

tions possibly. We choose $N_s = \text{round}[T/(10 \text{ ns})] + 2$ (rounding to the nearest integer) for two reasons. First, we want to maintain an approximately constant B-spline density when varying T for consistency between different optimization problems. Second, this choice guarantees a minimum envelope rise time of around 15 ns, which is realistic in experiment [169]. The term “+2” accounts for B-splines on the time domain boundaries, which are always set to zero in Juqbox’s implementation to ensure that the final pulse starts and ends at $f_j(0) = f_j(T) = 0$.

The parameter γ_{max} limits the amplitudes of the individual B-splines, effectively limiting the maximum amplitude of the final pulses. This is needed to ensure weak driving [114], which increases the accuracy of approximate models like shown in Equation (3.4). Additionally, limited power reduces the risk of leakage into guard levels due to off-resonant transitions. We aim for lab frame pulse amplitudes of at most 40 MHz and achieve this by tuning $\gamma_{\text{max}}/(2\pi) = 40/(2\sqrt{2}N_f)$ MHz.

We set `error_threshold`, the target gate infidelity, to 10^{-3} , corresponding to a fidelity of $\mathcal{F}_{\text{targ}} = 99.9\%$. The optimization terminates if the pulse fidelity reaches this target.

Finally, the parameter `max_iter` limits the number of iterations of the optimizer if it does not terminate early due to reaching `error_threshold`. This variable must be set sufficiently large to ensure that the optimizer has converged very nearly to a local minimum, and is generally set experimentally depending on the dimension and target unitary (typically in the range 200...1000).

3.2.2 Gate duration optimization

In this section we describe our method for finding the shortest-duration, high-fidelity pulse for a given target unitary. Optimal control software generally works with a fixed pulse duration, so finding the shortest high-fidelity pulse involves repeated optimizations at different durations.

By default, each optimization begins with a random guess for the coefficient vector α within some bounds for each coefficient (typically $\pm\gamma_{\max}/100$ or $\pm\gamma_{\max}/10$). Due to the non-convex nature of the optimization space, it is generally good practice to try multiple initial guesses if the first does not work and the fidelity seems relatively close to the goal fidelity.

Naive approaches to duration optimization

We first introduce two simple approaches to finding optimal gate durations, the exhaustive method and the binary search method, and explain their downsides. The exhaustive process begins at some low $T = T_{\text{start}}$ and increases T at a fixed interval, running an independent random-guess optimization for each duration, until one value of T yields a high-fidelity pulse. While this method will work (at least for qudits of small dimension), it is computationally inefficient, motivating the second approach: a binary search. This method involves specifying $T_{\min}$ and $T_{\max}$ and then optimizing in a binary search pattern within the bounds, decreasing T if optimization succeeds and increasing if it fails (still running an independent, random-guess optimization at each duration) until reaching the specified granularity.

Both methods struggle when transitioning to high-dimensional qudit gates. Larger qudit dimension significantly increases computational cost of solving the optimization, which means the exhaustive method can potentially be very computationally expensive if the starting duration T is far below the shortest duration possible. The binary search method is more efficient, but individual long-duration optimizations near $T_{\max}$ will still be quite expensive if $T_{\max}$ is large.

However, the more significant issue is the increased complexity of the optimization space for higher qudit dimensions. We find that random guesses become less effective at consistently converging to the optimal solution for a given T , as we show in the comparison at the end of this section.

Incremental Pulse Re-seeding

We developed an optimization scheme to address both of these problems. The key idea is to reuse the coefficient vector γ from failed optimizations as a seed for the next. For example, if a pulse of duration T converges to a fidelity of 99.5%, we can reuse this pulse as the starting point for an optimization at duration $T' = T + \tau$, which we find to have a better chance of converging to 99.9% fidelity than a random guess.

Depending on the success of the optimizer to find a high-fidelity pulse for a certain duration, our algorithm changes the duration for the next iteration by a discrete time step and reuses the previous pulse (by extending or truncating it to fit the new duration) as a seed. The step size decreases as the algorithm approaches the final solution. We therefore call this scheme *Incremental Pulse Re-seeding* (IPR). The variables used in IPR are listed in Table 3.2, and a flowchart of the method is shown in Figure 3.2. We denote an optimized pulse as “successful” if it has fidelity greater than our goal fidelity, and “failed” if not.

Table 3.2: Variables used in Incremental Pulse Re-seeding

Symbol	Description
T_{start}	First duration tried by optimizer
step	Time step to increase/decrease by
granularity	Minimum step size
γ_{guess}	B-spline coefficients of initial guess pulse
error_threshold	Target infidelity

The program starts from an initial duration T_{start} with a random guess γ_{guess} . A flag variable **found_time**, which indicates if a successful pulse has been found before, is set to **false**. The optimizer then converges to a solution γ with fidelity $\mathcal{F}(T; \gamma)$. If the converged pulse is of sufficiently high fidelity, the solution is stored as γ_{best} , the duration is stored as T_{best} , and **found_time** is set to **true**. We then reduce the gate duration by the step size and re-seed the optimizer with a truncated version of γ_{best} that fits the new duration.

If the target fidelity is not met, but a successful pulse has been found before, we de-

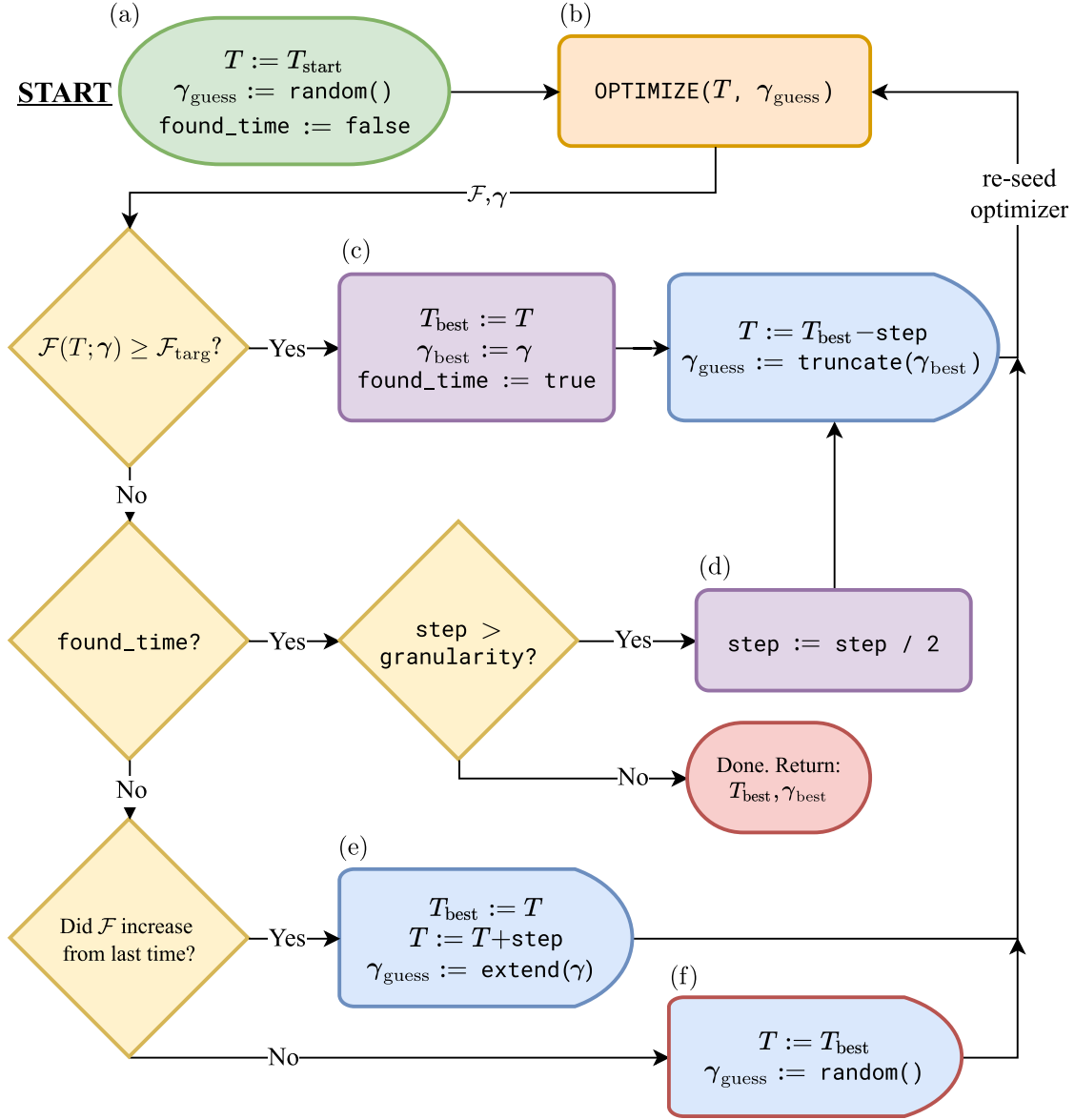


Figure 3.2: The Incremental Pulse Re-seeding process. (a) The problem is initialized with a random pulse and duration guess T_{start} . (b) The main optimization function, which searches for the highest-fidelity pulse for a given duration using Juqbox, returning an optimized pulse γ and a fidelity $\mathcal{F}(T; \gamma)$. (c) If an optimization succeeds, save the duration and pulse, then decrease T and start from a truncated version of the successful pulse. (d) If an optimization fails, but a successful pulse has been found in the past, decrease step size and work backwards from T_{best} again; if step size has reached granularity, return best pulse. (e) If no successful pulse has been found so far, but fidelity improved, mark this duration as the best so far, extend the duration, and re-seed with the extended previous pulse. (f) If no successful pulse has been found and fidelity decreased, restart the whole process with the highest-fidelity duration and a random guess.

crease the step size and repeat the above steps, starting from $T_{\text{best}} - \text{step}$ with $\gamma_{\text{guess}} = \text{truncate}(\gamma_{\text{best}})$. This only happens if the step size is larger than the predefined granularity limit (in our case 1 ns); otherwise, the algorithm terminates and returns the converged result with the shortest duration.

If the optimized pulse does not meet the target fidelity, and no successful pulse has been found before, we increase the duration and re-seed the optimizer with an extended version of the previous pulse, for as long as the fidelity increases with each optimization. If the fidelity decreases (compared to a shorter duration) before any solution is found, we start the entire procedure with a new random γ_{guess} . In this case we set $T_{\text{start}} = T_{\text{best}}$, the duration for which fidelity was highest.

Example: H_4 gate

As an example, we present an application of Incremental Pulse Re-seeding to find the optimal duration of the $d = 4$ Hadamard gate H_4 (see Equation (3.2)). Figure 3.3(a) shows the sequence of IPR steps, where the chronological order is given by the numbered arrows. Red arrows represent steps that failed to reach the target fidelity of 99.9%, while green arrows indicate success. We start at $T_{\text{start}} = 70$ ns with an initial time step size of 8 ns and a random guess for the coefficient vector γ . The fidelity of the optimized pulse falls short of the target, so IPR extends the output pulse and uses it to seed the optimizer at a longer duration (step 1). This occurs two more times (steps 2 and 3) before a pulse that reaches the target fidelity is found at $T = 94$ ns. In step 4, this solution is then truncated and used to revisit $T = 86$ ns. This time the optimization task is successful, so the duration can be further decreased (step 5), resulting in another high-fidelity pulse at 78 ns. $T = 70$ ns fails to converge to the target fidelity, so the time step is halved and IPR tries 74 ns instead (step 7). After few more steps, the minimum step granularity is reached, and $T_{\text{best}} = 76$ ns is found to be the shortest duration to realize the H_4 gate with 99.9% fidelity. It is important to

note that durations 78 ns and 86 ns both initially resulted in failed optimizations, but then succeeded when re-seeded with truncated successful pulses.

With the help of Figure 3.3(b), we take a closer look at steps 3 and 4 to explain how pulses are extended and truncated. Initially (plot A), we have an optimized 86 ns pulse that falls short of the target fidelity. The controls I_1 and Q_1 are numerically extended by appending 8 ns of idle time (zero amplitude pulse). The number of B-splines N_s changes due to the duration change, so we obtain the new coefficient vector γ_{guess} by applying a least-squares fit according to Equation (3.5) with the new number of B-splines. The approximation introduces small distortions, as can be seen in plot A*, because the manually modified control functions are not perfectly realizable with the given parameterization. The following optimization (plot B) meets the target fidelity, so the pulse is truncated to 86 ns and re-parameterized (plot B*). Plot C shows the converged pulse after step 4, which successfully reaches 99.9% fidelity. Together, the series of plots (where A and C are both at duration 86 ns) shows how a new minimum was found after visiting 94 ns and then revisiting 86 ns with a better guess.

Figure 3.4 visualizes the evolution of the basis state populations for the shortest pulse of 76 ns. It is clearly visible how an equal-population superposition emerges in every case, and the small imperfections at the final time come from the remaining 0.1% infidelity. The populations of the guard states cannot be seen because they are suppressed by multiple orders of magnitude. This guard state suppression is achieved by the carrier wave parameterization method, which allows us to selectively induce state transitions.

Comparison with random guessing

For the X_8 gate (see Equation (3.2)), the IPR scheme finds a shortest-duration pulse of 195 ns with target fidelity 99.9%. To compare our approach with a naive approach, we run 20 random-guess, single-duration Jupyter optimizations at every 5 ns between 180 ns and

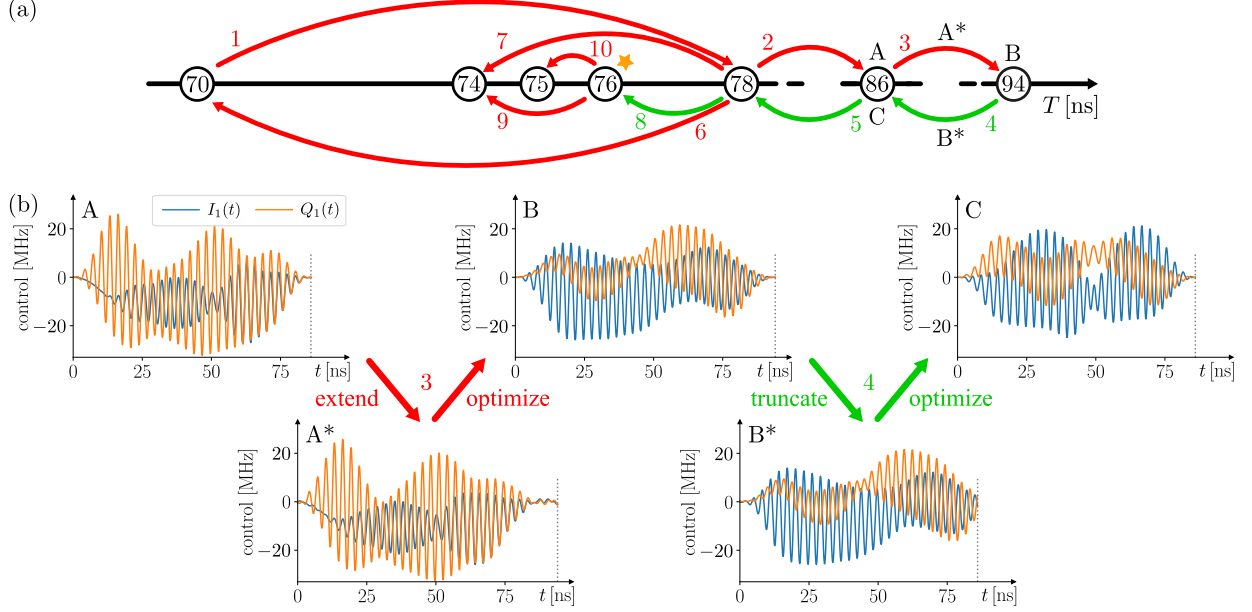


Figure 3.3: Applying Incremental Pulse Re-seeding (IPR) to find the shortest duration for the H_4 gate. (a) Visualization of attempted durations during the IPR procedure, with the chronological order of steps indicated by the colored numbers. Green (red) arrows represent steps that did (did not) lead to meeting the pulse target fidelity of $\mathcal{F}_{\text{targ}} = 99.9\%$. Starting from $T_{\text{start}} = 70$ ns and a time step of 8 ns, three pulse extension and re-optimization steps are required before the target fidelity is first reached at 94 ns. The duration is then reduced to 78 ns by successfully re-seeding with truncated pulses. After step 6 the time step is halved because re-optimizing at $T = 70$ ns with the truncated pulse does not meet the target fidelity. Repeating this scheme leads to the final duration of $T_{\text{best}} = 76$ ns, highlighted by the star, as after step 10 the time step falls below the granularity of 1 ns. (b) Plot A shows the optimized pulse after IPR step 2. A new seed for the optimization at $T = 94$ ns is generated by extending the previous result in duration (A^*). At 94 ns the target fidelity is met, so the optimized pulse is truncated (B^*) and used as a seed for the second $T = 84$ ns optimization, leading to the successful result shown in C.

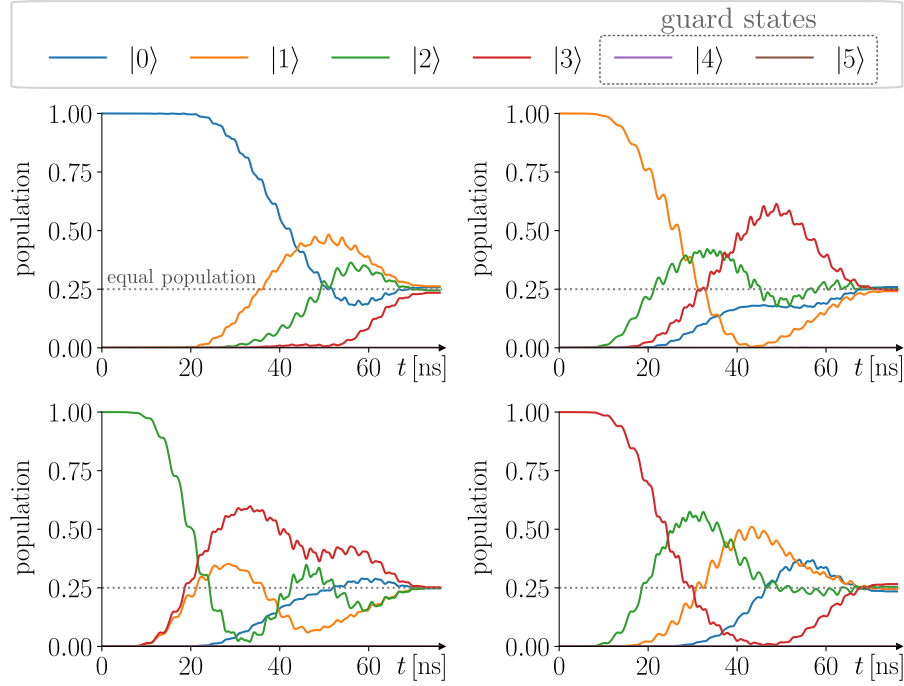


Figure 3.4: Evolution of state populations when applying the H_4 pulse of shortest duration (76 ns) to a qudit with $d = 4$. Every computational basis state is transformed into superposition of all four basis states with nearly equal population, deviations only arising from the remaining 0.1% infidelity. At all times the populations of the guard states $|4\rangle$ and $|5\rangle$ are below $2 \cdot 10^{-3}$ and 10^{-6} , respectively.

230 ns, where the components of γ_{guess} are sampled from $[-0.1\gamma_{\text{max}}, 0.1\gamma_{\text{max}}]$. Figure 3.5 shows the distribution of pulse fidelities achieved by this naive random guessing across a range of durations for the same gate, with 20 random attempts at each duration.

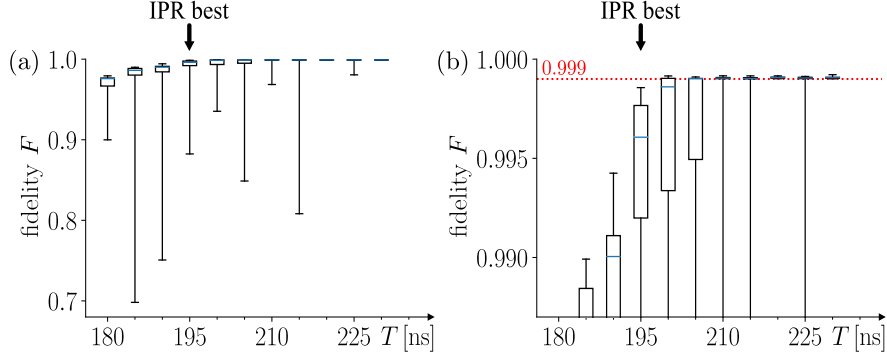


Figure 3.5: Distribution of naive random-guess optimizations of the X_8 gate for individual durations between 180 ns and 230 ns. Each boxplot consists of 20 fidelities from 20 individual Juqbox optimizations, with the blue line indicating the median value, the box indicating the middle 50% of fidelities, and the whiskers indicating the full range of values. (a) The full data, showing a general increase in fidelity as duration increases. (b) Focusing on the area of interest (fidelities close to 99.9%). The IPR scheme outperforms random guessing, finding a minimum duration of $T_{\text{best}} = 195$ ns at 99.9% fidelity, whereas none of the 20 random guesses at that duration reached the same fidelity.

It is evident that there is significant variation in the naive approach; no random-guess optimizations converged to 99.9 fidelity at 195 ns, and only 8 out of 20 random guesses converged to 99.9% fidelity at 200 ns. Even for larger durations such as 215 ns, not all guesses converged to the target fidelity. This is clearly a problem for binary-search-based methods, which implicitly assume that a given duration will either always succeed or never succeed. In contrast, over ten IPR runs (with T_{start} sampled from [160 ns, 240 ns]), we find an average duration of 198.6 ns and a standard deviation of 3.0 ns. This small standard deviation given the large range of starting times is evidence that our method is more stable than random guessing for large Hilbert space optimizations.

3.2.3 Qudit pulse results

Previous work [6, 7] has shown that breaking the qubit abstraction in quantum circuits can lead to substantially reduced circuit complexity involving fewer ancillae, making qudits promising candidates for quantum computation. However, these improvements can only be practical if the gate duration overhead does not scale poorly with the Hilbert space dimension while still maintaining high gate fidelity. In this section we investigate this relation by explicitly constructing short duration pulses of at least 99.9% fidelity for one- and two-qudit gates using Incremental Pulse Re-seeding (IPR).

Single-qudit gates

We consider the single-qudit gates X_d , X_d^s , H_d presented in Equation (3.2) as well as $T_d = Z_d^{1/4}$ for dimensions $2 \leq d \leq 8$. In each case we include two guard levels to account for and suppress leakage into higher energy states. Expecting short gate durations, we choose $T_{\text{start}} = 50$ ns for the qubit and qutrit ($d = 3$) cases, and use the resulting shortest durations to estimate T_{start} for higher d with a linear extrapolation. We set `step` to the power of two nearest to $0.1 T_{\text{start}}$ because this results in integer durations. Having found durations T_{best} for every dimension, we then run IPR repeatedly with starting durations randomly chosen from $[0.8 T_{\text{best}}, 1.2 T_{\text{best}}]$ to identify other local minima and reduce the gate time further, as well as to measure the uncertainty of the method.

Figure 3.6 shows the minimum single-qudit gate durations found over ten IPR runs for each dimension. The durations for the generalized gates X_d , X_d^s and the Hadamard gate, H_d , scale similarly with the qudit dimension. We suspect that the generalized T_d gate can be realized in lower time at higher dimension compared to the others because it only causes a change in phase, not in state populations. For this gate, we observe that the best pulse found for $d = 7$ is of shorter duration than the $d = 6$ pulse, when the reverse is intuitively expected. This discrepancy suggests shorter pulses may exist and indicates a

highly complicated optimization space.

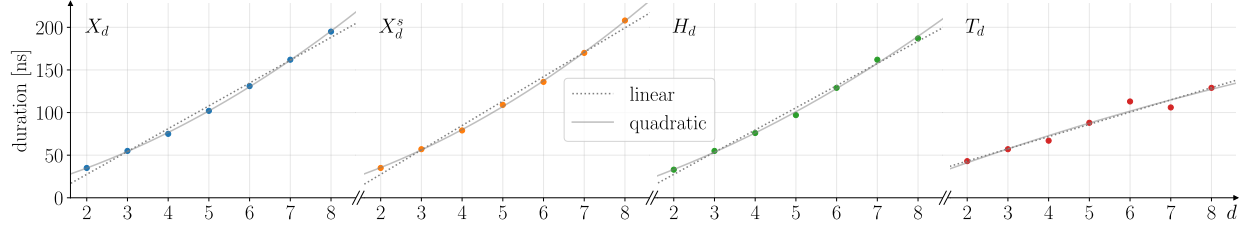


Figure 3.6: Dependence of single-qudit gate durations from qudit dimension d . The four gates are defined in Equation (3.2). Each data point shows the minimum duration of applying Incremental Pulse Re-seeding (IPR) from ten different starting times. The unexpected discrepancy between $d = 6$ and $d = 7$ for the generalized T_d gate suggests a highly complicated optimization space and shows that shorter pulses for some d values may be possible. For each gate the scaling of duration with dimension d is analyzed with quadratic and linear regressions, where the latter is less accurate but captures the essential behavior in this low-dimensional regime. The fit coefficients are presented in Table 3.3.

Theoretical studies [111] have shown that optimal gate durations scale asymptotically with $\mathcal{O}(d^{2N})$. Here we have $N = 1$, therefore we fit each sequence of data points with a quadratic polynomial $T_{\text{quad}}(d) = ad^2 + bd + c$ (solid gray lines) to compare our findings with the theory. Furthermore, the layout of the data also motivates linear regression, so we investigate the performance of a linear model $T_{\text{lin}}(d) = bd + c$ (dotted gray lines) as well. The fit coefficients and their standard deviation estimates are given in Table 3.3, where for each gate the first line corresponds to the linear fit and the second line to the quadratic fit. The quality of each regression is quantified with the coefficient of determination R^2 .

Table 3.3: Linear and quadratic fit coefficients for single-qudit gate durations

Gate	a [ns]	b [ns]	c [ns]	R^2
X_d	1.48 ± 0.12	26.79 ± 1.16	-26.07 ± 6.23	0.991
		12.02 ± 1.18	4.93 ± 2.67	1.000
X_d^s	1.60 ± 0.16	28.64 ± 1.26	-29.79 ± 6.78	0.990
		12.69 ± 1.60	3.71 ± 3.63	1.000
H_d	1.15 ± 0.36	26.04 ± 1.05	-24.61 ± 5.68	0.992
		14.49 ± 3.66	-0.36 ± 8.28	0.998
T_d	-0.38 ± 0.86	14.36 ± 1.37	14.36 ± 7.36	0.957
		18.17 ± 8.74	6.36 ± 19.78	0.959

The near-ideal R^2 values for quadratic fits to X_d , X_d^s and H_d indicate that the quadratic models describe the empirical scaling behaviors for these gates very well, which agrees with the theoretical predictions. However, we emphasize that linear scaling models capture the essential behavior in this low-dimensional regime (compared to asymptotic considerations) in good approximation too, deviating from the duration points by less than 5 ns on average for those three gates. Barely any difference between quadratic and linear regression is noticeable for the T_d gate, but the more scattered data points lead to a reduced R^2 value and larger uncertainty. Additionally, the fit parameter $a < 0$ indicates an unrealistic scaling of pulse duration because this leads to negative durations in the limit of large d . T_d is a good case study in the interplay between theoretical expectations and empirical realizations. Specifically, the poor fit obtained here indicates that for many values of d , these gate durations may overestimate the optimal gate times.

This result of essentially linear scaling over the practical range of qudit dimensions has promising implications for the current era where qudit experiments have been conducted successfully [14, 156, 80, 112, 125, 62, 160, 112, 106]. It suggests that the computational advantage single-qudit gates can provide is not outweighed by pulse time overhead.

Two-qudit SWAP gate

For the two-qudit case we consider the generalized SWAP gate, which fully swaps the state amplitudes between the qudits. In many architectures, this is an important operation for communication, as limited connectivity between devices requires quantum information to be moved around repeatedly. We apply IPR to find SWAP durations for two qudits of dimensions 2, 3, and 4, with corresponding Hilbert space dimensions 4, 9, and 16, respectively. We limit our analysis to these three cases because simulating the evolution of two qudits with $d \geq 5$ (plus guard states) becomes very resource-expensive. We use two guard states for dimensions 2 and 3 and only one guard state for 4 due to the large Hilbert space. As in the

single-qudit case, each data point is the minimum duration from ten IPR runs with different T_{start} values.

Figure 3.7 shows shortest-duration results for the SWAP gate with respect to the Hilbert space dimension d^2 . Given the small number of points, we do not attempt a fit, but note that the points appear nearly linear, as was the case with the single-qudit gates. It is worth noting that fully swapping the states of two ququarts ($d = 4$) is slightly less expensive (1128 ns) than performing two sequential SWAP operations on qubits ($581 \text{ ns} + 581 \text{ ns} = 1162 \text{ ns}$). Given that a single ququart has the same information capacity as two qubits, this represents an efficient way of moving quantum information. Additionally, given that the $d = 4$ SWAP optimization space is much more complex than the qubit case, we expect that this duration could be further decreased through more sampling or fine-tuning.

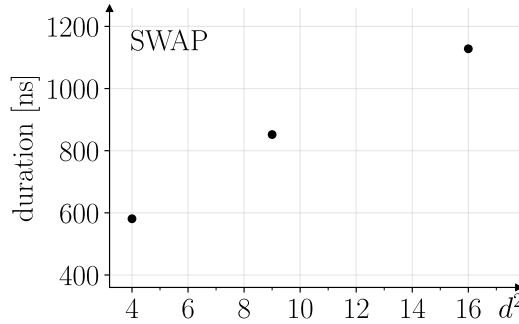


Figure 3.7: Gate durations of two-qudit SWAP operations for qudit dimensions $d = 2, 3, 4$. The horizontal axis corresponds to the dimension of the full Hilbert space d^2 . Due to the low number of data points, regression is not applied, but a near-linear trend is visible.

3.2.4 Conclusion

In this study we presented the Incremental Pulse Re-seeding (IPR) scheme, an efficient method for finding short-duration pulses for quantum optimal control tasks, which performs better than random-guess based methods for large Hilbert space gates. We used this method together with the optimal control software Jupyter to demonstrate that single-qudit gate duration scaling is nearly linear for several useful gates, with dimension up to $d = 8$. We

also show that the two-qudit SWAP gate appears to follow a similar trend for Hilbert space dimensions up to $d^2 = 16$.

We emphasize that while Incremental Pulse Re-seeding helps to make intelligent initial pulse guesses for the optimizer during intermediate steps, it does not eliminate the possibility of converging to a local infidelity minimum. As quantum optimal control problems are typically underconstrained, different initial guesses can converge to different solutions. In our case, this means starting IPR with different values for T_{start} and γ_{guess} may result in slightly different gate durations. However, as an example, our simulations show a relatively small variation in gate duration of 4.69 ns on average for single-qudit gates, underlining the advantage of our algorithm.

Our findings in this section prompt several directions of future research. For instance, the efficiency of qudit gates motivates further exploration into quantum circuits that explicitly use qudits for computation, such as the asymptotic ancilla reductions achieved in Refs. [6, 7]. Furthermore, it justifies and encourages the investigation of quantum error correction schemes based on qudits as we move towards fault-tolerant quantum computing.

We specifically considered the generalized SWAP two-qudit gate, which appears to scale nearly linearly for small qudit dimension. This result encourages analysis of duration scaling for more two-qudit gates, such as the SUM gate (a generalization of CNOT) from Ref. [65].

The presented results are based on a system Hamiltonian modeling superconducting transmons. We emphasize that our approach is not limited to this specific case and could be directly applied to optimal control problems across a wider class of quantum systems.

Our optimizations use an approximate Hamiltonian in a closed system (governed by the Schrodinger equation). For a more accurate measure of gate duration scaling, similar studies are needed in an open system, where errors can be modeled more realistically. This can be achieved by describing the dynamics with the GKSL master equation [108, 61]. In addition, benchmarking the performance of qudit pulses on actual noisy quantum architectures would

help to investigate the accuracy of models employed. Optimizing control pulses to specifically mitigate the increased sources of errors from operating on higher energy levels may lead to making qudit computation practical.

3.3 Exploring ququart computation on a transmon using optimal control

Given the encouraging results of the theoretical qudit gate study discussed in Section 3.2, we now address the last-mentioned future direction and ask how well qudit optimal control pulses perform on actual quantum hardware. To this end we collaborate with the experimental research group around David Schuster (formerly University of Chicago, now Stanford University).

We are specifically interested in operating a ququart ($d = 4$) as it can be viewed from two different perspectives: the four-level extension of the qubit as well as the encoding of two qubit’s worth of information. Theoretical works have looked at this encoding of qubit pairs into ququarts to study advantages for quantum circuit compilation [109, 110]. Increased hardware utilization, efficient internal two-qubit gates and reduced routing costs render this design highly promising. Recent experimental results have shown the efficient realization of a variational quantum algorithm [23] and the verification of the entropic inequality [49] under this scheme.

In the following, we discuss the realization of single-ququart operations on a superconducting transmon under both the qudit perspective and the encoded-qubit perspective. We compare two approaches for each gate: an optimized decomposition into natively supported gates on our system and direct implementation through quantum optimal control. Using the frameworks Jqbox [141] and Boulder Opal [151], we find optimal control pulses that drive all three transitions in parallel while respecting hardware constraints.

3.3.1 Hardware specifications

Before we discuss the pulse design and experimental analysis, we start by briefly summarizing the hardware parameters and fabrication process like it was described to us by our collaborators. While this is rather involved and not critical to understanding the results of this research project, it provides a nice view behind the scenes of an experimental lab and gives a taste of what engineering work at the very bottom of the quantum computing stack looks like.

We use the two-transmon processor presented in Ref. [160, 106], which is built from two superconducting transmons (see Section 2.3.2) with a flux-tunable coupler, and focus solely on one of the transmons (left device in Figure 3.8 below, marked in red). The other transmon remains in the ground state and can be disregarded when the coupler is not modulated and biased at the sweet spot. We use the lowest four energy levels $\{|0\rangle, |1\rangle, |2\rangle, |3\rangle\}$ of our superconducting transmon, shown in Figure 2.2(b), to represent the ququart.

The substrate used for the device fabrication is a $430\text{ }\mu\text{m}$ thick C-plane sapphire that was annealed at $1200\text{ }^\circ\text{C}$ for two hours. 200 nm thick Tantalum film was sputtered at $800\text{ }^\circ\text{C}$. Large features are written with optical lithography using Heidelberg MLA 150 Aligner and wet-etched using Transene Tantalum etchant 111. Ebeam lithography with Ratih EBPG5000 Plus E-Beam writer was used to create the junction mask with a double-layer resist consisting of MMA EL11-950 and PMMA A7. The Dolan-bridge junctions were evaporated in the Plassys electron beam evaporator through double-angle evaporation. $7 \times 7\text{ mm}$ chips are diced and lifted off. The resistance of the on-chip test junctions are measured at room temperature to help pre-selection. The selected chip was mounted on a printed circuit board, wire-bounded, and mounted inside a double-shielded μ -metal can.

The coherence times and frequency parameters are listed in Table 3.4. We note that we observe a larger T_1 relaxation time for the $|2\rangle \rightarrow |1\rangle$ transition than for the $|1\rangle \rightarrow |0\rangle$ transition, which seems to contradict the expected scaling $T_1[|j\rangle \rightarrow |j-1\rangle] \sim T_1/j$ [14]. This

is due to our choice of centering the filter stopbands around the $|1\rangle \leftrightarrow |2\rangle$ transition, which leads to less protection for the other transitions. Figure 3.8 shows our device’s false-colored optical image. Unused charge lines coupling to the transmon pad are visible, which cause Purcell-limited relaxation times between qudit levels. The cryogenic and room temperature measurement setup is illustrated in Figure 3.9.

Transitions	$ 0\rangle \leftrightarrow 1\rangle$	$ 1\rangle \leftrightarrow 2\rangle$	$ 2\rangle \leftrightarrow 3\rangle$
T_1 (μs)	22.62	25.96	10.19
T_R (μs)	40.60	42.32	2.68
Frequency $\omega_j/(2\pi)$ (GHz)	3.2222	3.1021	2.9717
Drive response r_j (MHz/AWG)	5.5528	5.6712	1.6749

Table 3.4: Ququart characterization. Relaxation times T_1 and Ramsey dephasing times T_R between neighboring energy levels $|j-1\rangle \leftrightarrow |j\rangle$ ($j = 1, 2, 3$) are listed, as well as transition frequencies ω_j and on-chip drive responses r_j to unit power AWG driving.

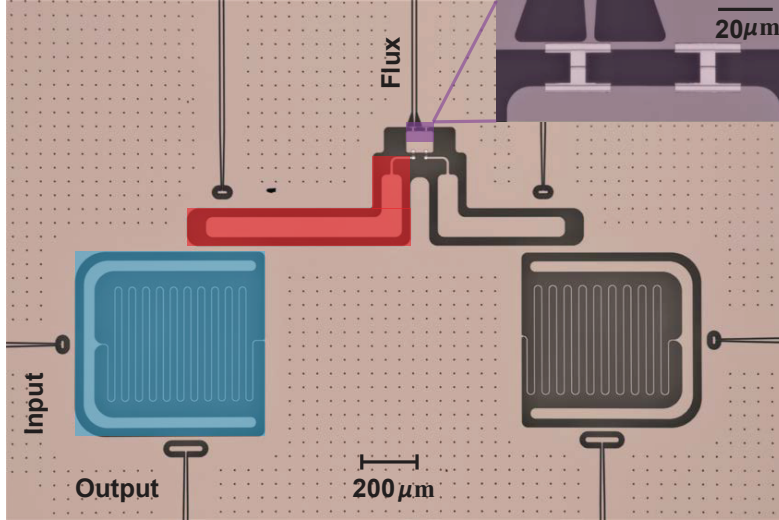


Figure 3.8: False-colored optical image of the device. The left transmon (red) is used as the ququart in the experiment. Single-ququart rotations are sent through the input line coupled to the resonator (blue). The SQUID coupler (purple, zoomed-in region) is kept biased to the flux sweet spot throughout the experiment.

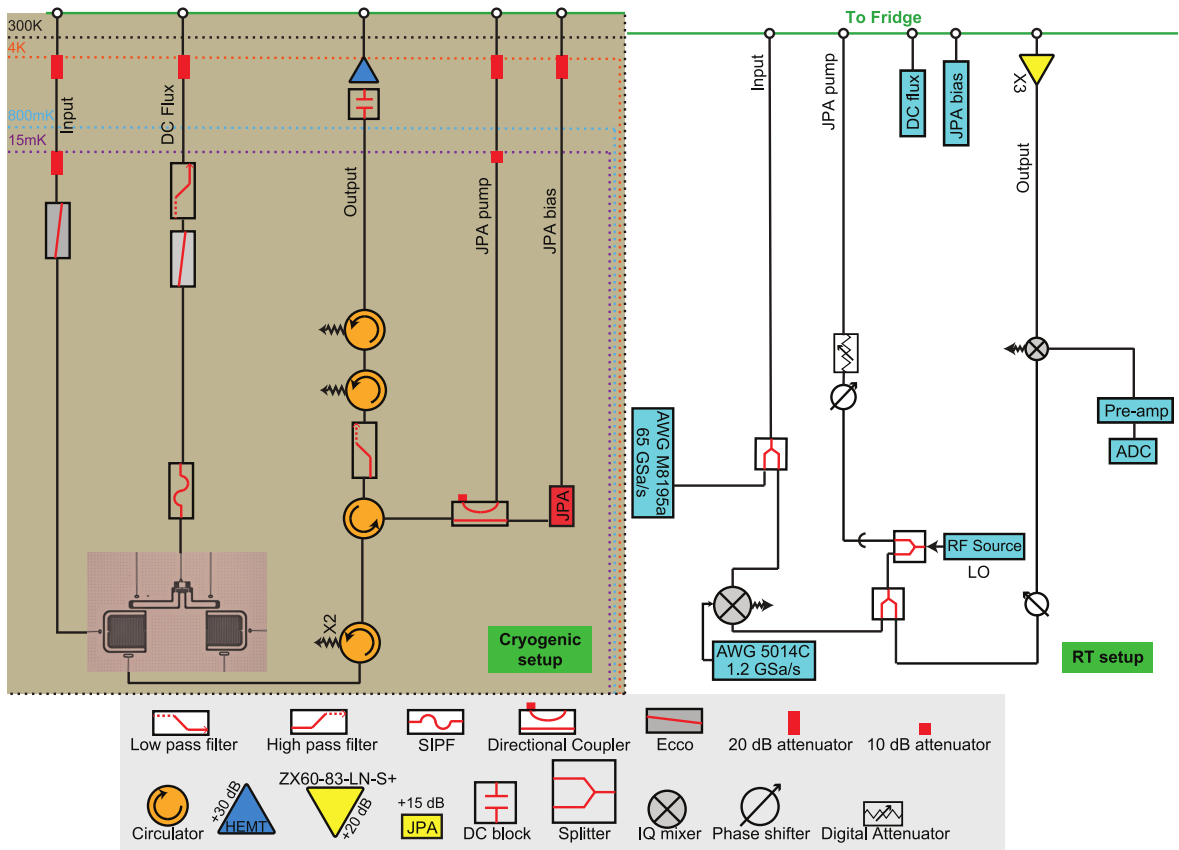


Figure 3.9: Measurement setup.

AWG drive calibration and native gates

In experiments, all components in the measurement and control chain have their unique frequency response matrix. To map the actual ququart drive strength observed in the experiment to the drive strength specified by the input signal, we need to extract the frequency response matrix between the Arbitrary Waveform Generator (AWG) and the transmon around the frequency band used in the control pulse. In our case, we only need to consider the frequencies around the three transitions $|0\rangle \leftrightarrow |1\rangle$, $|1\rangle \leftrightarrow |2\rangle$, and $|2\rangle \leftrightarrow |3\rangle$ each within a 25 MHz frequency span. We further assume that under such a narrow span the frequency response matrix is constant for each transition $|j-1\rangle \leftrightarrow |j\rangle$, which allows us to directly extract the on-chip drive response factors r_j . They describe the experimentally measured effective drive strength at the maximum AWG output power for each transition, determined from the observed Rabi rate. For arbitrary transition rates, as they appear in the optimal control pulses, the corresponding AWG output power can be scaled accordingly. The unit of these parameters is $[r_j] = \text{MHz}/\text{AWG}$, where we overload the abbreviation AWG to also refer to the amplitude unit of the waveform generator. We include the calibrated r_j values in Table 3.4.

We note that we do not observe the expected scaling $r_j \sim \sqrt{j}$, which we suspect is due to the filter choices and undesired couplings of the input line to external modes on the chip. Especially for the $|2\rangle \leftrightarrow |3\rangle$ transition, we suspect that the unused control lines, which are directly coupled to the transmon pads, cause a stray on-chip mode which blocks this input signal. We do not study this more within the scope of this work.

3.3.2 One ququart or two qubits

The concept of the ququart extends the computational unit of a qubit by two levels and describes the special qudit case $d = 4$. Recalling the generalized qudit gates from Equa-

tion (3.2), some special ququart gates are given by

$$X_4 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad H_4 = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{bmatrix}, \quad Z_4 = \text{diag}(1, i, -1, -i). \quad (3.7)$$

Due to the special Hilbert space dimension $d = 4 = 2^2$, another way to think about a ququart is as a pair of qubits. One possible encoding is given by mapping the qubit bitstrings to their corresponding decimal numbers:

$$|q_0 q_1\rangle = \left\{ \begin{array}{l} |00\rangle \leftrightarrow |0\rangle \\ |01\rangle \leftrightarrow |1\rangle \\ |10\rangle \leftrightarrow |2\rangle \\ |11\rangle \leftrightarrow |3\rangle \end{array} \right\} = |q^{(4)}\rangle. \quad (3.8)$$

Therefore, two-qubit operators can be directly translated to ququart gates given their matrix representation. For example, simultaneous bit flips on q_0 and q_1 described by $X \otimes X$ can be implemented by flipping $|0\rangle \leftrightarrow |3\rangle$ and $|1\rangle \leftrightarrow |2\rangle$. Furthermore, such an encoding provides efficient qubit CX and SWAP gates through state flips $|2\rangle \leftrightarrow |3\rangle$ and $|1\rangle \leftrightarrow |2\rangle$, respectively. Past theoretical and experimental works have studied qubit-based computation under this scheme to find potential advantages in resource requirements and circuit speed [109, 110, 23].

