

Error Detection in Bosonic Circuit Quantum Electrodynamics

James D. Teoh

Abstract

Error Detection in Bosonic Circuit Quantum Electrodynamics

James David Teoh

2023

Any useful quantum computer must be robust to errors on the individual qubits. This means a quantum algorithm can still reliably give the correct answer despite the qubits suffering from decoherence. Quantum error correction provides a means for achieving this, by encoding logical qubits redundantly in many additional physical qubits and/or additional energy levels. In the presence of errors the logical information is still preserved, but this remains a monumentally difficult task to achieve in practice. But what can we do with quantum error *detection*? Rather than trying to recover from a quantum error we instead reset qubits that are known to have suffered errors. If we know exactly when and which qubit suffered an error, this becomes an erasure, which is significantly easier to correct than Pauli errors. For erasure qubits to be viable, we must be able to detect errors that happen not just while idling as for a quantum memory, but during every operation needed to run a quantum computer. In this thesis we describe a way to realize the necessary error-detected operations in the qubit platform of bosonic circuit quantum electrodynamics. We choose to encode our qubits in 3D superconducting microwave cavity modes, and introduce the dual-rail cavity qubit, in which the dominant hardware errors can all be detected as erasure errors. With this new approach, practical quantum error correction is expected to be within reach for currently available experimental hardware and coherence times. Finally, we also investigate how to use error detection to circumvent loss errors in links that connect qubit modules together, experimentally realizing high-fidelity state transfer between modules.

Error Detection in Bosonic Circuit Quantum Electrodynamics

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

by
James David Teoh

Dissertation Director: Professor Robert J. Schoelkopf

December 2023

Copyright © 2023 by James David Teoh

All rights reserved.

Contents

1	Introduction	1
1.1	Overview of the Thesis	6
2	Qubits and Their Errors	9
2.1	Qubits and Their Errors	9
2.1.1	What is a qubit?	9
2.1.2	Pauli error channels	12
2.1.3	T_1 energy relaxation	13
2.1.4	T_ϕ dephasing and T_2	17
2.1.5	The Lindblad master equation	18
2.1.6	The fidelity metric	21
2.2	Gates, Measurements and Computation	22
2.2.1	Single qubit gates	23
2.2.2	Entangling gates	25
2.2.3	Three-qubit gates	28
2.2.4	Excitation preserving gates	29
2.2.5	The Clifford group	30
2.2.6	Projective measurements	31
2.2.7	End-of-the-line vs mid-circuit measurements	31
2.2.8	Changing the measurement basis	33
2.3	Entanglement	36

2.3.1	Correlations in entangled states	36
2.3.2	Measurements with an ancilla qubit	38
2.4	Quantum Teleportation	39
2.4.1	The protocol	39
2.4.2	Use in quantum communication	41
2.4.3	Gate teleportation	44
2.4.4	Measurement-free gate teleportation	46
3	Quantum Error Correction and Detection	47
3.1	Quantum Error Correction Fundamentals	47
3.1.1	Knill-Laflamme conditions	48
3.1.2	The 3-qubit repetition code	49
3.1.3	Fault tolerance	52
3.1.4	Correcting for Pauli errors	53
3.1.5	Erasure errors	54
3.1.6	Leakage errors	57
3.1.7	Erasure conversion	58
3.2	The Surface Code	62
3.2.1	Fault tolerance and the threshold theorem	66
3.2.2	Error hierarchy and structured noise	67
3.2.3	Erasure errors in the surface code	68
3.2.4	Error detection enables the erasure error hierarchy	71
3.3	Bosonic Quantum Error Correction	74
3.3.1	Quantum control of bosonic modes	75
3.3.2	Propagation of ancilla errors	76
3.3.3	The photon loss channel for a bosonic mode	77
3.3.4	Dephasing in a bosonic mode	80
3.3.5	Cat codes	81

3.3.6	The kitten code	83
3.3.7	The dual-rail code	84
3.3.8	The 0-2 code and biased erasures	85
3.3.9	The 0-1 code	87
3.3.10	Wigner functions	88
3.4	Hardware Fault-Tolerance	92
3.4.1	Path independence	93
3.4.2	Error transparency	98
3.4.3	Extension to error closure	98
4	The Circuit Quantum Electrodynamics Toolkit	101
4.0.1	The cQED toolkit: a quick reference	103
4.1	Microwave Cavities	108
4.1.1	3D cavities	110
4.1.2	Stripline resonators	111
4.1.3	The natural noise-bias of microwave cavities	113
4.2	Transmons	114
4.2.1	The transmon regime	115
4.2.2	Transmon decoherence	118
4.2.3	Controlling transmons	119
4.3	The Dispersive Coupling	121
4.3.1	χ -matching	124
4.4	Transmon Readout	125
4.5	Cavity control	129
4.5.1	Displacement drives	129
4.5.2	Universal control via the dispersive interaction	130
4.5.3	Photon-number preserving unitaries	132
4.5.4	Frequency combs for parity measurements	133

4.5.5	SNAP gates	134
4.5.6	Optimal control theory pulses	136
4.6	Controllable Beamsplitter Interactions	139
4.6.1	Realizing a parametric beamsplitter	139
4.6.2	Evolution from a parametric beamsplitter interaction	145
5	Error Detected Gates	148
5.1	Searching for an error-correctable two-qubit gate	149
5.1.1	Two-mode SNAP gates	150
5.1.2	Is eSWAP Hardware Fault-Tolerant?	153
5.1.3	Constructing eSWAP as an exponential gadget	154
5.1.4	Measurements are easier to error-correct?	156
5.1.5	Errors in photon number parity measurement	158
5.1.6	The Hardware Fault-tolerant parity measurement	161
5.1.7	Z measurements	163
5.2	The Hardware Fault-Tolerant ZZ measurement	164
5.2.1	Adding a parametric beamsplitter interaction	165
5.2.2	Why not just SWAP?	168
5.2.3	Evolution under a detuned beamsplitter interaction	169
5.2.4	The Operator Bloch Sphere	174
5.2.5	A QND ZZ measurement and its Hardware Fault-Tolerance	178
5.2.6	How to error-correct a ZZ measurement (is it worth it?)	181
5.3	The Error-Detected ZZ(θ) Gate	186
5.3.1	Why is the exponentiation gadget only error-detectable?	187
5.3.2	Checking it works: master equation simulations	188
5.3.3	Error scaling	194
5.3.4	Kerr and χ'	195
5.3.5	Realistic pulse shaping	199

5.4	Pole-to-Pole: other useful trajectories on the Operator Bloch Sphere .	203
5.4.1	A family of excitation-preserving two-qubit gates	206
5.5	Conclusions	207
6	The Dual-rail Code with Superconducting Cavities	209
6.1	Error-detection Enables Erasure Qubits	209
6.2	Engineering the Error Hierarchy	211
6.3	Dual-rail in Microwave Cavities	213
6.4	Expected Idling Error Hierarchy	214
6.4.1	No-jump backaction	217
6.5	Operating One Dual-rail Cavity Qubit	218
6.5.1	State preparation	218
6.5.2	Logical measurement with in-built erasure detection	220
6.5.3	Simulating logical measurement	223
6.5.4	Single qubit gates	226
6.5.5	Mid-circuit erasure detection	228
6.6	Entangling Gates	232
6.6.1	With in-built erasure detection	234
6.7	Operational Error Hierarchy	237
6.8	Scaling Up	239
6.8.1	The physical qubit unit cell	241
6.8.2	Comparison to transmon-based approaches	242
6.8.3	Higher-order nonlinearities and cross-talk	243
6.8.4	4-qubit erasure code	244
6.9	The Dual-rail Glossary	247
6.10	The Dual-rail FAQ	252
6.11	Conclusions	263

7	Error Detection for Quantum Networks	266
7.1	Regimes of Quantum Communication	270
7.2	Examples of Deterministic State Transfer Schemes	274
7.2.1	Pitch and catch	274
7.2.2	The quantum bus	277
7.2.3	Adiabatic schemes	279
7.3	The Crossover from Discrete to Continuum	284
7.4	Heralded Schemes	286
7.5	Review of Previous Experiments	288
7.6	Qubit Transfer in the ‘Impossible’ Regime	290
7.6.1	Classical bright and dark states	292
7.6.2	The Dark Mode Measurement scheme	293
7.6.3	The DMM scheme is robust to bus loss	295
7.6.4	Fundamental loss limits to the DMM scheme	296
7.7	Making the Scheme Experimentally-friendly	301
7.8	cQED Hardware for the DMM Scheme	305
7.8.1	Purcell filter for the bus mode	307
7.8.2	Detailed cQED hardware	310
7.8.3	Tuning up simultaneous parametric beamsplitters	312
7.8.4	Pump phase-locking	314
7.8.5	Measuring loss in the bus mode	315
7.8.6	Kerr-cancellation	317
7.9	Observing Bright and Dark State Dynamics	317
7.9.1	High-fidelity cat state preparation	321
7.9.2	Procedure for entanglement generation	323
7.9.3	System reset	323
7.10	Wigner Functions and Parity Measurements	324

7.10.1	Measuring Wigner functions	324
7.10.2	Frequency combs for parity measurements	325
7.10.3	Fine-tuning pump parameters with joint parity	326
7.11	Joint-cavity Tomography	326
7.11.1	Joint Wigner functions	328
7.11.2	Symmetrizing Joint Wigner functions	329
7.11.3	OCT decode + a measured Wigner function	329
7.12	Extracting Bell State Fidelity	331
7.12.1	Error budget	337
7.12.2	Bell state fidelities for different values of α	339
7.13	State Transfer via Teleportation	342
7.13.1	Bell measurements in a module	342
7.13.2	Average teleported state transfer fidelity	345
7.14	Conclusions	351
8	Outlook	355
	Appendices	359
A	Correcting for erasure errors in the 4-qubit erasure code	360
B	Magic values of α in the 4-legged cat code	364

List of Figures

1.1	Error-detection vs error correction	5
1.2	Diagram of results	8

2.1	The Bloch Sphere	10
2.2	Qubit Measurements	32
2.3	CNOT gate via measurements	34
2.4	Bell measurement circuits	35
2.5	Destructive Bell measurement circuits	35
2.6	QND measurements via an ancilla qubit	39
2.7	Gate teleportation	45
3.1	Surface code qubit layout	63
3.2	XXXX stabilizer circuit	64
3.3	ZZZZ stabilizer circuit	64
3.4	XZZX stabilizer circuit	64
3.5	Surface code scaling with erasures and Pauli errors	70
3.6	Resilience of the surface code to different error classes	72
3.7	Wigner functions for the Kitten code and two-legged cat code	90
3.8	Wigner functions for the 4-legged cat code	91
3.9	Path independence for an error-corrected gate	95
3.10	Path independence example	96
3.11	Error paths for an error-detectable gate	97
4.1	cQED hardware for 3D microwave cavities	107
4.2	LC lumped element circuit model	108
4.3	Energy levels of a Quantum Harmonic Oscillator	109
4.4	Types of microwave cavities	113
4.5	Circuit diagram for a transmon	115
4.6	Energy level diagram of a transmon	118
4.7	Driving g-f transitions in a transmon	121
4.8	Dispersive transmon readout circuit	125

4.9	Quasi-2D transmon and readout	129
4.10	3D stub cavity coupled to a transmon	131
4.11	Cavity-transmon-readout circuit	131
4.12	Frequency combs for measuring photon number parity	134
4.13	Example OCT pulses	138
5.1	Y-mon cQED Hardware	151
5.2	Gate construction for eSWAP	154
5.3	Gate construction for a photon number parity measurement	158
5.4	Gate construction for a hardware fault-tolerant parity measurement	162
5.5	cZZ gates via instant SWAP gates	169
5.6	cQED hardware for cZZ unitaries	171
5.7	Trajectories on the operator Bloch sphere	176
5.8	Conditional trajectories on the operator Bloch sphere	177
5.9	cZZ trajectories on the operator Bloch sphere	179
5.10	ZZ measurement gate construction	180
5.11	ZZ measurements for even and odd photon number parities	182
5.12	Measurement sequences for an error-corrected QND ZZ measurement	183
5.13	Gate construction for realizing an error-detected $ZZ(\theta)$ gate	186
5.14	Simulated $ZZ(\theta)$ performance	193
5.15	Simulated $ZZ(\theta)$ gate performance with higher order nonlinearities	197
5.16	OCT pulse shaping for the $ZZ(\theta)$ gate	200
5.17	Control-SWAP and unconditional-SWAP trajectories	204
5.18	Trajectories for an unconditional SWAP	206
6.1	cQED hardware for one dual-rail cavity qubit	215
6.2	Control manifold of a dual-rail cavity qubit	219
6.3	Logical measurement scheme for a dual-rail cavity qubit	221

6.4	Simulated logical measurements for a dual-rail cavity qubit	226
6.5	Joint-parity measurement circuit	229
6.6	cQED hardware for $ZZ(\theta)$ in dual-rail cavity qubits	233
6.7	Gate construction for $ZZ(\theta)$ gates in a dual-rail cavity qubit	233
6.8	Photon number dynamics during a cZZ unitary	235
6.9	Gate construction for $ZZ(\theta)$ gates that detects photon loss	236
6.10	Graphical depiction of the operational error hierarchy	238
6.11	Possible dual-rail surface code layout	240
6.12	Potential cQED hardware for realizing a 4-qubit erasure code.	245
7.1	Discrete and continuous quantum communication schemes	275
7.2	Population dynamics during on-resonance state transfer	280
7.3	Population dynamics for adiabatic state transfer	283
7.4	Simulating the crossover from a discrete mode to a continuum of modes	285
7.5	Setup for heralding entanglement in an optical quantum network	287
7.6	Classical pendulum analog for three coupled bosonic modes	293
7.7	The Dark Mode Measurement (DMM) scheme	296
7.8	The DMM scheme in different bus loss regimes	297
7.9	The effect coupling to off-resonant harmonics in the DMM scheme	300
7.10	Modifying the bosonic code basis for small α	303
7.11	Minimalist layout of the cQED hardware needed for the DMM scheme	306
7.12	Design of a 3D Purcell filter for the bus mode	309
7.13	Detailed cQED hardware for the DMM scheme	310
7.14	Pump calibration to account for stark shift	313
7.15	Wiring diagram for generating pairs of phase-locked pumps	315
7.16	Measuring κ_b in-situ.	316
7.17	Pulse sequence for characterizing bright and dark states	318
7.18	Measured dynamics of bright and dark states	319

7.19	Correcting for pulse distortions during cat state preparation	322
7.20	Pulse sequence for entanglement generation	323
7.21	Improving parity measurements with combs of selective π -pulses . . .	327
7.22	2D slices of the Joint Wigner function	330
7.23	Measured Wigner functions of cavity 1 in the DMM scheme	333
7.24	Reconstructed Pauli bars for $\alpha = 1.414$	335
7.25	Measured Bell state fidelities and success probabilities vs. α	339
7.26	Measured Wigner functions and Pauli bars for different values of α .	341
7.27	Circuit for quantum teleportation from transmon 2 to cavity 1. . . .	343
7.28	Bell measurements between transmon 2 and cavity 2	343
7.29	Measured Wigner functions after teleportation, Kerr cancellation . . .	349
7.30	Measured Wigner functions after teleportation, no Kerr cancellation	350

List of Tables

3.1	Correcting erasures	60
4.1	Hierarchy of Hamiltonian rates for control of bosonic modes in cQED	145
5.1	χ -matching conditions for two-mode SNAP	153
5.2	3 round majority voting with a QND measurement	157
5.3	Error closure for various Hamiltonians for single photon loss errors . .	168
5.4	Operating points for HFT \mathbf{cZZ}_L	174
6.1	The expected idling error hierarchy for a dual-rail cavity qubit.	216
6.2	Measurement decoding for dual-rail logical measurements	222
6.3	Collapse operators for logical measurement simulations	225

6.4	Operational error hierarchy for dual-rail cavity qubits.	237
6.5	Estimated coherence times for cavities and transmons	239
7.1	Comparing entanglement experiments in superconducting networks	289
7.2	Experimental parameters used for the DMM scheme	311
7.3	DMM scheme performance for different values of α	340
7.4	Teleportation fidelities for each cardinal state	347
7.5	Optimized logical basis for cavity 1 for state teleportation	348

Acknowledgements

There are many people without whom this thesis would not have come to exist. I am fortunate to have spent the previous years with amazing friends and colleagues both on the 4th floor of Becton and at YQI, both inside and outside of academic pursuits. I apologize now for any names I may have missed.

First and foremost, I must thank my advisor ROB SCHOELKOPF, who has mentored me every step of the way and helped me grow as a researcher over the past six years. He has given me the freedom and encouragement to explore many new ideas off the beaten path, and can near-instantly recognize when we are onto something. He has the impressive ability to cut through to the heart of the physics in almost any problem, and will more often than not extend discussions until we eventually reach some satisfying conclusion. I am especially grateful for his insistence on the necessity of clear science communication, which often proves just as valuable as the rest of the scientific journey. Although my discussions with MICHEL DEVORET have been much briefer, Michel never fails to make me stop and think with the challenging questions he asks, and I am often left pondering his deep insights long after our discussion has ended. I thank both Rob and Michel for their part in making the 4th floor of Becton a place where science and scientists can truly thrive.

On the theory side, SHRUTI PURI has been essential to the development of many ideas in this thesis. She really tests the mettle of our proposals with a trial by fire and has helped shape our proposals from the musings of experimentalists to theoretically sound manifestos. Shruti has taken the time and patience to help navigate me (and Rob!) through the minefield of quantum error correction theory and her sharp criticism and suggestions have been vital in guiding our efforts. STEVE GIRVIN has always impressed me with his approachability and ease at which he can bridge the gap between experimental and theory land. He is a master at explaining things in the simplest of terms while drawing from his formidable breadth and depth of knowledge. He also has my express gratitude for closely examining every sentence in this weighty tome! I also thank MICHAEL HATRIDGE for taking the time to act as my external reader (and for the occasional good chat about the future of remote entanglement.)

LUIGI FRUNZIO is one of the friendliest people I have met. Always eager to help, he will often interrupt whatever he is doing to assist you without a moment's hesitation. Time with Luigi is always time spent in good company, I am thankful for his advice and stories about lab and life in general, and for helping me through the often tedious and daunting thesis writing process. Now, onto the students (roughly in chronological order) with whom I've had the fortune of working with.

I start by thanking ANDREW LINGENFELTER, who first introduced me to the weird and wonderful world of microwave electronics and helped orient me in the lab. Back then as a talented undergrad, he took the time to teach me the basics when I knew next to nothing. I am eternally grateful to LUKE BURKHART, who took me under his wing when I was in my second year. He has taught me most everything I know about being an experimentalist and I could not have asked for a better mentor. From him I have learnt the meticulousness, the planning, the patience, and the creativity needed to tease out the beautiful data from our often finicky devices.

The infectious enthusiasm and ambition of NAT COTTET led to a lot of excitement in our corner of the lab, even in the face of a looming pandemic. His determination and genuine love of quantum physics has helped me appreciate the experiments we worked on. His ability to adapt experiments on the fly to solve problems as they arise has greatly influenced me. I worked briefly with LEV KRAYZMAN on suspended resonators and miss his dry humor and his insistence on doing things properly. During the latter half of my Ph.D., I have had the pleasure of working with two amazing postdocs – PATRICK WINKEL and TAKAHIRO TSUNODA, as well as the great opportunity to mentor two graduate students – NEEL THAKUR and HARSH BABLA. Without your help and support, I sometimes wonder if I would have ever made it to the end!

Patrick hopped aboard the dual-rail train when it was still in its infancy. His intense desire to understand the physics and all its fine details was instrumental in bringing this idea to the realms of experimental feasibility. I must also thank him for usually being the first one to hear my half-baked, sometimes fantastical ideas at any hour of the day, and being able to instantly respond with the most pertinent questions. His rigor and ability to juggle many projects at once is also something to be admired. Takahiro has proved himself as an invaluable collaborator on the error detected gates project. His persistence in searching and finding beautiful mathematical frameworks and visualizations is something I value immensely. Mentoring Neel and Harsh has been both fun and challenging. I have been impressed with Neel's ability to absorb complicated ideas and then immediately put them into practice when running the experiments. His world-class banana bread and lemon bars will indeed be sorely missed too. Working out the seemingly never-ending details with the 4-cat project with Harsh has been an invigorating experience. His rigor and determination to make sure our understanding is water-tight has been the main driving force of this work which, ultimately, was too big to fit in this thesis.

It goes without saying that I owe a lot to the two ‘Beamsplitter Crews’. Their excellent results and rigorous understanding have laid the groundwork that brings the bulk of the work in this thesis into the realm of experimental possibility. I worked briefly alongside the ‘SNAIL crew’ with BEN CHAPMAN, STIJN DE GRAAF and SOPHIA XUE. Ben’s leadership and perseverance has often times left me in awe. His kindness and approachability has helped foster a sense of camaraderie with everyone he works with. It has been a lot of fun working with Stijn and Sophia trying to realize these proposals. I am also grateful for many interesting discussions with the ‘SQUID crew’ — YAO LU, ANIKET MAITI and JOHN GARMON whose generous expertise has helped shape the finer details of my work and keeps me grounded to the constraints and surprising capabilities of our devices. Across the street at YQI, I must give my many thanks to JAHAN CLAES and KAAVYA SAHAY, who have been patient in answering my near-endless string of questions regarding surface codes and error correction. They are adept explainers and teachers and helped me out many times when I was stuck in a maze of confusing theory. Although we never worked directly on any projects together, I must mention NICK FRATTINI. Our discussions lamenting how transmon errors seemed to plague every experiment in the lab ultimately inspired much of the work in this thesis.

There are many more friends and colleagues that have made the past six years such a special experience. My many thanks goes to ALEC EICKBUSCH, ALEX READ, AKSHAY KOOTTANDAVIDA, ANDY DING, BEN BROCK, BILLY KALFUS, CHRIS WANG, CHRIS AXLINE, JACOB CURTIS, JAYA VENKATRAMAN, KEVIN CHOU, LIANG JIANG, SAL ELDER, SPENCER DIAMOND, SUHAS GANJAM, YANNI DAHMANI, and many other names I have missed. Things would not run nearly as smoothly without the work of ALEX BOZZI, GISELLE DEVITO, MARIA RAO, TERRI EVANGELISTE, NUCH GRAVES at Becton and FLORIAN CARLE and RACQUEL MILLER at YQI. They have helped this place feel like home.

My time in New Haven would have been ill spent were it not for the good company of Ako, April, Elad, Grant, Kimmy, Mary, Olivier, Sheridan, Tal, Varun, Xinyi and Yasemin. Likewise, my double life in Princeton was made all the merrier with Aaron, Cassidy, Christian, Fedor, Grace, Himanshu, Jaime, James, Katie, Katja, Peter, Sarah-Marie, Steven and Wayne. Cassidy was taken from us too soon and her kindness is dearly missed.

Finally, I wish to thank Mum and Dad, to whom this thesis is dedicated. Thanks for encouraging me every step of the way and fostering my love of science. I must also thank my brother, Andy and my pet hamster, Snoofie, and Steve and Jenny, for supporting me from afar. Lastly, I wish to thank my girlfriend, Diana for her continuous love and support throughout both of our Ph.D. journeys. I feel very lucky to have someone as wonderful as you in my life.

Six years is a surprisingly long time and I am very grateful to everyone who has been a part of this ordeal. Thank you all!!!

Introduction

Over the past four decades the field of quantum computing has grown from a theoretical curiosity [[Feynman, 1982](#)] into one of the juggernauts of modern research, spanning across multiple fields of physics, engineering, and computer science.

Why try and build a quantum computer? We know enough about quantum mechanics to simulate any quantum system on regular ‘classical’ computers after all, but quantum computers can make use of quantum interference to solve certain niche problems dramatically faster than any classical computer ever could. This is best exemplified with Shor’s algorithm [[Shor, 1994](#)], which efficiently solves the prime factorization problem on a hypothetical quantum computer. It is an incredibly important discovery, because it is an existence proof that quantum computers can solve problems which:

- (a) Have nothing to do with quantum mechanics.
- (b) Have a provable, dramatic speedup over the best known classical algorithms.
- (c) Have real-world consequences — widely used encryption protocols such as RSA would no longer be computationally secure.

To get an idea for how dramatic the speedup can be, cracking 2048-bit RSA on a classical computer will take an estimated 300 trillion years. For a quantum computer with on order $\sim 10,000$ logical qubits this can take just a single day [Gidney and Ekerå, 2021]. The search for more useful algorithms is itself a very active field of research which now ponders uses in quantum chemistry, material science, machine learning and many more. Naturally, a lot of these potential applications revolve around the simulation of quantum systems.

What we will be concerned with is *how* to make a quantum computer. We want this quantum computer to be *universal*, meaning it can run any possible quantum algorithm (the quantum equivalent of a universal Turing machine), and we need it to be *fault-tolerant*, meaning that it must reliably produce the correct answer, even if the underlying components are faulty and prone to errors. The difficulty is not just making and controlling a large number of qubits but that the qubits themselves have high error rates. Correcting these errors is notoriously hard, and quantum error correction is arguably the largest unsolved challenge in quantum computing. Why is this the case?

Unlike a transistor, which can perform many trillions of logic gates without suffering from any errors, qubits are much more fragile. It is very easy for unwanted interactions with a qubit's surrounding environment to decohere its quantum superposition. The kind of qubits we focus on in this thesis are superconducting qubits, specifically ones made from microwave resonators cooled down to ~ 10 mK — all in an effort to shield them from noise in the environment. Typically, our qubit states only

last for a millisecond at most before they suffer from an error, limiting us to ~ 1000 gates on a qubit. The other difficulty is that qubits can suffer from two types of error — bit flips and phase flips. (There is another, called leakage which will be very important later.) Bit flips are the same as for a classical bit. A 0 changes to a 1 or vice versa without us knowing. Quantum bits can also suffer from phase flips as well, which disturbs the superposition. We also are not allowed by the laws of quantum mechanics to make copies of the qubit states for extra redundancy by the ‘no-cloning’ theorem [Wootters and Zurek, 1982]. Finally, correcting for qubit errors is difficult because we cannot directly measure the qubits without collapsing the superposition.

Despite these added difficulties, quantum error correction works in theory [Shor, 1995]. It has been shown that once the error-per-operation of the qubits gets below a particular value, known as the threshold [Aharonov and Ben-Or, 1997], then in principle we can reduce the error rate of a logical qubit as much as we want, by using more and more ‘physical’ qubits to make up each logical qubit. This trade-off means we could tolerate, say 0.1% error rates in each of our physical qubits, but now would now need 1,000 physical qubits per logical qubit! That’s ~ 10 million qubits to implement Shor’s algorithm!

The march towards more and more qubits has been steady over the years. It is important to remember that we must maintain qubit error rates significantly below the threshold while increasing the number of qubits in a device. This year IBM announced a chip with 433 transmon qubits and Rydberg atom systems can now routinely reach ~ 100 -qubit devices [Singh et al., 2022]. However, no equivalent of Moore’s law has been shown for qubits and miniaturization is far from guaranteed like it was for transistors. To make more powerful processors, we could definitely benefit from a modular approach where we network smaller qubit processors together.

The central aim of this thesis is to apply the paradigm of error *detection* to superconducting qubits, specifically a variation known as bosonic circuit quantum

electrodynamics in which we encode the qubits in harmonic oscillators. We will show how this can make the task of large-scale error correction easier and we do this in two main ways. Firstly, by using error detection to realize a type of qubit called an erasure qubit, quantum error correction becomes easier. The most common errors are no longer bit flips or phase flips but a different error known as an erasure — an unknown error at a known qubit and at a known time. Practically, this means we can now tolerate $\sim 1\%$ error rates, which are now within reach for the devices we have today. Secondly, will be by using error detection to make communication between modules better, even when the inter-module links are very lossy. At some point we will hit a limit to the number of qubits we can reliably make in one module. Then, we must join modules together if we want more qubits. The links tend to be the most error-prone part of the entire system but with error-detection, we will show how we can efficiently circumvent these errors.

The new paradigm of error detection stems from the previous paradigm of ‘hardware-efficient error correction.’ Previously, the idea was to correct for the dominant qubit errors as they happened, without the need for additional physical qubits. Then, when we scale up, we already start with physical qubit error rates that are orders of magnitude below the threshold, giving us more insurance that we stay below the threshold as we add more physical qubits, as well as reducing the number of physical qubits needed per logical qubit.

The difference between correcting and detecting errors in physical qubits is illustrated by Fig. 1.1. Hardware efficient error-correction requires building in redundancy without adding more components. This is most often achieved by using additional energy levels in a harmonic oscillator to encode the qubit, as well as additional levels in any ancillary systems used to control the qubit. In addition, some operations need to be repeated several times until they are successful.

The paradigm of error detection argues that hardware-efficient error correction

is often more trouble than it is worth, and in practical implementations the benefit is not justified by the increase in hardware complexity. Instead of detecting and then correcting hardware errors to fully recover the qubit, we stop at the detection step and then let our control software know the qubit had an error. We make no effort to fix the error and instead reinitialize the qubit in a fresh new state. Error-detection still requires some redundancy to be built into our hardware but it is far more lenient. Moreover, we can achieve the same benefits promised by hardware-efficient error correction – large-scale error correction still becomes easier in the same way.

We will find this seemingly innocuous change in perspective has far reaching consequences in how we design our qubits, their gates and network them together; in many cases greatly simplifying the requirements for their experimental realization.

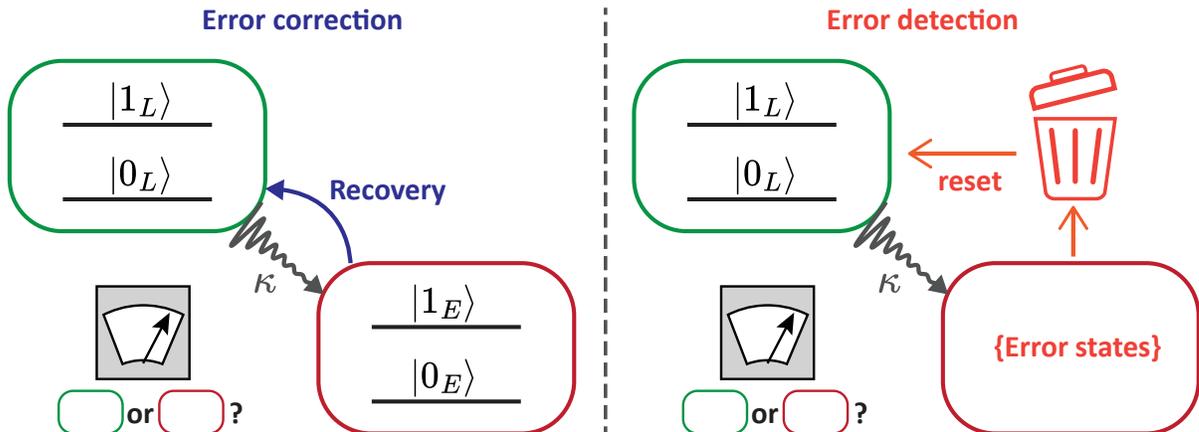


Figure 1.1: The subtle difference between error correction and error detection for a qubit. In both cases, we initially begin in the computational space (green) and suffer an error which takes us to the error space (red) and perform a measurement that tells us whether we are in the computational space or error space. **(Left)** For error correction, the qubit information is still preserved within the error space, and a recovery operation takes us back to the original qubit state – it is as though the error never occurred! **(Right)** Error detection is subtly different. Upon detecting we are in the error space, we make no attempt to recover the qubit information. Instead we elect to reset the qubit back into the computational space in a new state. Error detection is always easier than error correction, and this has far-reaching consequences throughout this thesis.

1.1 Overview of the Thesis

Chapters 2-4 contain mostly an in-depth review on error correction and bosonic circuit-QED. New results are presented in chapters 5-7 and summarized in Fig. 1.2.

Chapter 2 begins with a review of qubits in general and their error processes. [Nielsen and Chuang, 2010] is an indispensable reference. After discussing how errors are modelled and simulated we then detail the set of gates and measurements needed to run a quantum computer. Briefly we touch on the concept of entanglement and quantum teleportation which will be highly relevant for chapter 7.

Chapter 3 introduces methods for quantum error correction beginning with the 3-qubit repetition code which elucidates most features of a quantum error-correcting code. The concept of erasures and leakage errors are also explained and we show how error detection enables leakage errors to be converted to erasures. Two kinds of practical error-correcting codes are then reviewed. First is the surface code which is currently the most promising large-scale error correcting code and is adept at correcting for erasures. Second are bosonic codes, specifically the ones designed to detect/correct for photon loss. We end this chapter with formal discussion of hardware fault-tolerance, which ensures errors during a gate are still correctable/detectable after the gate. This is summarized by the ‘error-closure’ condition, a new result obtained in [Tsunoda et al., 2022] and builds on the work of [Rosenblum et al., 2018b] and [Ma et al., 2020].

Chapter 4 is a first look at the physical hardware we will be using for our qubits. 3D microwave cavities, transmons, coaxial cable links and most recently beamsplitter couplers make up these building blocks. The chapter begins with a quick reference section for their general properties such as their Hamiltonians, typical couplings and coherences. We devote some time to formalize conventions for the beamsplitter interaction which will be important in the proceeding chapters and describe some cavity control techniques which are relevant for Chapter 7.

Chapter 5 details the work in [Tsunoda et al., 2022], on which I am co-first author. We build on the work of [Rosenblum et al., 2018b] and [Reinhold et al., 2020] to extend hardware fault-tolerant control of bosonic qubits to two-qubit entangling operations. We show proposals for a new error-corrected ZZ measurement and an error-detected $ZZ(\theta)$ gate for a variety of bosonic codes. These gates form the bedrock of important operations in chapter 6. Crucially, the dominant hardware errors can be detected after these operations. The remaining errors, which go undetected and hence contribute to the overall error rate, are predicted to approach $\sim 0.01\%$.

Chapter 6 is the dual-rail chapter and presents the work from [Teoh et al., 2022]. Inspired by the proposal [Wu et al., 2022] in Rydberg atoms, we show how to realize erasure qubits with new hardware that we call dual-rail cavity qubits. As well as showing how to detect errors in idle dual-rail qubits, we give recipes for how to implement and error-detect every required gate and measurement. This proposal joins a recent wave of proposals for realizing erasure qubits in many different possible qubit platforms [Kubica et al., 2022, Scholl et al., 2023]. Of particular use may be Sec. 6.9, the ‘dual-rail glossary’ which defines many new terms we have introduced to describe and quantify the performance of dual-rail cavity qubits. Sec. 6.10 is the ‘Frequently Asked Questions’ that aims to answer some common questions I have encountered (or anticipate) regarding this work.

Chapter 7 switches gears to quantum networks for modular quantum computing. First is a review of quantum communication scheme in microwave links. Many different schemes have been proposed and demonstrated over the years. Here, we describe the physics of communication through a microwave channel in a unified way and reveal what dictates the fundamental limits to the fidelity of quantum state transfer. The rest of the chapter focuses on the work in [Teoh et al., 2023], the main experimental result of this thesis. Here we show how we can use interference effects unique to microwave links to teleport states between modules with high fidelity, even when the

links themselves are very lossy. We coin this the ‘Dark Mode Measurement’ scheme.

Lastly, in chapter 8 I give my outlook on the future of the new work presented in chapters 5-7 and what I see as the main challenges and opportunities facing error-detected qubits in general.

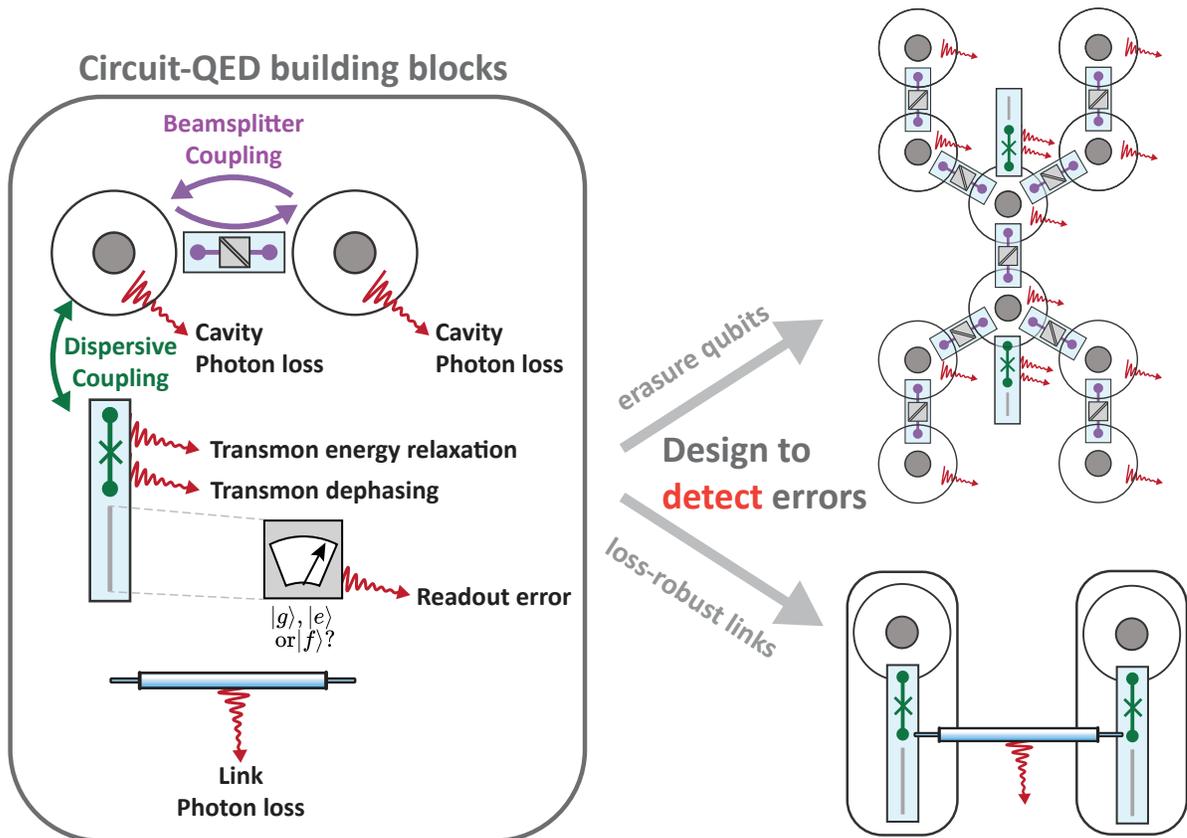


Figure 1.2: The new results of this thesis presented in chapters 5-7 take the well-established (and recently improved) building blocks of circuit-Quantum Electrodynamics and combines them in new ways that detect dominant hardware errors. In chapters 5 and 6 we explore how to realize erasure qubits, where error-detection can be performed in all required operations for quantum computing. In chapter 7 we see how error-detection can be used to circumvent link photon loss when networking quantum modules together, by far the dominant error in this system.

Qubits and Their Errors

2.1 Qubits and Their Errors

2.1.1 What is a qubit?

A qubit is the fundamental building block of quantum information and a quantum computer. They comprise a superposition of two orthogonal quantum states, which we denote as $|0\rangle$ and $|1\rangle$. For an isolated qubit state, its wave function can be written as

$$|\psi\rangle = c_0 |0\rangle + c_1 |1\rangle, \quad (2.1)$$

where c_0 and c_1 are complex numbers that satisfy $|c_0|^2 + |c_1|^2 = 1$ to ensure $|\psi\rangle$ is normalized. It is helpful to visualize these qubit states on the surface of the Bloch Sphere as shown in Fig. [2.1](#).

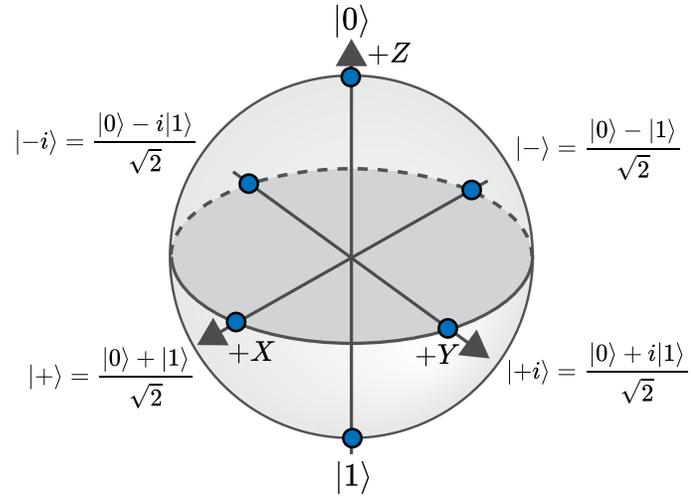


Figure 2.1: The Bloch sphere for a general qubit. Blue circles indicate the six cardinal states $\{|0\rangle, |1\rangle, |+\rangle, |-\rangle, |+i\rangle, |-i\rangle\}$

In the Bloch sphere representation, we can see a clear geometric mapping between the position of the state on the Bloch sphere if we rewrite our qubit state as

$$|\psi\rangle = \cos\left(\frac{\theta}{2}\right) |0\rangle + \sin\left(\frac{\theta}{2}\right) e^{i\phi} |1\rangle \quad (2.2)$$

θ is the polar angle and ϕ is the azimuthal angle. Perhaps confusingly, orthogonal qubit states are antipodal on the Bloch sphere!

Isolated qubit states are ‘pure’ and are confined to the surface of the Bloch sphere. When a qubit interacts with uncontrolled degrees of freedom in its environment, this results in the phenomenon of decoherence, the central challenge to realizing a useful quantum computer.

The composite state consisting of the qubit subspace and the environment subspace remains a pure state, but we represent our lack of knowledge concerning the environment’s state by tracing over the environment subspace and writing our qubit state as a mixed state. In the density matrix formalism, the state of the qubit can then be

written more generally as

$$\rho_L = \sum_i p_i |\psi_i\rangle \langle \psi_i| \quad (2.3)$$

where p_i represents the classical probability of being in the pure state $|\psi_i\rangle$, parametrized by θ_i and ϕ_i in Eq. 2.1. All such mixed states can be represented as (single) points in the interior of the Bloch sphere.

We can write down the density matrix for a general qubit state as

$$\rho_{\text{init}} = \begin{pmatrix} \rho_{00} & \rho_{01} \\ \rho_{10} & \rho_{11} \end{pmatrix} \quad (2.4)$$

Where ρ_{00} and ρ_{11} are called the ‘populations’ and the off diagonal terms, which satisfy $\rho_{10} = \rho_{01}^*$ are called the ‘coherences’. For a completely mixed state, the coherences are zero.

At the center of the Bloch Sphere lies the maximally mixed state,

$$\rho_{\text{mixed}} = \frac{1}{2} |0\rangle \langle 0| + \frac{1}{2} |1\rangle \langle 1| \quad (2.5)$$

Noise processes in the environment that uniformly take qubits to this state are called ‘depolarizing noise’. This is often a good description of what would happen if one were to attempt a long-depth algorithm on noisy qubits. Even though the physical error processes are not depolarizing noise, after performing many gates on our qubits, the overall effect of these errors is well-approximated by depolarizing noise.

2.1.2 Pauli error channels

Physical error processes that take our pure qubit states to mixed ones are mathematically described by linear maps known as error channels. For an initially pure qubit state $\rho_{\text{init}} = |\psi_L\rangle\langle\psi_L|$, this is written in the Kraus operator formalism as

$$\begin{aligned}\rho_{\text{final}} &= \mathcal{E}(\rho_{\text{init}}) \\ &= \sum_i \hat{E}_i \rho_{\text{init}} \hat{E}_i^\dagger\end{aligned}\tag{2.6}$$

where \hat{E}_i are the so-called Kraus operators that must satisfy the following completeness relation to conserve probability.

$$\sum_i \hat{E}_i^\dagger \hat{E}_i = \mathbf{1},\tag{2.7}$$

A Pauli channel is a special case where

$$\begin{aligned}\rho_{\text{final}} &= \mathcal{E}_{\text{pauli}}(\rho_{\text{init}}) \\ &= (1 - p_x - p_y - p_z)\rho_{\text{init}} + p_x X \rho_{\text{init}} X + p_y Y \rho_{\text{init}} Y + p_z Z \rho_{\text{init}} Z,\end{aligned}\tag{2.8}$$

where p_x, p_y, p_z are probabilities and X, Y, Z are Pauli operators acting on the qubit space. Explicitly, the Kraus operators for this channel can be written as

$$\begin{aligned}\hat{E}_0 &= \sqrt{1 - p_x - p_y - p_z} \mathbf{1} \\ \hat{E}_1 &= \sqrt{p_x} X \\ \hat{E}_2 &= \sqrt{p_y} Y \\ \hat{E}_3 &= \sqrt{p_z} Z\end{aligned}\tag{2.9}$$

Specific decoherence mechanisms such as energy relaxation and dephasing (See chapter 8 of [Nielsen and Chuang, 2010]) are well-approximated by a Pauli error

channel under what is known as the Pauli Twirling Approximation (PTA) [Geller and Zhou, 2013]. For the case $p_x = p_y = p_z = p$, the error channel describes ‘depolarization noise’. The ability to accurately model physical qubit errors as Pauli errors or more frequently, just depolarizing noise is a key assumption made when analyzing the performance of quantum error correcting codes, often because it is easy to simulate, and one we will challenge extensively in Chapter 6. Depolarization error models are often assumed because the action of many gates on the qubits transforms X -type errors into Z -type errors and vice-versa, quickly destroying any bias or structure in the noise. It has been shown that by designing hardware-specific gates that are bias-preserving, the task of quantum error correction is significantly easier compared to the case of simple depolarizing noise. [Puri et al., 2020, Guillaud and Mirrahimi, 2019]

Depolarizing noise is modelled by using Monte-Carlo simulations to draw Pauli error operators uniformly with probability p from the set of Pauli operators $\{X, Y, Z\}$ and applying them to the qubit state. This can be extended to errors on two qubits by using the set $\{\mathbb{1}, X, Y, Z\}^{\otimes 2}$ (and excluding $\mathbb{1}\mathbb{1}$ which is not an error) and so on and so forth for ensembles of N qubits.

We will now briefly review the two physical error mechanisms that cause the majority of errors in qubits.

2.1.3 T_1 energy relaxation

Energy relaxation refers to the fact that the two computational levels $|0\rangle$ and $|1\rangle$ are non-degenerate in energy with $|1\rangle$ often denoting the state with higher energy and $|0\rangle$ often the ground state of the system. We denote the difference between these energy levels as $\hbar\omega_{01}$. When left to thermalize with its environment, the equilibrium state of

the qubit will be the thermal state

$$\rho_{\text{therm}} = (1 - n_{\text{therm}}) |0\rangle \langle 0| + n_{\text{therm}} |1\rangle \langle 1| \quad (2.10)$$

Where $n_{\text{therm}} = \left(e^{\frac{\hbar\omega_{01}}{k_B T_{\text{eff}}}} + 1 \right)^{-1}$ for effective qubit temperature, T_{eff} . In the limit of zero temperature, the equilibrium state is the ground state and any other qubit states will decay towards the ground state via the amplitude damping channel.

We can define this phenomenological error channel, via the Kraus operator $\hat{E}_1 = \sqrt{p} |0\rangle \langle 1|$ which models decay through spontaneous emission of a photon, from $|1\rangle$ to $|0\rangle$ with probability p within some given time frame. The abrupt, inherently random ‘jump’ evolution is usually the picture we have in our heads when we think about quantum errors. However, from Eq. 2.7, we need another Kraus operator to form a complete Kraus map:

$$\begin{aligned} \hat{E}_0 &= |0\rangle \langle 0| + \sqrt{1-p} |1\rangle \langle 1| \\ \hat{E}_1 &= \sqrt{p} |0\rangle \langle 1| \end{aligned} \quad (2.11)$$

We can also write this channel in the basis of Pauli operators where it takes the form

$$\begin{aligned} \hat{E}_0 &= \frac{1 + \sqrt{1-p}}{2} \mathbb{1} + \frac{1 - \sqrt{1-p}}{2} Z \\ \hat{E}_1 &= \sqrt{p} \left(\frac{X - iY}{2} \right) \end{aligned} \quad (2.12)$$

which are linear combinations of Pauli operators. The physical interpretation of \hat{E}_1 and \hat{E}_0 tells us something interesting, that most error channels also include a continuous evolution that causes decoherence, called the *no-jump backaction*. Suppose we start in the superposition state $|+\rangle$ such that

$$\rho_{\text{init}} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad (2.13)$$

Note that the probability of the jump error occurring is included in the definition of the Kraus operators. The interpretation of \hat{E}_1 is that the jump error $|0\rangle\langle 1|$ occurs with probability $p\langle 1|\rho_{\text{init}}|1\rangle$. The probability depends both on p and the probability we are in the $|1\rangle$ state to begin with.

The other Kraus operator, \hat{E}_0 is *not* identity and polarizes our state towards the $|0\rangle$ state. This can be understood as a Bayesian update that occurs because the environment is *weakly measuring* the state of the qubit. If no jump errors have happened, we know we are more likely to have begun in the state $|0\rangle$. This polarizes any quantum superposition towards $|0\rangle$ and so we sometimes refer to this error as a *polarization* error. It is a non-unitary evolution that results in decoherence, even though no emission of energy occurs! Since \hat{E}_0 it can be written as a linear combination of the identity operator and a Z operator, this is a dephasing-type error channel.

Starting from $|+\rangle$, no-jump backaction alone leaves us in the state

$$|\psi_{\text{no-jump}}\rangle = \frac{1}{\sqrt{2}} \frac{|0\rangle + \sqrt{1-p}|1\rangle}{\sqrt{1-\frac{p}{2}}} \quad (2.14)$$

Where now we have renormalized the state, corresponding to the scenario where we ‘post-select’ out all jump errors. The channel that describes post-selected no-jump backaction cannot be described as a linear map due to this renormalization. It can be written as

$$\mathcal{E}_{\text{nl}}(\rho) = \frac{\hat{E}_0 \rho \hat{E}_0^\dagger}{\|(\hat{E}_0 \rho \hat{E}_0^\dagger)\|} \quad (2.15)$$

where $\|\rho\|$ is the norm of ρ . We can calculate the square of the overlap (the fidelity) of this state with the initial state to find the error associated with no-jump backaction

$$|\langle +|\psi_{\text{no-jump}}\rangle|^2 = 1 - \frac{p^2}{16} + O(p^4)\dots \quad (2.16)$$

As we can see, no-jump backaction is quite a small error for small p , and takes a

quadratic scaling with a $1/16$ prefactor. This is good news for us! Most of the time we will instead be focusing on how to correct/detect the jump errors instead. As we move away from the equator and further towards the poles of the Bloch sphere, no-jump backaction becomes an increasingly smaller source of infidelity (the prefactor decreases but the quadratic scaling is maintained). For the pole states, no-jump backaction vanishes completely. As p becomes large i.e., for longer times, no-jump backaction is no longer negligible. A trick that completely nullifies no-jump backaction at long times is to apply unitaries to the qubit that ensure every initial state spends equal amounts of time in the $|0\rangle$ and $|1\rangle$ state. This can be achieved for example by periodically applying X gates. No-jump backaction does not only occur in two level systems, but in any qubit where the decay rates of each codeword are different.

Energy decay is often an exponential process, and the rate of spontaneous decay can be denoted as Γ_{\downarrow} . Similarly, the rate of heating from $|0\rangle$ to $|1\rangle$ can be denoted as Γ_{\uparrow} and is relevant for realistic systems at non-zero temperature. We can also rewrite the equilibrium thermal population in terms of these rates as

$$n_{\text{therm}} = \frac{\Gamma_{\uparrow}}{\Gamma_{\downarrow} + \Gamma_{\uparrow}} = e^{-\frac{\hbar\omega_{01}}{k_B T_{\text{eff}}}}. \quad (2.17)$$

Physically, these rates are determined by the noise spectral density of the environment which couples to the qubit mode. (See [Schuster, 2007, Schoelkopf et al., 2003] for further reading) whereby the value of the noise spectral density at $+\omega_{01}$ ($-\omega_{01}$) sets Γ_{\downarrow} (Γ_{\uparrow}).

The energy relaxation time, T_1 , is what is measured in experiment and is defined as $T_1 = 1/(\Gamma_{\downarrow} + \Gamma_{\uparrow})$. It is the characteristic timescale over which the qubit relaxes to its thermal equilibrium state. The fraction of initial energy remaining decreases as e^{-t/T_1} towards its equilibrium value where T_1 is defined to be the energy relaxation time. Thus it is convenient to replace p by $1 - e^{-t/T_1}$. With this substitution we can

write the state of an arbitrary qubit (at zero temperature) as

$$\rho_{T_1} = \begin{pmatrix} 1 - \rho_{11}e^{-\Gamma_{\downarrow}t} & \rho_{01}e^{-\Gamma_{\downarrow}t/2} \\ \rho_{01}^*e^{-\Gamma_{\downarrow}t/2} & \rho_{11}e^{-\Gamma_{\downarrow}t} \end{pmatrix} \quad (2.18)$$

2.1.4 T_{ϕ} dephasing and T_2

Another error that is a major source of qubit decoherence is the dephasing error. Physically, this arises from fluctuations in ω_{01} , the frequency spacing between our qubit levels. Once again, this is highly dependent on the noise environment seen by the qubit. One way to model this error model is to assume ‘white noise’ or a constant noise spectrum where the qubit frequency, ω_{01} fluctuates about its average value as a Gaussian noise process with standard deviation $1/T_{\phi}$. It is important to realize the limitations of this noise model. Often we find the noise spectrum predominantly contains ‘low frequency’ noise, where fluctuations in the qubit’s frequency are slower than the unitary dynamics of interest. If this is the case, we can ‘echo out’ and mitigate some of the decoherence caused by the dephasing noise through dynamical decoupling sequences [Pokharel et al., 2018]. The main culprit for low frequency noise is $1/f$ noise, attributed to uncontrolled two-level systems interacting with our qubit (See [Burnett et al., 2014, Paladino et al., 2014]).