3.3.3 Numerical gate decomposition

The elementary gates available in the transmon are qubit-like rotations $R_{j-1,j}(\theta, \phi) = Z_{j-1,j}^\dagger(\phi)Y_{j-1,j}(\theta)Z_{j-1,j}(\phi)$ between two adjacent levels $|j-1\rangle$ and $|j\rangle$, with $j \in \{1, 2, 3\}$:

$$R_{01}(\theta, \phi) = \begin{bmatrix} \cos \frac{\theta}{2} & -e^{i\phi} \sin \frac{\theta}{2} & 0 & 0 \\ e^{-i\phi} \sin \frac{\theta}{2} & \cos \frac{\theta}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (3.9a)$$

$$R_{12}(\theta, \phi) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \frac{\theta}{2} & -e^{i\phi} \sin \frac{\theta}{2} & 0 \\ 0 & e^{-i\phi} \sin \frac{\theta}{2} & \cos \frac{\theta}{2} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (3.9b)$$

$$R_{23}(\theta, \phi) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos \frac{\theta}{2} & -e^{i\phi} \sin \frac{\theta}{2} \\ 0 & 0 & e^{-i\phi} \sin \frac{\theta}{2} & \cos \frac{\theta}{2} \end{bmatrix}. \quad (3.9c)$$

The $Y_{j-1,j}(\theta)$ rotations are implemented using microwave pulses resonant with the $|j-1\rangle \leftrightarrow |j\rangle$ transitions, where the pulse shape is a Gaussian flat-top pulse with a $2\sigma = 5$ ns ramp and a $2\sigma = 5$ ns tail, where σ is the standard deviation of the Gaussian edge. Arbitrary rotation angles θ are achieved by driving at the power limit of the waveform generator (1 AWG) and tuning the duration of the constant part τ_j of the flat-top section. The overall gate duration $T_j(\theta) = \tau_j(\theta) + 4\sigma$ is then given by:

$$T_j(\theta) = \frac{\theta}{2\pi\sqrt{j}r_j \cdot 1 \text{ AWG}} - \sqrt{2\pi}\text{erf}\left(\sqrt{2}\right)\sigma + 4\sigma \approx \frac{\theta}{2\pi\sqrt{j}r_j \cdot 1 \text{ AWG}} + 4 \text{ ns}. \quad (3.10)$$

The phase rotations $Z_{j-1,j}(\phi)$ are realized in software (known as “virtual $Z(\phi)$,” we also use the abbreviation VZ) by appropriately updating the phase of the respective carrier wave [119] and are therefore near-perfect. The concept of VZ gates can be generalized to the ququart case, allowing the direct implementation of relative phase rotations $Z_4(\phi) =$

$\text{diag}\left(1, e^{i\phi_1}, e^{i(\phi_1+\phi_2)}, e^{i(\phi_1+\phi_2+\phi_3)}\right)$ by updating all three carrier waves at once, where $\phi = (\phi_1, \phi_2, \phi_3)$ and ϕ_j denotes the phase shift to the j th microwave drive.

As derived in Ref. [48] and shown in Figure 3.10(a) at the end of this section, every single-ququart gate U can be decomposed in the following way:

$$U(\theta_i, \phi_i) = Z_4(\phi_4)Y_{01}(\theta_6)Y_{12}(\theta_5)Y_{23}(\theta_4)Z_4(\phi_3)Y_{01}(\theta_3)Y_{12}(\theta_2)Z_4(\phi_2)Y_{01}(\theta_1)Z_4(\phi_1). \quad (3.11)$$

This decomposition holds up to a global phase, which is inconsequential in the experiment. While at first, it looks like $4 \cdot 3$ phases + 6 angles = 18 real degrees of freedom, which contradicts the 15 real degrees of freedom for an arbitrary ququart gate, it is actually the case that one can fix $\phi_{1,2} = \phi_{1,3} = \phi_{2,3} = 0$ [48], leaving the right number of parameters. We changed the indexing of the subspaces compared to Ref. [48] such that the decomposition uses the fewest gates in the third subspace. This was motivated by the fact that we experimentally observed a significantly reduced achievable Rabi rate in this subspace as discussed in Section 3.3.1, which would likely cause much slower gate implementations when using the original decomposition under the same experimental constraints. We emphasize that Equation (3.11) represents the most general decomposition and in special cases not all gates might be needed. For instance, we have $X_4 = Z_4(\pi, 0, 0)Y_{01}(\pi)Y_{12}(\pi)Y_{23}(\pi)$ but H_4 requires all gates as shown in Table 3.5.

Given a target unitary U_{targ} , we determine the decomposition parameters $\theta_i \in [0, \pi]$ and $\phi_i \in [0, 2\pi)^3$ via numerical optimization, where we primarily minimize the gate infidelity (recall Equation (2.51))

$$\mathcal{I}(\theta_i, \phi_i) = 1 - \frac{1}{d^2} \left| \text{Tr}\left(U_{\text{targ}}^\dagger U(\theta_i, \phi_i)\right) \right|^2 \quad (3.12)$$

between the target and the decomposition $U(\theta_i, \phi_i)$. We speak of a *feasible decomposition* if $\mathcal{I}(\theta_i, \phi_i) \leq 10^{-8}$.

For a specific target, there may be many feasible decompositions, however, not all of them need to be reasonable with respect to experimental implementation. For instance, consider the identity operation I_4 on a ququart: $I_4 = Y_{01}(\pi)Y_{01}(\pi)$ would be compatible with the general formulation (3.11) but unnecessarily expensive to implement. Therefore, we also want to minimize the estimated overall gate duration

$$T^*(\theta_i) = \sum_{i=1}^6 \frac{\theta_i}{2\pi\sqrt{j} r_{s(i)} \cdot 1 \text{ AWG}}, \quad (3.13)$$

which is determined by the rotation angles. For simplicity, we assume rectangular pulse shapes (no Gaussian edges) in this model. The function s maps the angle index i to the corresponding subspace index j .

However, this approach is not ideal because it could lead to the unnecessary splitting of operations. For instance, if $U_{\text{targ}} = Y_{01}(\pi)$, feasible decompositions would be $U(\theta_i, \phi_i) = Y_{01}(\pi)$ (ideal) or $U(\theta_i, \phi_i) = Y_{01}(\frac{\pi}{3})Y_{01}(\frac{\pi}{3})Y_{01}(\frac{\pi}{3})$ (not ideal). Therefore, we actually use a different penalty of the form

$$\tilde{T}(\theta_i) = \sum_{i=1}^6 \sqrt{\frac{\theta_i}{2\pi\sqrt{j} r_{s(i)} \cdot 1 \text{ AWG}}}. \quad (3.14)$$

This makes use of the fact that, for $a, b \geq 0$, $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$, with equality if and only if a or b equals zero. Thus splitting of native rotations is discouraged. Other than that $\tilde{T}$ is still monotonically increasing in each θ_i and minimizing $\tilde{T}$ also leads to reducing T^* .

Additionally, it may be the case that (infinitely) many sets of phase angles ϕ_i lead to feasible decompositions, even for optimal rotation angles. In order to reduce this degeneracy and furthermore get a better intuition of the action for the gate sequence, we steer the optimization towards phase angles that are multiples of $\frac{\pi}{2}$. In conjunction with $Y_{j-1,j}(\theta)$ gates, these special phase shifts effectively lead to rotations about the X -axis and Y -axis in

the j th subspace's Bloch sphere. We achieve this by introducing another cost term

$$P(\phi_i) = \frac{1}{12} \sum_{i=1}^4 \sum_{j=1}^3 \sin^2(2\phi_{i,j}). \quad (3.15)$$

Overall, for a given target U_{targ} we solve the minimization problem

$$\hat{\theta}_i, \hat{\phi}_i = \arg \min_{\theta_i, \phi_i} \left(\mathcal{I}(\theta_i, \phi_i) + c_{\tilde{T}} \tilde{T}(\theta_i) + c_P P(\phi_i) \right) \quad (3.16)$$

to find a feasible decomposition. We optimize a batch of ten initial guesses for the optimization variables and check if the best-obtained infidelity is below the desired threshold; if not, we keep optimizing batches until we do. For a suitable choice of penalty coefficients $c_{\tilde{T}}$ and c_P this method is guaranteed to succeed eventually since any unitary can be realized with decomposition (3.11). We set $c_{\tilde{T}} = 0.2$ to strongly encourage fast sequences and $c_P = 0.05$ to gently push the optimizer towards preferred phase values, and find this setup to work consistently well in practice.

We show the decomposition for the ququart Hadamard H_4 as well as several gates acting on two qubits in Table 3.5 and further list the execution time computed by summing up the individual pulse times according to Equation (3.10), rounded to full ns. The numerically obtained parameter values suggest that they approximate multiples of $\frac{\pi}{2}$ or more complicated analytical expressions: For example, for the gate H_4 we find $\theta_1 \approx 2.09440$, which has the property $e^{i\theta_1/2} \approx \frac{1+\sqrt{3}i}{2}$. This analytical nature is reasonable given the structure of the gates we consider to factor unitaries, which motivates studying qudit gate decompositions further and finding better ways to obtain these expressions directly.

3.3.4 Quantum optimal control

We design open-loop control pulses using Juqbox [141] by LLNL as well as the proprietary tool Boulder Opal [151] from Q-CTRL. We model our superconducting transmon with the

Gate	$\theta_{1,2,3}$	$\theta_{4,5,6}$	ϕ_1	ϕ_2	ϕ_3	ϕ_4	$T[\text{ns}]$
H_4	$\frac{\pi}{2}$	$\tilde{\theta}_1$	$\frac{3}{2}\pi$	$\tilde{\phi}_1$	$\tilde{\phi}_1$	$\frac{\pi}{2}$	343
	$\tilde{\theta}_2$	$\tilde{\theta}_2$	0	$\frac{3}{2}\pi$	$\tilde{\phi}_2$	$\frac{\pi}{2}$	
	$\tilde{\theta}_3$	$\frac{\pi}{2}$	0	0	$\frac{3}{2}\pi$	$\frac{\pi}{2}$	
$H \otimes H$	$\frac{\pi}{2}$	$\tilde{\theta}_1$	π	π	π	0	365
	$\tilde{\theta}_2$	$\tilde{\theta}_2$	0	0	π	π	
	$\tilde{\theta}_1$	$\frac{\pi}{2}$	0	0	π	0	
$I \otimes H$	$\frac{\pi}{2}$	$\frac{\pi}{2}$	π	0	$\frac{\pi}{2}$	$\frac{3}{2}\pi$	139
	0	0	0	0	$\frac{\pi}{2}$	$\frac{\pi}{2}$	
	0	0	0	0	π	0	
$H \otimes I$	0	$\frac{\pi}{2}$	0	π	0	π	272
	π	π	0	π	π	π	
	$\frac{\pi}{2}$	0	0	0	0	0	
$\text{CX}_{c=q_0}^{t=q_1}$	0	π	0	0	0	0	176
	0	0	0	0	0	0	
	0	0	0	0	π	0	
$\text{CX}_{c=q_1}^{t=q_0}$	0	π	0	π	π	π	309
	π	π	0	$\frac{3}{2}\pi$	π	$\frac{\pi}{2}$	
	0	0	0	0	$\frac{\pi}{2}$	$\frac{3}{2}\pi$	

	$\tilde{\theta}_1$	$\tilde{\theta}_2$	$\tilde{\theta}_3$	$\tilde{\phi}_1$	$\tilde{\phi}_2$
value	2.09440	1.91063	1.31812	5.17604	5.03414
$e^{i\tilde{\theta}/2}$	$\frac{1+\sqrt{3}i}{2}$	$\frac{1+\sqrt{2}i}{\sqrt{3}}$	$\frac{\sqrt{5}+\sqrt{3}i}{\sqrt{8}}$	—	—
$e^{i\tilde{\phi}}$	—	—	—	$\frac{\sqrt{2}-\sqrt{8}i}{\sqrt{10}}$	$\frac{1-3i}{\sqrt{10}}$

Table 3.5: Numerically obtained decomposition parameters for several gates according to Equation (3.11), together with sequence execution durations T computed using Equation (3.10). All angles are measured in radians. We recognize decimal approximations of parameters as multiples of $\frac{\pi}{2}$ or special analytical expressions which we show in the lower table.

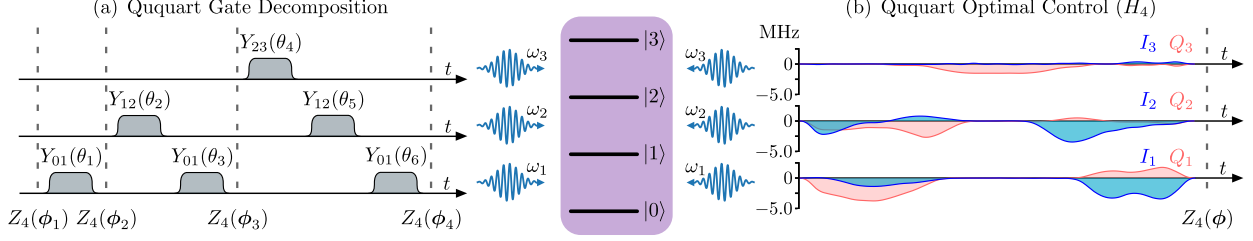


Figure 3.10: Two ways of realizing arbitrary quantum gates on a four-level ququart, which is controlled by three modulated carrier waves with resonant frequencies ω_j . (a) Any ququart gate can be decomposed into a series of at most six qubit-like $Y_{j-1,j}(\theta_i)$ rotations interleaved with ququart phase gates. Given a target unitary, the parameters θ_i and ϕ_i can be calculated numerically. (b) Lifting the gate abstraction, quantum optimal control can be used to determine a control signal that operates on all subspaces in parallel and manipulate the entire ququart state. A final virtual $Z_4(\phi)$ gate adjusts the relative phases. The shown pulses together with $\phi = (4.265, 4.590, 0.000)$ (radians) implement a ququart Hadamard H_4 in $T = 350$ ns.

Hamiltonian

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_c(t) = \sum_{n=0}^{\tilde{d}-1} \epsilon_n |n\rangle\langle n| + f(t)(-i)(a - a^\dagger), \quad (3.17)$$

where $\mathcal{H}_0$ and $\mathcal{H}_c$ describe the drift and the control Hamiltonian, respectively. $f(t)$ is the control pulse which we write as

$$f(t) = \sum_{j=1}^3 \Omega_j(t) \cos(\omega_j t + \varphi_j(t)) = \text{Re} \left[\sum_{j=1}^3 (I_j(t) + iQ_j(t)) e^{i\omega_j t} \right], \quad (3.18)$$

where for each subspace we define the quadratures $I_j(t) = \Omega_j(t) \cos(\varphi_j(t)) = \sum_{s=1}^{N_s} \gamma_{js}^I B_s(t)$ and $Q_j(t) = \Omega_j(t) \sin(\varphi_j(t)) = \sum_{s=1}^{N_s} \gamma_{js}^Q B_s(t)$. We employ the same parameterization as in the IPR study covered in Section 3.2, which is written in Equation (3.5).

We calibrate the resonant carrier frequencies $\omega_j = \epsilon_j - \epsilon_{j-1}$ for the ququart subspace from Rabi and Ramsey experiments and extrapolate the energies of higher levels as

$$\epsilon_n = \epsilon_3 + (n-3)\omega_3 + \frac{(n-2)(n-3)}{2}(\omega_2 - \omega_3) \quad (3.19)$$

for $n > 3$. This method essentially defines an anharmonicity $K = \omega_3 - \omega_2$ and extends the calibrated spectrum according to the anharmonic oscillator model that we introduced in Section 2.3.2. The difference in drive operator in Equation (3.18) $(-i(a - a^\dagger))$ compared to previously shown in Equation (2.60) $(a + a^\dagger)$ only corresponds to a different convention.

It is important to note that we want our optimal control pulses to implement a desired gate U_{targ} within duration T in the interaction frame defined by $|\psi_{\text{rot}}(t)\rangle = e^{i\mathcal{H}_0 t} |\psi_{\text{lab}}(t)\rangle$ with Hamiltonian

$$\mathcal{H}_{\text{int}} = \sum_{j=1}^3 \frac{\sqrt{J}}{2} (I_j(t)Y_j + Q_j(t)X_j).$$

This is the same frame that the native gate decomposition discussed in Section 3.3.3 realized the target gate in. However, optimal control using this model cannot take the leakage into higher energy levels into account, therefore we consider a different rotating frame $|\psi_{\text{rot}}(t)\rangle = e^{i\omega_{\text{rot}} t n} |\psi_{\text{lab}}(t)\rangle$ to solve the open-loop problem instead. This yields the Hamiltonian

$$\mathcal{H}_{\text{rot}}(t) = \sum_{n=0}^{\tilde{d}-1} (\epsilon_n - n\omega_{\text{rot}}) |n\rangle\langle n| + \frac{1}{2} \left(I(t)(-i)(a - a^\dagger) + Q(t)(a + a^\dagger) \right) \quad (3.20)$$

with $I(t) + iQ(t) = \sum_{j=1}^3 (I_j(t) + iQ_j(t))e^{i(\omega_j - \omega_{\text{rot}})t}$. The Hamiltonian is truncated at $\tilde{d} = 5$ levels to account for leakage and we find this to be sufficient in simulation (see final part of Section 3.3.5). Accordingly, the target gate has to be transformed to $U'_{\text{targ}} = W_T U_{\text{targ}}$, where $W_T = e^{i\omega_{\text{rot}} T n} e^{-i\mathcal{H}_0 T}$. The infidelity cost term becomes

$$\mathcal{I}(T; \gamma, \phi) = 1 - \frac{1}{d^2} \left| \text{Tr} \left[U'_{\text{targ}}{}^\dagger Z_4(\phi) U(T; \gamma) \right] \right|^2. \quad (3.21)$$

with the phase corrections ϕ and pulse parameters $\gamma = [\gamma_{js}^I, \gamma_{js}^Q]_{js}$ that are optimized over. This expression for the infidelity differs from our original definition in Equation (2.51) because we want to search for a pulse that implements the target up to a trailing phase gate $Z_4(\phi)$; this is why we also optimize over the phase vector ϕ . It effectively creates a manifold

of equally desirable optimization targets, some of which might be easier to realize than the original target. Since virtual phase gates come at zero cost in the experiment, this increases the freedom of the optimization and generally allows for faster pulse implementations. This trailing $Z_4(\phi)$ gate has the same effect as $Z_4(\phi_4)$ in the gate decomposition approach shown in Figure 3.10(a). Note that an explicit leading phase gate is not required as its effect is already captured by the phase relations of $I_j + iQ_j$.

Figure 3.10(b) shows optimized quadrature envelopes I_j and Q_j that modulate the carrier waves to directly realize the ququart Hadamard H_4 up to a final phase correction $\phi = (4.265, 4.590, 0.000)$ (radians). We similarly obtain controls for the Hadamard tensor product $H \otimes H$. For this gate the freedom of a final phase correction did not help to achieve a reduced gate duration, so we set $\phi = \mathbf{0}$.

Technical details of full QOC setup

We choose $\omega_{\text{rot}} = \omega_1$ to reduce the oscillating behavior of the functions I and Q , which allows the choice of larger time steps in the numerical solution of the dynamics. Typically we set the time step size $\Delta t = 0.03 \dots 0.3$ ns. We seek to find optimal control pulses of short durations as decoherence limits the capabilities of our device, especially due to increased dephasing in the $\{|2\rangle, |3\rangle\}$ subspace (see Table 3.4). Here we perform the duration optimization by hand, which is achieved by reducing the gate time further and further while making sure the numerical infidelity stays below the threshold $\mathcal{I}(T; \gamma, \phi) \leq 10^{-4}$. This simple approach could be substituted by a modified IPR scheme (see Section 3.2) which takes the freedom of the trailing phase gate $Z_4(\phi)$ into account. Alternatively, one could make the duration T an optimization variable, rescale all time variables with T^{-1} and all frequency variables with T , respectively, and then solve the optimal control task in the time interval $[0, 1]$. However, we observe worse convergence and higher obtained gate durations in this case. Despite its simpler nature, our method is sufficient for our purposes and manages to show the impact of

performing final phase corrections using a virtual $Z(\phi)$ (VZ) on the gate time T . Figure 3.11 shows H_4 pulse implementations both with (solid lines) and without a final VZ (dashed lines), where the difference in duration is highlighted by the solid gray line. Interestingly we find that for the $H \otimes H$ tensor product optimizing with a final VZ does not lead to a reduced gate duration, therefore we choose the optimization result without a VZ in our experiments.

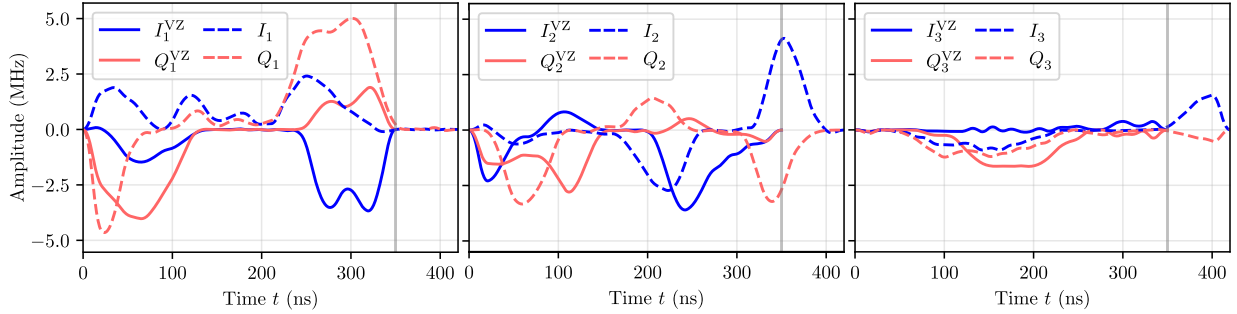


Figure 3.11: Optimal control pulses for ququart Hadamard H_4 , with final VZ correction (solid lines, $T = 350$ ns) and without final VZ correction (dashed lines, $T = 420$ ns). The vertical gray line marks the duration of the former.

The full optimal control problem formulation includes additional cost terms to ensure a narrow Fourier spectrum of the pulse f and to respect external hardware constraints. We first discuss the latter. The optimized control pulse needs to obey the power limit of the waveform generator (AWG), which imposes $|f(t)| \leq 1$ AWG. We recall that the conversion from the drive strength on-chip (in MHz) to AWG power is frequency-dependent and described by r_j (see Section 3.3.1). Using Equation (3.18), this constraint can be ensured by imposing

$$|f(t; \gamma)| \leq \sum_{j=1}^3 \frac{|\Omega_j(t; \gamma)|}{r_j} \equiv \tilde{\Omega}(t; \gamma) \stackrel{!}{\leq} 1 \text{ AWG} \quad \Rightarrow \quad \sum_{j=1}^3 \frac{1}{r_j} \sqrt{(\gamma_{js}^I)^2 + (\gamma_{js}^Q)^2} \leq 1 \text{ AWG}. \quad (3.22)$$

At this time Juqbox does not support constraints that include multiple optimization variables. Without any additional knowledge of the structure of the target gate, the most general approach is assigning one third of the power budget to each of the three carrier waves, which can be done by enforcing $|\gamma_{js}^I|, |\gamma_{js}^Q| \leq \frac{r_j}{3\sqrt{2}}$. However, this may drastically reduce the po-

tential speed of the gate implementation as no transition is allowed to be driven at more than a third of the maximum power.

This motivated the transition to Boulder Opal [151] as the optimization tool while sticking to the Jupyter pulse parameterization to overcome this limitation. While Boulder Opal currently cannot directly implement constraints of the form (3.22) either, it allows for a more flexible formulation of the quantum control task, construction of signals and definition of cost terms. Let us consider the functions

$$h(x) = 1 - \frac{\tanh\left(-\frac{1000x}{a}\right) + 1}{2}, \quad (3.23)$$

$$s(x) = (1 - h(x - a))x + h(x - a)a, \quad (3.24)$$

where h serves as a differentiable version of the Heaviside step function Θ that can be implemented in Boulder Opal's framework. s can be seen as a suppression function with threshold a , which means $s(x) \approx x$ for $x < a$ and $s(x) \approx a$ for $x \geq a$. Using this we can construct a suppressed signal

$$\tilde{\Omega}^{\text{sup}}(t; \gamma) = s\left(\tilde{\Omega}(t; \gamma)\right) \quad (3.25)$$

and a penalty term

$$P_{\text{amp}}(\gamma) = \int_0^T dt \left| \tilde{\Omega}(t; \gamma) - \tilde{\Omega}^{\text{sup}}(t; \gamma) \right|, \quad (3.26)$$

which captures how much the original signal $\tilde{\Omega}$ violates the power constraint. We set the threshold amplitude to $a = 0.95$ AWG to allow the optimizer some slack to the critical value of 1 AWG.

The other additional cost term addresses the assumption of a constant frequency response matrix in the range $[-25 \text{ MHz}, 25 \text{ MHz}]$ around the three transition frequencies, which we

discussed in Section 3.3.1. In order to be consistent with this assumption in the optimal control pulse design, we introduce a filter penalty to steer the optimization towards pulses with narrow peaks in the Fourier spectrum:

$$P_{\text{fil}}(\gamma) = \sum_{j=1}^3 \int_0^T dt \left| \Omega_j(t; \gamma) - \Omega_j^{\text{fil}}(t; \gamma) \right|. \quad (3.27)$$

Here Ω_j^{fil} is the convolution of Ω_j with a sinc-kernel to limit the bandwidth around each transition to $[-25 \text{ MHz}, 25 \text{ MHz}]$. The computation of the convolution is a built-in feature in Boulder Opal.

Ultimately, the full optimal control problem reads

$$\hat{\gamma}_{js}, \hat{\phi} = \arg \min_{\gamma, \phi} [\mathcal{I}(T; \gamma, \phi) + c_{\text{amp}} P_{\text{amp}}(\gamma) + c_{\text{fil}} P_{\text{fil}}(\gamma)], \quad (3.28)$$

where we empirically choose $c_{\text{amp}} = c_{\text{fil}} = \frac{1}{2}$. We show the obtained pulse implementations for our gates of interest H_4 and $H \otimes H$ in Figures 3.11 and 3.12, respectively.

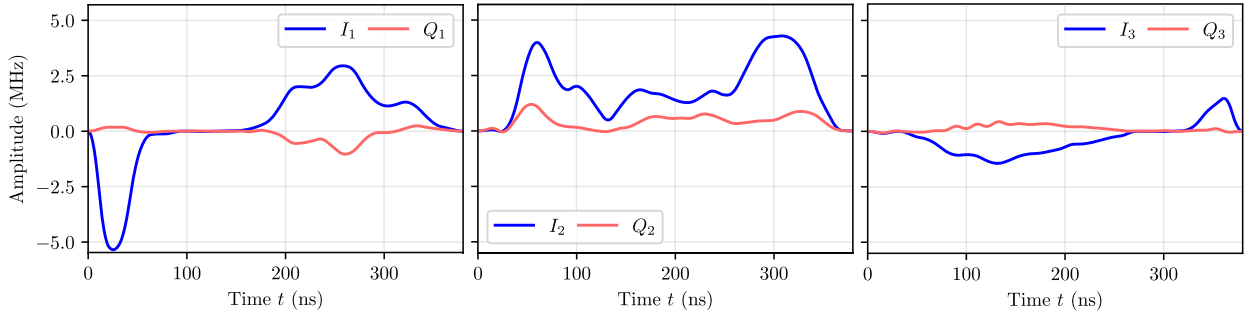


Figure 3.12: Optimal control pulses for the tensor product $H \otimes H$ without a final VZ correction, $T = 380 \text{ ns}$. Allowing a correction did not lead to a significantly reduced gate duration.

3.3.5 Experimental results

We now turn towards the experimental realization of qudit operations with our control pulses and describe how we employ quantum process tomography (QPT) as well as randomized

benchmarking (RB) to characterize their fidelities.

Quantum process tomography

We start by estimating the quality of our control pulses via quantum process tomography (QPT) [33], which allows us to fully characterize the corresponding quantum channel $\mathcal{E}(\rho) = \sum_{jk} \chi_{jk} A_j \rho A_k^\dagger$ (recall Section 2.4). We extract the process matrix χ of a specific gate by applying it to 16 differently prepared states $\{|k\rangle, (|l\rangle + |k\rangle)/\sqrt{2}, (|l\rangle - i|k\rangle)/\sqrt{2}\}$ ($k, l \in \{0, 1, 2, 3\}, k > l$) and reconstructing the resulting state using quantum state tomography (QST) [33]. This is achieved by sandwiching the gate pulse of interest between combinations of pre- and post-rotation gates from sets $\mathcal{S}$ and $\mathcal{T}$, respectively. These sets are shown in Table 3.6, where I_4 is the identity gate and the rotations $R_{j-1,j}(\theta, \phi)$ are defined in Equation (3.9).

Single-shot readouts are collected after applying the post-rotations. The readout histogram and heatmap of the assignment fidelity matrix are visualized in Figure 3.13. We apply the confusion matrix, which is the inverse matrix of the normalized readout heatmap, to the single-shot data to reduce the readout error. Maximum likelihood estimation (MLE) is then used to reconstruct the physical density matrix that describes the state prepared by a specific pre-rotation and gate pulse.

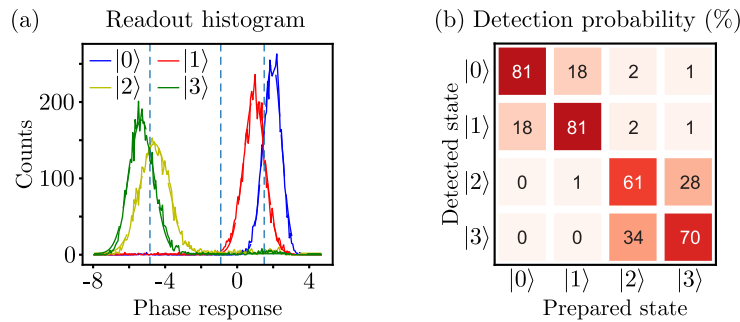


Figure 3.13: Readout characterization. (a) Single-shot histogram after phase rotation and (b) heatmap of detection probabilities (in %). Single-ququart basis states are prepared and measured 5000 times.

Pre-rotation set $\mathcal{S}$	Post-rotation set $\mathcal{T}$
I_4	I_4
$R_{01}(\frac{\pi}{2}, 0)$	$R_{01}(\frac{\pi}{2}, 0)$
$R_{01}(\frac{\pi}{2}, \frac{\pi}{2})$	$R_{01}(\frac{\pi}{2}, \frac{\pi}{2})$
$R_{01}(\pi, 0)$	$R_{01}(\pi, 0)$
$R_{12}(\frac{\pi}{2}, 0)$	$R_{12}(\frac{\pi}{2}, 0)$
$R_{12}(\frac{\pi}{2}, \frac{\pi}{2})$	$R_{12}(\frac{\pi}{2}, \frac{\pi}{2})$
$R_{12}(\pi, 0)R_{01}(\frac{\pi}{2}, 0)$	$R_{12}(\frac{\pi}{2}, 0)R_{01}(\pi, 0)$
$R_{12}(\pi, 0)R_{01}(\frac{\pi}{2}, \frac{\pi}{2})$	$R_{12}(\frac{\pi}{2}, \frac{\pi}{2})R_{01}(\pi, 0)$
$R_{12}(\pi, 0)R_{01}(\pi, 0)$	$R_{12}(\pi, 0)R_{01}(\pi, 0)$
$R_{23}(\frac{\pi}{2}, 0)$	$R_{23}(\frac{\pi}{2}, 0)$
$R_{23}(\frac{\pi}{2}, \frac{\pi}{2})$	$R_{23}(\frac{\pi}{2}, \frac{\pi}{2})$
$R_{23}(\pi, 0)R_{12}(\frac{\pi}{2}, 0)$	$R_{23}(\frac{\pi}{2}, 0)R_{12}(\pi, 0)$
$R_{23}(\pi, 0)R_{12}(\frac{\pi}{2}, \frac{\pi}{2})$	$R_{23}(\frac{\pi}{2}, \frac{\pi}{2})R_{12}(\pi, 0)$
$R_{23}(\pi, 0)R_{12}(\pi, 0)R_{01}(\frac{\pi}{2}, 0)$	$R_{23}(\frac{\pi}{2}, 0)R_{12}(\pi, 0)R_{01}(\pi, 0)$
$R_{23}(\pi, 0)R_{12}(\pi, 0)R_{01}(\frac{\pi}{2}, \frac{\pi}{2})$	$R_{23}(\frac{\pi}{2}, \frac{\pi}{2})R_{12}(\pi, 0)R_{01}(\pi, 0)$
$R_{13}(\pi, 0)R_{12}(\pi, 0)R_{01}(\pi, 0)$	$R_{23}(\pi, 0)R_{12}(\pi, 0)R_{01}(\pi, 0)$

Table 3.6: Sets of pre-rotation ($\mathcal{S}$) and post-rotation ($\mathcal{T}$) gates to implement quantum state tomography and quantum process tomography.

The full quantum channel

$$\mathcal{E}(\rho) = \sum_{mn, m'n'} \chi_{mn, m'n'} A_{mn} \rho A_{m'n'}^\dagger \quad (3.29)$$

is calculated based on Ref. [122] with either the basis $\{A_{mn} = Z_4^m X_4^n\}_{mn}$ (ququart scenario) or $\{A_{mn} = \sigma_m \otimes \sigma_n\}_{mn}$ (two-qubit scenario, ququart encoding), $m, n = \{0, 1, 2, 3\}$. Here Z_4 and X_4 are the single-ququart Pauli operators and σ_m denotes one of the qubit Pauli operators $\{I, X, Y, Z\}$. The operator basis elements are orthogonal under the Hilbert-Schmidt inner product: $\text{Tr}[A_{mn}^\dagger A_{m'n'}] = d \delta_{mm'} \delta_{nn'}$. The process fidelity of two channels with process matrices χ_1 and χ_2 is then given by

$$\mathcal{F} = |\text{Tr}[\chi_1 \chi_2]|. \quad (3.30)$$

Figure 3.14 shows the process matrices for both gates of interest H_4 and $H \otimes H$, comparing between ideal theory and implementation on our transmon. Defining the experimental process fidelity as $\mathcal{F} = |\text{Tr}[\chi_{\text{ideal}} \chi_{\text{expt}}]|$, we obtain fidelities of 86.19% and 84.06%, respectively. As a reference, we note that performing no gate (thus, characterizing the identity I_4) leads to a process fidelity of 89.88%. We trace back this loss in fidelity to measurement errors as the primary source, resulting from the challenge of implementing a four-state single-shot readout. A well-known disadvantage of quantum process tomography for benchmarking quantum gates is that it cannot separate gate errors from SPAM errors, thus yielding a lower bound on the true gate performance.

In Table 3.7 we include more QPT characterization results for a wider range of gates.

Randomized benchmarking

Randomized benchmarking (RB) allows extracting the average error per Clifford gate $r_{\mathcal{C}}$ without including SPAM errors [113]. In this protocol one constructs sequences of m Clifford

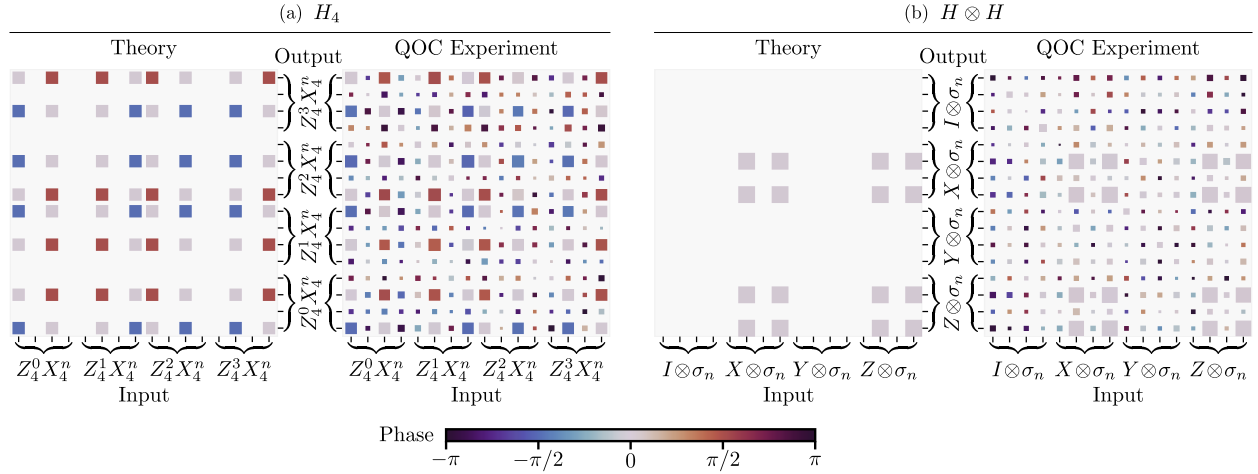


Figure 3.14: Process matrices for quantum optimal control implementations of the (a) ququart H_4 and (b) two-qubit $H \otimes H$. The area of each square encodes the magnitude of the corresponding element while the color encodes its phase. The H_4 process matrix is represented in the generalized Pauli basis while $H \otimes H$ is represented in the Pauli basis for two qubits. The exponent/index $n \in \{0, 1, 2, 3\}$ assumes increasing values from left to right or top to bottom for each tick in a label group.

Gate	DEC	QOC	
		no VZ	with VZ
I_4		89.88%	
Z_4		88.85%	
H_4	84.30%	81.35%	86.18%
$H \otimes H$	83.53%	84.06%	—*
$X \otimes H$	83.47%	—	—
$H \otimes I$	84.80%	—	—
$CX_{c=q_1}^{t=q_0}$	85.36%	—	—

Table 3.7: Single-ququart gate fidelity characterized by QPT, for decomposed gates (DEC) and optimally controlled gates (QOC). In the QOC implementation of $H \otimes H$ (*) we only considered the case without a final virtual Z (VZ).

gates $C_m C_{m-1} \dots C_1$ and further computes inversion gates C_{inv} , which invert the effect of the sequences:

$$C_{\text{inv}} C_m C_{m-1} \dots C_1 = I_d. \quad (3.31)$$

C_{inv} can be computed efficiently due to the Gottesman-Knill theorem [64]. In the experiment, one initializes the system in the ground state, applies the sequence, and measures the population in the ground state (survival probability) P_m .

By averaging over randomly sampled sequences of gates from the Clifford group $\mathcal{C}$, the error channel becomes effectively depolarizing; analyzing the exponential decay $\langle P_m \rangle = ap^m + b$ over sequences of different depths m allows extracting the depolarization parameter $p_{\mathcal{C}}$. The error per gate is then given as

$$r_{\mathcal{C}} = \frac{d-1}{d}(1 - p_{\mathcal{C}}) \quad (3.32)$$

and the average gate fidelity defined as $\mathcal{F}_{\mathcal{C}} = 1 - r_{\mathcal{C}}$.

Since RB relies on sampling from the Clifford group, choosing the appropriate Clifford group for a certain scenario is important. We recall from Section 2.2.5 that the Clifford group $\mathcal{C}$ is defined as the group which normalizes the Pauli group $\mathcal{P}$. In the ququart case, the Pauli group $\mathcal{P}_4$ consists of products of X_4 and Z_4 , which are normalized by H_4 and $S_4 = \text{diag}(1, \sqrt{i}, i, \sqrt{i})$ [79]. Z_4 trivially is an element of the ququart Clifford group $\mathcal{C}_4$ and we check numerically that $\{H_4, Z_4, S_4\}$ can generate the predicted number of 768 elements in $\mathcal{C}_4$ [180]. In the case of two qubits, the Pauli group $\mathcal{P}_2^{\otimes 2}$ is given by the tensor product of the single-qubit operators I, X, Y and Z , and the gates $\{H, S, CX\}$ generate the Clifford group $\mathcal{C}_2^{\otimes 2}$ with 11520 elements (up to scalars) [165].

For the ququart case, we sample from $\mathcal{C}_4$, which we construct explicitly, and for the two-qubit case we use Qiskit [152] to produce RB sequences. In either case, we decompose the matrices for each gate in a sequence according to decomposition (3.11) in order to implement

the RB experiment on our transmon. We consider RB sequences of depth m up to 100. The measured survival probabilities are visualized in Figure 3.15 and clearly show the exponential decay. We obtain RB fidelities $F_{C_4} = 96.22(14)\%$ and $F_{C_2^{\otimes 2}} = 95.84(05)\%$ for the ququart RB and two-qubit RB, respectively.

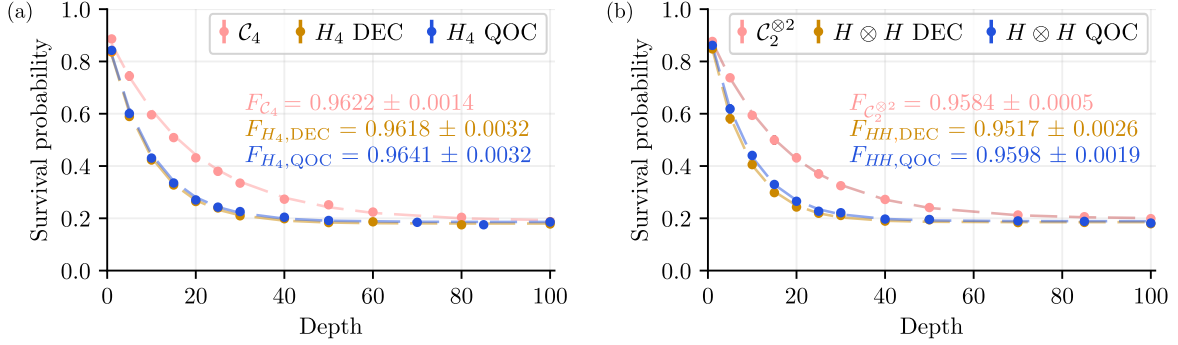


Figure 3.15: Experimental results for (a) ququart RB (C_4 Clifford group, red) and (b) two-qubit RB ($C_2^{\otimes 2}$ Clifford group, red) on a four-level transmon. Using IRB we benchmark (a) H_4 and (b) $H \otimes H$ in the appropriate setting, for both their gate-level decomposition (DEC, yellow) and their quantum optimal control (QOC, blue) implementation. Markers represent the mean survival probabilities and error bars (smaller than the marker size) are the standard deviations of these means.

Interleaved randomized benchmarking

Interleaved RB (IRB) builds upon standard RB to characterize the error of a specific Clifford group element G , which is achieved by interleaving the gate of interest between the randomly sampled RB gates [115]. Similar to standard RB, this yields the depolarization parameter p_{C+G} . Together with the previously determined parameter p_C , the gate fidelity can be computed as

$$\mathcal{F}_G = 1 - r_G = 1 - \frac{d-1}{d} \left(1 - \frac{p_{C+G}}{p_C} \right). \quad (3.33)$$

We apply IRB to characterize the ququart Hadamard H_4 and the Hadamard tensor product $H \otimes H$, where we interleave them between gates sampled from the appropriate Clifford groups. We do this for both the gate-level decomposition and the optimal control implementation of each gate and juxtapose the results in Figure 3.15. The open-loop optimized pulses

achieve an improvement over the gate-based realizations in both cases: The H_4 QOC pulse is benchmarked at 96.41(32)% while the decomposition achieves 96.18(32)% fidelity; the $H \otimes H$ QOC pulse yields 95.98(19)% fidelity compared to the decomposition with 95.17(26)%. This observation suggests a good agreement between the Hamiltonian model and the device. The main limitation is now caused by decoherence, dominated by the increased dephasing in the $\{|2\rangle, |3\rangle\}$ subspace.

In Table 3.8 we include additional RB and IRB results for special gates.

Gate	DEC	QOC
H_4	96.18(32)%	96.41(32)%
$H \otimes H$	95.17(19)%	95.98(19)%
$I \otimes H$	98.51(13)%	—
$H \otimes I$	96.70(29)%	—
$CX_{c=q_0}^{t=q_1}$	96.68(24)%	—
$CX_{c=q_1}^{t=q_0}$	98.84(17)%	—

Table 3.8: Single-ququart gate fidelity characterized by different types of RB experiments, for decomposed gates (DEC) and optimally controlled gates (QOC). The H_4 gate fidelity is characterized using the Clifford group $\mathcal{C}_4$, other gates are characterized using the Clifford group $\mathcal{C}_2^{\otimes 2}$. Base fidelity for RB with $\mathcal{C}_4$ is 96.22(14)%, and with $\mathcal{C}_2^{\otimes 2}$ is 95.84(5)%. We only consider the VZ implementation for the QOC H_4 pulse.

Error budget simulation

In order to simulate the gate errors we observe in experiments, we compute the system evolution under our designed control pulses using the Lindblad master equation [108, 61] and show the obtained infidelities of different ququart gates in Table 3.9. We compare the noise-free case with the noisy case, where for the latter we include decay and dephasing collapse operators for each subspace using the experimentally measured lifetimes shown in Table 3.4. The simulation time step is set to $dt = 1$ ps for high numerical precision, which is over an order of magnitude lower than what was used for the pulse optimization (see

Section 3.3.4). We perform all simulations with the lowest 5 energy levels of the transmon to account for a leakage into higher levels. By taking the trace of the 4×4 submatrix which describes the ququart state, we verify for each simulation result that leakage into the fifth state is below 0.01%, justifying the truncation of the Hamiltonian at $\tilde{d} = 5$ levels.

We observe that the optimal control pulses can theoretically reduce the decoherence-free infidelities by more than 30%, and future work will optimize pulse duration to reduce the decoherence error further. The difference in infidelity between the ideal simulation here and the optimization target $\mathcal{I} \leq 10^{-4}$ arises from the increased time step precision.

Gate	Decoherence-free	Decoherence
H_4 DEC	0.42%	4.53%
$H \otimes H$ DEC	0.41%	4.78%
$CX_{c=q_0}^{t=q_1}$ DEC	0.38%	4.11%
$CX_{c=q_1}^{t=q_0}$ DEC	0.03%	2.18%
$H \otimes I$ DEC	0.62%	3.89%
$I \otimes H$ DEC	0.11%	1.82%
H_4 QOC	0.08%	4.28%
$H \otimes H$ QOC	0.26%	4.81%

Table 3.9: Ququart gate infidelities extracted from simulation. DEC: decomposition, QOC: quantum optimal control.

3.3.6 Conclusion

We successfully demonstrated the coherent control of the lowest four levels in a superconducting transmon, where we explored the implementation of both high-level optimized gate sequences and low-level optimized control pulses to realize target unitaries. While in the first approach we drive only one transition at a time, the latter approach enables manipulating all internal states at once by using open-loop quantum optimal control tools.

We studied the quantum system from two angles: from the ququart perspective, as the four-dimensional generalization of a qubit, and from the perspective of two encoded qubits.