Under ‘white noise’ dephasing, the qubit state gradually loses all phase coherence between $|0\rangle$ or $|1\rangle$ but does not lose any energy. As such, we can write its state as

$$\rho_{T_{\phi}} = \begin{pmatrix} 1 - \rho_{11} & \rho_{01}e^{-t/T_{\phi}} \\ \rho_{01}^*e^{-t/T_{\phi}} & \rho_{11} \end{pmatrix} \quad (2.19)$$

We may choose the Kraus operator that describes dephasing to be the jump operator, $\hat{E}_1 = \sqrt{p}(|0\rangle\langle 0| - |1\rangle\langle 1|)$ which then constrains $\hat{E}_0 = \sqrt{1-p}\mathbf{1}$. This is

one way to see that no-jump backaction is absent for the dephasing error channel, since $\hat{E}_0 \propto \mathbb{1}$. (Normalizing the no-jump trajectory removes the $\sqrt{1-p}$ prefactor.)

We can describe the evolution of a qubit state under both dephasing and relaxation channels as

$$\rho_{T_1} = \begin{pmatrix} 1 - \rho_{11}e^{-t/T_1} & \rho_{01}e^{-t/T_2} \\ \rho_{01}^*e^{-t/T_2} & \rho_{11}e^{-t/T_1} \end{pmatrix} \quad (2.20)$$

where we have introduced T_2 , also sometimes called the transverse decoherence time, the timescale of the decay of the coherences of our density matrix. It is defined by the formula

$$\frac{1}{T_2} = \frac{1}{2T_1} + \frac{1}{T_\phi} \quad (2.21)$$

The ‘pure’ dephasing time, T_ϕ is rarely measured directly but instead extracted from separate measurements of T_1 and T_2 . When $T_\phi \gg T_1$ we are said to be ‘ T_1 limited’. In this regime $T_2 \approx 2T_1$ and there are larger uncertainties in our estimate of T_ϕ .

2.1.5 The Lindblad master equation

The Kraus map formalism is useful when our error channel can be described by a discrete set of Kraus operators. However, more often we wish to simulate the continuous evolution of an open quantum system, with error processes that occur at a characteristic rate in time, rather than known probabilities. For this we use the Lindblad master equation, given by the differential equation

$$\dot{\rho} = -\frac{i}{\hbar}[\mathcal{H}, \rho] + \sum_i \kappa_i \left(\hat{c}_i \rho \hat{c}_i^\dagger - \frac{1}{2} \left(\hat{c}_i^\dagger \hat{c}_i \rho + \rho \hat{c}_i^\dagger \hat{c}_i \right) \right) \quad (2.22)$$

Where the first term is the usual unitary evolution ¹ (the von Neumann equation) and the second term with the summation (the dissipation superoperator) gives rise to decoherence under the set of collapse operators, $\{\hat{c}_i\}$ which each describe a different jump error associated with an error process at rate κ_i . Within the dissipation superoperator, we can further subdivide into terms of the form $\hat{c}_i \rho \hat{c}_i^\dagger$ represent the action of quantum jump errors, whereas the remaining terms describe no-jump backaction and ensures ρ remains normalized. (By using the cyclic property of the trace one can prove the Lindblad master equation always preserves the norm of ρ).

The Lindblad Master Equation models both the dynamics and decoherence of quantum systems with several independent error processes. As an example, to model qubit relaxation and dephasing during system dynamics we can choose the set of collapse (i.e., jump) operators

$$\hat{c}_0 = |0\rangle \langle 1|, \hat{c}_1 = |1\rangle \langle 0|, \hat{c}_2 = |0\rangle \langle 0| - |1\rangle \langle 1|, \quad (2.23)$$

with rates

$$\kappa_0 = \Gamma_\downarrow, \kappa_1 = \Gamma_\uparrow, \kappa_2 = 1/T_\phi \quad (2.24)$$

The collapse operator formalism makes it particularly convenient to sprinkle in additional errors one by one, and examine the effects of specific errors on system dynamics in isolation. It is also very powerful, in the sense that it is extensible far beyond single-qubit systems. It is also valid for errors in quantum harmonic oscillators, and composite systems. It is used extensively throughout this thesis to model error processes in cQED systems. However, it is important to remember that the Lindblad Master equation contains underlying assumptions and should be used only when the following conditions are met:

1. note the minus sign in front of the commutator term, which is opposite to the Heisenberg operator evolution

- We can define a set of independent ‘collapse’ operators to describe each jump error, which are independent with well-defined error rates. A constant ‘white-noise’ spectral density is assumed. (Otherwise, the Bloch-Redfield master equation may be more appropriate).

- Each noise channel is Markovian - the environment does not ‘remember’ any previous interactions with the system
- We cannot obtain the quantities of interest from simpler models such as the quantum Langevin equation.

Solving this differential equation means we can now introduce the notion of order to the jump errors. Suppose we had some collapse operator \hat{c} with error rate κ . After some evolution for small finite time, δt , the system experiences the jump evolution associated with \hat{c} with probability $\kappa\delta t$. This is a ‘first-order’ jump error. The system also experiences the ‘second order’ jump error², \hat{c}^2 with a much smaller probability $(\kappa\delta t)^2$ and so on for higher order jump errors.

2.1.6 The fidelity metric

How do we quantify the effects of errors on our states? Throughout the thesis, we use the metric of state fidelity and state-transfer fidelity (as opposed to other metrics such as process fidelity, channel fidelity). For two pure states, one ideal and the other imperfect, this is defined as the square of their overlap

$$\mathcal{F} = |\langle \psi_{\text{ideal}} | \psi_{\text{imperfect}} \rangle|^2. \quad (2.25)$$

If our imperfect state is the result of an error channel then it can only be written as a density matrix and then we calculate the Fidelity to the ideal error-free pure state as

$$\mathcal{F} = \langle \psi_{\text{ideal}} | \rho_{\text{imperfect}} | \psi_{\text{ideal}} \rangle. \quad (2.26)$$

When defining a gate fidelity for gate \hat{U}_{gate} , we consider the gate application as an error-prone channel, $\mathcal{E}_{\text{gate}}$ and opt to use the average state-transfer fidelity metric.

2. If it can; e.g. if a qubit decays to $|0\rangle$, it cannot experience the decay jump operator again.

This is given for a single qubit gate as

$$\mathcal{F}_{\text{gate}} = \frac{1}{6} \sum_{i \in \{\text{cardinal}\}} \langle i | \hat{U}_{\text{gate}} \mathcal{E}_{\text{gate}}(|i\rangle\langle i|) \hat{U}_{\text{gate}}^\dagger |i\rangle \quad (2.27)$$

Where $\{\text{cardinal}\}$ is the set of 6 cardinal states $\{|0\rangle, |1\rangle, |+\rangle, |-\rangle, |+i\rangle, |-i\rangle\}$ we average over and $\mathcal{E}_{\text{gate}}(|i\rangle\langle i|)$ is the density matrix obtained after acting the error-prone gate on cardinal state $|i\rangle$. We can extend this to two-qubit gates by averaging over the 36 cardinal states from the set $\{|0\rangle, |1\rangle, |+\rangle, |-\rangle, |+i\rangle, |-i\rangle\}^{\otimes 2}$.

2.2 Gates, Measurements and Computation

So far we have only discussed qubits as two-level systems in which we can encode quantum information, and which can interact with uncontrolled degrees of freedom in the environment to cause decoherence. How do we do something useful with several qubits? Generally speaking, there are two ways to desirably control our qubits. The first are gates, desired unitary dynamics on our qubits within the computational subspace. The second are strong measurements on the qubits, from which we obtain useful classical measurement outcomes. Gate-based quantum algorithms are structured by first initializing individual qubits in the computational space, performing unitary evolution which is often broken down into single-qubit and two-qubit gate compilations and then finally reading out qubits in their logical basis to obtain a useful result. Of course, other variations exist where some of the qubits may be measured-out midway through the algorithm and the outcomes fed forward to conditionally change future gates and measurements.

In contrast to gate-based computation is the paradigm of measurement-based quantum computation. The initial proposal involved first creating a many-body entangled state between all the qubits and then performing the computation by performing projective measurements on the qubits. (A one-way quantum computer

[Raussendorf and Briegel, 2001]). Most quantum error correction approaches require frequent mid-circuit measurements to correct for errors, but follow a gate-based approach to algorithms. The subset of the qubits that are frequently measured out and reinitialized are called the ancilla qubits. The data qubits store the logical quantum information and will only be measured at the end of the computation to obtain the result of the algorithm (or midway through if the algorithm requires).

2.2.1 Single qubit gates

The simplest gates are single qubit gates. In general, these are arbitrary unitaries acting on qubits, corresponding to rotations on the qubit Bloch sphere by any angle and around any axis. It is helpful to define the Pauli operators, X, Y, Z which are themselves single qubit gates.

$$\begin{aligned} X &= |0\rangle\langle 1| + |1\rangle\langle 0| = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ Y &= i|0\rangle\langle 1| - i|1\rangle\langle 0| = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \\ Z &= |0\rangle\langle 0| - |1\rangle\langle 1| = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \end{aligned} \tag{2.28}$$

Together with the identity operator, these matrices form a basis for all 2×2 matrices. From these Pauli operators we can define rotations by arbitrary angles about the x, y, z axes of the Bloch sphere. We denote this by notation of the form

$$\hat{O}(\theta) = e^{i\frac{\theta}{2}\hat{O}}, \tag{2.29}$$

for any operator, \hat{O} that satisfies $\hat{O}^2 = \mathbb{1}$ (i.e., \hat{O} is Hermitian). For example, arbitrary rotations about the X axis can be denoted as

$$X(\theta) = e^{-i\frac{\theta}{2}X}. \quad (2.30)$$

The ability to perform rotations by any angle about two different axes of the Bloch sphere is sufficient to construct any single-qubit unitary. In other words, we have universal control of our qubit. However, it is also of interest to consider discrete sets of single qubit gates that are also universal. One such set is $\{H, T\}$ where H is the Hadamard gate and $T(\equiv Z(\pi/4))$ is the gate, defined as

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad (2.31)$$

$$T = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{pmatrix}$$

This small set of discrete rotations forms a universal set of gates for a single qubit. An arbitrary rotation can be implemented by repeated applications of H and T gates. Moreover, by the Solovay-Kitaev theorem, it does not take many gate applications to efficiently approximate an arbitrary gate unitary. Why would this be preferable over a continuous, parameterizable universal gate set such as $\{X(\theta), Z(\theta)\}$? At first glance, a continuous gate set seems like it would always be better, since only three gates at most (See Euler angles) are needed for an arbitrary rotation. However, there are often scenarios, particularly in error correction contexts where we do not have access to continuous rotations at the logical level and instead we must realize a discrete gate set. In general, a discrete universal gate set that contains more building blocks e.g. $\{X, Z, X(\pi/2), Z(\pi/2), T\}$ (where $Z(\pi/2) = S$ is sometimes called the phase gate),

will be able to perform an arbitrary rotation in fewer gate steps.

2.2.2 Entangling gates

The notion of unitaries extends to arbitrary numbers of qubits. Broadly speaking, the goal of quantum computation is to realize universal and fault-tolerant computation. Here universal refers the ability to perform any desired unitary on the computational space of an N -qubit system with no errors. Fortunately, we can decompose any unitary into a sequence of interleaved single-qubit and two-qubit gates. We do not need N -qubit gates!

Before we introduce the two-qubit gates we must first lay out the notation for two qubit states. When these states are separable (no entanglement), we may write the state as

$$|\psi\rangle = |\psi\rangle_1 \otimes |\psi\rangle_2 \quad (2.32)$$

and similarly for the unitaries acting on the system. Writing the state as a column vector is sufficient to describe any two-qubit state, including entangled ones

$$|\psi\rangle = c_0 |00\rangle + c_1 |01\rangle + c_2 |10\rangle + c_3 |11\rangle = \begin{pmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{pmatrix}. \quad (2.33)$$

We can write a general two-qubit gate as a 4×4 matrix. The most famous two-qubit gate is the CNOT gate, defined as

$$\text{CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \begin{array}{c} \text{---} \bullet \text{---} \\ | \\ \text{---} \oplus \text{---} \end{array} \quad (2.34)$$

In words, this gate flips the state of the second qubit (the target qubit) if and only if the control qubit is in the $|1\rangle$ state and is also known as a CX gate. The CNOT gate is maximally entangling. Only one application is needed to generate a maximally entangled state³ from initially separable states, e.g.

$$\text{CNOT}(|+\rangle_1 \otimes |0\rangle_2) = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle). \quad (2.35)$$

Other maximally entangling gates are the CZ gate and the $\text{ZZ}(\pi/2)$ gate, defined here as

$$\text{CZ} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \begin{array}{c} \text{---} \bullet \text{---} \\ | \\ \text{---} \bullet \text{---} \end{array} \quad (2.36)$$

$$\text{ZZ}(\pi/2) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & i & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{array}{c} \text{---} \text{---} \\ | \\ \text{---} \text{---} \\ \text{ZZ}(\pi/2) \end{array}$$

These gates are said to be locally equivalent. By applying single qubit rotations before and/or after the gates we can transform one entangling gate to another, e.g.

$$\begin{array}{c} \text{---} \bullet \text{---} \\ | \\ \text{---} \text{---} \\ \boxed{H} \quad \boxed{H} \end{array} = \begin{array}{c} \text{---} \bullet \text{---} \\ | \\ \text{---} \oplus \text{---} \end{array} \quad (2.37)$$

The CNOT and its local equivalents are powerful gates, because when combined with universal single qubit gates they form a universal gate set for a system of N qubits. A commonly example of such a gate set is $\{\text{CNOT}, H, T\}$.

³ A state is maximally entangled if its reduced density matrix is a maximally mixed state when we take the partial trace over either qubit 1 or qubit 2's subspace. A maximally entangled state also has unity concurrence [Wootters, 1998].

Gates such as the SWAP gate, defined as

$$\text{SWAP} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{array}{c} \text{---} \\ | \\ \text{SWAP} \\ | \\ \text{---} \end{array} = \begin{array}{c} \text{---} \times \\ | \\ \text{---} \times \\ | \\ \text{---} \end{array} \quad (2.38)$$

are not entangling whatsoever whereas the $\text{eSWAP}(\theta)$ gate [Gao et al., 2019] is maximally entangling⁴ at $\theta = \pi/2$ for which this gate takes the form

$$\text{eSWAP}(\pi/2) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} & 0 \\ 0 & \frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{array}{c} \text{---} \\ | \\ \text{eSWAP}(\pi/2) \\ | \\ \text{---} \end{array} \quad (2.39)$$

We also define the parametrized versions of these gates as

4. We are redefining $\theta \rightarrow \theta/2$ compared to the definition in [Gao et al., 2019]

$$\text{CPHASE}(\theta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & e^{i\theta} \end{pmatrix}$$

$$\text{ZZ}(\theta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & e^{i\theta/2} & 0 & 0 \\ 0 & 0 & e^{i\theta/2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (2.40)$$

$$\text{eSWAP}(\theta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\theta/2) & i \sin(\theta/2) & 0 \\ 0 & i \sin(\theta/2) & \cos(\theta/2) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

2.2.3 Three-qubit gates

For simplicity of experimental implementation and compilation, we do not perform N -qubit unitaries directly but instead decompose them into interleaved single-qubit and two-qubit gates. However, there are some three-qubit gates of interest. The Toffoli gate or CCNOT gate written as

$$\text{CCNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} = \begin{array}{c} \text{---} \bullet \text{---} \\ | \\ \text{---} \bullet \text{---} \\ | \\ \text{---} \oplus \text{---} \end{array} \quad (2.41)$$

In words, this gate flips the third qubit (the target qubit) only if both control qubits are in the $|1\rangle$ state. This gate is of interest because $\{\text{CCNOT}, H\}$ also forms a

universal gate set. In large scale quantum error correction, a fault-tolerant Toffoli gate (or CCZ gate) can have advantages over the T gate [Litinski, 2019b].

The other three-qubit gate we will introduce is the quantum Fredkin gate or controlled-SWAP gate. The Fredkin gate is universal for classical computation (but the quantum version is not universal). This gate is useful in the implementation of a quantum Random Access Memory (qRAM) [Giovannetti et al., 2008, Hann et al., 2019] and can reduce the number of two-qubit gates needed to implement quantum algorithms. It is defined as

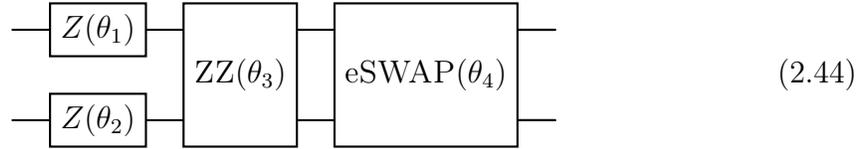
$$\text{CSWAP} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{array}{c} \text{---} \bullet \text{---} \\ | \\ \text{---} * \text{---} \\ | \\ \text{---} * \text{---} \end{array} \quad (2.42)$$

2.2.4 Excitation preserving gates

An important subset of two-qubit gates are excitation preserving (EP) gates. These gates are of interest because we will see they are much easier to implement on certain cQED hardware and in some encodings, such as the dual-rail code, every gate needed for quantum computation can be excitation preserving (at the hardware level). EP gates by themselves also find use in quantum simulation circuits [Google AI Quantum et al., 2020]. The most general excitation preserving gate can be parameterized by four angles: $\phi_1, \phi_2, \phi_3, \theta$ and decomposed into a $e\text{SWAP}(\theta)$ gate followed by an arbitrary phase gate on the two-qubit system

$$U_{\text{EP}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & e^{i\phi_1} & 0 & 0 \\ 0 & 0 & e^{i\phi_2} & 0 \\ 0 & 0 & 0 & e^{i\phi_3} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\theta/2) & i \sin(\theta/2) & 0 \\ 0 & i \sin(\theta/2) & \cos(\theta/2) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (2.43)$$

Which can alternately be parameterized by a different set of angles and the gate circuit



With the alternative parameterization

$$\phi_1 = (\theta_2 + \theta_3)/2$$

$$\phi_2 = (\theta_1 + \theta_3)/2$$

$$\phi_3 = (\theta_1 + \theta_2)/2$$

$$\phi_4 = \theta$$

2.2.5 The Clifford group

Which gates are needed in a quantum computer? The most basic gates are the Pauli group, $\{\mathbf{1}, X, Y, Z\}$. With the addition of the Hadamard and CNOT gate we form the Clifford group, $\{\mathbf{1}, X, Y, Z, H, \text{CNOT}\}$. Elements of the Clifford group acting (by conjugation) on the Pauli operators can only transform them into other Pauli operators. An example in the single-qubit case is $HXH^\dagger = Z$ and for the two-qubit case $\text{CNOT}(X_1 \otimes \mathbf{1}_2)\text{CNOT}^\dagger = X_1 \otimes X_2$.

The Gottesman-Knill theorem [Nielsen and Chuang, 2010] states that a quantum algorithm using only the Clifford group can be simulated efficiently on a classical computer. It is only when we add to this group a gate which is outside the Clifford group (a non-Clifford gate) that we can perform quantum computation with a clear edge over classical computers. The most common choice for this non-Clifford gate is the T gate or Toffoli gate, but a qubit rotation about any axis by an angle that is not an integer multiple of $\pi/2$ would also suffice.

2.2.6 Projective measurements

The last stop in our discussion of desired logical operations is a mathematical description of projective measurements. We will discuss their physical implementation briefly in Chapter 4. The kind of measurement we perform on individual qubits is a strong, binary-valued projective measurement in our basis of choice. Suppose we measure a single-qubit in the z basis (also called a Z measurement). With probability $|c_0|^2$ we obtain the the outcome, 0, a classical signal that tells us the qubit is in the $|0\rangle$ state. Evolution proceeds as if we applied the projector $|0\rangle\langle 0|$ to our qubit state (and then normalized the state). Similarly, with probability $|c_1|^2$ we obtain the outcome signal 1 and apply the projector $|1\rangle\langle 1|$ to our state.

This measurement is an example of a Quantum Non-Demolition (QND) measurement. The qubit state perturbed by the minimum amount required to obtain the measurement outcome, i.e., only the measurement projector has effectively been applied to the state. This is in contrast to a destructive measurement which we define to encompass all non-QND measurements. For example, we (or the environment) learn excessive information about the qubit state or the qubit may end up in a non-computational state (i.e. ‘destroyed’) after the measurement. The latter is often the case with fluorescence readout in neutral atoms, trapped ions and phase qubits [Martinis, 2009].

2.2.7 End-of-the-line vs mid-circuit measurements

There are two contexts in which we perform projective measurements on qubits. The first is an ‘End-Of-The-Line’ (EOTL) measurement. As the name suggests, this is the last operation we will perform on the qubit before it is either discarded or re-initialized in a fresh state. The goal of the measurement is to extract information about the state of the qubit prior to the measurement, usually a measurement in its logical basis. EOTL measurements do not have to be QND. We only care about how well

we can extract the information (the measurement fidelity). EOTL measurements can still benefit from QND measurements, as we can repeat the measurement many times and majority vote on the outcomes to suppress measurement errors, but QNDness is not a strict requirement.

An EOTL measurement may be the last time we ever use that particular qubit in the quantum circuit we are executing. Or, the qubit may be reinitialized which is the case if, for example, the qubit is an ancilla qubit used for a stabilizer measurement (see 3.2).

The other kind of measurement is a mid-circuit measurement. We will use the same qubit for gates/measurements later on so we want to preserve its state. These measurements are required to be QND, at least on the computational space and are denoted by continuing the quantum wire after the measurement. Circuits for EOTL and mid-circuit measurements are shown in the figure below.

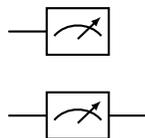
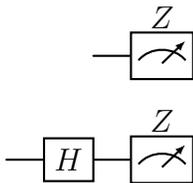


Figure 2.2: (Top) End-of-the-Line measurement on a single qubit. (Bottom) Mid-circuit measurement on a single qubit, presumed to be QND on states in the computational space.

2.2.8 Changing the measurement basis

The measurement basis is usually fixed by the hardware implementation but we can measure in other bases by performing additional single-qubit gates. For example, we can measure in the x basis by performing a $Y(\pi/2)$ or Hadamard gate prior to a Z measurement as shown below.



We can extend this type of measurement to more than one qubit and uncover an interesting observation. QND measurements on multiple qubits are also capable of generating entangled states and have the same computational power as a CNOT gate. As an example we introduce the ZZ measurement, a binary valued measurement where we obtain outcome 0 if the two-qubit state is in $|00\rangle$ or $|11\rangle$. The corresponding projector applied to the two-qubit state is $|00\rangle\langle 00| + |11\rangle\langle 11|$, where phase coherence between these two states is preserved. Obtaining outcome 1 signals we are in the $\{|01\rangle, |10\rangle\}$ manifold of states, with corresponding measurement projector $|01\rangle\langle 01| + |10\rangle\langle 10|$.

If we begin in the separable state $|+\rangle_1 \otimes |+\rangle_2$ then the ZZ measurement will always project us into a Bell state. Half the time we will obtain outcome 0 and the Bell state $(|00\rangle + |11\rangle)/\sqrt{2}$. The other possibility is we obtain outcome 1 and the different Bell state $(|01\rangle + |10\rangle)/\sqrt{2}$. This measurement is also referred to as a ‘joint parity measurement’. If we define the joint parity to be $(i + j) \bmod 2$ for two-qubit state $|ij\rangle$ then outcome 0 signifies ‘even’ joint parity and outcome 1 signifies ‘odd’. The circuit below in Fig. 2.2.8 shows how two QND ZZ measurements, Hadamard

gates and classical feedforward may be used to deterministically implement a logical CNOT gate between two qubits [Beenakker et al., 2004, Litinski and Oppen, 2018]. Note that in this circuit, we require a third auxiliary qubit.

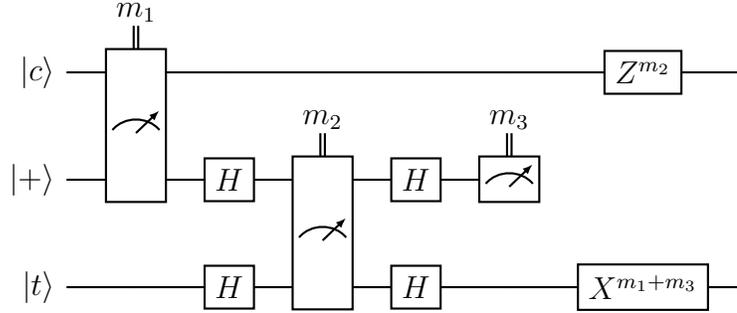


Figure 2.3: Performing a CNOT gate with ZZ measurements and feedforward. CNOT gate is ultimately performed between the top-qubit (control) and the bottom-rail qubit (target). m_1 is the classical outcome (0 or 1) from the first ZZ measurement, m_2 is the outcome of the second ZZ measurement and m_3 is the outcome when we measure out the middle-rail auxiliary qubit in its z basis. These three measurement outcomes are used to conditionally apply Z Pauli corrections to the control qubit and X Pauli corrections to the target. The gate is deterministic, but the three measurement outcomes are random.

The two-qubit measurement known as a Bell measurement will be of particular importance in chapter 7 for the quantum teleportation protocol. Unlike the ZZ measurement, this measurement gives four different outcomes, corresponding to projecting into one of the four different Bell states defined below:

$$\begin{aligned}
 |\Phi_+\rangle &= \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle) \\
 |\Phi_-\rangle &= \frac{1}{\sqrt{2}} (|00\rangle - |11\rangle) \\
 |\Psi_+\rangle &= \frac{1}{\sqrt{2}} (|01\rangle + |10\rangle) \\
 |\Psi_-\rangle &= \frac{1}{\sqrt{2}} (|01\rangle - |10\rangle).
 \end{aligned} \tag{2.45}$$

Such a measurement may be performed non-destructively by performing a QND ZZ measurement followed by a QND XX measurement, since these operators commute.

See the following circuits:

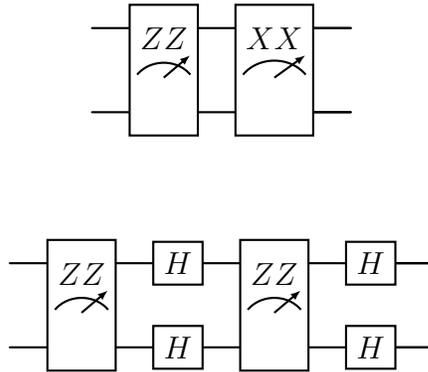


Figure 2.4: Circuits to perform a Bell measurement with non-destructive joint measurements. Two classical bits are extracted to project us into one of the four Bell states. These measurements are QND. If we began in a Bell state, we would measure which Bell state we are in and remain in that same Bell state.

Destructive Bell measurements are more common for quantum teleportation and ‘fusion’ measurements [Bartolucci et al., 2023] found in measurement-based quantum computing. Possible circuits are shown below.

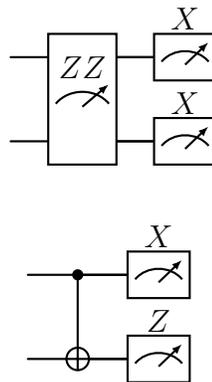


Figure 2.5: Circuits for destructive Bell measurements. At the end of the circuit, the qubits are no longer in a Bell state, but we still perform the Bell measurement. (which is equivalent to performing a QND Bell measurement to extract the two bits, then destroying the states). Destructive measurements are typically easier to realize than QND ones. In the bottom circuit, we could reconstruct the Bell state with subsequent CNOT gates. In the top two circuits, we have measured three bits of information and so the Bell states are impossible to recover.

2.3 Entanglement

2.3.1 Correlations in entangled states

Entangled states are quantum states with correlations not allowed by classical systems. Pretty much any useful quantum information task such as computation or communication relies on the ability to generate high-fidelity entanglement between qubits. The archetypal entangled state are two-qubit states known as Bell states defined in the previous section. These are examples of maximally-entangled states, showing the strongest non-local correlations allowed by quantum mechanics and no local correlations.

What do we mean by correlations? This is inherently tied to quantum measurements. A local correlation would be the expectation value of a Pauli operator for one qubit for example. The $\langle Z \rangle$ value can be measured by preparing N identical copies of a qubit state and measuring it in the z basis. The expectation value is then given by

$$\langle Z \rangle = \lim_{N \rightarrow \infty} \frac{\# \text{outcome } 0\text{s} - \# \text{outcome } 1\text{s}}{N} \quad (2.46)$$

We can measure the other single qubit correlators, $\langle X \rangle$, $\langle Y \rangle$ by repeatedly measuring copies in the x or y bases. The correlator $\langle \mathbf{1} \rangle$ refers to the probability we measure the qubit to be in its computational subspace (when there exists the possibility of leakage errors) and we obtain this by adding the two outcomes when relevant.

If we prepare a qubit in $|0\rangle$, we would obtain $\langle Z \rangle = 1$, $\langle X \rangle = 0$, $\langle Y \rangle = 0$. i.e., there is a maximally strong correlation when we measure in the z basis and no correlation in the other bases. For the maximally mixed single-qubit state, all three of these correlations are zero. From the expectation values (also known as the Stokes parameters), $\langle \mathbf{1} \rangle$, $\langle X \rangle$, $\langle Y \rangle$, $\langle Z \rangle$ it is possible to reconstruct the full qubit density matrix

as

$$\rho_{\text{recon}} = \frac{1}{2} (\langle \mathbb{1} \rangle \mathbb{1} + \langle X \rangle X + \langle Y \rangle Y + \langle Z \rangle Z) = \frac{1}{2} \sum_{i=0,1,2,3} \langle \sigma_i \rangle \sigma_i \quad (2.47)$$

Where σ_i is a Pauli matrix from the set $\{\mathbb{1}, X, Y, Z\}$. Things become more interesting when we look at the correlations in a two-qubit system. In the Pauli basis, there are now 16 correlations we can measure, which form the set $\{\mathbb{1}, X, Y, Z\}^{\otimes 2}$. These can be subdivided into two-qubit correlators such as $Z_1 \otimes Z_2$ or $X_1 \otimes Y_2$, single qubit correlators such as $\mathbb{1}_1 \otimes Z_2$ or $X_1 \otimes \mathbb{1}_1$ and the correlator $\mathbb{1}_1 \otimes \mathbb{1}_2$ which gives the probability we are in the computational space of the two-qubit system.

Single qubit correlators are measured as before in Eq. 2.48 for the respective qubits, ignoring the measurement outcome of the other qubit. To measure a two-qubit correlation, such as $\langle X_1 Y_2 \rangle$, we must measure qubit 1 in its x basis and qubit 2 in its y basis simultaneously to obtain four possible measurement outcomes, denoted by $(0_1, 0_2)$, $(1_1, 1_2)$, $(0_1, 1_2)$ and $(1_1, 0_2)$. The first two measurements are correlated whilst the second two measurements are anti-correlated. We can measure the expectation value for general two-qubit correlator $\langle \sigma_i \sigma_j \rangle$ for $i, j = 1, 2, 3$ with the procedure

$$\langle \sigma_i \sigma_j \rangle = \lim_{N \rightarrow \infty} \frac{\#(0_1, 0_2) + \#(1_1, 1_2) - \#(0_1, 1_2) - \#(1_1, 0_2)}{N} \quad (2.48)$$

The system has some degree of entanglement if the correlations cannot be reproduced by a classical state. The hallmark of a Bell state, is zero local correlations and perfect non-local correlations that persist throughout any joint measurement basis.

This last point is important. We can consider a classical state such as

$$\rho_{\text{classical}} = \frac{1}{2} |00\rangle \langle 00| + \frac{1}{2} |11\rangle \langle 11| \quad (2.49)$$

which can be thought of classically as the output of a machine that produces $|00\rangle$ or $|11\rangle$ depending on the output of a coin toss. This state also has no local correlations

and $\langle Z_1 Z_2 \rangle = +1$ (when we measure one qubit to be in $|0\rangle$ or $|1\rangle$, we always find the other qubit in the same state). But unlike a Bell state, the non-local correlations do not persist throughout different measurement bases, i.e., $\langle X_1 X_2 \rangle = \langle Y_1 Y_2 \rangle = 0$.

For the Bell state $|\Phi_+\rangle$ we find the only non-zero correlations are $\langle \mathbb{1}_1 \mathbb{1}_2 \rangle = \langle X_1 X_2 \rangle = \langle Y_1 Y_2 \rangle = \langle Z_1 Z_2 \rangle = +1$. For the Bell state $|\Psi_-\rangle$ we observe the measurements are maximally anti-correlated ($\langle X_1 X_2 \rangle = \langle Y_1 Y_2 \rangle = \langle Z_1 Z_2 \rangle = -1$).

2.3.2 Measurements with an ancilla qubit

Entanglement provides a pathway to performing QND measurements on one ‘data qubit’ via an extra ‘ancilla qubit’ whose purpose is to extract quantum information non-destructively from the data qubits, and to assist in performing gates on the data qubits. When used in ‘measurement mode’ the general approach is to entangle the data qubit observables with states of the ancilla qubit, and then measure the state of the ancilla qubit. Even if the ancilla measurement is destructive, this still amounts to a non-destructive measurement of the data qubit. It is also useful when we are unable to measure the data qubit directly, but able to entangle it with an ancilla, which is a situation frequently encountered when the data qubit is a bosonic mode, and the ancilla a two-level system.

Suppose we wish to measure some binary-valued discrete operator \hat{B} , which can take two values. This can be a Pauli operator or an observable such as the photon number parity in a bosonic mode, which can only be even or odd. The general circuit as shown below requires first preparing the ancilla in $|+\rangle$, entangling with the data qubit via a control- B unitary and finally measuring the ancilla in the x basis.

The state of the system just prior to the measurement can be written as

$$|\psi\rangle_{1,2} = \left(\frac{\mathbb{1} + \hat{B}}{2} \right) |\psi\rangle_1 |+\rangle_2 + \left(\frac{\mathbb{1} - \hat{B}}{2} \right) |\psi\rangle_1 |-\rangle_2 \quad (2.50)$$

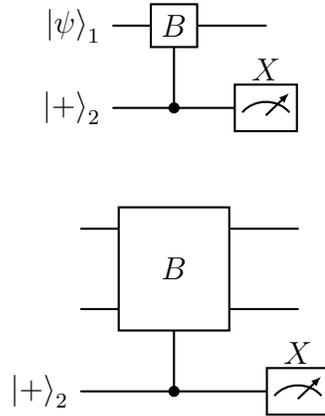


Figure 2.6: ‘Ramsey-like’ measurement circuit for performing QND measurements of \hat{B} with an ancilla qubit. The operator \hat{B} may be acting on one (top) or several (bottom) different qubits. For this circuit to work, operator \hat{B} must also satisfy $\hat{B}^2 = \mathbf{1}$.

With probability p_{\pm} we measure $|\pm\rangle_2$, corresponding to outcome ± 1 and apply projector $\hat{\Pi}_{\pm} = (\mathbf{1} \pm \hat{B})/2$ to $|\psi\rangle_1$. The probabilities are given by

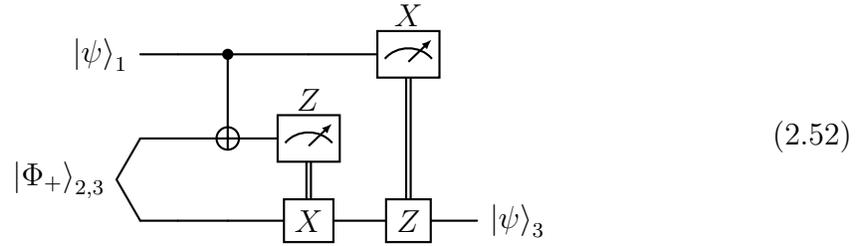
$$p_{\pm} = \langle \psi | \hat{\Pi}_{\pm} | \psi \rangle_1 \quad (2.51)$$

2.4 Quantum Teleportation

2.4.1 The protocol

Entangled states and specifically Bell states are commonly used as resource states for quantum information protocols. Quantum teleportation is a fundamental example of this, and shows that if two remote stations, Alice and Bob, share an entangled Bell state and a classical communication channel, then this can be used as a resource for Alice to transmit an unknown qubit state to Bob. The quantum teleportation

protocol was first put forth by [Bennett et al., 1993]



Suppose Alice wants to send a qubit to Bob, which is some arbitrary state⁵ in qubit 1 given by $|\psi\rangle_1 = c_0 |0\rangle_1 + c_1 |1\rangle_1$. First, Alice and Bob must prepare a Bell state (also called a Bell pair or EPR pair [Einstein et al., 1935, BELL, 1966]) between qubit 2 at Alice, and qubit 3 at Bob which we suppose is $|\Phi_+\rangle_{2,3}$. Alice then performs a Bell measurement between her qubits and obtains two classical bits of information which we denote as m_1 and m_2 . (since there are four measurement outcomes see 2.2.6). Alice then sends these outcomes via a perfect classical channel to Bob. Bob then performs the gate X on his qubit if $m_1 = 1$ and/or gate Z on his qubit if $m_2 = 1$. As soon as Alice performs her Bell measurement, Bob's qubit is unentangled with qubit 2. After these 'Pauli corrections' the state of Bob's qubit is then $|\psi\rangle_3 = c_0 |0\rangle_3 + c_1 |1\rangle_3$. In other words, Alice's qubit has been 'teleported' to Bob. Note that this does not violate the no-cloning theorem, since qubit 1 no longer contains any information about the original qubit state.

There are a few extra classifiers for the type of teleportation we can perform. The teleportation is deterministic if the Bell measurement always gives a valid outcome which Bob then uses. In some platforms, particularly optics, the Bell measurement may only give a valid outcome say 50% of the time [Braunstein and Mann, 1995]. For example, it may give outcomes (1,0) and (0,1) conclusively, but give a 'failed' outcome because, for platform-specific reasons, it is unable to work for Bell states

5. We can use any of the four Bell states in Eq. 2.45 and change the Pauli corrections accordingly.

$|\Psi_{\pm}\rangle$. This means on average half the time we successfully teleport the qubit, and half the time qubit 1 is lost.

The other classifier concerns whether we actively perform the Pauli corrections, or track them ‘in software’. We distinguish these two cases as *active* vs *passive* teleportation. Tracking Pauli corrections is common in error correction contexts, since actively performing the Pauli gates can introduce more errors. With this strategy, we use the information of the Pauli corrections to update which gates we perform next, or which basis to measure our qubits in. One interesting consequence of ‘in software’ tracking is that we may perform subsequent operations on qubit 3 before Alice has completed her Bell measurement. Eventually however, we must use the outcomes of Alice’s Bell measurement to actively change the gates and measurements we perform on Bob’s side, particularly for non-Clifford gates such as T gates. Overall, deterministic and active teleportation is the protocol that is most closely equivalent to ideal quantum state transfer [Pirandola et al., 2015].

2.4.2 Use in quantum communication

Quantum teleportation is vital to quantum communication protocols. The quantum communication problem concerns how to transfer quantum information such as qubit states between Alice and Bob when they are separated by some appreciable distance. Over these distances, energy losses and noise in the links or communication channel mean we are almost certainly doomed if we try and send our qubit state directly through the link. However, quantum teleportation says that all we need is a Bell state between Alice and Bob and a classical communication channel and then we can teleport our qubit. Although we may be almost certainly doomed if we try and create a Bell state between Alice and Bob via the lossy link, there is a subtle difference. We can attempt to create this Bell state several times, and we can perform non-destructive measurements that verify which attempts are successful, even if each attempt is very

unlikely to succeed. This approach is known as ‘heralding’ or ‘repeat until success’.

We might be waiting around for a while, but as soon as we get lucky and herald a Bell state, we can then teleport our qubit state. With this approach, the inefficiencies of our link greatly reduce our qubit data transfer rate, but we do not compromise transfer fidelity unlike in a direct approach. This approach is so ubiquitous, that it is often presumed to be the only way to send qubits over a long distance network and has been demonstrated in a plethora of experiments in different qubit platforms that either use fiber optics or free space as the communication medium [Pirandola et al., 2015, Xia et al., 2017].

There is another problem with this approach which is that energy remaining tends to decrease exponentially with distance. Eventually, the losses may be so great that the probability of heralding even a single Bell state becomes astronomically low. The solution to this is the quantum repeater [Briegel et al., 1998]. Suppose the losses between Alice and Bob are too large to herald a Bell state but we have another station at the halfway point, which we call Charlie. The losses are still large between Alice and Charlie, and Charlie and Bob but not so large that we cannot herald two Bell states between Alice and Charlie and Charlie and Bob. By using a protocol known as entanglement swapping, Charlie can perform a Bell measurement on their two qubits (one from each Bell pair) and classically communicate the outcomes to Alice. Once Alice is aware of these outcomes, Alice’s qubit is now entangled with Bob and not with Charlie’s anymore [Hermans et al., 2022]. Quantum teleportation can then proceed as normal. The key advantage of this quantum repeater protocol is that it scales favorably with distance, and does not suffer from the exponential decrease in successfully heralding Bell states.

Quantum teleportation can only be as good as the Bell state prepared between Alice and Bob. If we can only create noisy Bell states then the entanglement distillation protocol can get round this [Bennett et al., 1996]. The approach is to instead prepare

many noisy Bell state pairs between Alice and Bob and then use local operations such as Bell measurements to end up with one remaining two-qubit state which has a much higher fidelity with a perfect Bell state. The entangled pairs are distillable if their fidelity with a Bell state is $> 50\%$.

All together, these different strategies can be combined to achieve ever-better qubit transfer fidelity in the presence of faulty components. These can be summarised as the following

1. If we cannot send the qubit directly through the link (on-demand deterministic state transfer [Cirac et al., 1997]) because losses are too high, use quantum teleportation.
2. If it takes too long to herald a Bell state, use the quantum repeater protocol and entanglement swapping.
3. If the Bell states we herald are not high enough fidelity, prepare many copies and use the entanglement distillation protocol.

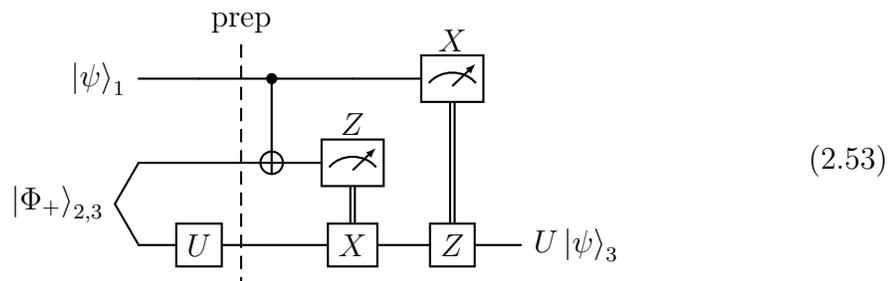
It is important to bear in mind that each time we adopt one of these protocols to make up for some imperfection in our system we are making a trade-off. By adopting teleportation we sacrifice the qubit transfer data rate, and the wait time between each qubit transfer is now inherently probabilistic. Our qubit state transfer (QST) fidelity will always be limited by the fidelity of our Bell state *and* our Bell measurement. With the repeater protocol, the number of Bell states we need to prepare and the number of Bell measurements scales linearly with the number of intermediate nodes between Alice and Bob. Finally, the entanglement distillation protocol can overcome the limitation of noisy Bell states but again requires a larger number of entangled pairs to be generated per each qubit transfer. In all these workarounds the fidelity of our local Bell measurements is still sets the ‘bottom line’ of our QST fidelity. On-demand deterministic state transfer is the most direct method of quantum communication,

with the minimum operation overhead but maximum susceptibility to errors in the communication channel.

The preceding discussion has implicitly assumed optical links, where quantum information is sent via photonic wavepackets. In Chapter 7, we show that these assumptions are not valid for all kinds of quantum communication channels, especially when the channel can be treated as a single standing wave mode.

2.4.3 Gate teleportation

First outlined in [Gottesman and Chuang, 1999], gate teleportation is based on the previous state teleportation circuit, but allows us to also perform a gate on qubit 3. The modification to the circuit is simple and shown below.



By performing gate unitary U on qubit 3 after the Bell pair creation but before Alice's Bell measurement, the final qubit 3 state (after Pauli corrections) will be $|\psi\rangle_3 = U(c_0|0\rangle + c_1|1\rangle)$. But why would we ever want to do this? What advantages does this provide over just performing U directly on qubit 3 after the whole teleportation sequence? Again, the benefit comes from being able to try multiple times until success. Suppose we can 'herald' the success of unitary U just as we herald successful Bell state generation (i.e., U is an *error-detected* gate). We now only proceed with the Bell measurement and teleportation if both the Bell state is heralded successfully and the gate is performed successfully (everything on qubits 2 and 3 before the dashed line counts as 'preparation'). Otherwise, we keep qubit 1 idle and try again. Now

when U fails, we do not have to throw away qubit 1, we only need to retry the Bell state production and the gate. This style of gate teleportation can be extended to two-qubit gates such as CNOT gates [Chou et al., 2018] although this requires more complex entangled states and more Bell measurements.

We will also examine gate teleportation as a flexible tool for indirectly implementing gate unitaries when we have limited control on some of our qubits. For this we will use the following circuit.

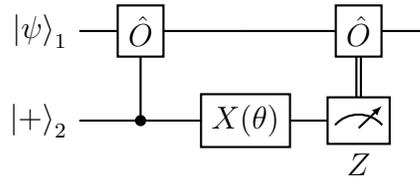


Figure 2.7: ‘Teleporting’ the gate $\hat{O}(\theta)$ onto qubit 1. When we measure outcome 1 on qubit 2, we must conditionally apply another \hat{O} gate.

This circuit has been used to perform teleported gates on GKP qubits [Flühmann et al., 2019, Campagne-Ibarcq et al., 2020]. Here we are interested in this circuit for a very hardware-specific context that forms the basis of the gates introduced in Chapter 6. We are interested in the case where qubit 1 and qubit 2 are ‘different species’ of qubits (such as qubit 1 being a bosonic mode and qubit 2 a two-level system). We can perform arbitrary qubit rotations on qubit 2 but not on qubit 1. On qubit 1 we can only perform Pauli gates and control Pauli gates with qubit 2 acting as the control. We can use gate teleportation techniques and feedback to perform rotations of the form $\hat{O}(\theta)$ on qubit 1. At the end of this circuit, if we measure qubit 2 to be in $|0\rangle$ (50% probability), we will have performed gate $\hat{O}(\theta)$ on qubit 1. But if we measure $|1\rangle$ we will have performed $\hat{O}(\theta + \pi)$ instead. Fortunately we are allowed to use feedforward to conditionally apply the Pauli gate \hat{O} to qubit 1, so overall we perform the desired gate.

2.4.4 Measurement-free gate teleportation

A drawback of this approach is the required measurement and feedforward. If the measurement has an error this directly causes an error on the data qubit(s). This is circumvented with the ‘measurement-free’ teleportation circuit below:

At the end of the sequence, qubit 2 returns to its original state, $|+\rangle$ and has been disentangled with the rest of the qubits whilst the gate $\hat{O}(\theta)$ is deterministically applied to qubit 1 as desired. This protocol can be extended to multiple qubits as shown in the circuit below:

The general circuit is also called the ‘exponentiation gadget’ and can be extended to Pauli-like operators ($\hat{O}^\dagger = \hat{O}$) that act on multiple qubits. At first glance, this circuit seems a little strange—why would you not be able to do $\hat{O}(\theta)$ directly on qubit 1?

It turns out this is often the case with bosonically encoded qubits where it can be significantly easier to do control-Pauli operations controlled by an ancilla qubit and perform the rotation by angle θ on the ancilla rather than the bosonic mode. In fact, amongst bosonic codes, this construction is becoming more ubiquitous. It forms the foundation of both the e SWAP gate [Gao et al., 2019] and ZZ(θ) gates that form the bulk of Chapter 5.

Quantum Error Correction and Detection

3.1 Quantum Error Correction Fundamentals

Quantum error correction (QEC) is arguably the central challenge that must be overcome to realize a useful quantum computer. Quantum algorithms must still reliably give a useful output even when the underlying qubits and gate operations are noisy. QEC is significantly harder than classical error correction. The no-cloning theorem forbids us from simply making ‘backups’ of our qubits and we cannot directly measure the logical information in our qubits without altering them via our projective measurements.

3.1.1 Knill-Laflamme conditions

Instead we must use a larger space of states, which have an inbuilt redundancy to errors. One subspace, known as the logical subspace (or computational subspace), encodes qubit information and decoherence mechanisms take us to an orthogonal space, known as the error subspace. All QEC protocols must realize a channel that takes us from the error subspace back to the logical subspace whilst preserving the quantum information (i.e., without disturbing or gaining knowledge of the logical qubit state). These requirements are formalized as the Knill-Laflamme conditions [Knill and Laflamme, 1997].

Suppose our qubit state undergoes an error given by channel

$$\rho \rightarrow \sum_i \hat{E}_i \rho \hat{E}_i^\dagger \quad (3.1)$$

Where $\{\hat{E}_i\}$ form the set of error processes. The Knill-Laflamme conditions can be written as

$$\begin{aligned} \langle 0_L | \hat{E}_i^\dagger \hat{E}_j | 0_L \rangle &= \langle 1_L | \hat{E}_i^\dagger \hat{E}_j | 1_L \rangle \\ \langle 0_L | \hat{E}_i^\dagger \hat{E}_j | 1_L \rangle &= \langle 1_L | \hat{E}_i^\dagger \hat{E}_j | 0_L \rangle = 0 \end{aligned} \quad (3.2)$$

In words, this means our error processes must preserve the orthogonality of our codewords once they have been transported to our error space, and that the rate at which a given error process takes us to the error space is the same for all codewords.

If these conditions are satisfied, then there exists a linear map that takes us back to ρ , our initial qubit state and it is as though the error never happened. This is the bare-bones theory—it does not tell us how best to implement the error correction step or whether our error correction itself will introduce more errors!

Next we discuss two main strategies for realizing error correction: with stabilizer codes and with bosonic QEC. In the first approach, our logical codewords reside

in the larger Hilbert space formed by an ensemble of many ‘physical qubits’. Many physical qubits are needed to realize each logical qubit. With bosonic error correction, we instead encode redundantly in the larger Hilbert space of a quantum harmonic oscillator. The two approaches may be complementary. Bosonic error correction may be employed to make better physical qubits at the ‘hardware’ level, and then many physical bosonic qubits can be assembled to form one logical qubit.

3.1.2 The 3-qubit repetition code

We can illustrate many of the core underlying concepts of practical error correction with the repetition code [Girvin, 2023], the simplest quantum error correcting code. A logical qubit in this code is comprised of three physical qubits as

$$|\psi_L\rangle = c_0 |0_L\rangle + c_1 |1_L\rangle = c_0 |000\rangle + c_1 |111\rangle \quad (3.3)$$

This is designed to correct for a single bit flip error that can occur on any of our three physical qubits so $\{\hat{E}_i\} = \{\mathbb{1}, X_1, X_2, X_3\}$.

Suppose the error X_1 has occurred. Our state is now

$$|\psi_{E_1}\rangle = c_0 |100\rangle + c_1 |011\rangle \quad (3.4)$$

We must detect the occurrence of this error, without destroying the qubit information encoded in c_0 and c_1 . We can do this through use of an additional ancilla qubit and a QND measurement called a stabilizer measurement. We non-destructively measure the stabilizer X_1X_2 by using the circuit in Fig. 2.6 and setting $B = X_1X_2$. Obtaining outcome -1 tells us the first two qubits are in the subspace $\{|01\rangle_{1,2}, |10\rangle_{1,2}\}$ but not whether we are in $|01\rangle_{1,2}$ or $|10\rangle_{1,2}$, and so we preserve the superposition. If we obtain outcome $+1$, we know neither of these qubits suffered a bit flip error (assuming two errors did not occur!).

With the single stabilizer measurement result of -1 , we know that either the first qubit or the second qubit had a bit flip, but we still don't know which one. We must perform a different stabilizer measurement, this time setting $B = X_2X_3$ to make this unambiguous. If we obtain outcome $+1$ (-1), we know it was qubit 1 (2) that had the error, and apply the correction unitary X_1 (X_2). With these two measurements, we can obtain 4 different measurement outcomes which unambiguously tell us whether to apply $\mathbb{1}, X_1, X_2, X_3$ to correct for the presence (or absence) of a bit flip error.