We characterized gates from either perspective in QPT and RB experiments, focusing in particular on the ququart H_4 and the two-qubit $H \otimes H$ gate. They have similar matrix representations but belong to different Clifford groups $\mathcal{C}_4$ and $C_2^{\otimes 2}$, respectively, and we observe IRB fidelities $\geq 95\%$ for both, finding that the optimal control implementations perform better.

The fundamental challenges of increased noise on higher levels as well as the difficulty of performing a four-state readout limit the capabilities of our device. One possible future direction includes exploring the application of data-driven methods like presented in Ref. [202] to incorporate hardware feedback into the pulse calibration and close the optimal control loop.

While recently presented related work also studied the two-qubit encoding scheme [23, 49], we were, to our knowledge, first to explore optimally controlled quantum gates on a transmon ququart and further perform ququart randomized benchmarking using the Clifford group $\mathcal{C}_4$.

CHAPTER 4

QUANTUM OPTIMAL CONTROL FOR QUANTUM SENSING

While the primary focus of this thesis lies with quantum computing, another quantum technology that has promising near-term applications and can benefit from advances control schemes are quantum sensors. In the following, we draw parallels to familiar concepts from quantum computing and introduce the control problem for a specific quantum sensing platform.

In the realm of quantum computing, the language of quantum gates has been developed to describe universal sets of operations that may be performed on a quantum computer (regardless of the platform) in order to implement quantum algorithms [136]. Such algorithms consist of state preparation, single-qubit rotations, entangling blocks, mid-circuit measurements, readout, etc., based on single-, two- and multi-qubit operations. Although this paradigm is intended for performing quantum computing, the high degree of control over quantum systems has opened up analog applications including quantum simulation of many-body Hamiltonians (particularly of lattice systems) [57, 38] and on-demand generation of quantum resources for quantum sensing [117, 89]. In the latter case, quantum computers have been explored as programmable quantum sensors, mainly exploring the preparation of states according to metrological cost functions.

These developments motivate the paradigm of “programmable” quantum sensing, which describes using quantum logic gates, in a hardware-agnostic manner, to perform sensing tasks. Much attention has been paid to using noisy intermediate-scale quantum (NISQ) [145] machines as quantum sensors of spin rotations due to, for example, stray electromagnetic fields [117, 89, 212, 88]. However, due to the tight confinement of qubits in quantum computing platforms, these machines are ill-suited for measuring inertial and gravitational forces, limiting their application as programmable inertial sensors. Conventional cold atom inertial sensing platforms are instead able to generate large momentum transfers and subsequently

split clouds of atoms to enclose large spacetime areas in free space, thereby yielding high sensitivities to external parameters [87, 128, 98]. The set of possible operations in such systems is however constrained due to the lack of reconfigurability of the platforms and the restriction to limited spatial regions of a vacuum chamber where the atoms can be addressed, which severely limits their programmability.

Recent experiments have demonstrated that cold atom inertial sensing can be performed using optical lattices, where the atoms remained always confined in a standing wave of light, and the basic steps of an interferometer — splitting, mirroring and recombining — can be achieved by time-dependent modulation of the phase of the light forming the lattice [196, 103, 102]. Such a protocol enables sensing the presence of external accelerations, thus this system can be regarded as a quantum accelerometer. Because the atoms remain always confined (albeit delocalized), the modulation waveforms for efficiently generating the traditional components required for interferometry are nontrivial. Previous studies have outlined how machine learning methods, including reinforcement learning, and (indirect) quantum optimal control can be used to perform state mappings and design the traditional components (SPLIT, MIRROR, RECOMBINE) as well as find modulation sequences spanning the entire operation time of the sensor which maximize the overall accelerometer sensitivity [103, 144, 50, 197, 195, 30, 168, 134, 3, 31, 29]. While much attention has been paid to these traditional component and end-to-end design problems, the reconfigurability of the optical lattice interferometer can also be used to perform a wider range of operations, making it a natural candidate for an inertial quantum sensing platform which can execute a broad array of user-defined programs.

In the following sections, we introduce the physics of the sensing system as well as its control, discuss a framework for programmable sensing using that system, and finally present an alternative approach to optimized quantum sensing that goes beyond the structure of fixed primitives. This work is in collaboration with Inflection, a cold-atom company focusing on

engineering quantum computers and sensors.

4.1 Sensing hardware

In this section, we describe both the cold atom optical lattice platform and the specific quantum optimal control method used to design quantum operations in this system. The description of cold atoms in an optical lattice is kept at a basic level, and we refer to Refs. [103, 101] for further details of this system, including the implementation of state preparation, control, and measurement in the lab.

4.1.1 Optical lattice interferometry

We consider a cloud of cold, spinless bosonic atoms in a one-dimensional optical lattice that is generated by two counter-propagating laser beams, as realized in Refs. [196, 103]. Furthermore, we consider density regimes where interaction effects between atoms can be neglected, reducing the description to the single-atom level. At sufficiently low temperatures, the cloud forms a Bose condensate [142] such that the wavefunction of the entire system of N atoms can be represented as a product $\Psi(x_1, \dots, x_N, t) = \prod_{i=1}^N \psi(x_i, t)$ of the single-particle wavefunction $\psi(x, t)$. The wavefunction $\psi(x, t)$ evolves according to the Schrödinger equation with time-dependent Hamiltonian

$$\mathcal{H}(t) = \frac{p^2}{2m} + V(x, t), \quad (4.1)$$

where $\hat{p}$ and $\hat{x}$ are, respectively, the momentum and position operators for an atom with mass m , and

$$V(x, t) = -\frac{V_0}{2} \cos(2k_L x + \varphi(t)), \quad (4.2)$$

is the potential generated by the optical lattice at time t . The depth V_0 of the lattice is determined by the dynamic atomic polarizability and the intensity of the laser via the AC Stark shift [82]. Along with the depth, the wavenumber k_L of the lattice is also held fixed while its phase φ is modulated according to a time-dependent control function whose waveform is determined using quantum optimal control as described in Section 4.1.2. Experimentally, the position of the lattice is altered using frequency shifts applying to acousto-optic modulators [103]. Previous works have developed control functions implementing **SPLIT**, **MIRROR** and **RECOMBINE** components of an interferometer [103, 30, 168, 134] as well as end-to-end protocols maximizing the overall sensitivity of the device [3, 31].

Expanding the lattice potential V in the plane-wave basis $\{|p\rangle\langle p|\}_p$, we find

$$V(t) = -\frac{V_0}{4} \left(e^{i\varphi(t)} b_- + e^{-i\varphi(t)} b_+ \right). \quad (4.3)$$

The operator $b_{\pm}$ is defined by $b_{\pm}|p\rangle = |p \pm 2k_L\rangle$, representing a two-photon Bragg scattering process in which an atom absorbs and emits a photon between counterpropagating laser beams, changing its momentum by $\pm 2k_L$ [36]. In the case of a stationary lattice ($\varphi \equiv 0$), the eigenstates of the lattice Hamiltonian are given by the Bloch states $|\nu, q\rangle$ with eigenenergies $E_{\nu}(q)$ parametrized by the band index ν and quasimomentum q . We refer to Refs. [94, 142] for a more detailed introduction. The Bloch states are of alternating parity with spatial dependence and eigenenergies shown in Figure 4.1. In particular, the quasimomentum is left unchanged by the action of V regardless of the specific control of the phase φ . However, the application of a linear force with Hamiltonian contribution $-F(t)x$ leads to a time-dependent $q(t)$.

In general, an inertial force leads to a final state with different expansion coefficients and quasimomentum than in an (ideal) unaccelerated frame. In experiment the cloud is typically split, mirrored and then recombined using modulated waveforms of the control function, with each component separated by a period of free evolution.

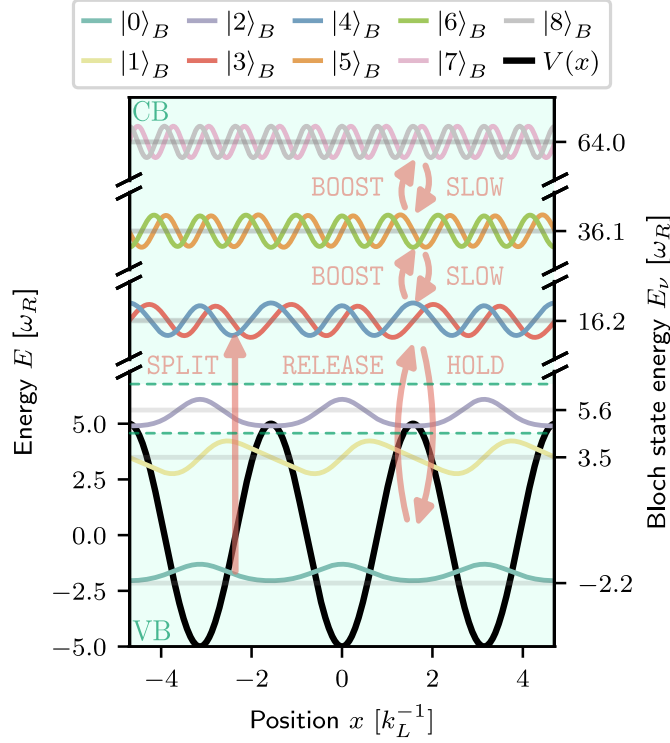


Figure 4.1: Band structure of the optical lattice Hamiltonian (see Equation (4.1)) for $\varphi \equiv 0$. The potential $V(x)$ for the specific choice $V_0 = 10 \omega_R$ is shown in black, while modulating φ corresponds to shifting the potential to the left or right. The Bloch states $\psi_\nu(x) = \langle x | \nu \rangle_B$ are shown at their corresponding energy eigenvalue, clearly defining a valence and a conduction band subspaces, with latter characterized by nearly degenerate pairs of opposite-parity states. A subset of the operations considered in this work are indicated by red arrows (see also Section 4.2.1 and Figure 4.2).

Measurement is performed in the final stage of the protocol, where the atomic cloud is allowed to ballistically expand for a fixed amount of time. This leads asymptotically to a separation of the wavefunction into the basis of discrete momentum states of the lattice $\{|2nk_L + q\rangle \mid n \in \mathbb{Z}\}$ [103]. Depending on the force F , different populations of the discrete momentum states are observed, producing a multi-fringed interference pattern, rather than the conventional interference signal of two recombined clouds in a Bragg interferometer, that reflects differential phase accumulation between the multiple arms of the interferometer. This is due to the possibility of occupying an array of discrete momentum states at integer multiples of the lattice wavenumber. In this chapter, we find control functions for the case $F = 0$, which can be straightforwardly generalized to $F \neq 0$ via the application of an offsetting bias field or through re-optimization.

It is convenient to work with the natural units of the lattice, where spatial dimensions are defined via the wave number k_L and energy scales are determined by the photon recoil frequency $\omega_R = k_L^2/2m$. In these units, the momentum-space representation of the Hamiltonian reads ($q = 0$):

$$\mathcal{H}(t) = 4\omega_R n^2 - \frac{V_0}{4} \left(e^{i\varphi(t)} b_- + e^{-i\varphi(t)} b_+ \right). \quad (4.4)$$

Here n denotes the operator that counts the multiples of $2k_L$ making up the momentum. Because q is a constant of motion, we therefore use the notations $|2nk_L\rangle_p$ for the discrete momentum states and $|\nu\rangle_B \equiv |\nu, q = 0\rangle$ for the Bloch states. Note that in this representation the operators $b_{\pm}$ effectively perform the mapping $n \rightarrow n \pm 1$.

For concreteness, we fix $V_0 = 10\omega_R$, which is a representative value of lattice depths used experimentally [103]. We note that the results in this chapter are specific to this choice of the lattice depth but the presented framework can be applied to any other setting in a straightforward fashion. As shown in Figure 4.1, the Bloch states $|0\rangle_B$ and $|1\rangle_B$ have harmonic-oscillator-like wavefunctions that are well localized in the lattice potential and

thus form the *valence band*. The higher-energy *conduction band* ($\nu \geq 3$) features quadratic dispersions with opposite-parity wavefunctions in the (approximate) form of plane waves. As the naming convention suggests, atoms in the conduction bands translate under the application of a weak external force, while atoms in the valence bands undergo (Bloch) oscillations [142, 11]. The state $|2\rangle_B$ can be viewed as an intermediate state with properties similar to both bands.

Formally, the fixed- q Bloch bands illustrated in Figure 4.1 form qudits, with the conduction and valence bands being infinite- and finite-dimensional subspaces, respectively. In practice, atoms are prepared in the ground Bloch state $|\nu = 0\rangle_B$, and then the dynamics of the control function drive metrologically useful transitions between various bands prior to measurement. Of particular importance are the (nearly) degenerate pairs of opposite-parity states in the conduction band. The sum and difference of these states (approximately) correspond to individual momentum eigenstates, i.e. $(|\nu\rangle_B \pm |\nu - 1\rangle_B)/\sqrt{2} \approx |\pm\nu k_L\rangle_p$. Given our specific choice of the lattice depth V_0 , we identify these pairs for even $\nu \geq 4$ and as can be seen in Figure 4.1. Here the parity of the state $|\nu\rangle_B$ is equivalent to the parity of the integer ν .

4.1.2 Direct trajectory optimization

Using optimal control, we are interested in designing components of the optical lattice interferometer that correspond to unitary operations U_{targ} on the whole system as well as simple state transfers $|\psi\rangle \rightarrow |\psi_{\text{targ}}\rangle$, which move the system from a specific initial state $|\psi\rangle$ to a desired final state $|\psi_{\text{targ}}\rangle$. While the quality of the former can be quantified by the gate infidelity $\mathcal{I}(T; \gamma)$ we introduced around Equation (2.51), the latter is characterized by the state infidelity

$$\mathcal{I}(T; \gamma) = 1 - |\langle \psi_{\text{targ}} | \psi(T; \gamma) \rangle|^2, \quad (4.5)$$

which relates to the overlap between the target state $|\psi_{\text{targ}}\rangle$ and the final state $|\psi(T; \gamma)\rangle$ at time T . As before, the evolution of the system is dependent on the control functions $\mathbf{f}(\gamma) = [f_j(\gamma_j)]_j$ which are expressed in terms of optimizable parameters $\gamma = [\gamma_j]_j$.

Most optimal control tools employ so-called *indirect* methods [141, 59, 22, 92], where the Schrödinger equation is solved numerically for each choice of control parameters, and the control parameters are adjusted iteratively to decrease the loss function. This is the approach we took in Chapter 3 as well. While being memory-efficient, it is restrictive if one wishes to also enforce constraints on the quantum state evolution or design cost functions that require information from intermediate time steps. Furthermore, reinforcement learning methods (as they have been applied specifically in the optical lattice interferometry context [103, 30]) are usually limited by their discrete, pre-defined action space, which typically leads to lower fidelities and requires further knowledge to determine the set of actions a priori.

In this chapter we look at a different approach known as *direct trajectory optimization*, which treats the quantum states at all times as free parameters, and treats equations of motion as constraints on these parameters. Minimization of the loss function can then be performed via interior-point optimization methods like IPOPT [188]. While this approach is more memory-intensive, as it must store the entire space-time history of a quantum state (or unitary) trajectory, it allows for a high degree of flexibility and control over the trajectory variables. For instance, in Section 4.2 we exploit the possibility to impose constraints on the quantum state parameters to guide the evolution through a certain part of Hilbert space. While this sounds rather restrictive at first, these constraints can be well motivated by the physical principles of the optical lattice interferometer and therefore reduce the search space and aid the optimizer in finding a solution. This methodology is enabled by the unique nature of direct trajectory optimization and would not be possible with the conventional optimal control tools and reinforcement learning methods mentioned above. In the following sections we explain in detail how we leverage constraints to solve for complex, high-dimensional

optical lattice operations.

Specifically, we make use of the `Piccolo.jl` framework introduced in Ref. [182]. In the remainder of this section, we introduce the fundamental language of direct trajectory optimization and apply it to the optical lattice problem.

We denote the “state” at the discrete time t_l abstractly with x_l , $l = 1, \dots, N_t$, which corresponds to either the quantum state $|\psi_l\rangle$ (for state transfer) or the unitary U_l (for gate design), and parameterize the controls as $f_j(t_l) = \gamma_{jl}$. In this setup, x_l become additional optimization variables in an extended nonlinear constrained optimization problem:

$$\hat{x}_l, \hat{\gamma}_{jl} = \arg \min_{x_l, \gamma_{jl}} \mathcal{L}(x_l, \gamma_{jl}) \quad (4.6)$$

$$\begin{aligned} \text{subject to: } & x_{l+1} - \exp(-i\mathcal{H}(\gamma_{jl})\Delta t)x_l = 0, \\ & x_1 = x(0) \end{aligned} \quad (4.7)$$

where Δt is the discrete time step. In the simplest case, the loss is given by the infidelity at the final time, $\mathcal{L}(x_l, \gamma_{jl}) = \mathcal{I}(x_{N_t})$. The advantage of this approach is that the cost function explicitly depends on a small subset of optimization variables, which allows for trivial gradient computation of $\mathcal{L}$ in these variables. The full gradient step for all variables x_l and γ_{jl} is then furthermore determined by constraints like in Equation (4.7), which enforces the unitary evolution according to the Schrödinger equation (2.14). Moreover, this approach gives us direct access to the quantum state or unitary at intermediate times of the evolution, which we exploit when designing control function waveforms. We refer the interested reader to Ref. [182] for further details.

In its current version `Piccolo.jl` supports Hamiltonians of the form shown in Equation (2.42), which is not directly compatible with the Hamiltonian in Equation (4.4) as the desired control function φ enters nonlinearly. Therefore we introduce auxiliary quadrature

controls

$$I(t) = \cos(\varphi(t)), \quad Q(t) = \sin(\varphi(t)), \quad (4.8)$$

so that we obtain

$$\mathcal{H}(t) = 4\omega_R n^2 - \frac{V_0}{4}(I(t)(b_- + b_+) + iQ(t)(b_- - b_+)). \quad (4.9)$$

After discretization, the quantum optimal control problem then takes the form

$$\hat{x}_l, \hat{I}_l, \hat{Q}_l, \hat{\varphi}_l = \arg \min_{x_l, I_l, Q_l, \varphi_l} \mathcal{L}(x_l, \varphi_l) \quad (4.10)$$

subject to the constraints in Equations (4.7) and (4.8) as well as

$$\varphi_l \in [-\pi, \pi], \quad \varphi_0 = \varphi(0) = \varphi_{N_t} = \varphi(T) = 0, \quad (4.11)$$

which constrains the range of the control function and resets the position of the lattice at the final time $t = T$.

In this current formulation there is nothing constraining the shape of the control function. However, for practical reasons of implementation on hardware, we do not want it to be discontinuous or oscillating too strongly. To this end we augment the cost function by a term

$$\begin{aligned} \mathcal{L}_\omega[\varphi] &= \int_0^T \left| \varphi(t) - \int_0^T K(t-s)\varphi(s) \, ds \right|^2 dt \\ \xrightarrow{\text{discretize}} \mathcal{L}_\omega(\varphi_l) &= \Delta t \sum_{l=1}^{N_t} \left| \varphi_l - \Delta t \sum_{m=1}^{N_t} K_{lm} \varphi_m \right|^2, \end{aligned} \quad (4.12)$$

that penalizes the Fourier spectrum of φ beyond a cutoff ω_c , which we choose to be $\omega_c =$

$70\omega_R$. $K_{lm} = K(t_l - t_m)$ is the convolution kernel and here we use a sinc kernel:

$$K(t) = \frac{\omega_c}{\pi} \text{sinc}(\omega_c t). \quad (4.13)$$

The final form of the optimal control problem is now given by Equation (4.10) with $\mathcal{L}(x_l, \varphi_l) = \mathcal{I}(x_{N_t}) + 100 \mathcal{L}_\omega(\varphi_l)$ and constraints given by Equations (4.7), (4.8) and (4.11). This general setup is used to find the results presented in the remainder of this chapter, and specific cases where further modifications are required or more constraints are added to improve the convergence of the numerical optimization are addressed as encountered.

It is important to note that our proposed framework does not require any modification to existing experimental implementations, such as that demonstrated in Ref. [103]. We merely suggest a different approach to designing high-fidelity control pulses that modulate the phase φ , therefore our results would be directly compatible with this experimental setup.

4.2 Programming an optical lattice interferometer

We use the optimal control framework introduced in the previous section to design control functions implementing various quantum gates for the optical lattice interferometer system previously demonstrated experimentally in works such as Ref. [103]. The direct trajectory optimization method, which has thus far been applied towards gate design for quantum computing [182], has distinct advantages including the ability to enforce constraints on the evolution of the quantum state and design cost functions that require information from intermediate time steps. We find that this method is able to produce high-fidelity gates, for instance for large-momentum two-level subspaces.

4.2.1 Gate design

In this section, we analyze the results for control functions producing the types of operations shown in Figure 4.2 that act on and between qubit subspaces defined within the system's full Hilbert space. A qubit subspace can be constructed from any pair of Bloch states, however, we find it useful to take neighboring pairs $(|\nu\rangle_B, |\nu-1\rangle_B)$ that are defined by projectors

$$\Pi_\nu = |\nu\rangle\langle\nu|_B + |\nu-1\rangle\langle\nu-1|_B. \quad (4.14)$$

For a given subspace Π_ν , we denote the logical qubit isometry in these subspaces as

$$P_\nu = |0\rangle_L^\nu\langle\nu|_B + |1\rangle_L^\nu\langle\nu-1|_B, \quad (4.15)$$

where we choose the convention that the even-parity state is assigned to the north pole of the logical Bloch sphere, associated with $|0\rangle_L$. For later convenience, the qubit subspace of the ground and first excited Bloch states are labeled following a different convention $P_{i_1} = |0\rangle\langle 0|_B + |1\rangle\langle 1|_B$ with isometry $P_1 = |0\rangle_L^1\langle 0|_B + |1\rangle_L^1\langle 1|_B$.

Figure 4.2 shows the qubit subspaces of interest as individual Bloch spheres as well as mapping operations between these subspaces. The mapping operations are also visualized in Figure 4.1 for reference. In the following discussion, we comment on the relevance of each operation for inertial sensing.

SPLIT $_\nu$, RECOMBINE $_\nu$

The Mach-Zehnder interferometer is a paradigmatic sensing scheme comprised of SPLIT $_\nu$, MIRROR $_\nu$ and RECOMBINE $_\nu$ operations. Given the central importance of this protocol we first analyze the results for the control function that produces the SPLIT $_\nu$ action. This operation maps the ground state $|0\rangle_B$, which corresponds to the initial state, to a Bloch state $|\nu\rangle_B$

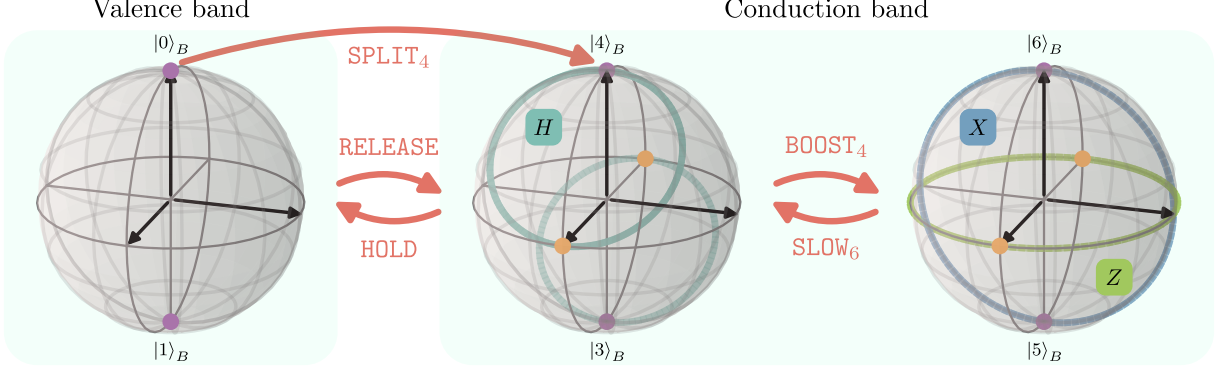


Figure 4.2: Visualization of a variety of operations that can be realized in the optical lattice interferometer. Neighboring pairs of Bloch states are identified as qubit subspaces and represented by separate Bloch spheres. Upon initialization the system is in $|0\rangle_B$ in the valence band (Π_1 qubit). SPLIT_ν maps to a Bloch state $|\nu\rangle_B$ in the conduction band, here shown for $\nu = 4$. Unitary $\text{BOOST}_\nu/\text{SLOW}_\nu$ operations map between adjacent conduction band qubits, for instance $\text{BOOST}_4 : \Pi_4 \rightarrow \hat{\Pi}_6$ and $\text{SLOW}_6 : \Pi_6 \rightarrow \Pi_4$. The mapping $\Pi_4 \rightarrow \Pi_1$ corresponds to bringing the atoms to rest (HOLD operation, inverse is RELEASE). Single-qubit operations on the conduction band qubits are visualized; we provide interpretations in terms of atom-cloud trajectories in the main text. The yellow circles in the conduction band Bloch spheres correspond to approximate momentum eigenstates $|\pm\nu k_L\rangle_p$.

in the conduction band. We note that this operation has been analyzed in several related works specifically for the choice $\nu = 3$, each finding different control functions [103, 30, 168]. As noted in Ref. [30], the inverse RECOMBINE_ν operation can be achieved through the time reversal of SPLIT_ν .

We briefly present our result for this well-studied task. In order to optimize the desired state transfer, we consider the optimization problem in Equation (4.10) with infidelity function Equation (4.5) and $|\psi(0)\rangle = |0\rangle_B$, $|\psi_{\text{targ}}\rangle = |3\rangle_B$. Figure 4.3 shows the obtained control function together with the Bloch state evolution. The even-parity $|2\rangle_B$ state is heavily populated at intermediate times as population is transported between valence and conduction bands.

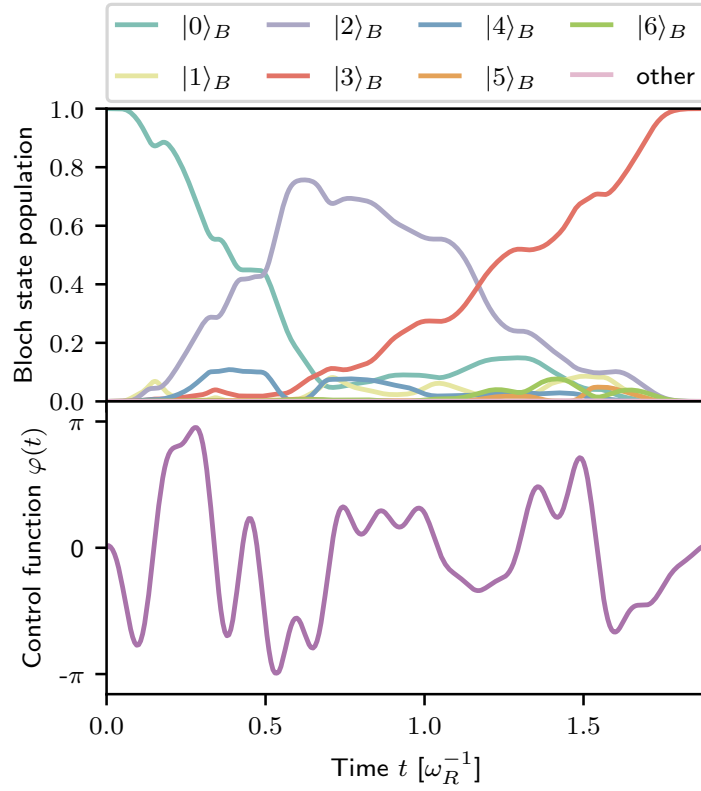


Figure 4.3: $\text{SPLIT}_3 : |0\rangle_B \rightarrow |3\rangle_B$ operation, $T = 1.88\omega_R^{-1}$, $\mathcal{I} = 3.27 \cdot 10^{-5}$. The top panel shows the state evolution in the Bloch basis while the lower panel shows the corresponding control function.

Qubit gates

Within a qubit subspace Π_ν , arbitrary unitary transformations can be performed. In order to optimize the control function producing these gates, we use the infidelity expression in Equation (2.54) with $U_{\text{targ}} = P_\nu^\dagger U_{\text{targ},L} P_\nu$, where $U_{\text{targ},L}$ is the logical gate of interest acting on Π_ν and P_ν is the isometry as defined in Equation (4.15).

We provide results for the control function achieving the common qubit gates X , Z , H and $T = Z^{1/4}$ that we introduced in the context of quantum computing in Section 2.2.2. These are of special interest because we recall that any arbitrary single-qubit gate can be approximated by a product of H and T gates [136]. We equip a gate label with the index ν to denote the qubit subspace it acts on, for instance X_ν acts on qubit subspace Π_ν . We obtain the control functions for the lowest conduction band qubit Π_4 and show the results together with the corresponding state evolutions in Figure 4.4. The bottom row shows the Fourier transformation of φ , providing information on the dominant spectral components of the control function.

Although this gate set is ubiquitous in circuit models of quantum computing (see Section 2.2.3), we provide their interpretation and utility for operations within the the conduction band in the context of inertial sensing. First, the Z_ν corresponds to the **MIRROR** $_\nu$ component in the Mach-Zehnder interferometer, which inverts the momenta of the interferometer arms. This can be seen by looking at the X -axis eigenstates (recall ν is even):

$$|\nu k_L\rangle_p \approx \frac{|\nu\rangle_B + |\nu - 1\rangle_B}{\sqrt{2}} \xleftrightarrow{Z_\nu} \frac{|\nu\rangle_B - |\nu - 1\rangle_B}{\sqrt{2}} \approx |-\nu k_L\rangle_p. \quad (4.16)$$

The other gates we consider cannot be found in the standard Mach-Zehnder protocol. The Hadamard gate H_ν transforms between Bloch and momentum bases. It maps the even-parity state $|\nu\rangle_B$, which describes a state with equally split momentum, to the positive momentum state $|\nu k_L\rangle_p$ (approximately) and vice versa. Similarly, it maps between the odd-parity state

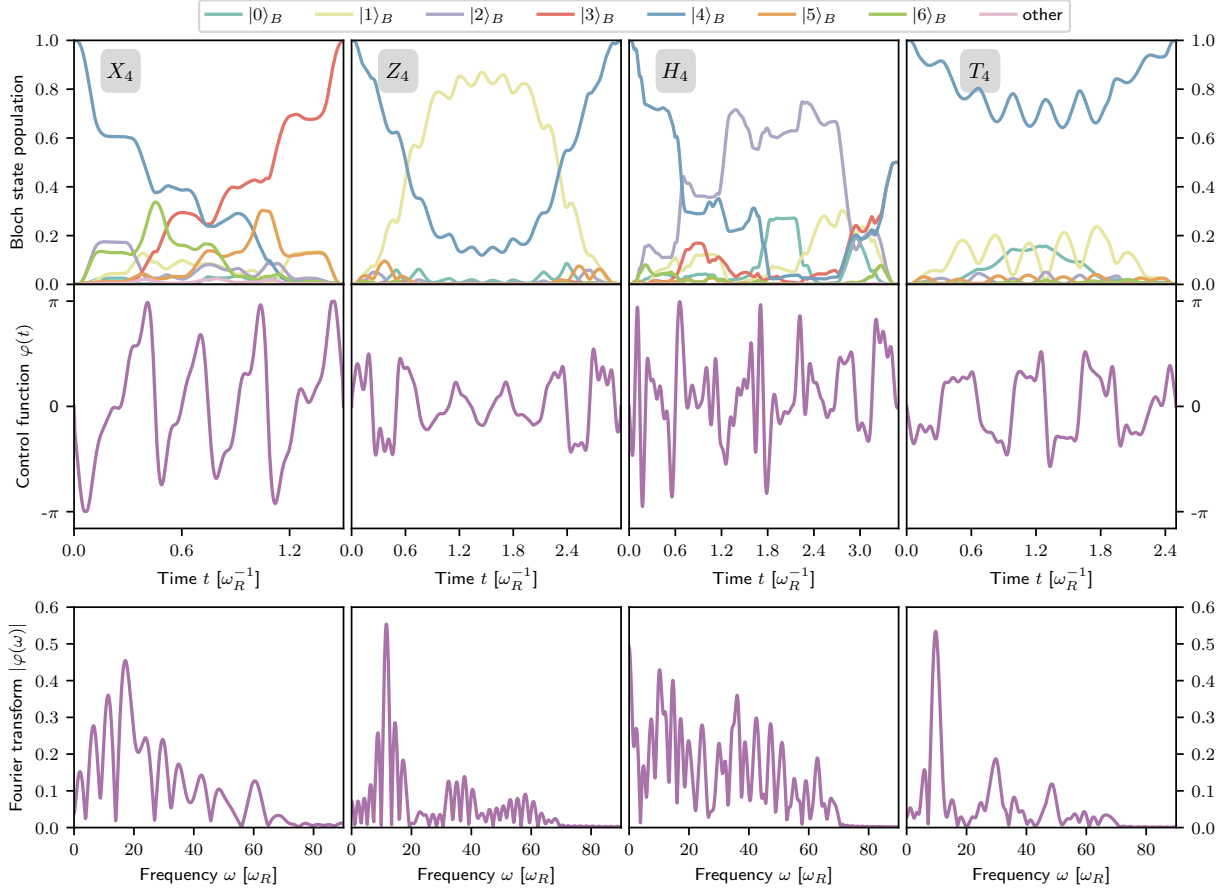


Figure 4.4: Realization of logical qubit gates on the Π_4 qubit. Each column shows the state evolution of one basis state, the control function as well as its Fourier transform. The desired frequency cutoff at $\omega_c = 70\omega_R$ is achieved through an additional term in the objective and enables smooth waveforms. The gate properties (duration T , infidelity $\mathcal{I}$) are $(1.50\omega_R^{-1}, 4.05 \cdot 10^{-3})$ for X_4 , $(3.00\omega_R^{-1}, 2.13 \cdot 10^{-3})$ for Z_4 , $(3.50\omega_R^{-1}, 9.39 \cdot 10^{-5})$ for H_4 and $(2.50\omega_R^{-1}, 3.61 \cdot 10^{-4})$ for T_4 .

$|\nu - 1\rangle_B$ and $|\nu k_L\rangle_p$. Therefore, if the atoms are in a split state and propagating apart, the application of H_ν splits each arm $|\pm\nu k_L\rangle_p$ again to then create four arms in total – two positive momentum and two negative momentum arms. However, when two arms come together, interfere and an H_ν is applied, they may merge and propagate in only one direction as they were mapped to an (approximate) momentum eigenstate.

The operations in Figure 4.4 are required to create branches of a multi-armed interferometer to measure, for instance, gravity gradients and rotations [181, 24] as well as manipulate the relative phases and populations between arms. The X_ν gate performs a π -rotation about the X -axis in the Bloch sphere, which leaves the population in the states $|\pm\nu k_L\rangle_p$ invariant but flips their relative phase. More generally, the design of arbitrary $X_\nu(\phi)$ rotations would allow to impart different relative phases on the interferometer arms in order to, for instance, shift the phase of the interference fringe pattern.

“Partial mirror” operations can be achieved with $Z_\nu(\theta)$ rotation gates as they move population between the momentum states:

$$Z_\nu(\theta) |\pm\nu k_L\rangle_p \approx \cos\left(\frac{\theta}{2}\right) |\pm\nu k_L\rangle_p - i \sin\left(\frac{\theta}{2}\right) |\mp\nu k_L\rangle_p. \quad (4.17)$$

Clearly $\theta = \pi$ recovers the standard **MIRROR** $_\nu$, but this also shows the effect of the T_ν gate, which corresponds to $\theta = \pi/4$, resulting in an unbalanced fraction of the interferometer arm ($\sim \sin^2(\pi/8)$) changing direction. The squared operation $T_\nu^2 = \sqrt{Z_\nu}$ can then be viewed as a “half mirror” since $\cos^2(\pi/4) = \sin^2(\pi/4) = 1/2$.

From Figure 4.4, we note the many distinct state transfer pathways between Bloch states achieved by the control function for each gate. The X_4 gate achieves population inversion in the Π_4 qubit subspace by temporarily transferring to the Π_6 qubit subspace as well as the lower energy state $|2\rangle_B$. The broad Fourier spectrum reflects the presence of these multiple transitions. The situation is different for the phase gates Z_4 and T_4 where specific transitions to the valence band dominate, leading to a dominant Fourier component at $\omega \approx 11 \omega_R$. This

supports the heuristic usage of this frequency in prior works to design control functions, for instance with reinforcement learning methods [103, 30]. The use of quantum optimal control makes this restrictive design assumption unnecessary, which is one of the advantages as described in Section 4.1.2. Last, the Hadamard gate H_4 employs many intermediate Bloch states in order to transform the basis due to a strict phase convention.

For completeness, and to demonstrate the advantages and generality of the direct optimal control method, we want to realize the Z_6 gate, which mirrors the interferometer arms at a higher momentum $|\pm 6k_L\rangle_p$. Computationally, this case is more complicated than the conventional mirror in the lowest conduction band qubit (Z_4) due to the involvement of a larger set of Bloch states. It is helpful to think about this problem classically to gain intuition for this task. A mirror inverts momentum, thus we aim to perform the transformations $|\pm 2nk_L\rangle_p \rightarrow |\mp 2nk_L\rangle_p$. The operators $b_{\pm}$ can only alter the momentum in steps of $\pm 2k_L$ ($n \rightarrow n \pm 1$), thus it is reasonable to expect that for larger momenta the complexity of control function achieving the Z_{ν} (MIRROR_{ν}) gate increases as a greater momentum difference has to be imparted in gradual steps. Furthermore, it is conceivable that a standard (indirect) optimal control method might struggle at the task as it would only optimize over the control parameters φ_l and could not immediately exploit any knowledge about the quantum state trajectory, highlighting the potential advantages of the direct trajectory optimization approach we employ here.

In particular, classical intuition allows us to form additional constraints for such high-momentum gates to aid their design. Classically, there must be a point in time during mirroring where the atoms are stationary. Due to time-reversal symmetry, we expect this to occur at time $T/2$. Since we are designing our gates in the basis of the Bloch states, we are interested in the highest-energy Bloch state that has a significant zero-momentum component. For our choice of lattice depth $V_0 = 10\omega_R$, this is the state $|2\rangle_B$, which is closest in energy to the initial ($|\nu\rangle_B$, $|\nu - 1\rangle_B$) and target split states ($Z_{\nu}|\nu\rangle_B$, $Z_{\nu}|\nu - 1\rangle_B$)

and, due to its zero-momentum component, allows for transition through the stationary state. This is why one could expect the state $|2\rangle_B$ to be populated at $T/2$. We incorporate this intuition via an additional constraint on the time evolution operator $U_{l'} = x_{l'}$ at time $t_{l'} = T/2$:

$$|\langle 2|_B x_{l'} |\nu\rangle_B|^2 \geq 0.1, \quad |\langle 2|_B x_{l'} |\nu-1\rangle_B|^2 \geq 0.1. \quad (4.18)$$

This forces the population of the state $|2\rangle_B$ to be at least 10% at time $T/2$. We choose 10% to inform the optimization that populating this state is required while aiming not to be too restrictive, otherwise the specific value is not of great relevance. It is important to keep in mind that this modification is directly enabled by the chosen optimal control method and would be nontrivial to realize with alternative frameworks that were used in Refs. [103, 30, 168, 134] for instance.

We now apply this extended problem formulation to design control functions for the Z_6 gate. Figure 4.5 shows the state evolution as well as the control function waveform. As expected, a significantly higher gate duration is required, and we note the expected large occupation of the state $|2\rangle_B$ at time $T/2$. The population of this state is much larger than our lower bound of 10%, thus verifying our reasoning for implementing the additional constraints.

Interqubit gates

With the introduction of the SPLIT_ν , we already provided examples of an operation that maps between qubit subspaces. This mapping considered only two states with one drawn from each qubit subspace. One can however imagine a broader set of useful operations $W : \Pi_\nu \leftrightarrow \Pi_{\nu'}$ that map between qubit subspaces. Within the conduction band, this includes $\text{BOOST}_\nu : \Pi_\nu \rightarrow \Pi_{\nu+2}$ or $\text{SLOW}_\nu : \Pi_\nu \rightarrow \Pi_{\nu-2}$ operations which accelerate or decelerate the atoms, respectively. Furthermore, a mapping between conduction and valence band qubit subspaces can serve to **HOLD** or **RELEASE** atoms, for instance, as steps in a lattice

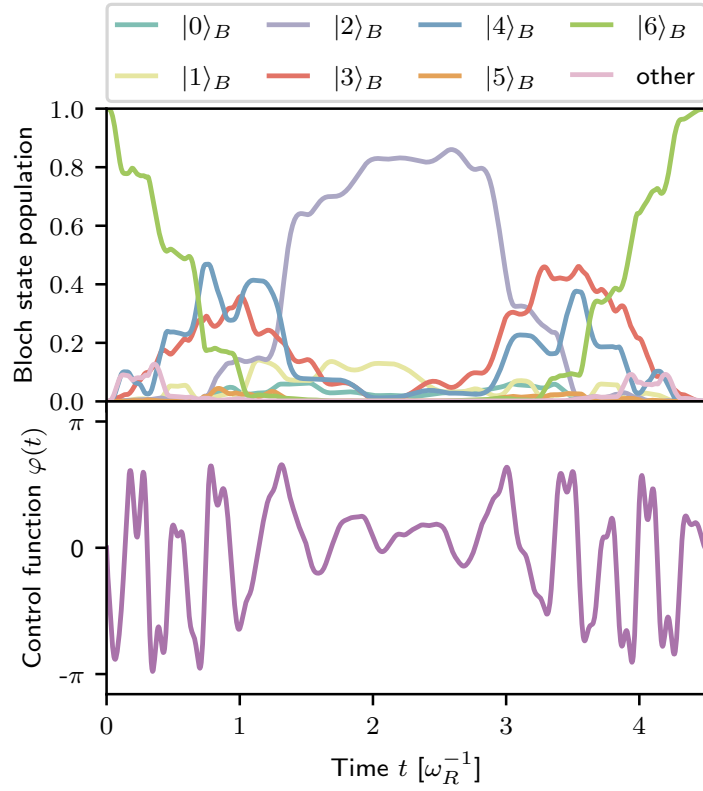


Figure 4.5: Z_6 gate acting on qubit Π_6 , $T = 4.50\omega_R^{-1}$, $\mathcal{I} = 2.90 \cdot 10^{-3}$. It is equivalent to a MIRROR operation for atoms in $|\pm 6k_L\rangle_p$. The top panel shows the state evolution in the Bloch basis while the lower panel shows the corresponding control function.

hold procedure designed to increase the space-time area enclosed within a spatially confined system [207]. We discuss results for these operations connecting the qubit subspaces Π_1 and Π_4 .

In order to optimize a subspace transfer, we propagate the initial qubit subspace spanned by $|\psi_0(t=0)\rangle = |\nu\rangle_B$ and $|\psi_1(t=0)\rangle = |\nu-1\rangle_B$ and project the final states into the target qubit subspace $\Pi_{\nu'}$. The infidelity

$$\mathcal{I}(T; \gamma) = 1 - \frac{1}{2} \sum_{k=0}^1 \langle \psi_k(T; \gamma) | \Pi_{\nu'} | \psi_k(T; \gamma) \rangle \quad (4.19)$$

describes the amount of population that is not transferred into the desired subspace, thus $\mathcal{I}(T; \gamma) = 0$ corresponds to a perfect transfer. For this operation it is not relevant where exactly in the target Bloch sphere the states in the initial qubit subspace are mapped to; when the transfer operation is optimized and fully characterized, additional corrections inside the target qubit subspace can be performed via single-qubit rotations, which were discussed above.

In Figure 4.6 we show the results for the **BOOST**₄ and **HOLD** operations. As expected, when evolving a basis state of the initial manifold we do not arrive at a basis state of the target manifold but instead in a superposition. This is due to the particular choice of infidelity function in Equation (4.19) which only fixes the target manifold but not a particular state. It allows for more freedom for the optimization compared to optimizing for a fixed unitary gate acting on both subspaces. We note that the related inverse actions **SLOW**₆ and **RELEASE** can be obtained via time-reversal symmetry of the corresponding control functions.

4.2.2 Conclusion

We demonstrated how the direct trajectory optimization can be used to design control functions for operations relevant to inertial sensing on various qubit subspaces of an atom con-

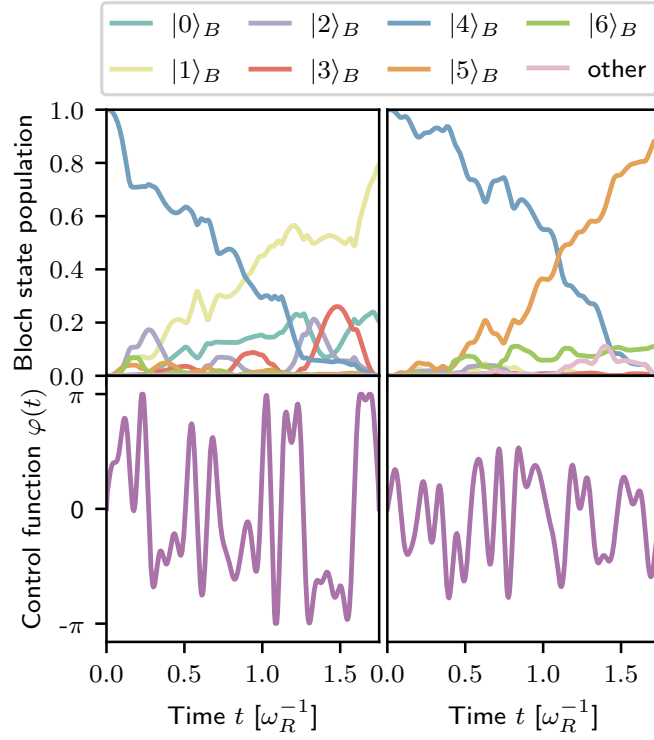


Figure 4.6: Control functions for the **HOLD** operation (left, $\mathcal{I} = 1.37 \cdot 10^{-3}$) and the **BOOST₄** operation (right, $\mathcal{I} = 9.13 \cdot 10^{-5}$). For both operations the gate time is $T = 1.75 \omega_r^{-1}$.