What happens if we do not have a discrete bit flip error but a continuous one such as $X_1(\theta)$ by some unknown angle θ ? Although this error (and realistic errors) is analog (θ can take any continuous value) the stabilizer measurements always give discrete outcomes. By this we mean $X_1(\theta) = \cos(\theta/2) \mathbb{1}_1 + i \sin(\theta/2) X_1$, is a superposition of error and no-error. The act of performing the stabilizer measurement collapses this superposition, so if we measure -1 , which happens with probability $\sin^2 \theta$, we definitely had a bit flip error and should correct for it! Conversely, if we measure $+1$, the act of measuring has 'collapsed away' the error and we do not need to correct for anything.

Now we suppose the bit flip errors all arise from the same error channel on each qubit where a bit flip happens with probability p on each qubit independently. How much has our error correction helped reduce the logical error probability? Without any error correction and a single qubit, our error rate would be p . When we encode in three data qubits, our error is now $3p(1-p)^2 \approx 3p$. Before doing error correction, we are now roughly three times more likely to suffer an error! This is a property of all error correction schemes, which we term the **error correction overhead**. The extra redundancy required for error correction increases our susceptibility to errors happening in the first place! The **hardware overhead** refers to the increase in the number of physical qubits per logical qubit, and the increased number/complexity of operations we must perform to implement error correction. After correcting for all

single qubit errors, there is a chance two or three errors occur, which are uncorrectable errors. This happens with probability $3p^2(1-p) + p^3 \approx 3p^2$, which sets the **logical error rate** of the code, p_L . We can only correct a single bit flip. We will mistake two bit flip errors on X_1 and X_2 for a bit flip on X_3 and so after the error correction we will have a logical bit flip error. If two bit flips occur, it is possible to detect this as an error from our stabilizer measurements, but we cannot correct this error, since both a single bit flip error and double bit flip error can give the same stabilizer measurement outcomes (and it is ambiguous what we should do to correct the error.)

Overall, our error correction helps if $3p^2(1-p) + p^3 < p$ which is true for $p < 1/2$. We call $p = 1/2$ the **breakeven** point, when the quantum error correction becomes beneficial, in spite of its overhead. In practice, we want to be well-past the breakeven point, because we eventually want our error correction to reduce our logical error rate by orders of magnitude compared to the physical error rate. For this reason, the scaling of the logical error rate with the physical error rate is incredibly important. In this case, the logical error rate scales $\sim O(p^2)$. Decreasing p by a factor of 10 leads to a hundredfold reduction in the logical error rate! For codes such as the surface code, this improvement can be even more dramatic. We can also define the **error correction gain**, $\Lambda = p/p_L$. When $\Lambda > 1$ we are beyond the breakeven point and our error correction is beneficial¹. For the repetition code, $\Lambda \approx \frac{1}{3p}$ for $p < 1$. This factor of 1/3 arises from the fact that using three physical qubits to encode one logical qubit increases the error rate by a factor of three.

It is also helpful to introduce the stabilizer formalism. We can succinctly summarize the bit flip repetition code as a three (data) qubit code stabilized by the Pauli operators X_1X_2 and X_2X_3 . (We could also swap out one of the stabilizers for X_1X_3)

There is one more important aspect we have swept under the rug. So far we have

1. We can also define breakeven for a quantum memory in a similar way, as the ratio of the coherence time of the logical qubit information to the physical lifetime of modes in the system, see [Ofek et al., 2016, Ni et al., 2023, Sivak et al., 2023].

assumed our ancilla qubit, all the stabilizer measurements and correction unitaries are perfect and error-free. In a realistic system these would be just as error prone as the data qubits. So how do we correct for errors in them as well? We must redundantly store information in the ancilla qubit, and check them for errors, which in turn requires more physical qubits and more gates to make a robust ancilla. We are starting to accumulate a lot of physical qubits and a lot of overhead!

3.1.3 Fault tolerance

We are also beginning to touch upon the concept of **fault-tolerance**. A rather nebulous concept that seems to mean something a little different to everyone. In spirit, fault tolerance means we are completely ‘water-tight’ to errors. Any errors that occur in any error-prone part of the system can be corrected and do not affect the output of whatever computation we are trying to run. Here we go through the various meanings of fault tolerance in more specific contexts:

- **A fault-tolerant quantum computer.** A quantum computer where the logical error rate (per gate) has been made low enough (presumably through QEC) to run useful algorithms. This error rate is often quoted somewhere between 10^{-10} – 10^{-18}
- **Fault-tolerant QEC.** In the context of performing many rounds of repeated QEC, errors introduced/missed from a previous round or between rounds can still be corrected in subsequent rounds of QEC. Overall this means the chances an error slips through all our rounds of QEC is low enough to perform a useful computation.
- **Fault-tolerant encoding.** When we do not assume unrealistically that some parts of our QEC procedure are error-free. That our chosen quantum error correction code fully consider errors that can occur in state preparation, gates,

ancillas, measurements, error correction unitaries and design our error correction to account for all error possibilities. (The repetition code we introduced earlier is not fault tolerant in this sense)

- **Hardware Fault Tolerance (HFT)**. The ability to tolerate specific jump errors occurring during gate unitaries and measurements on physical qubits. Tolerate means to detect/correct these errors as they occur, without needing excessive amounts of hardware and operations to do so. The goal of Chapter 5 is to extend this to two-qubit gates and measurements.

3.1.4 Correcting for Pauli errors

The bit-flip repetition code only corrects for a single X error² but we have seen most qubit errors can be modelled as Pauli errors—either an X or a Z error (or both since $Y = iXZ$) can occur. What encodings can deal with Pauli errors? We can modify the repetition code with codewords

$$|\psi_L\rangle = c_0 |+++ \rangle + c_1 |-- \rangle \quad (3.5)$$

which corrects for a single Z or phase-flip error but not any bit-flip errors. We can then use the process of **concatenation**. By replacing each physical qubit in the phase-flip code with three physical qubits encoded in the bit-flip code, we end up with a 9-qubit code known as the Shor code [Shor, 1995]:

$$|\psi_L\rangle = \frac{c_0}{2\sqrt{2}}(|000\rangle + |111\rangle)^{\otimes 3} + \frac{c_1}{2\sqrt{2}}(|000\rangle - |111\rangle)^{\otimes 3} \quad (3.6)$$

2. The bit-flip code can only correct for amplitude damping errors if we also physically implement Pauli twirling.

This is not the most qubit-efficient encoding to protect against a single Pauli error. The 5-qubit ‘perfect’ code [Laflamme et al., 1996] achieves this with the stabilizers

$$\begin{aligned} S_1 &= \mathbb{1}_1 X_2 Z_3 Z_4 X_5 \\ S_2 &= X_1 Z_2 Z_3 X_4 \mathbb{1}_5 \\ S_3 &= Z_1 Z_2 X_3 \mathbb{1}_4 X_5 \\ S_4 &= Z_1 X_2 \mathbb{1}_3 X_4 Z_5 \end{aligned}$$

Code distance quantifies the number of (Pauli) errors we can correct for. A distance d code corrects for up to $(d - 1)/2$ Pauli errors and can detect (but not correct) $d - 1$ Pauli errors. The 5-qubit perfect code is an example of a distance $d = 3$ code.

3.1.5 Erasure errors

There is an entirely different class of error known as erasure errors. An erasure error on a physical qubit occurs when we are able to determine where and when a physical qubit had an error, but we are not necessarily sure what happened to its quantum information. The qubit may have suffered a Pauli error, leakage error or even be physically destroyed! The key difference is that we (the agent trying to run a quantum computer) know that something happened to the information in the physical qubit, but we don’t know what.

These errors are almost like classical errors, and as such they are much easier to correct. Any code designed to correct for Pauli errors could also be used to correct for erasure errors without any modification of the hardware or gates. However, It is always easier to correct erasure errors than to correct than Pauli errors, because we can take advantage of the fact that we know the location of the error—there is less entropy in the noise. Other than the knowledge of the location of the erasure, no extra modification to the hardware or gates is required to correct for erasures.

For the error to be classified as erasure, we must also be able to reset that qubit

back into the computational space³ after the error has been detected. Any stabilizer code designed for Pauli errors can always correct double the number of erasure errors [Grassl et al., 1997], essentially doubling the code distance by allowing us to tolerate up to $d - 1$ erasure errors. A $d = 3$ code will have logical error rate due to Pauli errors that scales as $\sim (p^{\text{Pauli}})^2$ but the *same* code would be susceptible to erasure errors at rate $\sim (p^{\text{erasure}})^3$. The reason behind this is the removal of ambiguity in interpreting stabilizer measurements. Given a set of stabilizer measurement outcomes, what is the most likely physical errors on the qubits that gave us these outcomes? This is known as the **decoding** problem. In the bit-flip repetition code, we saw a single bit flip errors and two bit flip errors both gave the same error syndrome, and we assumed a single bit flip error since this was more likely. With erasure errors, we know exactly which and how many qubits had an error, allowing us to distinguish between and correct both of these cases.

Erasure errors must also satisfy their own version of the Knill-Laflamme conditions [Grassl et al., 1997]. For a set of Kraus operators, $\{\hat{E}'_i\}$ that describe the erasure channel, they must satisfy

$$\begin{aligned}\langle 0_L | \hat{E}'_i | 0_L \rangle &= \langle 1_L | \hat{E}'_i | 1_L \rangle \\ \langle 0_L | \hat{E}'_i | 1_L \rangle &= \langle 1_L | \hat{E}'_i | 0_L \rangle = 0\end{aligned}\tag{3.7}$$

In words, this states that erasure error processes should take our logical codewords out of the computational state to an error space that is orthogonal to our computational space, and with the same rate for each codeword⁴. The difference compared with error correction is that we do not need to preserve orthogonality in the error space. Usually this is required so that stabilizer measurements outcomes unambiguously tell us what

3. which may even be a mixed state!

4. This prevents a no-jump backaction error.

unitary we should apply to correct the error. Now, with error detection, we relax this requirement – stabilizer measurements by themselves are allowed to be ambiguous (and we will now use the additional knowledge of the erasure location to ultimately remove this ambiguity). If we compare this to the Knill-Laflamme conditions:

$$\begin{aligned}\langle 0_L | \hat{E}_i^\dagger \hat{E}_j | 0_L \rangle &= \langle 1_L | \hat{E}_i^\dagger \hat{E}_j | 1_L \rangle \\ \langle 0_L | \hat{E}_i^\dagger \hat{E}_j | 1_L \rangle &= \langle 1_L | \hat{E}_i^\dagger \hat{E}_j | 0_L \rangle = 0\end{aligned}$$

For any error correcting code that corrects for t errors in $\{\hat{E}_i\}$, the same code can detect $2t$ errors in the set $\{\hat{E}'_i\} = \{\hat{E}_i^\dagger \hat{E}_j\}$ which is always twice as large as $\{\hat{E}_i\}$. If we then have some additional method of determining which qubit suffered an unknown error from $\{\hat{E}'_i\}$ (i.e., error-detection on each physical qubit), then this becomes a fully error-correctable erasure error.

The smallest stabilizer code that can correct for a single erasure on any physical qubit is the 4-qubit erasure code with stabilizers. With two stabilizers and four

$$\begin{aligned}S_1 &= Z_1 Z_2 Z_3 Z_4 \\ S_2 &= X_1 X_2 X_3 X_4\end{aligned}$$

physical qubits, we have enough degrees of freedom to encode two logical qubits with codewords whilst still being able to correct a single erasure error on any qubit. The four codewords are

$$\begin{aligned}|0_L 0_L\rangle &= (|0000\rangle + |1111\rangle)/\sqrt{2} \\ |0_L 1_L\rangle &= (|1001\rangle + |0110\rangle)/\sqrt{2} \\ |1_L 0_L\rangle &= (|1100\rangle + |0011\rangle)/\sqrt{2} \\ |1_L 1_L\rangle &= (|1010\rangle + |0101\rangle)/\sqrt{2}.\end{aligned}\tag{3.8}$$

The procedure for correcting an erasure is to reset the erased qubit back into the codespace and then to remeasure the stabilizers S_1 and S_2 , applying Pauli updates

as needed. We walk through this procedure explicitly in Appendix A.

Because of the highly favorable scaling of erasure errors, it is worth investigating specific systems where erasure errors can naturally arise as the dominant error source and how to engineer our hardware and measurements such that erasure errors rather than Pauli errors are our dominant error channel. There are two main strategies to achieving this. The first is to error detect processes (to first-order) that are known to produce Pauli errors through use of a flag ancilla qubit or similar measurement, the second is to engineer our systems so our error processes predominantly cause leakage errors, which can then be efficiently detected and converted to erasure errors. We now describe the latter approach.

3.1.6 Leakage errors

Leakage errors take a physical qubit from its computational manifold $\{|0\rangle_L, |1\rangle_L\}$, to states outside of the computational manifold, $\{|\text{Leakage}\rangle\}$. These errors are usually the most damaging to a quantum error correction code and their occurrence is often hardware specific. For example, if $|0\rangle_L$ and $|1\rangle_L$ are both metastable states with a common ground state, then the energy relaxation channel now predominantly results in leakage to the ground state, which at first looks much more damaging compared to Pauli errors. However, a *detected* leakage error followed by qubit reset can be treated as an erasure error in a process known as erasure conversion.

The most benign type of leakage error is when qubits in leakage states do not interact at all with the qubits that remain in the computational subspace. Any attempts to interact with leakage qubits e.g. via a CNOT gate result in identity operations being applied to the non-leaked qubits. As the computation progresses, these leaked qubits act as a permanent, fixed source of Pauli errors on the surrounding non-leaked qubits whenever we try to perform an entangling gate. Leakage results in a permanent reduction in code distance, since within a round of error correction, the non-leaked

qubits that neighbour leaked ones will always suffer from Pauli errors. Code distance can only be restored by resetting the leaked qubits back to the codespace (and measuring stabilizers to re-entangle them with the rest of the physical qubits). This is the best case scenario since, depending on the exact hardware implementation, attempting entangling gates with a leaked qubit can result in the propagation of leakage errors to the surrounding qubits!

For a long enough computation, leakage errors reduce the effectiveness of our quantum error correction until the logical error rate is too large to run useful algorithms. Other error processes, may eventually take us from the leakage manifold back to the computational manifold are referred to as ‘Seepage’ errors [Wood and Gambetta, 2018]. Detecting leakage errors and converting them to erasure errors is one way to handle leakage as we will discuss later. The other way is to greatly increase the seepage rate, essentially by engineering a dissipative processes that takes us back to the computational subspace before leakage errors become disastrous. Effectively, this converts leakage errors to Pauli errors and the operations required to do this are referred to as ‘Leakage Reduction Units’ (LRUs) [Aliferis and Terhal, 2006]. This strategy (and also the erasure conversion strategy) ensures that only a constant, small fraction of qubits are in leakage states rather than steadily growing in time [Miao et al., 2022]. It is important to keep this fraction small, such that the Pauli error rate due to leakage errors is comparable or less than the Pauli error rate due to other physical error processes.

3.1.7 Erasure conversion

From a QEC standpoint, it is more advantageous to handle leakage errors by converting them to erasure errors instead. This approach has been recently proposed for many different qubit platforms [Wu et al., 2022, Kang et al., 2023, Scholl et al., 2023, Kubica et al., 2022, Teoh et al., 2022]. The key difference is that we perform a

measurement which tells us exactly which qubits have leaked before resetting them to the codespace. This additional knowledge is what allows us to remove the ambiguity in decoding future stabilizer measurements. Overall, the procedure has converted leakage errors to erasure errors.

We now walk through this process for the two contexts in which we convert leakage errors to erasure errors in a stabilizer code. Most often, what we want to do is a mid-circuit measurement, whereby we will still use the qubit afterwards in the event that no leakage error is detected. Any detection of the leakage error must be highly QND on states in the computational space; i.e., if the qubit has not leaked, we don't want the detection process to induce additional errors on the qubit! We call a detection procedure that satisfies this requirement a **mid-circuit erasure detection**⁵. If the detection flags a leakage has occurred, then we have successfully performed **erasure conversion**, because the leakage error channel can now be treated as an erasure channel and the leakage error can now be called an erasure error. But to be able to correct the erasure error in the stabilizer code, we must do two more steps. First, we must reset the leaked qubit back into the computational space, and secondly, we must inform our decoder as to which qubit suffered the erasure error and use this information when interpreting the subsequent stabilizer measurement outcomes.

The other context in which we can detect leakage errors is when we are performing an EOTL logical measurement on a qubit. This can occur at the very end of the computation on the data qubits, or during a computation regularly when measuring the ancilla qubits that measure the stabilizers.

As well as determining, e.g., if the state was in $|0_L\rangle$ or $|1_L\rangle$, we can also do a measurement that checks if the qubit has leaked. We call this an **EOTL erasure detection**. These are easier to perform, since they do not need to be QND on the

5. Erasure detection and leakage detection are almost used synonymously. Although a slight misnomer (erasures are by definition detected), erasure detection refers to leakage detection in the context of trying to convert leakage into erasure errors.

computational states. This can be achieved, for example, by a three-state readout that distinguishes between $|0_L\rangle, |1_L\rangle$ and $|\text{Leakage}\rangle$. If the physical qubit is reused again in the computation, it must be reset back into the computational space. All these definitions are a little confusing, so we summarize them in Table 3.1.

Measurement name	Erasure Conversion	Logical measurement w/ erasure detection
Type of erasure detection	Mid-circuit	EOTL
Measurement circuit	$\{ \text{leakage} \rangle\}$ or codespace? 	$ 0\rangle$ or $ 1\rangle$ or $\{ \text{leakage} \rangle\}$? 
How to then correct erasure?	Reset to codespace if leaked and inform decoder	Inform decoder if leaked, reinitialize to known state if reusing qubit later

Table 3.1: Procedure for correcting erasure errors. There are two different types of erasure detection measurements: Mid-circuit and EOTL. Each has a slightly different procedure to correct erasure errors. Realizing mid-circuit erasure detection is the most useful for correcting erasure errors.

If we know the history of the qubit’s interaction with other qubits just before an erasure is detected then we may have to reset multiple ‘tainted’ qubits back to the codespace. This is pertinent when performing erasure detection after two-qubit gates.

The two-qubit gates are often the most error-prone operations (in this case, our dominant errors are presumed to be detectable leakage errors). In the most conservative case we perform mid-circuit erasure detection after every two-qubit gate. In this scenario, if leakage is detected on any one of the two qubits, we must reset both, so leakage errors only affect at most two physical qubits before they are reset. In this case, leakage has resulted in erasure errors on two physical qubits. After resetting them both to the codespace and using the erasure information in our decoder, we have successfully corrected the erasure error.

We can define other metrics such as the erasure conversion efficiency [Wu et al.,

2022] which tells us the probability our measurement successfully flags a qubit as having leaked, given that it is in fact in the leakage state. What happens if this efficiency is less than 100%? Let us suppose that instead it is 98%, a reasonable value to assume given experimental measurement fidelities. There is a 2% chance the leakage is missed on the leakage detection measurement, but a 98% chance it is detected on the second one and so the chances of a leakage error sneaking through undetected decreases exponentially with the number of measurements. Overall, this results in a small, constant fraction of leaked qubits that have not yet been detected and in the meantime are a source of Pauli errors. If we denote the erasure conversion efficiency as η_{conv} , and the probability a physical qubit undergoes erasure during each two qubit gate as p^{erasure} , then the fraction of leaked qubits (assuming leakage does not spread), p_{leak} , is given by the geometric series

$$p^{\text{leak}} = p^{\text{erasure}} \left((1 - \eta_{\text{conv}}) + (1 - \eta_{\text{conv}})^2 + (1 - \eta_{\text{conv}})^3 + \dots \right) \quad (3.9)$$

Where the n^{th} term of the series gives the probability a leaked qubit has survived n rounds undetected, which decreases exponentially. For η_{conv} close to 1, we can approximate

$$p^{\text{leak}} \approx p^{\text{erasure}} (1 - \eta_{\text{conv}}) \quad (3.10)$$

We can also define the ‘false positive’ rate of erasure conversion, which is the probability we incorrectly flag a qubit as having leaked when it is in fact in the computational subspace. For each erasure detection step we say the probability of this occurring is $p_{\text{false}}^{\text{erasure}}$. Ideally we want this probability to be somewhat small, since it sets the erasure rate in the absence of any actual leakage errors and instead due to our faulty measurement. But as we shall see next, codes such as the surface code are highly tolerant of erasure errors.

3.2 The Surface Code

So you want to build a fault-tolerant quantum computer? As of this writing, the surface code [Bravyi and Kitaev, 1998] is probably your best bet. In a distance d surface code, a $d \times d$ grid of data qubits is interspersed with a $d \times d$ grid of ancilla qubits as shown in Fig. 3.1 for a total of $2d^2 - 1$ physical qubits per logical qubit. The encoding rate of this code is not very high, especially when compared to classical error correcting codes. We need $O(d^2)$ physical qubits to encode just one logical qubit!

The surface code is divided up into plaquettes - 4 data qubits with an ancilla at the center. The sole purpose of the ancilla is to measure stabilizers on its surrounding data qubits. The outcomes of the stabilizer measurements are decoded to correct for Pauli errors. There are two types of plaquettes (unless we are using the $XZZX$ variation [Bonilla Ataides et al., 2021], in which case all plaquettes are identical) that alternate in a checkerboard pattern. The layout of the surface code is shown in Fig. 3.1.

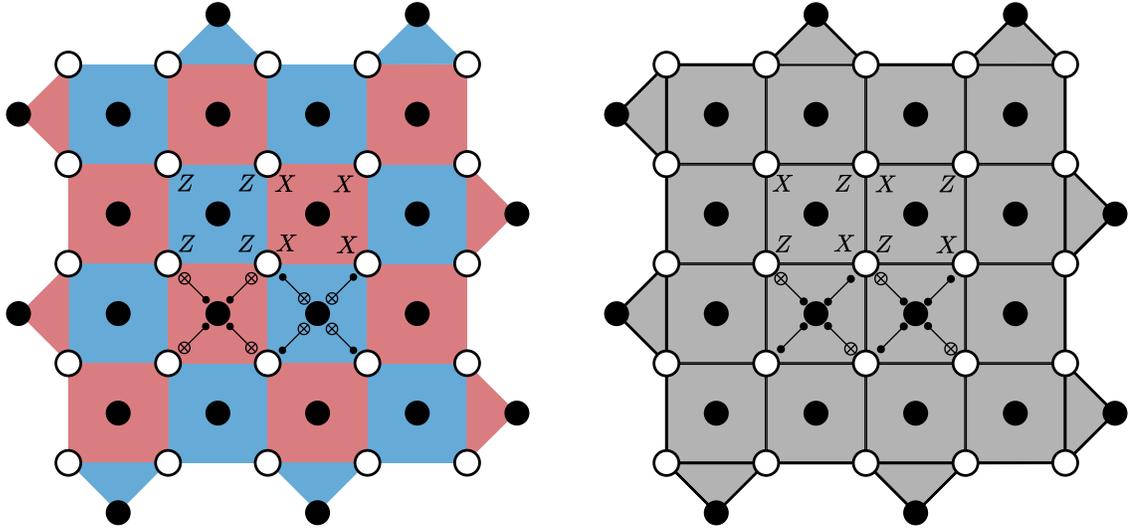


Figure 3.1: Physical qubit layout for the $d = 5$ (rotated) surface code. White circles are data qubits, black circles are ancilla qubits, for a total of 49 physical qubits per logical qubit. (Left) Traditional surface code. Ancillas in the blue (red) plaquette measure $ZZZZ$ ($XXXX$) stabilizers, except at the boundary where they measure weight two stabilizers. (Right) XZZX surface code. All plaquettes are identical and each ancilla measures its respective $XZZX$ stabilizer. Also shown are examples of the CX and CZ gates needed for the stabilizer measurement circuits.

Next we describe the circuits needed to measure stabilizers in variations of the circuit code. In the traditional surface code, we must measure both $XXXX$ and $ZZZZ$ stabilizers, which creates two different kinds of plaquettes. These are measured with the following circuits:

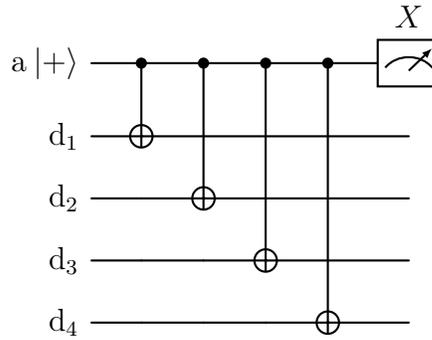


Figure 3.2: Ancilla qubits, a , at the center of X -type plaquettes measure the stabilizer $XXXX$ on the four surrounding data qubits, $d_{i=1,2,3,4}$.

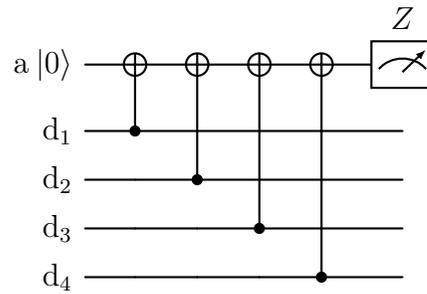


Figure 3.3: Ancilla qubits at the center of Z -type plaquettes measure the stabilizer $ZZZZ$ using a slightly different circuit with the control and target of the CNOT reversed

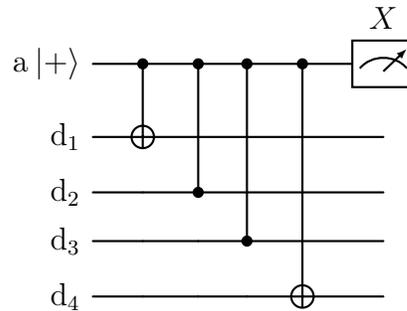


Figure 3.4: For the $XZZX$ surface code, the same stabilizer is measured on each plaquette and we only need one circuit which uses both CZ and CX gates.

The order of CZ and CX gates must be preserved in these circuits, as they are performed simultaneously with every ancilla qubit in the surface code. These weight-4 stabilizer measurements are the only operations needed to create and preserve a quantum memory in the surface code. The particularly nice properties of the surface code is that one only needs nearest-neighbour connectivity in a square lattice. All stabilizer measurements are local, and the distance of the surface code is extended by repeating the pattern of data and ancilla qubits to make a grid of larger size.

For computation, all Clifford gates in the surface code can be performed transversally via lattice surgery [Horsman et al., 2012, Litinski, 2019a] which makes their implementation fault tolerant. If we can stabilize a surface code memory, we have all the ingredients for fault-tolerant Clifford gates as well. The non-Clifford gate is more tricky. The Eastin-Knill theorem [Eastin and Knill, 2009] states that for a given quantum error correction code, it is impossible to perform a universal set of gates transversally (which guarantees they are fault-tolerant). In the case of the 2D surface code, it is the T gate that cannot be implemented transversally. The most popular workaround to implement T gates with error rates comparable to the other gates is to non-fault tolerantly prepare many logical $|T\rangle$ states and use ‘magic state distillation’ to reduce their errors followed by magic state injection to implement T gates. [Knill, 2005]. This procedure is expected to take up most of the computational time of a working quantum computer! At the physical qubit level, the only additional hardware operation is to perform a T -gate on (a small fraction of) the physical qubits⁶.

The last thing we need is at the classical level: a real-time decoder. Each round of stabilizer measurements produces many classical outcomes that need to be processed to deduce the most likely Pauli errors that led to the observed outcomes. This amounts to solving a hard graph matching problem, ideally before the next round

6. An alternative to magic state distillation is to use 3D surface codes for which the Eastin-Knill theorem does not apply. The just-in-time decoding scheme can achieve this by simulating a 3D surface code with a 2D qubit lattice

of error correction and definitely before any non-Clifford gates (See the ‘Backlog problem’ [Terhal, 2015]). The space of possible outcomes is so large that lookup tables quickly become unusable and we must use algorithms such as union find and Minimum-Weight-Perfect-Matching (MWPM) [Battistel et al., 2023].

3.2.1 Fault tolerance and the threshold theorem

The surface code is particularly promising because there exists a threshold for this code. Formally, this means as we concatenate our code with itself to increase the code distance, the logical error rate $p_L \rightarrow 0$, provided the physical qubit error rate is below a critical value, known as the threshold [Aharonov and Ben-Or, 1997], which we denote by p_{th} . We can define the Pauli error rate, p^{Pauli} (defined either *per gate* or *per EC cycle*⁷) and find that as we increase code distance, the logical error rate scales as

$$p_L \cong A \left(\frac{p^{\text{Pauli}}}{p_{\text{th}}} \right)^{\frac{d+1}{2}} \quad (3.11)$$

where A is a prefactor that scales weakly with distance and gives the error rate when $p^{\text{Pauli}} = p_{\text{th}}$. This prefactor is usually determined from Monte Carlo simulations. For the surface code, $A \approx 0.03$ and it is important that this formula only holds when we are well below the threshold and at large code distance. Additionally, only Pauli errors and a depolarizing noise channel is assumed throughout all operations, including ancilla qubit measurement and reset. Numerically, under this noise model $p_{\text{th}} \approx 1\%$ per gate. The threshold characterizes the overhead of quantum error correction with the surface code. We see it is much smaller than the 50% threshold for classical error correction! But achieving error rates below this threshold does not seem impossible in real systems. p_L decreases exponentially with code distance, so achieving p^{Pauli}

7. We also group ancilla measurement and reinitialization errors in the per gate errors. With four entangling gates per cycle, the per-cycle error is usually $\times 4$ the per-gate error.

significantly below threshold is important as it greatly reduces the code distance needed to realize a target p_L , and in turn the number of physical qubits per logical qubit. (Note: if we are close to threshold, we would not see any improvement at all if we increase code distance!)

We can plug some numbers in for context. For $p_L = 10^{-14}$, necessary for Shor’s algorithm [Gidney and Ekerå, 2021], with $p^{\text{Pauli}} = 0.1\%$, we would need $d = 27$ or 1,457 physical qubits per logical qubit (and some additional qubits for lattice surgery). That’s a large overhead! And because we need hundreds to thousands of logical qubits to perform useful algorithms, the estimated number of physical qubits required can stretch into the millions [Gidney and Ekerå, 2021]! Suppose we are a bit more optimistic, and say $p^{\text{Pauli}} = 0.01\%$, we would then need $d = 13$ or 337 physical qubits – about $4\times$ fewer physical qubits per logical qubits. A factor of 4 is a big deal. With the same number of physical qubits we could make 4 times the number of logical qubits, which increases our Hilbert space (and hence our computing power) by 2^4 . With leading industry devices surpassing the 400-qubit mark⁸, it may be possible to realize a single logical qubit, provided sufficiently low Pauli error rates can be achieved on such a device, the most daunting task.

3.2.2 Error hierarchy and structured noise

As noted by [Fowler et al., 2012] in their comprehensive study on the surface code, even when only Pauli errors are considered, the surface code has a strong hierarchy of errors. By this, we mean the logical error rate is more sensitive to some Pauli errors than others, depending on when they occur during the stabilization circuits. We can subdivide the Pauli errors into different classes, and each error class has a different per-gate threshold. We show values obtained from their simulations in our first error hierarchy example.

⁸. See IBM’s ‘Osprey’ processor.

Considering each error class in isolation, errors that occur on data qubits during idling (during ancilla measurement and reset) were found to have a 4.3% threshold. Ancilla measurement and reset errors were found to have a 12% threshold (close to 50% per cycle!), and finally the most damaging Pauli errors were during the two-qubit entangling gates, with only a 1.25% threshold.

How can we make use of such an error hierarchy? If uniformly reducing the Pauli error rates across all classes is not an option, then we should instead focus our efforts on reducing the error rates of the class that has the smallest threshold. In this example, achieving $< 1\%$ CNOT gate errors would be much more important than achieving $< 1\%$ ancilla readout, since we can suppress the effect of readout error by majority voting through multiple rounds of stabilizer measurements. In essence, we are engineering our physical error rates to ‘match’ the error hierarchy.

The error hierarchy should be thought of as a qualitative guide to improving errors. Consider two error classes (1) and (2). If $p_{\text{th}}^{(1)} \approx 5p_{\text{th}}^{(2)}$ then we should focus our efforts on reducing errors such that $p_{(1)}^{\text{Pauli}} \approx 5p_{(2)}^{\text{Pauli}}$ in order for each error to contribute similarly to the overall logical error rate p_L . We can stick to the rule of thumb

$$\frac{p_{(1)}^{\text{Pauli}}}{p_{\text{th}}^{(1)}} = \frac{p_{(2)}^{\text{Pauli}}}{p_{\text{th}}^{(2)}} \ll 1 \quad (3.12)$$

However, we want both error classes to be *well below* their respective isolated error thresholds, to ensure we are below threshold when we consider both error classes simultaneously as we would for a real system. No simple rule of thumb exists for estimating the overall threshold from isolated thresholds, e.g. $p_{\text{th}} \neq p_{\text{th}}^{(1)} + p_{\text{th}}^{(2)}$

3.2.3 Erasure errors in the surface code

The surface code is much more resilient to erasure errors than Pauli errors. The isolated per-cycle threshold to data qubit erasures is 50%. For erasures on both data

and ancilla qubits, the per-cycle threshold is $\approx 24.9\%$, set by the bond percolation threshold of a cubic lattice [Stace et al., 2009]. When we consider a more hardware-accurate erasure model, e.g., erasure checks after every two-qubit entangling gate, erasing both qubits, we find a per-gate threshold of 5.13% , roughly $5\times$ higher than the Pauli error threshold. This value was obtained by [Wu et al., 2022], and importantly, they found this threshold only drops to $\approx 4\%$ when finite erasure conversion efficiency (98%) and simultaneous Pauli errors ($\approx 10^{-4}$) per gate were considered. Once an erasure has been detected, we let the decoder ‘know’ by setting edge weights with erased qubits to zero when performing the decoding step. The favorable scaling and thresholds of a surface code with erasure and Pauli errors is shown in Fig. 3.5, with simulations by Yue Wu.

The higher threshold is convenient, but arguably the main benefit comes from the favorable scaling with code distance. The surface code tolerates up to $d - 1$ erasure errors in a round of stabilization, double the number of Pauli errors, and so

$$p_L \cong A \left(\frac{p^{\text{erasure}}}{p_{\text{th}}^{\text{erasure}}} \right)^d \quad (3.13)$$

To put this into perspective, with only erasure errors with 0.1% erasure error per gate, we can reach $\sim 10^{-14}$ error rates with $d = 8$ or 127 qubits, more than a factor of 10 fewer physical qubits!⁹

What may be more realistic to aim for is 1% erasure error rates and 0.01% Pauli error rates. With this error model, $d = 18$ (647 qubits) would be needed to reach 10^{-14} . Whilst the reduction in the number of qubits is not quite as dramatic, in the next section we show how this is achievable through error-detection at the hardware level.

What does it mean to align our error rates with the error hierarchy? This is slightly

9. with erasure errors, the code distance can be even.

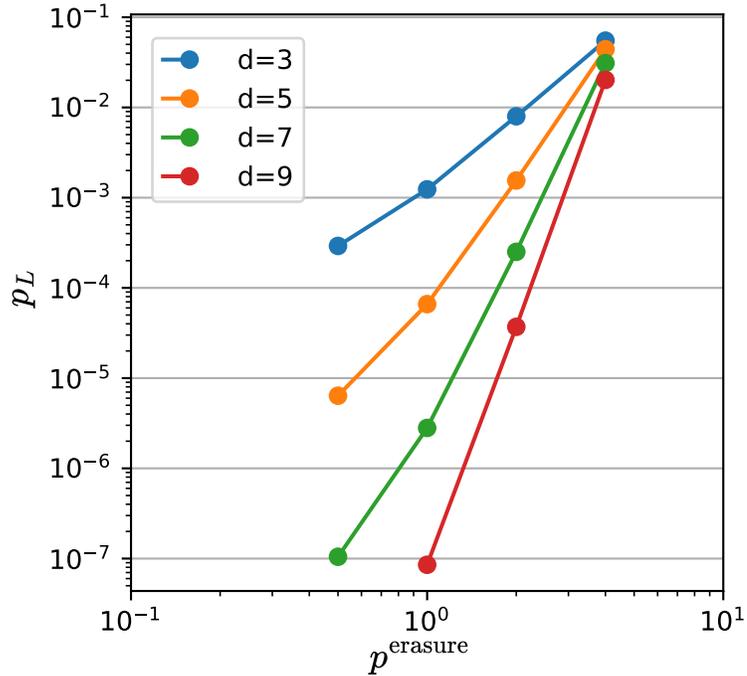


Figure 3.5: Logical error in a surface code with both Pauli and erasure errors, x axis in percent. Simulations by **Yue Wu**. For a fixed Pauli rate of 10^{-4} per gate (see [Wu et al., 2022] for the detailed error model). We find a threshold $p^{\text{erasure}} \approx 5\%$ and $A \approx 0.08$. This model assumes $\eta_{\text{conv}} = 1$ and mid-circuit erasure detection is performed after every two-qubit gate in an XZZX surface code. At $p^{\text{erasure}} = 1\%$, the error correction gain is $\Lambda \approx 20$ when we increase the distance from $d \rightarrow d + 2$. The value of Λ is hard to predict at larger code distances since it is set by both erasure and Pauli errors. At larger code distances, erasure errors become increasingly less relevant for the overall error rate, causing Λ to increase. Currently, predicting Λ vs code distance can be found only via numerical simulations. When estimating the distance required to reach a particular error rate, we make the assumption $\Lambda = 20$ and is constant vs distance which seems to give a lower bound.

different than with the Pauli error classes we examined earlier, because erasure errors effectively double the code distance. Instead, (assuming the prefactors A are equal) we should aim for

$$\frac{p^{\text{Pauli}}}{p_{\text{th}}^{\text{Pauli}}} = \left(\frac{p^{\text{erasure}}}{p_{\text{th}}^{\text{erasure}}} \right)^2 \ll 1 \quad (3.14)$$

For example, if our Pauli error rate is $100\times$ below threshold, our erasure rate can be just $10\times$ below threshold.

Figure 3.6 shows a multitude of different error classes including measurement, erasure and pauli errors and their respective threshold. This is not a direct comparison between classes, since the exact threshold depends on the specifics of the error model, but serves as a qualitative guide to illustrate the hierarchy of errors in the surface code.

3.2.4 Error detection enables the erasure error hierarchy

How can we engineer an error hierarchy like this in a physical system? And are such low Pauli rates physically realistic? The optimal hierarchy of erasures naturally arises when we attempt to detect (not correct!) Pauli errors at the hardware level. A striking new insight from this strategy is that we can now tolerate $\sim 1\%$ hardware error rates in our system and still achieve surface code performance well below threshold.

Consider a set of qubit hardware that suffers from (first-order) jump errors at rate p^{phys} . These jump errors traditionally set our Pauli error rate per gate i.e. $p^{\text{Pauli}} = p^{\text{phys}}$. The strategy is to engineer our gates and physical qubits such that all first-order jump errors can be detected after each gate or measurement operation, and converted into erasure errors at rate $p^{\text{erasure}} = p^{\text{phys}}$ per gate. The Pauli error rate is then set by the much smaller remaining undetected second order (and higher) jump errors at rate $O((p^{\text{phys}})^2)$ per gate, assuming perfect erasure conversion efficiency. Error detection naturally gives us a hierarchy of erasure and Pauli errors that is

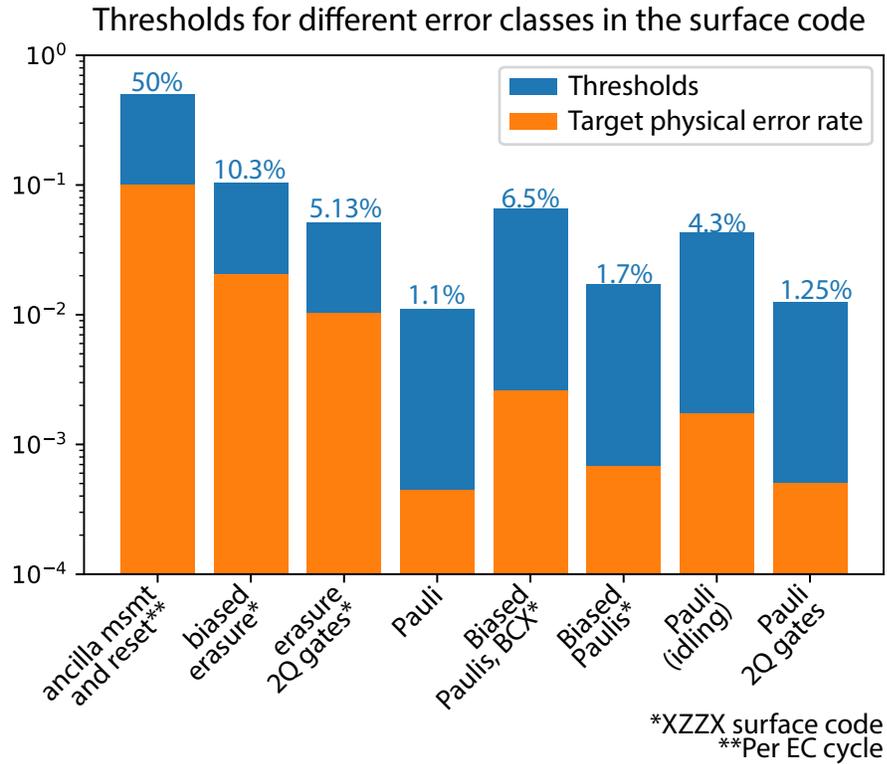


Figure 3.6: Resilience of the surface code to different error classes. In blue we show thresholds (per-gate unless noted) as determined from the different error models in [Fowler et al., 2012, Sahay et al., 2023, Wu et al., 2022, Wang et al., 2011, Darmawan et al., 2021] where each class of error was considered in isolation. In orange we show the physical error rate required to be a factor of 25 below threshold (factor of 5 below threshold for the first three errors since $p_L \propto p^d$). The chosen physical error rates are used to highlight the susceptibility to different error classes in a speculative surface code. We observe that the surface code (as expected for any error correcting code) is more resilient to structured noise (measurements, erasures, biased-Paulis) and the least resilient to unstructured noise, e.g., Pauli errors/depolarizing noise, which is reflected in the thresholds. BCX refers to bias-preserving control-X gates and Biased Paulis are where all Pauli errors are Z errors but gates are not bias-preserving.

aligned to the hierarchy of errors in the surface code and any other stabilizer code.

To achieve this our erasure error detection operation must be extremely good with respect to its ‘half QND-ness’. If no error occurred, we do not want our error detection to introduce Pauli errors at a rate greater than $O((p^{\text{phys}})^2)$, otherwise the act of error detection will limit the Pauli error rate!. The erasure detection also needs to have high efficiency (how many erasures do we miss?) and low false positive rates (how many good qubits do we erase?) since both of these set the minimum erasure rate.

This is the main motivation behind why we should be engineering error-detected gates at the hardware level rather than fully error corrected ones. Ironically, by forgoing full error correction at the hardware level, the task of error correction at the logical level now looks much easier!

For the surface code specifically, after error detection and erasure conversion we can expect

$$\begin{aligned}
 p_L^{\text{erasure}} &= \left(\frac{p^{\text{phys}}}{p_{\text{th}}^{\text{erasure}}} \right)^d \\
 p_L^{\text{Pauli}} &= \left(\frac{O((p^{\text{phys}})^2)}{p_{\text{th}}^{\text{Pauli}}} \right)^{\frac{d+1}{2}} \\
 p_L &\propto (p^{\text{phys}})^d
 \end{aligned} \tag{3.15}$$

Compare this to a physical qubit where we have error *corrected* all jump errors to first-order, which would suffer from just Pauli errors at reduced rate $O((p^{\text{phys}})^2)$ per gate. The logical error rate in this surface code would also scale as $p_L \propto (p^{\text{phys}})^d$. Up to a constant of proportionality (which is similar or even smaller for the error-detected case!), the logical error rate is the same. Due to the relaxed Knill-Laflamme conditions for erasure errors, we generally expect error detection of jump errors at the hardware level to always be easier to implement than full error correction.

3.3 Bosonic Quantum Error Correction

Bosonic QEC is a Hardware way to achieve the redundancy required for quantum error correction or detection at the hardware-level. Rather than increasing the size of our Hilbert space through additional physical qubits, we can use more levels within a single physical qubit. The premise of bosonic error correction is to use Quantum Harmonic Oscillators (QHO) in place of two-level systems¹⁰ as our physical qubit. QHOs have an infinite number of energy levels! (at least in theory.) Qubit information is encoded in a subspace of the QHO, the computational subspace and error processes take us to another subspace within the oscillator, which we can use as our orthogonal error subspace. Correcting or detecting errors brings the qubit back to the computational subspace. In this sense, bosonic QEC is very similar to standard error correction with stabilizer codes.

This approach to error correction/detection is often presented as being hardware efficient—we can perform the first layer of error correction without needing to introduce more physical qubits. Moreover, the error channels in a QHO can be incredibly simple and are often well-described by the photon loss channel. The collapse operator is the annihilation operator \hat{a} that characterizes energy loss processes in all the energy levels at rate $\kappa = 1/T_1$. For the case when the QHO is a microwave cavity mode, there are very few sources of intrinsic dephasing¹¹.

But one should always feel a bit uneasy when encoding qubits in an oscillator. True QHOs are purely linear systems, and so the thought of using them as qubits seems almost paradoxical [Nielsen and Chuang, 2010]. A consequence of the equal spacing between energy levels is that Gaussian control (the easiest to engineer in practice) by itself is not enough to prepare quantum states in QHOs or perform

10. Not to be conflated with TLS systems that cause $1/f$ noise in transmons!

11. A cavity can inherit dephasing errors from many sources such as energy jumps in a dispersively coupled transmon.

quantum error correction¹². The typical example to illustrate this is a displacement drive with Hamiltonian $\hat{\mathcal{H}}_d = \epsilon_d(t)\hat{a} + \epsilon_d(t)^*\hat{a}^\dagger$, which can only create displaced states of the form $|\alpha(t)\rangle$ in a QHO. By contrast, the same drive applied to a two-level system can perform universal single qubit gates.

The importance of this trade-off cannot be understated. For what we gain in Hardware error correction and a simplified error channel we risk losing entirely in our lack of quantum control over the states in our bosonic modes.

3.3.1 Quantum control of bosonic modes

The paradigm that reconciles these opposing requirements is to introduce ‘weak’ non-linearity into QHO by weakly coupling a QHO to a highly non-linear system such as a two-level system. We refer to this two level system as the ‘control ancilla’. Two resources arise that now allow quantum control of a QHO through ‘classical’ control drives. The first is that the QHO is no longer an exact harmonic oscillator and inherits a weak anharmonicity that perturbs the equal spacing between energy levels, but its error channel remains well-described by photon loss, \hat{a} . (This is true when the QHO’s loss rate is larger than its anharmonicity, and the environment is unable to tell which energy level of the cavity omitted the photon.) The second is that the coupling Hamiltonian between the QHO and control ancilla is also itself non-linear, and we can achieve quantum control by manipulating the control ancilla [Krastanov et al., 2015, Heeres et al., 2017]. Unlike ancillas that measure stabilizers, which are qubit-like, the control ancilla exists solely as an auxiliary system that broadens our control of a QHO. In order to maximize the speed of our operations, it is more common to use this non-linear coupling for quantum control, rather than the inherited anharmonicity.

However, the non-linear coupling will also be weaker than the non-linearity of the

12. In GKP codes, we can perform all Clifford gates with Gaussian operations, but not state preparation, stabilization and non-Clifford operations.

control ancilla, and so we can manipulate states in the ancilla faster than states in the QHO. When designing operations on QHOs with a control ancilla, a question we should always keep in the back of our minds is, ‘*why not just use the control ancilla as the physical qubit instead?*’ In Chapter 6, we show that we have good prospects to overcome this ‘gate time’ overhead with the dual-rail encoding.

3.3.2 Propagation of ancilla errors

The non-linear coupling between a QHO and control ancilla is the bridge that allows quantum control, enabling the state preparation, measurements and unitaries needed for bosonic QEC and gates. However, this bridge also allows jump errors in the control ancilla to propagate onto the QHO as Pauli errors (or worse!), threatening to undermine all of our error correction efforts. Solutions to this problem rely on the fact that the control ancilla is not a qubit. This means we can engineer redundancy to protect from errors without needing to satisfy the full Knill-Laflamme conditions. We will see later that the 0-2 encoding makes a bad bosonic QEC code but is a viable subspace for a (three-level) control ancilla.

The way to prevent the propagation of these errors is through use of a ‘noise-biased’ ancilla. Although non-trivial to realize, this is based on the observation that the non-linear coupling always takes the form

$$\mathcal{H}_{\text{nl}} = f(\hat{a}, \hat{a}^\dagger) \hat{O}_{\hat{q}} \quad (3.16)$$

An example of this is the dispersive interaction for which

$$\mathcal{H}_{\text{nl}} = \chi_{aq} \hat{a}^\dagger \hat{a} \hat{q}^\dagger \hat{q} \equiv -\frac{\chi_{aq}}{2} \hat{a}^\dagger \hat{a} \hat{Z}_q \quad (3.17)$$

for control ancilla mode, \hat{q} . Errors of the form $\hat{O}_{\hat{q}}$ acting on our ancilla do not propagate onto the bosonic mode, because they commute with \mathcal{H}_{nl} i.e. $([\mathcal{H}_{\text{nl}}, \hat{O}_{\hat{q}}] = 0)$

and are ‘error transparent’. For the dispersive interaction example, a noise-biased ancilla which only suffers from \hat{Z}_q errors during the control Hamiltonian but not \hat{X}_q errors will not propagate any errors to the bosonic mode. We can achieve this either by engineering an ancilla with true noise bias [Puri et al., 2019], or through use of a three-level ancilla [Rosenblum et al., 2018b, Reinhold et al., 2020]. The purpose of Hardware Fault-Tolerance (HFT) is to design operations with a control ancilla that prevents the propagation of these errors and to ensure they remain detectable/correctable after the operation.

In summary, current bosonic QEC realizations are in a constant state of tension. The introduction of ancillas for control means we are not quite as hardware-efficient anymore, and we are now susceptible to a larger set of errors. In fact, the control ancilla errors often occur at a higher rate than the bosonic errors we are trying to correct for! Control ancillas are by no means trivial and require a noise-bias to be engineered either directly or through redundant levels in the ancilla mode. The exact implementation is the subject of HFT. We will see that with this hybrid system, achieving full bosonic error correction is not impossible, but incredibly difficult. However, *error-detection* of both control ancilla and bosonic errors is much easier to realize. We now discuss bosonic error channels and specific bosonic error correction schemes in more detail.

3.3.3 The photon loss channel for a bosonic mode

The photon loss channel is often the dominant error channel and is described by collapse operator \hat{a} with rate κ_a . Evolution under the photon loss channel is readily calculated from the Lindblad master equation but it is also intuitive to examine the Kraus map for this error [Michael et al., 2016]. For time interval δt we find

$$\hat{E}_n = \sqrt{\frac{(1 - e^{-\kappa\delta t})^n}{n!}} e^{-\frac{\kappa\delta t}{2}\hat{a}^\dagger\hat{a}} \hat{a}^n \quad (3.18)$$

for $n = 0, 1, 2, 3, \dots$. This formalism lets us order our error channel by the number of photon jumps. $n = 0$ corresponds to no photon loss, but the system still experiences no-jump backaction. This amounts to weak measurement of the photon number, which polarizes a superposition of Fock states towards states with lower photon number. To see how this arises consider the Kraus operator $\hat{E}_0 = e^{-\frac{\kappa\delta t}{2}\hat{a}^\dagger\hat{a}}$ acting on the state $|\psi\rangle = c_0|0\rangle + c_1|1\rangle$. We find $\hat{E}_0|\psi\rangle \propto c_0|0\rangle + c_1e^{-\frac{\kappa\delta t}{2}}|1\rangle$, which decreases the amplitude of the $|1\rangle$ state relative to the $|0\rangle$ state, polarizing us towards vacuum.

\hat{E}_1 is the Kraus operator for single photon loss, which occurs with probability $n\kappa\delta t$ for small δt for Fock state $|n\rangle$ in the QHO. Double photon loss is a less likely, second-order error¹³ and occurs with probability $\approx n(n-1)(\kappa\delta t)^2/2$ for Fock state $|n\rangle$ and $n \geq 2$. Discrete variable (DV) bosonic QEC codes are designed to correct for a discrete number of photon loss events. Continuous variable (CV) encodings are bosonic codes that can correct for continuous errors such as small displacements of an oscillator [Gottesman et al., 2001]. We do not discuss CV encodings much here but, as with all bosonic error correction, a non-linear control resource is required and CV encodings suffer from the same control ancilla limitations as DV encodings [Campagne-Ibarcq et al., 2020, Sivak et al., 2023].

The overhead of the bosonic QEC is similar to the overhead of a stabilizer encoding. Rather than having more physical qubits which results in the probability of a single error that scales linearly with the number of physical qubits, encoding in higher levels of an oscillator for the required redundancy to photon loss means the probability of a photon loss is greater, because the n^{th} level of the QHO has decay rate $n\kappa$. We call this the \bar{n} overhead since the average decay rate is linear in the average number of photons \bar{n} in the encoding.

For the simplest 0-1 encoding, $\bar{n} = 0.5$. The smallest code that can exactly correct for single photon loss in a bosonic mode is the Kitten code, with $\bar{n} = 2$, a factor of 4

13. second order in $(\kappa\delta t)^2$

overhead. The \bar{n} overhead (and increasing control complexity) makes it impractical to increase the code distance (in this context the number of photon jumps we can correct for) within a single oscillator indefinitely, which requires increasing \bar{n} . Instead, the approach is to perform some degree of bosonic QEC that reduces our Pauli error rate due to photon loss to $O((\kappa\delta t)^2)$ and then concatenate with a stabilizer code such as the surface code to achieve fault tolerant computation.

The procedure for correcting jump errors in DV bosonic codes is to first perform a QND measurement that tells us whether the jump error occurred. This is the bosonic version of a stabilizer measurement. Conditioned on this outcome, we perform the appropriate unitary that either corrects for the jump error or the no-jump error. This whole procedure can be made ‘autonomous’ [Gertler et al., 2021, Li et al., 2023] by applying conditional correction unitaries followed by reset of the control ancilla, which is ultimately a dissipative process. (Entropy must be evacuated from the system one way or another—either by measurement or engineered dissipation!)

Just as we saw for two-level qubits, the system may have a non-zero temperature. In the case of a bosonic mode, $n_{\text{therm}} \neq 0$ results in photon loss described by collapse operator \hat{a} at rate κ_{\downarrow} and photon gain described by collapse operator \hat{a}^{\dagger} at rate κ_{\uparrow} . Once again we find (for $n_{\text{therm}} \ll 1$):

$$n_{\text{therm}} = \frac{\kappa_{\uparrow}}{\kappa_{\downarrow} + \kappa_{\uparrow}} = \kappa_{\uparrow} T_1 \quad (3.19)$$

bosonic codes designed to correct for single photon loss are not able to correct (but can detect) single photon gain. Since $n_{\text{therm}} < 1\%$ is typical in the microwave cavities we consider, single photon gain occurs at a similar rate to double photon loss and occurs at a similar rate to ‘second-order’ error processes.

3.3.4 Dephasing in a bosonic mode

The other bosonic error to consider is dephasing described by collapse operator $\hat{a}^\dagger \hat{a} = \hat{n}$ where \hat{n} is the number operator. Dephasing has a strong ‘lever-arm’ effect in a bosonic mode [Eickbusch et al., 2022]. Under this error channel the superposition $(|0\rangle + |2\rangle)/\sqrt{2}$ will dephase *four times* faster than the superposition $(|0\rangle + |1\rangle)/\sqrt{2}$, which can be seen by considering the jump evolution in the Lindblad master equation, given by $\kappa \hat{n} \rho \hat{n}$, hence the n^2 dependence of the dephasing rate. Fortunately, the bosonic modes found in cQED platforms in the form of microwave cavities are rather unique in the sense that their intrinsic dephasing error is so far immeasurably small [Rosenblum et al., 2018b, Sivak et al., 2023]. We say that microwave cavities have a ‘natural’ noise-bias, since the dominant error channel is photon loss. Such dephasing mechanisms would need to alter the capacitance or inductance of the oscillator, which involves changing the relative dimensions of the metallic conductors that make up the resonator.

The same cannot be said for bosonic motional degrees of freedom found in trapped ion or neutral atom platforms where laser/trap stability sets the dephasing rate. In microwave cavities, any dephasing errors are instead inherited from non-linear elements coupled to our system (e.g. energy jumps in a dispersively coupled control ancilla) and so should not be modelled as white-noise dephasing but rather as low frequency noise.

We now give an overview of specific DV bosonic codes for error correction and detection.

3.3.5 Cat codes

2-legged cat code

Cat codes are formed from superpositions of coherent states [Cochrane et al., 1999, Mirrahimi et al., 2014]. It is the quantum version of classical phase-keying codes. A coherent state can be written in the Fock basis as

$$|\alpha\rangle = e^{-\frac{|\alpha|^2}{2}} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle \quad (3.20)$$

The simplest is the 2-legged cat code with codewords

$$\begin{aligned} |0_L\rangle &= \frac{|\alpha\rangle + |-\alpha\rangle}{\sqrt{2(1 + e^{-2|\alpha|^2})}} \\ |1_L\rangle &= \frac{|\alpha\rangle - |-\alpha\rangle}{\sqrt{2(1 - e^{-2|\alpha|^2})}} \end{aligned} \quad (3.21)$$

parameterized by α . In the limit $\alpha \rightarrow \infty$, $\bar{n} = |\alpha|^2$ for both codewords. Deviations from this scale as $e^{-2|\alpha|^2}$ and at $\alpha = 2$ this deviations are on the 10^{-4} level. From the Knill-Laflamme conditions, we are unable to correct photon loss errors via e.g., photon number parity measurements – $|0_L\rangle$ contains an even number of photons whereas $|1_L\rangle$ contains an odd number of photons. In principle, we can correct for cavity dephasing errors of the form $e^{i\theta\hat{n}}$ for $\theta \in \{-\pi/2, \pi/2\}$. In practice this is achieved by stabilizing a meta-potential in the oscillator (through dissipative [Mirrahimi et al., 2014] or Hamiltonian processes [Puri et al., 2017, Grimm et al., 2020]), with minima centered at α and $-\alpha$ in the oscillator’s phase space. This stabilized bosonic qubit shows a strong noise-bias and may itself be useful as a noise-biased control ancilla [Puri et al., 2019]. Without stabilization, no-jump backaction causes loss of energy and $\alpha \rightarrow \alpha e^{-\kappa t}$, altering the definition of the codewords.

4-legged cat code (4-Cat code)

The 4-Cat code is defined from superpositions of four coherent states:

$$\begin{aligned} |0_L\rangle &= |C_0^\alpha\rangle = \frac{|\alpha\rangle + |i\alpha\rangle + |-\alpha\rangle + | -i\alpha\rangle}{N_0} \\ |1_L\rangle &= |C_2^\alpha\rangle = \frac{|\alpha\rangle - |i\alpha\rangle + |-\alpha\rangle - | -i\alpha\rangle}{N_2} \end{aligned} \quad (3.22)$$

with normalization factors N_2 and N_0 . Both the $|0_L\rangle$ and $|1_L\rangle$ codewords contain an even number of photons. The codewords are distinguishable by their 4-parity (or superparity). By this we mean $n \bmod 4 = 0$ for the photon number distribution of $|0_L\rangle$ and $n \bmod 4 = 2$ for $|1_L\rangle$. A single photon loss takes us to the error space

$$\begin{aligned} |0_E\rangle &\propto \hat{a} |C_0^\alpha\rangle = |C_3^\alpha\rangle = \frac{|\alpha\rangle + i|i\alpha\rangle - |-\alpha\rangle - i|-i\alpha\rangle}{N_3} \\ |1_E\rangle &\propto \hat{a} |C_2^\alpha\rangle = |C_1^\alpha\rangle = \frac{|\alpha\rangle - i|i\alpha\rangle - |-\alpha\rangle + i|-i\alpha\rangle}{N_1} \end{aligned} \quad (3.23)$$

These codewords all have odd photon number parity. As such, a QND photon number parity measurement can detect single photon loss events. Double photon loss \hat{a}^2 results in an uncorrectable X error on the original codewords. Explicitly, this means $\hat{a} |0_E\rangle \propto |1_L\rangle$ and $\hat{a} |1_E\rangle \propto |0_L\rangle$.

A convenient property of the four-legged cat code is that the error space is also a valid error correcting code, with the same error correcting properties. Similar to our initial codewords, the error codewords are distinguishable by their 4-parity. $n \bmod 4 = 3$ for $|0_E\rangle$ and $n \bmod 4 = 1$ for $|1_E\rangle$. Rather than performing a correction unitary, after detecting photon loss we can instead update the definition of our logical basis from $\{|0_L\rangle, |1_L\rangle\}$ to $\{|0_E\rangle, |1_E\rangle\}$. This holds true for subsequent photon jumps. We can correct for single photon loss jumps by tracking them with repeated QND photon number parity measurements, provided we perform these measurements faster than the rate of photon jumps (i.e., intervals shorter than T_1/\bar{n}). Further subtleties regarding the choice of α in the 4-Cat code are discussed in Appendix B.

3.3.6 The kitten code

The kitten code achieves similar benefits to the 4-legged cat code, without as large a \bar{n} overhead. It is the lowest order Binomial code [Michael et al., 2016] defined with codewords

$$\begin{aligned} |0_L\rangle &= \frac{|0\rangle + |4\rangle}{\sqrt{2}} \\ |1_L\rangle &= |2\rangle \end{aligned} \tag{3.24}$$

and is able to correct for a single photon loss error, which takes codewords to the error states

$$\begin{aligned} |0_E\rangle &= |3\rangle \\ |1_E\rangle &= |1\rangle \end{aligned} \tag{3.25}$$

Like the 4-legged cat code, single photon loss is detected through QND photon number parity measurements but unlike the 4-legged cat code, the error states do not form a good new basis since $\Delta\bar{n} = 2 \neq 0$. Upon detecting odd parity or even, we should actively apply the correction unitaries¹⁴

$$\begin{aligned} \hat{U}_E &= \left(\frac{|0\rangle + |4\rangle}{\sqrt{2}} \right) \langle 3| + |2\rangle \langle 1| + \text{h.c.} \\ \hat{U}_L &= e^{2\kappa\delta t(|4\rangle\langle 0| - |0\rangle\langle 4|)} \end{aligned} \tag{3.26}$$

No-jump backaction is also rather annoying for the Kitten code. Although a second-order effect, when we do not detect any photon jumps, the $|1_L\rangle$ codeword polarizes towards the vacuum state $|0\rangle$. This is the rare case where no-jump backaction can be corrected (to first-order in δt) with a unitary operation, in this case, a rotation

14. Technically, these do not have to be unitaries, as long as they are linear maps that take the error states back to the initial codewords

in the $\{|0\rangle, |4\rangle\}$ manifold of the cavity. The \hat{U}_L unitary must be implemented with high fidelity since any errors here will result in a first-order error on the logical qubit. By contrast, any errors during the \hat{U}_E unitary will be second order on the logical qubit, since we require both a photon loss error and a correction unitary error to occur. The related two-mode [Chuang et al., 1997] version of this code does not suffer from no-jump backaction when the loss rates of both modes are equal. The complexity of the correction unitaries and state preparation means one must use numerically optimized pulses to implement this code, which are themselves prone to propagating ancilla errors to the bosonic mode.

3.3.7 The dual-rail code

The dual-rail (DR) code is defined with two modes per physical qubit, with codewords

$$\begin{aligned} |0_{\text{DR}}\rangle &= |01\rangle = |0\rangle_1 |1\rangle_2 \\ |1_{\text{DR}}\rangle &= |10\rangle = |1\rangle_1 |0\rangle_2. \end{aligned} \tag{3.27}$$

The two modes may be bosonic or two-level systems. This code has been used extensively in linear quantum optics for more than two decades [Chuang and Yamamoto, 1995, Knill et al., 2001]. In this context, and for the case we are interested in with microwave cavities, the loss channel is now photon loss in either bosonic mode, with collapse operators \hat{a}_1 and \hat{a}_2 and rates κ_1 and κ_2 , and with negligible dephasing errors. In the case $\kappa_1 = \kappa_2$, the error channels satisfy the Knill-Laflamme conditions for erasure error and both error channels take both codewords with equal probability to the ground state $|00\rangle$, a leakage state.

We cannot error-correct photon loss in a dual-rail qubit. The goal is instead to convert all photon loss errors to erasure errors and then concatenate physical dual-rail qubits with a stabilizer code to correct the erasure errors. Detection of the leakage error followed by reset back into the dual-rail codespace allows us to

realize erasure conversion. This requires performing a binary measurement described by projectors $\hat{\Pi}_{\text{DR}} = |0_{\text{DR}}\rangle\langle 0_{\text{DR}}| + |1_{\text{DR}}\rangle\langle 1_{\text{DR}}|$ and $\hat{\Pi}_{\text{Leakage}} = |00\rangle\langle 00|$. The Half-QND requirement necessitates that the measurement is QND when we obtain the $\hat{\Pi}_{\text{DR}}$ outcome. Implementing the dual-rail encoding in cQED with superconducting cavities [Teoh et al., 2022] is the main focus of Chapter 6.

When $\kappa_1 - \kappa_2 = \Delta\kappa \neq 0$, this encoding is susceptible to no-jump backaction. If $\kappa_1 > \kappa_2$, for short times, this polarizes codewords towards the $|0_{\text{DR}}\rangle$ state at rate $O((\Delta\kappa\delta t)^2)$ and can be mitigated e.g. by periodically applying X gates as we saw for the qubit case.

The strength of the dual-rail code lies in its simplicity. Compared to other bosonic codes, the unitary operations required for state preparation, readout, leakage detection and gates are much easier to engineer, largely in part to its simple codewords and gates that are excitation preserving. (There is always a single photon in one mode or the other.)

3.3.8 The 0-2 code and biased erasures

An error-detecting code with

$$\begin{aligned} |0_L\rangle &= |0\rangle \\ |1_L\rangle &= |2\rangle \end{aligned} \tag{3.28}$$

Photon loss is described by collapse operator \hat{a} . Single photon loss takes us to the leakage error state $|1\rangle$ whereas double photon loss is a second order error that takes $|1_L\rangle$ to $|0_L\rangle$. At first glance, the 0-2 code has advantages over the dual-rail code. Only one mode rather than two is needed per physical qubit, making this encoding more hardware efficient. first-order photon loss in this code is an example of a ‘biased-erasure’ error—only one of the codewords ($|2\rangle$) can decay to the leakage state. When we detect $|1\rangle$, we know the state must have previously been $|2\rangle$.

Biased erasures are even easier to correct than normal erasures [Sahay et al., 2023]. To illustrate this, we can fully correct a biased erasure error by concatenating with the 2-bit phase flip detection code such that

$$\begin{aligned} |0_L\rangle &= |++\rangle = \frac{1}{2}(|0\rangle_1 + |2\rangle_1)(|0\rangle_2 + |2\rangle_2) \\ |1_L\rangle &= |--\rangle = \frac{1}{2}(|0\rangle_1 - |2\rangle_1)(|0\rangle_2 - |2\rangle_2) \end{aligned} \quad (3.29)$$

A general qubit will be in the state

$$|\psi_L\rangle = c_0 |++\rangle + c_1 |--\rangle \quad (3.30)$$

After a photon loss event, we will detect one of the modes to be in $|1\rangle$. Suppose this is mode 1, then we will be in the state

$$|\psi_L\rangle = c_0 |1\rangle_1 |+\rangle_2 - c_1 |1\rangle_1 |-\rangle_2 \quad (3.31)$$

If the loss is detected in mode 2 we will be in mode

$$|\psi_L\rangle = c_0 |+\rangle_1 |1\rangle_2 - c_1 |-\rangle_1 |1\rangle_2 \quad (3.32)$$

where we see the qubit information (i.e. the coefficients) is preserved in either case. Extending further to the 3-bit phase flip code allows the correction of single biased erasure and a single phase flip error, which still requires fewer physical qubits than the 5-bit perfect code.

Biased erasures typically arise when only one of the codewords can decay to the leakage state, but this is not strictly necessary. For an erasure to be ‘biased’ we must gain information as to *which* codeword decayed to the leakage state. When photon loss decay happens in the dual-rail code, the environment knows which cavity had the photon loss (the photon’s energy must be absorbed *somewhere*) but with the standard

leakage detection measurement this information is lost.

A side-by-side comparison shows the 0-2 encoding is more hardware efficient than the DR code. To correct for a single photon loss error we require two physical 0-2 qubits and two bosonic modes. In comparison, the dual-rail code requires four physical DR qubits and eight bosonic modes!

The 0-2 encoding does have some major drawbacks:

- Gates must be performed in the $\{|0\rangle, |2\rangle\}$ manifold and are *not* excitation preserving, making implementation difficult in a bosonic mode.
- No-jump backaction is more pronounced, $\Delta\bar{n} = 2$ and so $\Delta\kappa = 2\kappa$. This is a consequence of not exactly satisfying the Knill-Laflamme conditions for the erasure channel. (Can be mitigated with X gates)
- Double photon loss is a possible error and results in a logical error (a Pauli error if we concatenate with a stabilizer code.)

Control in the $\{|0\rangle, |2\rangle\}$ is easier in more anharmonic systems, such as three-level transmons. However, systems with sufficiently strong anharmonicity often have non-negligible dephasing errors that need to be corrected. These states are used in three-level control ancilla as well, where the ability to detect decay to $|1\rangle$ allows us to error-detect control ancilla errors.

3.3.9 The 0-1 code

This code is also called the Fock encoding or single-rail encoding and has the lowest possible \bar{n} overhead with $\bar{n} = 0.5$ but no way to detect photon loss at the hardware level. The codewords are

$$\begin{aligned} |0_L\rangle &= |0\rangle \\ |1_L\rangle &= |1\rangle \end{aligned} \tag{3.33}$$

This is the code we compare other codes against in breakeven quantum memory experiments, since for the amplitude damping channel, this encoding has the minimal number of photons and hence the slowest possible decay rate. Arguably the simplest code with one mode per physical qubit, when the mode is bosonic, single qubit gates can be surprisingly hard to implement due to the linearity of the bosonic mode which makes it difficult to confine manipulation to the $\{|0\rangle, |1\rangle\}$ manifold. In this respect, the DR encoding is advantageous.

3.3.10 Wigner functions

The Wigner function can be thought of as the ‘phase space portrait’ of a bosonic state. The Wigner function is a quasi-probability function defined for a pure or mixed bosonic state that has the property:

$$|\psi_x(x)|^2 = \int_{-\infty}^{\infty} W(x, p) dp \quad (3.34)$$

$$|\psi_p(p)|^2 = \int_{-\infty}^{\infty} W(x, p) dx \quad (3.35)$$

Where $\psi_x(x)$ is the representation of the QHO state in the x basis and $\psi_p(p)$ is the representation of the QHO state in the p basis. Often we divide the coordinates x and p by their respective zero-point energy fluctuations for the QHO such that x and p become dimensionless coordinates which we refer to as the I and Q quadratures respectively. The argument of the Wigner function can be replaced by β , a complex scalar.

The Wigner function is also something that can be measured directly in experiment and is vital for performing state tomography on a QHO. This is achieved by measuring the expectation value of the ‘displaced parity operator’, i.e.,

$$W(\beta) = \frac{2}{\pi} \text{Tr} \left(\hat{D}_{\beta} \rho \hat{D}_{\beta}^{\dagger} \hat{P} \right), \quad (3.36)$$

where $\hat{P} = e^{i\pi\hat{a}^\dagger\hat{a}}$ is the photon number parity operator. In an experiment this means displacing the QHO by amplitude β , measuring the photon number parity and averaging the results for many repeated copies of the state in order to measure a single point of the Wigner function at $W(\beta)$. We then repeat the procedure scanning over different values of β to form a 2D ‘portrait’ of the state that can quantitatively be analyzed. The Wigner function contains complete information about the state of the QHO and from this data we can reconstruct the full density matrix. In Fig. 3.7 and 3.8 we show the simulated Wigner functions for various bosonic codes where properties like rotational symmetry become apparent.

Wigner functions are also helpful to illustrate what we mean by ‘classical’, ‘Gaussian’ and ‘quantum’ control. We define ‘classical’ control to mean control that transforms coherent states into other coherent states, i.e., displacements, rotations and beamsplitter couplings. We define Gaussian control to be unitaries that transform a state with a 2D Gaussian Wigner function into other states which also have a (different) 2D Gaussian Wigner function. This includes all of classical control but also single-mode and two-mode squeezing. Gaussian control on Gaussian states by itself is not sufficient for universal quantum computing.

Lastly, quantum control refers to unitaries that can create negativity in the state’s Wigner function, if we started from an only positive Wigner function. The amount of coherence in a superposition of states appears as ‘fringes’ in the Wigner function, which are sub-Planckian ($< \frac{1}{2}$) in width and contain negative regions. If decoherence leaves us in a completely mixed state, we will no longer observe any negative regions and the fringes will have faded out. Visually inspecting Wigner functions can thus give us a quick ‘diagnosis’ of the QHO state and are a helpful visual aid as well as a means for more quantitative analysis.

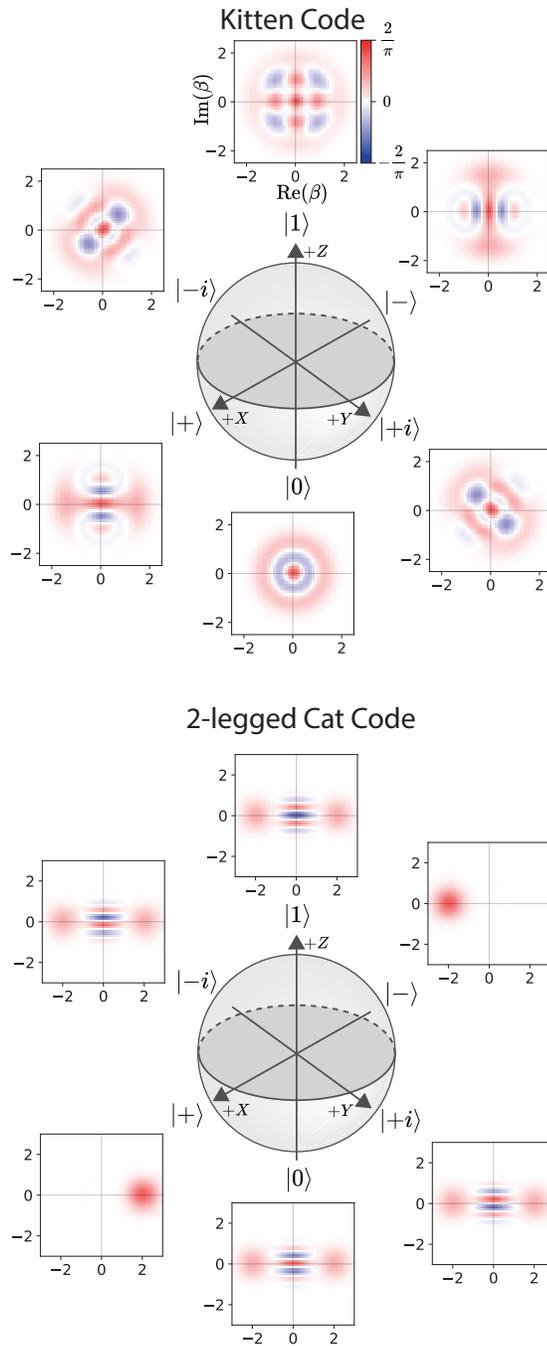


Figure 3.7: Wigner functions for the cardinal states of the Kitten code, and the 2-legged cat code with $\alpha = 2$. Color bar and axes shown in the $|1_L\rangle$ Kitten state only.

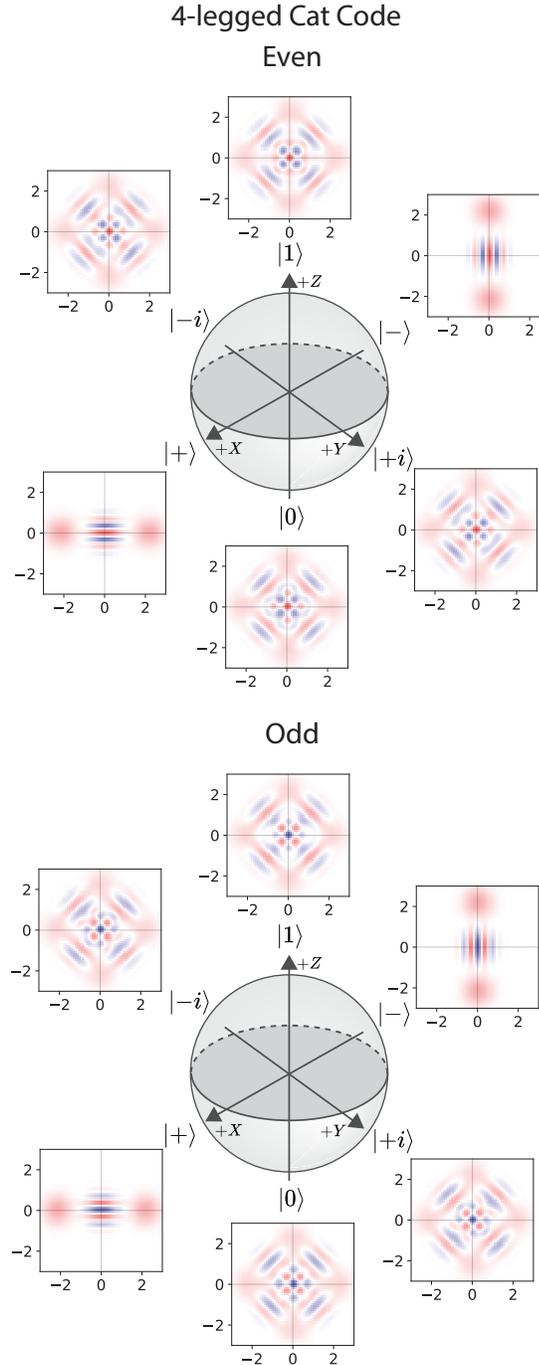


Figure 3.8: Wigner functions for the cardinal states of the 4-legged cat code at $\alpha = \sqrt{1.5\pi} \approx 2.171$, an example of a ‘magic’ α for even codewords (See Appendix B). A single photon loss events takes us from the even photon number codespace (top) to the odd photon number codespace (bottom). An additional photon loss takes us back to the even photon number codespace but with a logical bit-flip, X_L .

3.4 Hardware Fault-Tolerance

So far our discussion of bosonic QEC has been confined to memory operations, where the goal is to preserve the quantum information in an idle qubit for as long as possible. For universal quantum computation, we must be able to perform a universal gate set on the qubits. This is true for bosonically encoded qubits as well where not only must we find a way to engineer the gates themselves, we must also ensure they are *fault-tolerant* (at the hardware level). Performing gates and measurements on our bosonic qubits should not introduce errors that cannot be corrected in subsequent bosonic QEC operations, and the bosonic QEC operations themselves should also not introduce uncorrectable errors.

Depending on the choice of encoding, gates and/or measurements on bosonic qubits require engineering highly non-linear, photon number-specific interactions where using a control ancilla is unavoidable. Unfortunately, errors that propagate from the control ancilla to the bosonic mode(s) are not usually correctable through the standard bosonic QEC, which is designed to correct only for ‘natural’ errors in the bosonic mode such as photon loss. bosonic QEC also tells us which measurements and unitaries we need to realize, but not how to realize them fault tolerantly in a given set of hardware. This is the goal of ‘Hardware Fault-Tolerance’ (HFT).

We need to handle errors that occur in the control ancilla as well as errors that occur in the bosonic mode. Moreover, gate unitaries are physically realized by enacting a control Hamiltonian for a set amount of time. Jump errors can occur at any time during the continuous evolution and we must ensure we can still correct for these errors *after* the operation, even though the jumps happened at an unknown time during the operation (denoted by T_{jump}).

The scope of a HFT operation revolves around implementing a desired unitary with a control Hamiltonian, $\hat{\mathcal{H}}_c$ that acts on both the bosonic qubit and control ancilla, denoted by modes \hat{a} and \hat{q} respectively. This unitary may act on the bosonic

qubit, e.g. $(\hat{U}_{\text{gate}} \otimes \mathbb{1}_q)$ such as in the case of a gate, or on the composite bosonic qubit-control ancilla system, e.g. $(\mathbb{1}_a \otimes \mathbb{1}_q + \hat{U}_{\text{gate}} \otimes \hat{X}_q)$ as for a measurement where we want to map information from the bosonic qubit onto the control ancilla using the circuit in Eq. 2.6 with $\hat{B} = \hat{U}_{\text{gate}}$. The desired unitary can be further classified as a unitary that performs a logical gate or measurement on a bosonic qubit, or a unitary that performs a gate or measurement for bosonic QEC purposes. Both are needed for computation with bosonic qubits.

It is helpful to also define a finite set of hardware errors, $\{\hat{\epsilon}\}_{\text{hardware}}$ that we wish to be fault-tolerant to. This set usually contains only the dominant first-order hardware errors such as single-photon loss, and a single decay or dephasing jump in the control ancilla. We also do not typically consider no-jump backaction, or other second-order jump errors. On the timescales we consider, where gate operations are much faster than decoherence times, these errors scale as $(\sim (\kappa T_{\text{gate}})^2)$.

All operations on bosonic qubits that require a control ancilla must satisfy HFT, even when we wish to convert hardware errors to erasure errors. The conditions are more relaxed since the operation must preserve the error-detectability rather than the error-correctability of errors that happened during the unitary evolution. (Jump errors from $\{\hat{\epsilon}\}_{\text{hardware}}$ that occur during the evolution must be error-detectable after the gate). We now discuss the HFT criteria for realizing gates on bosonic qubits.

3.4.1 Path independence

When performing an error-corrected gate on a bosonic qubit of the form $(\hat{U}_{\text{gate}} \otimes \mathbb{1}_q)$, the hardware implementation must be *path independence* to satisfy HFT. The concept of path independence is that ultimately we perform the desired unitary on the bosonic mode, regardless of the errors that happen during the application of the control Hamiltonian required for its implementation. The occurrence of errors takes us on a different ‘path’, but ultimately all paths converge and the same desired unitary is

performed on the bosonic qubit regardless of which path we took. This is illustrated diagrammatically in Fig. 3.9. An example of a control Hamiltonian for which we can satisfy path independence is

$$\hat{\mathcal{H}}_c/\hbar = \Omega \left(\hat{U}_{\text{gate}} |f\rangle \langle g| + \hat{U}_{\text{gate}}^\dagger |g\rangle \langle f| \right) \quad (3.37)$$

which is the general case of the Hamiltonian used to perform the fault-tolerant gate in [Reinhold et al., 2020] where $|g\rangle, |e\rangle, |f\rangle$ are the first three levels of the control ancilla. The set of hardware errors we want to be fault-tolerant to is $\{\hat{a}, |f\rangle \langle f|, |e\rangle \langle f|\}$, corresponding physically to single photon loss, ancilla dephasing and ancilla decay. We presume the bosonic qubit can correct for single photon loss. The path independence diagram is shown for this specific case in Fig. 3.10 but we can also see why the dynamics are path independent by considering the unitary generated by this Hamiltonian.

This unitary is given by

$$\hat{U}_c = e^{-i\hat{\mathcal{H}}_c t/\hbar} = \cos(\Omega t) \mathbb{1} - i \sin(\Omega t) \left(\hat{U}_{\text{gate}} |f\rangle \langle g| + \hat{U}_{\text{gate}}^\dagger |g\rangle \langle f| \right). \quad (3.38)$$

In the absence of errors, after time $\Omega t = \pi/2$ we will have implemented \hat{U}_{gate} if the ancilla begins in state $|g\rangle$ (and ends in $|f\rangle$). If we stop the evolution at some intermediate time, we can probabilistically perform the gate. Measuring the ancilla to be in $|f\rangle$ after applying $\hat{\mathcal{H}}_c$ for some non-zero time always ‘heralds’ successful implementation of \hat{U}_{gate} . Similarly, measuring $|g\rangle$ means we have performed identity and must attempt the gate again. Transmon dephasing midway through the gate results in the same effect – after time $\Omega t = \pi/2$ there is now a probability we measure $|g\rangle$ at the end of the gate, which also indicates we performed $\mathbb{1}$ and must try again to implement the gate. Transmon decay from $|f\rangle$ to $|e\rangle$ has the effect of turning off the dynamics at some unknown time. However, if we measure the transmon in $|e\rangle$ at

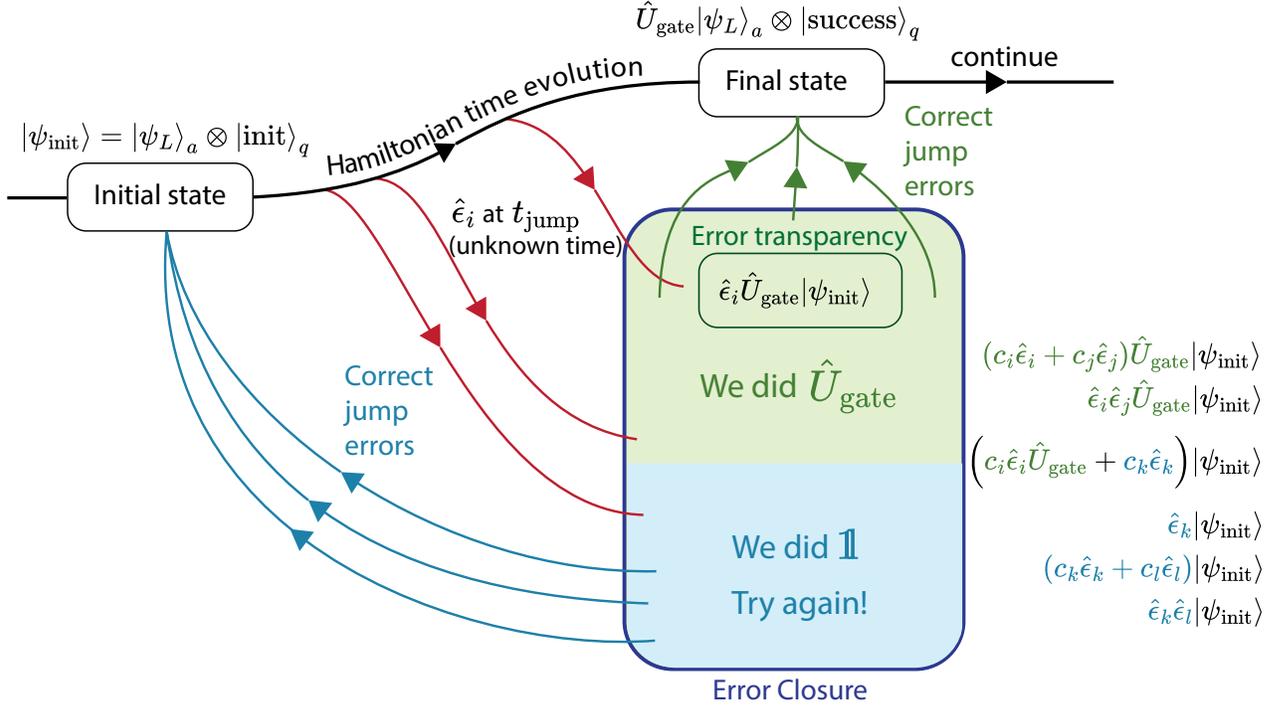


Figure 3.9: Path independence for an error-corrected gate. The control Hamiltonian takes the initial state to the final state in the absence of errors. Errors, \hat{e}_i happen at unknown time t_{jump} during the evolution and take us on different paths. At the end of the gate we end up in the space of correctable errors (Error closure). This space can be subdivided into errors where we applied the gate \hat{U}_{gate} , and suffered jump errors (light green background) and states where $\mathbb{1}$ was applied to the bosonic qubit and a jump error occurred (light blue background). Examples of final states that fall into one of these two categories are shown on the far right, which include superpositions and products of jump errors. After performing a syndrome measurement, we know which jump errors happened and whether \hat{U}_{gate} or $\mathbb{1}$ was applied. If we were in a superposition of different errors and/or a superposition of doing \hat{U}_{gate} and $\mathbb{1}$, this superposition is collapsed by the measurement. If \hat{U}_{gate} was applied, we correct for the jump errors (using a separate operation that must also satisfy path independence or is otherwise presumed to be error-free) as shown by the green arrows and end up in the desired final state. If $\mathbb{1}$ was applied, we must correct the jump errors (blue arrows) and attempt the gate again, repeating as necessary until we reach the final state. Failing the second attempt is a second-order error and so we only at most two gate attempts to ensure hardware errors are suppressed to second order. Also shown in green is the subset of error-correctable errors which are the error-transparent error. These jump errors commute with the control Hamiltonian and so it is as if we performed \hat{U}_{gate} perfectly and then applied the jump error.

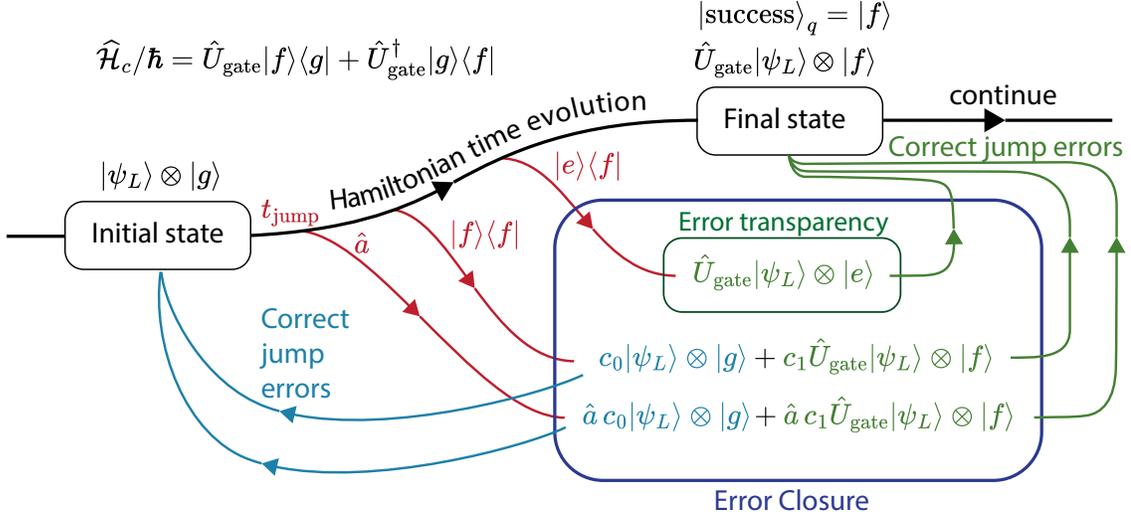


Figure 3.10: Path independence for a specific control Hamiltonian. The decay error $|e\rangle\langle f|$ is error transparent. If we detect $|e\rangle$ after the gate attempt, we know \hat{U}_{gate} was applied successfully. If $|f\rangle\langle f|$ occurred then after the gate we are in a superposition of having applied $\mathbb{1}$ with the ancilla remaining in $|g\rangle$ and having applied \hat{U}_{gate} with the ancilla now in $|f\rangle$. Measuring the ancilla collapses the superposition allowing us to correct this error accordingly. If we measure $|g\rangle$, try again. If $|f\rangle$, we were successful. The complex coefficients c_0 and c_1 depend on t_{jump} , which is unknown to us, but this only affects our gate success probability and not our ability to correct this error. Single photon loss, \hat{a} has the same effect as ancilla dephasing, with the same superposition after the gate, with an additional \hat{a} operator. Correcting this error is the same as for dephasing, but we must also correct for single photon loss. In this implementation, the control ancilla measurement is presumed perfect. Ancilla measurement error results in a first-order error since we will do the wrong correction step!

the end of the gate, this also indicates we have performed \hat{U}_{gate} successfully (since we must reach $|f\rangle$ in order to decay to $|e\rangle$).

Path independence does not apply to error-detected gates, which is one of the key reasons why the latter are easier to realize. Instead, independent of when the jump error happens, all ‘error paths’ must lead us to the space of error states which can be error-detected with measurements after the gate. This difference is shown in Fig. 3.11. These diagrams serve as a useful way to track the effects of jump errors, but what is the mathematical formalism which tells us whether path independence or its error-detected equivalent can be satisfied?

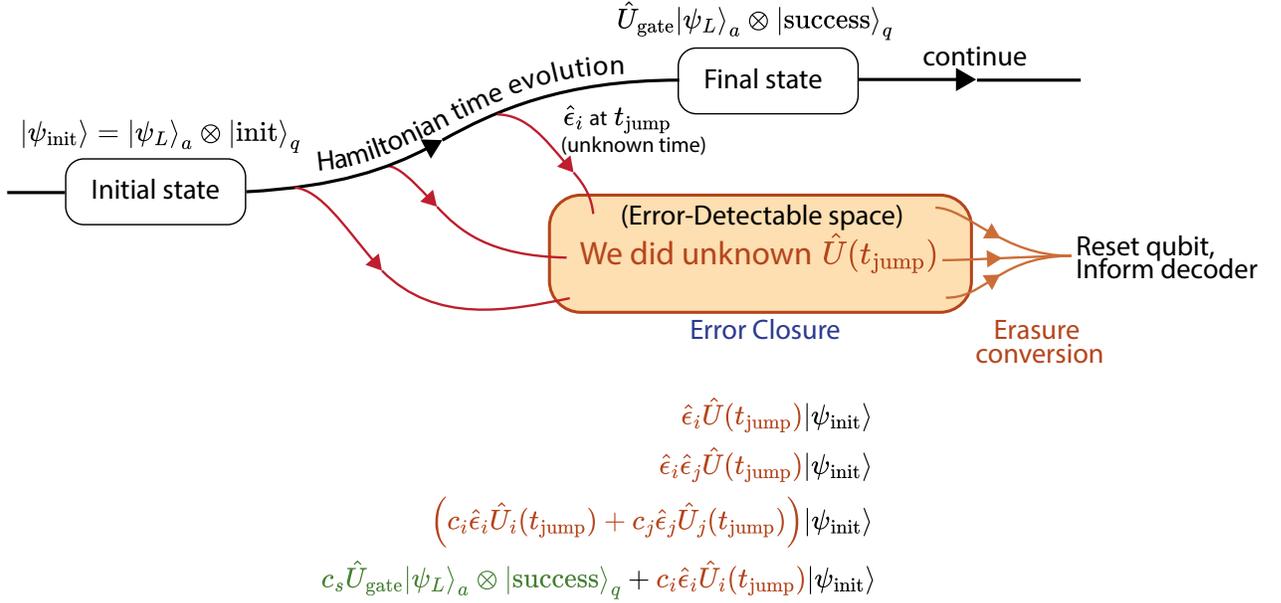


Figure 3.11: Error paths for an error-detectable gate. The conditions for a jump error during the gate attempt to still be error-detectable after the gate are more relaxed compared to the path independence conditions. Now it is acceptable to have performed an unknown unitary $\hat{U}(t_{\text{jump}})$, whose exact form depends on t_{jump} and is thus unknown. So long as we can still detect the jump error, we can reset the qubit to convert this error into an erasure error. Examples of final states in the error space that remain error detectable (orange) are shown at the bottom of the figure. Just as we saw with path-independence, jump errors that happen during the evolution may look like a superposition of jump error and no error after the gate attempt.

3.4.2 Error transparency

Error transparency is a stricter than necessary condition that ensures path independence. Formally, we are error-transparent to jump error $\hat{\epsilon}_i$ if $[\widehat{\mathcal{H}}_c, \hat{\epsilon}_i] = 0$. The desired unitary is generated by

$$\hat{U}_{\text{gate}} \otimes \mathbf{1}_q = e^{-i\widehat{\mathcal{H}}_c\tau} \quad (3.39)$$

by applying $\widehat{\mathcal{H}}_c$ for gate time τ . If error transparency is satisfied then the effect of the jump error occurring at unknown time t_{jump} is given by

$$e^{-i\widehat{\mathcal{H}}_c(\tau-t_{\text{jump}})}\hat{\epsilon}_i e^{-i\widehat{\mathcal{H}}_c\tau} = \hat{\epsilon}_i e^{-i\widehat{\mathcal{H}}_c\tau} = \hat{\epsilon}_i U_{\text{gate}} \quad (3.40)$$

All we need is to be able to detect and correct for jump error $\hat{\epsilon}_i$ to satisfy path independence.

3.4.3 Extension to error closure

The error closure condition we derive here is more general than error transparency and asks ‘what conditions do errors in $\{\hat{\epsilon}\}_{\text{hardware}}$ need to satisfy so they remain error-correctable after the gate?’ The set of errors we can correct is usually larger than just the set of individual hardware jump operators, $\{\hat{\epsilon}\}_{\text{hardware}}$. For example, if we can correct for individual hardware jump operators, $\hat{\epsilon}_i$ and $\hat{\epsilon}_j$, then we can also correct for arbitrary superpositions of these errors such as $c_i\hat{\epsilon}_i + c_j\hat{\epsilon}_j$ or products such as $\hat{\epsilon}_i\hat{\epsilon}_j$ (but not $(\hat{\epsilon}_i)^2$). We call this larger set of correctable errors $\{\hat{\epsilon}\}_{\text{corr}}$. A set of hardware jump errors, $\{\hat{\epsilon}\}_{\text{hardware}}$ satisfies error closure if any error, $\hat{\epsilon}$ during the gate evolution results in a correctable error after the unitary. i.e.

$$e^{-i\widehat{\mathcal{H}}_c(\tau-t_{\text{jump}})}\hat{\epsilon} e^{-i\widehat{\mathcal{H}}_c t_{\text{jump}}} = \hat{\epsilon}_{\text{corr}} \hat{U}_{\text{gate}} \quad (3.41)$$

for $\hat{\epsilon}_{\text{corr}} \in \{\hat{\epsilon}\}_{\text{corr}}$. Note that errors of the form $\hat{\epsilon}_i \hat{U}_{\text{gate}}^\dagger$ are also in $\{\hat{\epsilon}\}_{\text{corr}}$, provided we can distinguish them from their $\hat{\epsilon}_i$ counterpart. We call these errors ‘try again’ errors. Detecting this error signifies that the gate ‘failed’ and we did the identity operation instead of \hat{U}_{gate} . After correcting the jump error, we should attempt the gate again until we no longer detect a ‘try again’ error.

We now derive the minimal conditions that ensure this is the case, the error closure conditions. First we generate a new set of errors $\{\hat{\epsilon}\}_{\text{ext}}$ from every possible commutator between $\hat{\mathcal{H}}_c$ and the elements of $\{\hat{\epsilon}\}_{\text{hardware}}$:

$$[\hat{\mathcal{H}}_c, \hat{\epsilon}_{\text{hardware}}] = \hat{\epsilon}_{\text{ext}}, \quad (3.42)$$

as well as linear combinations of these commutators. Error closure is satisfied if:

1. $\hat{\epsilon}_{\text{ext}} \in \{\hat{\epsilon}\}_{\text{corr}}$,
2. $[\hat{\mathcal{H}}_c, \hat{\epsilon}_{\text{ext}}] \in \{\hat{\epsilon}\}_{\text{ext}}$.

These conditions state that the errors generated by Eq. 3.42 must form a closed set of correctable errors and ensure that hardware errors during the gate remain correctable errors after the gate. We now sketch the proof.

Jump evolution for a hardware error $\hat{\epsilon}_i$ occurring at time $T - t$ during the gate evolution can be written as

$$e^{-i\hat{\mathcal{H}}_c t} \hat{\epsilon}_i e^{-i\hat{\mathcal{H}}_c(\tau-t)} = e^{-i\hat{\mathcal{H}}_c t} \hat{\epsilon}_i e^{i\hat{\mathcal{H}}_c t} e^{-i\hat{\mathcal{H}}_c \tau}. \quad (3.43)$$

The condition for hardware errors to be correctable is:

$$e^{-i\hat{\mathcal{H}}_c t} \hat{\epsilon}_i e^{i\hat{\mathcal{H}}_c t} \in \{\hat{\epsilon}\}_{\text{corr}}, \quad (3.44)$$

$$\forall \hat{\epsilon}_i \in \{\hat{\epsilon}\}_{\text{hardware}}, \quad (3.45)$$

$$\forall t \in [0, \tau]. \quad (3.46)$$

Our error closure conditions ensure this is satisfied. From the Baker–Campbell–Hausdorff (BCH) theorem, we may write:

$$e^{-i\widehat{\mathcal{H}}_c t} \hat{\epsilon}_i e^{i\widehat{\mathcal{H}}_c t} = \tag{3.47}$$

$$\hat{\epsilon}_i + i[\widehat{\mathcal{H}}_c, \hat{\epsilon}_i]t - [\widehat{\mathcal{H}}_c, [\widehat{\mathcal{H}}_c, \hat{\epsilon}_i]]\frac{t^2}{2!} - i[\widehat{\mathcal{H}}_c, [\widehat{\mathcal{H}}_c, [\widehat{\mathcal{H}}_c, \hat{\epsilon}_i]]]\frac{t^3}{3!} + \dots \tag{3.48}$$

The error closure conditions ensure the nested commutation relations only produce errors that are in $\{\hat{\epsilon}\}_{\text{corr}}$ and hence the entire Taylor series is also in $\{\hat{\epsilon}\}_{\text{corr}}$ since it is a superposition of correctable errors. If $\widehat{\mathcal{H}}_c^{(1)}$ and $\widehat{\mathcal{H}}_c^{(2)}$ both satisfy error closure then so does $\widehat{\mathcal{H}}_c^{(1)} + \widehat{\mathcal{H}}_c^{(2)}$. Note that satisfying the error transparency condition automatically satisfies the error closure conditions. The error closure conditions may be trivially reformulated for error detected unitaries by replacing the set of errors we can correct for, $\{\hat{\epsilon}\}_{\text{corr}}$ with the set of errors we can detect for $\{\hat{\epsilon}\}_{\text{det}}$.

The Circuit Quantum Electrodynamics Toolkit

Circuit Quantum Electrodynamics (cQED) is a relatively new field derived from the much older field of Cavity Quantum Electrodynamics (CQED) where quantized states of an atom interact with a quantized light field. Traditionally, CQED uses atoms interacting with photons in optical frequency cavities. cQED realizes the same light-matter interactions where superconducting circuits play the role of ‘artificial atoms’ which interact with individual photons in microwave frequency cavities. What sets cQED apart from CQED is the ease in realizing *strong coupling*¹. The strength of the light-matter coupling can far exceed the decoherence rates of the cavities and artificial atoms, giving us access to a unique suite of Hamiltonians and controllable couplings that form the cQED toolkit.

This suite of tools is well-suited for quantum information processing. The goal of this chapter is to introduce and highlight the parts of the cQED toolkit we have chosen

1. Why is this the case? The effective electric dipole moment of an artificial atom is *huge* compared to a real atom. See [Devoret et al., 2007]

to use for our experiments (and proposed experiments) in the rest of this thesis. We have elected to store quantum information in microwave cavities and used transmon ‘artificial atoms’ as our control ancillas. Moreover, we have zeroed in on a particular subset of Hamiltonians and control techniques to design operations and protocols that can be easily *error-detected*. The work in Chapter 5 and Chapter 6 relies on being able to distill our cQED systems and control techniques down to simple Hamiltonians and error processes, enabling us to design gate operations that effectively detect the dominant errors. These are the same tools we have at our disposal in Chapter 7, with a different set of constraints when designing error-detected operations for a quantum communication problem.

We begin by summarizing the hardware, the Hamiltonians and dominant error processes we wish to error-detect. This serves as a reference guide for subsequent chapters. The rest of the chapter briefly describes the origins and properties of these building blocks, and why we have chosen them for their particular roles in our error-detected schemes. For a more rigorous introduction to cQED I recommend "A Quantum Engineer’s Guide to Superconducting Qubits" [Krantz et al., 2019]. For an overview of control with bosonic qubits in microwave cavities, see [Gao et al., 2021].

4.0.1 The cQED toolkit: a quick reference

Hardware:

Microwave cavities - Our QHOs in which we encode quantum information. Each oscillator mode is denoted by \hat{a} or \hat{b} etc.

Transmon control ancillas - the system which we manipulate to perform quantum operations on the QHOs, via the dispersive interaction. We consider up to the first four levels $|g\rangle, |e\rangle, |f\rangle, |h\rangle$. The transmon, denoted by mode operator \hat{q} , is the main source of non-linearity for quantum control and is an anharmonic oscillator $\omega_{ge} \neq \omega_{ef}$

Readout resonators - A QHO dispersively coupled to the transmon and used only for measuring the state of the transmon, performing a strong projective measurement that collapses the transmon into $|g\rangle, |e\rangle, |f\rangle \dots$ etc.

Non-linear cavity coupling elements - Non-linear systems that exist solely to actuate parametric couplings between cavity modes. We will only focus on the **beamsplitter coupling**.

Drive ports - ports that allow us to couple microwave drives to the above elements, for the purpose of actuating couplings, control Hamiltonians or readout.

Coaxial cable resonators - a QHO used for quantum communication between (in this case) microwave cavities by serving as a communication bus

Static Hamiltonians:

Bare cavity Hamiltonian

$$\widehat{\mathcal{H}}_a/\hbar = \omega_a \hat{a}^\dagger \hat{a} \quad (4.1)$$

where ω_a is the angular frequency of the cavity (1-10 GHz). Note we have omitted the $\hbar\omega_a/2$ zero-point energy term.