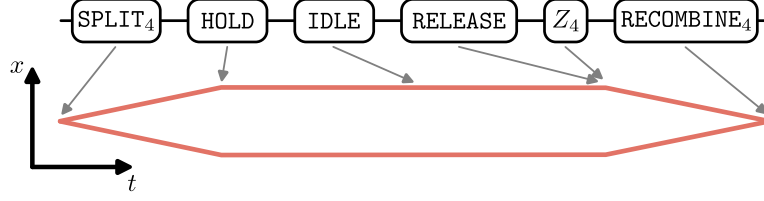


Figure 4.7: Space-time representation of a simple circuit for a lattice-hold interferometer sequence written using the gates proposed in this work.

finned in a one-dimensional optical lattice. We presented a framework that breaks the complex control space of the optical lattice interferometer architecture into fundamental computational units that can be acted on by quantum gates in analogy to quantum computing. We identified two-level subspaces which are embedded in the infinite-dimensional optical lattice system and can be thought of as logical qubit subspaces that describe a specific momentum configuration of atoms in the lattice. In this section we demonstrated the design of control functions that implement physically motivated transformations on and between these logical qubit subspaces.

The larger aim of this work is to introduce “programmability” into this platform by gathering a catalog of useful operations, which allow abstraction and consequently sensing-protocol design which rises above the level of the control function dynamics themselves. This has a direct analogy to the kinds of abstractions which occur in the field of quantum computing where algorithms are expressed in terms of quantum circuits. Using such primitives, various sensing protocols for the optical lattice interferometer can be expressed in the circuit language. Figure 4.7 shows an example of a circuit of gates that realize an interferometry protocol where the atom cloud is split, propagated, held, mirrored and eventually recombined, allowing for the enclosure of a large space-time area (and thus high sensitivity) without increasing the spatial size of the lattice. More complex protocols that involve interqubit operations and qubit subspace operations can be developed without having to return to the (non-intuitive) level of the control function and Bloch state transfer dynamics.

Much as quantum computing, this transition from pulse to circuit level design has obvious

practical advantages also for quantum sensing. For inertial sensors, one can easily imagine a catalog of operations (those discussed in Section 4.2.1 constituting only a small subset) that can be used to assemble sensing protocols in one, two, and three dimensions in order to sense vector components of acceleration, rotation, gravity, and force gradients [102, 203]. This also includes calibrating protocols which are sensitive to internal settings of a device that are subject to drift as well as the changing inertial and electromagnetic environment of a device [3]. Just like in quantum computing, one could imagine utilizing the catalog of operations available on an inertial sensor in order to characterize and verify performance of a device [78]. This could include the equivalent of mid-circuit measurements via (for instance) nondestructive measurements or the transfer of atomic populations into untrapped states [13, 166]. Here the prospect of correcting errors in quantum sensors has been the subject of recent interest [210, 91], highlighting the need for protocols to generate quantum resources (such as entanglement and squeezing) that are useful for inertial sensing. [176, 70, 200, 25, 18, 90].

Towards the conclusion of this project, we became aware of related work [104] taking a similar approach to gate sets as well as their implementation for optical lattice interferometers. However, they employ a different control optimization framework and do not draw parallels to the qubit language as we did.

4.3 End-to-end optimization of interferometer sensitivity

The effectiveness of any sensing technology is ultimately determined by its sensitivity to the physical quantity being measured. Given a parameter θ we wish to sense, we can think of the sensor’s sensitivity as related to the variance $\text{Var}(\hat{\theta})$ of an unbiased estimator $\hat{\theta}$ obtained from measurement outcomes. The Cramér-Rao bound [135]

$$\text{Var}(\hat{\theta}) \geq \frac{1}{k \mathcal{I}(\theta)} \tag{4.20}$$

sets a lower bound on the sensitivity that we can infer after k trials. It is determined by the (classical) Fisher information of a single sample

$$\mathcal{I}(\theta) = \mathbb{E}_{X;\theta} \left[\left(\frac{\partial \log p(X; \theta)}{\partial \theta} \right)^2 \right] = \int dX \frac{1}{p(X; \theta)} \left(\frac{\partial p(X; \theta)}{\partial \theta} \right)^2, \quad (4.21)$$

where X denotes the random variable describing the stochastic measurement with distribution $p(X; \theta)$.

In the context of optical lattice interferometry, this motivates the design of control pulses that, instead of realizing predefined primitives as in Section 4.2, are directly optimized to boost the Fisher information in order to enable a very fine sensitivity of the sensor. In this final part of the chapter, we discuss how to exploit the structure of direct trajectory optimization and adapt the Piccolo.jl framework [182] to find such control pulses in an end-to-end fashion.

4.3.1 Sensing accelerations

As a reminder, we are specifically interested to construct an accelerometer from our cold atom platform, i.e., a quantum sensor that can sense (1D) accelerations of the system. This is modeled by an additional force $F(t)$ acting on the atoms (mass m), which yields the single-particle Hamiltonian

$$\mathcal{H}(t) = \frac{p^2}{2m} + V(x, t) - F(t)x. \quad (4.22)$$

The term $V(x, t)$ describes the optical lattice potential defined in Equation (4.2). The TDSE (2.14) in momentum space reads

$$i \frac{\partial \psi(p, t)}{\partial t} = \left(\frac{p^2}{2m} + V \left(-i \frac{\partial}{\partial p}, t \right) - i F(t) \frac{\partial}{\partial p} \right) \psi(p, t). \quad (4.23)$$

Here p denotes the total momentum of a particle. With the help of intuition from classical mechanics, we expect an evolution of the momentum that is related to the external force $\dot{P}(t) = F(t)$, thus we perform a variable transformation

$$\tilde{p} = p - P(t), \quad \tilde{t} = t. \quad (4.24)$$

This lets us rewrite the partial derivatives,

$$\begin{aligned} \frac{\partial}{\partial p} &= \frac{\partial \tilde{p}}{\partial p} \frac{\partial}{\partial \tilde{p}} + \frac{\partial \tilde{t}}{\partial p} \frac{\partial}{\partial \tilde{t}} = \frac{\partial}{\partial \tilde{p}} \\ \frac{\partial}{\partial t} &= \frac{\partial \tilde{p}}{\partial t} \frac{\partial}{\partial \tilde{p}} + \frac{\partial \tilde{t}}{\partial t} \frac{\partial}{\partial \tilde{t}} = -\dot{P}(\tilde{t}) \frac{\partial}{\partial \tilde{p}} + \frac{\partial}{\partial \tilde{t}} = -F(\tilde{t}) \frac{\partial}{\partial \tilde{p}} + \frac{\partial}{\partial \tilde{t}}, \end{aligned} \quad (4.25)$$

and obtain the TDSE in the new variables:

$$i \left(-F(\tilde{t}) \frac{\partial}{\partial \tilde{p}} + \frac{\partial}{\partial \tilde{t}} \right) \psi(\tilde{p} + P(\tilde{t}), \tilde{t}) = \left(\frac{(\tilde{p} + P(\tilde{t}))^2}{2m} + V \left(-i \frac{\partial}{\partial \tilde{p}}, \tilde{t} \right) - iF(\tilde{t}) \frac{\partial}{\partial \tilde{p}} \right) \psi(\tilde{p} + P(\tilde{t}), \tilde{t}) \quad (4.26)$$

$$i \frac{\partial}{\partial \tilde{t}} \psi(\tilde{p} + P(\tilde{t}), \tilde{t}) = \left(\frac{(\tilde{p} + P(\tilde{t}))^2}{2m} + V \left(-i \frac{\partial}{\partial \tilde{p}}, \tilde{t} \right) \right) \psi(\tilde{p} + P(\tilde{t}), \tilde{t}) \quad (4.27)$$

$$i \frac{\partial}{\partial \tilde{t}} \tilde{\psi}(\tilde{p}, \tilde{t}) = \left(\frac{(\tilde{p} + P(\tilde{t}))^2}{2m} + V \left(-i \frac{\partial}{\partial \tilde{p}}, \tilde{t} \right) \right) \tilde{\psi}(\tilde{p}, \tilde{t}). \quad (4.28)$$

In the last step we relabeled $\tilde{t} \rightarrow t$ (essentially used $\tilde{t} = t$) and introduced the new wavefunction $\psi(\tilde{p} + P(t), t) = \tilde{\psi}(\tilde{p}, t)$. $\tilde{p}$ can be understood as the momentum deviation from the classical momentum $P(t)$ which is fully determined by the force trajectory $F(t)$, and thus the wavefunction $\tilde{\psi}(\tilde{p}, t)$ now captures the quantum state with respect to this moving reference frame. In order to draw analogies to derivations in the previous sections, we relabel $\tilde{p} \rightarrow p$ and $\tilde{\psi} \rightarrow \psi$ but need to keep the change in reference frame (i.e., the transformations discussed) in mind. Going forward we assume the initial condition $P(0) = 0$.

Using a similar analysis that lead to Equation (4.4), we can write down the effective Hamiltonian that takes into account that the momentum (relative to $P(t)$) can only take values $p = 2nk_L$ for $n \in \mathbb{Z}$,

$$\begin{aligned}\mathcal{H}(t) &= \frac{(2k_L + P(t))^2}{2m} - \frac{V_0}{4} \left(e^{i\varphi(t)} b_- + e^{-i\varphi(t)} b_+ \right) \\ &= 4\omega_R n^2 + k_L n \frac{P(t)}{m} + \frac{P(t)^2}{2m} - \frac{V_0}{4} \left(e^{i\varphi(t)} b_- + e^{-i\varphi(t)} b_+ \right) \\ &\cong 4\omega_R n^2 + k_L n \frac{P(t)}{m} - \frac{V_0}{4} \left(e^{i\varphi(t)} b_- + e^{-i\varphi(t)} b_+ \right).\end{aligned}\tag{4.29}$$

We denote the single-particle state with $|\psi(t)\rangle$ as usual and remember its meaning in the moving reference frame. Equation (4.29) holds up to the “constant” (non-operator) term $P(t)^2/2m$, which would only influence the global phase of the quantum state $|\psi(t)\rangle$ and is therefore not observable.

Constant accelerations

Going forward we restrict our discussion to the case of constant accelerations a , thus we have $F = ma = \text{const.}$ and $P(t) = mat$. The final Hamiltonian reads

$$\mathcal{H}(t; a) = 4\omega_R n^2 + 2k_L ant - \frac{V_0}{4} (I(t)(b_- + b_+) + iQ(t)(b_- - b_+)),\tag{4.30}$$

where we also introduced the quadrature controls $f_1 = I$ and $f_2 = Q$ from Equation (4.8) again so that the Hamiltonian is linear in all control terms. Technically, the second component $2k_L ant$ corresponds to a non-controllable drift term of the system, however, in QOC tools like Piccolo.jl, time-dependent Hamiltonian terms are considered controls. We integrate this by considering the additional control “pulse” f_3 that couples to the operator $2k_L an$ and enforcing the constraint

$$f_3(t) = t.\tag{4.31}$$

Such constraints can easily be configured in `Piccolo.jl`.

4.3.2 Fisher information optimization

Measurement of the atom cloud after a sensing protocol yields information about the population of different momentum states. We denote the probability of observing a particle with (relative) momentum $p = 2nk_L$ at time t given the system acceleration a by $P(n|a; t) = |\langle n|\psi(t; a)\rangle|^2$, which corresponds to a probability distribution. The Fisher information (4.3.2) for our sensing parameter of interest $\theta = a$ is thus given by

$$\mathcal{I}(a; t) = \sum_n \frac{1}{P(n|a; t)} \left(\frac{\partial P(n|a; t)}{\partial a} \right)^2. \quad (4.32)$$

Ultimately we are interested in the final time of the sensing protocol $t = T$ in order to make statements about the sensitivity, but it shall prove useful for the optimization to consider intermediate times as well.

A critical component of the Fisher information expression (4.32) is the partial derivative of the population with respect to the sensing parameter. It can be computed via

$$\frac{\partial P(n|a; t)}{\partial a} = \partial_a |\langle n|\psi(t; a)\rangle|^2 = 2 \operatorname{Re}[\langle n|\partial_a \psi(t; a)\rangle \langle \psi(t; a)|n\rangle]. \quad (4.33)$$

Thus we require the dynamics of the derivative state $\partial_a \psi(t; a)$, which can be obtained by differentiating the TDSE (2.14) with Hamiltonian $\mathcal{H}(t; a)$ from Equation (4.30):

$$i\partial_t |\partial_a \psi(t; a)\rangle = (\partial_a \mathcal{H}(t; a)) |\psi(t; a)\rangle + \mathcal{H}(t; a) |\partial_a \psi(t; a)\rangle. \quad (4.34)$$

The dynamics of $\partial_a \psi(t; a)$ couples to the dynamics of $\psi(t; a)$, leading to non-unitary dynam-

ics of $\partial_a \psi(t; a)$. By defining an *augmented state vector*

$$|\underline{\psi}(t; a)\rangle = \begin{bmatrix} |\psi(t; a)\rangle \\ |\partial_a \psi(t; a)\rangle \end{bmatrix}, \quad (4.35)$$

we can write the dynamics succinctly in TDSE form as

$$i\partial_t |\underline{\psi}(t; a)\rangle = \underline{\mathcal{H}}(t; a) |\underline{\psi}(t; a)\rangle \quad (4.36)$$

with *augmented Hamiltonian*

$$\underline{\mathcal{H}}(t; a) = \begin{bmatrix} \mathcal{H}(t; a) & 0 \\ \partial_a \mathcal{H}(t; a) & \mathcal{H}(t; a) \end{bmatrix}. \quad (4.37)$$

From Equation (4.30) we can deduce $\partial_a \mathcal{H}(t; a) = 2k_L n t$. The augmented Hamiltonian $\underline{\mathcal{H}}(t; a)$ is not Hermitian, therefore the evolution determined by Equation (4.36) is not unitary. This is because the norm of the derivative state $\partial_a \psi(t; a)$ is not preserved (it is not a physical state), which is a desired feature as we want its components to grow such that the Fisher information increases (see Equations (4.32) and (4.33)). We can compute the Fisher information at the final time T of a sensing protocol by propagating the initial augmented state

$$|\underline{\psi}(0; a)\rangle = \begin{bmatrix} |\psi(0)\rangle \\ 0 \end{bmatrix} \quad (4.38)$$

under the generalized TDSE (4.36) and evaluating Equation (4.32) at $t = T$:

$$\mathcal{I}(a; T) = \sum_n \frac{4(\text{Re}[\langle n | \partial_a \psi(T; a) \rangle \langle \psi(T; a) | n \rangle])^2}{|\langle n | \psi(T; a) \rangle|^2}. \quad (4.39)$$

Using direct trajectory optimization (see Section 4.1.2), we seek to find sensing protocols, i.e., determine control functions for the phase φ through the helper quadratures I and Q , that maximize the Fisher information $\mathcal{I}(a; T)$. The state variables x_l now describe the augmented

state vector at discrete times t_l , $|\underline{\psi}(t_l; a)\rangle = x_l$, while the control variables γ_{jl} parameterize the φ , I and Q as before, but also the “time control”:

$$\varphi(t_l) = \gamma_{0l}, \quad I(t_l) = \gamma_{1l}, \quad Q(t_l) = \gamma_{2l}, \quad t_l = \gamma_{3l}. \quad (4.40)$$

The last expression serves as the constraint (4.31) that is required to implement the effect of static accelerations a . The phase modulation φ , the control that we ultimately care about, is included in the optimization through auxiliary parameters γ_{0l} , which do not appear in the Hamiltonian but are crucial for the constraint (4.8) that connects the quadratures. Furthermore, this way the desired solution for $\varphi(t)$ can be directly extracted after the optimization.

The objective function $\mathcal{L}$ we seek to minimize is primarily given by the negative Fisher information, $\mathcal{L}(x_l, \gamma_{jl}) = -\mathcal{I}(a; T) = \mathcal{L}(x_{N_t})$. Just like in Section 4.1.2, the “directness” of the optimization approach is helpful as the Fisher information and therefore the objective function are solely determined by the augmented state at the final time. Thus they can be computed from a small subset of directly accessible optimization variables. Just like before, we furthermore add an Fourier penalty term $\mathcal{L}_\omega$ as shown in Equation (4.12) to the optimization in order to encourage smooth control solutions. The final form of the optimization problem now reads

$$\hat{x}_l, \hat{\gamma}_{jl} = \arg \min_{x_l, \gamma_{jl}} -\mathcal{I}(a; T) + \alpha \mathcal{L}_\omega(\gamma_{0l}) \quad (4.41)$$

$$\begin{aligned} \text{subject to: } & x_{l+1} - \exp(-i\mathcal{H}(\gamma_{jl})\Delta t)x_l = 0, & x_1 &= x(0), \\ & \gamma_{1,l} = \cos(\gamma_{0,l}), & \gamma_{2,l} &= \sin(\gamma_{0,l}), \\ & \gamma_{3,l+1} = \gamma_{3,l} + \Delta t, & \gamma_{3,1} &= 0, \\ & \gamma_{0,1} = \gamma_{0,N_t} = 0, & & \end{aligned} \quad (4.42)$$

where α is a hyperparameter that determines the weight of the Fourier penalty term and is chosen empirically.

Insertion of free propagation windows

Due to fine time discretization and large number of optimization variables typically encountered when employing QOC, trying to solve for sensing protocols in this end-to-end fashion becomes infeasible as the duration T grows. In conventional sensing setups using optical lattice interferometry, the time period of active control (e.g., during a **SPLIT** or **MIRROR** operation) is much smaller than the time period when the atoms propagate freely [30]. It is during these free propagation windows when the arms of the interferometer accrue significant relative phase differences due to the influence of the external parameter that we want to sense, and thus those time windows contribute strongly to the final Fisher information.

For this reason we want to impose this structure on our end-to-end problem as well. While this restricts the set of possible control schemes we can optimize over, it allows us to push to long-duration sensing protocols. We achieve this by once again exploiting the design of direct trajectory optimization, here specifically the fact that the dynamics of the augmented state $|\underline{\psi}(t; a)\rangle$ is implemented via constraints as is shown in Equations (4.42). In Piccolo.jl [182], the propagator $\exp(-i\underline{\mathcal{H}}(\gamma_{jI})\Delta t)$ over the small time step Δt is implemented using Padé approximants [123], which leads to approximation errors as ΔT is larger. We insert free propagation windows by replacing these approximated propagators in a small number of constraints with the exact exponential propagator $\underline{U}(t', t' + t_{\text{free}}; a, \varphi = 0)$ over a longer time interval $t_{\text{free}} \gg \Delta t$, which is obtained from numerically integrating the “TDSE” (recall that the evolution of the augmented state is non-unitary)

$$i \frac{\partial \underline{U}(t', t; a, \varphi = 0)}{\partial t} = \underline{\mathcal{H}}(t; a) \underline{U}(t', t; a, \varphi = 0), \quad (4.43)$$

over the domain $[\tilde{t}, \tilde{t} + t_{\text{free}}]$. The time windows for the free propagation are set a priori. We furthermore impose the additional constraint that the lattice is in its unperturbed state during this stage, hence the specification $\varphi = 0$ above. These two design choices bring the

advantage that the propagator $\underline{U}(t', t' + t_{\text{free}}; a, \varphi = 0)$ can be computed independently before the actual optimization because it does not depend on any control parameters, therefore this more costly computation is only performed once and the result stored for use when evaluating the modified constraints.

In order to implement such a window at index l' , we modify the following optimization constraints from Equations (4.42):

$$x_{l'+1} - U(t_{l'}, t_{l'} + t_{\text{free}}; a, \varphi = 0)x_{l'} = 0, \quad (4.44)$$

$$\gamma_{0,l'} = 0, \quad (4.45)$$

$$\gamma_{3,l'+1} = \gamma_{3,l'} + t_{\text{free}}. \quad (4.46)$$

Constraint (4.44) shows the replacement of the small-time-step propagator with the window's precomputed propagator, while constraint (4.46) accounts for the larger jump in time. In constraints (4.45) we capture the additional problem structure of $\varphi = 0$ during the window.

End-to-end optimization result

The standard Mach-Zehnder interferometer sequence, which has been studied extensively in the context of optical lattice tomography [30, 29, 103], is given by (see Section 4.2 for terminology)

$$\text{SPLIT} - \text{FREE PROP.} - \text{MIRROR} - \text{FREE PROP.} - \text{RECOMBINE} (= \text{SPLIT}^{-1}). \quad (4.47)$$

Usually the nontrivial components **SPLIT** and **MIRROR** are optimized with QOC individually, but here we want to use the end-to-end approach to boost the sensitivity while preserving the structure of the protocol. Like in the previous works, we are specifically interested in sensing accelerations around the point $a = 0$.

We consider three actively controlled components of duration $0.3 \cdot 2\pi\omega_R^{-1}$, $0.6 \cdot 2\pi\omega_R^{-1}$

and $0.3 \cdot 2\pi\omega_R^{-1}$, respectively, interspersed with two free propagation windows of duration $2.0 \cdot 2\pi\omega_R^{-1}$ each. Therefore, the duration of the full protocol is $T = 5.2 \cdot 2\pi\omega_R^{-1} \approx 32.7 \omega_R^{-1}$. The components are initialized with control solutions for **SPLIT**₃, **MIRROR**₄ and **RECOMBINE**₃, respectively, which we obtain like described in Section 4.2. Thus, our starting point is a Mach-Zehnder interferometer sequence that puts the atoms in a superposition of momentum states $|\pm 4k_L\rangle$ (approximately) during propagation, and therefore the momentum difference between the interferometer arms is $\Delta p \approx 8k_L$. For the idealized Mach-Zehnder interferometer sequence with instant state manipulation (control components have zero duration) and total duration T , the Fisher information is given by [104]

$$\mathcal{I}_{\text{ideal}}(a = 0; T) = \left(\Delta p \left(\frac{T}{2} \right)^2 \right)^2, \quad (4.48)$$

and we compare our optimization result to this baseline value.

Starting from the Mach-Zehnder initialization, we apply the end-to-end trajectory optimization, where we employ the free-propagation modification to Piccolo.jl in order to simulate the two non-controlled windows of duration $2.0 \cdot 2\pi\omega_R^{-1} \approx 12.6 \omega_R^{-1}$. Figure 4.8 shows the evolution of (a) the quantum state as well as the (b) Fisher information for the obtained solution, where the latter is normalized to the idealized case $\mathcal{I}_{\text{ideal}}(a = 0; T)$. We observe that, while the individual control components still look similar to the starting configuration (compare to Figure 4.3 for **SPLIT**₃ and Figure 4.4 for **MIRROR**₄ = Z_4), the end-to-end optimization finds modifications to the components that ultimately yield a Fisher information that is three times as high as in the standard ideal case, thus improving the sensor's sensitivity by a factor of three.

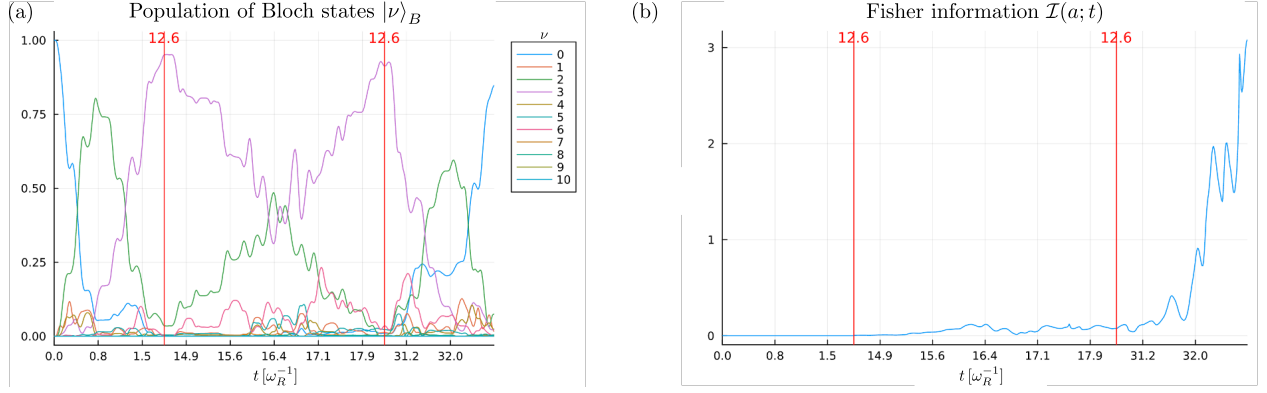


Figure 4.8: Optimization of the accelerometer’s sensitivity in an end-to-end fashion. The red vertical lines indicate the insertion of free-propagation windows of duration $2.0 \cdot 2\pi\omega_R^{-1} \approx 12.6 \omega_R^{-1}$ in order to mimic the structure of a Mach-Zehnder interferometer. (a) Evolution of the Bloch states throughout the protocol. (b) Evolution of the (classical) Fisher information, normalized to the Fisher information of the idealized Mach-Zehnder interferometer $\mathcal{I}_{\text{ideal}}(a = 0; T)$ shown in Equation (4.48).

4.3.3 Conclusion

Leveraging the framework of direct trajectory optimization, we demonstrated the implementation of an end-to-end search to boost the sensitivity of an accelerometer based on optical lattice interferometry. By including both the quantum state $|\psi\rangle$ as well as the derivative state $|\partial_a\psi\rangle$ as optimization variables, we were able to compute their time evolution efficiently in an augmented system and directly evaluate the relevant objective function, the Fisher information, at the final time T . Through manipulation of the original set of constraints, we imposed additional structure on the optimization problem (free-propagation windows) that was inspired by existing work on interferometry. This aided the optimization process and enabled finding sensing sequences of longer duration.

So far these additional modifications were designed by hand. In our implementation the free-propagation windows have to be pre-computed and fixed in time a priori, which allows the optimization to run fast, however, this also limits the achievable improvement. This opens the door for many opportunities to extend this framework. Making the window duration tunable or allowing for a flexible number of windows, thus relying on less assumptions about

the concrete sensing protocol, may lead to further improvements but also adds complexity to the optimization task. This could require combining direct trajectory optimization with other optimization techniques, such as reinforcement learning, to explore the parameter space.

We particularly focus on the task of sensing 1D accelerations, therefore two natural extensions are given by increasing the system dimensionality or sensing multiple parameters at the same time. Recent work [3] has started to look at the latter, however the authors employ reinforcement learning to solve for the control functions, which is less expressive than the direct trajectory optimization approach we discussed.

CHAPTER 5

CONTROL OF A NOISE-PROTECTED QUBIT SYSTEM

To build a useful quantum computer which can solve problems of practical importance, it is crucial to increase the size of quantum computing systems while ensuring sufficiently low component error rates. In recent years, the field of quantum computing has witnessed significant progress as larger systems with lower error rates have been realized with various hardware platforms [15, 2, 150, 140, 189, 127]. Despite the rapid progress, however, it remains a challenge to produce a large quantum computing system in a reliable manner. Realizing quantum error correction (QEC) will be a vital step towards achieving this goal as it enables logical pushing error rates below the individual components' physical error rates.

As we discussed in Section 2.4.2, QEC usually follows the approach of encoding logical qubit information within a great number of physical qubits and exploiting this redundancy to detect and correct errors during computation. However, there exists an alternative approach: Instead of using many copies of small qubit systems, one can also consider a single system with a much larger Hilbert space. For instance, in *bosonic error correction* [68, 19], one encodes a logical qubit in a single bosonic mode, that is, a quantum harmonic oscillator (see Section 2.1.3).

Cat codes [19] are an example of such an error correction scheme. In their simplest form, the qubit states are chosen to be *coherent states* $|\pm\alpha\rangle$, $\alpha \in \mathbb{C}$, where

$$|\alpha\rangle = e^{-\frac{|\alpha|^2}{2}} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle. \quad (5.1)$$

We do not go into further background and details on bosonic codes here as this is not relevant for this chapter. However, we like to point out the special property of this qubit encoding:

$$a |\pm\alpha\rangle = \pm\alpha |\pm\alpha\rangle. \quad (5.2)$$

Coherent states are eigenstates of the annihilation operator a , therefore they are robust to excitation loss, which is one of the dominant error sources in many systems such as superconducting transmons (recall Section 2.4). Therefore, by choosing this encoding, we obtain a qubit that is inherently robust to this type of error. This allows for additional QEC schemes to focus on other error sources and thus potentially reduces the resource overhead required for QEC.

In this chapter, we discuss the use of such noise-protected qubits to mitigate another source of error encountered in superconducting systems: frequency noise. In the following sections we argue how an encoding similar to the one mentioned above enables the suppression of this type of noise and we furthermore demonstrate gate schemes for universal computation. The latter poses a challenge in these systems, which we explain in Section 5.2.1. This work represents a joint effort with the theory team at the AWS Center for Quantum Computing.

5.1 Mitigating frequency noise on the physical level

The performance of a large-scale superconducting quantum device is often limited by fabrication disorders and spurious two-level systems (TLSs). In particular, TLSs often move close to a subset of qubits in an unpredictable way and significantly degrade the performance of the affected qubits (e.g., substantially reducing qubit lifetimes) [43, 95, 209, 32, 118].

The negative effects of TLSs are often mitigated by using frequency-tunable transmons [96, 163, 8] and operating them at a frequency that is off-resonant with any nearby TLSs. However, this might require tuning the qubit away from its “sweet spot,” where the frequency is insensitive to its control knob (usually the external flux Φ_x) to first order, as illustrated in Figure 5.1(a). Thus, reduction in the TLS-induced errors on a frequency-tunable qubit comes at the cost of increased frequency noise and dephasing caused by undesired fluctuations in the control knob, without any protection from a first-order insensitivity.

Taking inspiration from this example, throughout this chapter, we seek to propose a general approach for robustly operating qubits encoded in Kerr-nonlinear oscillators that are subject to undesired fluctuations in their frequency (e.g., frequency-tunable transmons). Our proposal is based on Kerr-cat qubits (KCs) [35, 72, 147, 148, 205, 161, 76, 4, 186, 185, 56, 41, 81, 52, 149, 69, 63]. In Section 5.2 we discuss how KCs enable the inherent suppression of dephasing errors due to unknown frequency shifts by increasing the mean photon number $\bar{n}$ of the encoded qubit states. However, a critical consequence is the reduced controllability of KCs in the protected regime. Therefore, in Section 5.3, we propose gate schemes that overcome this tradeoff between protection and controllability and enable universal computation with KCs, while being highly robust to dephasing errors that emerge from undesired frequency noise. We are driven by a fundamentally different motivation compared to past studies on KCs, which typically focused on engineering a strong noise bias through large mean photon numbers and realizing bias-preserving gates [205, 149, 148]. Here we specifically consider the regime of smaller mean photon numbers ($\bar{n} \lesssim 3$) in order to limit the negative impact of increasing bit-flip rates ($\propto \bar{n}$) in KCs. We further note that, compared to previous approaches to realizing frequency-noise-insensitive control of superconducting qubits (e.g., via spin-locking [211]), our approach allows for always-on frequency noise protection across a universal set of single- and two-qubit gates.

5.2 Kerr-cat qubits

We consider a Kerr-nonlinear oscillator subject to a two-photon squeezing drive with frequency ω_{2p} . In the frame rotating with $\omega_{2p}/2$, the system Hamiltonian reads

$$\mathcal{H} = \delta a^\dagger a - \frac{K}{2} a^{\dagger 2} a^2 + \frac{\varepsilon_2}{2} (a^2 + a^{\dagger 2}), \quad (5.3)$$

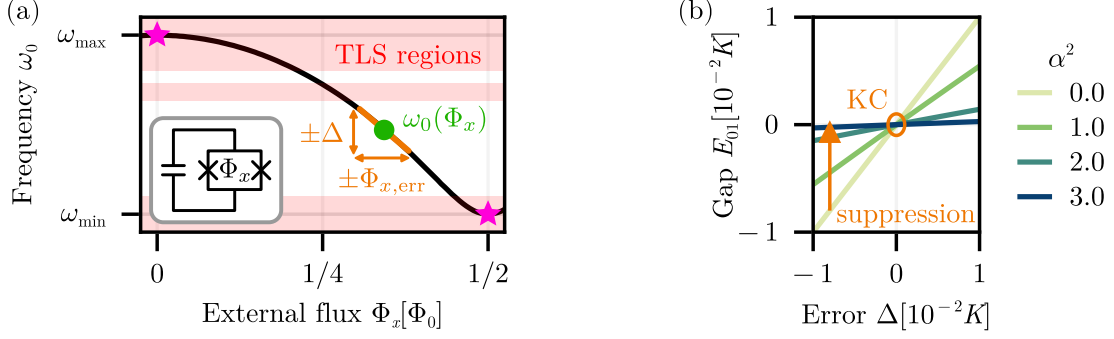


Figure 5.1: (a) In frequency-tunable transmons (inset), uncertainties in the flux $\Phi_{x,\text{err}}$ lead to unknown frequency shifts Δ (orange arrows around green point) when operated away from sweet spots (pink stars). (b) Shifts Δ break the degeneracy of the KC qubit (orange circle), causing phase errors. The gap opening is exponentially suppressed with the cat size α^2 , creating robustness to Δ .

where $\delta = \omega_0 - \omega_{2p}/2$ is the detuning between frequency of the bare qubit ω_0 and the rotating frame, ε_2 is the two-photon drive strength, and K is the Kerr nonlinearity. Without loss of generality we assume $K > 0$.

Equation (5.3) shows that the system has a symmetry in the photon number parity, therefore the eigenstates are superpositions of either only even or only odd photon number states. We denote by $|2k\rangle$ the $(k+1)$ th even parity state and by $|2k+1\rangle$ the $(k+1)$ th odd parity state, $k \geq 0$, sorted by energy. The computational subspace is spanned by the states $|0\rangle$ and $|1\rangle$ with energy gap $E_{01}(\delta, \alpha^2) = E_1 - E_0$. On resonance at $\delta = 0$, these states are degenerate ($E_{01} = 0$) and furthermore correspond to Schrödinger cat states $|0\rangle = |C_\alpha^+\rangle = (|\alpha\rangle + |-\alpha\rangle)/\mathcal{N}_+$ and $|1\rangle = |C_\alpha^-\rangle = (|\alpha\rangle - |-\alpha\rangle)/\mathcal{N}_-$ with normalization constants $\mathcal{N}_\pm = \sqrt{2(1 \pm \exp(-2\alpha^2))}$, thus they are superpositions of coherent states (see Equation (5.1)). This defines the standard Kerr-cat (KC) with cat size $\alpha^2 = \varepsilon_2/K \geq 0$ and mean photon number $\bar{n} \simeq \alpha^2$. The isometry that maps from the full KC Hilbert space to the logical qubit states is then given by

$$P_\alpha = |0\rangle_L \langle C_\alpha^+| + |1\rangle_L \langle C_\alpha^-|. \quad (5.4)$$

Figure 5.2(a) shows the Bloch sphere of the KC qubit, where the Wigner functions (phase space representations) of the X and Z eigenstates are shown at the respective axes.

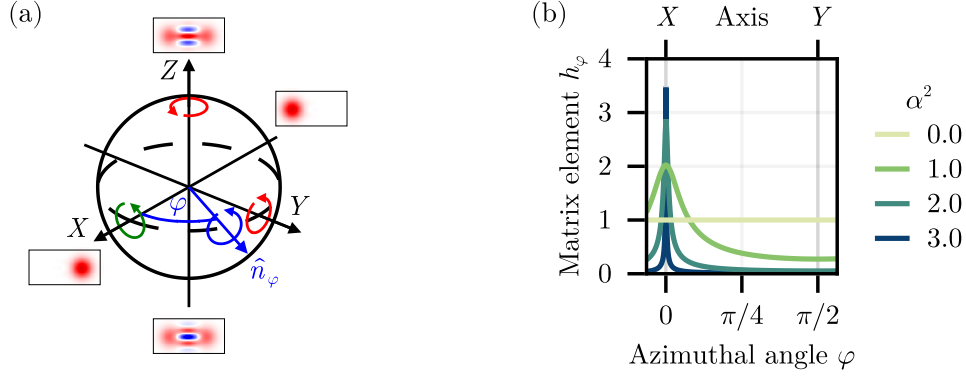


Figure 5.2: (a) Bloch sphere of the KC qubit showing Wigner functions of the X and Z eigenstates as well as rotations about the axes X (easy, green), Y and Z (hard, red), and an arbitrary axis in the X - Y plane $\hat{n}_\varphi$ (blue). (b) Dependence of the magnitude of the drive matrix element h_φ on the angle φ . As the cat size and protection increase, rotations about all axes in the X - Y plane except for the X axis get suppressed.

5.2.1 Protection versus controllability

Unknown shifts Δ in the oscillator frequency break the KC's degenerate qubit manifold in the basis $\{|C_\alpha^+\rangle, |C_\alpha^-\rangle\}$ and open the energy gap to first order like $E_{01}(\Delta, \alpha^2) \simeq \partial_\delta E_{01}(0, \alpha^2) \Delta$, leading to the accumulation of undesired phase errors. However, the appeal of the KC is that such phase errors are exponentially suppressed with the cat size via $\partial_\delta E_{01}(0, \alpha^2) \simeq 4\alpha^2 \exp(-2\alpha^2)$. Figure 5.1(b) visualizes the energy gap $E_{01}(\Delta, \alpha^2)$ as a function of the frequency uncertainty Δ , which shows the exponential suppression as the cat size α^2 increases. This illustrates the inherent protection of KCs to phase errors induced by unknown frequency shifts when idling. In the remainder of this article we address the question of whether this robustness can be retained while simultaneously realizing universal control of the KC.

The key challenge arises from the difficulty to implement arbitrary single-qubit gates. While the inherent protection mitigates unwanted perturbations, it also limits the controllability. For example, rotations about axes in the X - Y plane are realized by applying a

single-photon drive with frequency $\omega_{1p} = \omega_{2p}/2$, which adds the Hamiltonian term

$$\mathcal{H}_c(t) = \frac{\varepsilon(t)}{2} \left(e^{i\tilde{\varphi}} a + e^{-i\tilde{\varphi}} a^\dagger \right) \rightarrow P_\alpha \mathcal{H}_c(t) P_\alpha^\dagger = \frac{\varepsilon(t)}{2} (\cos(\tilde{\varphi}) h_x X - \sin(\tilde{\varphi}) h_y Y). \quad (5.5)$$

The projection onto the KC subspace gives rise to the drive coupling terms

$$h_x = \frac{2\alpha}{\sqrt{1 - e^{-4\alpha^2}}}, \quad h_y = \frac{2\alpha e^{-2\alpha^2}}{\sqrt{1 - e^{-4\alpha^2}}} = h_x e^{-2\alpha^2}, \quad (5.6)$$

which shows that rotations about the Y axis are exponentially suppressed in the cat size α^2 while rotations about the X axis are (approximately) linearly enhanced. More generally, the relevant matrix element for such KC rotations is given by

$$h_\varphi e^{i\varphi} \equiv \langle 1 | (\cos(\tilde{\varphi}) h_x X - \sin(\tilde{\varphi}) h_y Y) | 0 \rangle = h_x (\cos(\tilde{\varphi}) - i \sin(\tilde{\varphi}) e^{-2\alpha^2}), \quad (5.7)$$

where $\hat{n}_\varphi = (\cos(\varphi), \sin(\varphi), 0)^\top$ describes the rotation axis in the Bloch sphere and the magnitude h_φ determines the achievable Rabi rate $\Omega \propto h_\varphi$ for rotations about that axis. We find an exponential suppression of the Rabi rate when $\tilde{\varphi} \neq m\pi$, $m \in \mathbb{N}$.

We visualize this dependence of the matrix element h_φ on the azimuthal angle φ in Figure 5.2(b). In order to obtain an analytical expression, we first conclude the relationship between the azimuthal angle φ and the drive phase $\tilde{\varphi}$ from Equation (5.7):

$$\varphi = \arg(\cos(\tilde{\varphi}) - i \sin(\tilde{\varphi}) e^{-2\alpha^2}) \iff \sin(\varphi) \cos(\tilde{\varphi}) = -\sin(\tilde{\varphi}) \cos(\varphi) e^{-2\alpha^2}. \quad (5.8)$$

Using this we can determine the magnitude of the matrix element h_φ as a function of φ :

$$\begin{aligned}
h_\varphi &= h_x \sqrt{\cos^2(\tilde{\varphi}) + \sin^2(\tilde{\varphi})e^{-4\alpha^2}} \\
&= h_x \frac{e^{-2\alpha^2}}{\sqrt{1 - \cos^2(\varphi)(1 - e^{-4\alpha^2})}}
\end{aligned} \tag{5.9}$$

$$= h_x \begin{cases} 1 - 2\alpha^2 \sin^2(\varphi) & (\text{for } \alpha^2 \ll 1), \\ \begin{cases} 1 & (\varphi = m\pi), \\ \sqrt{1 - \cos^2(\varphi)}^{-1} e^{-2\alpha^2} + \mathcal{O}(e^{-4\alpha^2}) & (\text{otherwise}), \end{cases} & (\text{for } \alpha^2 \gg 1). \end{cases} \tag{5.10}$$

Equation (5.9) is derived using the square of the right part of Equation (5.8) and the Pythagorean trigonometric identity. Equation (5.10) shows the exponential suppression of rotations about all axes except for the $\pm X$ axis in the large cat size limit $\alpha^2 \gg 1$. We indicate this observation with differently colored axis rotations in Figure 5.2(a). The limit $\alpha^2 \rightarrow 0$ recovers the familiar property $h_\varphi \rightarrow h_x \rightarrow 1$ of the bare qubit, where rotations about all axes are equally possible and there is no dependence on φ .

5.3 Frequency-noise-insensitive universal control of Kerr-cat qubits

In this main section, we analyze the impact of this special nature of the KC on the performance of rotation gates about the principal axes X ($\varphi = 0$, $h_\varphi \rightarrow h_x$) and Y ($\varphi = \pi/2$, $h_\varphi \rightarrow h_y$) as well as Z . Furthermore, we briefly comment on the realization of two-qubit gates in order to complete the universal set of gates.

5.3.1 $X(\theta)$ rotations

Figure 5.2(b) underlines the well-known fact that for large cat sizes only X rotations are not suppressed [72, 76, 52]. In fact, the relevant matrix element grows like $h_x \simeq 2\alpha$ as we saw in Equation (5.6). This enables the straightforward realization of arbitrary $X(\theta)$ gates under the KC protection, so we refer to X as the “easy” rotation axis.

We validate the robustness of a $X(\pi/2)$ rotation (target gate U_{targ}) to frequency noise by simulating the dynamics (actual gate $U(T; \Delta)$) and computing the infidelity (recall Equation (2.51))

$$\mathcal{I}(T; \Delta, \gamma) = 1 - \frac{1}{4} \left| \text{Tr} \left[U_{\text{targ}}^\dagger P_\alpha U(T; \Delta, \gamma) P_\alpha^\dagger \right] \right|^2 \quad (5.11)$$

for gate times $T \in [10 K^{-1}, 50 K^{-1}]$ and frequency shifts $\Delta \in [-\Delta_{\text{max}}, \Delta_{\text{max}}]$ with $\Delta_{\text{max}} = 5 \cdot 10^{-3} K$. Here γ denotes parameters that determine the gate pulse. Were there no protection, our choice for T and Δ_{max} would correspond to coherent errors in the range $(\Delta_{\text{max}} T)^2 \sim 10^{-3} - 10^{-1}$. In the case of superconducting platforms with typical Kerr values $K/2\pi \sim 10 - 100$ MHz, we would have characteristic time units of $K^{-1} \sim 16 - 1.6$ ns and frequency shifts up to $\Delta_{\text{max}}/2\pi \sim 50 - 500$ kHz. For now we focus on static shifts Δ which allow for conclusions about errors with low-frequency signatures (“quasistatic”), and we discuss time-dependent errors $\Delta(t)$ in Section 5.3.5.

For the pulse shape ε_x we consider the truncated Gaussian

$$f(t) = \left[\frac{e^{-(t-T/2)^2/2\sigma^2} - e^{-(T/2)^2/2\sigma^2}}{1 - e^{-(T/2)^2/2\sigma^2}} \right]^m = \left[\frac{e^{-(t/T-1/2)^2/2} - e^{-1/8}}{1 - e^{-1/8}} \right]^2. \quad (5.12)$$

that was used in Ref. [205] and write $\varepsilon_x(t) = \varepsilon_{x,0} f(t)$, thus we have $\gamma = [\varepsilon_{x,0}]$. The last equation holds for our specific choice of parameters $\sigma = T$ and $m = 2$, which is the same as in Ref. [205].

Given a gate time T , we find the optimal pulse parameters (here the drive amplitude

$\varepsilon_{x,0}$) that minimizes the average infidelity:

$$\bar{\mathcal{I}}(T) = \min_{\gamma} \frac{1}{2\Delta_{\max}} \int_{-\Delta_{\max}}^{\Delta_{\max}} d\Delta \mathcal{I}(T; \Delta, \gamma). \quad (5.13)$$

The integral is evaluated numerically by computing the infidelity at eleven equally spaced points in the interval $[-\Delta_{\max}, \Delta_{\max}]$ and applying the Simpson rule for a better approximation. We perform the optimization exhaustively, which is feasible for the low-dimensional pulse parameterization we apply.