Bare transmon Hamiltonian

$$\widehat{\mathcal{H}}_q/\hbar = \omega_q \hat{q}^\dagger \hat{q} + \frac{\alpha_q}{2} \hat{q}^\dagger \hat{q}^\dagger \hat{q} \hat{q} \quad (4.2)$$

where ω_q is the frequency of the transmon g-e transition (1-10 GHz) α_q is the anharmonicity of the transmon (100-300 MHz)

Ideal cavity-transmon dispersive interaction

$$\widehat{\mathcal{H}}_\chi/\hbar = \chi_{ge} \hat{a}^\dagger \hat{a} |e\rangle \langle e| + \chi_{gf} \hat{a}^\dagger \hat{a} |f\rangle \langle f| \quad (4.3)$$

where χ_{ge} is the dispersive shift in the g - e manifold and χ_{gf} is the dispersive shift in the g - f manifold. (0.5-5 MHz)

Cavity-transmon dispersive interaction in the g - e manifold

$$\widehat{\mathcal{H}}_{\text{disp}}/\hbar = \chi_{ge} \hat{a}^\dagger \hat{a} |e\rangle \langle e| + \frac{K}{2} \hat{a}^\dagger \hat{a}^\dagger \hat{a} \hat{a} + \chi'_{ge} \hat{a}^\dagger \hat{a}^\dagger \hat{a} \hat{a} |e\rangle \langle e| \quad (4.4)$$

Where K is the self-Kerr of the cavity, the anharmonicity ‘inherited’ from the transmon control ancilla. χ'_{ge} is a higher-order non-linear coupling that can be thought of as the change in cavity self-Kerr when the transmon is in $|e\rangle$. ($K, \chi' \sim 1$ -20 kHz) A similar Hamiltonian exists for the g - f manifold.

Control Hamiltonians:

Cavity displacement drive

$$\widehat{\mathcal{H}}_{\epsilon_a} = \epsilon_a(t)\hat{a} + \epsilon_a^*(t)\hat{a}^\dagger \quad (4.5)$$

Written on resonance in the ‘drive frame’, detuning this drive by Δ_a from the resonance condition results in the additional term $\Delta_a\hat{a}^\dagger\hat{a}$. $\epsilon_a(t)$ is the complex classical drive amplitude which we vary in time e.g. turning it on or off. The fact that this amplitude is complex indicates we have phase control.

Transmon (charge) drive

$$\widehat{\mathcal{H}}_{\epsilon_q} = \epsilon_q(t) |g\rangle \langle e| + \epsilon_q^*(t) |e\rangle \langle g| \quad (4.6)$$

Where $\epsilon_q(t)$ is again a complex ‘classical’ drive field that couples to the transmon, causing transitions in the g - e manifold (Replace $|g\rangle \langle e| \rightarrow |g\rangle \langle f|$ for control in the g - f manifold)

Parametric beamsplitter interaction

$$\widehat{\mathcal{H}}_{\text{BS}} = g_{\text{BS}}(t)\hat{a}^\dagger\hat{b} + g_{\text{BS}}^*(t)\hat{a}\hat{b}^\dagger \quad (4.7)$$

By driving a non-linear coupling element, we can actuate the (bilinear) beamsplitter coupling between two QHOs. The classical drive amplitude $g_{\text{BS}}(t)$ is analogous to $\epsilon(t)$

Hardware error processes

Transmon Relaxation

Collapse operator, \hat{c}_i	Rate, κ_i	κ_i^{-1} Typical Lifetime Values
$ g\rangle\langle e $	Γ_{\downarrow}^{ge}	10-500 $\mu\text{s} \approx T_1^q$
$ e\rangle\langle f $	$\Gamma_{\downarrow}^{ef} \approx 2\Gamma_{\downarrow}^{ge}$	$= T_1^q/2$
$ e\rangle\langle g $	$\Gamma_{\uparrow}^{ge} = n_{\text{therm}}\Gamma_{\downarrow}^{ge}$	T_1^q/n_{therm}
$ f\rangle\langle e $	$\Gamma_{\uparrow}^{ef} \approx 2\Gamma_{\uparrow}^{ge}$	$2T_1^q/n_{\text{therm}}$

Where T_1^q is the relaxation time of the transmon. We do not list all possible collapse operators, only the dominant error channels when operating the transmon in the g - e or g - f manifold.

Transmon Dephasing

Collapse operator, \hat{c}_i	Rate, κ_i	κ_i^{-1} Typical Lifetime Values
$ g\rangle\langle g - e\rangle\langle e $	$\Gamma_{\phi}^{ge}/2$	$(\Gamma_{\phi}^{ge})^{-1} = 10\text{-}1000 \mu\text{s} = T_{\phi}^q$
$ g\rangle\langle g - f\rangle\langle f $	$\Gamma_{\phi}^{gf} \approx 4\Gamma_{\phi}^{ge}$	$\approx T_{\phi}^q/4$

Note the factor of 1/2 between the dephasing rate Γ_{ϕ}^{ge} and the rate for the dephasing collapse operator $\kappa_{\phi}^{ge} = \Gamma_{\phi}^{ge}/2$

Cavity Decoherence

Collapse operator, \hat{c}_i	Rate, κ_i	κ_i^{-1} Typical Lifetime Values
\hat{a}	Γ_{\downarrow}^a	0.5-10 ms = T_1^a
\hat{a}^\dagger	Γ_{\uparrow}^a	T_1^a/n_{therm}
$\hat{a}^\dagger\hat{a}$	$2\Gamma_{\phi}^a$	$(\Gamma_{\phi}^a)^{-1} = T_{\phi}^a > 10$ ms

Again note factor of 2 in the relation between the dephasing time, $T_{\phi}^a = (\Gamma_{\phi}^a)^{-1}$ and the rate for the collapse operator $\hat{a}^\dagger\hat{a}$. This operator represents white noise ‘intrinsic’ dephasing in a microwave cavity and is thought to be extremely small.

cQED hardware summary

The diagram below in Fig. 4.1 summarizes the cQED building blocks used in this thesis

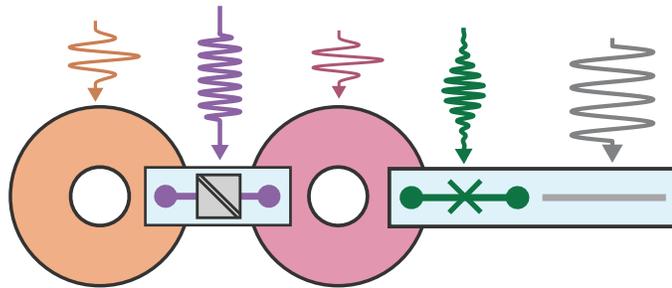


Figure 4.1: The cQED hardware necessary for the error detected gates in Chapter 5 and a dual-rail qubit in 6. Two cavities (orange and pink) are coupled together via a switchable parametric beamsplitter coupler (purple). The pink cavity also has a transmon control ancilla and readout resonator for transmon readout. Microwave drives (resonant) and pumps (off-resonant) can be applied to each of the five modes

4.1 Microwave Cavities

Microwave cavities are our go-to QHOs. As shown by Fig. 4.2 They are essentially just quantized LC oscillators with frequency $f_a = 1/\sqrt{LC}$ with frequencies typically chosen to be in the range 1-10 GHz (what we define to be the microwave regime) in order to couple sufficiently to our transmon qubits.

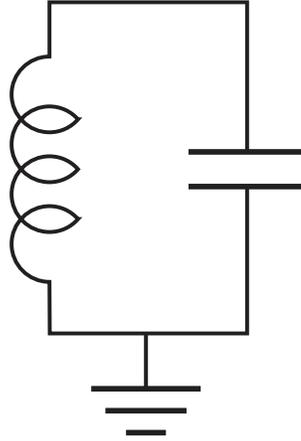


Figure 4.2: LC lumped element circuit model for a microwave cavity. A 3D geometry may support many standing wave modes, but each one can effectively be modelled as an LC oscillator with the inductance and capacitance determined only by the geometry and any dielectric material present

The Hamiltonian for a ‘bare’ microwave cavity is given by

$$\hat{\mathcal{H}}_a = \hbar\omega_a \left(\hat{a}^\dagger \hat{a} + \frac{1}{2} \right) \quad (4.8)$$

where $\omega_a = 2\pi f_a$ is the angular frequency of the cavity mode. For a formal treatment on how to quantize an LC oscillator, see [Vool and Devoret, 2017]. Often we represent the mode of a bosonic cavity as a QHO with a quadratic energy potential shown in Fig. 4.3.

Microwave cavities form the simplest part of the cQED toolkit and have the best coherence times and often negligible dephasing. Using superconducting metals (most commonly 99.999% 5N purity Aluminum but Tantalum and Niobium can also be

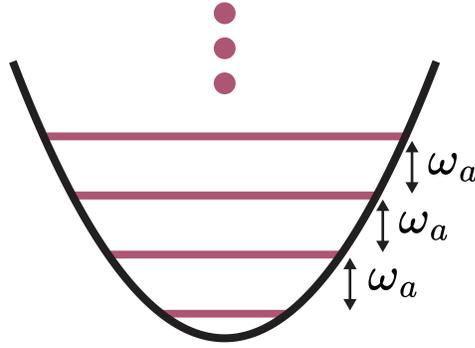


Figure 4.3: The energy levels of a QHO are evenly spaced by $\hbar\omega_a$

used) ensures low dissipation and long T_1 times. A dilution refrigerator at 10-20 mK is required to ensure n_{therm} is small ($<1\%$). The thermal population at low temperature is given by

$$n_{\text{therm}} = e^{-\frac{\hbar\omega_a}{k_B T_{\text{eff}}}} \quad (4.9)$$

One must be careful to ensure the cavity and the hardware that houses it is well thermalized to the baseplate of the dilution refrigerator. It is not uncommon to infer $T_{\text{eff}} \approx 100$ mK! This also sets a limit on how low frequency we can make our cavities, as this will increase n_{therm} for a given effective temperature. It is also helpful to define the Quality factor for a cavity as

$$Q = \frac{\omega_a}{\kappa_a} = 2\pi f_a T_1^a \quad (4.10)$$

Where f_a is the frequency of the cavity (the thing with units GHz! Don't forget the 2π !) This dimensionless number is useful for comparing the loss rates between different cavities independent of their frequency. A Q of say 1 million means the cavity can undergo 1 million oscillations before its energy has dropped to $1/e$ of the initial energy.

4.1.1 3D cavities

The highest Q microwave cavities are known as 3D cavities. These are essentially just a hollow box machined into one chunk of solid conductor such as 5N aluminum. These cavities are still well-described by the lumped element LC oscillator model, except now the capacitance and inductance is distributed over a 3D structure (but can easily be calculated from finite-element simulations).

3D cavities also usually require a ‘lid’ on one side which creates a seam². Care needs to be taken in where this seam is placed as it can be a source of energy loss. These cavities are fairly bulky, with dimensions on the order of centimetres (set by the typical wavelength of microwave light in vacuum in the 1-10GHz range). The reason for high Q in these cavities is that the dielectric is mostly vacuum and the geometries can be designed to reduce loss from other known mechanisms such as dielectric loss in the few nm thick metal oxide layer that forms on the metal-air interface. (Chemical etching is known to improve the Q but exactly why this works is still uncertain³.)

Many 3D cavity geometries have been explored over the years including rectangular, ellipsoidal, panflute and coaxial stub cavities [Reagor et al., 2013]. We chose to focus solely on the coaxial stub cavity (sometimes called a post cavity). A cavity can have many modes, only some of which may be useful as a controllable QHO. We use the fundamental or $\lambda/4$ mode of the coaxial stub cavity, shown in Fig. 4.4. By choosing the dimensions such that the upper section functions as a circular wave guide with a cutoff frequency above the frequency of the $\lambda/4$, placing the lid at the end of this wave guide results in negligible seam loss.

Apart from the relative ease in machining, the main advantage of stub cavities is that we can easily couple transmons and other non-linear coupling elements to the $\lambda/4$

2. This can be circumvented with the ‘panflute’ architecture see [Chakram et al., 2021]

3. Possible reasons include removing surface contaminants, reducing surface roughness, removing defects from CNC machining...

mode, without spoiling the coherence times of the cavity mode ($T_1^a \sim 1$ ms [Reagor et al., 2016]). This is achieved as shown below in Fig. 4.10 where the transmon or other coupling element has an antenna that juts out into the cavity field, slightly below the top of the post and gives a static capacitive coupling on the order of 10-50 MHz (The further it pokes in, the larger the coupling).

4.1.2 Stripline resonators

The other style of microwave cavity we are interested in is the stripline resonator⁴. This consists of an inner conductor, patterned (often meandered) on a dielectric substrate usually made of sapphire (or silicon). These materials have a higher dielectric constant, typically $\epsilon_r = 10-12$. This combined with the ability to meander the resonator make stripline resonators much more compact than 3D stub cavities. By suspending the sapphire chip within a tunnel made of another superconductor (such as Aluminum), we realize another coaxial geometry, this time with an open at both ends. The chip must be held in place on one end by use of a clamp, where the seam is also located.

The fundamental mode is the $\lambda/2$ mode and the tunnel functions as the ground plane. Since the mode consists of a ‘2D’ on-chip inner conductor and 3D tunnel-like outer conductor, this is often referred to as a quasi-2D resonator. True 2D cQED architectures have a ground plane on one side of the chip (or similar).

These cavities can conveniently be patterned on the same chip as other elements such as transmons or couplers but tend to have lower T_1 times due to the increased number of interfaces and higher participation of the mode in the (more lossy) substrate. Stripline resonators are ubiquitous in their use as readout resonators for transmons and other artificial atoms, where high Q is not a requirement (in fact, lowering the Q is desirable for fast readout!). Recent material developments where

⁴. resonator and cavity are synonymous

the stripline resonator is patterned with tantalum [Place et al., 2021] rather than aluminum [Axline et al., 2016] can increase T_1 to be comparable with stub cavities with lifetimes of 1 ms.

Coaxial cable resonators

The last instance of a microwave cavity of interest to us is formed from a microwave coaxial (coax) cable. These are commercially available and used to route microwave signals to and from our quantum devices. A length of coax terminated by an open (or a short) at both ends functions as a $\lambda/2$ resonator, just like a stripline resonator. The cable can host a number of standing wave mode harmonics, spaced evenly in frequency by Δ_{FSR} , called the Free Spectral Range. The Free Spectral Range is determined by the length of the cable, L and speed of light in the cable, c as $\Delta_{\text{FSR}} = \frac{c}{2L}$.

For sections of coaxial cable between 0.01-10m, there exists standing wave modes we can couple to in the desirable 1-10 GHz window, and Δ_{FSR} is sufficiently large such that we can couple to individual standing wave modes, rather than a continuum.

When the cable is made out of superconductor, such as Niobium Titanium, its Q factor can be large: $\sim 10,000$ - $100,000$. Q factors of order $\sim 1\text{M}$ have recently been reported in all-Aluminum cables [Niu et al., 2023]. Although these resonators have lower Quality factors than 3D cavities, they are convenient for linking 3D cavities together across separate modules, and we may easily couple to them at E-field antinodes, e.g., by placing transmon pad antennae near the ends of the cable.

It is important that the outer conductor of these cables shares a common ground with the 3D cavities and transmons. Contact resistance between the outer conductor and common ground is thought to limit Q factors rather than the Teflon dielectric that fills the space between the inner and outer conductor.

The three types of microwave cavity we have introduced are summarized in Fig. 4.4 below.

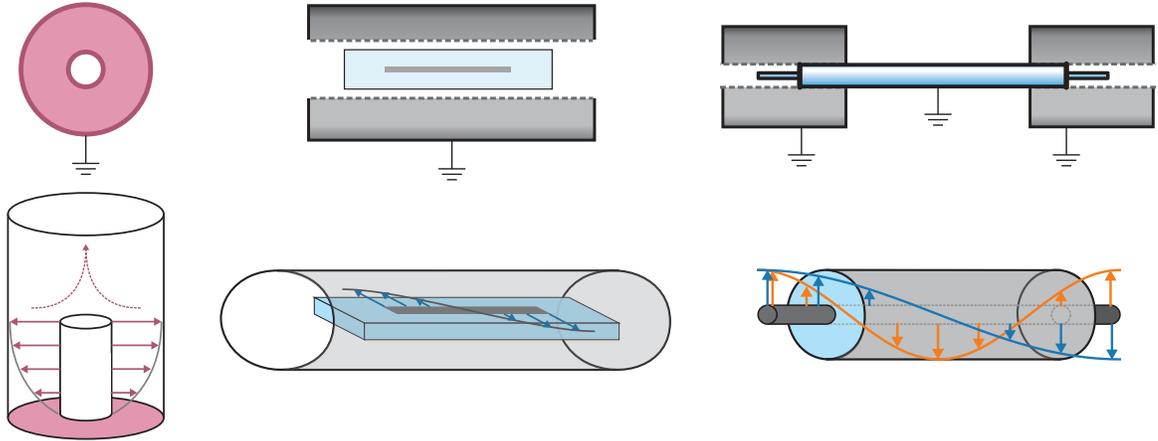


Figure 4.4: The three types of microwave cavities of interest in this thesis. (Top row) 2D projections showing ground planes (Bottom row) 3D geometry showing the E-field profile for the cavity mode(s). (Left) 3D stub cavity showing the E-field profile for the $\lambda/4$ mode. (Middle) on-chip stripline resonator suspended in a cylindrical metal tunnel (typically Aluminum), showing the $\lambda/2$ fundamental mode. (Right) Coaxial cable resonator showing the $\lambda/2$ and λ fundamental (blue) and first harmonic (orange). The exposed inner conductor allows capacitive coupling to other modes of interest. Both ends of the cable are terminated as an ‘open’ by slotting them into metal tunnels as indicated in the top projection.

4.1.3 The natural noise-bias of microwave cavities

A property of microwave cavities we will exploit heavily in Chapters 5 and 6 is their frequency stability. By this, we mean it is hard to change the frequency of cavity modes once the package has been assembled. This results in negligible intrinsic dephasing and a large natural noise-bias (dominant error channel is photon loss). In fact, the intrinsic dephasing time of 3D microwave cavities is one of the highest coherence times measured in the cQED platform.

Why are microwave cavities so robust to dephasing noise? In short, it is very hard for the environment to change the length of the inner conductor, which is largely what sets ω_a . For example, a cryogen-free dilution refrigerator is not terribly good at damping vibrations [Olivieri et al., 2017] (in part due to the pulse tubes) and the entire cavity may suffer from mechanical vibrations at frequencies up to a few kHz. However, the whole cavity is vibrating and so this does not alter ω_a , especially if the

inner conductor and outer conductor are machined from the same block of metal as with a stub cavity.

One source of dephasing we may encounter is from the so-called ‘diving board mode’ which is more prominent in stripline resonators. If the chip is not clamped securely (or not clamped at both ends), or if the chip is excessively long then mechanical vibrations may cause the chip to wobble like a diving board. Changes in the distance between the inner conductor and tunnel walls of the outer conductor only weakly affect the cavity frequency.

A curiosity observed in microwave cavities is the increase in Q with the number of photons in the resonator. This is attributed to TLS defects (the same which are responsible for $1/f$ noise) which become saturated at high photon numbers but act as an additional source of loss at low photon numbers. Since we typically operate our cavities where $\bar{n} < 10$, we must characterize our cavities at low photon numbers.

4.2 Transmons

The source of non-linearity in the cQED toolkit is the Josephson junction (J.J.), and one of the simplest ‘artificial atoms’ that can be made from a Josephson junction is the transmon. The transmon will be the system we use as our control ancilla. The other use in this thesis for Josephson junctions is in non-linear couplers⁵, elements designed to actuate parametric coupling between other modes in our systems.

Why do we use Josephson junctions as our source of non-linearity? And what does it mean for a mode to be non-linear?

A non-linear mode has an anharmonic, i.e., not evenly spaced energy spectrum which allows us to manipulate its quantum states via classical drives. Josephson junctions are tunnel junctions consisting of two superconductor separated by a thin

5. which may also themselves be transmons

barrier of insulating metal oxide. The Josephson junction is unique in the sense that it functions as a non-linear, lossless circuit element. This is best exemplified by the Josephson relations [Josephson, 1962] which show the junction's unique I-V characteristics.

$$\begin{aligned} I &= I_c \sin(\phi) \\ V &= \frac{\hbar}{2e} \frac{d\phi}{dt} \end{aligned} \tag{4.11}$$

Where I_c is the critical current of the junction set by both the width of the tunnel barrier (typically few nm) and the choice of superconducting material. ϕ is the superconducting phase difference across the junction.

4.2.1 The transmon regime

The transmon is derived from the ‘Cooper-pair box’, and is formed by shunting a single Josephson junction with a large capacitance [Koch et al., 2007] as shown by the circuit in Fig. 4.5. Its Hamiltonian is given by:

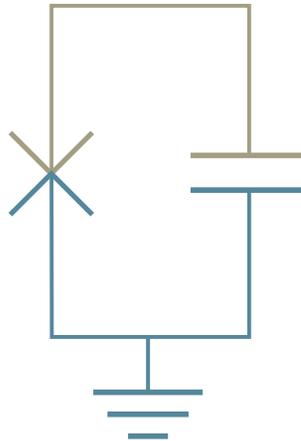


Figure 4.5: Circuit diagram for a transmon. The Josephson junction can be modelled as a non-linear inductance with $L = L_J / \cos(\phi)$. Where L_J is the Josephson inductance $L_J = 2\pi\Phi_0/I_c$. The top half of the circuit is colored dark green to show the ‘superconducting island’.

$$\hat{\mathcal{H}}_{\text{Transmon}} = 4E_c \hat{n}_c^2 - E_J \cos(\hat{\phi}) \quad (4.12)$$

Where E_c is the charging energy, given by $E_c = e^2/2C_\Sigma$, where C_Σ is the total capacitance of the circuit. E_J is the Josephson energy, related to the critical current as $E_J = I_c \Phi_0/2\pi$ where $\Phi_0 = h/2e$ is the magnetic flux quantum. \hat{n}_c is the number of Cooper pairs on the superconducting island (the top half of Fig. 4.5). This Hamiltonian is analogous to a quantum pendulum.

In a superconductor, ideally the only mobile charge carriers are Cooper pairs. The first term in this Hamiltonian is the capacitive contribution to the energy i.e. the charging energy to add Cooper pairs to the island which takes this same form for an LC oscillator as well. The second term is the energy cost for a Cooper pair to tunnel through the barrier of the Josephson junction.

A transmon must satisfy the ‘transmon regime’, for which $E_J/E_c \gtrsim 50$. By engineering a large shunting capacitance, we can reduce E_c to reach the transmon regime. We can engineer E_c easily from the geometry of the antenna pads either side of the junction. The transmons we consider are quasi-2D. They are patterned on sapphire chips suspended in Aluminum tunnels. The total effective lumped element capacitance is determined from all the capacitances in the distributed 3D geometry. E_c is largely set by the distance between the transmon antenna pads and their dimensions, but also their capacitance to the outer tunnel which functions as the ground plane (and the junction capacitance). E_J (10-50 GHz) is set only by the barrier width controlled in fabrication.

We opt to use transmons as our control ancillas because they are naturally insensitive to both charge and flux noise. The other reason is that it is easy to engineer a dispersive coupling and QND readout. Without any flux loops in the circuit, the (fixed frequency) transmon is insensitive to flux noise. The main benefit of the transmon regime is that, as we increase the E_J/E_c ratio, the charge dispersion (which

dictates our sensitivity to charge noise) decreases exponentially whilst the anharmonicity (which dictates the speed of control) only decreases as a weak power law. Compared to other artificial atoms, such as Fluxonium, the transmon is remarkably linear! However the remaining non-linearity, α_q that sets the ‘quantum speed limit’ for our quantum control⁶ is still sufficiently fast to perform transmon gate unitaries on the order ~ 10 ns.

We carefully change basis (see [Girvin, 2014]) to the ladder operators \hat{q} and \hat{q}^\dagger to arrive at the Transmon Hamiltonian

$$\hat{\mathcal{H}}_q/\hbar = \omega_q \hat{q}^\dagger \hat{q} - \frac{E_J}{\hbar} \cos_{\text{nl}}(\varphi_q(\hat{q} + \hat{q}^\dagger)) \quad (4.13)$$

Where $\omega_q = \omega_{ge}$ is the frequency of the g - e transition, given by $\hbar\omega_q = \sqrt{8E_c E_J} - E_c$ and $\varphi_q = (2E_c/E_J)^{1/4}$, the zero-point fluctuations of the phase across the junction and \cos_{nl} is the 4th order and higher terms of the cosine (we do this so the cosine terms don’t keep renormalizing our mode frequencies) i.e.

$$\cos_{\text{nl}}(x) = \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \dots \quad (4.14)$$

This Hamiltonian keeps the ‘full cosine’ of the transmon. We can instead truncate the Hamiltonian, keeping only the 4th order non-linearity to arrive at the Hamiltonian

$$\hat{\mathcal{H}}_q/\hbar = \omega_q \hat{q}^\dagger \hat{q} + \frac{\alpha_q}{2} \hat{q}^\dagger \hat{q}^\dagger \hat{q} \hat{q} \quad (4.15)$$

which is a quantum Duffing oscillator. The anharmonicity, $\alpha_q = -E_c$ is set only by the charging energy. The anharmonicity gives the difference in frequency spacing between the levels. i.e. $\omega_{ef} = \omega_{ge} + \alpha_q$, $\omega_{fh} = \omega_{ef} + \alpha_q$. As a consequence of the cosine potential, the anharmonicity is always negative and results in the higher trans-

6. if we use pulses much faster than $1/\alpha_q$, our transmon behaves like a QHO and the individual levels are no longer addressable

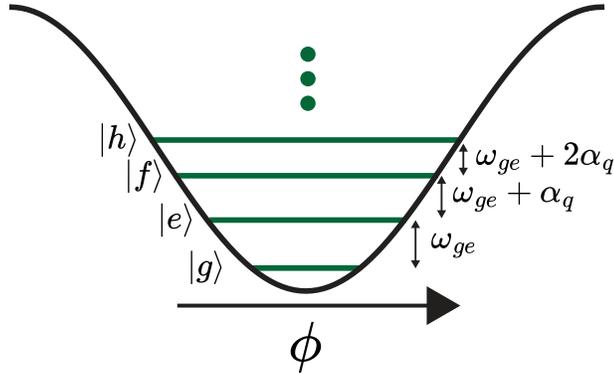


Figure 4.6: The ‘cosine potential’ for a transmon showing the lowest four energy states and frequency spacing. The anharmonicity, $\alpha_q = -E_c$ is always negative for a transmon.

mon states bunching closer together in energy. The different transmon transitions are susceptible to noise at different parts of the noise spectrum, which can lead to unexpected decoherence rates. For example, in a QHO we expect the decay rate from $|2\rangle$ to $|1\rangle$ to be exactly double the decay rate from $|1\rangle$ to $|0\rangle$. However, we often find experimentally that $\Gamma_{\downarrow}^{ef} \neq 2\Gamma_{\downarrow}^{ge}$, due to differing noise spectral densities at ω_{ge} and ω_{ef} . The same goes for dephasing in the g - e and g - f manifolds and the exact dephasing rate ratios are highly device dependent. (Although assuming the rates we would expect for a QHO is still a good rule of thumb [Peterer et al., 2015]).

The energy spectrum or first four energy levels, $|g\rangle, |e\rangle, |f\rangle, |h\rangle$ of the transmon is shown in Fig. 4.6

4.2.2 Transmon decoherence

The T_1 and T_2 coherence times of a transmon are bounded by the coherence times of stripline resonators, and so it is safe to say that microwave cavities will always have coherence times better than what can be achieved in transmons. (For the best transmon design, we could skip the Josephson junction fabrication step to realize a stripline resonator which is just as good.)

As well as being sensitive to the usual sources of loss in a stripline resonator, such

as dielectric loss in the substrate, the presence of the junction tunnel barrier means transmons are susceptible to more noise sources, such as charge noise, quasiparticle tunnelling and photon-induced dephasing. Whilst $T_\phi^q > T_1^q$ is typical in a transmon, they do not exhibit the strong natural noise bias found in stripline resonators.

Improvements in materials, reducing sources of noise and suppressing sensitivity to noise sources have all contributed to the steady rise in transmon coherence times over the past 15 years. One of the most recent developments, the use of tantalum antenna pads [Place et al., 2021] (which allow more aggressive cleaning) have resulted in state-of-the-art T_1 times around 300-500 μs and dephasing times up to 1 ms.

4.2.3 Controlling transmons

The transmon is controlled by capacitively coupling a drive pin (attached to a microwave transmission line) to the transmon antenna. This is known as ‘charge driving’ and results in the Hamiltonian

$$\hat{\mathcal{H}}/\hbar = \omega_q \hat{q}^\dagger \hat{q} + \frac{\alpha_q}{2} \hat{q}^\dagger \hat{q}^\dagger \hat{q} \hat{q} + \epsilon_q(t) \hat{q} + \epsilon_q^*(t) \hat{q}^\dagger \quad (4.16)$$

Where we have assumed the stiff pump approximation (the drive field is large enough so it can be considered a ‘classical’ rather than as a quantum variable. Removing a photon from the drive has a negligible effect on the drive amplitude.) By choosing this drive to be on resonance with the ω_{ge} transition i.e. $\epsilon_q(t) \rightarrow \epsilon_q(t) e^{-i\omega_q t}$ where we can rewrite Eq. 4.16 in the rotating frame of the transmon as

$$\hat{\mathcal{H}}/\hbar = \frac{\alpha_q}{2} \hat{q}^\dagger \hat{q}^\dagger \hat{q} \hat{q} + \epsilon_q(t) \hat{q} + \epsilon_q^*(t) \hat{q}^\dagger. \quad (4.17)$$

If the bandwidth of the $\epsilon_q(t)$ is smaller than the anharmonicity α_q then in the frame of the drive, the Hamiltonian becomes

$$\widehat{\mathcal{H}}_{\epsilon_q} = \epsilon_q(t) |g\rangle \langle e| + \epsilon_q^*(t) |e\rangle \langle g| \quad (4.18)$$

where we only address the g - e manifold. By changing the phase of the drive, we can rotate the transmon in its g - e Bloch sphere about any equatorial axis. By tuning our drive frequency to ω_{ef} , on resonance with the e - f level, we can perform the same control in the e - f manifold. However, things are a little different if we tune on resonance with $\omega_{gf} = \omega_{ge} + \omega_{ef}$. There is no matrix element between $|g\rangle$ and $|f\rangle$, this is a forbidden transition. However, if the transmon absorbs two drive photons then it is possible to drive this transition, via a virtual transmon level detuned from the $|e\rangle$ level by $\alpha_q/2$.

The need for a virtual intermediate level means control in the g - f manifold, necessary for chapters 6 and 5 is inevitably slower than control in the g - e manifold. It takes 50-100 ns to perform a π -pulse in the g - f manifold using this method whereas π pulses in the g - e manifold can be as fast as 4 ns.

In practice, it is difficult to implement a drive that only addresses a particular manifold of the transmon. As we shorten the pulses for g - e control, the presence of the $|f\rangle$ level first causes an additional phase shift to our g - e rotations and then causes a leakage error at shorter pulses. Leakage results in unwanted population transfer to the $|f\rangle$ state due to the large bandwidth of the pulse. It is mitigated by DRAG pulses [Motzoi et al., 2009] or by using numerical optimized pulses for the Hamiltonian in Eq. 4.17 which allow us to get close to the $1/\alpha_q$ speed limit. DRAG and similar pulses may also be employed for control in the g - f manifold, where the goal is to minimize leakage to the $|e\rangle$ level (and higher levels).

Control in the g - f manifold may also be achieved with two drive frequencies ω_1

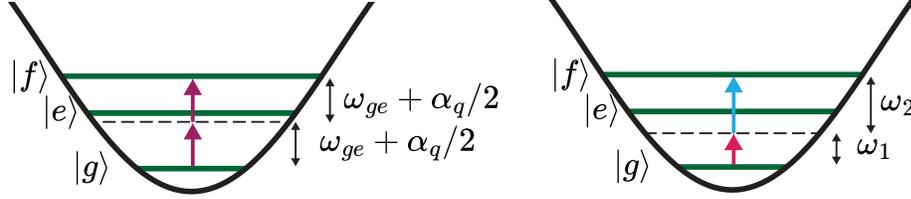


Figure 4.7: Two schemes for directly driving $|g\rangle \leftrightarrow |f\rangle$ transitions in a transmon. The virtual intermediate level is shown as a dashed line. (Left) driving with a single pump at frequency $\omega = \omega_{ge} + \alpha_q/2$. Two pump photons must be absorbed to excite the transmon from $|g\rangle$ to $|f\rangle$. The virtual level is detuned from the $|e\rangle$ level by $\alpha_q/2$. (Right) Two-photon driving with two drives at frequencies ω_1 and ω_2 . Detuning from the virtual level can be varied and is given by $\omega_{ge} - \omega_1$

and ω_2 which satisfy $\omega_1 + \omega_2 = \omega_{ge} + \omega_{ef}$. This is similar to the drive scheme we introduced earlier, except the virtual level is detuned from the $|e\rangle$ level by $\omega_{ge} - \omega_1$, a detuning that can be controlled by our choice of drive frequencies. The bandwidth of our drive must be smaller than this detuning to avoid directly populating the $|e\rangle$ level. An advantage of this approach is we can set the amplitude of one drive to be constant and the amplitude of the other to change the angle of rotation. Unlike the first approach, this approach does not have a strong quadratic dependence on drive power. The degenerate case when $\omega_1 = \omega_2$ recovers the first approach. Both possible drive schemes are shown in Fig. 4.7 For the purpose of chapters 6 and 5 we assume we can engineer the Hamiltonian

$$\hat{\mathcal{H}}_{gf} = \epsilon_{gf}(t) |g\rangle \langle f| + \epsilon_{gf}^*(t) |f\rangle \langle g| \quad (4.19)$$

where leakage to $|e\rangle$ is a detectable error but still rare ($<1\%$ per π -pulse)

4.3 The Dispersive Coupling

The dispersive coupling is the primary non-linear resource we use for controlling bosonic qubits via an ancilla transmon, and also used in transmon readout. It will

be utilized extensively throughout the rest of this thesis.

To derive the dispersive Hamiltonian we begin with the Jaynes-Cummings Hamiltonian:

$$\widehat{\mathcal{H}}_{\text{JC}}/\hbar = \omega_a \hat{a}^\dagger \hat{a} + \omega_q \hat{q}^\dagger \hat{q} + \frac{\alpha_q}{2} \hat{q}^\dagger \hat{q}^\dagger \hat{q} \hat{q} + g_c (\hat{a}^\dagger \hat{q} + \hat{a} \hat{q}^\dagger) \quad (4.20)$$

where we have truncated the transmon's cosine term at 4th order in \hat{q} and g_c is the static capacitive (vacuum Rabi) coupling between a microwave cavity mode, \hat{a} and transmon mode \hat{q} . Typically $\omega_a/2\pi$ is 1-10 GHz, $\omega_q/2\pi$ is 3-6 GHz, $\alpha_q/2\pi$ is 100-300 MHz and g_c is 10-50 MHz. This is the central Hamiltonian of both CQED and cQED. In cQED, g_c far exceeds the decoherence rates of the cavity and transmon (strong-coupling) and we reach the dispersive interaction in the strong-dispersive limit.

To reach this limit, we chose our mode frequencies such that $\Delta = |\omega_a - \omega_q| \gg g_c$. Taking the Schrieffer-Wolff (S-W) transformation and the rotating wave approximation (RWA) and keeping up to 4th order terms we arrive at the dispersive Hamiltonian:

$$\widehat{\mathcal{H}}_{\text{disp}}/\hbar = \omega_a \hat{a}^\dagger \hat{a} + \omega_q \hat{q}^\dagger \hat{q} + \frac{\alpha_q}{2} \hat{q}^\dagger \hat{q}^\dagger \hat{q} \hat{q} + \chi \hat{a}^\dagger \hat{a} \hat{q}^\dagger \hat{q} + \frac{K}{2} \hat{a}^\dagger \hat{a}^\dagger \hat{a} \hat{a} \quad (4.21)$$

Where $\chi = \frac{2g_c^2 \alpha_q}{\Delta(\Delta + \alpha_q)}$ is the strength of the dispersive interaction. When $\Delta \gg \alpha_q$ we can write $\chi \approx 2 \left(\frac{g_c}{\Delta}\right)^2 \alpha_q$ and $K \approx \left(\frac{g_c}{\Delta}\right)^4 \alpha_q$ is the self-Kerr inherited by the cavity. One subtlety with this transformation is that we have actually changed our mode basis from the 'bare modes' to the 'dressed modes' (approximately, $\hat{a} \rightarrow \hat{a} + \left(\frac{g_c}{\Delta}\right) \hat{q}$, $\hat{q} \rightarrow \hat{q} - \left(\frac{g_c}{\Delta}\right) \hat{a}$) This weak hybridization slightly perturbs the mode frequencies and transmon anharmonicity and results in the dressed cavity mode becoming slightly anharmonic and inheriting a self-Kerr term.

The parameter $\left(\frac{g_c}{\Delta}\right)$ quantifies the entire dispersive interaction. It sets the cavity-transmon mode hybridization, and relative strengths of the wanted and unwanted terms. The strength of the dispersive interaction scales as $\chi \sim \left(\frac{g_c}{\Delta}\right)^2 \alpha_q$. Higher-order corrections to this Hamiltonian are usually undesired such as cavity self-Kerr, cavity-

cavity cross-Kerr, and χ' terms⁷ and all scale $\sim \left(\frac{g_c}{\Delta}\right)^4 \alpha_q$. We see that as $\left(\frac{g_c}{\Delta}\right) \rightarrow 0$, these unwanted terms vanish faster than the desired dispersive interaction. However, we want χ , which sets the speed of our quantum control of the cavity to still be faster than the decoherence rates of our system! For this reason, we typically work with $\left(\frac{g_c}{\Delta}\right) \sim 0.05-0.1$, with the hierarchy of non-linearities $\alpha_q/2\pi \approx 200$ MHz, $\chi/2\pi \approx 0.5-2$ MHz, $K/2\pi \approx 5-20$ kHz.

It is convenient to write the dispersive Hamiltonian for the first three transmon levels, since these are usually the only levels we plan to use. This Hamiltonian is

$$\begin{aligned} \widehat{\mathcal{H}}_{\text{disp}}/\hbar = & \chi_{ge} \hat{a}^\dagger \hat{a} |e\rangle \langle e| + \chi'_{ge} \hat{a}^\dagger \hat{a}^\dagger \hat{a} \hat{a} |e\rangle \langle e| + \frac{K}{2} \hat{a}^\dagger \hat{a}^\dagger \hat{a} \hat{a} \\ & + \chi_{gf} \hat{a}^\dagger \hat{a} |f\rangle \langle f| + \chi'_{gf} \hat{a}^\dagger \hat{a}^\dagger \hat{a} \hat{a} |f\rangle \langle f| \end{aligned} \quad (4.22)$$

Often we find in experiments $\chi_{gf} \approx 2\chi_{ge}$ but rarely exactly due to higher order corrections. By aiming for a $\omega_q - \omega_a \approx \alpha_q$ we can get naturally close to the ‘ χ -matching’ regime where $\chi_{gf} = \chi_{ge}$ without needing additional Hamiltonian engineering. We touch on this later in Chapter 5. All proposed and existing error-corrected operations with a transmon ancilla have required χ -matching.

Note how the dispersive interaction is necessarily always smaller than α_q . This has important consequences when we chose to encode our qubit in a cavity instead of a transmon. By opting to use the dispersive interaction rather than the transmon’s anharmonicity as our main resource for cavity control we naturally incur a ‘time overhead’ of $\left(\frac{\Delta}{g_c}\right)^2 \sim 100$ in the gate operation speed. This unavoidable slowdown means our cavity error rates had better be good! And if we do nothing to protect ourselves from transmon ancilla errors propagating onto the bosonic qubit, our gates will have error rates $\times 100$ worse than (single qubit) transmon gates! The spirit behind the work in Chapter 6 is to find the simplest-to-engineer system for which encoding

7. χ' is a 6th order term that comes from keeping higher orders in the cosine expansion before the S-W transformation, it results in the extra term $\chi' \hat{q}^\dagger \hat{q} \hat{a}^\dagger \hat{a}^\dagger \hat{a} \hat{a}$

a qubit in a cavity (or rather pair of cavities) is clearly advantageous compared to encoding the qubit in a transmon.

4.3.1 χ -matching

To date, decay in the control transmon ancilla has been error-corrected only by operating in the g - f manifold and engineering χ -matching. We briefly mention how χ -matching is engineered but refer the reader to [Rosenblum et al., 2018b] for a comprehensive guide. If we first rewrite the dispersive Hamiltonian in a different frame as

$$\widehat{\mathcal{H}}_{\text{disp}}/\hbar = \chi_g \hat{a}^\dagger \hat{a} |g\rangle \langle g| + \chi_e \hat{a}^\dagger \hat{a} |e\rangle \langle e| + \chi_f \hat{a}^\dagger \hat{a} |f\rangle \langle f| \quad (4.23)$$

Where $\chi_{ge} = \chi_e - \chi_g$ and $\chi_{gf} = \chi_f - \chi_g$. The goal of χ -matching is to achieve $\chi_e = \chi_f$, which can be done by applying a strong off-resonant pump to the transmon close to the $|e, n\rangle \leftrightarrow |h, n-1\rangle$ sideband transition, which excites the transmon from $|e\rangle$ to $|h\rangle$ whilst removing a photon from the cavity and vice versa. This shifts the value of χ_e such that

$$\chi_e \rightarrow \chi_e + \frac{g_p^2}{4\Delta_p} \quad (4.24)$$

Where g_p is the Rabi rate of the sideband transition if the pump were on resonance and Δ_p is the detuning of the pump from resonance (and we can change its sign).

This technique comes with the following drawbacks

1. Applying a strong, off resonant pump to the transmon can ‘heat’ the transmon out of its ground state, an uncorrectable error
2. If g_p is not strong enough χ -matching can only be achieved for a fairly small $n < 5$ number of photons in the cavity
3. As with any pump applied to a transmon, we run the risk of unwanted frequency collisions/multiphoton processes [Xiao et al., 2023] which tends to get worse as

g_p increases.

χ -matching is only required for error-correction of transmon decay. It is not required for error-detection, which allows us to circumvent these limitations.

4.4 Transmon Readout

The dispersive interaction is vital for QND readout of the state of a transmon qubit. The electrical circuit that models the transmon-readout resonator system that is ubiquitously used for readout is shown in Fig 4.8.

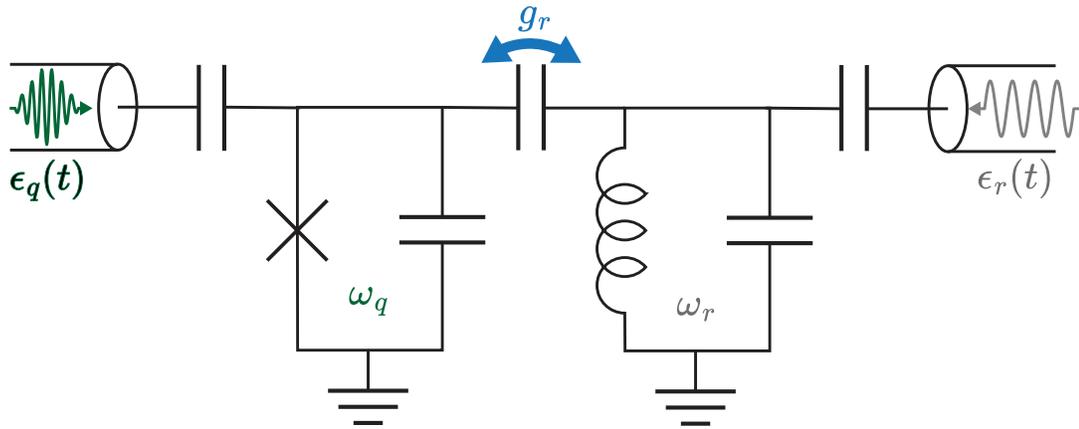


Figure 4.8: Electrical circuit for a transmon capacitively coupled to a readout resonator, also showing coupled transmission lines for delivering pulses to the transmon mode (green) and for displacing the readout resonator (grey). The value of the coupling capacitor between the transmon and readout resonator is characterized by g_r , which gives the rate of vacuum-Rabi oscillations if $\omega_q = \omega_r$, typically $g_r = 10\text{-}50$ MHz and can be found/designed from finite element simulations. When $|\omega_r - \omega_q| \gg g_r$, we are in the dispersive regime. The length of the drive pins (see Fig. 4.9) sets the value of the other coupling capacitances. The transmission lines that carry pulses present a lossy $50\ \Omega$ environment. For the transmon drive, we must ensure a large capacitance, so the presence of the drive port does not degrade transmon T_1 . For the readout drive, we want a smaller capacitance (we want to be overcoupled to the readout resonator) which ensures fast readout.

In this context a QND readout means when we measure the transmon outcome to be i , for $i = g, e, f, \dots$ it was as if the projector $|i\rangle\langle i|$ was applied to the transmon

state and no other information was gained. The transmon state should be perturbed as little as possible by our measurement. e.g. if the transmon was in $|e\rangle$ before the measurement, it should remain in $|e\rangle$ afterwards. We can quantify the QNDness as

$$\text{QNDness} = (P(|g\rangle_{\text{after}} | |g\rangle_{\text{before}}) + P(|e\rangle_{\text{after}} | |e\rangle_{\text{before}})) / 2 \quad (4.25)$$

We also want the measurement to be high fidelity, with readout fidelity defined as

$$\mathcal{F}_{\text{readout}} = (P(\text{declare } g | |g\rangle_{\text{before}}) + P(\text{declare } e | |e\rangle_{\text{before}})) / 2 \quad (4.26)$$

These metrics can be trivially extended to include readout of the $|f\rangle$ level.

We pattern the transmon and a stripline resonator, which functions as the readout resonator on the same chip such that there is a dispersive interaction between the transmon and readout resonator mode. How can we use the dispersive interaction, $\chi = \hat{a}^\dagger \hat{a} \hat{q}^\dagger \hat{q}$, to perform QND readout? The dispersive interaction is reciprocal. We can either view it as a shift in the transmon frequency by χ per photon, or a shift in the resonator frequency by χ per excitation in the transmon. Taking the first point of view makes it clear that (a distribution) of photons in the readout resonator dephases the transmon. The second point of view makes dispersive readout more intuitive. If the transmon is in $|g\rangle$, the readout resonator is at frequency ω_r . If the transmon is in $|e\rangle$, the readout resonator shifts to $\omega_r + \chi_{ge}$. (And if the transmon is in $|f\rangle$, the readout resonator shifts to $\omega_r + \chi_{gf}$ and so on...) Thus QND readout is achievable by probing the frequency of the readout resonator.

To do this (in reflection), we displace the readout resonator mode with an external drive field at a frequency near ω_r . The readout resonator is over-coupled to its drive pin such that photons leak out of the resonator at rate κ_r and back into the drive line where the signal can be boosted through several stages of amplification (starting with a quantum-limited amplifier for optimal Signal-to-Noise ratio) until we can record

the phase shift from the time-integrated output signal. By using a quantum-limited amplifier which only adds half a photon of noise at the first stage we are able to reliably achieve single-shot readout routinely in experiments. This means our overall SNR is good enough to determine the state of the transmon without needing to average many experimental shots.

The important thing to know is that this measurement is not instantaneous. It takes time $1/\kappa_r$ for the readout resonator to ring-up to a steady (coherent) state with average photon number $\bar{n} = 10\text{-}50$. Each photon weakly measures the state of the transmon due to the dispersive shift, and gradually this forms a strong projective measurement.

The measurement rate is give by

$$\Gamma^r = \frac{4\chi^2\kappa_r}{4\chi^2 + \kappa_r^2}\bar{n}, \quad (4.27)$$

where the measurement rate is maximized for $\kappa_r = 2\chi$. We typically choose $\chi \sim 1$ MHz, which is small enough to prevent the resonator from Purcell-limiting the transmon. (Purcell filters on the readout resonator also help with this.) We also aim for $\kappa_r \approx \chi$. At first glance, it seems as though we can arbitrarily increase the measurement rate by displacing the readout to larger \bar{n} . This is true up to a point. At excessively large \bar{n} ($> 50?$) we see a drop in T_1^q which degrades $\mathcal{F}_{\text{readout}}$ and QNDness despite reading out quickly. This effect is known as T_1 vs. \bar{n} , thought to be caused by unavoidable multiphoton transitions between the readout resonator and the transmon [Xiao et al., 2023, Shillito et al., 2022]. This sets a limit on typical readout times at 0.5-2 μs .

This time is important because it largely quantifies the ‘time cost’ of performing a QND measurement. These measurements can be responsible for most of the ‘idling time’ when running a quantum circuit with frequent mid-circuit measurements. Dur-

ing the measurement, other parts of the system (Microwave cavities, transmons etc.) will be accumulating errors.

Certain errors, such as decay of the transmon during readout reduce both the QNDness and $\mathcal{F}_{\text{readout}}$. Other errors, e.g. poor SNR only reduce $\mathcal{F}_{\text{readout}}$. Whilst transmon decay during the readout usually limits QNDness and $\mathcal{F}_{\text{readout}}$ to 98-99.5%, the SNR is often much better than this, with overall errors below 10^{-4} . One trick to boost readout fidelity when trying to distinguish between $|e\rangle$ and $|f\rangle$ is to play an e - f π -pulse to shelve the population in $|e\rangle$ to $|f\rangle$. Provided we have good SNR between the g and not g outcomes (e or f), we will still declare the correct measurement outcome when a single decay error occurs during readout, requiring two decays to make it likely we declare the wrong outcome [Elder et al., 2020].

We must be careful to minimize the number of residual thermal photons in the readout resonator, as these will cause dephasing at rate

$$\Gamma_{\phi}^r = \frac{4\chi^2\kappa_r}{4\chi^2 + \kappa_r^2} n_{\text{therm}}^r \quad (4.28)$$

This is the same physics that occurs during dispersive readout, except now unwanted photons in the readout are weakly ‘measuring’ the transmon all the time!

Below in Fig. 4.9 we see the quasi-2D cQED implementation of a controllable transmon with transmon readout.

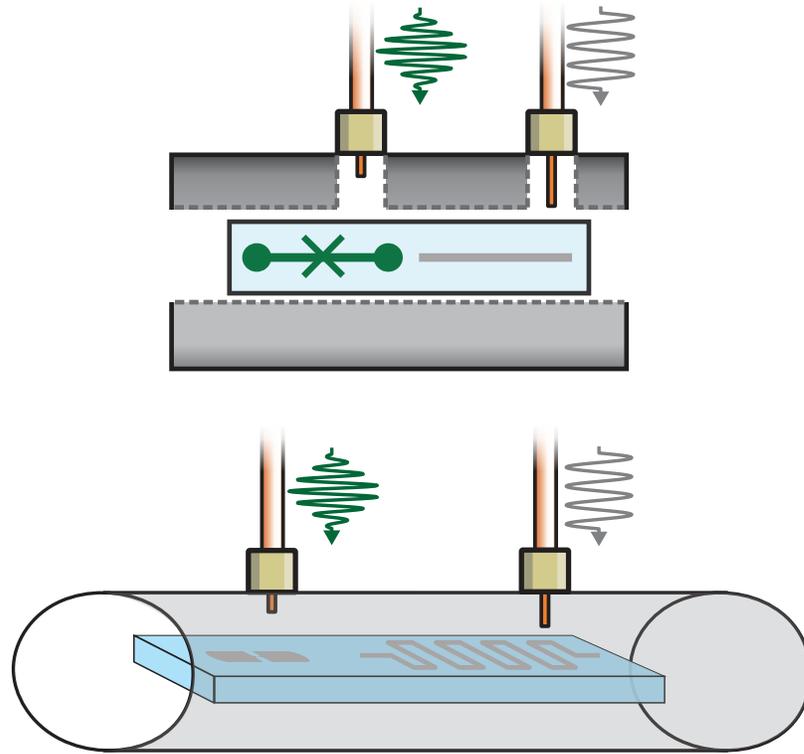


Figure 4.9: Quasi-2D cQED implementation of a transmon and readout resonator on the same chip. Two coupling ports are shown, one for the transmon drives (green) used for transmon control and another for the readout resonator pulses (grey) for readout of the transmon state. In practice, these may be the same physical port. A Purcell filter is typically used as well but is not shown. (Top) simplified 2D projection. (Bottom) 3D geometry with meandered readout resonator.

4.5 Cavity control

4.5.1 Displacement drives

We may displace a cavity via microwave drives near the cavity's resonant frequency. Just as with transmon control for quasi-2D transmons, we place a coupling pin at the end of the transmission line to capacitively couple to the E field of the cavity mode. The general strategy with drive pins is to couple weakly to the field and use strong drive powers. This ensures our drives are in the stiff-pump approximation and prevents our drive-pin from limiting the lifetime of the mode we are trying to couple to via the Purcell effect (Purcell-limiting). (Essentially, we get a high on-off ratio,

and our drives can be treated as classical variables.) The Hamiltonian for cavity displacement drive $\epsilon_a(t)$ on resonance is given by

$$\widehat{\mathcal{H}}_{\epsilon_a} = \epsilon_a(t)\hat{a} + \epsilon_a^*(t)\hat{a}^\dagger \quad (4.29)$$

This drive is only capable of displacing the initial vacuum state of the cavity to coherent states of the form $|\alpha(t)\rangle$. These states are often referred to as ‘classical’. (The only thing ‘quantum’ about them is that they still obey the Heisenberg uncertainty relation.)

4.5.2 Universal control via the dispersive interaction

To create more interesting ‘quantum’ states in the cavities, we usually make use of the non-linearities afforded by the dispersive coupling to the transmon. A popular Hamiltonian for cavity control via a transmon ancilla is

$$\widehat{\mathcal{H}}_c = \chi_{ge}\hat{a}^\dagger\hat{a}|e\rangle\langle e| + \epsilon_a(t)\hat{a} + \epsilon_a^*(t)\hat{a}^\dagger + \epsilon_q(t)|g\rangle\langle e| + \epsilon_q^*(t)|e\rangle\langle g| \quad (4.30)$$

which includes the static dispersive interaction and drives on the cavity and transmon mode near the ω_a and ω_{ge} frequencies respectively. This Hamiltonian is capable of universal control. Any unitary may be generated on the joint Hilbert space of the cavity and g - e manifold of the transmon, and thus we can achieve quantum control.

Finding the controls for the drives, $\epsilon_a(t)$ and $\epsilon_q(t)$ that generates these unitaries is not trivial though and requires numerical optimization techniques. Moreover, this Hamiltonian (for the most general controls) does not have any inbuilt fault-tolerance to ancilla transmon or cavity errors. There is no guarantee that transmon errors will not propagate onto the cavity or vice versa. The cQED hardware used to realize this control Hamiltonian is illustrated in Fig. 4.10 and the corresponding electrical circuit in Fig. 4.11

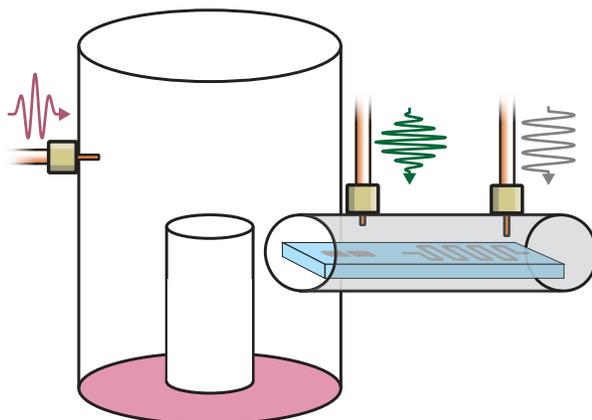


Figure 4.10: The ‘Swiss army knife’ of microwave cavity control. A transmon (with its own readout resonator) is dispersively coupled to a 3D stub cavity mode. The three drive ports used for pulse delivery to their respective modes.

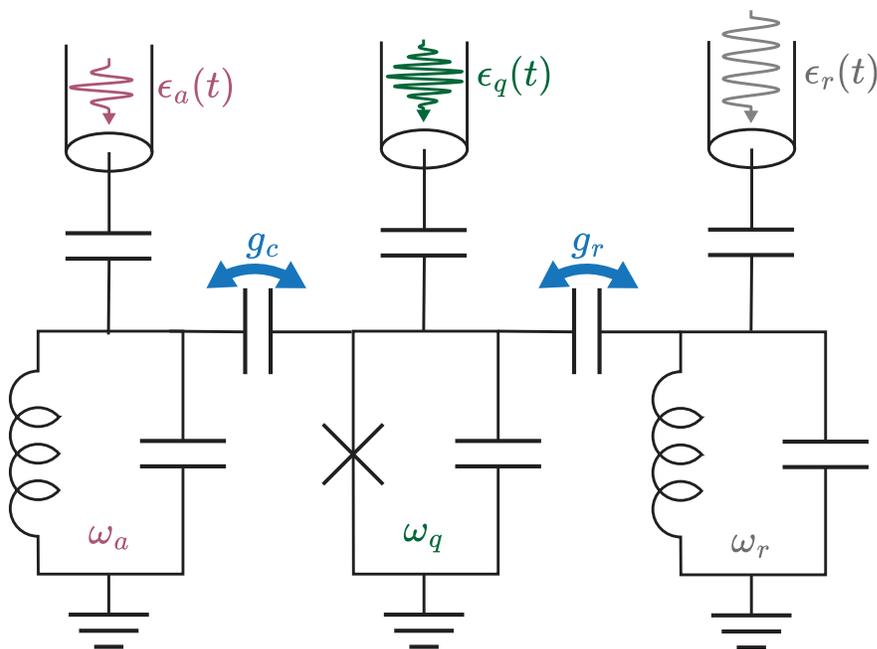


Figure 4.11: Electrical circuit for a transmon dispersively coupled to both a 3D cavity mode and a readout resonator, showing control lines for quantum control and transmon readout.

4.5.3 Photon-number preserving unitaries

We can examine the special case of this Hamiltonian where $\epsilon_a(t) = 0$. Any unitaries generated by this Hamiltonian will be excitation-conserving on the cavity. The transmon only interacts with the cavity via the dispersive interaction, which does not change the photon number. There are two broad classes of unitary we can generate when $\epsilon_a(t) = 0$. The first are unitaries that leave the cavity and transmon disentangled at the end. These unitaries must take the form.

$$\hat{U}_{\text{SNAP}} = \sum_{n=0}^N e^{i\phi_n} |n\rangle \langle n|_a \otimes \hat{U}_q \quad (4.31)$$

These are called Selective-Number Arbitrary-Phase (SNAP) gates because we can engineer them to impart arbitrary phases on each number state $|n\rangle$ in the cavity. We usually choose $\hat{U}_q = \mathbb{1}_q$ or $\hat{U}_q = |e\rangle \langle g| + |g\rangle \langle e| = X_q$.

To see how to construct $\epsilon_q(t)$ without resorting to numerical optimization we can make use of the phenomena of ‘photon number splitting’. Suppose we play a π -pulse on the transmon at the ω_{ge} frequency but chose a long duration/small bandwidth pulse such that $T_{\text{pulse}} > 1/\chi$. We can use the example of a ‘ 4σ ’ π -pulse, truncated Gaussian pulse envelope for the $\epsilon_q(t)$ drive. We can write the time-dependent Hamiltonian as.

$$\mathcal{H}(t)/\hbar = \chi_{ge} \hat{a}^\dagger \hat{a} |e\rangle \langle e| + A e^{-\frac{(t-2\sigma)^2}{2\sigma^2}} (|g\rangle \langle e| + |e\rangle \langle g|) \quad (4.32)$$

Where σ is the time width of the pulse (with bandwidth $1/\sigma$) and A is the normalized drive strength calibrated such that the area under the pulse is π after applying the pulse for time $T_{\text{pulse}} = 4\sigma$ (and the pulse flips the transmon). We call such a pulse a ‘selective’ π -pulse since it enacts the unitary.

$$\hat{U} = |0\rangle \langle 0|_a \otimes X_q + \sum_{n=1}^N |n\rangle \langle n|_a \otimes \mathbb{1}_q \quad (4.33)$$

If the transmon begins in $|g\rangle$, we only excite the transmon to $|e\rangle$ if the cavity is in $|0\rangle$. If we then measure the transmon, this amounts to a QND measurement of whether the cavity is in $|0\rangle$ or not.

4.5.4 Frequency combs for parity measurements

By detuning our pulse to frequency $\omega_{ge} + n\chi$, we only excite the transmon if the cavity is in state $|n\rangle$. By playing a simultaneous frequency comb of selective pulses at frequencies $\omega_{ge} + n_s\chi$ for $n_s \in \{n_s\}$, we realize the unitary.

$$\hat{U} = \sum_{n \in \{n_s\}}^N |n\rangle \langle n|_a \otimes X_q + \sum_{m \notin \{n_s\}}^N |m\rangle \langle m|_a \otimes \mathbb{1}_q \quad (4.34)$$

Which is realized by enacting the following Hamiltonian for time $T_{\text{pulse}} = 4\sigma$

$$\mathcal{H}(t)/\hbar = \chi_{ge} \hat{a}^\dagger \hat{a} |e\rangle \langle e| + A \sum_{n \in \{n_s\}}^N e^{-in\chi t} e^{-\frac{(t-2\sigma)^2}{2\sigma^2}} |g\rangle \langle e| + \text{h.c.} \quad (4.35)$$

Where h.c. is the Hermitian conjugate. Now, measuring the transmon is a QND measurement of whether the cavity is in the manifold $\{|n_s\rangle\}$. If we choose $\{n_s\} = \{0, 2, 4, 6, 8, \dots\}$, this amounts to a QND photon number parity measurement of the cavity.⁸

We must be careful in our choice of σ . Too long and the long duration of the pulse means we are more susceptible to decoherence in the transmon and cavity. Too short and we are no longer ‘selective’ enough and there is a larger probability the transmon is excited when $n \notin \{n_s\}$. Setting $\sigma = \pi/\chi$ is usually sufficient to avoid the ‘teeth’ of the frequency comb from overlapping whilst not being excessively slow. Numerical optimized pulses with the constraint $\epsilon_a(t) = 0$ can account for the teeth

⁸. we could also have chosen $\{n_s\} = \{1, 3, 5, 7, 9, \dots\}$ and then the measurement outcomes are flipped

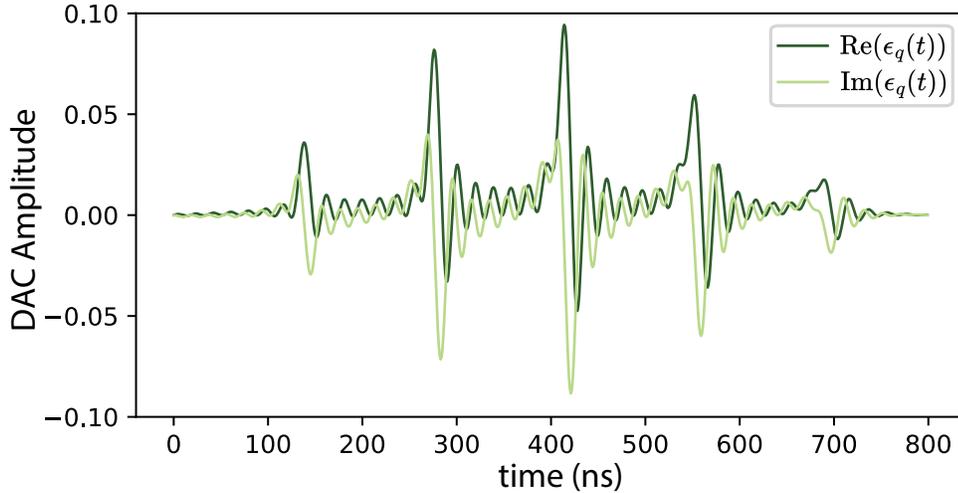


Figure 4.12: Example of a frequency comb on the transmon that implements a parity mapping unitary. The qubit will be excited to $|e\rangle$ if the number of photons in the cavity is even and $n < 14$. This pulse is used in Chapter 7 to measure photon number parity. DAC amplitude is essentially an arbitrary unit. We see the time domain waveform of a frequency comb also appears as a comb with spacing $\approx \pi/\chi_{ge}$ between the ‘teeth’.

of the frequency comb overlapping and can generate pulses shorter than $4\pi/\chi$ in duration to realize unitaries as in Eq. 4.34.

In the limit $N \rightarrow \infty$ and for constant pulse envelopes with duration $T_{\text{pulse}} = \pi/\chi$, $\epsilon_q(t)$ converges to the usual parity sequence:

$$\text{(unselective) } \frac{\pi}{2} \text{ pulse} \text{ — wait } \frac{\pi}{\chi} \text{ — (unselective) } \frac{\pi}{2} \text{ pulse}$$

Although as we shall see in Chapter 7, there are some advantages to constructing parity measurements via a frequency comb. An example of an analytically generated frequency comb for a parity measurement is shown in Fig 4.12.

4.5.5 SNAP gates

The frequency comb can be used to generate a SNAP gate by playing selective π -pulses for $\{n_s\} = \{0, 1, 2, 3, 4, \dots\}$ but with varying phases such that we enact the Hamiltonian

$$\mathcal{H}_{\text{SNAP}}(t)/\hbar = \chi_{ge} \hat{a}^\dagger \hat{a} |e\rangle \langle e| + A \sum_{n=0,1,2,\dots}^N e^{i\phi_n} e^{-inx t} e^{-\frac{(t-2\sigma)^2}{2\sigma^2}} |g\rangle \langle e| + \text{h.c.} \quad (4.36)$$

Which generates the SNAP unitary in Eq. 4.31. SNAP gates are useful for implementing logical gates on a variety of bosonic codes. For the family of codes that are rotationally symmetric, i.e. $Z_L = e^{i\frac{\pi}{m}\hat{a}^\dagger\hat{a}}$ for positive integer integer m (this includes cat codes and the kitten code with $m = 2$), the logical gate $Z_L(\theta)$ can be written as

$$Z_L(\theta) = \cos\left(\frac{\theta}{2}\right) \mathbb{1} + i \sin\left(\frac{\theta}{2}\right) e^{i\frac{\pi}{m}\hat{a}^\dagger\hat{a}} \quad (4.37)$$

A bosonic code with m -fold rotational symmetry, will only have support on the number states $|n'\rangle$ where $n' = n \times m$ for integer n . E.g. two-fold rotationally symmetric states only have support on even photon number states i.e. $\{|0\rangle, |2\rangle, |4\rangle \dots\}$ Our SNAP pulse to the $Z_L(\theta)$ gate thus only needs frequencies at $\omega_{ge} - n'\chi$ with phases

$$\phi_{n'} = (-1)^{n'} \times \frac{\theta}{2} \quad (4.38)$$

Interleaved SNAP and cavity displacements operations can generate any unitary on the cavity's Hilbert space but in general such constructions are not fault-tolerant to transmon or photon loss errors.

In the limit where $\sigma > 1/\chi$, we can write this Hamiltonian as

$$\mathcal{H}_{\text{SNAP}}/\hbar = \sum_n e^{i\phi_n} |n\rangle \langle n| \otimes |g\rangle \langle e| + \text{h.c.} = \hat{U}_{\text{SNAP}} |g\rangle \langle e| + \hat{U}_{\text{SNAP}}^\dagger |e\rangle \langle g| \quad (4.39)$$

This Hamiltonian takes the form of Eq. 3.37 in Chapter 3 when we discussed path independence. This means the SNAP gate is fault tolerant at the hardware level to both transmon dephasing errors and photon loss errors during the gate. If we measure the transmon in $|e\rangle$ at the end of the gate, we did \hat{U}_{SNAP} , otherwise we did $\mathbb{1}$.

(And we should error-correct any photon loss errors if we detect any after the gate). Performing selective π -pulses in the g - f manifold and χ -matching are necessary to error-correct transmon decay to first-order. (i.e. $|e\rangle\langle f|$).

We may combine SNAP gates and photon-number measurements into a single operation using a frequency comb of the form

$$\mathcal{H}(t)/\hbar = \chi_{ge} \hat{a}^\dagger \hat{a} |e\rangle\langle e| + A \sum_{n \in \{n_s\}}^N e^{i\phi_n} e^{-in\chi t} e^{-\frac{(t-2\sigma)^2}{2\sigma^2}} |g\rangle\langle e| + \text{h.c.} \quad (4.40)$$

which enacts the unitary

$$\hat{U} = \sum_{n \in \{n_s\}}^N e^{i\phi_n} |n\rangle\langle n|_a \otimes X_q + \sum_{m \notin \{n_s\}}^N |m\rangle\langle m|_a \otimes \mathbb{1}_q \quad (4.41)$$

4.5.6 Optimal control theory pulses

The other paradigm we introduce for achieving control of a single cavity mode are optimal control theory pulses. Starting from the same control Hamiltonian in Eq. 4.30, we numerically optimize the pulses for the control drives, $\epsilon_q(t)$ and $\epsilon_a(t)$ to implement a desired unitary on the cavity-transmon Hilbert space.

Because 4.30 is a universal control, in principle we can realize any valid unitary, making this a very powerful tool. The pulse shape to implement a desired unitary can be found via the GRAPE algorithm (GRAdient Ascent Pulse Engineering [Heeres et al., 2017]) and, given suitable constraints on the time, amplitude and bandwidth of the pulse, this can find the pulse shapes to implement any valid unitary. Such pulses are referred to as OCT pulses.

Rather than specify the full unitary, we instead satisfy the state transfer we wish to achieve. (Making sure that a valid unitary exists to perform the state transfer!) Broadly, there are three types of unitaries for which OCT pulses are routinely used. These are:

- State Preparation. e.g. preparing a single photon state in the cavity

$$|0\rangle \otimes |g\rangle \rightarrow |1\rangle \otimes |g\rangle \quad (4.42)$$

- Logical gates on a bosonic qubit encoded in the cavity (If we allow $\epsilon_a(t) \neq 0$, these no-longer have to be excitation preserving). E.g. the X gate for the Binomial code

$$\begin{aligned} |2\rangle \otimes |g\rangle &\rightarrow \frac{|0\rangle + |4\rangle}{\sqrt{2}} \otimes |g\rangle \\ \frac{|0\rangle + |4\rangle}{\sqrt{2}} \otimes |g\rangle &\rightarrow |2\rangle \otimes |g\rangle \end{aligned} \quad (4.43)$$

- Mapping states of the cavity to the transmon for measurement (after we could readout the transmon state) e.g. reading out $|\pm\alpha\rangle$

$$\begin{aligned} |\alpha\rangle \otimes |g\rangle &\rightarrow |0\rangle \otimes |g\rangle \\ |-\alpha\rangle \otimes |g\rangle &\rightarrow |0\rangle \otimes |e\rangle \end{aligned} \quad (4.44)$$

The length of the pulse is predetermined at the start of the optimization. The quantum speed limit of the pulse is again qualitatively set by $\sim 1/\chi$. If the pulse duration is too short, GRAPE will never find a pulse that performs the state transfer to high fidelity. For further reading on the GRAPE algorithm and OCT pulses, see [Heeres et al., 2017, Reinhold, 2019]. In Fig. 4.13 we show two examples of OCT pulses used in Chapter 7.

There are three main disadvantages of OCT pulses. The first is that the pulse waveforms are highly counter-intuitive. A lot of the pulse shapes look random. OCT pulses can perform complex unitaries on the cavities quickly (much quicker than the SNAP + displacement constructions), but we rarely understand *why* they work.

The second is that OCT pulses must be custom-generated for every physical sys-

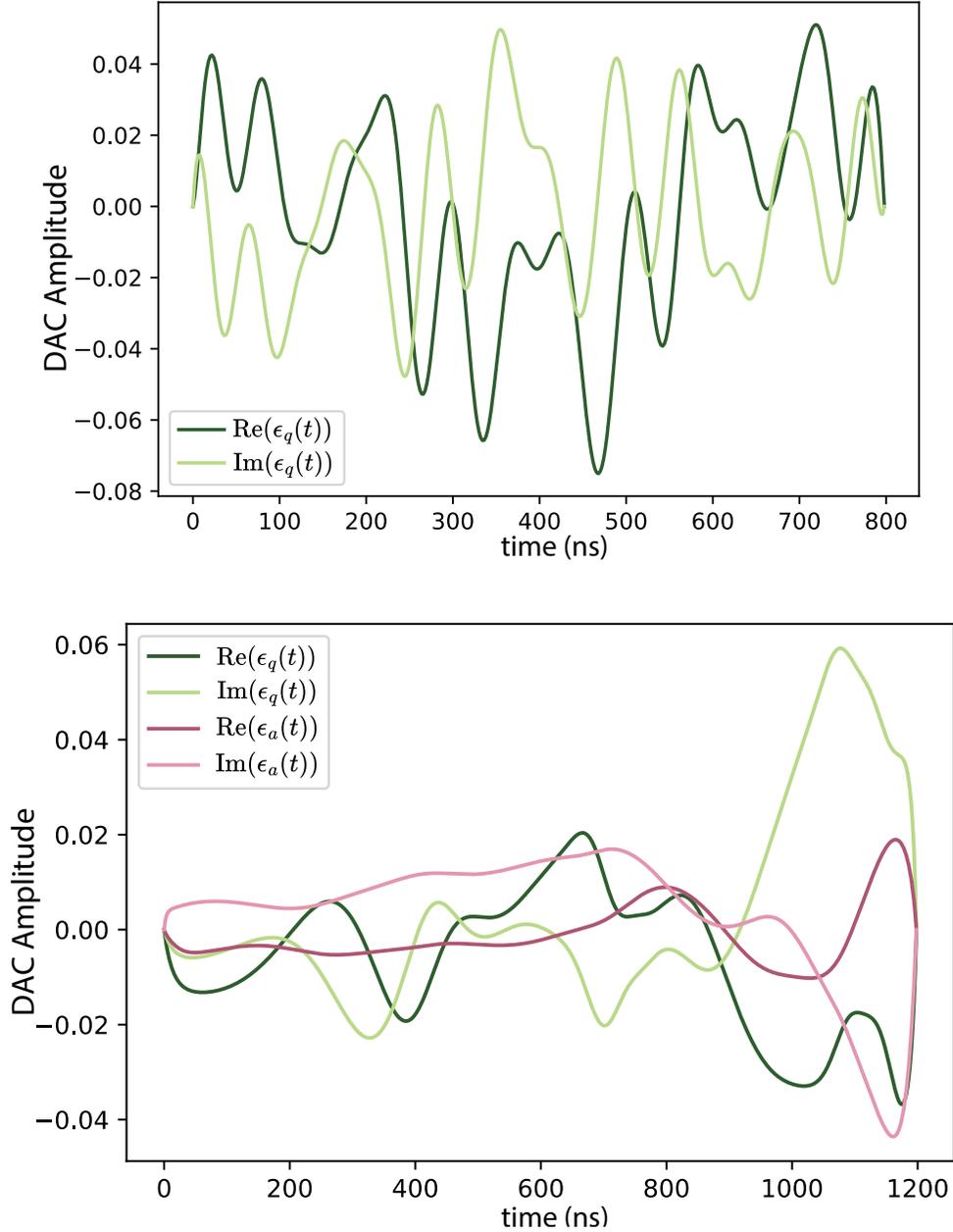


Figure 4.13: Example OCT pulse waveforms. Pulses are centered around ω_{ge} or ω_a depending on the mode being driven. (Top) OCT SNAP pulse optimized to perform the state transfer $|\alpha\rangle \otimes |g\rangle \rightarrow (|\alpha\rangle + i|-\alpha\rangle)/\sqrt{2} \otimes |g\rangle$ for $\alpha = \sqrt{2}$. OCT SNAP pulses fair better (shorter duration, higher fidelity) than analytical SNAP pulse waveforms since OCT does not need to assume a comb of infinitely selective π -pulses. Only drives on the transmon are needed since the cavity photon number distribution is unchanged. (Bottom) OCT pulse to achieve the state transfer $|0\rangle \otimes |g\rangle \rightarrow (|\alpha\rangle + |-\alpha\rangle)/\sqrt{N_\alpha^+} \otimes |g\rangle$. Although the pulse ‘does the job’ its waveform is not at all intuitive. Changing the photon number distribution of the cavity requires simultaneous drives on the transmon and cavity modes. These pulses are generally not HFT to any jump errors in the transmon or cavity.

tem. An OCT pulse will only work for the specific values of the control Hamiltonian we use to optimize it (in particular, the value of χ , K and χ'). This is because OCT pulses are not parameterizable in terms of these system variables, unlike SNAP pulses.

Thirdly and arguably most importantly is that OCT pulses are not inherently hardware fault-tolerant (HFT). Neither of the three dominant hardware errors (transmon dephasing, transmon decay and photon loss) are correctable or even detectable after an OCT pulse. In general, the coherence-limited fidelity of an OCT pulse will be $\sim T_2^q/T_{\text{pulse}}$. In Chapter 5 we explore in more detail why standard GRAPE does not produce pulses that are HFT, and how the optimization constraints can be modified to improve on this.

4.6 Controllable Beamsplitter Interactions

4.6.1 Realizing a parametric beamsplitter

The last ‘tool’ in the cQED toolkit we will introduce is a controllable parametric beamsplitter interaction between two bosonic modes \hat{a} and \hat{b} . This is described by the Hamiltonian

$$\hat{\mathcal{H}}_{\text{BS}} = \frac{g_{\text{BS}}(t)}{2} \hat{a}^\dagger \hat{b} + \frac{g_{\text{BS}}^*(t)}{2} \hat{a} \hat{b}^\dagger \quad (4.45)$$

Where $g_{\text{BS}}(t)$ denoted the strength of the beamsplitter interaction. This interaction represents a bilinear coupling between the two bosonic cavity modes. This Hamiltonian looks similar to the vacuum-Rabi coupling term in the Jaynes-Cummings. However, the advantage of a parametric coupling is its tunability. We cannot turn off g_c in the Jaynes-Cummings Hamiltonian, it is static. However, with a parametric beamsplitter interaction we can control its phase and amplitude in time.

The key to being able to engineer a beamsplitter interaction is that the (dressed) cavity modes \hat{a} and \hat{b} should participate in the same non-linear element. By driving

the phase across one (or more) Josephson junction with off-resonant classical drive tones (or a single tone depending on the element) we can activate parametric processes that are otherwise off-resonant under the RWA. This is typically achieved via 4-wave or 3-wave mixing. 4-wave mixing relies on a 4th order non-linearity (e.g., from a transmon or SQUID coupler) and the beamsplitter process is actuated when two off-resonant drives at frequencies ω_x and ω_y , which we now refer to as the pumps, satisfy the frequency matching condition

$$|\omega_y - \omega_x| = |\omega_a - \omega_b| \quad (4.46)$$

where the frequency difference between the pumps must match the frequency difference between the cavities. In a 3-wave mixing element, such as a SNAIL we can use a 3rd order non-linearity instead and a single pump at frequency

$$\omega_x = |\omega_a - \omega_b| \quad (4.47)$$

Recent advances in coupler design have resulted in the advent of extremely high-fidelity parametric beamsplitter interactions, with very low errors during beamsplitter operations (99.98% fidelity has been demonstrated! [Lu et al., 2023]). The works in chapters 5 and 6 presume near-perfect beamsplitters as a starting point. We will not discuss in detail the operation of these next-gen beamsplitter couplers. Instead, we elucidate the general physics principles by showing how 4-wave mixing in a transmon (called the transmon coupler) can generate a parametric beamsplitter interaction.

We begin with the static Hamiltonian in the dressed mode basis for two cavity modes \hat{a} and \hat{b} dispersively coupled to the same transmon and add two terms for the (constant) pump tones applied off resonantly that displace the transmon mode:

$$\begin{aligned}
\mathcal{H}/\hbar &= \omega_a \hat{a}^\dagger \hat{a} + \omega_b \hat{b}^\dagger \hat{b} + \omega_q \hat{q}^\dagger \hat{q} \\
&\quad - \frac{E_J}{\hbar} \cos_{\text{nl}} \left(\varphi_q (\hat{q} + \hat{q}^\dagger) + \varphi_a (\hat{a} + \hat{a}^\dagger) + \varphi_b (\hat{b} + \hat{b}^\dagger) \right) \\
&\quad + \epsilon_x e^{-i\omega_x t} \hat{q} + \epsilon_y e^{-i\omega_y t} \hat{q} + \text{h.c.}
\end{aligned} \tag{4.48}$$

Where now $\varphi_q \approx (2E_c/E_J)^{1/4}$ since we are in the dressed mode basis and now we have additional terms φ_a and φ_b because the dressed cavity modes now participate in the Josephson junction. These values are determined by how much each bare cavity mode hybridizes with the bare transmon mode i.e. $\varphi_a \approx \frac{g_a}{\Delta_a} \varphi_q$ and $\varphi_b \approx \frac{g_b}{\Delta_b} \varphi_q$

To see how a resonance condition arises in this Hamiltonian we must first go to the rotating frame for each of the dressed modes, and then the displaced frame of the drives. In this frame the Hamiltonian takes the form

$$\begin{aligned}
\mathcal{H} &= -E_J \cos_{\text{nl}} \left(\varphi_q (\hat{q} e^{-i\omega_q t} + \hat{q}^\dagger e^{i\omega_q t} + \xi_x e^{-i\omega_x t} + \xi_x^* e^{i\omega_x t} + \xi_y e^{-i\omega_y t} + \xi_y^* e^{i\omega_y t}) \right. \\
&\quad \left. + \varphi_a (\hat{a} e^{-i\omega_a t} + \hat{a}^\dagger e^{i\omega_a t}) + \varphi_b (\hat{b} e^{-i\omega_b t} + \hat{b}^\dagger e^{i\omega_b t}) \right)
\end{aligned} \tag{4.49}$$

Where we have rewritten the pump terms as dimensionless pump strengths⁹

$$\xi_i = \frac{\epsilon_i}{\omega_q - \omega_i} \tag{4.51}$$

We can truncate the cosine at 4th order such that

$$\begin{aligned}
\mathcal{H}^{(4)} &= -\frac{E_J}{24} \left(\varphi_q (\hat{q} e^{-i\omega_q t} + \hat{q}^\dagger e^{i\omega_q t} + \xi_x e^{-i\omega_x t} + \xi_x^* e^{i\omega_x t} + \xi_y e^{-i\omega_y t} + \xi_y^* e^{i\omega_y t}) \right. \\
&\quad \left. + \varphi_a (\hat{a} e^{-i\omega_a t} + \hat{a}^\dagger e^{i\omega_a t}) + \varphi_b (\hat{b} e^{-i\omega_b t} + \hat{b}^\dagger e^{i\omega_b t}) \right)^4
\end{aligned} \tag{4.52}$$

9. This formula breaks down when $\omega_q \approx \omega_i$ and needs to be modified to be

$$\xi_i = \frac{\epsilon_i}{i\kappa_q/2 + \omega_q - \omega_i} \tag{4.50}$$

where $\kappa_q = 1/T_1^q$

Writing out this expansion produces many terms, but most are fast rotating terms i.e. they are rotating at fast frequencies and so are effectively off-resonant, and do not contribute significantly to the dynamics of interest.

The term of interest for the beamsplitter interaction is given by

$$\hat{\mathcal{H}}_{\text{BS}} = -E_J \varphi_q^2 \varphi_a \varphi_b \xi_x \xi_y^* \hat{a} \hat{b}^\dagger e^{-i(\omega_a - \omega_b - \omega_x + \omega_y)t} + \text{h.c.} \quad (4.53)$$

This term is usually off resonance, unless the frequency matching condition $|\omega_a - \omega_b| = |\omega_y - \omega_x|$ is satisfied. In this case, the Hamiltonian takes the form of the resonant beamsplitter interaction.

$$\begin{aligned} \hat{\mathcal{H}}_{\text{BS}} &= -E_J \varphi_q^2 \varphi_a \varphi_b \left(\xi_x \xi_y^* \hat{a} \hat{b}^\dagger + \xi_x^* \xi_y \hat{a}^\dagger \hat{b} \right) \\ &= \frac{\hbar g_{\text{BS}}}{2} \hat{a} \hat{b}^\dagger + \frac{\hbar g_{\text{BS}}^*}{2} \hat{a}^\dagger \hat{b} \end{aligned} \quad (4.54)$$

The parametric beamsplitter interaction takes the form

$$\frac{\hat{\mathcal{H}}_{\text{BS}}}{\hbar} = \frac{g_{\text{BS}}(t)}{2} \hat{a} \hat{b}^\dagger + \frac{g_{\text{BS}}(t)^*}{2} \hat{a}^\dagger \hat{b} \quad (4.55)$$

Which can be achieved by either varying pump x or pump y in time i.e. $\xi_x \rightarrow \xi_x(t)$ or $\xi_y \rightarrow \xi_y(t)$. Note how we only need to modulate one of the two pumps to modulate the amplitude and phase of $g_{\text{BS}}(t)$.

What sets the strength of g_{BS} ? There is a maximum amount of ‘phase’ that can be driven across the junction by our parametric pumps. If the total phase across the junction exceeds π then the state of the transmon mode will no-longer be confined to a single well of the cosine potential. A good indicator for when we are approaching this limit is when the total stark shift is of the order of the transmon’s anharmonicity. For two pumps of equal (normalized) pump strength, this occurs when

$$|\xi_x|^2 + |\xi_y|^2 = 0.5 \quad (4.56)$$

or when $\xi_x = \xi_y = 0.5$. We can predict how fast the beamsplitter rate will be by writing g_{BS} in terms of the dispersive shifts between each cavity mode and the transmon:

$$\frac{g_{\text{BS}}}{2} = E_J \varphi_q^2 \varphi_a \varphi_b \xi_x \xi_y^* = \sqrt{\chi_{aq} \chi_{bq}} \xi_x \xi_y^* \quad (4.57)$$

Choosing typical values $\chi_{aq} = \chi_{bq} = 2\pi \times 2$ MHz gives a maximum $g_{\text{BS}} \sim 2\pi \times 1$ MHz, which will be important throughout this thesis as it sets the timescale for interactions between cavity modes. Next-gen beamsplitter couplers can reach values of $g_{\text{BS}} \sim 2\pi \times 10$ MHz [Lu et al., 2023, Chapman et al., 2022].

There are other terms that are always resonant, even when the pumps are off. Examples of such terms are the self-Kerr terms

$$\begin{aligned} \mathcal{H}_{\text{self-Kerr}} &= -\frac{E_J}{4} \varphi_q^4 \hat{q}^\dagger \hat{q} \hat{q} \hat{q} - \frac{E_J}{4} \varphi_a^4 \hat{a}^\dagger \hat{a} \hat{a} \hat{a} - \frac{E_J}{4} \varphi_b^4 \hat{b}^\dagger \hat{b} \hat{b} \hat{b} \\ &= \frac{\hbar \alpha_q}{2} \hat{q}^\dagger \hat{q} \hat{q} \hat{q} + \frac{\hbar K_a}{2} \hat{a}^\dagger \hat{a} \hat{a} \hat{a} + \frac{\hbar K_b}{2} \hat{b}^\dagger \hat{b} \hat{b} \hat{b} \end{aligned} \quad (4.58)$$

and cross-Kerr terms

$$\begin{aligned} \mathcal{H}_{\text{cross-Kerr}} &= -E_J \varphi_q^2 \varphi_a^2 \hat{q}^\dagger \hat{q} \hat{a}^\dagger \hat{a} - E_J \varphi_q^2 \varphi_b^2 \hat{q}^\dagger \hat{q} \hat{b}^\dagger \hat{b} - E_J \varphi_a^2 \varphi_b^2 \hat{a}^\dagger \hat{a} \hat{b}^\dagger \hat{b} \\ &= \hbar \chi_{aq} \hat{q}^\dagger \hat{q} \hat{a}^\dagger \hat{a} + \hbar \chi_{bq} \hat{q}^\dagger \hat{q} \hat{b}^\dagger \hat{b} + \hbar \chi_{ab} \hat{a}^\dagger \hat{a} \hat{b}^\dagger \hat{b} \end{aligned} \quad (4.59)$$

Where χ_{aq} and χ_{bq} are the dispersive shifts and χ_{ab} is the often unwanted Cross-Kerr term between the cavities.

Some terms are always resonant whenever the pump strengths are non-zero. These are Stark shift terms that cause the mode frequency to shift with total pump power:

$$\begin{aligned} \mathcal{H}_{\text{Stark}} &= -E_J \varphi_q^4 (|\xi_x|^2 + |\xi_y|^2) \hat{q}^\dagger \hat{q} \\ &\quad - E_J \varphi_q^2 \varphi_a^2 (|\xi_x|^2 + |\xi_y|^2) \hat{a}^\dagger \hat{a} \\ &\quad - E_J \varphi_q^2 \varphi_b^2 (|\xi_x|^2 + |\xi_y|^2) \hat{b}^\dagger \hat{b} \end{aligned} \quad (4.60)$$

The Stark shift on mode \hat{q} is always significantly larger than the Stark shifts on the

bosonic modes. These Stark shifts also shift the resonance condition for the beam splitter interaction at different pump strengths and thus often needs to be calibrated in experiment whenever the pump amplitudes are changed.

The Stark shift is in some ways both a blessing and a curse. It gives us an easy way to find ξ_i directly in experiment, by measuring the frequency shift on the transmon mode (and dividing by the transmon’s anharmonicity which is also measurable). However, the Stark shift can prove quite a nuisance. The mode frequencies of the system shift depending on the amplitude of the pumps, which makes predicting the exact beamsplitter resonance condition difficult. In practice we must retune the beamsplitter resonance condition whenever we change the pump amplitudes.

Even more nefarious is the problem of frequency collisions during pump ‘ring up’. When we activate a beamsplitter interaction, we must ‘ring up’ the pumps from zero amplitude to their constant amplitude somewhat gradually (in around 50 ns, otherwise the large pump bandwidth is also problematic). During this time, the transmon mode gradually shifts down in frequency due to the Stark shift term, often by ~ 100 MHz. It is very easy to activate unwanted parametric resonances, if only for a short time. The unwanted dynamics that occur when we ‘cross’ one of these spurious resonances can significantly reduce the fidelity of our beamsplitter operations and/or the size of g_{BS} we can engineer. An advantage of couplers 3-wave mixing based on 3-wave mixing such as the SNAIL is that the Stark shift can be made negligible.

Sometimes we are interested in a more accurate model of the system that goes beyond 4th order and RWA approximations. These dynamics can be found through Floquet simulations, which can study more complex effects such as decoherence inherited by the modes when the parametric resonance is activated. For more reading see [Zhang et al., 2019].

Expanding the cosine reveals a hierarchy of rates for wanted and unwanted terms in our cavity-transmon system, set by the all-important $\left(\frac{g_c}{\Delta}\right)$ parameter. This can be

summarised succinctly in the table below

Symbol	Rate	Typical value ($\times 2\pi$)	Use
α_q	$\sim \alpha_q$	200 MHz	Rotations of control ancilla
$\chi_{aq}, \chi_{bq}, g_{BS}$	$\sim \left(\frac{g_c}{\Delta}\right)^2 \alpha_q$	0.5-5 MHz	Quantum control of cavities, Cavity-cavity interactions
K_a, K_b, χ_{ab}	$\sim \left(\frac{g_c}{\Delta}\right)^4 \alpha_q$	1-20 kHz	Unwanted cavity non-linearities

Table 4.1: Hierarchy of Hamiltonian rates for control of bosonic modes in cQED

4.6.2 Evolution from a parametric beamsplitter interaction

There are two possible conventions for defining the beamsplitter rate, g_{BS} . For theoretical analysis, we adopt the convention

$$\widehat{\mathcal{H}}_{BS} = \frac{g_{BS}}{2} \hat{a}^\dagger \hat{b} + \frac{g_{BS}}{2} \hat{a} \hat{b}^\dagger \quad (4.61)$$

which is chosen such that a π -pulse e.g. applying $\widehat{\mathcal{H}}_{BS}$ for time $T = \pi/g_{BS}$ enacts a SWAP operation. The experimentalist definition, used in Chapter 7 and [Lu et al., 2023, Chapman et al., 2022, Gao et al., 2018] is instead

$$\widehat{\mathcal{H}}_{BS} = g_{BS} \hat{a}^\dagger \hat{b} + g_{BS} \hat{a} \hat{b}^\dagger \quad (4.62)$$

Taking the theorist's definition, if we instead apply this Hamiltonian for half the time, $T = \pi/2g_{BS}$ we realize the mode transformations (in the Heisenberg picture)

$$\begin{aligned} \hat{a} &\rightarrow \frac{\hat{a} + i\hat{b}}{\sqrt{2}} \\ \hat{b} &\rightarrow \frac{\hat{b} + i\hat{a}}{\sqrt{2}} \end{aligned} \quad (4.63)$$

which are the same mode transformations as for a 50-50 beamsplitter in quantum optics, hence the name beamsplitter interaction. Realizing these mode transformations via a parametric process rather than passing light through a physical beamsplitter gives us more capabilities.

By adjusting the phase of one of the pumps, we can control the phase of the beamsplitter transformation, e.g., applying

$$\hat{\mathcal{H}}_{\text{BS}} = \frac{g_{\text{BS}}}{2} e^{i\phi} \hat{a}^\dagger \hat{b} + \frac{g_{\text{BS}}^*}{2} e^{-i\phi} \hat{a} \hat{b}^\dagger \quad (4.64)$$

for time $T = \pi/2g_{\text{BS}}$ realizes the mode transformations.

$$\begin{aligned} \hat{a} &\rightarrow \frac{\hat{a} + ie^{i\phi}\hat{b}}{\sqrt{2}} \\ \hat{b} &\rightarrow \frac{\hat{b} + ie^{i\phi}\hat{a}}{\sqrt{2}} \end{aligned} \quad (4.65)$$

Moreover, this transformation is not implemented instantaneously, but continuously. Writing the evolution of the mode operators in the Heisenberg picture is a convenient way to track evolution in a state-independent way. These mode transformations at time t during the evolution (for constant, complex $g_{\text{BS}}e^{i\phi}$) are given by

$$\begin{aligned} \hat{a}(t) &\rightarrow e^{i\hat{\mathcal{H}}_{\text{BS}}t} \hat{a}(0) e^{-i\hat{\mathcal{H}}_{\text{BS}}t} = \cos\left(\frac{g_{\text{BS}}t}{2}\right) \hat{a} + i \sin\left(\frac{g_{\text{BS}}t}{2}\right) e^{i\phi} \hat{b} \\ \hat{b}(t) &\rightarrow e^{i\hat{\mathcal{H}}_{\text{BS}}t} \hat{b}(0) e^{-i\hat{\mathcal{H}}_{\text{BS}}t} = \cos\left(\frac{g_{\text{BS}}t}{2}\right) \hat{b} + i \sin\left(\frac{g_{\text{BS}}t}{2}\right) e^{i\phi} \hat{a} \end{aligned} \quad (4.66)$$

We designate the beamsplitter unitary

$$\hat{U}_{\text{BS}}(t) = e^{-i\hat{\mathcal{H}}_{\text{BS}}t} \quad (4.67)$$

For $t = \pi/2g_{\text{BS}}$ and $\phi = 0$ this takes the form

$$\hat{U}_{\text{BS}}^{50-50} = e^{i\frac{\pi}{4}(\hat{a}^\dagger\hat{b} + \hat{b}^\dagger\hat{a})} \quad (4.68)$$

and for the choice of phase $\phi = \pi/2$ this takes the form.

$$\hat{U}_{\text{BS}}^{50-50} = e^{i\frac{\pi}{4}(\hat{a}^\dagger\hat{b} - \hat{b}^\dagger\hat{a})} \quad (4.69)$$

Going back to the Schrödinger picture, the evolution of states in the cavity modes is given by

$$|\psi(t)\rangle_{a,b} = \hat{U}_{\text{BS}}(t) |\psi(0)\rangle_{a,b} \quad (4.70)$$

Another useful feature of a parametric beamsplitter Hamiltonian (which is not possible to realize with an optical beamsplitter) is that we may detune one of the pumps from resonance to realize the Hamiltonian

$$\hat{\mathcal{H}}_{\text{BS}}(t)/\hbar = \frac{g_{\text{BS}}}{2} e^{i\phi} \hat{a}^\dagger \hat{b} + \frac{g_{\text{BS}}^*}{2} e^{-i\phi} \hat{a} \hat{b}^\dagger + \Delta \hat{a}^\dagger \hat{a} \quad (4.71)$$

for constant, complex g_{BS} with a fixed detuning, Δ from the resonance condition. The sign of Δ may be positive or negative depending on whether our pump frequency is above or below its resonance condition frequency. The mode we choose to pick up this detuning is arbitrary. In a different rotating frame, we could write this same Hamiltonian as

$$\hat{\mathcal{H}}_{\text{BS}}(t)/\hbar = \frac{g_{\text{BS}}}{2} e^{i\phi} \hat{a}^\dagger \hat{b} + \frac{g_{\text{BS}}^*}{2} e^{-i\phi} \hat{a} \hat{b}^\dagger - \Delta \hat{b}^\dagger \hat{b} \quad (4.72)$$

Error Detected Gates

This chapter summarizes and contextualizes the work presented in [Tsunoda et al., 2022], on which I am co-first author. The broad goal is to find a way to realize two-qubit entangling gates on discrete-variable (DV) bosonic codewords designed to correct for single photon loss. The implementation must be experimentally feasible and satisfy Hardware Fault-tolerance (HFT). We must still be able to correct single photon loss that occurs before/during the gate as well as first-order control ancilla errors (decay and dephasing).

Ultimately, we find two solutions to this problem. One is to realize two-qubit gates via a measurement-based approach, which requires a QND HFT ZZ measurement. However, this solution is not entirely complete since we would also need an HFT Hadamard gate (which presents a major problem in and of itself for discrete variable bosonic codewords). We show that a fully-error corrected ZZ measurement is possible but requires physically performing a long string of measurements in time to ensure all first-order errors can be caught and corrected, which introduces a substantial overhead.

We then show that error-*detected* two-qubit gates are much easier to realize and satisfy HFT. Whilst error-detection is not the same as error-correction, in Chapter 3 we saw how error-detection at the hardware level gives a path for realizing erasure qubits. When we concatenate erasure qubits to form a larger error-correction code, we saw they have much the same benefit as if we had error-corrected the dominant hardware errors instead.

This approach naturally leads one to ask: why don't we use error-detectable operations with error-detectable bosonic codes? (Rather than error-correctable ones like the Kitten code or 4-legged cat). This leads us to Chapter 6, where we build up an entire set of error-detectable operations for the dual-rail code, an encoding in which loss errors can only be detected but not corrected. Two crucial operations in Chapter 6: the mid-circuit erasure detection and the entangling gate will be based on the gates we present in this chapter.

5.1 Searching for an error-correctable two-qubit gate

For bosonic QEC to be viable with a control ancilla, we must be able to perform *all* operations¹ for universal quantum computing whilst satisfying HFT. In our case, this means being able to correct the dominant first-order errors on the control ancilla: transmon decay and dephasing, as well as the bosonic error the bosonic code is designed to correct for. We will focus on bosonic codes designed to correct for discrete photon loss events, the so-called rotationally-symmetric bosonic codes. We will further restrict our bosonic codes to those designed to correct for a single photon loss event (such as the Kitten and 4-legged cat code.)

We will also use a three-level ancilla, which effectively functions as a noise-biased ancilla by reserving the $|e\rangle$ level to detect and correct for a single decay event in

1. State preparation, readout, single-qubit, two-qubit gates, error correction operations.

the transmon. This forms the set of Hardware errors $\{\hat{\epsilon}\}_{\text{hardware}} = \{|e\rangle\langle f|, |g\rangle\langle g| - |f\rangle\langle f|, \hat{a}\}$ we wish to correct.

We have seen previously how engineering the general Hamiltonian

$$\hat{\mathcal{H}}_{\text{HFT}} = \hat{U}_{\text{gate}} |f\rangle\langle g| + \hat{U}_{\text{gate}}^\dagger |g\rangle\langle f| \quad (5.1)$$

realizes \hat{U}_{gate} on the bosonic mode in a path-independent manner, allowing us to correct for the hardware errors in $\{\hat{\epsilon}\}_{\text{hardware}}$ after the gate has been attempted. In order to correct for transmon decay events, it is important that $\hat{\mathcal{H}}_{\text{HFT}} |e\rangle = 0$ so that we do not incur any unwanted, time-dependent evolution after a transmon decay event has occurred.

Engineering such a Hamiltonian is difficult; remember our error model assumes the only source of errors are jump errors from $\{\hat{\epsilon}\}_{\text{hardware}}$ and that otherwise the $\hat{\mathcal{H}}_{\text{HFT}}$ Hamiltonian is implemented perfectly. We are not robust to unwanted terms or interactions that arise when trying to engineer this Hamiltonian.

The SNAP unitary construction only approximately engineers this Hamiltonian, in the regime where χ -matching is achieved (either $\chi_{ge} = \chi_{gf}$ or $\chi_{ge} = 0$ are satisfactory) and in the limit of long selective π -pulses, where the n^{th} frequency in the frequency comb only acts on the space $|n\rangle\langle n| \otimes (|g\rangle\langle g| + |f\rangle\langle f|)$. SNAP unitaries are capable of performing $Z(\theta)$ gates with HFT on rotationally symmetric bosonic codes, but can we find an extension of this construction to two-qubit gates? We shall find this is in principle possible but comes with a daunting set of experimental challenges.

5.1.1 Two-mode SNAP gates

The first natural thought may be to engineer this Hamiltonian as a two-mode SNAP unitary, which can be realized by dispersively coupling a transmon to two cavity modes. Such a transmon is often called a ‘Y-mon’ owing to its Y-shaped antenna

geometry. As shown in Fig. 5.1, a Y-mon transmon needs to couple to three cavity modes, the two modes that store bosonic qubits and its readout resonator.

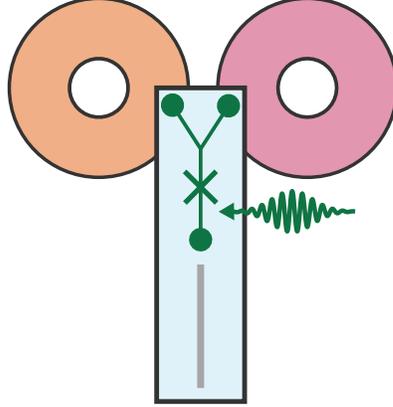


Figure 5.1: Y-mon transmon for coupling two cavities together. By driving only the transmon mode, it is possible, in theory to perform an HFT two-mode SNAP gate. In practice, χ -matching conditions make this difficult.

The Hamiltonian needed for a two-mode SNAP unitary is written as

$$\begin{aligned} \widehat{\mathcal{H}}_{2\text{-mode}} = & \chi_{aq}^{ge} \hat{a}^\dagger \hat{a} |e\rangle \langle e| + \chi_{bq}^{ge} \hat{b}^\dagger \hat{b} |e\rangle \langle e| + \chi_{aq}^{gf} \hat{a}^\dagger \hat{a} |f\rangle \langle f| + \chi_{bq}^{gf} \hat{b}^\dagger \hat{b} |f\rangle \langle f| \\ & + \epsilon_q(t) \hat{q} + \epsilon_q^*(t) \hat{q}^\dagger + \frac{\alpha_q}{2} \hat{q}^\dagger \hat{q}^\dagger \hat{q} \hat{q} \end{aligned} \quad (5.2)$$

Where the subscripts now denote which modes the dispersive interaction acts on and the superscripts indicate the manifold of the transmon. The shift in the transmon transition frequency now depends on the number of photons in both modes, e.g., the frequency of the g - e transition is given by $\omega_{ge} + \chi_{aq}^{ge} n_a + \chi_{bq}^{ge} n_b$ for n_a photons in mode a and n_b photons in mode b .

In order to perform a two-mode SNAP gate, we want these frequencies to be uniquely addressable for each joint number state $|n_a\rangle_a \otimes |n_b\rangle_b = |n_a, n_b\rangle$. Herein lies a problem. No matter our choice of χ_{aq}^{ge} and χ_{bq}^{ge} , sooner or later we will find two different joint number states which have the same frequency shift and are indistinguishable. Ideally we want our χ s to be incommensurate², such that these frequency overlaps

2. The optimal ratio of χ_{aq}^{ge} to χ_{bq}^{ge} may be the golden ratio to avoid these overlaps

occur at high photon number states that are not used in our bosonic codewords.

If we take the example of two bosonic qubits in the kitten code, then all possible joint-photon number states are $|0, 0\rangle, |0, 2\rangle, |2, 0\rangle, |2, 2\rangle, |4, 2\rangle, |2, 4\rangle, |4, 4\rangle, |0, 4\rangle, |4, 0\rangle$. We must engineer our χ s such that all 9 of these states have distinct frequencies. The minimum frequency spacing sets how selective our π -pulses must be (and in turn the length of our 2-mode SNAP gate).

What happens if we instead try to engineer $\chi_{aq}^{ge} = \chi_{bq}^{ge}$ through χ -matching? If this were the case we cannot distinguish for example the states $|2, 2\rangle, |0, 4\rangle, |4, 0\rangle$ in the frequency comb. Whilst it is not possible to perform an arbitrary SNAP unitary, we can still perform SNAP unitaries that implement logical two-qubit entangling gates on this code. This is because it is still possible to engineer a subset of SNAP gates where one SNAP unitary is applied to the manifold $\{|0_L, 0_L\rangle, |1_L, 1_L\rangle\}$ and another to the $\{|0_L, 1_L\rangle, |1_L, 0_L\rangle\}$ manifold of the Kitten code. For example, if we acquire a phase of $\pi/2$ on joint-number states $|2, 2\rangle, |0, 4\rangle, |4, 0\rangle, |4, 4\rangle, |0, 0\rangle$ and a phase of 0 on the joint-number states $|0, 2\rangle, |2, 0\rangle, |4, 2\rangle, |2, 4\rangle$ (all of which is possible to engineer when $\chi_{aq}^{ge} = \chi_{bq}^{ge}$), then we will have performed the $ZZ(\pi/2)$ entangling gate on the Binomial code.

Now, to make the two-mode SNAP gate fault tolerant to transmon decay we must use the g - f manifold which necessitates even more χ -matching. Overall, it is possible to engineer a two-mode SNAP gate with the same HFT properties as the one-mode SNAP gate, but many different χ -matching conditions must be simultaneously satisfied which would be incredibly difficult to engineer in practice. Valid χ -matching conditions that allow us to perform, say the $ZZ(\pi/2)$ gate are shown in Table 5.1.