In Figure 5.3(a) we graph the pulse shape for $T = 40 K^{-1}$, which achieves infidelities $\mathcal{I} \lesssim 10^{-4}$ over the Δ range as shown in Figure 5.3(b). This constitutes one data point in Figure 5.3(c), where we show the average infidelity $\bar{\mathcal{I}}$ over the range of gate durations. We observe that $\bar{\mathcal{I}}$ decreases (the robustness increases) for greater α^2 , as expected, reaching values orders of magnitude below the naive coherent error limit $(\Delta_{\max} T)^2$ thanks to the KC's robustness. We note that infidelities increase with larger gate time as the phase errors are not perfectly eliminated and accumulate more over longer durations. The comparatively large infidelity in the short gate time regime for small cat sizes is because the required drive amplitudes become large while the gap to leakage states is rather small ($\simeq 2\alpha^2 K$ [72]), leading to off-resonant excitation. A more sophisticated pulse design or the introduction of a DRAG component (see discussion in Section 5.3.3, where this is more relevant) could mitigate this effect, however here we do not optimize the well-understood $X(\theta)$ protocol further. For larger cat sizes the leakage issue does not emerge because the gap to leakage states grows linearly with α^2 and the drive amplitude required for the gate decreases since $h_x \simeq 2\alpha$ increases.

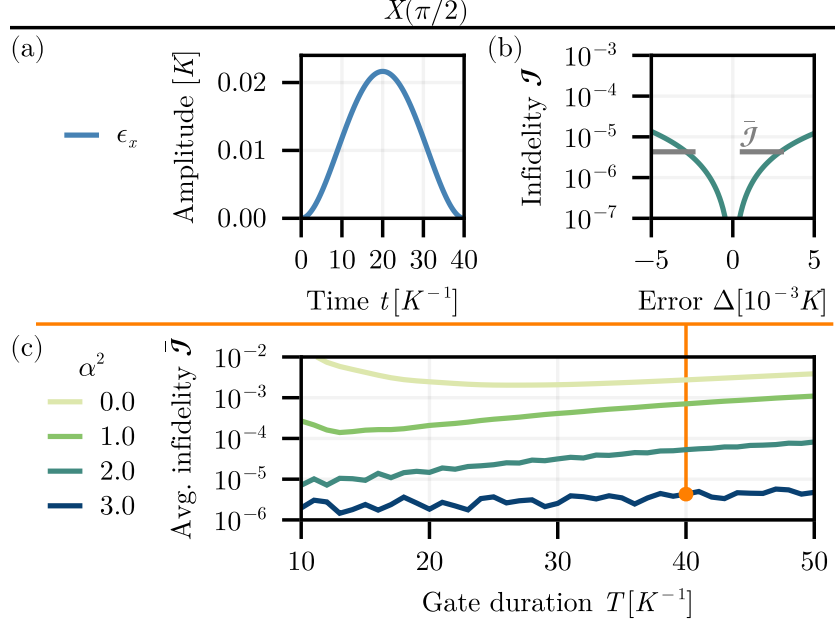


Figure 5.3: Pulse implementation and infidelity robustness for the $X(\pi/2)$ gate. (a) The $X(\theta)$ rotation is native to the KC and is realized through a simple single-photon drive pulse. (b) The infidelity of the pulse (a) when simulated over a range of detuning errors Δ . (c) The average infidelity $\bar{\mathcal{J}}$ of pulses of type (a) for different durations and cat sizes. The infidelity decreases when the cat size is increased

5.3.2 Two-qubit gate

Before we turn to the more challenging single-qubit gates (the main focus of this chapter), we briefly discuss the realization of a two-qubit gate that is analogous to the $X(\theta)$ gate. We consider two KC qubits $i = A, B$ with cat sizes α_i^2 coupled via a tunable beamsplitter interaction [5, 149]:

$$\begin{aligned} \mathcal{H}_{\text{int}}(t) &= g(t) \left(e^{i\tilde{\varphi}} a_A^\dagger a_B + e^{-i\tilde{\varphi}} a_B^\dagger a_A \right) \\ &= g(t) \left(\cos(\tilde{\varphi}) \left(a_A^\dagger a_B + a_B^\dagger a_A \right) + i \sin(\tilde{\varphi}) \left(a_A^\dagger a_B - a_B^\dagger a_A \right) \right). \end{aligned} \quad (5.14)$$

On a superconducting architecture this could be realized by driving a tunable transmon coupler [4] or SNAIL coupler [27] at the difference frequency $|\omega_0^A - \omega_0^B|$. In the KC basis

the effective Hamiltonian reads

$$P_{\alpha_A} P_{\alpha_B} \mathcal{H}_{\text{int}}(t) P_{\alpha_A}^\dagger P_{\alpha_B}^\dagger = \frac{g(t)}{2} \left(\cos(\tilde{\varphi}) \left(h_x^A h_x^B X_A X_B + h_y^A h_y^B Y_A Y_B \right) + i \sin(\tilde{\varphi}) \left(h_x^A h_y^B X_A Y_B - h_y^A h_x^B Y_A X_B \right) \right). \quad (5.15)$$

Depending on the drive phase $\tilde{\varphi}$, different rotation axes can be accessed, similar to the single-qubit case. However, we note that only the XX term does not contain a suppressed factor, as each other term contains at least one factor $h_y^i \propto e^{-2\alpha_i^2}$ (see Equation (5.6)), significantly reducing the achievable Rabi rate of rotations about those axes. Hence, we focus the on the $\tilde{\varphi} = 0$ case:

$$P_{\alpha_A} P_{\alpha_B} \mathcal{H}_{\text{int}}(t) P_{\alpha_A}^\dagger P_{\alpha_B}^\dagger = \frac{g(t)}{2} \left(h_x^A h_x^B X_A X_B + h_y^A h_y^B Y_A Y_B \right). \quad (5.16)$$

For the bare qubit with $\alpha^2 = 0$ this reduces to an iSWAP interaction. In the KC case we instead target an $XX(\theta)$ gate as the vanishing $h_y^i \propto e^{-2\alpha_i^2}$ elements suppress the YY term. However, since the YY term may be non-negligible in the regime of smaller α^2 , an echo sequence

$$XX(\theta) = X_i R_{\text{int}}(\theta/2) X_i R_{\text{int}}(\theta/2) \quad (5.17)$$

can be employed to fully cancel the YY term. Here either $i = A$ or $i = B$ is possible and $R_{\text{int}}(\eta)$ is the rotation generated by $\mathcal{H}_{\text{int}}$ with angle

$$\eta = \int_0^T dt g(t) h_x^A h_x^B \quad (5.18)$$

The special choice $XX(\pi/2)$ is equivalent to the Mølmer-Sørensen gate on trapped-ion platforms [172], which is locally equivalent to CNOT and CZ. Thus, similarly to the $X(\theta)$ case, an entangling gate can be directly realized in a robust way as the KCs have an exponentially suppressed sensitivity to detuning-induced dephasing while offering a large matrix element

product $h_x^A h_x^B \simeq 4\alpha_A \alpha_B$ to drive the interaction.

5.3.3 $Y(\theta)$ rotations

In order to realize a universal gate set, single-qubit rotations about at least one “nontrivial” axis are required. A natural first candidate are $Y(\theta)$ gates as they can also be realized with a single-photon drive, similarly to $X(\theta)$ gates. However, Y rotations are exponentially suppressed in the KC (see Figure 5.2(d)). To overcome this issue, we temporarily reduce the cat size $\alpha^2(t)$ by adiabatically ramping down the two-photon drive amplitude $\varepsilon_2(t) = \varepsilon_{2,0} + \tilde{\varepsilon}_2(t)$, where $\tilde{\varepsilon}_2(t) \leq 0$. This trades off an increased matrix element $h_y(t)$ for smaller inherent protection, which we expect to be partially compensated by a dynamical robustness effect emerging from the gate drive itself, similar to a spin-locking effect [211]. Both the Y gate drive ϵ_y and the ramping $\tilde{\varepsilon}_2$ are given by truncated Gaussians (5.12) with optimizable amplitudes $\gamma = [\varepsilon_{y,0}, \tilde{\varepsilon}_{2,0}]$.

It turns out that in order to produce a high-fidelity gate, we need to include a corrective component ϵ_x in the other quadrature, inspired by the Derivate Removal by Adiabatic Gate (DRAG) scheme [126, 54]. This is due to non-negligible coupling to leakage states when driving rotations about the Y axis, and we provide the derivation of the ε_x component as well as a detailed discussion at the end of this section.

We perform the robustness study introduced in Section 5.3.1 for $Y(\pi/2)$ and present the results in Figures 5.4. We note that, while the robustness generally grows with α^2 and improves over the spin-locking-only case at $\alpha^2 = 0$, the average infidelity $\bar{\mathcal{I}}$ does not decrease significantly as $\alpha^2 = 2 \rightarrow 3$. This is in stark contrast to the $X(\pi/2)$ case where the exponentially growing robustness is evident, which is because the KC protection has to be partially sacrificed in order to realize the $Y(\pi/2)$ gate when starting with a greater cat size, and the spin-locking effect of the gate drive provides only limited robustness. While this scheme can produce average infidelities well below 10^{-4} on short timescales for these

intermediate cat sizes, it requires possibly prohibitively large drive amplitudes $\varepsilon_{y,0}$ (see discussion below). In Section 5.3.4, we present an alternative gate scheme for a “nontrivial” axis and show that a greater level of robustness can be achieved by harnessing the structure of KC qubit’s energy spectrum, going beyond the limitation of a spin-locking mechanism while also avoiding the necessity for strong gate drives.

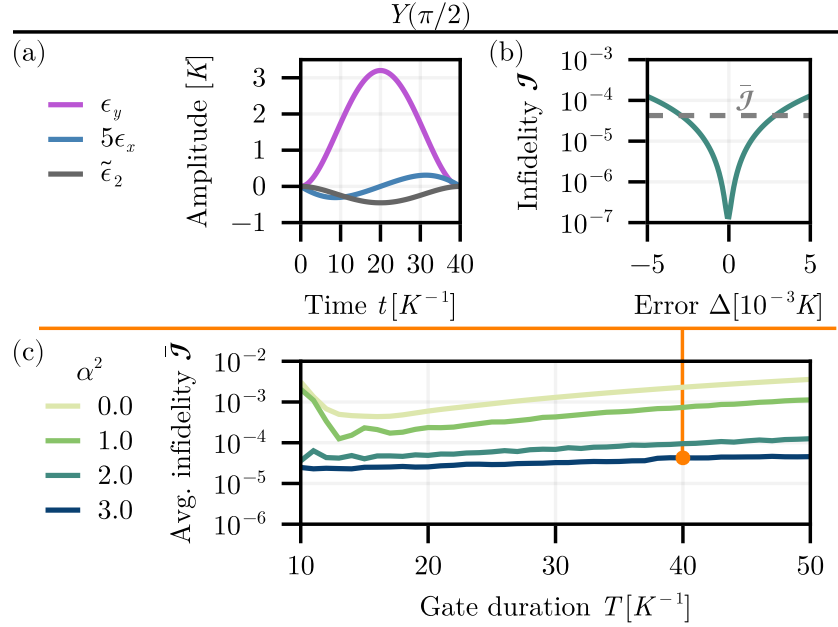


Figure 5.4: (a) Pulse scheme for the Y rotation, which includes tuning the two-photon the pump strength $\tilde{\epsilon}_2$ to lower the cat size and a DRAG term ϵ_x . (b) The infidelity of the pulse (a) when simulated over a range of detuning errors Δ . (c) The average infidelity $\bar{\mathcal{J}}$ of pulses of type (a) for different durations and cat sizes. The increase in robustness diminishes when increasing the cat size beyond $\alpha^2 \gtrsim 2$.

Importance of the DRAG correction

The exponentially small matrix element $h_y \simeq 2\alpha e^{-2\alpha^2}$ necessitates a large drive component ε_y , which leads to the undesired off-resonant coupling to states outside the computational subspace. In particular, we observe Stark shifts of the computational states $|0\rangle$ and $|1\rangle$, causing a tilted rotation axis $\hat{n}$ with a nonzero n_z component. We aim to cancel this effect by introducing a DRAG (Derivate Removal by Adiabatic Gate [126, 54]) term ε_x in the X

quadrature, which we derive in the following. The system Hamiltonian reads

$$\begin{aligned}
\mathcal{H}(t) &= -\frac{K}{2}a^{\dagger 2}a^2 + \frac{\varepsilon_{2,0} + \tilde{\varepsilon}_2(t)}{2}(a^2 + a^{\dagger 2}) + \frac{\varepsilon_y(t)}{2}(-i)(a - a^{\dagger}) + \frac{\varepsilon_x(t)}{2}(a + a^{\dagger}) \\
&= \mathcal{H}_0 + \frac{\tilde{\varepsilon}_2(t)}{2}\mathcal{H}_2 + \frac{\varepsilon_y(t)}{2}\mathcal{H}_y + \frac{\varepsilon_x(t)}{2}\mathcal{H}_x \\
&= \mathcal{H}_{02y}(t) + \frac{\varepsilon_x(t)}{2}\mathcal{H}_x.
\end{aligned} \tag{5.19}$$

Defining the instantaneous eigenstates $|\psi_j(t)\rangle$ with respect to the main drive Hamiltonian $\mathcal{H}_{02y}$ via $\mathcal{H}_{02y}(t)|\psi_j(t)\rangle = \tilde{E}_j(t)|\psi_j(t)\rangle$, we can conclude for the time evolution of an arbitrary state $|\psi(t)\rangle = \sum_j c_j(t)|\psi_j(t)\rangle$:

$$\dot{c}_j(t) = -i\tilde{E}_j(t)c_j(t) - \sum_k c_k \left(i\frac{\varepsilon_x(t)}{2} \langle \psi_j(t) | \mathcal{H}_x | \psi_k(t) \rangle + \langle \psi_j(t) | \dot{\psi}_k(t) \rangle \right). \tag{5.20}$$

All controls start at 0, therefore we have $|\psi_0(0)\rangle = |+\rangle = (|0\rangle + i|1\rangle)/\sqrt{2}$ and $|\psi_1(0)\rangle = |-i\rangle = (|0\rangle - i|1\rangle)/\sqrt{2}$ due to the presence of $\mathcal{H}_y$, where $|0/1\rangle = |C_{\alpha}^{\pm}\rangle$ and $\alpha^2 = \varepsilon_{2,0}/K$. In the ideal case the desired $Y(\theta)$ rotation in the computational basis is realized through the accrual of a phase difference between $|\psi_0(t)\rangle$ and $|\psi_1(t)\rangle$ without any transitions in the process. However, the sum in Equation (5.20) includes terms that describe such transitions between instantaneous eigenstates, and using the DRAG control ε_x we aim to specifically cancel the $|\psi_0(t)\rangle \leftrightarrow |\psi_1(t)\rangle$ transition. Thus we require

$$\varepsilon_x(t) = \frac{2i \langle \psi_1(t) | \dot{\psi}_0(t) \rangle}{\langle \psi_1(t) | \mathcal{H}_x | \psi_0(t) \rangle}. \tag{5.21}$$

In order to rewrite $\langle \psi_1(t) | \dot{\psi}_0(t) \rangle$, we consider

$$\begin{aligned}
\partial_t (\mathcal{H}_{02y}(t) |\psi_0(t)\rangle) &= \partial_t (\tilde{E}_0(t) |\psi_0(t)\rangle) \\
\dot{\mathcal{H}}_{02y}(t) |\psi_0(t)\rangle + \mathcal{H}_{02y}(t) |\dot{\psi}_0(t)\rangle &= \dot{\tilde{E}}_0(t) |\psi_0(t)\rangle + \tilde{E}_0(t) |\dot{\psi}_0(t)\rangle \\
\implies \langle \psi_1(t) | \dot{\psi}_0(t) \rangle &= \frac{\langle \psi_1(t) | \dot{\mathcal{H}}_{02y}(t) | \psi_0(t) \rangle}{\tilde{E}_0(t) - \tilde{E}_1(t)} \\
&= \frac{\tilde{\varepsilon}_2(t) \langle \psi_1(t) | \mathcal{H}_2 | \psi_0(t) \rangle + \dot{\varepsilon}_y(t) \langle \psi_1(t) | \mathcal{H}_y | \psi_0(t) \rangle}{2(\tilde{E}_0(t) - \tilde{E}_1(t))}.
\end{aligned} \tag{5.22}$$

Therefore the full DRAG correction reads

$$\varepsilon_x(t) = \frac{i}{\tilde{E}_0(t) - \tilde{E}_1(t)} \frac{\tilde{\varepsilon}_2(t) \langle \psi_1(t) | \mathcal{H}_2 | \psi_0(t) \rangle + \dot{\varepsilon}_y(t) \langle \psi_1(t) | \mathcal{H}_y | \psi_0(t) \rangle}{\langle \psi_1(t) | \mathcal{H}_x | \psi_0(t) \rangle}. \tag{5.23}$$

This is the exact solution that cancels the desired transition at every point in time t . It can be simplified by treating the control ε_y perturbatively and approximating the energies and matrix elements. Expanding the numerator and denominator to first order ultimately yields

$$\varepsilon_x(t) = \frac{1}{4h_y(t)h_x(t)} \left(\frac{h_y^{1 \rightarrow 2}(t)^2}{E_2(t)} - \frac{h_y^{0 \rightarrow 3}(t)^2}{E_3(t)} \right) \dot{\varepsilon}_y(t), \tag{5.24}$$

where $h_y^{0 \rightarrow 3}(t) = |\langle 3 | H_y | 0 \rangle(t)|$ and $h_y^{1 \rightarrow 2}(t) = |\langle 2 | H_y | 1 \rangle(t)|$ describe the coupling of the computational states to the first two excited states $|2\rangle$ (energy $E_2(t)$) and $|3\rangle$ (energy $E_3(t)$) with respect to the instantaneous KC of size $\alpha^2(t)$. Equation (5.24) recovers the familiar structure where the DRAG drive ε_x is proportional to the derivative of the primary drive ε_y , here for the case of two relevant transitions, which was also discussed in Ref. [54].

We compare the robustness study for the $Y(\pi/2)$ gate for different settings. To reiterate, the drive's Y component $\varepsilon_y(t) = \varepsilon_{y,0}f(t)$ as well as the two-photon ramp $\tilde{\varepsilon}_2(t) = \tilde{\varepsilon}_{2,0}f(t)$ are given by truncated Gaussians, where the amplitudes $\varepsilon_{y,0}$ and $\tilde{\varepsilon}_{2,0}$ are optimized over. We include the results in Figure 5.5, which shows the optimized average infidelities $\bar{\mathcal{I}}$ alongside

the optimized pulse parameters for the different settings, where instead of $\tilde{\varepsilon}_{2,0}$ we plot the relative cat size $\eta = \alpha^2(T/2)/\alpha^2 = 1 + \tilde{\varepsilon}_{2,0}/(\alpha^2 K)$ at the most ramped-down point ($t = T/2$).

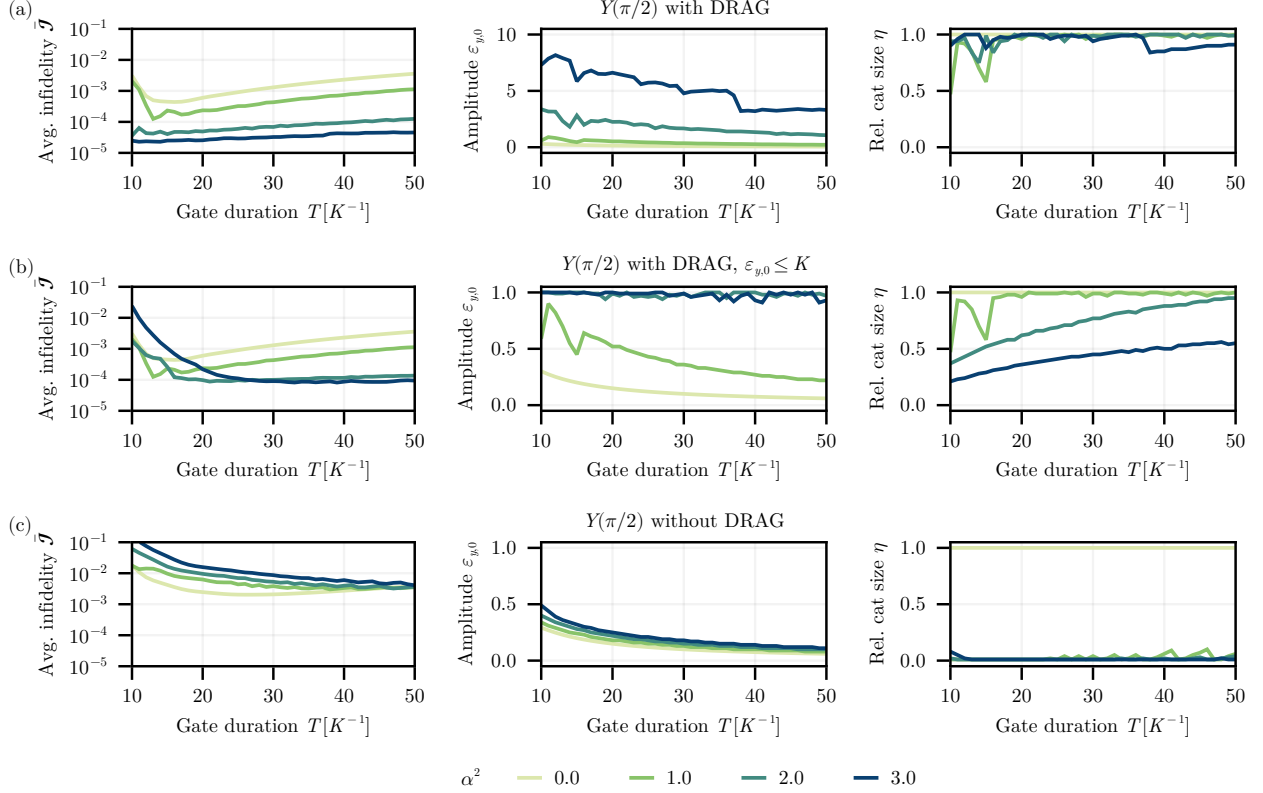


Figure 5.5: Optimization results of the $Y(\pi/2)$ gate for different settings, showing average infidelity $\bar{\mathcal{I}}$, drive amplitude $\varepsilon_{y,0}$ and relative cat size $\eta = \alpha^2(T/2)/\alpha^2$ at the most ramped-down point. (a) Gate scheme including the DRAG correction ε_x . This is the main result, which is also included in Figure 5.4. The cat size is barely ramped down to retain the KC protection, however this makes large drive amplitudes necessary, which can become a limiting factor. (b) When constraining the drive amplitude, the cat size has to be ramped increasingly more for larger initial α^2 in order to counteract the suppressed h_y element. Sacrificing the KC protection limits the achievable average infidelity. (c) Without the implementation of a DRAG correction, the coupling to excited states combined with the suppressed h_y element leads to a Stark shift and ultimately to a loss in fidelity, hence the found solution is to ramp the cat size down all the way to zero. No robustness improvement over the bare qubit ($\alpha^2 = 0$) is achieved.

Figure 5.5(a) reiterates the results from above for the $Y(\pi/2)$ gate with DRAG correction and further shows the pulse parameters. We note that under the assumption of an unconstrained drive strength, the optimal solution is to keep the cat size high in order to retain

as much KC protection as possible and achieve improving infidelities as the cat size grows. However, this necessitates the application of a rapidly growing drive strength to compensate for the exponentially suppressed matrix element h_y . Therefore, as long as the required power can be provided, this approach is feasible and can produce average infidelities well below $\bar{\mathcal{I}} < 10^{-4}$ for cat sizes $\alpha^2 \gtrsim 2$ with short gate times.

As the supply of drive power can be a physical limitation, we consider the case where the dominant Y component is upper-bounded, and for this study we choose $\varepsilon_{y,0} \leq K$ as an example. Figure 5.23(b) shows how the robustness and pulse parameters are affected by this constraint. While there are no changes for the small cat sizes $\alpha^2 \lesssim 1$ as these systems do not require that much drive power in the unconstrained case, we observe a clear loss of robustness for the larger cat sizes we consider, especially in the short gate time regime. Due to the upper limit on the drive amplitude, the cat size has to be ramped down significantly during the gate protocol to increase the matrix element h_y and speed up the gate angle accumulation. The associated reduction of KC protection leads to an increase of the average infidelity, and we essentially find no improvement when increasing the cat size beyond $\alpha^2 \gtrsim 2$.

In order to emphasize the importance of the DRAG correction, we include in Figure 5.5(c) the (unconstrained) gate optimization without a DRAG term ε_x . We observe that in this setting the best performing solution corresponds to ramping the KC down all the way to the bare qubit at $\alpha^2 = 0$ (equivalent to $\eta = 0$) to realize the rotation. This is because at nonzero cat size the KC suffers from a Stark shift that emerges from the strong drive (due to exponentially suppressed matrix element h_y) coupling to the excited states, which is not suppressed with the cat size. This coherent error hurts the gate fidelity more than the influence of the frequency noise Δ , hence the reduction to the bare qubit and no robustness improvement is achieved by making the KC larger. Therefore, the addition of the DRAG correction is absolutely crucial.

5.3.4 $Z(\theta)$ rotations

The Z axis is the other “nontrivial” principal axis as rotations are directly suppressed by the KC’s phase error protection. In many quantum computing platforms, rotations about the Z axis are implemented “virtually” by adjusting the phase $\tilde{\varphi}$ of the single-photon drive for subsequent gates [119]. However, this relies on the assumption that a change in $\tilde{\varphi}$ induces a significant change in the azimuthal angle φ of the rotation axis, which does not apply in the case of the KC as can be seen in Equation (5.8). This is why $Z(\theta)$ gates have to be implemented physically (i.e., with an actual gate pulse) beyond a certain cat size as the imbalance between rotation axes becomes too dominant, and therefore we consider $\alpha^2 \gtrsim 1$ in the following.

Prior works [147, 63, 84] suggested intentionally detuning the KC ($\delta(t) \neq 0$). However, due to the KC protection mechanism and the resulting suppressed energy gap $E_{01}(\delta, \alpha^2)$ (see Figure 5.1(b) and inset of Figure 5.6 below), this leads to gate times that are exponentially slow in α^2 , which is impractical for larger cat sizes. To overcome this controllability issue, we consider an approach where the cat size is adjusted during the $Z(\theta)$ gate in addition to the detuning. In particular, we propose an adiabatic $Z(\theta)$ scheme that can be implemented on a similar timescale as $X(\theta)$ gates while retaining robustness to first order, thereby circumventing the usual tradeoff between protection and controllability.

Robust $Z(\theta)$ gates enabled by the robust line

Our scheme relies on the key observation that for a detuned KC there exists a “robust line” $\delta_{\text{rob}}(\alpha^2) \in [0, K]$ where the energy gap $E_{01}(\delta_{\text{rob}}(\alpha^2), \alpha^2) > 0$ is positive but also first-order-insensitive to detuning errors, i.e. $\partial_{\delta} E_{01}(\delta_{\text{rob}}, \alpha^2) = 0$. This is visualized in Figure 5.6(a) with black dotted lines. The solid orange line highlights an example gate trajectory of our proposed $Z(\theta)$ scheme, which works as follows: Starting from the initial KC with size α^2 (1), we adiabatically change the detuning $\delta : 0 \rightarrow \delta_{\text{rob}}(\alpha_1^2)$ over a time τ . Next, over a time

$(T - \tau)/2$, we adiabatically trace the robust line by ramping down the cat size $\alpha^2 \rightarrow \alpha'^2 < \alpha^2$ while adjusting the detuning $\delta(t) = \delta_{\text{rob}}(\alpha^2(t))$ (midpoint ②, cat size α'^2). This increases the energy gap E_{01} to accrue a logical phase difference more quickly while the system remains first-order-insensitive to unwanted detuning errors Δ . Finally, we reverse the process to complete the gate (①). Figure 5.6(a) shows how the gate trajectory is confined to regions of parameter space where $\partial_\delta E_{01} \ll 1$, suggesting a gate with high robustness and average fidelity overall.

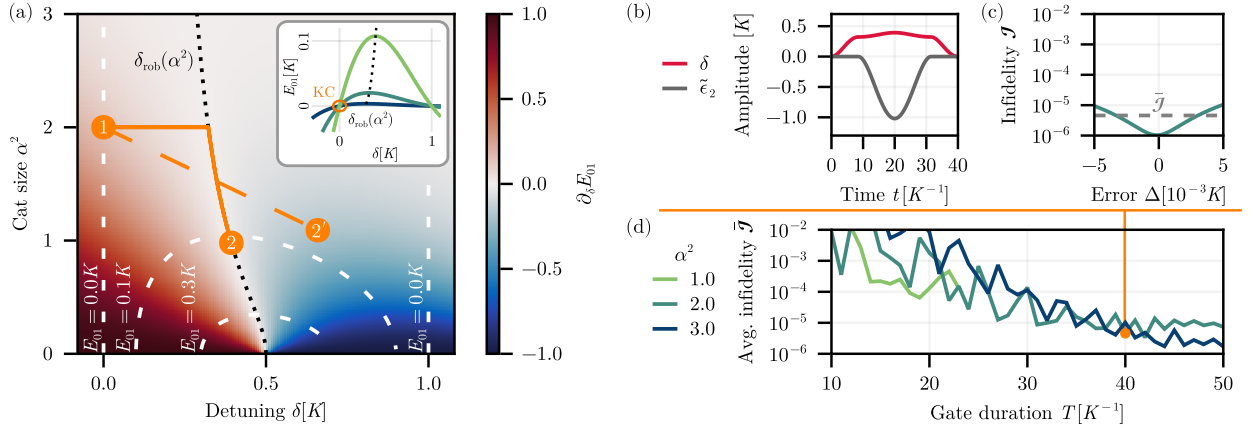


Figure 5.6: Adiabatic realization of the $Z(-\pi/2)$ gate using the scheme ①↔②. (a) The energy gap derivative $\partial_\delta E_{01}(\delta, \alpha^2)$ as a function of the detuning δ and cat size α^2 . The robust line $\delta_{\text{rob}}(\alpha^2)$ is highlighted by a black dashed line. An example gate trajectory is shown with an orange solid line. The inset (same legend as (d)) shows the robust line in the energy spectrum. (b) Control pulse for the solid trajectory shown in (a), which simultaneously tunes the detuning δ and the two-photon pump strength via $\tilde{\epsilon}_2$. (c) The infidelity of the pulse (b) when simulated over a range of detuning errors Δ . (d) The average infidelity $\bar{\mathcal{I}}$ of pulses of type (b) for different durations and cat sizes. We note that the $\alpha^2 = 1$ line ends at $T \approx 22 \text{ K}^{-1}$ because our scheme ①↔② is not applicable in the case of small cat sizes ($\alpha^2 \lesssim 1$), which is also why $\alpha^2 = 0$ is excluded here. We discuss an alternative scheme ①↔②' below, which works for all cat sizes but is less suited for dynamical errors $\Delta(t)$.

We perform the robustness study for a $Z(-\pi/2)$ gate following this scheme. The negative rotation angle is due to the nonnegative energy gap $E_{01} \geq 0$ (see Equation (5.28) below). The detuning ramps $\delta : 0 \leftrightarrow \delta_{\text{rob}}$ as well as the pump modulation $\tilde{\epsilon}_2$ are based on truncated Gaussians (see Equation (5.12)), and the optimizable parameters are the ramp time τ and the

amplitude $\tilde{\varepsilon}_{2,0}$. An example pulse is included in Figure 5.6(b). We observe in Figures 5.6(c)-(d) that this gate scheme can achieve average infidelities $\bar{\mathcal{I}}$ in the $10^{-4} - 10^{-6}$ range, which means similar performance to native $X(\theta)$ rotations at gate times that are within the same order of magnitude (tens of K^{-1}). This underlines that utilizing the robust line $\delta_{\text{rob}}(\alpha^2)$ is crucial for realizing a protected gate.

We provide analytical reasoning for the emergence of robustness using a simplified model of the adiabatic evolution. We start with the Hamiltonian

$$\mathcal{H}(t) = \delta(t)a^\dagger a - \frac{K}{2}a^{\dagger 2}a^2 + \frac{\varepsilon_2(t)}{2}(a^2 + a^{\dagger 2}), \quad (5.25)$$

where we now assume control over the detuning $\delta(t)$ and, like for $Y(\theta)$ gates discussed in Section 5.3.3, the cat size $\alpha^2(t) = \varepsilon_2(t)/K$. Considering the evolution of a quantum state $|\psi(t)\rangle = \sum_j c_j(t) |\psi_j(t)\rangle$ expressed in the instantaneous eigenbasis defined by $\mathcal{H}(t) |\psi_j(t)\rangle = E_j(t) |\psi_j(t)\rangle$, we obtain the equation of motion (similarly to Equation (5.20))

$$\dot{c}_j(t) = -iE_j(t)c_j(t) - \sum_k c_k(t) \langle \psi_j(t) | \dot{\psi}_k(t) \rangle. \quad (5.26)$$

We define the gauge such that $\arg(c_j(0)) = 0$. In our idealistic and simplified analysis of an adiabatic evolution, where we assume no transitions and geometric phases, we can disregard the right term and only consider the phase accumulation of each eigenstate $|\psi_j(t)\rangle$ due to its energy $E_j(t)$:

$$c_j(t) = \exp\left(-i \int_0^t dt' E_j(t')\right) c_j(0). \quad (5.27)$$

The gate schemes we consider start and end with the non-detuned KC ($\delta = 0$, cat size α^2 , point ①) and then trace the parameter space $(\delta(t), \alpha^2(t))$ approximately adiabatically in a time T . The number operator $a^\dagger a$ breaks the KC's degeneracy in the computational basis $\{|C_\alpha^+\rangle, |C_\alpha^-\rangle\}$ basis, therefore we have $|\psi_0(0)\rangle = |\psi_0(T)\rangle = |C_\alpha^+\rangle$ and $|\psi_1(0)\rangle = |\psi_1(T)\rangle =$

$|C_\alpha^- \rangle$. This is why the relative phase accrued between the instantaneous eigenstates $|\psi_0(t)\rangle$ and $|\psi_1(t)\rangle$ leads to a $Z(\theta)$ rotation in the computational basis, with rotation angle

$$\theta = \arg \left(\frac{c_1(T)}{c_0(T)} \right) = - \int_0^T dt E_{01}(t). \quad (5.28)$$

Here $E_{01}(t) = E_0(t) - E_1(t) = E_{01}(\delta(t), \alpha^2(t))$ denotes the instantaneous energy gap along the trajectory $(\delta(t), \alpha^2(t))$. In the detuning range $\delta \in [0, K]$ we consider in this work, the emerging energy gap is nonnegative, i.e. $E_{01}(t) \geq 0$, thus the rotation occurs clockwise and the gate angle θ grows in the negative direction.

Detuning errors Δ lead to shifts in the energy gap $E_{01}(t; \Delta)$ along the trajectory and therefore cause erroneous rotation angle $\theta(\Delta)$. To first order in Δ , we can write

$$\theta(\Delta) = - \int_0^T dt E_{01}(\delta(t) + \Delta, \alpha(t)) \simeq - \int_0^T dt \left(E_{01}(\delta(t), \alpha(t)) + \Delta \frac{\partial E_{01}}{\partial \delta}(\delta(t), \alpha(t)) \right) \quad (5.29)$$

$$= \theta(\Delta = 0) + \Delta \partial_\delta \theta(\Delta = 0) \quad (5.30)$$

$$\implies \partial_\delta \theta(\Delta = 0) = \int_0^T dt \frac{\partial E_{01}}{\partial \delta}(\delta(t), \alpha(t)). \quad (5.31)$$

The robustness of our primary scheme $\textcircled{1} \leftrightarrow \textcircled{2}$ originates from the fact that by design the energy gap derivative $\partial_\delta E_{01}$ is either vanishingly small or exactly zero along the entire trajectory. During the initial detuning ramp $\delta : 0 \rightarrow \delta_{\text{rob}}(\alpha^2)$ and final detuning ramp $\delta : \delta_{\text{rob}}(\alpha^2) \rightarrow 0$ the gap derivative is exponentially small in α^2 due to the KC protection, and while traversing the robust line $(\delta_{\text{rob}}(\alpha^2(t)), \alpha^2(t))$ we have $\partial_\delta E_{01}(t) = 0$ by construction. This enables average gate infidelities $\bar{\mathcal{I}} < 10^{-5}$ as can be seen in Figure 5.6(d).

Since this gate scheme relies on adiabatic evolution, nonadiabatic errors are a relevant source of infidelity, in particular in the low duration regime. This reason for the spikes in that regime in Figure 5.6(d). More carefully designed pulse shapes could help mitigate this

effect. Especially the two-photon pulse $\tilde{\varepsilon}_2$ should be looked at in this context, as the relevant spectral gaps, which critically influence the validity of the adiabatic approximation, strongly depend on the cat size.

We note that Equations (5.30) and (5.31) are only valid for static detuning errors as we take Δ out of the integral. However, the limitation to static errors is not strictly necessary for the scheme $\textcircled{1} \leftrightarrow \textcircled{2}$, because the error term under the integral in Equation (5.29) is small or zero for all t . This is why high robustness can be expected for certain dynamical errors $\Delta(t)$ as well, and we expand on the extension to time-dependent errors in Section 5.3.5.

Alternative gate scheme for static shifts Δ

If we strictly restrict ourselves to static detuning errors Δ , then a different gate scheme with a simpler implementation becomes possible. It allows to relax the constraint of $\partial_\delta E_{01}(t)$ (approximately) vanishing at all times t , and instead we can only require $\partial_\delta \theta(\Delta = 0) = 0$ over the entire gate evolution in order to have first-order robustness against static detuning errors. We conclude from Equation (5.31) that this corresponds to a vanishing *time-averaged* energy gap.

The robust line $\delta_{\text{rob}}(\alpha^2)$ traces the maxima of the energy gap and therefore separates parameter space into two parts, which is visualized in Figure 5.7(a). On the left side we have $\partial_\delta E_{01}(\delta, \alpha^2) > 0$, while on the right side we have $\partial_\delta E_{01}(\delta, \alpha^2) < 0$. Therefore, there are many gate trajectories that lead to first-order robustness because they start with a non-detuned KC on the left side and cross over to the right side before returning to the initial point on the left side, which, when pulse parameters are chosen properly, results in a vanishing time-averaged energy gap.

To illustrate this scheme, we consider the subset of such gate trajectories that traverse the parameter space in a straight line from an initial KC ($\delta = 0$, cat size α^2) at point $\textcircled{1}$ to the midpoint $\textcircled{2}$ and back to $\textcircled{1}$. The midpoint $\textcircled{2}$ with parameters $(\delta_{\text{max}}, \alpha'^2)$ may lie

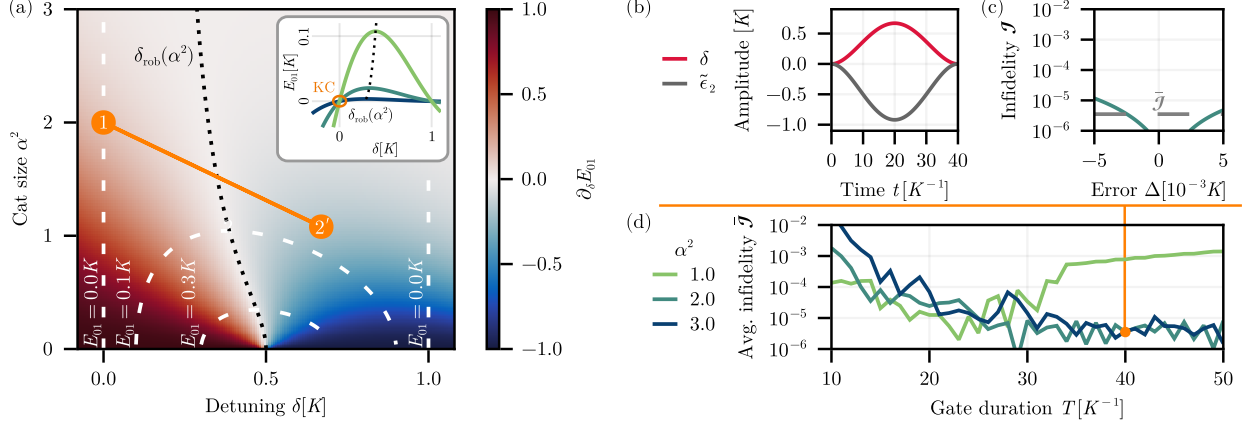


Figure 5.7: Alternative realization of the $Z(-\pi/2)$ gate using the $1 \leftrightarrow 2'$ scheme. (a) The energy gap derivative $\partial_\delta E_{01}(\delta, \alpha^2)$ as a function of the detuning δ and cat size α^2 . The robust line $\delta_{\text{rob}}(\alpha^2)$ (highlighted by a black dashed line) separates the parameter space into two regions with different signs of the energy gap derivative $\partial_\delta E_{01}(\delta, \alpha^2)$, which enables first-order insensitivity to static errors Δ on averaging over the trajectory. An example trajectory is drawn with an orange solid line. The inset shows the energy spectrum. (b) Control pulse for the solid trajectory shown in (a), which simultaneously tunes the detuning δ and the two-photon pump strength via $\tilde{\varepsilon}_2$. (c) The infidelity of the pulse (b) when simulated over a range of detuning errors Δ . (d) The average infidelity $\bar{\mathcal{I}}$ of pulses of type (b) for different durations and cat sizes.

anywhere in parameter space for this scheme, however, as discussed above, robustness is only possible when it is located on the right side of the robust line ($\delta_{\text{max}} > \delta_{\text{rob}}(\alpha'^2)$). An example trajectory is highlighted in Figure 5.7(a) with an orange line. We specifically choose the truncated Gaussian pulse shape $f(t)$ to trace the trajectory, so that we have:

$$\begin{pmatrix} \delta(t) \\ \alpha^2(t) \end{pmatrix} = f(t) \begin{pmatrix} \delta_{\text{max}} \\ \alpha'^2 \end{pmatrix} + (1 - f(t)) \begin{pmatrix} 0 \\ \alpha_1^2 \end{pmatrix} \quad (5.32)$$

$$\Rightarrow \begin{pmatrix} \delta(t) \\ \tilde{\varepsilon}_2(t) \end{pmatrix} = f(t) \begin{pmatrix} \delta_{\text{max}} \\ \tilde{\varepsilon}_{2,0} \end{pmatrix} \quad (5.33)$$

Here we introduced the two-photon pump difference $\tilde{\varepsilon}_2(t)/K \leq 0$ again, which describes the change in cat size from the starting point via $\alpha^2(t)K = \varepsilon_2(t) = \varepsilon_{2,0} + \tilde{\varepsilon}_2(t)$. The amplitudes δ_{max} and $\tilde{\varepsilon}_{2,0}$ are the optimizable parameters.

We perform the $Z(-\pi/2)$ robustness study for this $1 \leftrightarrow 2'$ scheme like we did for the

①↔② above and present the results in Figures 5.7(b)-(d). We find highly robust pulses that achieve infidelities well below 10^{-5} for all cat sizes studied. Similarly to the primary scheme, we note higher infidelities in the low duration regime with stronger spikes due to nonadiabatic errors. The infidelity for $\alpha^2 = 1$ grows significantly for $T \gtrsim 30 K^{-1}$ because feasible trajectories that implement the right rotation angle $\theta = \pi/2$ do not extend far enough into the blue (right) region (as most of the phase is accumulated in the red (left) region) to average out the energy gap $\partial_\delta E_{01}$.

We observe small infidelities over a greater range of gate times T compared to the primary scheme, which is due to the greater freedom in pulse trajectories and simpler pulse shapes. Regarding the latter, we emphasize that this scheme ①↔②' is more practical and easier to implement compared to ①↔② because it only uses analytical pulse shapes (truncated Gaussians) that can be easily calibrated in experiment (amplitude tuning). The scheme ①↔② is more challenging to realize in this regard as the computation of the robust line $\delta_{\text{rob}}(\alpha^2)$ requires numerical diagonalization of the Hamiltonian in theory and a more involved calibration in practice. However, it is important to keep in mind that this simpler nature of the ①↔②' scheme comes at the cost of being effective to static errors Δ only.

Brief comparison to the Kerr gate

Due to the retained protection, our schemes offer a better error scaling than the fundamentally different Kerr gate scheme, which implements a discrete $Z(\pi/2)$ rotation and has been demonstrated in experiment [72, 52]. That approach involves rapidly turning off the two-photon pump ($\varepsilon_2 = 0$) and letting the system evolve under the Kerr Hamiltonian for a time $T = \pi/K$. During the gate evolution the qubit is fully unprotected and $\mathcal{I} \simeq \alpha^2 \Delta^2 T^2$ for $\alpha^2 \gtrsim 1$, thus the error grows with cat size. Our schemes enable $\mathcal{I} = \mathcal{O}(\Delta^4 T^2 K^{-2})$ as they cancel the lowest-order error term. This comes at the cost of gate times $T = \pi/(2\bar{E}_{01})$ (for $Z(-\pi/2)$) that are about an order of magnitude longer than for the Kerr gate ($\bar{E}_{01} \sim 0.1 K$

is the energy gap averaged over the gate trajectory, see Figure 5.6(a)).

5.3.5 Extension to time-dependent errors

While so far we have mostly focused on static frequency shifts Δ , here we give qualitative and quantitative reasoning why we can expect our KC control framework to be robust to (well-behaved) time-dependent shifts $\Delta(t)$ as well. In the case of idling, $X(\theta)$ and $XX(\theta)$ rotations, where we have the full KC protection related to the cat size α^2 , this is easy to see because of the suppressed energy gap between the computational states:

$$E_{01}(\Delta(t), \alpha^2) \simeq \partial_\delta E_{01}(0, \alpha^2) \Delta(t). \quad (5.34)$$

This holds for slow-varying perturbations $\Delta(t)$ where we can consider the instantaneous energy spectrum of the system. More generally, we can look at the action of the perturbation operator (the photon number operator in this case) in the computational subspace

$$P_\alpha a^\dagger a P_\alpha^\dagger \simeq \alpha^2 I - 2\alpha^2 e^{-2\alpha^2} Z \quad (5.35)$$

and note the exponential suppression of phase errors ($\sim Z$).