χ s commensurate?	yes (equal)	yes (all equal)	no	no
χ -matching conditions	$\chi_{aq}^{gf} = \chi_{bq}^{gf}$ $\chi_{aq}^{ge} = \chi_{bq}^{ge} = 0$	$\chi_{aq}^{gf} = \chi_{bq}^{gf}$ $= \chi_{aq}^{ge} = \chi_{bq}^{ge}$	$\chi_{aq}^{gf} \neq \chi_{bq}^{gf}$ $\chi_{aq}^{ge} = \chi_{bq}^{ge} = 0$	$\chi_{aq}^{gf} \neq \chi_{bq}^{gf}$ $\chi_{aq}^{ge} = \chi_{bq}^{ge}$ $\chi_{aq}^{ge} \neq \chi_{bq}^{ge}$

Table 5.1: Four of the allowed χ -matching conditions for two-mode SNAP with a three-level transmon as the control ancilla.

5.1.2 Is eSWAP Hardware Fault-Tolerant?

Can other easier-to-realize two-qubit gate constructions satisfy HFT and be error-corrected instead? One prominent two-qubit entangling gate for bosonic qubits is the e SWAP(θ) gate that was realized in [Gao et al., 2019]. When $\theta = \pi/2$ it is capable of generating Bell states even when the bosonic modes are in different encodings. But how does this gate fare against errors? The cQED construction of the e SWAP gate uses the exponentiation gadget as shown in Fig. 5.2, which is known to only be able to *detect* errors rather than correct them. Specifically, when we use the exponentiation gadget to implement gate $\hat{O}(\theta)$ for $\hat{O}^2 = \mathbb{1}$, we widen our error set. If our hardware jump error was originally $\hat{\epsilon}_i$, then using the exponentiation gadget means we would now experience errors from the set $\{\hat{\epsilon}_i, \hat{\epsilon}_i \hat{O}\}$ (and superpositions of these errors) after the gate. Generally it is not possible to correct for errors of the form $c_0 \hat{\epsilon}_i \mathbb{1} + c_1 \hat{\epsilon}_i \hat{O}$ at the hardware level (we can detect $\hat{\epsilon}_i$ but not correct since we don't know whether $\mathbb{1}$ or \hat{O} was applied to the bosonic mode.)

For our e SWAP gate example (with $\hat{O} = \text{SWAP}$), suppose we wish to correct for photon loss on either cavity \hat{a} or \hat{b} during the gate. At the end of the gate, we can *detect* this error by looking for photon loss in both cavities via photon number parity measurements. However, upon detecting a photon loss, it is not possible to correct this error, since there is no way of knowing whether to apply $\mathbb{1}$ or SWAP to the cavities to correct the error. This is the first example of a wider family of error-detectable (and excitation-preserving) two-qubit gates, with the other being the $ZZ(\theta)$ gate.

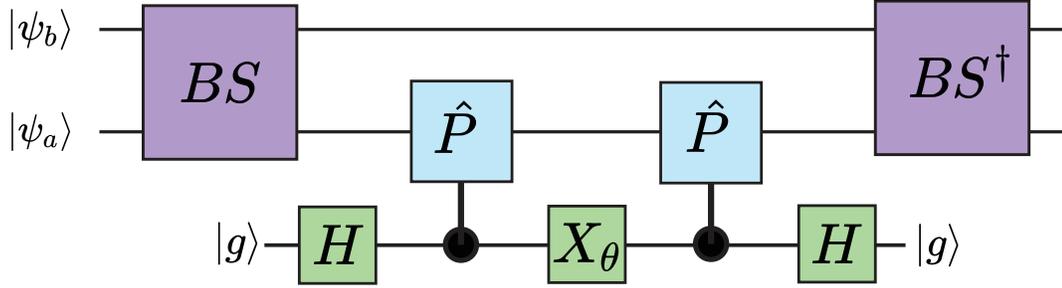


Figure 5.2: Gate construction for e SWAP. This is essentially the exponentiation gadget to implement the gate $\hat{P}(\theta)$ on cavity mode \hat{a} where \hat{P} is the parity operator. By sandwiching this operation between two 50-50 beamsplitter unitaries, we ultimately perform the e SWAP(θ) gate on both cavity modes \hat{a} and \hat{b} . The e SWAP(θ) gate thus inherits all the fault-tolerant properties of an exponentiation gadget. In particular, it is good at detecting first-order jump errors but not correcting them. The transmon ancilla starts and ends in $|g\rangle$ when no errors occur and can be measured at the end to detect ancilla errors during the gate sequence.

5.1.3 Constructing e SWAP as an exponential gadget

For completeness, we show explicitly how the e SWAP unitary is realized from the circuit in Fig. 5.2. The starting point is an exponentiation gadget on a single bosonic cavity, where $\hat{O} = e^{i\pi\hat{a}^\dagger\hat{a}} = \hat{P}$ where \hat{P} is the parity operator. The parity operator represents a 180° rotation of the cavity's phase space and has the property $\hat{P}|n\rangle = (-1)^n|n\rangle$, i.e., eigenvalue $+1$ if $n = \text{even}$, -1 if $n = \text{odd}$. To perform the exponentiation gadget, we must realize the control-parity unitary,

$$\mathbf{cP} = |g\rangle\langle g| + e^{i\pi\hat{a}^\dagger\hat{a}}|e\rangle\langle e| \quad (5.3)$$

Fortunately, this is almost trivial to realize with the dispersive interaction. If the control ancilla is dispersively coupled to a cavity mode via the Hamiltonian

$$\hat{\mathcal{H}}_\chi = \chi_{ge}\hat{a}^\dagger\hat{a}|e\rangle\langle e| \quad (5.4)$$

evolution for time $T = \pi/\chi_{ge}$ generates the control-parity unitary. Embedding this in the exponentiation gadget shown in Eq. 2.54 realizes the $\hat{P}(\theta)$ gate

$$\begin{aligned}\hat{P}(\theta) &= \cos\left(\frac{\theta}{2}\right) \mathbb{1} + i \sin\left(\frac{\theta}{2}\right) \hat{P} \\ &= \cos\left(\frac{\theta}{2}\right) \mathbb{1} + i \sin\left(\frac{\theta}{2}\right) e^{i\pi\hat{a}^\dagger\hat{a}}\end{aligned}\tag{5.5}$$

So far, this gate only acts on one bosonic mode. e SWAP acts on two bosonic modes. By sandwiching the $\hat{P}(\theta)$ gate between a 50-50 beamsplitter and inverse 50-50 beamsplitter, we realize the e SWAP unitary

$$\begin{aligned}e\text{SWAP}(\theta) &= \left(\hat{U}'_{\text{BS}}{}^{50-50}\right)^\dagger \hat{P}(\theta) \left(\hat{U}'_{\text{BS}}{}^{50-50}\right) \\ &= \cos\left(\frac{\theta}{2}\right) \mathbb{1} + i \sin\left(\frac{\theta}{2}\right) e^{i\pi\left(\frac{\hat{a}^\dagger - \hat{b}^\dagger}{\sqrt{2}}\right)\left(\frac{\hat{a} - \hat{b}}{\sqrt{2}}\right)}\end{aligned}\tag{5.6}$$

Where we have made use of the mode transformations $\hat{a} \rightarrow (\hat{a} - \hat{b})/\sqrt{2}$, $\hat{a}^\dagger \rightarrow (\hat{a}^\dagger - \hat{b}^\dagger)/\sqrt{2}$ for this particular 50-50 beamsplitter convention. Is the operator in the second term in fact the SWAP operator?

We can see this could indeed be the SWAP operator by rewriting

$$\hat{U}' = e^{i\pi\left(\frac{\hat{a}^\dagger - \hat{b}^\dagger}{\sqrt{2}}\right)\left(\frac{\hat{a} - \hat{b}}{\sqrt{2}}\right)} = e^{i\frac{\pi}{2}(\hat{a}^\dagger\hat{a} + \hat{b}^\dagger\hat{b})} e^{-i\frac{\pi}{2}(\hat{a}^\dagger\hat{b} + \hat{a}\hat{b}^\dagger)}\tag{5.7}$$

and we verify its action is indeed to SWAP the bosonic modes by confirming

$$\begin{aligned}\hat{U}'\hat{a}\hat{U}'^\dagger &= \hat{b} \\ \hat{U}'\hat{b}\hat{U}'^\dagger &= \hat{a}\end{aligned}\tag{5.8}$$

which means this gate construction does in fact realize the e SWAP unitary.

5.1.4 Measurements are easier to error-correct?

The key to realizing an HFT two-qubit gate might be to build one out of measurements instead. From Fig. 2.3 QND ZZ measurements and Hadamard gates are sufficient to realize deterministic CNOT gates.

First, we begin by arguing why measurements of operators satisfy HFT more easily than logical gates on a bosonic qubit. Let us briefly summarize the requirements for a gate to be error correctable to a set of Hardware jump errors, $\{\hat{\epsilon}\}_{\text{hardware}}$

1. We must engineer the Hamiltonian $\hat{\mathcal{H}}_{\text{HFT}} = \hat{U}_{\text{gate}} |f\rangle \langle g| + \hat{U}_{\text{gate}}^\dagger |g\rangle \langle f|$ very well
2. We must realize three-state QND readout of the control ancilla, to distinguish between the states $|g\rangle, |e\rangle, |f\rangle$ with high-fidelity. Any readout misassignment errors will result in us performing the wrong recovery operation and so result in an error on the bosonic qubit.

Compare this to the requirements for a measurement on the bosonic qubit to be Hardware Fault-tolerant:

1. If any errors from $\{\hat{\epsilon}\}_{\text{hardware}}$ occur during the measurement procedure, they must be QND on the bosonic qubit.
2. We must be able to repeat the measurement multiple times. Any readout misassignment errors can be mitigating by majority voting on a repeated set of measurements. Errors from $\{\hat{\epsilon}\}_{\text{hardware}}$ may result in a particular measurement giving the wrong outcome, but this can always be suppressed to an n^{th} order error if we are able repeat the measurement $2n - 1$ times in a row.

The way majority voting works is very similar to a repetition code. We illustrate this with the simple example of three-round majority voting. Suppose we wish to measure binary-valued operator \hat{B} on a bosonic mode (or modes) using the circuit from Fig. 2.6. We label the classical outcome of this measurement to be 0_i or 1_i for

the i^{th} round of measurement. We also suppose the measurement is entirely QND on the bosonic mode. An error from $\{\hat{\epsilon}\}_{\text{hardware}}$ occurs with probability p during the measurement, and in our example flips the measurement outcomes causing a measurement error. In one round of measurement, with probability p we declare the wrong outcome (and when we feed forward this outcome to future computations this causes an error on the bosonic qubit).

If we instead perform three rounds of measurement then we obtain strings of outcomes $(0_1, 0_2, 0_3)$ or $(1_1, 1_2, 1_3)$ if no errors (or three errors if we are really unlucky) occur but also mixed strings of outcomes such as $(0_1, 1_2, 0_3)$, where we can majority vote to still declare the correct outcome when a single measurement is faulty. The probability we declare the wrong outcome is $3p^2 + p^3$, just as we saw for the 3-bit repetition code and the error on our bosonic qubit is now suppressed to second order. We summarize these outcomes in Table 5.2. Repeating the measurement

Declare 0	Declare 1
$(0_1, 0_2, 0_3)$	$(1_1, 1_2, 1_3)$
$(1_1, 0_2, 0_3)$	$(0_1, 1_2, 1_3)$
$(0_1, 1_2, 0_3)$	$(1_1, 0_2, 1_3)$
$(0_1, 0_2, 1_3)$	$(1_1, 1_2, 0_3)$

Table 5.2: 3 round majority voting with a QND measurement

several times comes with its own trade-offs. There is a linear time overhead which can be inconvenient if these are mid-circuit measurements. We also cannot repeat the measurement an infinite number of times. Eventually (on the timescale of cavity decoherence) the value of \hat{B} we are trying to measure will actually change, e.g., if we keep repeating parity measurements, eventually we will see the parity value switch due to a photon jump! This sets the timescale on which we can obtain useful information from the outcomes of our repeated measurements.

5.1.5 Errors in photon number parity measurement

We now go through an example of a useful measurement which is QND to jump errors from $\{\hat{e}\}_{\text{hardware}}$. This is the HFT photon number parity measurement realized with the construction in Fig. 5.3 and its QNDness (to cavity photon number) was studied in [Sun et al., 2014]. The photon number parity measurement (parity measurement for short) is useful as the error syndrome measurement for single photon loss in the kitten code/4-legged cat code. (Although these codes are also somewhat robust in the Knill-Laflamme sense to cavity dephasing as well, parity measurements only detect photon loss events.) We study the parity measurement in detail, since the ZZ measurements we introduce later are effectively a two-qubit extension of this measurement operation and are used to construct both error-correctable HFT ZZ measurements and the error-detectable HFT $ZZ(\theta)$ gates.

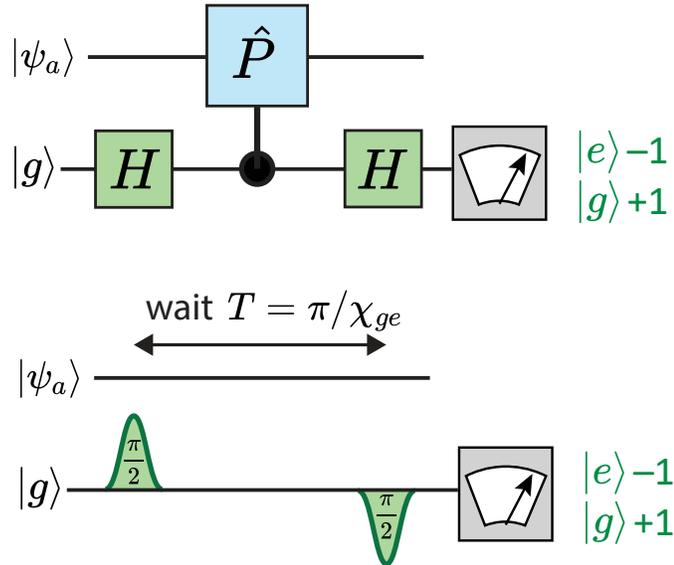


Figure 5.3: Gate construction (top) and pulse sequence (bottom) for a standard parity measurement. The two transmon pulses are assumed to be ‘unselective’, performing the same rotation on the transmon regardless of the cavity state. Here +1 designates the even photon number parity outcome and -1 designates odd photon number parity. By inverting the phase of the final $\pi/2$ -pulse, we can invert these outcomes.

The control-parity unitaries are realized with wait time $T = \pi/\chi_{ge}$, just as they

were for the e SWAP construction. No matter what happens in the control ancilla, the measurement will be QND on the cavity photon number. This is because the ancilla only interacts with the cavity via the dispersive interaction. Even in the presence of errors, this interaction cannot change cavity photon number. The only error that can be imparted on the cavity mode via the dispersive interaction must act on the cavity phase, since it cannot act on cavity photon number. We walk through the example of an HFT parity measurement to show how such a measurement can be made QND to $\{\hat{\epsilon}\}_{\text{hardware}}$.

To begin, we consider the non-HFT parity measurement given by the same circuit construction but in the g - e manifold of the transmon. Our first-order Hardware error set is $\{\hat{\epsilon}\}_{\text{hardware}} = \{\hat{a}, |g\rangle\langle e|, |g\rangle\langle g| - |e\rangle\langle e|\}$. For now, we focus on errors that happen during the longest part of the measurement, the $T = \pi/\chi_{ge}$ wait time in the control-parity unitary.

Transmon Dephasing

First we examine transmon dephasing at unknown time $T - t$ for $t \in [0, T]$ (we choose this convention to simplify the maths slightly). This is the easiest case since it commutes with the dispersive interaction such that

$$\begin{aligned} & e^{-i\hat{\mathcal{H}}_\chi t}(|g\rangle\langle g| - |e\rangle\langle e|)e^{-i\hat{\mathcal{H}}_\chi(T-t)} \\ & = (|g\rangle\langle g| - |e\rangle\langle e|)e^{-i\hat{\mathcal{H}}_\chi T} = (|g\rangle\langle g| - |e\rangle\langle e|)\mathbf{cP} \end{aligned} \tag{5.9}$$

The dispersive Hamiltonian is error-transparent to dephasing. The end result is that it is as if we did the correct \mathbf{cP} unitary, and then applied jump $(|g\rangle\langle g| - |e\rangle\langle e|)$. Ultimately, this results in a flipped measurement outcome ($0_i \leftrightarrow 1_i$) and is QND on the cavity. As a result, this error can be suppressed easily by majority voting.

Cavity photon loss

The next error we consider is photon loss in the cavity. This error is rather peculiar, since it is the error we are usually trying to detect with a parity measurement anyways! What we are interested in is the effect of photon loss during the measurement itself. QND in this case means photon loss during the wait time does not dephase the cavity and can be detected in a subsequent parity measurement, although it may scramble this measurement outcome. The analysis follows the same steps as transmon dephasing.

$$\begin{aligned} e^{-i\hat{\mathcal{H}}_x t} \hat{a} e^{-i\hat{\mathcal{H}}_x (T-t)} &= e^{-i\hat{\mathcal{H}}_x t} \hat{a} e^{i\hat{\mathcal{H}}_x t} e^{-i\hat{\mathcal{H}}_x T} \\ &= \hat{a} (|g\rangle \langle g| + e^{-i\chi_{ge} t} |e\rangle \langle e|) \mathbf{cP} \end{aligned} \quad (5.10)$$

We see that a photon loss in the cavity also dephases the transmon, rotating it by unknown angle $\theta = -\chi_{ge} t$ about its Z axis. (The time of the photon loss is unknown). The overall result is that there is a chance the measurement will give the wrong outcome. Other than the photon loss operator, \hat{a} , the error is completely QND on the cavity state. This means we can detect the photon loss error with subsequent parity measurements—unless of course we suffer another photon loss! (but this is a second order error). Overall, with repeated measurements, we are HFT to photon loss during the parity measurement.

Transmon decay

Now transmon decay is the worst error. It is the only one in $\{\hat{e}\}_{\text{hardware}}$ that is non-QND on the cavity mode. With the same analysis we now find

$$e^{-i\hat{\mathcal{H}}_x t} |g\rangle \langle e| e^{-i\hat{\mathcal{H}}_x (T-t)} = |g\rangle \langle e| e^{-i\hat{\mathcal{H}}_x (T-t)} = |g\rangle \langle e| e^{-i\chi (T-t) \hat{a}^\dagger \hat{a}} \quad (5.11)$$

Before the last Hadamard gate (or $\pi/2$ pulse, they are equivalent in this construction), the transmon will always be in its ground state, meaning, that after the last Hadamard, we measure $|g\rangle$ or $|e\rangle$ with equal probability and our measurement doesn't contain any information about the cavity photon number. This error can be suppressed with majority voting. The disastrous error is the fact we have performed unitary $e^{-i\chi(T-t)\hat{a}^\dagger\hat{a}}$ on the cavity mode, which is a rotation in the cavity's phase space by *unknown* angle $\theta = -\chi(T-t)$. Since $t \in [0, \pi/\chi_{ge}]$, it follows $\theta \in [0, \pi]$ with an approximately uniform distribution. This is a large angle (or 'complete') dephasing and is usually impossible to recover from. In fact, the only difference between the non-HFT and HFT parity measurement is that we can mitigate the effects of transmon decay.

5.1.6 The Hardware Fault-tolerant parity measurement

In the HFT-parity measurement [Rosenblum et al., 2018b], the only difference is that we have a way to correct for first-order transmon decay by using a three-level ancilla and engineering the χ -matched Hamiltonian

$$\widehat{\mathcal{H}}'_\chi = \chi(|e\rangle\langle e| + |f\rangle\langle f|)\hat{a}^\dagger\hat{a} \quad (5.12)$$

Now if we redo our analysis for the decay jump operator $|e\rangle\langle f|$ we find

$$e^{-i\widehat{\mathcal{H}}'_\chi t} |e\rangle\langle f| e^{-i\widehat{\mathcal{H}}'_\chi(T-t)} = e^{-i\chi t\hat{a}^\dagger\hat{a}} |e\rangle\langle f| e^{-i\chi(T-t)\hat{a}^\dagger\hat{a}} = |e\rangle\langle f| e^{i\pi\hat{a}^\dagger\hat{a}} \quad (5.13)$$

Now, independent of the time of the decay, we always do unitary $e^{i\pi\hat{a}^\dagger\hat{a}}$ on the cavity mode and this error is corrected by first reading out $|e\rangle$ on the transmon (the final Hadamard or $\pi/2$ -pulse only acts on the g - f manifold, and we cannot end up in $|e\rangle$ unless we were in $|f\rangle$ before, guaranteeing the dispersive coupling was on during the

entire wait time) and then updating the cavity phase by π which can be done in software (by updating phases of future operations by π). The gate construction is shown below in Fig. 5.4

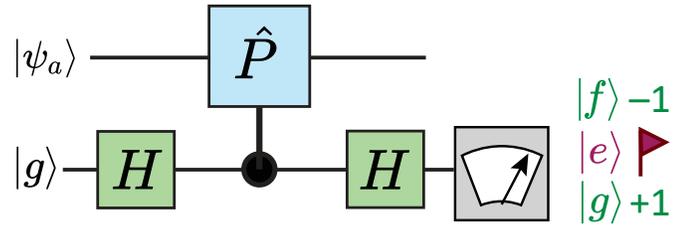


Figure 5.4: Gate construction for an HFT parity measurement. Transmon ancilla is operated in its g - f manifold. Three-state readout of the ancilla allows us to detect and correct for a single transmon decay during the gate. Upon reading out $|e\rangle$, we should update the cavity phase by π

5.1.7 Z measurements

For encodings such as the 0-1 encoding or dual-rail encoding, measuring the photon number parity of either cavity mode that comprises the dual-rail qubit amounts to a measurement of the dual-rail qubit in the Z basis. The same construction can be used for a Z basis in any bosonic code where $Z_L = e^{i\frac{\pi}{n}\hat{a}^\dagger\hat{a}}$ for some integer n . We must adjust the wait time such that $T = \pi/n\chi$ but otherwise the HFT properties carry on over to this Z logical measurement with one exception.

Consider an odd photon number parity 2-fold rotationally symmetric bosonic code, i.e., when

$$|0_L\rangle = \sum_{n=0}^N c_n |4n+1\rangle \quad (5.14)$$

$$|1_L\rangle = \sum_{m=0}^M c_m |4m+3\rangle \quad (5.15)$$

In this case, $e^{i\frac{\pi}{n}\hat{a}^\dagger\hat{a}} = iZ_L$. This factor of i is important! If we don't update the phase properly in the final transmon rotation (e.g., change the Hadamard rotation axis to $\frac{\hat{y}+\hat{z}}{\sqrt{2}}$ instead of $\frac{\hat{x}+\hat{z}}{\sqrt{2}}$) then we will obtain random measurement outcomes! For a code like the Kitten code or 4-legged cat code we do not know when we end up in this scenario if we suffer from a single photon loss. For this reason it is essential that we interleave Z measurements with parity measurements in order to track single photon loss and ensure we can trust the Z measurement outcomes.

If we use the wrong phase for some reason, rather than the entire gate construction realizing the unitary

$$\frac{\mathbf{1} + \mathbf{Z}_L}{2} \otimes \mathbf{1}_q + \frac{\mathbf{1} - \mathbf{Z}_L}{2} \otimes X_q \quad (5.16)$$

acting on initial state $|\psi_a\rangle \otimes |g\rangle$, we instead unknowingly realize the unitary

$$\frac{\mathbb{1} + i\mathbf{Z}_L}{\sqrt{2}} \otimes \mathbb{1}_q + \frac{\mathbb{1} - i\mathbf{Z}_L}{\sqrt{2}} \otimes X_q \quad (5.17)$$

The backaction on the cavity state at first glance looks bad. We were trying to do a Z measurement, but due to an unnoticed single-photon loss event we are now randomly applying the gates $\mathbf{Z}_L(\pm\pi/2)$ without knowing it! This error is actually benign, if we can eventually catch this photon loss error, e.g., with subsequent parity measurements. Then we know not to trust the prior Z measurements, to update the phase and to trust the next set of Z measurements. And because $\mathbf{Z}_L(\pm\pi/2)$ does not change $\langle \mathbf{Z}_L \rangle$, this measurement will still ultimately be QND and we will not dephase the cavity state.

5.2 The Hardware Fault-Tolerant ZZ measurement

We now show how we can extend the error-correction properties of the HFT parity measurement to a QND ZZ measurement on bosonic codewords. This enables CNOT in Fig. 2.3 and is in principle valid for any rotationally symmetric bosonic codeword for which $Z_L = e^{i\frac{\pi}{n}\hat{a}^\dagger\hat{a}}$ for $n = 1, 2, 4, 6, \dots$. A ZZ measurement is very different from a ZZ gate. A ZZ gate on two bosonic is merely two Z gates, i.e., $ZZ = Z_a \otimes Z_b$ and is not entangling. A ZZ measurement is entangling (it can be used to make Bell states from separable states). The ZZ measurement we wish to implement is QND. For initial state $|\psi_{ab}\rangle$ we obtain outcome ± 1 with probability

$$p_{\pm 1} = \left\langle \psi_{ab} \left| \frac{\mathbb{1} \pm ZZ}{2} \right| \psi_{ab} \right\rangle \quad (5.18)$$

and end up in final state

$$|\psi_{ab}^{(\pm 1)}\rangle = \frac{1}{\sqrt{p_{\pm 1}}} \frac{\mathbb{1} \pm ZZ}{2} |\psi_{ab}\rangle \quad (5.19)$$

From the general circuit for QND measurements of binary valued operator, \hat{B} , we can realize a QND ZZ measurement if we can realize the ancilla-controlled unitary cZZ . From our HFT parity measurement, we can make this error-correctable if we operate in the g - f manifold of the transmon. Then at the end of this circuit, measuring the transmon in $|g\rangle$ ($|f\rangle$) gives us outcome $+1$ (-1). Thus in order to realize an error-correctable HFT ZZ measurement, we must realize the unitary

$$\begin{aligned} \mathbf{cZZ}_L &= |g\rangle \langle g| + \mathbf{ZZ}_L (|e\rangle \langle e| + |f\rangle \langle f|) \\ &= |g\rangle \langle g| + e^{i\frac{\pi}{n}(\hat{a}^\dagger \hat{a} + \hat{b}^\dagger \hat{b})} (|e\rangle \langle e| + |f\rangle \langle f|) \end{aligned} \quad (5.20)$$

We now show how such a unitary (or rather an equivalent unitary) is surprisingly easy to construct by combining the parametric beamsplitter with the dispersive interaction simultaneously.

5.2.1 Adding a parametric beamsplitter interaction

When extending our control Hamiltonian, we must be careful to still satisfy error closure. We show that the parametric beamsplitter interaction is particularly adept at this. In this scenario, our bosonic code is designed to correct for single photon loss so $\{\hat{\epsilon}\}_{\text{hardware}} = \{\hat{a}, \hat{b}\}$. The set of correctable errors (via HFT parity measurements + recovery unitaries) after the gate is larger, given by $\{\hat{\epsilon}\}_{\text{corr}} = \{\hat{a}, \hat{b}, \hat{a}\hat{b}, c_0\hat{a} + c_1\hat{b} + c_2\hat{a}\hat{b}\}$

We briefly summarize the conditions for error closure for a given control Hamiltonian, $\hat{\mathcal{H}}_0$, $\{\hat{\epsilon}\}_{\text{hardware}}$ and $\{\hat{\epsilon}\}_{\text{corr}}$

1. Generate a new set of errors, $\{\hat{\epsilon}\}_{\text{ext}}$ by commuting errors from $\{\hat{\epsilon}\}_{\text{hardware}}$ with $\hat{\mathcal{H}}_0$

$$[\hat{\mathcal{H}}_0, \hat{\epsilon}_{\text{hardware}}] = \hat{\epsilon}_{\text{ext}} \quad (5.21)$$

2. Check $\{\hat{\epsilon}\}_{\text{ext}}$ is indeed a closed set of errors by commuting errors from this set with $\hat{\mathcal{H}}_0$ and checking

$$[\hat{\mathcal{H}}_0, \hat{\epsilon}_{\text{ext}}] \in \{\hat{\epsilon}\}_{\text{ext}} \quad (5.22)$$

3. check that $\{\hat{\epsilon}\}_{\text{ext}}$ is a subset of $\{\hat{\epsilon}\}_{\text{corr}}$

$$\{\hat{\epsilon}\}_{\text{ext}} \subseteq \{\hat{\epsilon}\}_{\text{corr}} \quad (5.23)$$

We now walk through this example for photon loss during a parametric beamsplitter. For this purpose, we can set $\hat{\mathcal{H}}_0 = \hat{a}^\dagger \hat{b} + \hat{a} \hat{b}^\dagger$ and these results will hold for any phase or amplitude of g_{BS} .

We take commutators and see we do indeed generate the closed set of errors $\{\hat{\epsilon}\}_{\text{ext}} = \{\hat{a}, \hat{b}\}$

$$[\hat{a}^\dagger \hat{b} + \hat{a} \hat{b}^\dagger, \hat{a}] = -\hat{b}, \quad (5.24)$$

$$[\hat{a}^\dagger \hat{b} + \hat{a} \hat{b}^\dagger, \hat{b}] = -\hat{a}, \quad (5.25)$$

What this tells us is that photon loss errors that happen during the beamsplitter evolution can be corrected after the gate. We can actually check this is indeed the case with the SWAP gate, which is realized with the Hamiltonian

$$\hat{\mathcal{H}}_{\text{BS}}/\hbar = \frac{g_{\text{BS}}}{2} (\hat{a}^\dagger \hat{b} + \hat{a} \hat{b}^\dagger) \quad (5.26)$$

applied for time $T_{\text{SWAP}} = \pi/g_{\text{BS}}$. Suppose we suffer from a photon loss event \hat{a} at intermediate time $t_{\text{jump}} = T_{\text{SWAP}} - t$.

Going through the usual jump-error analysis we find

$$e^{-i\hat{\mathcal{H}}_{\text{BS}}t} \hat{a} e^{i\hat{\mathcal{H}}_{\text{BS}}t} e^{-i\hat{\mathcal{H}}_{\text{BS}}T_{\text{SWAP}}} = \left(\cos\left(\frac{g_{\text{BS}}t}{2}\right) \hat{a} + i \sin\left(\frac{g_{\text{BS}}t}{2}\right) \hat{b} \right) \text{SWAP} \quad (5.27)$$

So photon loss during the SWAP gate is equivalent to applying the SWAP gate perfectly, and then applying a jump error of the form $c_0\hat{a} + c_1\hat{b}$, which is a correctable error! The unknown time of photon loss is now imprinted on the probability we detect cavity \hat{a} or cavity \hat{b} to have a photon loss (the earlier it happens, the more likely we detect photon loss in cavity \hat{a}). When we check for a photon loss error via HFT parity measurements, we collapse this superposition. Given that photon loss happens during the SWAP operation, we are guaranteed to detect a photon loss in either cavity a or b but this is uncertain until we measure. We then correct for the photon loss error the appropriate cavity just as we would for a photon loss error during idling.

Note that not all error-correctable errors satisfy error closure. For instance, for two cavities each using a single-mode bosonic encoding, loss in both cavities $\hat{a}\hat{b}$ is correctable if it happens whilst the qubits are idling, but if it happens during the implementation of the SWAP gate itself, it results in an uncorrectable error. To quickly see why, we can calculate

$$\left[\hat{a}^\dagger\hat{b} + \hat{a}\hat{b}^\dagger, \hat{a}\hat{b}\right] = -\hat{b}^2 - \hat{a}^2 \quad (5.28)$$

In other words, photon loss $\hat{a}\hat{b}$ during the gate can look like double photon loss, e.g., \hat{a}^2 or \hat{b}^2 after the gate! An uncorrectable error! Interestingly, there are surprisingly few cavity-cavity Hamiltonians that describe cavity-cavity interactions that have this error closure property. We show this in Table 5.3 below

What error closure tells us is that for bosonic codes designed to correct for photon loss we should only use Hamiltonians derived from beamsplitter couplings to couple qubits together. For Hamiltonians on a single cavity we can use rotations and displacements but should steer clear of squeezing since this can convert photon loss into a heating event.

We can redo this analysis for another set of $\{\hat{e}\}_{\text{hardware}}$ and $\{\hat{e}\}_{\text{corr}}$. For instance,

Name	$\widehat{\mathcal{H}}_0$	$[\widehat{\mathcal{H}}_0, a]$	$\widehat{\mathcal{H}}_0$ and $\{a, b\}$ satisfy error closure?
Rotation	$a^\dagger a$	a	✓
Displacement	$a + a^\dagger$	1	✓
Beamsplitter	$ab^\dagger + a^\dagger b$	b	✓
Two-mode squeezing	$a^\dagger b^\dagger + ab$	b^\dagger	✗
Squeezing	$a^{\dagger 2} + a^2$	$2a^\dagger$	✗
optomechanical coupling	$a^\dagger a(b + b^\dagger)$	$a(b + b^\dagger)$	✗
	$(a + a^\dagger)b^\dagger b$	$b^\dagger b$	✗
Cross-Kerr	$a^\dagger ab^\dagger b$	$ab^\dagger b$	✗

Table 5.3: **Error closure for candidate Hamiltonians for bosonic codes designed to protect from single photon loss.** For bosonic codes designed to correct against single photon loss, we can evaluate the commutator $[\widehat{\mathcal{H}}_0, \hat{a}]$ to see if the Hamiltonian could satisfy error closure. We find only the lowest order Hamiltonians make good candidates for constructing gates. In the table, if $\widehat{\mathcal{H}}_0$ is a good candidate, so is $\widehat{\mathcal{H}}_0 \otimes \hat{\sigma}_z^{gf}$

in the GKP code small displacement errors are correctable. Suitable control Hamiltonians must convert small displacement errors during the gate to small displacement errors after the gate. The number of control Hamiltonians that satisfy these error-closure conditions is still quite small, although single and two mode squeezing are now allowed (and rotations are no longer allowed!).

5.2.2 Why not just SWAP?

Given that SWAP gates generated via beamsplitter interactions satisfy error closure to photon loss, one can presume it is easy to engineer an HFT \mathbf{cZZ}_L unitary with the circuit in Fig. 5.5. Swaps are used to ensure that the transmon ancilla state is entangled with (and measures) both cavities before readout at the end.

However, we run into a problem. Most of the time, the transmon ancilla is on the equator of the g - f Bloch sphere throughout the measurement. The dispersive interaction is static, we cannot turn it off (easily). When we activate the parametric

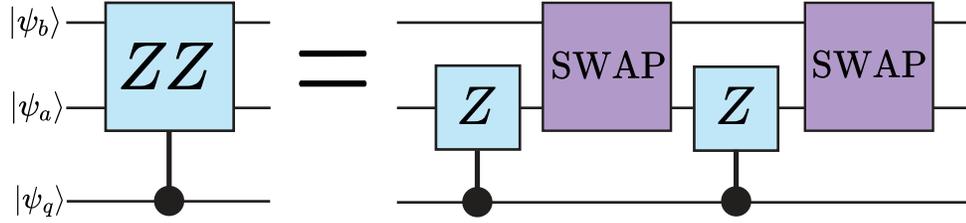


Figure 5.5: Realizing \mathbf{cZZ}_L would be easy if you had instantaneously fast SWAPs! In practice, we would need to realize a SWAP that is independent of the transmon ancilla state.

beamsplitter to realize the SWAP gate, if we are in the ground state, our beamsplitter will be on resonance and we will apply the Hamiltonian

$$\hat{\mathcal{H}}_{\text{BS}}/\hbar = \frac{g_{\text{BS}}}{2} (\hat{a}^\dagger \hat{b} + \hat{a} \hat{b}^\dagger), \quad (5.29)$$

but if we are in the $|f\rangle$ state it is as though our beamsplitter is detuned from resonance by χ_{gf} ! Since typically $g_{\text{BS}} \sim \chi$ and it is hard to engineer $g_{\text{BS}} \gg \chi$ we find ourselves in trouble. What we would need is an *unconditional* beamsplitter, one that is on-resonance regardless of the state of the transmon ancilla.

5.2.3 Evolution under a detuned beamsplitter interaction

The Eureka moment is that the same ‘problem’ we are trying to fix — a static dispersive interaction that cannot be turned off during a parametric beamsplitter interaction, can also be used to realize a \mathbf{cZZ}_L unitary! What’s even better, is that we will see this realization is HFT and is in fact even shorter than our original proposed gate construction with the SWAPs in Fig. 5.5! To show how to engineer \mathbf{cZZ}_L , it is helpful to first consider general evolution under a detuned beamsplitter Hamiltonian given by

$$\hat{\mathcal{H}}_{\Delta}/\hbar = \frac{g_{\text{BS}}}{2} (\hat{a}^\dagger \hat{b} + \hat{a} \hat{b}^\dagger) + \Delta \hat{a}^\dagger \hat{a} \quad (5.30)$$

Where we have chosen the convention where cavity mode \hat{a} is the mode that is detuned from resonance. We examine the dynamics of this Hamiltonian in the Heisenberg picture, and see how the cavity mode operators transform according to

$$\hat{a}(t) = U_{\Delta}^{\dagger} \hat{a} U_{\Delta} \quad (5.31)$$

$$\hat{b}(t) = U_{\Delta}^{\dagger} \hat{b} U_{\Delta} \quad (5.32)$$

where $U_{\Delta} = e^{-i\hat{H}_{\Delta}t/\hbar}$ from which we find the rather long-winded analytical expressions

$$\hat{a}(t) = e^{-\frac{i\Delta t}{2}} \left(\left(\cos\left(\frac{\Omega t}{2}\right) - \frac{i\Delta}{\Omega} \sin\left(\frac{\Omega t}{2}\right) \right) \hat{a} - \frac{ig_{\text{bs}}}{\Omega} \sin\left(\frac{\Omega t}{2}\right) \hat{b} \right) \quad (5.33)$$

$$\hat{b}(t) = e^{-\frac{i\Delta t}{2}} \left(\left(\cos\left(\frac{\Omega t}{2}\right) + \frac{i\Delta}{\Omega} \sin\left(\frac{\Omega t}{2}\right) \right) \hat{b} - \frac{ig_{\text{bs}}}{\Omega} \sin\left(\frac{\Omega t}{2}\right) \hat{a} \right) \quad (5.34)$$

where

$$\Omega = \sqrt{g_{\text{bs}}^2 + \Delta^2}$$

At time $t = \frac{2\pi}{\Omega}$, we find the cavity operators first 'return' to their original value with an additional phase shift:

$$\hat{a}\left(t = \frac{2\pi}{\Omega}\right) = -e^{-\frac{i\Delta\pi}{\Omega}} \hat{a} \quad (5.35)$$

$$\hat{b}\left(t = \frac{2\pi}{\Omega}\right) = -e^{-\frac{i\Delta\pi}{\Omega}} \hat{b} \quad (5.36)$$

This is not a global phase shift! It corresponds to a measurable rotation of both cavity modes corresponding to the unitary

$$U_{\Delta}\left(\frac{2\pi}{\Omega}\right) = e^{i\phi(a^{\dagger}a + b^{\dagger}b)}, \quad (5.37)$$

where $\phi = \pi(1 - \frac{\Delta}{\Omega})$

Now consider a control Hamiltonian consisting a parametric beamsplitter and a

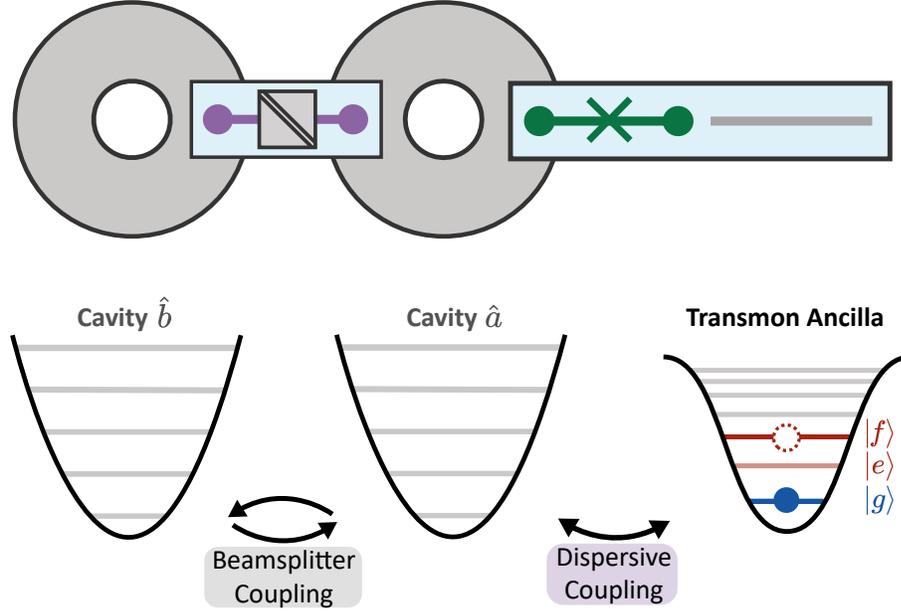


Figure 5.6: cQED hardware and interactions needed to realize an HFT \mathbf{cZZ}_L unitary. Only a parametric beamsplitter coupling between the two cavities and a dispersively coupled transmon ancilla is required. The dispersively coupling only needs to be to one of the cavities. The $|e\rangle$ level of the transmon is reserved only for detecting/correcting first-order transmon decay

χ -matched dispersive interaction between cavity \hat{a} and a control transmon.

$$\hat{\mathcal{H}}/\hbar = \frac{g_{\text{BS}}}{2} (\hat{a}^\dagger \hat{b} + \hat{a} \hat{b}^\dagger) + \frac{\chi}{2} \hat{a}^\dagger \hat{a} (|g\rangle \langle g| - |e\rangle \langle e| - |f\rangle \langle f|) \quad (5.38)$$

We engineer this Hamiltonian if we detune our beamsplitter pumps to satisfy $|\omega_1 - \omega_2| = |\omega_a + \chi/2 - \omega_b|$. The cQED Hardware needed for this control Hamiltonian is shown in Fig. 5.6

If the control transmon is in $|g\rangle$, it is as if $\Delta = +\chi/2$ in the detuned beamsplitter dynamics. If the transmon is in $|f\rangle$ or $|e\rangle$, it is as though $\Delta = -\chi/2$. (And to make things a bit more confusing χ is always negative!) This means, after applying the Hamiltonian for time $t = 2\pi/\Omega$, we have implemented the unitary

$$\hat{U} = e^{i\phi + (a^\dagger a + b^\dagger b)} |g\rangle \langle g| + e^{i\phi - (a^\dagger a + b^\dagger b)} (|e\rangle \langle e| + |f\rangle \langle f|) \quad (5.39)$$

Where $\phi_{\pm} = \pi(1 \mp \frac{\chi}{2\Omega})$. This is almost the unitary we are after! We can re-factor this unitary as

$$\hat{U} = e^{i\phi_+(a^\dagger a + b^\dagger b)} \left(|g\rangle \langle g| + e^{i\frac{\pi\chi}{\Omega}(a^\dagger a + b^\dagger b)} (|e\rangle \langle e| + |f\rangle \langle f|) \right) \quad (5.40)$$

Which can be decomposed into a control unitary followed by an unconditional rotation of both cavities by angle ϕ_+ . This unconditional rotation is not simply a global phase but can be easily compensated by updating the phase on subsequent cavity pulses.

Now we just need to find the amplitude g_{BS} for our parametric beamsplitter interaction such that the control unitary is \mathbf{cZZ}_L for our n -fold symmetric bosonic code. For this we require

$$\frac{\pi}{n} = \frac{\pi\chi}{\Omega} \quad (5.41)$$

Which is solved for

$$g_{\text{BS}} = \sqrt{n^2 - \frac{1}{4}} |\chi| \quad (5.42)$$

and applying the Hamiltonian for time

$$T = \frac{2\pi}{\Omega} = \frac{2\pi}{n\chi} \quad (5.43)$$

Equations 5.42 and 5.43 give the conditions for the fastest possible \mathbf{cZZ}_L unitaries under this Hamiltonian. However, there are other ‘operating points’ that also realize \mathbf{cZZ}_L which may be more experimentally convenient, e.g., if we cannot reach a sufficient g_{BS} amplitude to satisfy Eq. 5.42, what should we do? One compromise is to go around the loop ‘ m ’ number of times to consecutively build up the required phase shift, acquiring phase $\frac{\pi}{nm}$ on each loop. This means we can engineer \mathbf{cZZ}_L no matter how slow g_{BS} is, at the cost of an increase in total gate time.

Another degree of freedom we can use to find more valid operating points is to notice the required total phase shift is not π/n , but instead $(2k - 1)\pi/n$ for positive

integer $k = 1, 2, \dots, n$ (we restrict $k \leq n$ to avoid repeated values). We are taking advantage of the fact $e^{i\frac{2k\pi}{n}\hat{a}^\dagger\hat{a}} = \mathbb{1}_L$ on the bosonic codewords.

This gives us the more general set of conditions to realize \mathbf{cZZ}_L unitaries parameterized by positive integers (n, m, k) where $k \leq n$, n is the n -fold rotational symmetry of the bosonic codeword, m is the number of orbits we complete and k is our choice of phase difference. We also introduce the parameter $r = \Omega/|\chi|$ for notational convenience.

$$r = \frac{nm}{2km - 1} \quad (5.44)$$

$$T = \frac{2\pi m}{r\chi} = \frac{2\pi}{\chi} \frac{2km - 1}{n} \quad (5.45)$$

$$g_{\text{BS}} = \sqrt{r^2 - \frac{1}{4}} |\chi| \quad (5.46)$$

As well as implementing \mathbf{cZZ}_L , we also implement an additional ‘unconditional’ rotation on both cavities which must be tracked ‘in software’, given by

$$\hat{U}_u = e^{i\phi_u(a^\dagger a + b^\dagger b)} \quad (5.47)$$

Where

$$\phi_u = m\pi \left(1 - \frac{1}{2r}\right) \quad (5.48)$$

We give examples \mathbf{cZZ}_L operating points in Table 5.4. The naming of different \mathbf{cZZ}_L unitaries relates to their physical control operators. For instance, the parity operator is $\hat{P} = e^{i\pi\hat{a}^\dagger\hat{a}}$, the joint parity operator is $\widehat{JP} = e^{i\pi(a^\dagger a + b^\dagger b)}$ and the control-joint parity operator (in the g - e manifold) would be $\mathbf{cJP} = |g\rangle\langle g| + \widehat{JP}|e\rangle\langle e|$. The convention continues for the ‘4-parity’ operator, given by $\widehat{4P} = e^{i\frac{\pi}{2}\hat{a}^\dagger\hat{a}}$.

In practice, realizing \mathbf{cZZ}_L seems remarkably easy! All we need to do is tune the amplitude of g_{BS} to one of the valid operating points with detuning $\chi/2$ from the usual resonance condition. After applying a constant value of g_{BS} for the appropriate

duration, we will performed the \mathbf{cZZ}_L unitary. In Sec. 5.2.5, we will show this unitary satisfies error closure and is therefore HFT.

\mathbf{cZZ}_L name	(n, m, k)	r	g_{BS}	Duration	ϕ_u
Control joint-parity	(1, 1, 1)	1	$\frac{\sqrt{3}}{2} \chi $	$\frac{2\pi}{\chi}$	$\frac{\pi}{2}$
Control joint 4-parity (Fast)	(2, 1, 1)	2	$\frac{\sqrt{15}}{2} \chi $	$\frac{\pi}{\chi}$	$\frac{3\pi}{4}$
Control joint 4-parity (Slow)	(2, 1, 2)	$\frac{2}{3}$	$\frac{\sqrt{7}}{6} \chi $	$\frac{3\pi}{\chi}$	$\frac{\pi}{4}$
Control-joint parity (2 loops)	(1, 2, 1)	$\frac{2}{3}$	$\frac{\sqrt{7}}{6} \chi $	$\frac{6\pi}{\chi}$	$\frac{3\pi}{2}$
General	(n, m, k)	$\frac{nm}{2km-1}$	$\sqrt{r^2 - \frac{1}{4}} \chi $	$\frac{2\pi m}{r\chi}$	$m\pi \left(1 - \frac{1}{2r}\right)$

Table 5.4: Operating points for realizing HFT \mathbf{cZZ}_L for an n -fold symmetric bosonic encoding.

So far our analysis of the dynamics has been analytical evolution in the Heisenberg picture. However, it is possible to map these dynamics to a new kind of ‘Bloch sphere’ which allows us to use our intuition from the qubit Bloch sphere to explore all possible dynamics generated from $\widehat{\mathcal{H}}_c$ and to design new control unitaries.

5.2.4 The Operator Bloch Sphere

The dynamics of the detuned beamsplitter are reminiscent of a detuned Rabi drive on a two-level qubit. In fact they are completely analogous, and we can derive a ‘Bloch Sphere’ for the cavity mode operators.

First we rewrite our detuned beamsplitter Hamiltonian as

$$\widehat{\mathcal{H}}_{\text{BS}}/\hbar = \frac{g_{\text{BS}}(t)}{2} \left(e^{i\varphi(t)} \hat{a}^\dagger \hat{b} + e^{-i\varphi(t)} \hat{a} \hat{b}^\dagger \right) + \Delta(t) \hat{a}^\dagger \hat{a}, \quad (5.49)$$

Inspired by Schwinger's angular momentum formalism of bosonic operators [Schwinger, 2001], we rewrite $\widehat{\mathcal{H}}_{\text{BS}}$ with the 'angular momentum operators' $\widehat{L}_I = \frac{1}{2}(\hat{a}^\dagger \hat{a} + \hat{b}^\dagger \hat{b})$, $\widehat{L}_X = \frac{1}{2}(\hat{a}^\dagger \hat{b} + \hat{a} \hat{b}^\dagger)$, $\widehat{L}_Y = \frac{1}{2i}(\hat{a}^\dagger \hat{b} - \hat{a} \hat{b}^\dagger)$ and $\widehat{L}_Z = \frac{1}{2}(\hat{a}^\dagger \hat{a} - \hat{b}^\dagger \hat{b})$, which allows us to rewrite $\widehat{\mathcal{H}}_{\text{BS}}$ as the parameterized Hamiltonian

$$\widehat{\mathcal{H}}_{\text{BS}}(g_{\text{BS}}, \varphi, \Delta)/\hbar = g_{\text{BS}} \left(\cos \varphi \widehat{L}_X - \sin \varphi \widehat{L}_Y \right) + \Delta \left(\widehat{L}_Z + \widehat{L}_I \right).$$

For now we consider static $(g_{\text{BS}}, \varphi, \Delta)$ but this Bloch Sphere picture holds even when these vary in time. We call these angular momentum operators since they obey the commutation relations $[\widehat{L}_X, \widehat{L}_Y] = i\widehat{L}_Z$ up to cyclic permutations of X, Y, Z . This is the same commutation relation obeyed by the Pauli operators, i.e., $[\widehat{X}, \widehat{Y}] = i\widehat{Z}$. To make the analogy with a qubit Bloch sphere more explicit, a detuned constant Rabi drive on a qubit has the Hamiltonian

$$\widehat{\mathcal{H}}_{\text{Rabi}} = g_{\text{Rabi}} \left(\cos \varphi \widehat{X} - \sin \varphi \widehat{Y} \right) + \Delta \widehat{Z}. \quad (5.50)$$

Evolution under the detuned beamsplitter Hamiltonian corresponds to rotation about a fixed axis of a Bloch sphere. To see how this arises we transform the mode operators via the unitary operator $\widehat{U} = \exp(-i\widehat{\mathcal{H}}_{\text{BS}}t/\hbar)$,

$$\begin{pmatrix} \hat{a}(t) \\ \hat{b}(t) \end{pmatrix} = \begin{pmatrix} \widehat{U} \hat{a} \widehat{U}^\dagger \\ \widehat{U} \hat{b} \widehat{U}^\dagger \end{pmatrix} = e^{-i\frac{\Delta}{2}t} R_{\vec{n}}(\Omega t) \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix}, \quad (5.51)$$

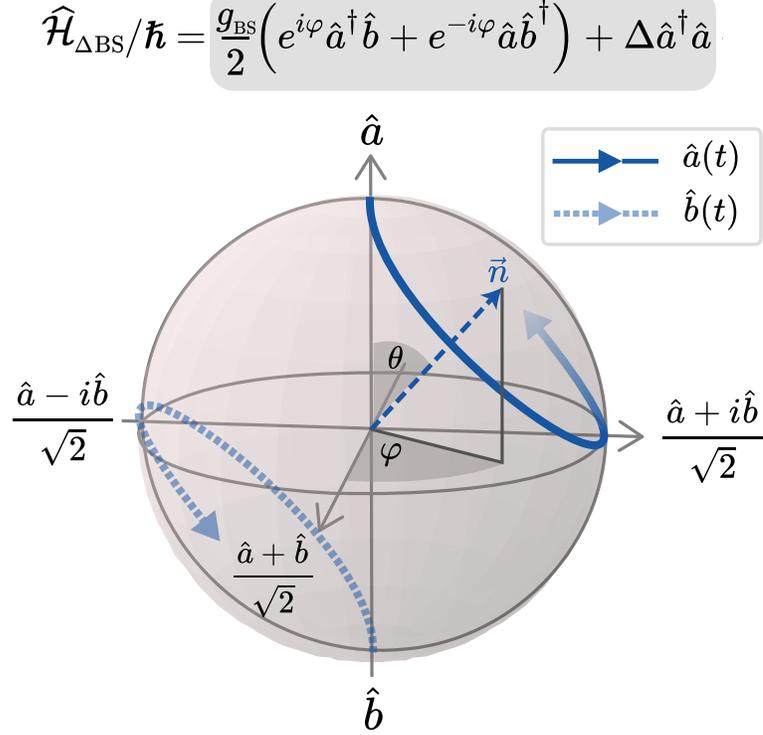


Figure 5.7: Applying the beamsplitter interaction, causes the bosonic operators to evolve in time in the Heisenberg picture. The new modes are linear combinations of the original mode operators \hat{a} and \hat{b} and this time evolution can be represented geometrically by the trajectories shown in blue. The evolution of $\hat{b}(t)$ is always the ‘mirror image’ of $\hat{a}(t)$. Trajectories orbit the precession vector \vec{n} , which is fully determined by Hamiltonian parameters: g_{BS} and Δ set the polar angle, θ , whereas φ sets the azimuthal angle.

where $R_{\vec{n}}(\Omega t) = (\cos \frac{\Omega t}{2} I - i \sin \frac{\Omega t}{2} \vec{n} \cdot \vec{\sigma})$ is a *symplectic* matrix in $\text{SU}(2)$, which can be interpreted as a rotation around a precession vector $\vec{n} = [\sin \theta \cos \varphi, -\sin \theta \sin \varphi, \cos \theta]$ at rate Ω . The polar angle of the precession vector is determined by the ratio of the coupling strength g_{BS} to Ω such that $\cos \theta = \Delta/\Omega$ and $\sin \theta = g_{\text{BS}}/\Omega$. Analogous to state evolution on a qubit Bloch sphere, we plot the mode transformations at each point in time to form trajectories on the Operator Bloch Sphere (OBS) as shown in Fig. 5.7.