In the following we focus the discussion on our adiabatic $Z(\theta)$ schemes, where we assume slowly-varying perturbations $\Delta(t)$, so we can apply the adiabatic modeling we already introduced in Section 5.3.4. Employing the same derivation that lead to Equation (5.29), we can approximate the error of the gate angle due to the shift $\Delta(t)$ as

$$\tilde{\theta}[\Delta] = \theta[\Delta] - \theta_{\text{id}} \simeq - \int_0^T dt \Delta(t) \partial_\delta E_{01}(t), \quad (5.36)$$

where $\partial_\delta E_{01}(t) = \partial_\delta E_{01}(\delta(t), \alpha^2(t))$ denotes the energy gap derivative along the trajectory. We use the notation $g[\Delta]$ to indicate that g is a functional that maps the function $\Delta(t)$ to

a scalar. The gate infidelity is given by

$$\begin{aligned}
\mathcal{I}[\Delta] &= 1 - \frac{1}{4} \left| \text{Tr} \left[Z \left(\tilde{\theta}[\Delta] \right) \right] \right|^2 \\
&= \sin^2 \left(\frac{\tilde{\theta}[\Delta]}{2} \right) \\
&\simeq \frac{\tilde{\theta}[\Delta]^2}{4} \\
&= \frac{1}{4} \int_0^T dt_1 \int_0^T dt_2 \Delta(t_1) \Delta(t_2) \underbrace{\partial_\delta E_{01}(t_1) \partial_\delta E_{01}(t_2)}_{\equiv F(t_1, t_2)}.
\end{aligned} \tag{5.37}$$

Ultimately we are interested in the average fidelity over different realizations of the frequency noise $\Delta(t)$,

$$\bar{\mathcal{I}} \simeq \frac{1}{4} \int_0^T dt_1 \int_0^T dt_2 \langle \Delta(t_1) \Delta(t_2) \rangle F(t_1, t_2), \tag{5.38}$$

where we have introduced the autocorrelation function of $\Delta(t)$

$$\langle \Delta(t_1) \Delta(t_2) \rangle = \int \mathcal{D}[\Delta] \mathcal{P}[\Delta] \Delta(t_1) \Delta(t_2) \tag{5.39}$$

with the probability distribution $\mathcal{P}[\Delta]$ and integration measure $\mathcal{D}[\Delta]$. We make the additional assumption that the frequency noise has no memory effect, i.e., the autocorrelation is a function of the time difference only: $\langle \Delta(t_1) \Delta(t_2) \rangle = \langle \Delta(t_1 - t_2) \Delta(0) \rangle$. This lets us rewrite

Equation (5.38):

$$\begin{aligned}\bar{\mathcal{I}} &\simeq \frac{1}{4} \int_0^T dt_1 \int_0^T dt_2 \langle \Delta(t_1 - t_2) \Delta(0) \rangle F(t_1, t_2) \\ &= \frac{1}{4} \int_0^T dt_1 \int_0^T dt_2 \int dt \hat{\delta}(t - (t_1 - t_2)) \langle \Delta(t) \Delta(0) \rangle F(t_1, t_2)\end{aligned}\quad (5.40)$$

$$= \frac{1}{4} \int_0^T dt_1 \int_0^T dt_2 \int dt \int \frac{d\omega}{2\pi} e^{-i\omega(t-(t_1-t_2))} \langle \Delta(t) \Delta(0) \rangle F(t_1, t_2) \quad (5.41)$$

$$\begin{aligned}&= \int \frac{d\omega}{2\pi} \left(\int dt e^{-i\omega t} \langle \Delta(t) \Delta(0) \rangle \right) \left(\frac{1}{4} \int_0^T dt_1 \int_0^T dt_2 e^{i\omega(t_1-t_2)} F(t_1, t_2) \right) \\ &= \int \frac{d\omega}{2\pi} S_\Delta(\omega) W(\omega).\end{aligned}\quad (5.42)$$

In Equation (5.40) we inserted a Dirac- $\hat{\delta}$ distribution, which was then expanded to its exponential form in Equation (5.41). In its final form the average infidelity is expressed as a convolution of the power spectral density $S_\Delta(\omega)$ of the frequency noise and the weighting function $W(\omega)$, which is solely determined by the energy gap derivative along the gate trajectory:

$$S_\Delta(\omega) = \int dt e^{-i\omega t} \langle \Delta(t) \Delta(0) \rangle, \quad (5.43)$$

$$W(\omega) = \frac{1}{4} \int_0^T dt_1 \int_0^T dt_2 e^{i\omega(t_1-t_2)} \partial_\delta E_{01}(t_1) \partial_\delta E_{01}(t_2). \quad (5.44)$$

This result supports our robustness analysis for the $Z(\theta)$ gate, which was discussed in Section 5.3.4. Our primary scheme $\textcircled{1} \leftrightarrow \textcircled{2}$ is designed such that we have $\partial_\delta E_{01}(t) \simeq 0$ along the trajectory, therefore we have $W(\omega) \simeq 0$ for all ω and, due to Equation (5.42), we can expect very small infidelities independent of the structure ($\sim S_\Delta(\omega)$) of the frequency noise.

For the other scheme $\textcircled{1} \leftrightarrow \textcircled{2'}$, however, the weighting function $W(\omega)$ does not vanish for most values of ω , which makes this scheme less robust to time-varying errors $\Delta(t)$. In the special case of static frequency shifts Δ the power spectral density simplifies to $S_\Delta(\omega) = 2\pi\hat{\delta}(\omega)\Delta^2$ and the average infidelity becomes $\bar{\mathcal{I}} \simeq \Delta^2 W(0)$. The value $W(0)$

is proportional to the square of the trajectory-averaged energy gap derivative, which can be zero for certain trajectories as discussed in Section 5.3.4, leading to the first-order robustness to static frequency shifts that we observed in the study (see Figure 5.7).

All in all we note that good robustness properties for a universal set of gates (in particular, $X(\theta)$, $XX(\theta)$, $Z(\theta)$) can be expected in the presence of time-dependent frequency shifts $\Delta(t)$ as well. However, it is important to keep in mind that the above analysis relied on an idealized adiabatic model and the assumption of slow-varying noise, which is helpful to gain intuition for the origin of robustness in the KC, but it might not hold for highly oscillating or less well-behaved noise.

5.3.6 Conclusion

We introduced a robust approach to universally control Kerr-cat qubits (KCs) with resiliency to frequency fluctuations, which circumvents the KC’s usual tradeoff of inherent protection at the cost of reduced controllability. We discuss gate schemes for rotations about the principal axes X , Y and Z in the Bloch sphere as well as the two-qubit XX rotation and analyze their robustness against frequency shifts Δ . The “easy” gates $X(\theta)$ and $XX(\theta)$ are native to the KC as they can be directly implemented for arbitrary cat size and therefore are trivially robust due to the exponential KC protection. While $Y(\theta)$ rotations can be realized with low infidelity (to a certain limit) but require large drive power, our carefully designed $Z(\theta)$ scheme demonstrates high robustness without facing any physical barriers. It leverages tracing an error-insensitive region $\delta_{\text{rob}}(\alpha^2)$ in parameter space (δ, α^2) , which makes the gate robust throughout the entire state evolution.

This ability to perform arbitrary rotations about the axes $\{X, Z, XX\}$ suffices for universal computation, motivating the concept of a robust KC-based platform. For instance, in the context of tunable superconducting hardware, this could enable the design of a TLS-resistant processor where individual physical devices are operated away from their sweet spots. Alter-

natively, one could imagine a hybrid platform where devices are used as transmons at their sweet spots ($\alpha^2 = 0$) by default and are only tuned away and converted to KCs ($\alpha^2 > 0$) when an interfering TLS is observed. Note that $XX(\theta)$ rotations are natively possible for any pair of cat sizes (see Equation (5.16)), which allows to directly interact a bare transmon with a KC. This hybrid approach would also reduce the impact of increased bit-flip rates in KCs. Lastly, we remark that since our work provides strategies for circumventing the usual trade-off between protection and controllability in protected qubits, it opens up interesting future research directions for enhancing the controllability of various types of protected qubits such as fluxonium qubits and $0 - \pi$ qubits without significantly compromising the protection.

CHAPTER 6

APPLICATION-TAILORED ERROR MITIGATION

In the past three chapters we targeted the lower layers of the quantum stack and discussed how research on the pulse level leads to opportunities for optimization in different platforms. Now we turn to the top layer of the stack, the algorithm layer, where the problems are formulated that we are interested in solving with quantum hardware.

In the current noisy intermediate-scale quantum (NISQ) era, quantum computers are prone to a variety of errors as outlined in Section 2.4. They propagate up the stack and impact the performance of quantum algorithms. While in the long term scientists would like to implement quantum error correction (QEC) schemes that can detect and correct for these distortions in real time, the adaption of quantum algorithms to hardware noise profiles can enable practical improvements in the near future and furthermore reduce the gap to the fault-tolerant era. There are certain classes of algorithms where even current noisy quantum computers are postulated to outperform the most powerful classical computers, and one such class is variational quantum algorithms (VQAs). In this chapter, we model the noise that affects the VQA performance and propose a pre-processing error mitigation scheme that exploits the noise’s structure and is efficient to implement.

6.1 Variational quantum algorithms

A variational quantum algorithm (VQA) consists of a parameterized circuit $A(\boldsymbol{\theta})$, called an *ansatz*, and a Hermitian operator of interest $\mathcal{H}$, called the problem’s *Hamiltonian*. The *ansatz* is usually constructed from parameterized single-qubit rotations interspersed with entangling gates and it prepares a quantum state $|\psi(\boldsymbol{\theta})\rangle$ depending on the parameters $\boldsymbol{\theta}$. A scalar loss (objective function) is given by the expectation value $\mathcal{L}(\boldsymbol{\theta}) = \langle \mathcal{H} \rangle_{\boldsymbol{\theta}}$ of the problem Hamiltonian for this quantum state (that is, the state’s average *energy* with respect to $\mathcal{H}$),

which can be obtained from measurement on quantum hardware after preparing the state. The goal is to obtain an optimal set of parameters that minimizes the loss. This is achieved by employing a classical minimization algorithm which feeds ansatz parameters to the quantum device and processes the resulting loss. Gradient-based methods like SPSA [173] and gradient-free methods like Nelder-Mead [131] have been used for this task in the past [179]. Figure 6.1 summarizes the process.

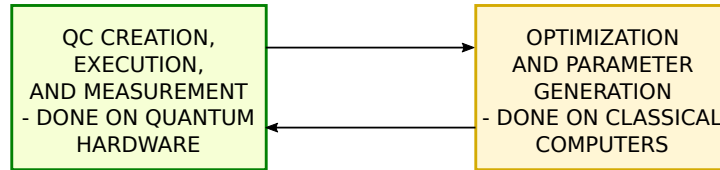


Figure 6.1: VQAs combine a classical optimizer which adjusts the parameters in each iteration with a quantum device producing function evaluations by running the corresponding quantum circuits.

VQAs exhibit *optimal parameter resilience (OPR)* [192]: If the noise is below a certain threshold, the parameters that minimize the loss on noisy hardware also often minimize the loss on a noiseless device. This makes VQAs robust to errors and leading candidates for demonstrating quantum advantage on NISQ devices. However, often the noise on NISQ devices is too high to observe OPR. Thus, we need to suppress errors using error mitigation techniques tailored for VQAs in order to realize OPR.

6.1.1 Variational Quantum Eigensolver

A popular VQA, which involves the estimation of the minimum eigenvalue of a given Hamiltonian $\mathcal{H}$, is called the *Variational Quantum Eigensolver (VQE)* [179]. VQE is extensively used in physics and chemistry applications for finding the minimum energy of Ising models or computing dissociation energies of molecules as functions of their bond lengths. This is why VQE has been studied extensively from an architectural perspective [153, 73, 191, 154] to ensure its reliable execution.

6.1.2 CAFQA

For VQAs, the optimal circuit parameters $\hat{\theta}$ and the target objective is unknown initially. However, prior work has proposed effective classical methods to find good initial circuit parameters θ_0 for VQA problems [77, 153]. Of particular interest to this work is CAFQA [153], which uses classical noiseless simulation to explore the Clifford space of a target VQA problem to find a good initial state. CAFQA exploits the key idea that circuits made up of only Clifford operations can be exactly simulated in polynomial time [64], which is due to their special properties mentioned in Section 2.2.5.

Clifford operations do not provide a universal set of quantum gates. Hence, the stabilizer states produced by Clifford-only circuits are limited in how effectively they can explore the quantum space. However, CAFQA showed that the Clifford-based initial state has high overlap with the problem’s ground state, and often achieved 90-99% accuracy finding the ground state energy, even before running VQA on the actual quantum device. While CAFQA only targeted VQA initialization, our work Clapton pursues the idea that these known good initial states provide excellent opportunity for state-dependent error mitigation.

CAFQA’s main shortcoming is that it does not take noise characteristics of the underlying hardware into account, which makes it susceptible to device errors that can significantly degrade the quality of the found solution when executed on a quantum machine, therefore not guaranteeing high-fidelity results in experiment.

6.2 Clapton: Clifford-assisted problem transformation for VQA error mitigation

VQAs and other NISQ algorithms can sustain noise strengths only up to certain limits, and techniques that can reduce the effects of noise-induced errors are of paramount importance. As we briefly discussed in Section 2.4.2, several error mitigation schemes have been proposed

in the past and they have considerably improved the result quality of quantum algorithms. However, there is still a long way to go — despite the implementation of error mitigation, the accuracy of VQAs on current NISQ machines often falls notably short of the stringent precision requirements essential in fields like molecular chemistry, particularly as the scale of problems grows. Thus, novel error mitigation techniques tailored to VQAs can substantially enhance their real-world utility.

We begin with the important observation that the noise affecting quantum devices has structure, stemming from physical properties of the qubits and their interaction with the environment. The noise itself may vary temporally (over time) and spatially (across qubits), but a significant portion of the noise can be broken down into key components that can be modeled accurately and scalably with classical compute. If such noise characteristics are well understood, then it is intuitive that quantum algorithms can be transformed so that they gain robustness to a target’s device noise profile.

A simple example of a noise bias is the higher fidelity of the ground state $|0\rangle$ compared to the excited state $|1\rangle$ on most qubits due to state decay (recall Section 2.4). Thus, one could expect that quantum circuits that mostly operate on and navigate through quantum states with many excitations (basis states with large Hamming weight) are likely to show reduced fidelity compared to circuits which operate close to the $|0\rangle$ states throughout. Taking this a step further, if one has prior knowledge that a circuit is expected to produce outputs with large Hamming weight, and if they have the ability to efficiently transform this circuit to produce low Hamming weight outputs instead, then the execution fidelity of the transformed circuit is expected to be higher than the former. Applying the inverse transformation to the output should then yield a good solution to the original problem.

While state-dependent biases in quantum device noise and bias-aware circuit transformations are interesting, their utility is not obvious. First, useful quantum circuits produce output states unknown to the user, which is why the circuit is executed on the device in the

first place. The classical computational complexity of identifying the correct output quantum state a priori would be exponential. Second, naive classical simulation of the effects of device noise on a quantum circuit, to gauge its impact on a given quantum state, is exponentially hard as well. Thus, naively thinking, it would be implausible to use information about the correct output state and state-dependent noise biases to avoid the impact of detrimental execution on quantum devices.

However, there is unique opportunity in the domain of VQAs. A VQA allows identifying noise-free “good” output states for the target problem (i.e., states having a high overlap with the unknown correct output state of the circuit) which can be computed in polynomial time using classical methods. Furthermore, a VQA can be efficiently transformed to an equivalent VQA problem with a different quantum state search space but an *unchanged* solution (the minimum energy), enabling a more noise-resilient search that is tailored to the quantum device characteristics. This is because there exists the freedom of a unitary basis change which only changes the representation of a problem (by defining new unit basis vectors in state space) but not its solution. As a fundamental operation in linear algebra this is universally applicable to any VQA problem.

Therefore, if we have a classically simulable model of some components of the device noise, then we are able to classically estimate the noisy fidelity of a known good state of a VQA problem and furthermore transform the VQA problem such that this state gets mapped to a good hardware state, like the all-qubit ground state $|0\rangle$, in a more noise-robust environment. This is the essence behind our proposed error mitigation scheme *Clapton: Clifford-assisted problem transformation for VQA error mitigation*. The Clapton hypothesis is that as long as the known good states of the VQA problem are fairly close to the ground state and the device noise modeling is reasonably accurate, our proposed problem transformations decrease the impact of device noise on the ground state of the VQA problem, thereby increasing the accuracy of the VQA solution. We design an end-to-end framework that implements

this hypothesis and achieve tremendous improvements in VQA quality in noisy quantum environments. Importantly, all the steps in the Clapton framework can be performed in polynomial time and are, therefore, classically efficient. The high-level idea is illustrated in Figure 6.2.

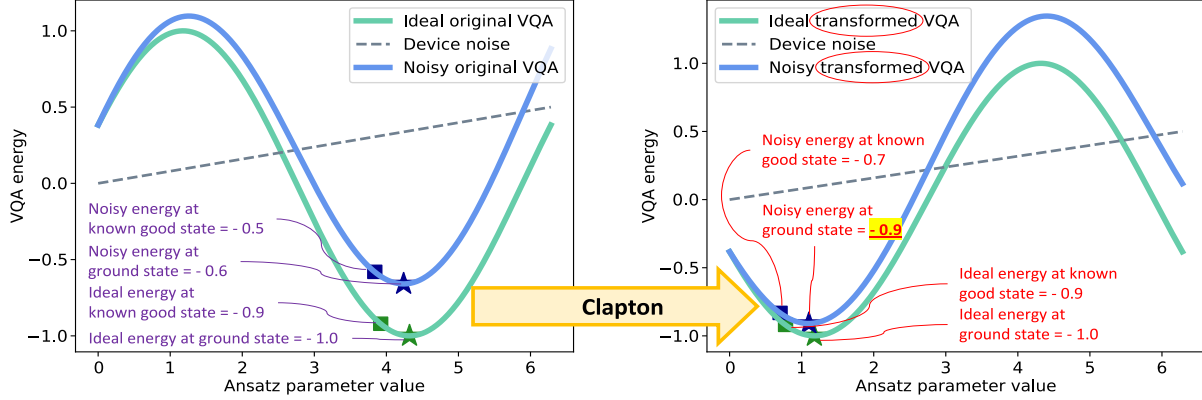


Figure 6.2: The fundamental idea behind Clapton: Transforming a VQA problem structure to match the noise structure on a quantum device. Having identified a good algorithmic state, transforming the VQA problem such that it is mapped to a good hardware state improves the fidelity under noisy evaluation, leading to improved performance of the entire VQA instance.

Going forward, we specifically focus on the special case of VQE, which was introduced in Section 6.1.1, to benchmark our error mitigation scheme Clapton. However, it is important to note that Clapton can be applied to other VQAs as well. Figure 6.3 summarizes the key result we observe with Clapton, and throughout the next sections we go over the design choices that enabled it.

6.2.1 Clapton proposal

Clapton addresses the issues of CAFQA described in Section 6.1.2 by incorporating hardware information from the lowest layer of the stack and modeling relevant sources of noise in Clifford-based efficient simulations. Given a VQE problem $\mathcal{H}$, it searches for a transformation $\hat{\mathcal{H}} = \hat{C}^\dagger \mathcal{H} \hat{C}$ to an equivalent problem $\hat{\mathcal{H}}$ with improved resilience to realistic hardware errors. $\mathcal{H}$ and $\hat{\mathcal{H}}$ are equivalent formulations of the same optimization problem because they share

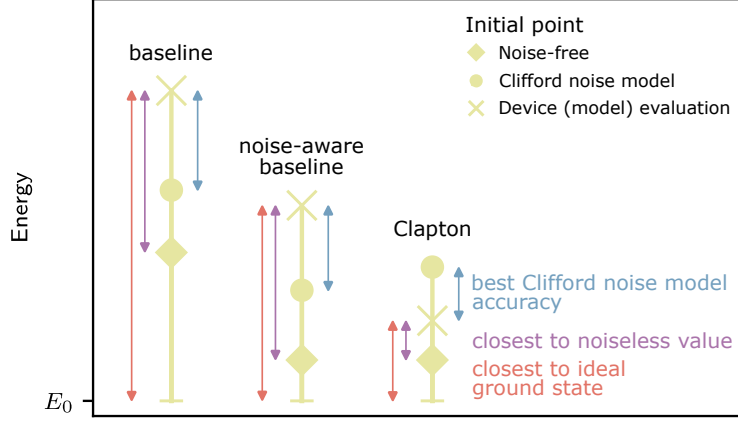


Figure 6.3: Key result: Clifford operations enable the combination of classically simulable noise and VQA Hamiltonian transformation, which ultimately gives Clapton its advantage over other initialization schemes in terms of solution quality (lower energy) and modeling accuracy.

the same eigenvalues (in particular, the ground state energy which is given by the smallest eigenvalue) due to the unitary transformation via the Clifford operation $\hat{C}$.

The big picture of the Clapton transformation

Let us consider a VQE problem $\mathcal{H}$ on N qubits with a variational ansatz $A(\theta)$ over d rotation parameters $\theta \in [0, 2\pi)^d$. The VQE is solved by finding the minimum expectation value over all quantum states that can be produced by the ansatz:

$$E^* = \min_{\theta} \langle \psi(\theta) | \mathcal{H} | \psi(\theta) \rangle = \min_{\theta} \langle 0 | A^\dagger(\theta) \mathcal{H} A(\theta) | 0 \rangle. \quad (6.1)$$

Here $E^* \geq E_0$ denotes the best approximation to the true ground state energy E_0 achievable in the variational setting. CAFQA limits the search over discrete sets of parameters θ such that $A(\theta)$ becomes a Clifford operation and is efficiently simulable on classical hardware.

Clapton changes perspective by searching over Hamiltonian transformations $\mathcal{H}(\gamma) = C^\dagger(\gamma) \mathcal{H} C(\gamma)$ where the qubit ground state $|0\rangle^{\otimes N} \equiv |0\rangle$ corresponds to the best initial state for the VQE problem. The full post-Clapton VQE algorithm on the quantum device is then

initialized with ansatz parameters $\boldsymbol{\theta} = \mathbf{0}$. Clapton finds this transformation by primarily focusing on the following optimization problem:

$$\begin{aligned}\hat{\boldsymbol{\gamma}} &= \arg \min_{\boldsymbol{\gamma}} \langle 0 | \tilde{A}^\dagger(\mathbf{0}) \mathcal{H}(\boldsymbol{\gamma}) \tilde{A}(\mathbf{0}) | 0 \rangle \\ &= \arg \min_{\boldsymbol{\gamma}} \langle 0 | \tilde{A}^\dagger(\mathbf{0}) C^\dagger(\boldsymbol{\gamma}) \mathcal{H} C(\boldsymbol{\gamma}) \tilde{A}(\mathbf{0}) | 0 \rangle.\end{aligned}\tag{6.2}$$

$C(\boldsymbol{\gamma})$ corresponds to a parametrized Clifford operation with parameters $\boldsymbol{\gamma} \in \Gamma$ from a space Γ . The set of Clifford operations on N qubits scales exponentially with N [165], therefore it is useful to choose an ansatz for C as well to limit the search space. $\tilde{A}$ denotes the noisy version of the ansatz A . We model this noise with stochastic Pauli operations (see Section 2.4). The goal is to resemble the hardware noise as accurately as possible to account for its effect during the online VQE optimization that follows Clapton. However, it is crucial to note that this design choice makes $\tilde{A}(\mathbf{0})$ a composition of (in part stochastic) Clifford operations, meaning it can be efficiently evaluated with classical compute. Clapton's output is the transformation parameter vector $\hat{\boldsymbol{\gamma}}$ which yields the transformation operator $\hat{C} = C(\hat{\boldsymbol{\gamma}})$ and new problem formulation $\hat{\mathcal{H}} = \mathcal{H}(\hat{\boldsymbol{\gamma}})$.

Transformation via Pauli conjugation

The VQE Hamiltonian is an observable, and we showed in Equation (2.35) that therefore it can be expressed as a weighted sum of M Pauli terms $\mathcal{H} = \sum_{i=1}^M c_i P_i$, where c_i is the energy coefficient corresponding to the Pauli string P_i . Clapton's transformation with a Clifford operation C maps every Pauli string P_i to its *anticonjugated* Pauli string $C^\dagger P_i C = \pm P'_i$. The sign is absorbed into the coefficient $c_i \rightarrow c'_i = \pm c_i$, therefore the transformed Hamiltonian reads

$$C^\dagger \mathcal{H} C = \sum_{i=1}^M c'_i P'_i.\tag{6.3}$$

This applies one of the appealing properties of Clifford gates that we introduced in Section 2.2.5: Since they map Pauli strings to (signed) Pauli strings, the structure of the VQE problem is preserved and the transformed Hamiltonian can be implemented in the VQE framework just like the original one.

Anticonjugation is the inverse operation compared to the formal definition introduced in Equation (2.37). The choice of conjugation method is arbitrary because $C^\dagger$ is Clifford if and only if C is, so the same properties hold. We prefer the anticonjugation because it is more intuitive in the grand scheme of VQE: When running the post-Clapton VQE algorithm, a quantum state $|\hat{\psi}(\boldsymbol{\theta})\rangle = A(\boldsymbol{\theta})|0\rangle$ with respect to the Clapton Hamiltonian $\hat{\mathcal{H}}$ can be directly translated back to a corresponding quantum state $|\psi(\boldsymbol{\theta})\rangle = \hat{C}|\hat{\psi}(\boldsymbol{\theta})\rangle = \hat{C}A(\boldsymbol{\theta})|0\rangle$ with respect to the original Hamiltonian $\mathcal{H}$. The states are equivalent as far as VQE is concerned because they share the same energy, $\langle\psi(\boldsymbol{\theta})|\mathcal{H}|\psi(\boldsymbol{\theta})\rangle = \langle\psi(\boldsymbol{\theta})|\hat{C}^\dagger\hat{\mathcal{H}}\hat{C}|\psi(\boldsymbol{\theta})\rangle = \langle\hat{\psi}(\boldsymbol{\theta})|\hat{\mathcal{H}}|\hat{\psi}(\boldsymbol{\theta})\rangle$. Therefore, $\hat{C}$ can be used to construct a good solution with respect to the original Hamiltonian, which is particularly easy in experiment if $\hat{C}$ is given by a circuit using single- and two-qubit Clifford gates like we are suggesting.

6.2.2 Clapton design

Within the scope of this paper we focus on specific choices for the VQE ansatz $A(\boldsymbol{\theta})$ and the Clapton transformation ansatz $C(\boldsymbol{\gamma})$. For the VQE ansatz we consider the circular hardware-efficient ansatz. It is composed of parameterized rotation gates $Y(\theta_j)$ and $Z(\theta_j)$ introduced in Equation (2.29) as well as controlled-NOT gates CX. The number of parameters is $d = 4N$. If $\boldsymbol{\theta} \in \{0, \pi/2, \pi, 3\pi/2\}^d$, then all rotation gates are Clifford, turning the full ansatz circuit into an efficiently simulable Clifford operation. CAFQA searches over this discrete space directly to find the best Clifford initialization for $A(\boldsymbol{\theta})$ in noiseless simulation. For Clapton we are especially interested in the circuit $A(\mathbf{0})$, where all the rotation gates become identity gates I and only the CX skeleton remains.

The transformation ansatz $C(\gamma)$ is inspired by the VQE ansatz, however it treats the two-qubit gates differently. While it still includes the parameterized rotation gates, it replaces the CX operations with parameterized two-qubit gates, which, depending on the added parameters γ_j , can each be one of the following (acting on qubits k and l):

$$\begin{aligned}\gamma_j = 0 &: I^k I^l, & \gamma_j = 1 &: \text{CX}_{k \rightarrow l}, \\ \gamma_j = 2 &: \text{CX}_{l \rightarrow k}, & \gamma_j = 3 &: \text{SWAP}_{k \leftrightarrow l}.\end{aligned}\tag{6.4}$$

We do this to increase the expressiveness of the ansatz and include the possibility to move around the components P_i^k of the Hamiltonian Pauli strings P_i by the means of SWAP operations ($\gamma_j = 3$). Furthermore, given the properties of CX that we discussed after Equation (2.39), having a CX ($\gamma_j = 1, 2$) or not ($\gamma_j = 0$) as well as the choice of direction ($k \rightarrow l$ versus $l \rightarrow k$) has a significant impact on the transformed Hamiltonian and its characteristics in a noisy environment. Our Clapton transformation ansatz for four qubits is visualized as part of the right flowchart shown in Figure 6.4.

Clapton's parameter space Γ for this ansatz has dimension $5N$ and each parameter γ_j can take four values, either one of four rotation angles for a single-qubit gate or one of the four choices for a two-qubit gate.

The implementation of the Clapton optimization

Just like CAFQA, Clapton focuses on finding a good initial point for a given VQE problem. Unlike CAFQA, it achieves this by transforming the Hamiltonian of the problem and restricting the initial parameters for the post-Clapton VQE to be $\boldsymbol{\theta} = \mathbf{0}$. Thus, one part of Clapton's cost function is given by the noisy estimate of the energy of the transformed Hamiltonian $\mathcal{H}(\gamma)$ at the initial VQE point:

$$\mathcal{L}_{\mathcal{N}}(\gamma) = \langle 0 | \tilde{A}^\dagger(\mathbf{0}) \mathcal{H}(\gamma) \tilde{A}(\mathbf{0}) | 0 \rangle.\tag{6.5}$$

However, only searching for a low noisy energy is not sufficient. There can be solutions that, when evaluated without noise, are relatively far from the true ground state energy E_0 but they are also rather error-resilient, therefore their noisy evaluations appear deceptively good. Thus, a transformation γ will be useful if it shows low noisy *and* noiseless energy. This is why Clapton furthermore adds the noiseless energy as another term (note $A(\mathbf{0})|0\rangle = |0\rangle$)

$$\mathcal{L}_0(\gamma) = \langle 0|A^\dagger(\mathbf{0})\mathcal{H}(\gamma)A(\mathbf{0})|0\rangle = \langle 0|\mathcal{H}(\gamma)|0\rangle, \quad (6.6)$$

so that the full cost function reads $\mathcal{L}(\gamma) = \mathcal{L}_{\mathcal{N}}(\gamma) + \mathcal{L}_0(\gamma)$.

Clapton evaluates both terms using the open-source framework Stim [58] that has been developed for fast Clifford simulations. It provides functionality to compute the anticonjugation of Pauli strings under Clifford operations as well as tools to simulate noisy Clifford circuits with depolarizing errors. Clapton then solves the discrete optimization problem

$$\hat{\gamma} = \arg \min_{\gamma} \mathcal{L}(\gamma) \quad (6.7)$$

by searching the parameter space Γ using genetic algorithms (GAs) implemented with PyGAD [53].

The complete optimization engine is visualized in Figure 6.4. Feeding in the VQE problem $(\mathcal{H})$ and setup $(\tilde{A}(\mathbf{0}), C(\gamma))$, Clapton spawns $\mathbf{s}$ GA instances that individually search Γ starting from random populations S_i . They each compute $\mathbf{m}$ iterations using the loss function $\mathcal{L}$, yielding final populations S_i^* . Now Clapton combines good features from different GA instances. Therefore it chooses the best $\mathbf{k}$ solutions from each S_i^* , mixes them and randomly draws $\mathbf{s}$ new initial populations, while adding new random guesses to fill back up to the original population size $|S|$. This is repeated until the global loss does not decrease between rounds, where Clapton typically allows for two retry rounds in this case before terminating. For the benchmarks we describe in the following sections, we ran Clapton with the hyper-

parameters $\mathbf{s} = 10$, $\mathbf{m} = 100$, $\mathbf{k} = 20$ and $|S| = 100$, which we found to work consistently well.

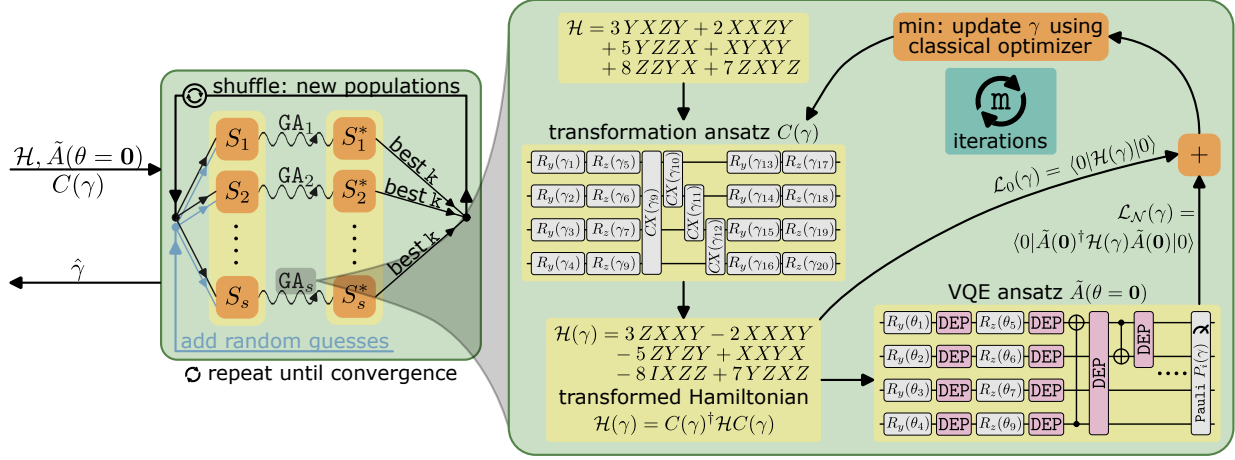


Figure 6.4: The Clapton transformation engine. Given a problem Hamiltonian $\mathcal{H}$, a noise-equipped version of the VQE ansatz (at the zero point) $\tilde{A}(\mathbf{0})$ and a Clifford transformation ansatz $C(\gamma)$, Clapton employs genetic algorithms (GAs) to search over transformation parameters γ to find a solution $\hat{\gamma}$. It instantiates $\mathbf{s}$ processes $\text{GA}_1, \dots, \text{GA}_s$, starting from random populations $S_1, \dots, S_s$, that search the parameter space independently. Internally, each GA instance performs $\mathbf{m}$ iterations of transformation optimization, where for different γ the Hamiltonian is transformed and the noisy ($\mathcal{L}_{\mathcal{N}}$) as well as noiseless ($\mathcal{L}_0$) energy of the initial point is evaluated. The GAs produce final populations S_i^* , of which Clapton takes the top $\mathbf{k}$ solutions, mixes them and randomly constructs new starting populations S_i , additionally adding new random guesses that are not derived from the previous round. This entire scheme is repeated until the global loss value does not decrease between rounds (convergence).

Clapton is designed to account for realistic noise

Here we describe how the Clapton design addresses the three relevant sources of error introduced in Section 2.4.

Thermal relaxation — As this error channel causes excitations $|1\rangle$ to decay, the qubit ground state $|0\rangle$ is the favored state in its presence. This is why Clapton chooses the initial point $\theta = \mathbf{0}$ for the post-Clapton VQE: In the idealized case of no gate errors, the VQE circuit $A(\mathbf{0})$ keeps all qubits in their ground state because the rotation gates reduce to

identity operations and the CX gates do not have an effect, $A(\mathbf{0})|0\rangle = |0\rangle$. Furthermore, recalling the discussion from the end of Section 6.2.1, the ground state $|0\rangle$ translates to the equivalent state $|\hat{\psi}\rangle = \hat{C}|0\rangle$ with respect to the original Hamiltonian $\mathcal{H}$. The work on CAFQA has shown that Clifford states like $|\hat{\psi}\rangle$ can approximate the true solution of $\mathcal{H}$ very well. This implies that $|0\rangle$ represents a good approximation to the solution of $\hat{\mathcal{H}}$ as well. Thus, it is reasonable to expect that the true solution of $\hat{\mathcal{H}}$ shows a large overlap with the ground state and the measurement results will be dominated by the ground state during the VQE execution. Therefore, by prioritizing the $|0\rangle$ state this design leads to increased resilience to thermal relaxation when performing VQE on the transformed problem $\hat{\mathcal{H}}$.

Gate errors — However, in reality quantum instructions are imperfect and gate errors must be considered. This requires replacing the ansatz circuit $A(\mathbf{0})$ by its noisy version $\tilde{A}(\mathbf{0})$. As we discussed in Section 2.4, a common way to model gate errors is with depolarizing noise, which is implemented in simulation via the stochastic application of Pauli strings. Each time the circuit is simulated (a shot is executed) the random gates get sampled which ultimately leads to the noisy output distribution of the circuit. As $\tilde{A}(\mathbf{0})$ is only constructed from (in part stochastic) Clifford operations, it is efficiently simulable in Stim. By incorporating gate errors, Clapton becomes aware of which qubits might accumulate a significant amount of noise and adjust the Hamiltonian accordingly to minimize the effect on the evaluation of Pauli terms with large energy coefficients. In particular, CX gates propagate and spread errors due to their special nature (see Equation (2.39)), which especially affects qubits at the end of long CX chains.

Measurement errors — Measurement is typically performed in the computational basis and yields classical bit values 0 or 1 (corresponding to the Z eigenvalues $+1$ and -1 , respectively) for each measured qubit as output. Readout errors correspond to the event of misidentifying the underlying state, meaning the state $|0\rangle$ yields output 1 or vice versa. While this misassignment is not perfectly symmetric in 0 and 1 in real hardware,

approximating this error as a random bit flip usually serves as a good model. In order to measure the transformed Pauli strings $P_i(\gamma)$ when evaluating the energy cost $\mathcal{L}_C(\gamma)$, we first append single-qubit gates to the noisy circuit $\tilde{A}(\mathbf{0})$ (including gate errors) to prepare the measurement basis. We then add random X gates with probabilities p_k on the qubits prior to the ideal readout, where p_k describes the misassignment error on qubit k . Readout errors can be significantly larger than gate errors, making this addition essential to build a more accurate model.

6.2.3 Methodology

In this section we introduce the benchmarks we use to analyze Clapton’s performance and furthermore describe the metric of interest.

Selection of VQE problems as benchmarks

In order to study the impact of Clapton we focus on a selection of physics and chemistry Hamiltonians as VQE problems.

Physics models — In statistical mechanics, two popular types of theoretical descriptions of magnetic systems are given by Ising models [34] and Heisenberg models [17]. They are known to be able to show phase transitions in certain dimensions as well as complex non-classical ground states. Here we focus on the simple case of 1D models with constant couplings. The (transverse field) *Ising model* is given by

$$\mathcal{H} = J \sum_{i=1}^{N-1} X_i X_{i+1} + \sum_{i=1}^N Z_i, \quad (6.8)$$

which describes how N spins (qubits) coupled in X -direction with strength J interact with an external field of unit strength pointing in Z -direction.

We further look at a special case of the (field-free) Heisenberg model, which couples the

spins in all directions:

$$\mathcal{H} = \sum_{i=1}^{N-1} (JX_iX_{i+1} + JY_iY_{i+1} + Z_iZ_{i+1}). \quad (6.9)$$

Here the energy scale is set by choosing the ZZ -interaction to be of unit strength. Since the XX - and YY -interactions are constrained to have the same coupling J , this model is also referred to as the *XXZ model*.

Here we study the cases $J = 0.25, 0.50, 1.00$ for both the Ising model and the XXZ model.

Chemistry models — In order to evaluate Clapton for more complex Hamiltonians with a greater number of terms, we select three different molecules: H_2O , H_6 and LiH . We use Qiskit Nature [46] (which uses PySCF [175] behind the scenes) to construct the corresponding Hamiltonians. They are computed in the STO-3G basis and mapped to qubits using the parity mapping [164] with the two-qubit reduction applied. We limit the qubit count of all Hamiltonians to ten by restricting the active space of the molecules to six spatial orbitals. For each molecule we consider two different bond lengths l , one where classical methods yield high accuracy (small bond length) and one where the accuracy of classical methods is known to be less good (high bond length) [153]. The specific configurations are:

- H_2O : $l = 1.0 \text{ \AA}, 3.0 \text{ \AA}$ — 367 Hamiltonian terms
- H_6 : $l = 1.0 \text{ \AA}, 3.0 \text{ \AA}$ — 919 Hamiltonian terms
- LiH : $l = 1.5 \text{ \AA}, 4.5 \text{ \AA}$ — 631 Hamiltonian terms

Evaluation Method

Since Clapton proposes a transformation scheme which is based on the initial point of a VQA, we compare its performance against the current state-of-the-art initialization method CAFQA. While the original CAFQA framework was implemented using Bayesian optimiza-

tion [153], we build our own adapted version that employs an optimization engine similar to the one shown Figure 6.4. The main difference is that CAFQA optimizes the VQE ansatz A over Clifford-compatible angles $\boldsymbol{\theta}$ and only considers a noiseless cost term similar to $\mathcal{L}_0$.

Since CAFQA does not have a sense for noise, we furthermore implement an improved baseline to compare to: *noise-aware CAFQA*, abbreviated as *nCAFQA*. Just like CAFQA, nCAFQA finds an initialization for the original problem $\mathcal{H}$ by searching over Clifford-compatible angles, but it uses the noise-equipped ansatz $\tilde{A}$. It computes cost terms similar to Clapton’s $\mathcal{L}_{\mathcal{N}}$ and $\mathcal{L}_0$ and also uses an adapted optimization engine. Comparing Clapton to nCAFQA highlights the relevance of the Hamiltonian transformation step. While we refer to nCAFQA as a baseline, we would like to emphasize that it does not correspond to any prior art and gains its improvement over CAFQA through one of our major contributions, the classically efficient noise modeling.

We are comparing initialization schemes, thus we are mostly interested in the quality of the noisy evaluation of initial points. For Clapton this means $\boldsymbol{\theta} = \mathbf{0}$ on the transformed problem and for CAFQA/nCAFQA $\boldsymbol{\theta} = \boldsymbol{\theta}_{\text{CAFQA/nCAFQA}}$ on the original problem. However, since for practical applications we care about the converged solutions after the application of VQE, we also run the full VQE for all problems to show the impact of the initialization schemes on convergence and the final result. We build the VQE simulation within the Qiskit framework and use the SPSA optimizer [173].

Evaluation metric — For all studies, our primary metric is given by the performance of Clapton relative to CAFQA/nCAFQA. We compute the *relative improvement* η which describes to what extent Clapton can reduce the gap to the problem’s ideal ground state energy E_0 under noisy evaluation compared to a baseline:

$$\eta = \frac{E_0 - E_{\text{noisy}}(\text{baseline})}{E_0 - E_{\text{noisy}}(\text{Clapton})}. \quad (6.10)$$

Here “baseline” can either be CAFQA or nCAFQA. Since we consider benchmarks with up

to ten qubits, it is possible to compute E_0 exactly by diagonalizing the Hamiltonian. As an example, $\eta = 2$ means that Clapton was able to reduce the gap by half.

Evaluation on quantum backends — For our main studies we use both realistic noise models (not Clifford-only simulable) and real quantum hardware to obtain energy values for our Hamiltonians of choice. To this end we use Qiskit [152] and its built-in simulators as well as the IBM Quantum Platform [1] to access real devices. In particular, we select noise model snapshots (known as fake backends) of the IBM machines `nairobi` (seven qubits), `toronto` (27 qubits) and `mumbai` (27 qubits). We furthermore choose the cloud device `hanoi` (27 qubits) for our quantum experiments. For `nairobi` we constrain the set of benchmarks to only the physics models ($N = 7$) while for the larger machines we investigate all models ($N = 10$). Since quantum devices have a restricted topology, every VQE ansatz circuit A needs to be mapped and routed on the machine before execution is possible. For Clapton this means that this so-called *transpilation* step happens first to produce the transpiled ansatz A' (now consistent with the hardware constraints), which is then fed to the Clapton scheme and used to evaluate the loss $\mathcal{L}_{\mathcal{N}}$ (Equation (6.5)). Clapton extracts the required parameters for gate and measurement errors from the noise models or machine calibration data accessible through Qiskit.

Evaluation for specific sources of noise – Additionally, for a subset of the $N = 10$ benchmarks, we conduct studies for depolarizing noise and measurement noise individually, where we sweep a noise parameter p . In the case of depolarizing noise, p represents the single-qubit gate error and ranges in $5 \cdot 10^{-4} \dots 5 \cdot 10^{-3}$. The two-qubit gate errors are set to $10p$. When studying measurement noise, p describes the misassignment error (see Section 6.2.2) and we vary it within $5 \cdot 10^{-3} \dots 9.5 \cdot 10^{-2}$. For both studies we add thermal relaxation errors with characteristic decay times $T_1 = 50 \mu s, 100 \mu s, 150 \mu s$. The noise parameters are global across all qubits for this study. These chosen values are representative of what we find in the noise models of the quantum backends introduced above. In order to respect the

realistic constraint of a limited hardware connectivity, we transpile the ten-qubit ansatz A to match `toronto`’s topology for optimization and evaluation. The latter is performed using Qiskit’s AerSimulator.

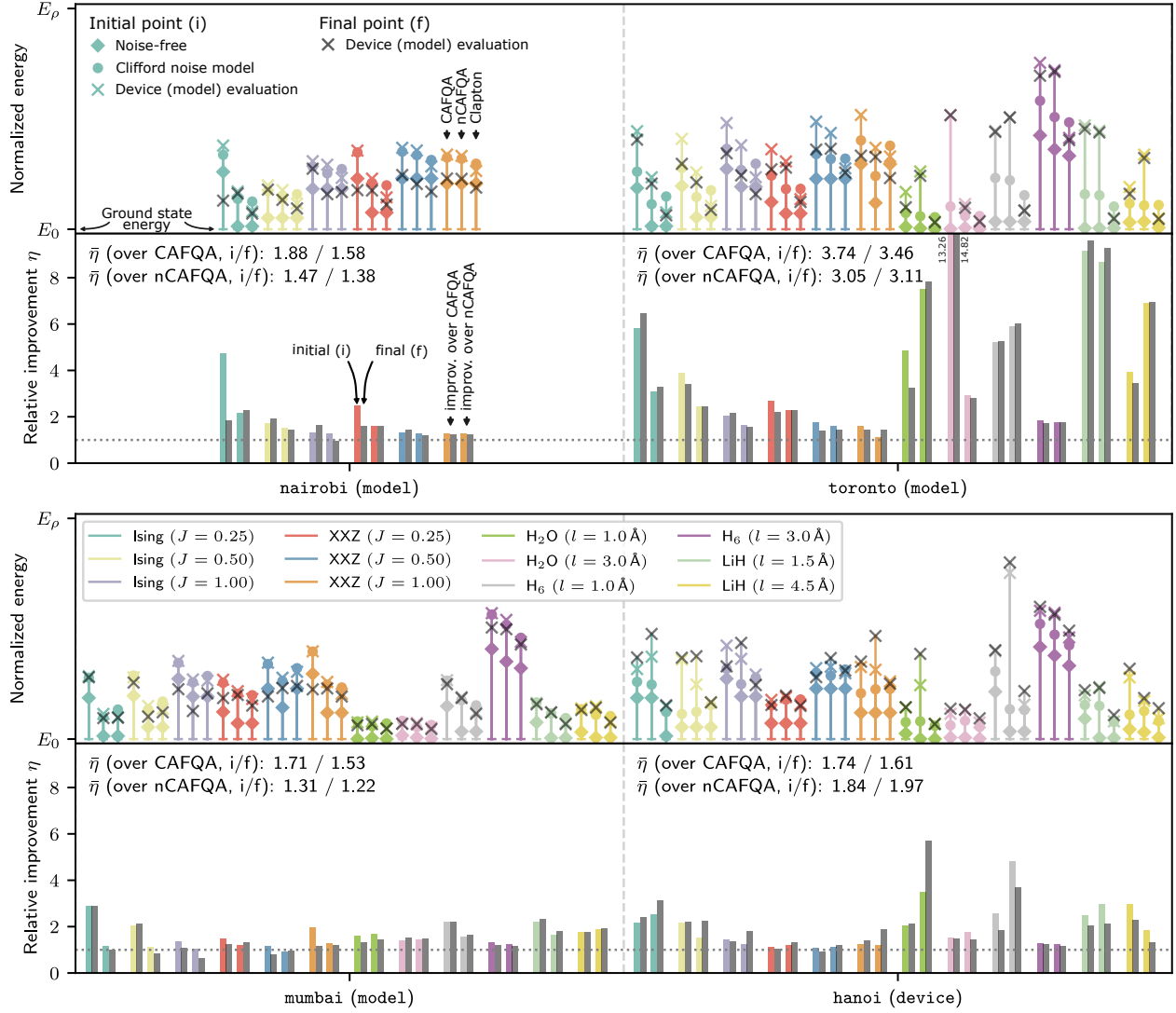
6.2.4 Evaluation

Having introduced the methodology, we now dive into the results and highlight the suitability of Clapton account for device errors in VQE applications.

VQE initial and final points on backends

For the general evaluation of our scheme under realistic noise effects, we start by running Clapton, nCAFQA and CAFQA for all VQE problems on all the backends as described in Section 6.2.3. In order to find the quality of the initial VQE points, we need to evaluate the realistic energy estimates for the specific problems and inputs. We achieve this by using Qiskit’s built-in simulators for the fake backends (“models”) `nairobi`, `toronto` and `mumbai`, while for the real machine `hanoi` we send jobs to the IBM Quantum Platform. Afterwards we run up to a thousand VQE iterations from each initial point on those backend noise models and find the quality of the converged solution. Therefore we only evaluate the final points on the actual hardware. Finally we determine relative improvements η for initial and final points and combine all findings in Figure 6.5.

Figure 6.5 consists of two similar parts showing results for two backends respectfully. The top rows present the quality analysis of found solutions from each method for each benchmark. Different colors correspond to different benchmarks. For easier visualization they are normalized such that two fixpoints align: the ground state energy E_0 and the energy of the fully mixed state $E_\rho = \text{tr}[\rho\mathcal{H}] = \text{tr}[\mathcal{H}]/2^N$ ($\rho = I_N/2^N$). We show three different energy values for each initial point: (1) The noiseless case ($\diamond$), which is a lower bound for any noisy value and used by all three methods in their optimization (Clapton:



$\mathcal{L}_0$). (2) The energy under the approximate Clifford noise model ($\circ$), which Clapton and nCAFQA use to evaluate the other part of their cost function (Clapton: $\mathcal{L}_N$). (3) The energy under the full complex noise model derived from a fake backend or, in the case of `hanoi`, the result from the hardware experiment ($\times$). Clearly the most relevant property is (3). We observe that for almost all benchmarks Clapton finds significantly better starting points than the other two methods, while nCAFQA often performs better than CAFQA as one would expect. Another interesting observation is that the discrepancy between (2) and (3) is typically the lowest for Clapton (see also Figure 6.3 for a magnified version of one benchmark), indicating that the Clifford noise modeling becomes more accurate when we transform to the $|0\rangle$ state. This aligns with the Clapton hypothesis that exploring around the $|0\rangle$ is favorable as the influence of relaxation (which is non-Clifford) is mitigated.

We mark the final point after the application of VQE with a dark $\times$ and find that the advantage of Clapton persists through the convergence in nearly all cases. In few cases a CAFQA or nCAFQA instance is able to catch up and overtake a Clapton instance. We address this further in the next section. In the case of `hanoi` these results were obtained by extracting the noise model from calibration data and perform the VQE iterations on the classical simulator. The final points are then evaluated using hardware experiments. The overhead of submitting jobs to the cloud for every iteration in combination with long queue times would render full VQE on the real devices infeasible.

The bottom rows visualize the relative improvements that Clapton enables, both for the initial and the final points. When comparing to CAFQA, we find (geometric) mean improvements ranging from 1.7x to 3.7x for initial points, with improvements for individual VQE problems up to 13.3x. For the final points the mean improvements range 1.5x to 3.5x (14.8x max). As expected, the improvements over nCAFQA are smaller but still significant. This is because nCAFQA already benefits strongly from modeling gate and readout errors during its optimization which is one of our key contributions. The Clapton transformation

then further adds resilience to state decay which yields its advantage over nCAFQA. However, one should keep in mind that even a relative improvement of (for example) 1.3 is remarkable as it corresponds to a $\sim 23\%$ gap reduction.

Looking across devices we observe that experiments executed on the `toronto` model particularly benefit from Clapton, while looking across problem types shows that chemistry problems profit greatly. We suspect the latter to be caused by the substantially larger number of Pauli terms in the respective Hamiltonians, allowing for a greater impact of the Clapton transformation – this is promising for real world applications. We emphasize the observed improvements for the real quantum device `hanoi`: Device execution of noise-model-optimized transformations show improvements across all benchmarks except for one case where the final point was slightly worse. This should not be taken for granted as there can be no guarantee that simplified noise models represent effects on real hardware well.

VQE convergence — In Figure 6.6, we visualize the VQE convergence of the ten-qubit XXZ model with (a) $J = 0.25$ and (b) $J = 1.00$ when run on the models of the four 27-qubit machines `toronto` and `hanoi`. $J = 0.25$ represents a case where stabilizer states can provide a good approximation to the ground state while for $J = 1.00$ that is not the case. We see here that Clapton finds good starting points and converges well in both cases. We note that even though the Clapton solution starts better than the nCAFQA solution for $J = 1.00$ on `hanoi`, the latter overtakes quickly and finds a better final point. This suggests that the cause for these two observations lies with SPSA and its configuration. Nondeterminism in noise, barren plateaus and uncertainty in sampled energy estimates can be detrimental to its performance. This issue could be addressed by configuring the optimizer hyperparameters more carefully or studying other optimization frameworks, which is outside the scope of this work. We emphasize that while Clapton is able to produce high-quality starting points, it does not come with the guarantee of best convergence every time, as this is influenced by factors Clapton (and any other VQA initialization method) does not have control over.

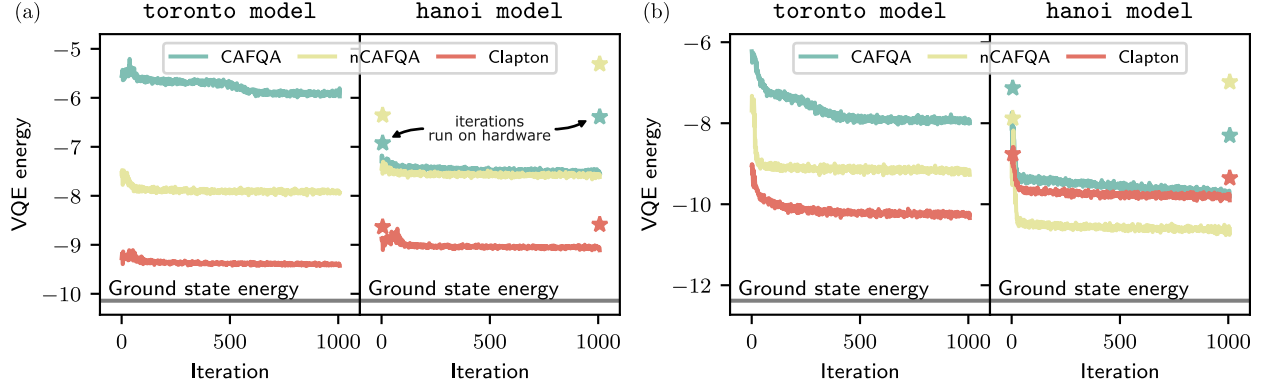


Figure 6.6: Visualizations of the VQE convergence of the ten-qubit XXZ model with (a) $J = 0.25$ and (b) $J = 1.00$, run on two different noise models. The Clapton initialization starts better than the baselines and typically converges better. Initial and final evaluations on real machine `hanoi` are shown with stars. Hardware experiments yield worse final points in some cases, suggesting a device-model discrepancy that Clapton does not seem to be strongly affected by.

The experimental results of the initial and final points on `hanoi` are visualized by stars. It is interesting to observe that in some cases experiments on real hardware yield worse final points that are worse than the initial points. This suggests a discrepancy between the extracted noise model which is used for the simulated VQE and the actual device. This is unsurprising as accurately capturing the quantum effects on many-qubit devices is very challenging. We note that even though Clapton did not find the best final point for $J = 1.00$ on `hanoi` according to the simulation, it achieved the best experimental result.

Varying the strength of individual noise channels

Having presented the positive impact on Clapton when applied to full machine noise models, we now turn to studying individual noise channels in an isolated environment in particular. As described towards the end of Section 6.2.3, we sweep gate errors and measurement errors, which are both essential Clapton ingredients, for different decay times T_1 . Results are presented for the initial point of selected VQE problems and we focus on comparing to the modified baseline nCAFQA.

In Figure 6.7 we show the relative improvements for four different benchmarks when sweeping the gate errors. We note that the Clapton transformation is especially useful for lower error rates when solving the complex chemistry models, while we observe an upwards trend for the Ising model as the gate error increases. As expected, shorter T_1 times, meaning stronger thermal relaxation, typically lead to better improvements as the nCAFQA initial point suffers from more state decays. This influence is actively counteracted by Clapton by transforming to the qubit ground state $|0\rangle$.

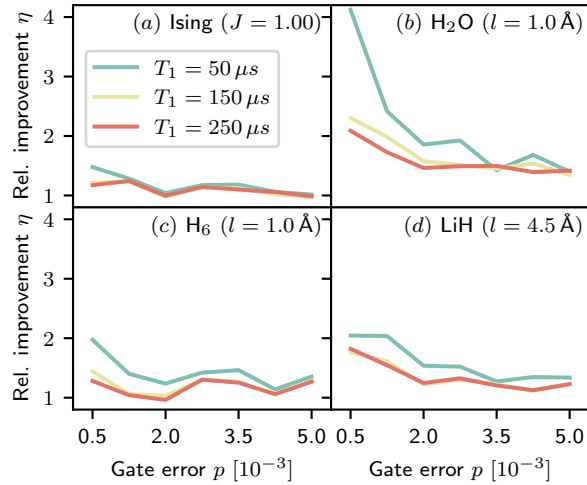


Figure 6.7: The impact of Clapton when solely sweeping the single-qubit gate error p (two-qubit gate error $10p$), for three different decay times and a selection of benchmarks. Relative improvement is evaluated at the initial VQE point.

We find similar results when varying measurement errors on the readout qubits as shown in Figure 6.8. The noticeable difference is that the Ising model seems to be more robust to this kind of error as only small improvements over the baseline can be observed. The chemistry problems still profit significantly from the transformation, underlining that Clapton is well suited for this type of Hamiltonian structure.

Clapton's performance scaling

In this last results section we study the classical compute Clapton requires for its optimization. We investigate the performance scaling with respect to the qubit count N and to this

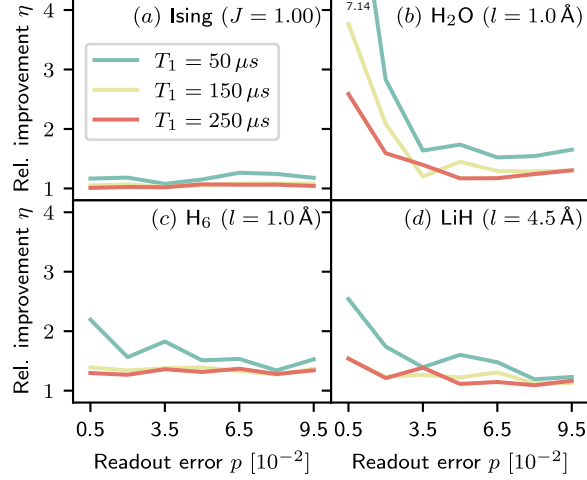


Figure 6.8: The impact of Clapton when solely sweeping the measurement error p , for three different decay times and a selection of benchmarks. Relative improvement is evaluated at the initial VQE point.

end consider the Ising model with $\mathcal{O}(N)$ terms. For the purpose of this study transpilation is not required so we keep the ansatz structure A , which grows in depth and width with $\mathcal{O}(N)$. Thus the complexity of the circuit that needs to be simulated scales with $\mathcal{O}(N^2)$.

For $N = 11, \dots, 40$ we run Clapton five times from random initial configurations and measure the time until convergence (that is, no loss function improvement even after two retry rounds). We specify $\mathfrak{s} = 10$ and use ten processes, thus we only parallelize over the genetic algorithm instances $\mathbf{GA}_1, \dots, \mathbf{GA}_{10}$ and not over the Pauli terms or population size.

Figure 6.9 shows the scaling of the total compute time t with qubit count N . Up to 20 qubits we barely observe any variation in optimization time across the five guesses, noting an essentially linear increase with N . For higher qubit counts it shows that an increasing number of rounds is required to achieve convergence, which is expected due to the growing number of parameters ($\mathcal{O}(N)$) and the vast search space Γ (dimension exponential in N). While the right half of Figure 6.9 hints at this exponential scaling, it is important to note that the compute time never exceeded ten hours even for $N = 40$, which is already well beyond what would be classically simulable if we considered more accurate quantum models.

The inset visualizes the average time per round τ , which scales like the circuit spacetime

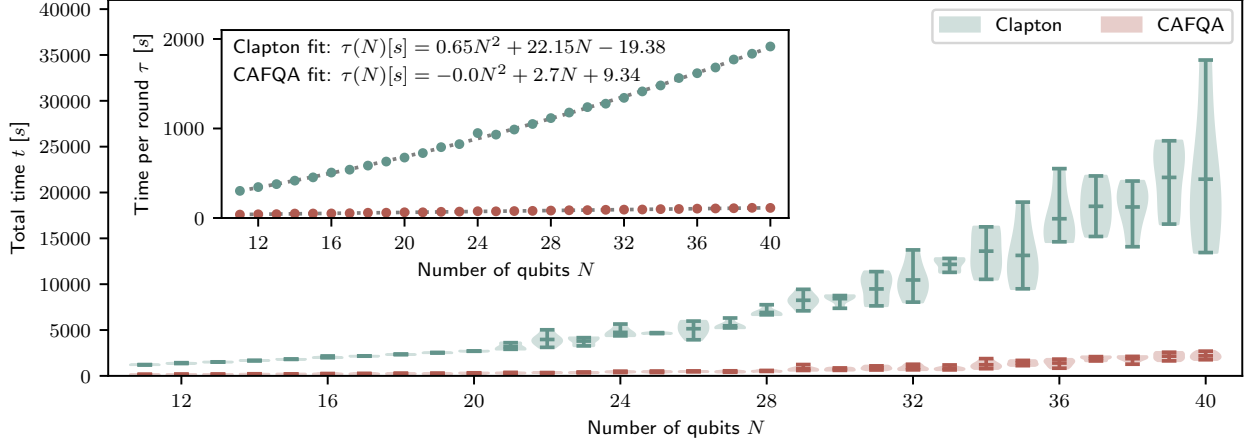


Figure 6.9: Scaling of Clapton optimization time with the number of qubits N for the Ising model ($J = 0.25$). Each instance is run from five initial guesses using $\mathbf{s} = 10$. The inset presents the scaling of the average optimization time per round, which follows an $\mathcal{O}(N^2)$ trend. The baseline CAFQA method shows a considerably faster performance as it only computes noiseless evaluations.

volume with $\mathcal{O}(N)$. This makes sense as rounds have a fixed number of iterations and their speed is therefore determined by Stim’s simulation efficiency. The round duration could be reduced by a constant factor by furthermore parallelizing over the Pauli term count and GA population size in each iteration, which would speed up Clapton’s performance by a constant factor. This can be realized in a straightforward way due to the parallel nature of the Clapton design with the use of genetic algorithms (see Figure 6.4), underlining Clapton’s efficient and scalable implementation.

We perform the same study for the baseline methods CAFQA and nCAFQA. While nCAFQA yields a performance scaling that is essentially equivalent to Clapton due to its very similar design (we thus do not show it in Figure 6.9), our CAFQA implementation runs in much less time to achieve convergence. This is because only noiseless Clifford simulations are required where no sampling is necessary, every Pauli expectation value can be determined in one circuit evaluation. Figure 6.9 includes these results and the inset shows how Stim’s fast Clifford simulation leads to linear scaling of round time τ over the considered qubit count range.

This study as well as all other Clapton or CAFQA optimizations demonstrated in this chapter were performed on Intel Gold 6248R CPUs.

6.2.5 *Related work*

Although Clifford group operations and Pauli measurements do not suffice for universal quantum computing, many quantum domains such as quantum networks [184], error correction codes [157], teleportation [67] and error mitigation [37, 174] focus on applications within the Clifford space.

In terms of related work in the space of VQA optimization, Mitarai et al. [120] advocate a perturbative expansion of the cost function from Hartree-Fock initialization, potentially yielding improved performance. Sauvage et al. [162] introduce FLIP using machine learning for initialization. MetaVQE by Cervera et al. [26] encodes Hamiltonian parameters into the ansatz for a more reliable estimate of the ground state. Wurtz et al. [204] use classical solutions to aid some special sets of problem-specific circuits. Wang et al. propose QuantumNAS [191], a framework for noise-adaptive co-search of the variational circuit structure and qubit mapping. Their approach is substantially different to the Clapton proposal as they leave the degrees of freedom that come from transforming the Hamiltonian untouched and focus on the manipulation of the original problem implementation.

Various orthogonal error mitigation strategies are also available to address quantum errors at different stages of a quantum program: Pre-processing techniques (e.g., noise-aware mapping [129]), post-processing techniques (e.g., measurement error mitigation [177]) or in-situ techniques (e.g., dynamical decoupling [12, 83, 93, 143]). Previous studies [85, 192] emphasize error mitigation’s role in enhancing VQA accuracy and trainability, with techniques like zero-noise extrapolation (ZNE) and addressing noise-induced barren plateaus. The VAQEM approach [154] tailors existing error mitigation techniques dynamically to real-time execution characteristics of VQAs on a quantum machine.

6.2.6 Conclusion

In this current NISQ era VQAs represent a promising avenue to obtain useful insights from quantum devices. However, noise is the defining characteristic of this period and will always have a detrimental effect on any quantum computing results, rendering adjustment to this noise through sophisticated error mitigation techniques indispensable. The Clapton proposal directly addresses this issue for VQAs by exploiting the freedom that can be found in the problem formulation (the Hamiltonian $\mathcal{H}$). It makes use of the observation that, given a specific quantum device, some Hamiltonians are less susceptible to its noise pattern than others when running the VQA. Clapton performs a classically simulable search over transformations that make the problem formulation more compatible with the hardware noise characteristics, yielding a new Hamiltonian $\hat{\mathcal{H}}$ that is then used during the full VQA execution. This work shows the beneficial impact of this approach across a variety of benchmarks.

Clapton can be viewed as a pre-processing error mitigation technique, which may be combined with other popular error mitigation methods outlined in Section 6.2.5; an interesting idea for future work. It is a pure software tool which models and simulates the target quantum machine efficiently by expressing all operations in terms of Clifford gates.

As quantum computing research moves closer to fault-tolerance and error correction, it is important to point out that the Clapton is not limited to the NISQ regime. Error-corrected quantum devices will also run VQEs and errors on (logical) qubits will still occur, albeit with a significantly reduced rate. In fact, errors on these machines are discretized and typically expressed in terms of bit flips (X operations) and phase flips (Z operations) [20], which directly suggests the depolarizing error model to describe these effects. Therefore, given the implementation of Clapton, it might prove itself to be even more relevant in the future.

CHAPTER 7

CONCLUSION AND FUTURE DIRECTIONS

The fundamental insight of this thesis is that leveraging the physical features and characteristics of a quantum technology enable optimizations across its entire stack. We particularly identify the pulse level to be a crucial layer that can motivate architectural design decisions and we provide demonstrations in three distinct settings under different perspectives. Using analytical device modeling and numerical pulse optimization, we validate the feasibility of higher-dimensional qudit systems, introduce a gate language for optical lattice interferometers and present universal gate schemes for noise-protected Kerr-cat qubits. While the approaches to addressing these research problems are conceptually similar (i.e., minimizing objective functions over parameters of the control functions of interest), the specific implementations differ depending on the respective physical system and underlying constraints.

On the matter of qudit systems, we used open-loop quantum optimal control methods to show that qudit gate durations scale more favorably with dimension than originally assumed, which encourages their use as architectural components. In particular, ququarts, which can store the information content of two qubits, become promising candidates as fundamental units on quantum processors, and their architectural implications have been studied in related follow-up works [109, 110]. This directly showcases how low-level optimization enables high-level design decisions on the systems level. In order to verify the feasibility of qudits in practice, we implemented optimal control solutions in experiment on a superconducting transmon processor and demonstrated the coherent control of a ququart.

Our studies of qudit systems only scratched the surface and motivate many avenues of future research. The pulse optimizations we presented were performed without the assumption of noise, therefore including noise contributions in the model of the system is a natural next step, in particular given the fact that higher additional levels in qudits introduce increasingly more noise. An alternative strategy is given by switching to closed-loop optimal

control methods where the pulses are directly optimized using the feedback obtained from execution on the quantum hardware, which allows for the most accurate representation of the noisy physical system. However, this opens up other challenges such as the efficient evaluation of the objective function from experimental data. Another direction is to realize qudit computation in larger systems with multiple units.

In the context of quantum sensing, we employed a different open-loop optimal control technique, known as direct trajectory optimization, that allowed us to exploit the structures of the optical lattice system as well as the optimization targets. Specifically, we are able to design highly complex gate primitives for programmable sensing by imposing additional constraints on the optimization problem that guide the numerical solver and are reasonably motivated by physical intuition. Thus a key observation is that identifying such intuition and incorporating it as supplemental structure, which is very specific to the underlying system, can lead to significant improvements in the optimization results or enable them in the first place. In the second study we applied this theme in a similar way: We directly optimized the quantum sensor’s sensitivity in an end-to-end fashion and built the interferometer structure into the problem using constraints.

Our studies on quantum sensing were purely theoretical, therefore it would be interesting to evaluate the performance of the optimal control solutions on hardware. Furthermore, this particular sensing platform using cold atoms is usually analyzed using the single-particle wavefunction under the assumption of no interactions, however, that is not the case for realistic Bose-condensed atom clouds. Thus extending the control optimization framework to the more general description including interactions seems necessary.

In our work on Kerr-cat qubits we were primarily interested in the conceptual realization of and intuition behind the proposed gate schemes, which is why we did not employ optimal control methods and instead limited our exploration to simple pulse shapes with few tunable parameters. Here the crucial insight came from modeling the system and energy

landscape over a broader parameter regime and identifying local noise-insensitive points, which ultimately lead to the proposal of the protected adiabatic $Z(\theta)$ scheme. Therefore, this is another example of exploiting device-specific (Hamiltonian-specific) structure. With the $Z(\theta)$ scheme we were able to recover universal computation with Kerr-cat qubits while protected to frequency-noise-induced phase flips like they occur in superconducting architectures. We thereby successfully addressed the critical limitations imposed by the tradeoff between noise protection and controllability, making Kerr-cat qubits an even more attractive candidate as quantum processing units.

Now that we have gained this intuition, applying optical control to the Kerr-cat system would be a viable next step. In particular, for our adiabatic gate schemes one could define the optimal control problem such that the gate time and robustness is optimized while not sacrificing adiabaticity. Another direction could extend the discussion on the impact of time-dependent errors that we only briefly touched on. Furthermore, as we based our entire analysis on an effective Kerr oscillator model, a more realistic modeling within the language of circuit quantum electrodynamics would be crucial in order to understand how our findings would translate to realistic hardware implementations.

The optimizations we proposed on the algorithm layer of the stack are not based on pulse control but observations about the noise properties of quantum hardware instead. With our framework Clapton we demonstrated that adapting quantum programs to the noise characteristics of the underlying hardware can lead to significant improvements in the performance, and we showed that variational algorithms are particularly well suited for this error mitigation technique. We leverage the observations that certain families of gates (in this case Clifford gates) are efficiently simulable on classical computers and that noise models constrained to these families capture the actual hardware noise sufficiently well. Clapton is a pre-processing error mitigation technique that could work in conjunction with other error mitigation techniques (e.g., measurement error mitigation [39]), and we showed its positive

impact for a variety of benchmarks in simulation and experiment.

In our demonstrations of Clapton we used standard ansatz structures (in this case the hardware-efficient circular ansatz) that were not tailored to the specific problem Hamiltonian. Investigating the performance of Clapton for different ansatz structures and the dependence of ansatz and problem symmetries in connection with Clapton would be worthwhile future research directions. Furthermore, extending the noise modeling beyond depolarizing noise or considering other efficiently simulable gate families such as matchgates [121] are interesting candidates for follow-up work. Alternatively, it would be interesting to study if the transformation idea behind Clapton can be transferred to other quantum applications beyond variational algorithms.

As quantum technologies become larger in size and grow in complexity, it will remain crucial to understand the interplay between the different layers of the stack and how optimizations at one layer can impact the performance of the entire system. For the case of quantum computing, even though the community moves towards the pursuit of full fault-tolerance that could imply our strategies to become less relevant eventually, optimizations like the ones we propose can push the limits of current devices and thereby help to bridge the gap to the fault-tolerant era.

REFERENCES

- [1] IBM Quantum.
- [2] Rajeev Acharya, Dmitry A. Abanin, Laleh Aghababaie-Beni, Igor Aleiner, Trond I. Andersen, Markus Ansmann, Frank Arute, Kunal Arya, Abraham Asfaw, Nikita As-trakhantsev, Juan Atalaya, Ryan Babbush, Dave Bacon, Brian Ballard, Joseph C. Bardin, Johannes Bausch, Andreas Bengtsson, Alexander Bilmes, Sam Blackwell, Sergio Boixo, Gina Bortoli, Alexandre Bourassa, Jenna Bovaird, Leon Brill, Michael Broughton, David A. Browne, Brett Buchea, Bob B. Buckley, David A. Buell, Tim Burger, Brian Burkett, Nicholas Bushnell, Anthony Cabrera, Juan Campero, Hung-Shen Chang, Yu Chen, Zijun Chen, Ben Chiaro, Desmond Chik, Charina Chou, Jahan Claes, Agnetta Y. Cleland, Josh Cogan, Roberto Collins, Paul Conner, William Courtney, Alexander L. Crook, Ben Curtin, Sayan Das, Alex Davies, Laura De Lorenzo, Dripto M. Debroy, Sean Demura, Michel Devoret, Agustin Di Paolo, Paul Donohoe, Ilya Drozdov, Andrew Dunsworth, Clint Earle, Thomas Edlich, Alec Eickbusch, Aviv Moshe Elbag, Mahmoud Elzouka, Catherine Erickson, Lara Faoro, Edward Farhi, Vinicius S. Ferreira, Leslie Flores Burgos, Ebrahim Forati, Austin G. Fowler, Brooks Foxen, Suhas Ganjam, Gonzalo Garcia, Robert Gasca, Élie Genois, William Giang, Craig Gidney, Dar Gilboa, Raja Gosula, Alejandro Grajales Dau, Dietrich Graumann, Alex Greene, Jonathan A. Gross, Steve Habegger, John Hall, Michael C. Hamilton, Monica Hansen, Matthew P. Harrigan, Sean D. Harrington, Francisco J. H. Heras, Stephen Heslin, Paula Heu, Oscar Higgott, Gordon Hill, Jeremy Hilton, George Holland, Sabrina Hong, Hsin-Yuan Huang, Ashley Huff, William J. Huggins, Lev B. Ioffe, Sergei V. Isakov, Justin Iveland, Evan Jeffrey, Zhang Jiang, Cody Jones, Stephen Jordan, Chaitali Joshi, Pavol Juhas, Dvir Kafri, Hui Kang, Amir H. Karamlou, Kostyantyn Kechedzhi, Julian Kelly, Trupti Khairé, Tanuj Khattar, Mostafa Khezri, Seon Kim, Paul V. Klimov, Andrey R. Klotz, Bryce Kobrin, Pushmeet Kohli, Alexander N. Korotkov, Fedor Kostritsa, Robin Kothari, Borislav Kozlovskii, John Mark Kreikebaum, Vladislav D. Kurilovich, Nathan Lacroix, David Landhuis, Tiano Lange-Dei, Brandon W. Langley, Pavel Laptev, Kim-Ming Lau, Loïck Le Guevel, Justin Ledford, Joonho Lee, Kenny Lee, Yuri D. Lensky, Shannon Leon, Brian J. Lester, Wing Yan Li, Yin Li, Alexander T. Lill, Wayne Liu, William P. Livingston, Aditya Locharla, Erik Lucero, Daniel Lundahl, Aaron Lunt, Sid Madhuk, Fionn D. Malone, Ashley Maloney, Salvatore Mandrà, James Manyika, Leigh S. Martin, Orion Martin, Steven Martin, Cameron Maxfield, Jarrod R. McClean, Matt McEwen, Seneca Meeks, Anthony Megrant, Xiao Mi, Kevin C. Miao, Amanda Mieszala, Reza Molavi, Sebastian Molina, Shirin Montazeri, Alexis Morvan, Ramis Movassagh, Wojciech Mruczkiewicz, Ofer Naaman, Matthew Neeley, Charles Neill, Ani Nersisyan, Hartmut Neven, Michael Newman, Jiun How Ng, Anthony Nguyen, Murray Nguyen, Chia-Hung Ni, Murphy Yuezhen Niu, Thomas E. O'Brien, William D. Oliver, Alex Opremcak, Kristoffer Ottosson, Andre Petukhov, Alex Pizzuto, John Platt, Rebecca Potter, Orion Pritchard, Leonid P. Pryadko, Chris Quintana, Ganesh Ramachandran, Matthew J. Reagor, John Redding, David M. Rhodes, Gabrielle Roberts, Elliott Rosenberg, Emma Rosenfeld,

- Pedram Roushan, Nicholas C. Rubin, Negar Saei, Daniel Sank, Kannan Sankaragomathi, Kevin J. Satzinger, Henry F. Schurkus, Christopher Schuster, Andrew W. Senior, Michael J. Shearn, Aaron Shorter, Noah Shutt, Vladimir Shvarts, Shradha Singh, Volodymyr Sivak, Jindra Skruzny, Spencer Small, Vadim Smelyanskiy, W. Clarke Smith, Rolando D. Somma, Sofia Springer, George Sterling, Doug Strain, Jordan Suchard, Aaron Szasz, Alex Szein, Douglas Thor, Alfredo Torres, M. Mert Torunbalci, Abeer Vaishnav, Justin Vargas, Sergey Vdovichev, Guifre Vidal, Benjamin Villalonga, Catherine Vollgraft Heidweiller, Steven Waltman, Shannon X. Wang, Brayden Ware, Kate Weber, Travis Weidel, Theodore White, Kristi Wong, Bryan W. K. Woo, Cheng Xing, Z. Jamie Yao, Ping Yeh, Bicheng Ying, Juhwan Yoo, Noureldin Yosri, Grayson Young, Adam Zalcman, Yaxing Zhang, Ningfeng Zhu, Nicholas Zobrist, and Google Quantum AI and Collaborators. Quantum error correction below the surface code threshold. *Nature*, pages 1–7, December 2024.
- [3] Shah Saad Alam, Victor E. Colussi, John Drew Wilson, Jarrod T. Reilly, Michael A. Perlin, and Murray J. Holland. Robust Quantum Sensing with Multiparameter Decorrelation. May 2024.
- [4] Takaaki Aoki, Taro Kanao, Hayato Goto, Shiro Kawabata, and Shumpei Masuda. Control of the ZZ coupling between Kerr cat qubits via transmon couplers. *Physical Review Applied*, 21(1):014030, January 2024.
- [5] Takaaki Aoki, Akiyoshi Tomonaga, Kosuke Mizuno, and Shumpei Masuda. Residual ZZ-coupling suppression and fast two-qubit gate for Kerr-cat qubits based on level-degeneracy engineering. *Applied Physics Letters*, 126(4):044004, January 2025.
- [6] Jonathan M. Baker, Casey Duckering, and Frederic T. Chong. Efficient Quantum Circuit Decompositions via Intermediate Qudits. In *2020 IEEE 50th International Symposium on Multiple-Valued Logic (ISMVL)*, pages 303–308, November 2020.
- [7] Jonathan M. Baker, Casey Duckering, Pranav Gokhale, Natalie C. Brown, Kenneth R. Brown, and Frederic T. Chong. Improved Quantum Circuits via Intermediate Qutrits. *ACM Transactions on Quantum Computing*, 1(1):2:1–2:25, October 2020.
- [8] R. Barends, J. Kelly, A. Megrant, A. Veitia, D. Sank, E. Jeffrey, T. C. White, J. Mutus, A. G. Fowler, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, C. Neill, P. O’Malley, P. Roushan, A. Vainsencher, J. Wenner, A. N. Korotkov, A. N. Cleland, and John M. Martinis. Superconducting quantum circuits at the surface code threshold for fault tolerance. *Nature*, 508(7497):500–503, April 2014.
- [9] Yuval Baum, Mirko Amico, Sean Howell, Michael Hush, Maggie Liuzzi, Pranav Mundada, Thomas Merkh, Andre R.R. 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, November 2021.
- [10] J. S. Bell. On the Einstein Podolsky Rosen paradox. *Physics Physique Fizika*, 1(3):195–200, November 1964.

- [11] Maxime Ben Dahan, Ekkehard Peik, Jakob Reichel, Yvan Castin, and Christophe Salomon. Bloch Oscillations of Atoms in an Optical Potential. *Physical Review Letters*, 76(24):4508–4511, June 1996.
- [12] M J Biercuk, A C 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, July 2011.
- [13] Immanuel Bloch, Theodor W. Hänsch, and Tilman Esslinger. Atom Laser with a cw Output Coupler. *Physical Review Letters*, 82(15):3008–3011, April 1999.
- [14] M. S. Blok, V. V. Ramasesh, T. Schuster, K. O’Brien, J. M. Kreikebaum, D. Dahlen, A. Morvan, B. Yoshida, N. Y. Yao, and I. Siddiqi. Quantum Information Scrambling on a Superconducting Qutrit Processor. *Physical Review X*, 11(2):021010, April 2021.
- [15] Dolev Bluvstein, Simon J. Evered, Alexandra A. Geim, Sophie H. Li, Hengyun Zhou, Tom Manovitz, Sepehr Ebadi, Madelyn Cain, Marcin Kalinowski, Dominik Hangleiter, J. Pablo Bonilla Ataides, Nishad Maskara, Iris Cong, Xun Gao, Pedro Sales Rodriguez, Thomas Karolyshyn, Giulia Semeghini, Michael J. Gullans, Markus Greiner, Vladan Vuletić, and Mikhail D. Lukin. Logical quantum processor based on reconfigurable atom arrays. *Nature*, 626(7997):58–65, February 2024.
- [16] Alex Bocharov, Martin Roetteler, and Krysta M. Svore. Factoring with qutrits: Shor’s algorithm on ternary and metaplectic quantum architectures. *Physical Review A*, 96(1):012306, July 2017.
- [17] F. Bonechi, E. Celeghini, R. Giachetti, E. Sorace, and M. Tarlini. Heisenberg XXZ model and quantum Galilei group. *Journal of Physics A: Mathematical and General*, 25(15):L939, August 1992.
- [18] Gavin K. Brennen, Carlton M. Caves, Poul S. Jessen, and Ivan H. Deutsch. Quantum Logic Gates in Optical Lattices. *Physical Review Letters*, 82(5):1060–1063, February 1999.
- [19] Weizhou Cai, Yuwei Ma, Weiting Wang, Chang-Ling Zou, and Luyan Sun. Bosonic quantum error correction codes in superconducting quantum circuits. *Fundamental Research*, 1(1):50–67, January 2021.
- [20] A. R. Calderbank and Peter W. Shor. Good quantum error-correcting codes exist. *Physical Review A*, 54(2):1098–1105, August 1996.
- [21] Earl Campbell. A series of fast-paced advances in Quantum Error Correction. *Nature Reviews Physics*, 6(3):160–161, March 2024.
- [22] Tommaso Caneva, Tommaso Calarco, and Simone Montangero. Chopped random-basis quantum optimization. *Physical Review A*, 84(2):022326, August 2011.

- [23] Shuxiang Cao, Mustafa Bakr, Giulio Campanaro, Simone D. Fasciati, James Wills, Deep Lall, Boris Shteynas, Vivek Chidambaram, Ivan Rungger, and Peter Leek. Emulating two qubits with a four-level transmon qudit for variational quantum algorithms. *Quantum Science and Technology*, 9(3):035003, April 2024.
- [24] Olivier Carraz, Christian Siemes, Luca Massotti, Roger Haagmans, and Pierluigi Silvestrin. A Spaceborne Gravity Gradiometer Concept Based on Cold Atom Interferometers for Measuring Earth’s Gravity Field. *Microgravity Science and Technology*, 26(3):139–145, October 2014.
- [25] Christophe Cassens, Bernd Meyer-Hoppe, Ernst Rasel, and Carsten Klempt. Entanglement-Enhanced Atomic Gravimeter. *Physical Review X*, 15(1):011029, February 2025.
- [26] Alba Cervera-Lierta, Jakob S. Kottmann, and Alán Aspuru-Guzik. Meta-Variational Quantum Eigensolver: Learning Energy Profiles of Parameterized Hamiltonians for Quantum Simulation. *PRX Quantum*, 2(2):020329, May 2021.
- [27] Benjamin J. Chapman, Stijn J. De Graaf, Sophia H. Xue, Yaxing Zhang, James Teoh, Jacob C. Curtis, Takahiro Tsunoda, Alec Eickbusch, Alexander P. Read, Akshay Kootandavida, Shantanu O. Mundhada, Luigi Frunzio, M.H. Devoret, S.M. Girvin, and R.J. Schoelkopf. High-On-Off-Ratio Beam-Splitter Interaction for Gates on Bosonically Encoded Qubits. *PRX Quantum*, 4(2):020355, June 2023.
- [28] Senrui Chen, Zhihan Zhang, Liang Jiang, and Steven T. Flammia. Efficient self-consistent learning of gate set Pauli noise, October 2024.
- [29] Liang-Ying Chih. *Machine-Learning-Based Design of Quantum Systems for Extreme Sensing*. PhD thesis, University of Colorado Boulder, Boulder, 2022/2022-12.
- [30] Liang-Ying Chih and Murray Holland. Reinforcement-learning-based matter-wave interferometer in a shaken optical lattice. *Physical Review Research*, 3(3):033279, September 2021.
- [31] Liang-Ying Chih and Murray Holland. Reinforcement learning for rotation sensing with ultracold atoms in an optical lattice. *Physical Review Research*, 6(4):043191, November 2024.
- [32] Yujin Cho, Dipti Jasrasaria, Keith G Ray, Daniel M Tennant, Vincenzo Lordi, Jonathan L DuBois, and Yaniv J Rosen. Simulating noise on a quantum processor: Interactions between a qubit and resonant two-level system bath. *Quantum Science and Technology*, 8(4):045023, September 2023.
- [33] Isaac L. Chuang and M. A. Nielsen. Prescription for experimental determination of the dynamics of a quantum black box. *Journal of Modern Optics*, 44(11-12):2455–2467, November 1997.

- [34] Barry A. Cipra. An introduction to the Ising model. *Am. Math. Monthly*, 94(10):937–959, December 1987.
- [35] P. T. Cochrane, G. J. Milburn, and W. J. Munro. Macroscopically distinct quantum superposition states as a bosonic code for amplitude damping. *Physical Review A*, 59(4):2631–2634, April 1999.
- [36] Claude Cohen-Tannoudji and David Guéry-Odelin. *Advances in Atomic Physics*. WORLD SCIENTIFIC, 2011.
- [37] Piotr Czarnik, Andrew Arrasmith, Patrick J. Coles, and Lukasz Cincio. Error mitigation with Clifford quantum-circuit data. *Quantum*, 5:592, November 2021.
- [38] Andrew J. Daley, Immanuel Bloch, Christian Kokail, Stuart Flannigan, Natalie Pearson, Matthias Troyer, and Peter Zoller. Practical quantum advantage in quantum simulation. *Nature*, 607(7920):667–676, July 2022.
- [39] Siddharth Dangwal, Gokul Subramanian Ravi, Poulami Das, Kaitlin N. Smith, Jonathan Mark Baker, and Frederic T. Chong. VarSaw: Application-tailored Measurement Error Mitigation for Variational Quantum Algorithms. In *Proceedings of the 28th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 4*, ASPLOS ’23, pages 362–377, New York, NY, USA, February 2024. Association for Computing Machinery.
- [40] Siddharth Dangwal, Gokul Subramanian Ravi, Lennart Maximilian Seifert, Poulami Das, James Sud, and Frederic T. Chong. COMPASS: Compiler Pass Selection For Improving Fidelity Of NISQ Applications. In *2024 IEEE International Conference on Rebooting Computing (ICRC)*, pages 1–14, December 2024.
- [41] Andrew S. Darmawan, Benjamin J. Brown, Arne L. Grimsmo, David K. Tuckett, and Shruti Puri. Practical Quantum Error Correction with the XZZX Code and Kerr-Cat Qubits. *PRX Quantum*, 2(3):030345, September 2021.
- [42] Christopher M. Dawson and Michael A. Nielsen. The Solovay-Kitaev algorithm. *Quantum Info. Comput.*, 6(1):81–95, January 2006.
- [43] S. E. de Graaf, L. Faoro, L. B. Ioffe, S. Mahashabde, J. J. Burnett, T. Lindström, S. E. Kubatkin, A. V. Danilov, and A. Ya. Tzalenchuk. Two-level systems in superconducting quantum devices due to trapped quasiparticles. *Science Advances*, 6(51):eabc5055, December 2020.
- [44] C. L. Degen, F. Reinhard, and P. Cappellaro. Quantum sensing. *Reviews of Modern Physics*, 89(3):035002, July 2017.
- [45] Yannick Deller, Sebastian Schmitt, Maciej Lewenstein, Steve Lenk, Marika Federer, Fred Jendrzejewski, Philipp Hauke, and Valentin Kasper. Quantum approximate optimization algorithm for qudit systems. *Physical Review A*, 107(6):062410, June 2023.

- [46] The Qiskit Nature developers and contributors. Qiskit nature 0.6.0. Zenodo, April 2023.
- [47] P. a. M. Dirac. A new notation for quantum mechanics. *Mathematical Proceedings of the Cambridge Philosophical Society*, 35(3):416–418, July 1939.
- [48] P. Dita. Factorization of unitary matrices. *Journal of Physics A: Mathematical and General*, 36(11):2781, March 2003.
- [49] Yuqian Dong, Qiang Liu, Jianhua Wang, Qingshi Li, Xiaoyan Yu, Wen Zheng, Yong Li, Dong Lan, Xinsheng Tan, and Yang Yu. Simulation of Two-Qubit Gates with a Superconducting Qudit. *physica status solidi (b)*, 259(3):2100500, 2022.
- [50] N. Dupont, G. Chatelain, L. Gabardos, M. Arnal, J. Billy, B. Peaudecerf, D. Sugny, and D. Guéry-Odelin. Quantum State Control of a Bose-Einstein Condensate in an Optical Lattice. *PRX Quantum*, 2(4):040303, October 2021.
- [51] Austin G Fowler, Matteo Mariantoni, John M Martinis, and Andrew N Cleland. Surface codes: Towards practical large-scale quantum computation. *Physical Review A*, 86(3):032324, 2012.
- [52] Nicholas E. Frattini, Rodrigo G. Cortiñas, Jayameenakshi Venkatraman, Xu Xiao, Qile Su, Chan U. Lei, Benjamin J. Chapman, Vidul R. Joshi, S. M. Girvin, Robert J. Schoelkopf, Shruti Puri, and Michel H. Devoret. Observation of Pairwise Level Degeneracies and the Quantum Regime of the Arrhenius Law in a Double-Well Parametric Oscillator. *Physical Review X*, 14(3):031040, September 2024.
- [53] Ahmed Fawzy Gad. PyGAD: An intuitive genetic algorithm Python library. *Multimedia Tools and Applications*, 83(20):58029–58042, June 2024.
- [54] J. M. Gambetta, F. Motzoi, S. T. Merkel, and F. K. Wilhelm. Analytic control methods for high-fidelity unitary operations in a weakly nonlinear oscillator. *Physical Review A*, 83(1):012308, January 2011.
- [55] Yvonne Y. Gao, Brian J. Lester, Kevin S. Chou, Luigi Frunzio, Michel H. Devoret, Liang Jiang, S. M. Girvin, and Robert J. Schoelkopf. Entanglement of bosonic modes through an engineered exchange interaction. *Nature*, 566(7745):509–512, February 2019.
- [56] Ignacio García-Mata, Rodrigo G. Cortiñas, Xu Xiao, Jorge Chávez-Carlos, Victor S. Batista, Lea F. Santos, and Diego A. Wisniacki. Effective versus Floquet theory for the Kerr parametric oscillator. *Quantum*, 8:1298, March 2024.
- [57] I. M. Georgescu, S. Ashhab, and Franco Nori. Quantum simulation. *Reviews of Modern Physics*, 86(1):153–185, March 2014.
- [58] Craig Gidney. Stim: A fast stabilizer circuit simulator. *Quantum*, 5:497, July 2021.

- [59] Michael Goerz, Daniel Basilewitsch, Fernando Gago-Encinas, Matthias G. Krauss, Karl P. Horn, Daniel M. Reich, and Christiane Koch. Krotov: A Python implementation of Krotov’s method for quantum optimal control. *SciPost Physics*, 7(6):080, December 2019.
- [60] H. Goldstein, C.P. Poole, and J.L. Safko. *Classical Mechanics*. Addison Wesley, 2002.
- [61] Vittorio Gorini. Completely positive dynamical semigroups of N-level systems. *Journal of Mathematical Physics*, 17(5):821, 1976.
- [62] Noah Goss, Alexis Morvan, Brian Marinelli, Bradley K. Mitchell, Long B. Nguyen, Ravi K. Naik, Larry Chen, Christian Jünger, John Mark Kreikebaum, David I. Santiago, Joel J. Wallman, and Irfan Siddiqi. High-fidelity qutrit entangling gates for superconducting circuits. *Nature Communications*, 13(1):7481, December 2022.
- [63] Hayato Goto. Universal quantum computation with a nonlinear oscillator network. *Physical Review A*, 93(5):050301, May 2016.
- [64] Daniel Gottesman. The Heisenberg Representation of Quantum Computers, July 1998.
- [65] Daniel Gottesman. Fault-Tolerant Quantum Computation with Higher-Dimensional Systems. In Colin P. Williams, editor, *Quantum Computing and Quantum Communications*, Lecture Notes in Computer Science, pages 302–313, Berlin, Heidelberg, 1999. Springer.
- [66] Daniel Gottesman. *Surviving as a Quantum Computer in a Classical World*. 2024.
- [67] Daniel Gottesman and Isaac L. Chuang. Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations. *Nature*, 402(6760):390–393, November 1999.
- [68] Daniel Gottesman, Alexei Kitaev, and John Preskill. Encoding a qubit in an oscillator. *Physical Review A*, 64(1):012310, June 2001.
- [69] Luca Gravina, Fabrizio Minganti, and Vincenzo Savona. Critical Schrödinger Cat Qubit. *PRX Quantum*, 4(2):020337, June 2023.
- [70] Graham P. Greve, Chengyi Luo, Baochen Wu, and James K. Thompson. Entanglement-enhanced matter-wave interferometry in a high-finesse cavity. *Nature*, 610(7932):472–477, October 2022.
- [71] David J. Griffiths and Darrell F. Schroeter. *Introduction to Quantum Mechanics*. Cambridge University Press, Cambridge, 3 edition, 2018.
- [72] A. Grimm, N. E. Frattini, S. Puri, S. O. Mundhada, S. Touzard, M. Mirrahimi, S. M. Girvin, S. Shankar, and M. H. Devoret. Stabilization and operation of a Kerr-cat qubit. *Nature*, 584(7820):205–209, August 2020.

- [73] Harper R. Grimsley, Sophia E. Economou, Edwin Barnes, and Nicholas J. Mayhall. An adaptive variational algorithm for exact molecular simulations on a quantum computer. *Nature Communications*, 10(1):3007, July 2019.
- [74] Christian Gross and Immanuel Bloch. Quantum simulations with ultracold atoms in optical lattices. *Science*, 357(6355):995–1001, September 2017.
- [75] Stefanie Günther, N. Anders Petersson, and Jonathan L. DuBois. Quandary: An open-source C++ package for high-performance optimal control of open quantum systems. In *2021 IEEE/ACM Second International Workshop on Quantum Computing Software (QCS)*, pages 88–98. IEEE Computer Society, November 2021.
- [76] Ahmed Hajr, Bingcheng Qing, Ke Wang, Gerwin Koolstra, Zahra Pedramrazi, Ziqi Kang, Larry Chen, Long B. Nguyen, Christian Jünger, Noah Goss, Irwin Huang, Bibek Bhandari, Nicholas E. Frattini, Shruti Puri, Justin Dressel, Andrew N. Jordan, David I. Santiago, and Irfan Siddiqi. High-Coherence Kerr-Cat Qubit in 2D Architecture. *Physical Review X*, 14(4):041049, November 2024.
- [77] Douglas Rayner Hartree and W. Hartree. Self-consistent field, with exchange, for beryllium. *Proceedings of the Royal Society of London. Series A - Mathematical and Physical Sciences*, 150(869):9–33, January 1997.
- [78] Akel Hashim, Long B. Nguyen, Noah Goss, Brian Marinelli, Ravi K. Naik, Trevor Chistolini, Jordan Hines, J. P. Marceaux, Yosep Kim, Pranav Gokhale, Teague Tomesh, Senrui Chen, Liang Jiang, Samuele Ferracin, Kenneth Rudinger, Timothy Proctor, Kevin C. Young, Robin Blume-Kohout, and Irfan Siddiqi. A Practical Introduction to Benchmarking and Characterization of Quantum Computers. August 2024.
- [79] Erik Hostens, Jeroen Dehaene, and Bart De Moor. Stabilizer states and Clifford operations for systems of arbitrary dimensions and modular arithmetic. *Physical Review A*, 71(4):042315, April 2005.
- [80] Pavel Hřmó, Benjamin Wilhelm, Lukas Gerster, Martin W. van Mourik, Marcus Huber, Rainer Blatt, Philipp Schindler, Thomas Monz, and Martin Ringbauer. Native qudit entanglement in a trapped ion quantum processor. *Nature Communications*, 14(1):2242, April 2023.
- [81] Daisuke Iyama, Takahiko Kamiya, Shiori Fujii, Hiroto Mukai, Yu Zhou, Toshiaki Nagase, Akiyoshi Tomonaga, Rui Wang, Jiao-Jiao Xue, Shohei Watabe, Sangil Kwon, and Jaw-Shen Tsai. Observation and manipulation of quantum interference in a superconducting Kerr parametric oscillator. *Nature Communications*, 15(1):86, January 2024.
- [82] D. Jaksch, C. Bruder, J. I. Cirac, C. W. Gardiner, and P. Zoller. Cold Bosonic Atoms in Optical Lattices. *Physical Review Letters*, 81(15):3108–3111, October 1998.

- [83] Petar Jurcevic, Ali Javadi-Abhari, Lev S Bishop, Isaac Lauer, Daniela F Bogorin, Markus Brink, Lauren Capelluto, Oktay Günlük, Toshinari Itoko, Naoki Kanazawa, Abhinav Kandala, George A Keefe, Kevin Krsulich, William Landers, Eric P Lewandowski, Douglas T McClure, Giacomo Nannicini, Adinath Narasgond, Hasan M Nayfeh, Emily Pritchett, Mary Beth Rothwell, Srikanth Srinivasan, Neereja Sundaresan, Cindy Wang, Ken X Wei, Christopher J Wood, Jeng-Bang Yau, Eric J Zhang, Oliver E Dial, Jerry M Chow, and Jay M Gambetta. Demonstration of quantum volume 64 on a superconducting quantum computing system. *Quantum Science and Technology*, 6(2):025020, March 2021.
- [84] Taro Kanao, Shumpei Masuda, Shiro Kawabata, and Hayato Goto. Quantum Gate for a Kerr Nonlinear Parametric Oscillator Using Effective Excited States. *Physical Review Applied*, 18(1):014019, July 2022.
- [85] Abhinav Kandala, Kristan Temme, Antonio D. Córcoles, Antonio Mezzacapo, Jerry M. Chow, and Jay M. Gambetta. Error mitigation extends the computational reach of a noisy quantum processor. *Nature*, 567(7749):491–495, March 2019.
- [86] Eliot Kapit. Hardware-Efficient and Fully Autonomous Quantum Error Correction in Superconducting Circuits. *Physical Review Letters*, 116(15):150501, April 2016.
- [87] Mark Kasevich and Steven Chu. Atomic interferometry using stimulated Raman transitions. *Physical Review Letters*, 67(2):181–184, July 1991.
- [88] Raphael Kaubruegger, Athreya Shankar, Denis V. Vasilyev, and Peter Zoller. Optimal and Variational Multiparameter Quantum Metrology and Vector-Field Sensing. *PRX Quantum*, 4(2):020333, June 2023.
- [89] Raphael Kaubruegger, Pietro Silvi, Christian Kokail, Rick van Bijnen, Ana Maria Rey, Jun Ye, Adam M. Kaufman, and Peter Zoller. Variational Spin-Squeezing Algorithms on Programmable Quantum Sensors. *Physical Review Letters*, 123(26):260505, December 2019.
- [90] Harry C. P. Kendell, Giacomo Ferranti, and Carrie A. Weidner. Error Correcting States in Ultracold Atoms, December 2023.
- [91] E. M. Kessler, I. Lovchinsky, A. O. Sushkov, and M. D. Lukin. Quantum Error Correction for Metrology. *Physical Review Letters*, 112(15):150802, April 2014.
- [92] 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, February 2005.
- [93] Kaveh Khodjasteh and Daniel A. Lidar. Performance of deterministic dynamical decoupling schemes: Concatenated and periodic pulse sequences. *Physical Review A*, 75(6):062310, June 2007.

- [94] C. Kittel. *Introduction to Solid State Physics*. Wiley, 2004.
- [95] P. V. Klimov, J. Kelly, Z. Chen, M. Neeley, A. Megrant, B. Burkett, R. Barends, K. Arya, B. Chiaro, Yu Chen, A. Dunsworth, A. Fowler, B. Foxen, C. Gidney, M. Giustina, R. Graff, T. Huang, E. Jeffrey, Erik Lucero, J. Y. Mutus, O. Naaman, C. Neill, C. Quintana, P. Roushan, Daniel Sank, A. Vainsencher, J. Wenner, T. C. White, S. Boixo, R. Babbush, V. N. Smelyanskiy, H. Neven, and John M. Martinis. Fluctuations of Energy-Relaxation Times in Superconducting Qubits. *Physical Review Letters*, 121(9):090502, August 2018.
- [96] Jens Koch, Terri M. Yu, Jay Gambetta, A. A. Houck, D. I. Schuster, J. Majer, Alexandre Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf. Charge-insensitive qubit design derived from the Cooper pair box. *Physical Review A: Atomic, Molecular, and Optical Physics*, 76(4):042319, October 2007.
- [97] Jens Koch, Terri M. Yu, Jay Gambetta, A. A. Houck, D. I. Schuster, J. Majer, Alexandre Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf. Charge-insensitive qubit design derived from the Cooper pair box. *Physical Review A*, 76(4):042319, October 2007.
- [98] T. Kovachy, P. Asenbaum, C. Overstreet, C. A. Donnelly, S. M. Dickerson, A. Sugarbaker, J. M. Hogan, and M. A. Kasevich. Quantum superposition at the half-metre scale. *Nature*, 528(7583):530–533, December 2015.
- [99] P. Krantz, M. Kjaergaard, F. Yan, T. P. Orlando, S. Gustavsson, and W. D. Oliver. A quantum engineer’s guide to superconducting qubits. *Applied Physics Reviews*, 6(2):021318, June 2019.
- [100] T. D. Ladd, F. Jelezko, R. Laflamme, Y. Nakamura, C. Monroe, and J. L. O’Brien. Quantum computers. *Nature*, 464(7285):45–53, March 2010.
- [101] Catie LeDesma. *Programmable Bloch-Band Atom Interferometry in a Multidimensional Optical Lattice*. PhD thesis, University of Colorado at Boulder, 2024.
- [102] Catie LeDesma, Kendall Mehling, and Murray Holland. Vector Atom Accelerometry in an Optical Lattice. July 2024.
- [103] Catie LeDesma, Kendall Mehling, Jieqiu Shao, John Drew Wilson, Penina Axelrad, Marco Nicotra, Dana Z. Anderson, and Murray Holland. Demonstration of a programmable optical lattice atom interferometer. *Physical Review Research*, 6(4):043120, November 2024.
- [104] Catie LeDesma, Kendall Mehling, John Drew Wilson, Marco Nicotra, and Murray Holland. Universal gate set for optical lattice based atom interferometry. *Physical Review Research*, 7(1):013246, March 2025.

- [105] Nelson Leung, Mohamed Abdelhafez, Jens Koch, and David I. Schuster. Speedup for quantum optimal control from automatic differentiation based on graphics processing units. *Physical Review A*, 95(4):042318, April 2017.
- [106] Ziqian Li, Tanay Roy, David Rodríguez Pérez, David I. Schuster, and Eliot Kapit. Hardware-efficient autonomous error correction with linear couplers in superconducting circuits. *Physical Review Research*, 6(1):013171, February 2024.
- [107] Ziqian Li, Tanay Roy, David Rodríguez Pérez, Kan-Heng Lee, Eliot Kapit, and David I. Schuster. Autonomous error correction of a single logical qubit using two transmons. *Nature Communications*, 15(1):1681, February 2024.
- [108] G. Lindblad. On the generators of quantum dynamical semigroups. *Communications in Mathematical Physics*, 48(2):119–130, June 1976.
- [109] Andrew Litteken, Lennart Maximilian Seifert, Jason Chadwick, Natalia Nottingham, Frederic T. Chong, and Jonathan M. Baker. Qompress: Efficient Compilation for Ququarts Exploiting Partial and Mixed Radix Operations for Communication Reduction. In *Proceedings of the 28th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 2*, ASPLOS 2023, pages 646–659, New York, NY, USA, January 2023. Association for Computing Machinery.
- [110] Andrew Litteken, Lennart Maximilian Seifert, Jason D. Chadwick, Natalia Nottingham, Tanay Roy, Ziqian Li, David Schuster, Frederic T. Chong, and Jonathan M. Baker. Dancing the Quantum Waltz: Compiling Three-Qubit Gates on Four Level Architectures. In *Proceedings of the 50th Annual International Symposium on Computer Architecture*, ISCA '23, pages 1–14, New York, NY, USA, June 2023. Association for Computing Machinery.
- [111] Seth Lloyd and Reevu Maity. Efficient implementation of unitary transformations. *arXiv:1901.03431 [quant-ph]*, January 2019.
- [112] Kai Luo, Wenhui Huang, Ziyu Tao, Libo Zhang, Yuxuan Zhou, Ji Chu, Wuxin Liu, Biying Wang, Jiangyu Cui, Song Liu, Fei Yan, Man-Hong Yung, Yuanzhen Chen, Tongxing Yan, and Dapeng Yu. Experimental Realization of Two Qutrits Gate with Tunable Coupling in Superconducting Circuits. *Physical Review Letters*, 130(3):030603, January 2023.
- [113] Easwar Magesan, J. M. Gambetta, and Joseph Emerson. Scalable and Robust Randomized Benchmarking of Quantum Processes. *Physical Review Letters*, 106(18):180504, May 2011.
- [114] Easwar Magesan and Jay M. Gambetta. Effective Hamiltonian models of the cross-resonance gate. *Physical Review A: Atomic, Molecular, and Optical Physics*, 101(5):052308, May 2020.

- [115] Easwar Magesan, Jay M. Gambetta, B. R. Johnson, Colm A. Ryan, Jerry M. Chow, Seth T. Merkel, Marcus P. da Silva, George A. Keefe, Mary B. Rothwell, Thomas A. Ohki, Mark B. Ketchen, and M. Steffen. Efficient Measurement of Quantum Gate Error by Interleaved Randomized Benchmarking. *Physical Review Letters*, 109(8):080505, August 2012.
- [116] Ritajit Majumdar, Saikat Basu, Shibashis Ghosh, and Susmita Sur-Kolay. Quantum error-correcting code for ternary logic. *Physical Review A*, 97(5):052302, May 2018.
- [117] Christian D. Marciniak, Thomas Feldker, Ivan Pogorelov, Raphael Kaubruegger, Denis V. Vasilyev, Rick van Bijnen, Philipp Schindler, Peter Zoller, Rainer Blatt, and Thomas Monz. Optimal metrology with programmable quantum sensors. *Nature*, 603(7902):604–609, March 2022.
- [118] John M. Martinis, K. B. Cooper, R. McDermott, Matthias Steffen, Markus Ansmann, K. D. Osborn, K. Cicak, Seongshik Oh, D. P. Pappas, R. W. Simmonds, and Clare C. Yu. Decoherence in Josephson Qubits from Dielectric Loss. *Physical Review Letters*, 95(21):210503, November 2005.
- [119] David C. McKay, Christopher J. Wood, Sarah Sheldon, Jerry M. Chow, and Jay M. Gambetta. Efficient Z gates for quantum computing. *Physical Review A*, 96(2):022330, August 2017.
- [120] Kosuke Mitarai, Yasunari Suzuki, Wataru Mizukami, Yuya O. Nakagawa, and Keisuke Fujii. Quadratic Clifford expansion for efficient benchmarking and initialization of variational quantum algorithms. *Physical Review Research*, 4(3):033012, July 2022.
- [121] Avinash Mocherla, Lingling Lao, and Dan E. Browne. Extending Matchgate Simulation Methods to Universal Quantum Circuits, June 2024.
- [122] M. Mohseni, A. T. Rezakhani, and D. A. Lidar. Quantum-process tomography: Resource analysis of different strategies. *Physical Review A*, 77(3):032322, March 2008.
- [123] Cleve Moler and Charles Van Loan. Nineteen Dubious Ways to Compute the Exponential of a Matrix, Twenty-Five Years Later. *SIAM Review*, 45(1):3–49, January 2003.
- [124] C. Monroe, D. M. Meekhof, B. E. King, and D. J. Wineland. A “Schrödinger Cat” Superposition State of an Atom. *Science*, 272(5265):1131–1136, May 1996.
- [125] A. Morvan, V. V. Ramasesh, M. S. Blok, J. M. Kreikebaum, K. O’Brien, L. Chen, B. K. Mitchell, R. K. Naik, D. I. Santiago, and I. Siddiqi. Qutrit Randomized Benchmarking. *Physical Review Letters*, 126(21):210504, May 2021.
- [126] F. Motzoi, J. M. Gambetta, P. Rebentrost, and F. K. Wilhelm. Simple Pulses for Elimination of Leakage in Weakly Nonlinear Qubits. *Physical Review Letters*, 103(11):110501, September 2009.

- [127] Clemens Müller, Jared H Cole, and Jürgen Lisenfeld. Towards understanding two-level-systems in amorphous solids: Insights from quantum circuits. *Reports on Progress in Physics*, 82(12):124501, October 2019.
- [128] Holger Müller, Sheng-wei Chiow, Sven Herrmann, Steven Chu, and Keng-Yeow Chung. Atom-Interferometry Tests of the Isotropy of Post-Newtonian Gravity. *Physical Review Letters*, 100(3):031101, January 2008.
- [129] Prakash Murali, Jonathan M. Baker, Ali Javadi-Abhari, Frederic T. Chong, and Margaret Martonosi. Noise-Adaptive Compiler Mappings for Noisy Intermediate-Scale Quantum Computers. In *Proceedings of the Twenty-Fourth International Conference on Architectural Support for Programming Languages and Operating Systems*, ASPLOS '19, pages 1015–1029, New York, NY, USA, April 2019. Association for Computing Machinery.
- [130] Sreraman Muralidharan, Chang-Ling Zou, Linshu Li, Jianming Wen, and Liang Jiang. Overcoming erasure errors with multilevel systems. *New Journal of Physics*, 19(1):013026, January 2017.
- [131] John A Nelder and Roger Mead. A simplex method for function minimization. *The Computer Journal*, 7(4):308–313, 1965.
- [132] Duc Manh Nguyen and Sunghwan Kim. Quantum Key Distribution Protocol Based on Modified Generalization of Deutsch-Jozsa Algorithm in d-level Quantum System. *International Journal of Theoretical Physics*, 58(1):71–82, January 2019.
- [133] Long B. Nguyen, Noah Goss, Karthik Siva, Yosep Kim, Ed Younis, Bingcheng Qing, Akel Hashim, David I. Santiago, and Irfan Siddiqi. Empowering a qudit-based quantum processor by traversing the dual bosonic ladder. *Nature Communications*, 15(1):7117, August 2024.
- [134] Marco M. Nicotra, Jieqiu Shao, Joshua Combes, Anne Cross Theurkauf, Penina Axelrad, Liang-Ying Chih, Murray Holland, Alex A. Zozulya, Catie K. LeDesma, Kendall Mehling, and Dana Z. Anderson. Modeling and Control of Ultracold Atoms Trapped in an Optical Lattice: An Example-driven Tutorial on Quantum Control. *IEEE Control Systems Magazine*, 43(1):28–43, February 2023.
- [135] Frank Nielsen. Cramér-Rao Lower Bound and Information Geometry. In Rajendra Bhatia, C. S. Rajan, and Ajit Iqbal Singh, editors, *Connected at Infinity II: A Selection of Mathematics by Indians*, pages 18–37. Hindustan Book Agency, Gurgaon, 2013.
- [136] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, Cambridge, 2010.
- [137] Anastasiia S. Nikolaeva, Evgeniy O. Kiktenko, and Aleksey K. Fedorov. Efficient realization of quantum algorithms with qudits. *EPJ Quantum Technology*, 11(1):1–25, December 2024.

- [138] Nissim Ofek, Andrei Petrenko, Reinier Heeres, Philip Reinhold, Zaki Leghtas, Brian Vlastakis, Yehan Liu, Luigi Frunzio, S. M. Girvin, L. Jiang, Mazyar Mirrahimi, M. H. Devoret, and R. J. Schoelkopf. Extending the lifetime of a quantum bit with error correction in superconducting circuits. *Nature*, 536(7617):441–445, August 2016.
- [139] A. Omran, H. Levine, A. Keesling, G. Semeghini, T. T. Wang, S. Ebadi, H. Bernien, A. S. Zibrov, H. Pichler, S. Choi, J. Cui, M. Rossignolo, P. Rembold, S. Montangero, T. Calarco, M. Endres, M. Greiner, V. Vuletić, and M. D. Lukin. Generation and manipulation of Schrödinger cat states in Rydberg atom arrays. *Science*, 365(6453):570–574, August 2019.
- [140] A. Paetznick, M. P. da Silva, C. Ryan-Anderson, J. M. Bello-Rivas, J. P. Campora III, A. Chernoguzov, J. M. Dreiling, C. Foltz, F. Frachon, J. P. Gaebler, T. M. Gatterman, L. Grans-Samuelsson, D. Gresh, D. Hayes, N. Hewitt, C. Holliman, C. V. Horst, J. Johansen, D. Lucchetti, Y. Matsuoka, M. Mills, S. A. Moses, B. Neyenhuis, A. Paz, J. Pino, P. Siegfried, A. Sundaram, D. Tom, S. J. Wernli, M. Zanner, R. P. Stutz, and K. M. Svore. Demonstration of logical qubits and repeated error correction with better-than-physical error rates, November 2024.
- [141] N. Anders Petersson and Fortino Garcia. Optimal Control of Closed Quantum Systems via B-Splines with Carrier Waves. *SIAM Journal on Scientific Computing*, 44(6):A3592–A3616, December 2022.
- [142] L. P. Pitaevskiĭ and S. Stringari. *Bose-Einstein Condensation and Superfluidity*. Number 164 in International Series of Monographs on Physics. Oxford University Press, Oxford, United Kingdom, first edition edition, 2016.
- [143] Bibek Pokharel, Namit Anand, Benjamin Fortman, and Daniel A. Lidar. Demonstration of Fidelity Improvement Using Dynamical Decoupling with Superconducting Qubits. *Physical Review Letters*, 121(22):220502, November 2018.
- [144] Sierk Pötting, Marcus Cramer, and Pierre Meystre. Momentum-state engineering and control in Bose-Einstein condensates. *Physical Review A*, 64(6):063613, November 2001.
- [145] John Preskill. Quantum Computing in the NISQ era and beyond. *Quantum*, 2:79, August 2018.
- [146] Thomas Propson, Brian E. Jackson, Jens Koch, Zachary Manchester, and David I. Schuster. Robust Quantum Optimal Control with Trajectory Optimization. *Physical Review Applied*, 17(1):014036, January 2022.
- [147] Shruti Puri, Samuel Boutin, and Alexandre Blais. Engineering the quantum states of light in a Kerr-nonlinear resonator by two-photon driving. *npj Quantum Information*, 3(1):1–7, April 2017.

- [148] Shruti Puri, Alexander Grimm, Philippe Campagne-Ibarcq, Alec Eickbusch, Kyungjoo Noh, Gabrielle Roberts, Liang Jiang, Mazyar Mirrahimi, Michel H. Devoret, and S. M. Girvin. Stabilized Cat in a Driven Nonlinear Cavity: A Fault-Tolerant Error Syndrome Detector. *Physical Review X*, 9(4):041009, October 2019.
- [149] Shruti Puri, Lucas St-Jean, Jonathan A. Gross, Alexander Grimm, Nicholas E. Fratini, Pavithran S. Iyer, Anirudh Krishna, Steven Touzard, Liang Jiang, Alexandre Blais, Steven T. Flammia, and S. M. Girvin. Bias-preserving gates with stabilized cat qubits. *Science Advances*, 6(34):eaay5901, August 2020.
- [150] Harald Putterman, Kyungjoo Noh, Connor T. Hann, Gregory S. MacCabe, Shahriar Aghaeimeibodi, Rishi N. Patel, Menyoung Lee, William M. Jones, Hesam Moradinejad, Roberto Rodriguez, Neha Mahuli, Jefferson Rose, John Clai Owens, Harry Levine, Emma Rosenfeld, Philip Reinhold, Lorenzo Monceli, Joshua Ari Alcid, Nasser Ali-doust, Patricio Arrangoiz-Arriola, James Barnett, Przemyslaw Bienias, Hugh A. Carson, Cliff Chen, Li Chen, Harutiun Chinkejian, Eric M. Chisholm, Ming-Han Chou, Aashish Clerk, Andrew Clifford, R. Cosmic, Ana Valdes Curiel, Erik Davis, Laura DeLorenzo, J. Mitchell D’Ewart, Art Diky, Nathan D’Souza, Philipp T. Dumitrescu, Shmuel Eisenmann, Essam Elkhoully, Glen Evenbly, Michael T. Fang, Yawen Fang, Matthew J. Fling, Warren Fon, Gabriel Garcia, Alexey V. Gorshkov, Julia A. Grant, Mason J. Gray, Sebastian Grimberg, Arne L. Grimsmo, Arbel Haim, Justin Hand, Yuan He, Mike Hernandez, David Hover, Jimmy S. C. Hung, Matthew Hunt, Joe Iverson, Ignace Jarrige, Jean-Christophe Jaskula, Liang Jiang, Mahmoud Kalae, Rassul Karabalin, Peter J. Karalekas, Andrew J. Keller, Amirhossein Khalajhedayati, Aleksander Kubica, Hanho Lee, Catherine Leroux, Simon Lieu, Victor Ly, Keven Villegas Madrigal, Guillaume Marcaud, Gavin McCabe, Cody Miles, Ashley Milsted, Joaquin Minguzzi, Anurag Mishra, Biswaroop Mukherjee, Mahdi Naghiloo, Eric Oblepias, Gerson Ortuno, Jason Pagdilao, Nicola Pancotti, Ashley Panduro, J. P. Paquette, Minje Park, Gregory A. Pairs, David Perello, Eric C. Peterson, Sophia Ponte, John Preskill, Johnson Qiao, Gil Refael, Rachel Resnick, Alex Retzker, Omar A. Reyna, Marc Runyan, Colm A. Ryan, Abdulrahman Sahmoud, Ernesto Sanchez, Rohan Sanil, Krishanu Sankar, Yuki Sato, Thomas Scaffidi, Salome Siavoshi, Prasahnt Sivarajah, Trenton Skogland, Chun-Ju Su, Loren J. Swenson, Stephanie M. Teo, Astrid Tomada, Giacomo Torlai, E. Alex Wollack, Yufeng Ye, Jessica A. Zerrudo, Kailing Zhang, Fernando G. S. L. Brandão, Matthew H. Matheny, and Oskar Painter. Hardware-efficient quantum error correction via concatenated bosonic qubits. *Nature*, 638(8052):927–934, February 2025.
- [151] Q-CTRL. Boulder Opal, 2025.
- [152] Qiskit contributors. Qiskit: An Open-source Framework for Quantum Computing, 2023.
- [153] Gokul Subramanian Ravi, Pranav Gokhale, Yi Ding, William Kirby, Kaitlin Smith, Jonathan M. Baker, Peter J. Love, Henry Hoffmann, Kenneth R. Brown, and Fred-eric T. Chong. CAFQA: A Classical Simulation Bootstrap for Variational Quantum

- Algorithms. In *Proceedings of the 28th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 1*, ASPLOS 2023, pages 15–29, New York, NY, USA, December 2022. Association for Computing Machinery.
- [154] Gokul Subramanian Ravi, Kaitlin N. Smith, Pranav Gokhale, Andrea Mari, Nathan Earnest, Ali Javadi-Abhari, and Frederic T. Chong. VAQEM: A Variational Approach to Quantum Error Mitigation. In *2022 IEEE International Symposium on High-Performance Computer Architecture (HPCA)*, pages 288–303, April 2022.
 - [155] Chad Rigetti and Michel Devoret. Fully microwave-tunable universal gates in superconducting qubits with linear couplings and fixed transition frequencies. *Physical Review B*, 81(13):134507, April 2010.
 - [156] Martin Ringbauer, Michael Meth, Lukas Postler, Roman Stricker, Rainer Blatt, Philipp Schindler, and Thomas Monz. A universal qudit quantum processor with trapped ions. *Nature Physics*, 18(9):1053–1057, September 2022.
 - [157] Joschka Roffe. Quantum error correction: An introductory guide. *Contemporary Physics*, 60(3):226–245, July 2019.
 - [158] Neil J. Ross and Peter Selinger. Optimal ancilla-free Clifford+T approximation of z-rotations. *Quantum Info. Comput.*, 16(11-12):901–953, September 2016.
 - [159] Jared Rovny, Zhiyang Yuan, Mattias Fitzpatrick, Ahmed I. Abdalla, Laura Futamura, Carter Fox, Matthew Carl Cambria, Shimon Kolkowitz, and Nathalie P. de Leon. Nanoscale covariance magnetometry with diamond quantum sensors. *Science*, 378(6626):1301–1305, December 2022.
 - [160] Tanay Roy, Ziqian Li, Eliot Kapit, and David I. Schuster. Two-Qutrit Quantum Algorithms on a Programmable Superconducting Processor. *Physical Review Applied*, 19(6):064024, June 2023.
 - [161] Diego Ruiz, Ronan Gautier, Jérémie Guillaud, and Mazyar Mirrahimi. Two-photon driven Kerr quantum oscillator with multiple spectral degeneracies. *Physical Review A*, 107(4):042407, April 2023.
 - [162] Frederic Sauvage, Sukin Sim, Alexander A. Kunitsa, William A. Simon, Marta Mauri, and Alejandro Perdomo-Ortiz. FLIP: A flexible initializer for arbitrarily-sized parametrized quantum circuits, May 2021.
 - [163] J. A. Schreier, A. A. Houck, Jens Koch, D. I. Schuster, B. R. Johnson, J. M. Chow, J. M. Gambetta, J. Majer, L. Frunzio, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf. Suppressing charge noise decoherence in superconducting charge qubits. *Physical Review B*, 77(18):180502, May 2008.

- [164] Jacob T. Seeley, Martin J. Richard, and Peter J. Love. The Bravyi-Kitaev transformation for quantum computation of electronic structure. *The Journal of Chemical Physics*, 137(22):224109, December 2012.
- [165] Peter Selinger. Generators and relations for n-qubit Clifford operators. *Logical Methods in Computer Science*, Volume 11, Issue 2, June 2015.
- [166] Simone Serafini, Luca Galantucci, Elena Iseni, Tom Bienaimé, Russell N. Bisset, Carlo F. Barenghi, Franco Dalfovo, Giacomo Lamporesi, and Gabriele Ferrari. Vortex Reconnections and Rebounds in Trapped Atomic Bose-Einstein Condensates. *Physical Review X*, 7(2):021031, May 2017.
- [167] R. Shankar. *Principles of Quantum Mechanics*. Springer US, New York, NY, 1994.
- [168] Jieqiu Shao, Liang-Ying Chih, Mantas Naris, Murray Holland, and Marco M. Nicotra. Application of Quantum Optimal Control to Shaken Lattice Interferometry. In *2023 American Control Conference (ACC)*, pages 4593–4598, May 2023.
- [169] Sarah Sheldon, Easwar Magesan, Jerry M. Chow, and Jay M. Gambetta. Procedure for systematically tuning up cross-talk in the cross-resonance gate. *Physical Review A*, 93(6):060302, June 2016.
- [170] Kevin Singh, Shraddha Anand, Andrew Pocklington, Jordan T. Kemp, and Hannes Bernien. Dual-Element, Two-Dimensional Atom Array with Continuous-Mode Operation. *Physical Review X*, 12(1):011040, March 2022.
- [171] Chao Song, Kai Xu, Hekang Li, Yu-Ran Zhang, Xu Zhang, Wuxin Liu, Qiujiang Guo, Zhen Wang, Wenhui Ren, Jie Hao, Hui Feng, Heng Fan, Dongning Zheng, Da-Wei Wang, H. Wang, and Shi-Yao Zhu. Generation of multicomponent atomic Schrödinger cat states of up to 20 qubits. *Science*, 365(6453):574–577, August 2019.
- [172] Anders Sørensen and Klaus Mølmer. Quantum Computation with Ions in Thermal Motion. *Physical Review Letters*, 82(9):1971–1974, March 1999.
- [173] J.C. Spall. Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. *IEEE Transactions on Automatic Control*, 37(3):332–341, March 1992.
- [174] Armands Strikis, Dayue Qin, Yanzhu Chen, Simon C. Benjamin, and Ying Li. Learning-Based Quantum Error Mitigation. *PRX Quantum*, 2(4):040330, November 2021.
- [175] Qiming Sun, Timothy C. Berkelbach, Nick S. Blunt, George H. Booth, Sheng Guo, Zhendong Li, Junzi Liu, James D. McClain, Elvira R. Sayfutyarova, Sandeep Sharma, Sebastian Wouters, and Garnet Kin-Lic Chan. PySCF: The Python-based simulations of chemistry framework. *WIREs Computational Molecular Science*, 8(1):e1340, 2018.

- [176] Stuart S. Szigeti, Onur Hosten, and Simon A. Haine. Improving cold-atom sensors with quantum entanglement: Prospects and challenges. *Applied Physics Letters*, 118(14):140501, April 2021.
- [177] Swamit S. Tannu and Moinuddin K. Qureshi. Mitigating Measurement Errors in Quantum Computers by Exploiting State-Dependent Bias. In *Proceedings of the 52nd Annual IEEE/ACM International Symposium on Microarchitecture*, MICRO '52, pages 279–290, New York, NY, USA, October 2019. Association for Computing Machinery.
- [178] Kristan Temme, Sergey Bravyi, and Jay M. Gambetta. Error Mitigation for Short-Depth Quantum Circuits. *Physical Review Letters*, 119(18):180509, November 2017.
- [179] Jules Tilly, Hongxiang Chen, Shuxiang Cao, Dario Picozzi, Kanav Setia, Ying Li, Edward Grant, Leonard Wossnig, Ivan Rungger, George H. Booth, and Jonathan Tenynson. The Variational Quantum Eigensolver: A review of methods and best practices. *Physics Reports*, 986:1–128, November 2022.
- [180] J Tolar. On Clifford groups in quantum computing. *Journal of Physics: Conference Series*, 1071:012022, August 2018.
- [181] A. Trimeche, B. Battelier, D. Becker, A. Bertoldi, P. Bouyer, C. Braxmaier, E. Charron, R. Corgier, M. Cornelius, K. Douch, N. Gaaloul, S. Herrmann, J. Müller, E. Rasel, C. Schubert, H. Wu, and F. Pereira dos Santos. Concept study and preliminary design of a cold atom interferometer for space gravity gradiometry. *Classical and Quantum Gravity*, 36(21):215004, October 2019.
- [182] Aaron Trowbridge, Aditya Bhardwaj, Kevin He, David I. Schuster, and Zachary Manchester. Direct Collocation for Quantum Optimal Control. In *2023 IEEE International Conference on Quantum Computing and Engineering (QCE)*, volume 01, pages 1278–1285, September 2023.
- [183] E Urban, T A Johnson, T Henage, L Isenhower, D D Yavuz, T G Walker, and M Saffman. Observation of Rydberg blockade between two atoms. *NATURE PHYSICS*, 5, 2009.
- [184] Victor Veitch, S A Hamed Mousavian, Daniel Gottesman, and Joseph Emerson. The resource theory of stabilizer quantum computation. *New Journal of Physics*, 16(1):013009, January 2014.
- [185] Jayameenakshi Venkatraman, Rodrigo G. Cortiñas, Nicholas E. Frattini, Xu Xiao, and Michel H. Devoret. A driven Kerr oscillator with two-fold degeneracies for qubit protection. *Proceedings of the National Academy of Sciences*, 121(24):e2311241121, June 2024.
- [186] Jayameenakshi Venkatraman, Xu Xiao, Rodrigo G. Cortiñas, and Michel H. Devoret. Nonlinear dissipation in a driven superconducting circuit. *Physical Review A*, 110(4):042411, October 2024.

- [187] Lorenza Viola, Emanuel Knill, and Seth Lloyd. Dynamical decoupling of open quantum systems. *Physical Review Letters*, 82(12):2417–2421, March 1999.
- [188] Andreas Wächter and Lorenz T. Biegler. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical Programming*, 106(1):25–57, March 2006.
- [189] C. Wang, C. Axline, Y. Y. Gao, T. Brecht, Y. Chu, L. Frunzio, M. H. Devoret, and R. J. Schoelkopf. Surface participation and dielectric loss in superconducting qubits. *Applied Physics Letters*, 107(16):162601, October 2015.
- [190] Chen Wang, Yvonne Y. Gao, Philip Reinhold, R. W. Heeres, Nissim Ofek, Kevin Chou, Christopher Axline, Matthew Reagor, Jacob Blumoff, K. M. Sliwa, L. Frunzio, S. M. Girvin, Liang Jiang, M. Mirrahimi, M. H. Devoret, and R. J. Schoelkopf. A Schrödinger cat living in two boxes. *Science*, 352(6289):1087–1091, May 2016.
- [191] Hanrui Wang, Yongshan Ding, Jiaqi Gu, Yujun Lin, David Z. Pan, Frederic T. Chong, and Song Han. QuantumNAS: Noise-Adaptive Search for Robust Quantum Circuits. In *2022 IEEE International Symposium on High-Performance Computer Architecture (HPCA)*, pages 692–708, April 2022.
- [192] Samson Wang, Piotr Czarnik, Andrew Arrasmith, M. Cerezo, Lukasz Cincio, and Patrick J. Coles. Can Error Mitigation Improve Trainability of Noisy Variational Quantum Algorithms? *Quantum*, 8:1287, March 2024.
- [193] Yuchen Wang, Zixuan Hu, Barry C. Sanders, and Sabre Kais. Qudits and high-dimensional quantum computing. *Frontiers in Physics*, 8:589504, November 2020.
- [194] Matthew Ware, Blake R. Johnson, Jay M. Gambetta, Thomas A. Ohki, Jerry M. Chow, and B. L. T. Plourde. Cross-resonance interactions between superconducting qubits with variable detuning. *arXiv:1905.11480 [quant-ph]*, May 2019.
- [195] C. A. Weidner and D. Z. Anderson. Simplified landscapes for optimization of shaken lattice interferometry. *New Journal of Physics*, 20(7):075007, July 2018.
- [196] C. A. Weidner and Dana Z. Anderson. Experimental Demonstration of Shaken-Lattice Interferometry. *Physical Review Letters*, 120(26):263201, June 2018.
- [197] C. A. Weidner, Hoon Yu, Ronnie Kosloff, and Dana Z. Anderson. Atom interferometry using a shaken optical lattice. *Physical Review A*, 95(4):043624, April 2017.
- [198] M. Werninghaus, D. J. Egger, F. Roy, S. Machnes, F. K. Wilhelm, and S. Filipp. Leakage reduction in fast superconducting qubit gates via optimal control. *npj Quantum Information*, 7(1):14, January 2021.
- [199] J. Werschnik and E. K. U. Gross. Quantum optimal control theory. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 40(18):R175, September 2007.

- [200] John Drew Wilson, Jarrod T. Reilly, Haoqing Zhang, Chengyi Luo, Anjun Chu, James K. Thompson, Ana Maria Rey, and Murray J. Holland. Entangled matter waves for quantum enhanced sensing. *Physical Review A*, 110(4):L041301, October 2024.
- [201] Karen Wintersperger, Florian Dommert, Thomas Ehmer, Andrey Hoursanov, Johannes Klepsch, Wolfgang Mauerer, Georg Reuber, Thomas Strohm, Ming Yin, and Sebastian Luber. Neutral atom quantum computing hardware: Performance and end-user perspective. *EPJ Quantum Technology*, 10(1):1–26, December 2023.
- [202] Re-Bing Wu, Bing Chu, David H. Owens, and Herschel Rabitz. Data-driven gradient algorithm for high-precision quantum control. *Physical Review A*, 97(4):042122, April 2018.
- [203] Xuejian Wu, Zachary Pagel, Bola S. Malek, Timothy H. Nguyen, Fei Zi, Daniel S. Scheirer, and Holger Müller. Gravity surveys using a mobile atom interferometer. *Science Advances*, 5(9):eaax0800, September 2019.
- [204] Jonathan Wurtz and Peter J. Love. Classically Optimal Variational Quantum Algorithms. *IEEE Transactions on Quantum Engineering*, 2:1–7, 2021.
- [205] Qian Xu, Joseph K. Iverson, Fernando G. S. L. Brandão, and Liang Jiang. Engineering fast bias-preserving gates on stabilized cat qubits. *Physical Review Research*, 4(1):013082, February 2022.
- [206] Qian Xu, Guo Zheng, Yu-Xin Wang, Peter Zoller, Aashish A. Clerk, and Liang Jiang. Autonomous quantum error correction and fault-tolerant quantum computation with squeezed cat qubits. *npj Quantum Information*, 9(1):78, August 2023.
- [207] Victoria Xu, Matt Jaffe, Cristian D. Panda, Sofus L. Kristensen, Logan W. Clark, and Holger Müller. Probing gravity by holding atoms for 20 seconds. *Science*, 366(6466):745–749, November 2019.
- [208] Lia Yeh and John van de Wetering. Constructing All Qutrit Controlled Clifford+T gates in Clifford+T. In Claudio Antares Mezzina and Krzysztof Podlaski, editors, *Reversible Computation*, Lecture Notes in Computer Science, pages 28–50, Cham, 2022. Springer International Publishing.
- [209] Xinyuan You, Ziwen Huang, Ugur Alyanak, Alexander Romanenko, Anna Grassellino, and Shaojiang Zhu. Stabilizing and Improving Qubit Coherence by Engineering the Noise Spectrum of Two-Level Systems. *Physical Review Applied*, 18(4):044026, October 2022.
- [210] Sisi Zhou, Mengzhen Zhang, John Preskill, and Liang Jiang. Achieving the Heisenberg limit in quantum metrology using quantum error correction. *Nature Communications*, 9(1):78, January 2018.

- [211] Ido Zuk, Daniel Cohen, Alexey V. Gorshkov, and Alex Retzker. Robust gates with spin-locked superconducting qubits. *Physical Review Research*, 6(1):013217, February 2024.
- [212] Juan C. Zuñiga Castro, Jeffrey Larson, Sri Hari Krishna Narayanan, Victor E. Colussi, Michael A. Perlin, and Robert J. Lewis-Swan. Variational quantum state preparation for quantum-enhanced metrology in noisy systems. *Physical Review A*, 110(5):052615, November 2024.