Here, the north pole represents the initial mode operator \hat{a} and the solid arrow represents the trajectory of the transformed mode operator $\hat{a}(t)$. Similarly, the south pole represents the initial mode operator \hat{b} and the dashed arrow represents the trajec-

$$\hat{\mathcal{H}}_{\chi \text{ BS}}/\hbar = \frac{g_{\text{BS}}}{2} \left(e^{i\varphi} \hat{a}^\dagger \hat{b} + e^{-i\varphi} \hat{a} \hat{b}^\dagger \right) + \Delta \hat{a}^\dagger \hat{a} - \frac{\chi}{2} \hat{a}^\dagger \hat{a} (|g\rangle\langle g| - |e\rangle\langle e| - |f\rangle\langle f|)$$

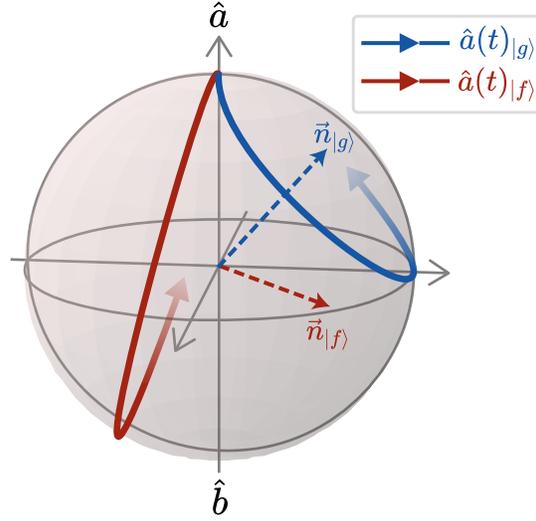


Figure 5.8: When an ancilla is dispersively coupled to one of the bosonic modes, the precession vector \vec{n} becomes dependent on the ancilla state, yielding distinct vectors \vec{n}_g and \vec{n}_f . Furthermore, if the ancilla is in a superposition of $|g\rangle$ and $|f\rangle$, the bosonic modes will be in a superposition of the two trajectories shown in the figure. The $\hat{b}(t)$ trajectories are antipodal to the $\hat{a}(t)$ trajectories and are not shown on this Bloch sphere.

tory of the transformed mode operator $\hat{b}(t)$. The trajectory can be fully controlled by modulating the complex amplitude of the beamsplitter interaction. The trajectories for $\hat{a}(t)$ and $\hat{b}(t)$ are antipodal to one another, a consequence of the unitary nature of the mode transformations. For simplicity we show only the trajectory of $\hat{a}(t)$ moving forwards. The end points of the trajectories indicate the final mode transformations of the original \hat{a}, \hat{b} operators.

When our detuning Δ becomes ancilla state dependent, so do our precession vectors (their z components will differ) and trajectories on the OBS. This is shown in Fig. 5.8 The geometric phase enclosed (the solid angle) by a trajectory has a physical meaning. On a qubit Bloch sphere enclosing solid angle ϕ results in a global phase shift, $e^{i\phi}$ on the qubit state which is typically ignored. However, enclosing phase ϕ on the OBS corresponds to the unitary $\hat{R}_\phi = e^{i\frac{\phi}{2}(\hat{a}^\dagger \hat{a} + \hat{b}^\dagger \hat{b})}$, generated by the \hat{L}_I term

in our Hamiltonian. The direction (clockwise or anticlockwise) in which we traverse a loop on the OBS dictates the sign of ϕ . With this knowledge we can now make sense of the \mathbf{cZZ}_L trajectories plotted on the OBS in Fig. 5.9 Having found how to implement the \mathbf{cZZ}_L unitary, we can either use it as in Fig. 2.6 to perform a QND ZZ measurement, or in Eq. 2.54 to realize the $ZZ(\theta)$ entangling gate. The rest of this chapter will be mostly spent analyzing the Hardware Fault-tolerance properties of both approaches.

5.2.5 A QND ZZ measurement and its Hardware Fault-Tolerance

The following gate construction in Fig. 5.10 performs a QND ZZ measurement. What we are interested in is checking whether this measurement construction is HFT. We first check that our general control Hamiltonian, $\widehat{\mathcal{H}}_c$ satisfies error closure.

$$\widehat{\mathcal{H}}_c = \frac{g_{\text{BS}}}{2} (\hat{a}^\dagger \hat{b} + \hat{a} \hat{b}^\dagger) + \frac{\chi}{2} \hat{a}^\dagger \hat{a} (|g\rangle \langle g| - |e\rangle \langle e| - |f\rangle \langle f|) \quad (5.52)$$

The set of hardware errors is $\{\hat{\epsilon}\}_{\text{hardware}} = \{\hat{a}, \hat{b}, |e\rangle \langle f|, |g\rangle \langle g| - |f\rangle \langle f|\}$. For convenience we define $\hat{\sigma}_z^{gf} = |g\rangle \langle g| - |f\rangle \langle f|$. The set of correctable errors is $\{\hat{\epsilon}\}_{\text{corr}} = \{\hat{a}, \hat{b}, \hat{\sigma}_z^{gf}, \hat{a} \hat{\sigma}_z^{gf}, \hat{b} \hat{\sigma}_z^{gf}, \hat{a} \hat{b}, \hat{a} \hat{b} \hat{\sigma}_z^{gf}, |e\rangle \langle f|, \hat{a} |e\rangle \langle f|, \hat{b} |e\rangle \langle f|, \hat{a} \hat{b} |e\rangle \langle f|\}$. By taking the commutation relations we start finding elements of $\{\hat{\epsilon}\}_{\text{ext}}$.

$$\left[\widehat{\mathcal{H}}_c, \hat{a} \right] = -\frac{g}{2} \hat{b} - \frac{\chi}{2} \hat{a} \hat{\sigma}_z^{gf}, \quad (5.53)$$

$$\left[\widehat{\mathcal{H}}_c, \hat{b} \right] = -\frac{g}{2} \hat{a}, \quad (5.54)$$

$$\left[\widehat{\mathcal{H}}_c, \hat{\sigma}_z^{gf} \right] = 0 \quad (5.55)$$

$$\left[\widehat{\mathcal{H}}_c, |e\rangle \langle f| \right] = 0 \quad (5.56)$$

So far, $\{\hat{\epsilon}\}_{\text{ext}}$ consists of our original hardware errors and one additional error type, $\hat{a} \hat{\sigma}_z^{gf}$ (ancilla dephasing due to cavity photon loss). To find all elements in $\{\hat{\epsilon}\}_{\text{ext}}$, we

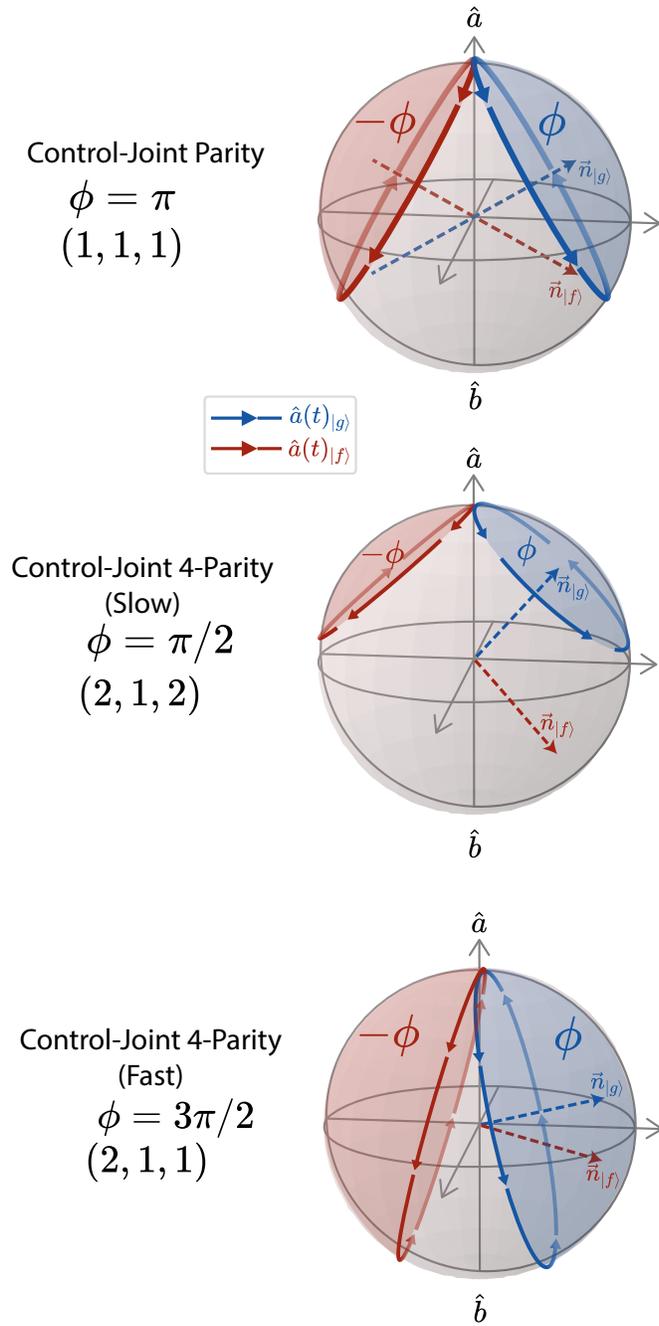


Figure 5.9: Operator Bloch Sphere trajectories for cavity mode operator $\hat{a}(t)$ for various \mathbf{cZZ}_L operating points. We show trajectories conditioned on the ancilla being in the $|g\rangle$ and $|f\rangle$ state. The $|e\rangle$ state trajectory is omitted. If $\chi_{ge} = \chi_{gf}$, $\hat{a}(t)|_e = \hat{a}(t)|_f$

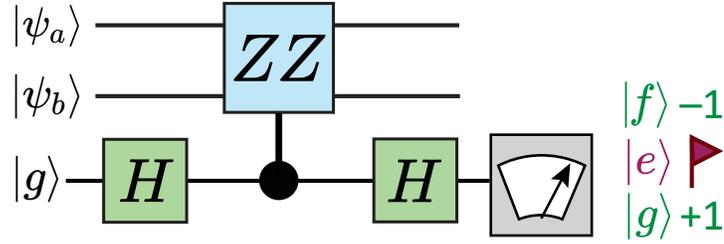


Figure 5.10: QND measurement of the ZZ operator via a transmon ancilla. We use the g - f manifold of the ancilla. The two Hadamard gates (which may be replaced with $X_{\frac{\pi}{2}}$ or $Y_{\frac{\pi}{2}}$ gates) act only on the g - f manifold and do not affect the $|e\rangle$ level.

need to calculate the next order of commutators, $[\hat{\mathcal{H}}_c, [\hat{\mathcal{H}}_c, \hat{\epsilon}]]$ which adds one more element to $\{\hat{\epsilon}\}_{\text{ext}} = \{\hat{\epsilon}\}_{\text{hardware}} \cup \{\hat{a}\hat{\sigma}_z^{gf}, \hat{b}\hat{\sigma}_z^{gf}\}$. Recalculating the commutators at higher orders does not populate $\{\hat{\epsilon}\}_{\text{ext}}$ with any more error-types and since $\{\hat{\epsilon}\}_{\text{ext}} \subset \{\hat{\epsilon}\}_{\text{corr}}$ we satisfy error closure.

Satisfying the error-closure condition tells us mathematically, everything should be sound, but what are the physical consequences of each type of hardware error occurring during the ZZ measurement?

Transmon Dephasing

The effects of transmon ancilla errors are almost identical to that of the HFT parity measurement. Transmon dephasing, $\hat{\sigma}_z^{gf}$ is error transparent, and so only inverts the measurement outcomes but is otherwise completely QND on both the bosonic qubits. The ‘bad’ measurement outcome can only be corrected by repeating the measurement at least three times and majority voting. Dephasing during the short transmon ancilla rotations is unlikely but still possible, and results in a measurement error but is otherwise QND.

Transmon Decay

With χ -matching, Transmon decay, $|e\rangle\langle f|$ is error transparent to first-order. When we detect the transmon ancilla in $|e\rangle$ at the end of the measurement, we know we

will have performed the unitary $\hat{R}_\phi = e^{i\frac{\phi}{2}(\hat{a}^\dagger\hat{a}+\hat{b}^\dagger\hat{b})}$ on the cavities (depending on our operating point) and need to attempt the measurement again. Transmon decay that happens during the transmon rotations is also correctable.

Photon loss

As with an HFT parity measurement, photon loss during the unitary dephases the transmon ancilla, resulting in a random measurement outcome. To correct for this error, we must employ additional HFT parity measurements after the ZZ measurement and then any recovery unitaries.

Another subtlety present in the ZZ measurement is that undetected single photon loss *prior* to a ZZ measurement results in random measurement outcomes for the same reason discussed for the Z measurements in Sec.5.1.7.

For this purpose it is helpful to define two ‘modes of operation’ for a ZZ measurement as in Fig 5.11. The first is for when we believe both bosonic qubits have the same parity (both even or both odd) and the second for when we believe the parities even,odd or odd,even. If we are unable to properly track single photon loss then we may end up using the wrong mode of operation, and the ZZ measurements will give us 50-50 random outcomes. In each case, the physical unitary we perform is the same, but whether it is \mathbf{cZZ}_L or \mathbf{ciZZ}_L depends on the joint parity of the cavities.

5.2.6 How to error-correct a ZZ measurement (is it worth it?)

So far we have arrived at a scheme that satisfies HFT for a QND ZZ measurement for bosonic qubits. HFT in this case states that the measurement is error-correctable, i.e., it *can* be error-corrected but *how* would one actually perform error correction in practice?

We can show how by presuming we are in the 4-legged cat code, where detecting single photon loss is equivalent to correcting it through an update of the logical basis

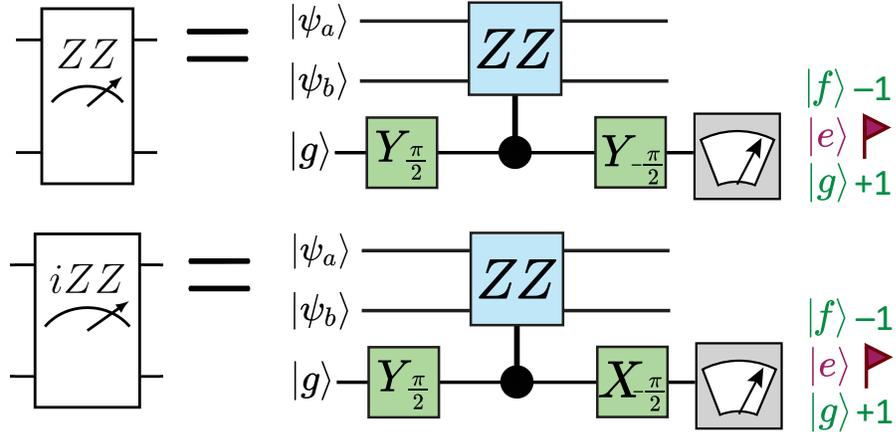


Figure 5.11: Two modes of operation for ZZ measurements. By default, we use the gate circuit shown in the top row. If we suspect a single photon loss has occurred in one of the modes, we use the construction in the bottom row. Physically, these gate constructions are identical except the phase of the last $\pi/2$ pulse has been changed by 90° . We have also replaced the Hadamard gates with $\pi/2$ transmon rotations to match what is typically done in experiments. The iZZ measurement will measure the iZZ logical operator if the joint parity is even, and the ZZ logical operator if the joint parity is odd.

(and for now we ignore the effects of no-jump backaction). How would one perform a QND ZZ measurement? Such a measurement must tolerate one of the errors from $\{\hat{\epsilon}\}_{\text{hardware}}$ occurring at any time during the whole measurement procedure. This means we must **a)** still obtain the correct ZZ measurement **b)** be QND on the cavity states.

A single error-corrected QND ZZ measurement will on average require around six different measurements of the cavities. The large number of measurements arise because we have a ‘who watches the watchman?’ conundrum. Each measurement we introduce to check the other measurements for faults is itself faulty! Fortunately, we can devise strings of measurements as shown in Fig. 5.12 that are finite in length that all check each other for (first-order) errors sufficiently, but these measurement chains can get rather long and incur substantial overhead (we still correct the first-order errors in $\{\hat{\epsilon}\}_{\text{hardware}}$ but the probability of second-order errors increases).

The rationale behind these measurement strings is the following: the outcome of

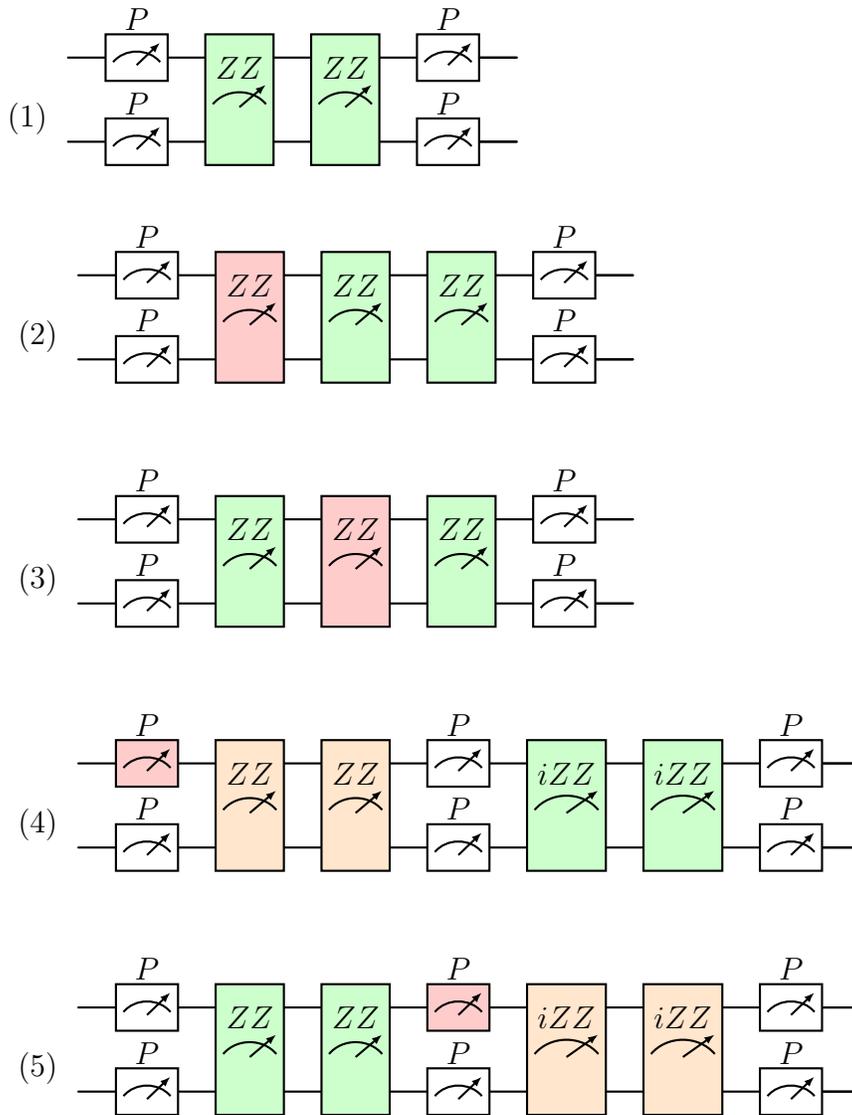


Figure 5.12: Circuits (1-5) to perform an error-corrected QND ZZ measurement. Such an operation has the same ‘entangling power’ as a CNOT gate. Key: Red: readout misassignment error due to readout error, transmon dephasing or cavity photon loss. Orange — wrong type of ZZ measurement performed because of a previous readout error. White: error-free measurement. Green: ZZ outcome info used to sway majority vote in the correct direction.

the first two parity measurements tells us the type of ZZ measurement to perform (ZZ or iZZ). If the wrong type of ZZ measurement is selected, e.g., due to an error in the first set of parity measurements, then we obtain random outcomes. Book-ending the string of ZZ measurements with parity measurements ensures no photon loss occurred during the ZZ measurements, which would scramble the measurement outcomes. Measurement sequences are adaptive. Most of the time, only two ZZ measurements are needed to obtain a majority vote, but sometimes three are required if there is a measurement error. A set of ZZ measurements must be ‘validated’ by the two sets of book-ended parity measurements agreeing. If not, we perform the other type of ZZ measurement. We can see we are accumulating a lot of measurements!

In words, we describe the algorithm for an error-corrected ZZ measurement:

1. For any measurement, if $|e\rangle$ is detected, repeat the measurement until $|g\rangle$ or $|f\rangle$ is obtained
2. Perform two HFT parity measurements on cavities \hat{a} and \hat{b} . If (even,even) or (odd,odd) outcomes are obtained, proceed with ZZ measurements, if (even,odd) or (odd,even) are obtained, proceed with iZZ measurements.
3. Perform repeated HFT ZZ or HFT iZZ measurements. If the first two outcomes agree (row 1 Fig. 5.12) of , proceed to the next step, otherwise, repeat once more for a total of three measurement outcomes (rows 2 and 3 of Fig. 5.12).
4. Perform HFT parity measurements on the cavities again. If the parity measurements agree with the first set, we have finished. Majority vote on the string of ZZ measurements, and correct any photon loss errors in the cavity if needed (e.g., if we detected (odd,even) and we are using the kitten code, correct photon loss in cavity \hat{a}). If however the second set of parity measurements disagree, we must continue measuring ZZ info.

5. Perform two more ZZ measurements. If we performed ZZ in the first set, now we perform iZZ and vice versa. If there is disagreement in the first two results, repeat up to three times.
6. Perform a third and final set of parity measurements on both cavities. If this set agrees with the second set of parity measurements, use the last set of ZZ measurements to majority vote (row 4 of Fig. 5.12). If the third set agrees with the first set of parity measurements, use the first set of ZZ measurements for the majority vote (row 5 of Fig. 5.12).

In undertaking the accounting for what is needed for an error-corrected ZZ measurement (which could then be used in an error-corrected CNOT gate), we have seen various overheads appear at the hardware level that are not just an increase in the number of physical qubit modes. Remember if without error correction our error-rate per operation is p^2 , and with error correction our error rate is Ap^2 because we are only sensitive to second order hardware errors, then A is the number that quantifies the overhead. These overheads are in an increase in the average photon number per qubit, the number of measurements per error-corrected logical measurements and the increased time taken to perform error-corrected logical measurement sequences. These overheads increase our susceptibility to second order hardware errors, to the point where our error correction does not give us any tangible benefit compared to, e.g., the 0-1 encoding without any HFT or encoding the qubits in the transmons instead.

This is the main motivation behind the error-detection approach. Rather than correcting errors at the hardware level, we will detect them. This greatly reduces the difficulty in physical implementation and reduces the hardware overhead. The task of correcting the hardware errors is off-loaded to the next layer of error-correction in the form of erasure errors.

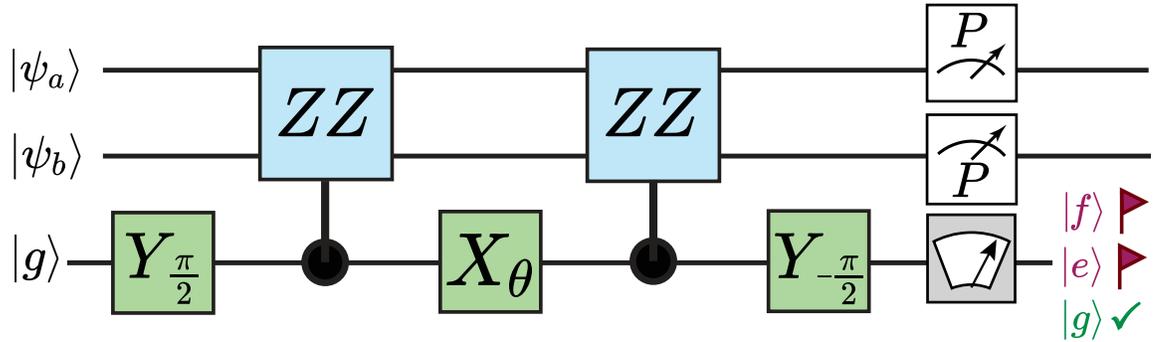


Figure 5.13: Gate construction for realizing an error-detected $ZZ(\theta)$ gate for 2-fold rotationally symmetric bosonic qubits. We detect first-order ancilla decay if we read-out $|e\rangle$, first-order ancilla dephasing if we readout $|f\rangle$ and single photon loss errors from the two QND HFT parity measurements performed after the gate. Once errors are detected, the logical qubit information cannot be recovered. The gate angle θ is parameterized by the intermediate ancilla rotation angle about the X axis.

5.3 The Error-Detected $ZZ(\theta)$ Gate

We now move on to the main result of this chapter, a scheme to realize a Hardware fault-tolerant, error-detected $ZZ(\theta)$ gate for rotationally-symmetric bosonic code-words.

The first thing we note when we set off down this path of error-detection is that χ -matching is no longer needed. Removing χ -matching means we are no longer error-transparent to $|e\rangle\langle f|$ but we can still detect this error. Upon detecting $|e\rangle$ we will know we will have performed some unknown detuned beamsplitter unitary that in general takes us out of the codespace but preserves the total number of excitations in the system. As such, it is safest to say we have detected a leakage error and must completely reinitialize the affected qubits.

With the relaxed requirements of error-detection we can use the exponentiation gadget with the \mathbf{cZZ}_L unitary to realize the error-detectable $ZZ(\theta)$ gate with the gate construction shown in Fig. 5.13. Unlike the error-corrected ZZ measurement, which required multiple repeated measurements, the total sequence of gates here is much shorter. What we are not showing are the steps needed to reset the qubits back

into the codespace which would be needed if this error-detected gate were used in an erasure qubit. This reset would be hardware and encoding specific.

At $\theta = \pi/2$ this gate is locally equivalent to a CNOT. With very little hardware modification, we have transferred from an error-correctable ZZ measurement that looked rather tedious to actually error-correct to an error-detected $ZZ(\pi/2)$ gate which looks much simpler to experimentally realize!

5.3.1 Why is the exponentiation gadget only error-detectable?

We show briefly why such a gate construction is fundamentally only error-detectable and not correctable. For this we switch back to the more general gate sequence and replace \mathbf{cZZ}_L with the more general control unitary \widehat{U}_c and the sequence realizes the gate $\widehat{O}(\theta)$.

The loss of error-correctability can be traced to the two \widehat{U}_c unitaries either side of the $X(\theta)$ transmon rotation. Errors from $\{\hat{\epsilon}\}_{\text{hardware}}$ that happen anywhere in the gate sequence will always be flagged at the end through three-state transmon readout and/or HFT parity measurements on the cavities. However, the unitary that was performed is now dependent on whether the jump error happened during the first \widehat{U}_c or second \widehat{U}_c , information known by the environment but lost to us. We can illustrate this with $\hat{\sigma}_z^{gf}$, a transmon dephasing jump. The gate construction can be written out as the sequence of unitaries

$$e^{-i\frac{\pi}{4}\hat{Y}_{\text{gf}}}\widehat{U}_c e^{-i\frac{\theta}{2}\hat{X}_{\text{gf}}}\widehat{U}_c e^{i\frac{\pi}{4}\hat{Y}_{\text{gf}}}. \quad (5.57)$$

which are applied sequentially from right to left on our initial states and includes the $\pi/2$ transmon pulses. First we consider dephasing that occurs during the first \widehat{U}_c

unitary (which can be commuted through because it is error-transparent) and find

$$e^{-i\frac{\pi}{4}\hat{Y}_{\text{gf}}}\hat{U}_c e^{-i\frac{\theta}{2}\hat{X}_{\text{gf}}}\hat{U}_c \hat{\sigma}_z^{gf} e^{i\frac{\pi}{4}\hat{Y}_{\text{gf}}} \quad (5.58)$$

$$= e^{-i\frac{\pi}{4}\hat{Y}_{\text{gf}}}\hat{U}_c e^{-i\frac{\theta}{2}\hat{X}_{\text{gf}}}\hat{U}_c e^{i\frac{\pi}{4}\hat{Y}_{\text{gf}}}\hat{\sigma}_x^{gf} \quad (5.59)$$

$$= \hat{O}(\theta)\hat{\sigma}_x^{gf} \quad (5.60)$$

$$= \hat{\sigma}_x^{gf}\hat{O}(-\theta). \quad (5.61)$$

Dephasing that occurs during the second \hat{U}_c instead results in the overall gate unitary $\hat{\sigma}_x^{gf}\hat{O}(\theta)$. The ancilla ends in $|f\rangle$ and the correct gate unitary $\hat{O}(\theta)$ is applied. From these two scenarios, if the ancilla is observed to be in $|f\rangle$ it is unknown whether the unitary $\hat{O}(-\theta)$ or the unitary $\hat{O}(\theta)$ was applied, this error is only detectable. Just because we used error-transparent building blocks does not guarantee the full gate construction will be error-transparent! Photon loss errors also cause dephasing on the ancilla and thus is also only an error-detectable error. Once we detect photon loss via HFT parity measurements after the gate, we don't know if $\hat{O}(\theta)$ or $\hat{O}(-\theta)$ was applied. Finally, even with χ -matching $|e\rangle\langle f|$ is also only a detectable error. Upon detecting $|e\rangle$, we are again unsure of whether the decay happened during the first or second \hat{U}_c unitaries. Without χ -matching, transmon decay becomes a detectable leakage error (applying a detuned beamsplitter Hamiltonian for unknown time will in general take us out of the codespace). However, we can detect these leakages with high probability by measuring the transmon in $|e\rangle$ followed by qubit resets. In summary, gates made from the exponentiation gadget are really bad at correcting hardware jump errors, but they are very adept at detecting them!

5.3.2 Checking it works: master equation simulations

Now we must test how well our $ZZ(\theta)$ gate works, at least in simulation. We wish to verify that the gate **a**) fails at rate $\sim p$ where p is the probability of an error from

$\{\hat{\epsilon}\}_{\text{hardware}}$ occurring during the gate and \mathbf{b}) has a strongly suppressed gate error $\sim p^2$ on attempts where no errors are flagged. We define the error-prone channel that implements our gate as $\mathcal{E}_{\text{gate}}(\rho)$ and the (perfect) measurement projector that detects no errors for the kitten code to be

$$\hat{M}_{\text{pass}} = |g\rangle \langle g| \otimes \left(\frac{\mathbf{1} + \mathbf{P}}{2} \right)_a \otimes \left(\frac{\mathbf{1} + \mathbf{P}}{2} \right)_b \quad (5.62)$$

and for the dual-rail code

$$\hat{M}_{\text{pass}} = |g\rangle \langle g| \otimes \left(\hat{\Pi}_L \right) \quad (5.63)$$

where $\hat{\Pi}_L$ is the projector onto the dual-rail logical subspace (which acts on 4 cavity modes.)

In words, we declare a successful gate attempt after measuring the transmon in $|g\rangle$ and no photon loss in either cavity from the parity measurements. In chapter 6 we consider the effects of imperfect error-detection measurements but for now we model them as perfect.

We also define the measurement projector for which we detect errors to be

$$\hat{M}_{\text{fail}} = \mathbf{1}_g \otimes \mathbf{1}_a \otimes \mathbf{1}_b - \hat{M}_{\text{pass}} \quad (5.64)$$

Overall gate performance can be quantified with two useful metrics: the average gate failure probability p_{fail} defined as

$$p_{\text{fail}} = \frac{1}{36} \sum_k P_k(\text{declare failure}) = \frac{1}{36} \sum_k \text{Tr} \left(\hat{M}_{\text{fail}} \mathcal{E}_{\text{gate}}(\rho_k) \hat{M}_{\text{fail}}^\dagger \right) \quad (5.65)$$

Where P_k is the probability we declare failure for the k^{th} cardinal state. The cardinal states are defined as $|\psi_k\rangle = |g\rangle \otimes |\psi\rangle_a \otimes |\phi\rangle_b \forall |\psi\rangle, |\phi\rangle \in \{|0_L\rangle, |1_L\rangle, |\pm+L\rangle, |\pm i_L\rangle\}$ for the logical bosonic codewords, of which there are 36 in total.

Similar we can define $\varepsilon_{\text{pass}}$ to be the average state transfer infidelity of the gate

given that no errors were detected, defined as

$$\varepsilon_{\text{pass}} = \frac{1}{36} \sum_k 1 - \langle \psi_k | U_{\text{gate}}^\dagger \frac{\hat{M}_{\text{pass}} \mathcal{E}_{\text{gate}}(\rho_k) \hat{M}_{\text{pass}}^\dagger}{\text{Tr}(\hat{M}_{\text{pass}} \mathcal{E}_{\text{gate}}(\rho_k) \hat{M}_{\text{pass}}^\dagger)} U_{\text{gate}} | \psi_k \rangle \quad (5.66)$$

where U_{gate} is the perfect $ZZ(\theta)$ gate unitary. In words this says $\varepsilon_{\text{pass}}$ is the average state transfer fidelity of the gate, given that it passed with no errors detected.

To perform master equation simulations we use the Hamiltonian

$$\hat{\mathcal{H}}_{\chi\text{BS}} = \hat{\mathcal{H}}_\chi + \hat{\mathcal{H}}_{\text{BS}}(t) + \hat{\mathcal{H}}_{\text{T}}(t). \quad (5.67)$$

For the dispersively coupling in the static Hamiltonian we use

$$\begin{aligned} \hat{\mathcal{H}}_\chi / \hbar = & -\hat{a}^\dagger \hat{a} \left(\frac{\chi_{gf}}{2} |g\rangle \langle g| \right. \\ & + \left(\frac{\chi_{gf}}{2} - \chi_{ge} \right) |e\rangle \langle ge| \\ & \left. - \frac{\chi_{gf}}{2} |f\rangle \langle f| \right), \end{aligned} \quad (5.68)$$

where we choose $\chi_e/2\pi = -0.5$ MHz, $\chi_{gf}/2\pi = -1$ MHz, i.e., no χ -matching and we are working in the frame where our beamsplitter drive is detuned from resonance by $\chi_{gf}/2$

The time-dependent beamsplitter and transmon drives are piecewise-constant throughout the gate sequence and constant when realizing each unitary in the gate construction. This allows us to neglect the effects of any particular choice of pulse shape, since we wish to highlight how the overall protocol fidelity scales with various error rates. Furthermore, because the beamsplitter and transmon pulses are never simultaneous, we can define distinct time-independent Hamiltonians corresponding

to beamsplitter operations and transmon operations:

$$\begin{aligned}\widehat{\mathcal{H}}_{\text{BS}}/\hbar &= \frac{g_{\text{BS}}}{2} \left(\hat{a}^\dagger \hat{b} + \hat{a} \hat{b}^\dagger \right) \\ \widehat{\mathcal{H}}_T/\hbar &= \epsilon_x \hat{\sigma}_x^{gf} + \epsilon_y \hat{\sigma}_y^{gf},\end{aligned}\tag{5.69}$$

where ϵ_x and ϵ_y are the drive strengths of the two control quadratures coupled to the g - f manifold Pauli operators $\hat{\sigma}_x^{gf} = |f\rangle\langle g| + |g\rangle\langle f|$ and $\hat{\sigma}_y^{gf} = i|f\rangle\langle g| - i|g\rangle\langle f|$. For each of the 5 gates that make up the gate sequence, we pick constant values for g_{BS} , ϵ_x and ϵ_y . During transmon operations, for now we neglect the dispersive coupling, as this can also be compensated for with appropriate pulse shaping.

We then use this Hamiltonian in the Lindblad master equation:

$$\frac{d\rho}{dt} = -i \left[\widehat{\mathcal{H}}, \rho \right] + \Gamma_1^T \mathcal{D}[\hat{t}]\rho + \Gamma_\phi^T \mathcal{D}[\hat{t}^\dagger \hat{t}]\rho + \Gamma_1^C \mathcal{D}[\hat{a}]\rho,\tag{5.70}$$

where $\hat{t} = |g\rangle\langle e| + \sqrt{2}|e\rangle\langle f|$ is the annihilation operator for the transmon mode and $\mathcal{D}[\hat{\mathcal{L}}]\rho = \hat{\mathcal{L}}\rho\hat{\mathcal{L}}^\dagger - \frac{1}{2}\hat{\mathcal{L}}^\dagger\hat{\mathcal{L}}\rho - \frac{1}{2}\rho\hat{\mathcal{L}}^\dagger\hat{\mathcal{L}}$ is the usual Lindblad dissipator superoperator where $\Gamma_1^T = 1/T_1^{|g\rangle\langle e|}$, $\Gamma_\phi^T = 1/T_\phi^{|e\rangle\langle e|}$, $\Gamma_1^C = 1/T_1^{\text{loss}}$, where T_1^{loss} is the T_1 relaxation time of the cavity.

The Lindblad equation can also be expressed in terms of a Liouvillian $\widehat{\mathcal{L}}$ as

$$\frac{d\rho}{dt} = -i\widehat{\mathcal{L}}\rho,\tag{5.71}$$

where $\widehat{\mathcal{L}}\rho = \left[\widehat{\mathcal{H}}, \rho \right] + i\Gamma_1^T \mathcal{D}[\hat{t}]\rho + i\Gamma_\phi^T \mathcal{D}[\hat{t}^\dagger \hat{t}]\rho + i\Gamma_1^C \mathcal{D}[\hat{a}]\rho$. The collapse operator for dephasing also includes dephasing of the $|f\rangle$ level, which is expected to dephase at rate $4/T_\phi^{|e\rangle\langle e|}$, four times faster than the $|e\rangle$ level although this is not exact because transmons are anharmonic. Because the transmon and beamsplitter drives are either enabled or disabled for each of the five steps in a given protocol, we can express the

final state density matrix after the whole sequence described in Fig. 5.13 as

$$\rho^f = \widehat{\mathcal{U}}_{T_3} \widehat{\mathcal{U}}_{B_2} \widehat{\mathcal{U}}_{T_2} \widehat{\mathcal{U}}_{B_1} \widehat{\mathcal{U}}_{T_1} \rho^i = \mathcal{E}_{\text{gate}}(\rho^i) \quad (5.72)$$

where each $\widehat{\mathcal{U}}_j = e^{-i\widehat{\mathcal{L}}_j t}$ is the (generally *non-unitary*) propagator under the time-independent Liouvillian corresponding to a transmon pulse $\widehat{\mathcal{U}}_T$ or beamsplitter pulse $\widehat{\mathcal{U}}_B$ in the presence of errors.

Now that we have a way to simulate the gate channel with errors, we can calculate p_{fail} and $\varepsilon_{\text{pass}}$ as a function of coherence times in the transmon ancilla and cavity lifetimes as shown in Fig. 5.14 for the $ZZ(\pi/2)$ gate. (No intrinsic cavity dephasing is modelled but the cavities still undergo dephasing due to transmon decay errors). All $ZZ(\theta)$ gate simulations were performed by **William Kalfus** using the QuTiP python library.

Gate failure, p_{fail} results from a first-order jump error from $\{\hat{\epsilon}\}_{\text{hardware}}$ happening during the gate, and therefore this probability scales linearly with decoherence rate. Since we can detect these errors because the $ZZ(\theta)$ is HFT, $\varepsilon_{\text{pass}}$ is only limited by double jump errors (second-order errors) and hence scales quadratically, allowing us to reach extremely low error-detected gate infidelities with typical cQED coherence times.

Whether we can actually reach these coherence times in experiment depends on how well our error-model actually captures the limiting experimental errors, and how well we can engineer the required Hamiltonian. For instance, we have assumed perfect parametric beamsplitter interactions. We can see readout errors do not pose too much of a problem by replacing our ‘pass’ projector to be

$$\hat{M}_{\text{pass}} = (1 - \eta_{ge} - \eta_{gf}) |g\rangle \langle g| + \eta_{ge} |e\rangle \langle e| + \eta_{gf} |f\rangle \langle f| \otimes \left(\frac{\mathbf{1} + \mathbf{P}}{2} \right)_a \otimes \left(\frac{\mathbf{1} + \mathbf{P}}{2} \right)_b \quad (5.73)$$

Where η_{ge} is the probability of reading out state g given that the transmon is actually

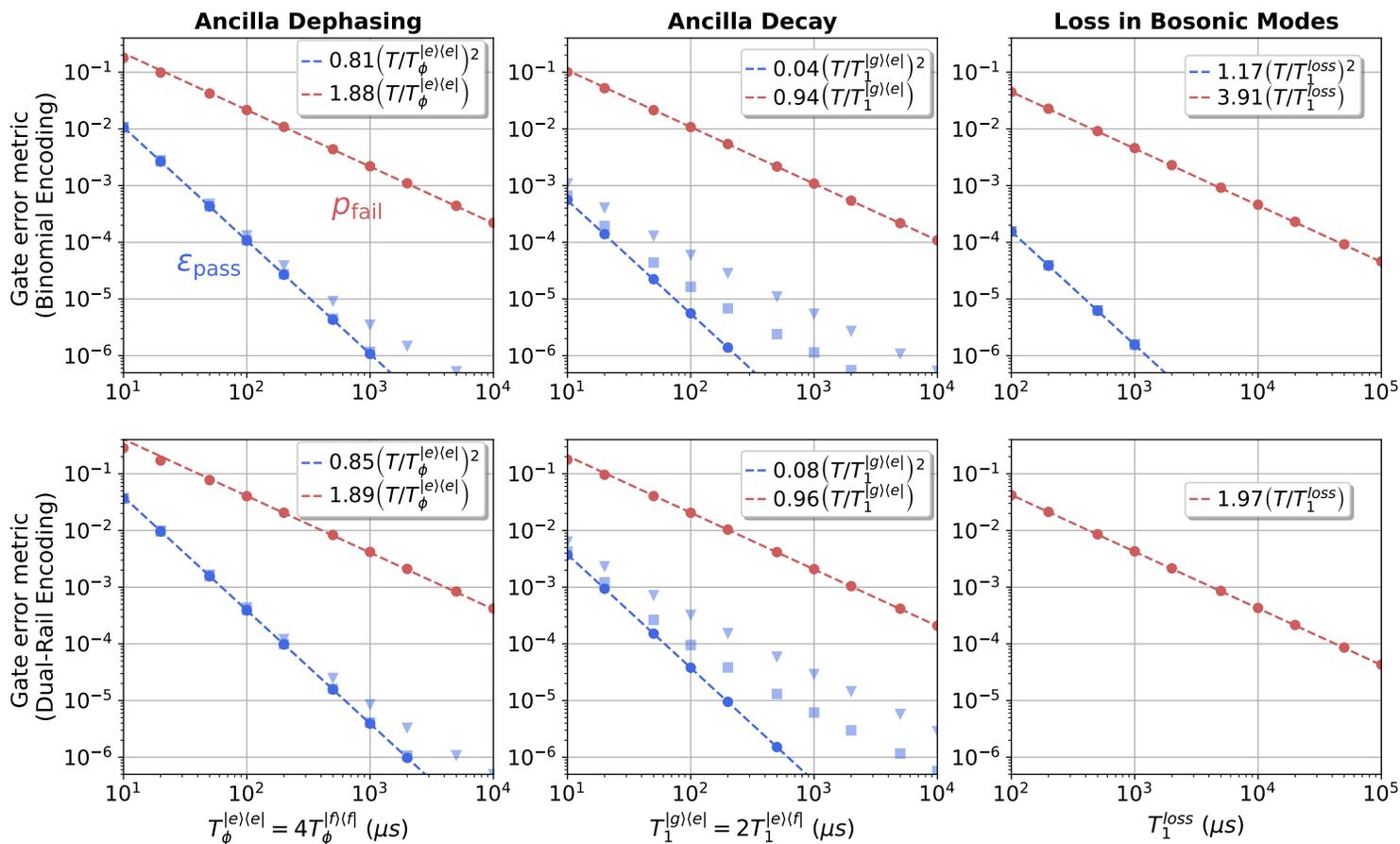


Figure 5.14: **Verifying Hardware Fault-tolerance in the simulated $ZZ_L(\frac{\pi}{2})$ gate construction, with ϵ_{pass} below 0.01%.** Numerical simulation of error-detected $ZZ_L(\frac{\pi}{2})$ gate on two logical qubits in both the binomial encoding (top row) and the dual-rail encoding (bottom row) under three separate error channels—ancilla decay, ancilla dephasing, and photon loss in the bosonic modes. Here we plot p_{fail} (red circles) and ϵ_{pass} (blue markers) vs coherence time. We also calculate ϵ_{pass} in the presence of 1% (blue squares) and 5% (blue triangle) ancilla readout errors. Numerical fits to illustrate these scalings are shown by the blue and red dashed lines. No photon loss data is shown for dual-rail since with perfect error-detection, this error is always detected.

in $|e\rangle$ and η_{gf} is the probability of reading out state g given that the transmon is actually in $|f\rangle$. This is a simple error model shows that readout missassignment error is also a second order error—we need two errors to go wrong to affect $\varepsilon_{\text{pass}}$, an ancilla error followed by a readout assignment error. In Fig. 5.14 we analyze the cases where $\eta_{ge} = 0.01$ and $\eta_{ge} = 0.05$ to model 1% and 5% readout errors. In all cases we also assume $\eta_{gf} = \eta_{ge}^2$, since misassigning $|f\rangle$ to g is usually only possible due to double transmon decay events during readout.

5.3.3 Error scaling

The simulation results shown in Fig. 5.14 show that gate failure scales linearly whilst error-detected gate infidelity scales quadratically, with each hardware decoherence rate considered. The fits take the form $A_n \left(\frac{\tau_{\text{gate}}}{T_{\text{coh}}} \right)^n$. For each of our error channels we find $A_2 < A_1$. How can this be? The ratio $\frac{A_2}{A_1}$ quantifies the likelihood we suffer a second error, given an error has already occurred. We find this number can be less than 1? How can this be?

We begin with the case of photon loss. The probability of a single photon loss occurring determines A_1 . For the binomial code, there are an average of 2 photons in each bosonic mode for a total of 4 photons. Hence $A_1 \approx 4$ for the gate failure probability. Similarly, for the dual-rail code $A_1 \approx 2$. Double photon loss in the kitten code sets A_2 for the error-detected gate infidelity. The probability of double photon loss is $\frac{1}{2} \left(\frac{\bar{n}\tau_{\text{gate}}}{T_{\text{coh}}} \right)^2$, where the factor of 1/2 comes from the fact that photon loss must occur sequentially in a given time window. Half of the time, double photon loss results in the detectable error $\hat{a}\hat{b}$. Overall, this means that $A_2 = \frac{1}{4}\bar{n}^2 \approx 1$

For the gate failure probability resulting from ancilla errors, $A_1 \approx 1$. A single ancilla error results in a failed gate. The values of A_2 for ancilla errors require more detailed analysis. For ancilla dephasing, if two σ_z errors occur within the same ancilla-controlled unitary, they cancel each other out. Only if they occur in different

ancilla-controlled unitaries do they cause a gate error and hence we pick up a factor of $1/2$. When this error happens, the applied gate is $ZZ_L(-\pi/2)$ which causes an error on half of the cardinal states, yielding another factor of $1/2$. Overall, this makes $A_2 \approx 1/4$ for ancilla dephasing.

Finally, A_2 for ancilla decay is the most involved to calculate. Double decay errors require decay to $|e\rangle$, then to $|g\rangle$. The $|g\rangle\langle e|$ decay rate is presumed to be half the $|e\rangle\langle f|$ decay rate, yielding an initial factor of $1/2$. Decay to $|e\rangle$ must happen before decay to $|g\rangle$ in the same time window, giving the next factor of $1/2$. Most of the time, double decay to $|g\rangle$ will leave the system in a random ancilla state in the gf -manifold due to the ancilla rotations in the sequence, and hence double decays are detected as $|f\rangle$ at the end of the sequence half of the time. Finally, we assume that when double decay happens, the bosonic modes may be outside of the codespace, but still have some overlap with the target states. This quantity is difficult to calculate. We define it as $A_{\text{leak}} < 1$. Overall, this means that $A_1 = A_{\text{leak}}/8 < 1/8$ for double ancilla decay errors. In general, these combinatoric factors help further suppress the effects of second-order ancilla errors.

These initial simulations are encouraging, and set a lower-bound on how low $\varepsilon_{\text{pass}}$ can be from these gates at the cost of a percent-level p_{fail} . If used in an erasure code, p_{fail} would bound the erasure rate and $\varepsilon_{\text{pass}}$ would bound the Pauli error rate (from below). We can continue to refine our error-model and Hamiltonian modelling to see what realistic experimental imperfections can impact gate performance.

5.3.4 Kerr and χ'

We know the pure dispersive interaction is only an approximation, with known higher-order corrections. The next-highest-order terms are the self-Kerr and cross-Kerr between bosonic modes, as well as higher-order corrections to the dispersive interaction itself. Although these are typically a factor of 100–1000 weaker than χ , set by $(\frac{g_c}{\Delta})^2$,

they present a source of coherent error unaccounted for in the proposed gate design. Here we quantify their effects on gate performance for realistic experimental parameters.

The higher-order corrections are modelled with the Hamiltonian

$$\hat{\mathcal{H}}_{NL}/\hbar = \frac{K_a}{2} \hat{a}^\dagger \hat{a}^\dagger \hat{a} \hat{a} + \frac{K_b}{2} \hat{b}^\dagger \hat{b}^\dagger \hat{b} \hat{b} \quad (5.74)$$

$$+ (\chi'_{gf} |f\rangle \langle f| + \chi'_{ge} |e\rangle \langle e|) \hat{a}^\dagger \hat{a}^\dagger \hat{a} \hat{a} \quad (5.75)$$

$$+ \chi_{ab} \hat{a}^\dagger \hat{a} \hat{b}^\dagger \hat{b}, \quad (5.76)$$

where K_a, K_b are the self-Kerrs of each bosonic mode, χ'_{ge}, χ'_{gf} are higher-order corrections to the dispersive interaction, and χ_{ab} is the cross-Kerr between the bosonic modes, which may result from both modes participating in the nonlinear element used to actuate the beamsplitter coupling.

What is the expected error on $\varepsilon_{\text{pass}}$ induced by these terms? We can write the rates associated with these corrections as $\tilde{K} \sim (K_a, K_b, \chi'_{gf}, \chi_{ab})$. As a rough approximation, these terms add a state-dependent detuning to the bosonic modes of order $\sim \bar{n} \tilde{K}$, where \bar{n} is the average photon number in the modes. From our intuition for single qubit gates on a Bloch sphere, this detuning will cause the state to miss its target by a small distance proportional to $\sim \bar{n} \tilde{K}$, which includes the possibility of leakage out of the logical codespace. Since fidelity is quadratic in the state overlap, our ancilla-controlled unitaries should only be quadratically sensitive to this detuning, with an infidelity expected to scale as $\sim (\bar{n} \tilde{K} \tau_{\text{gate}})^2 \sim \left(\frac{\bar{n} \tilde{K}}{\chi_{gf}} \right)^2 \sim \bar{n}^2 \left(\frac{g_c}{\Delta} \right)^4$.

When designing a device for experiment, we can control χ_{gf} through the geometry of the device. Increasing this parameter causes τ_{gate} to decrease linearly but \tilde{K} to increase *quadratically*. As such, reducing χ_{gf} will always reduce the gate error from $\hat{\mathcal{H}}_{NL}$. On the other hand, reducing χ_{gf} will always increase the errors due to deco-

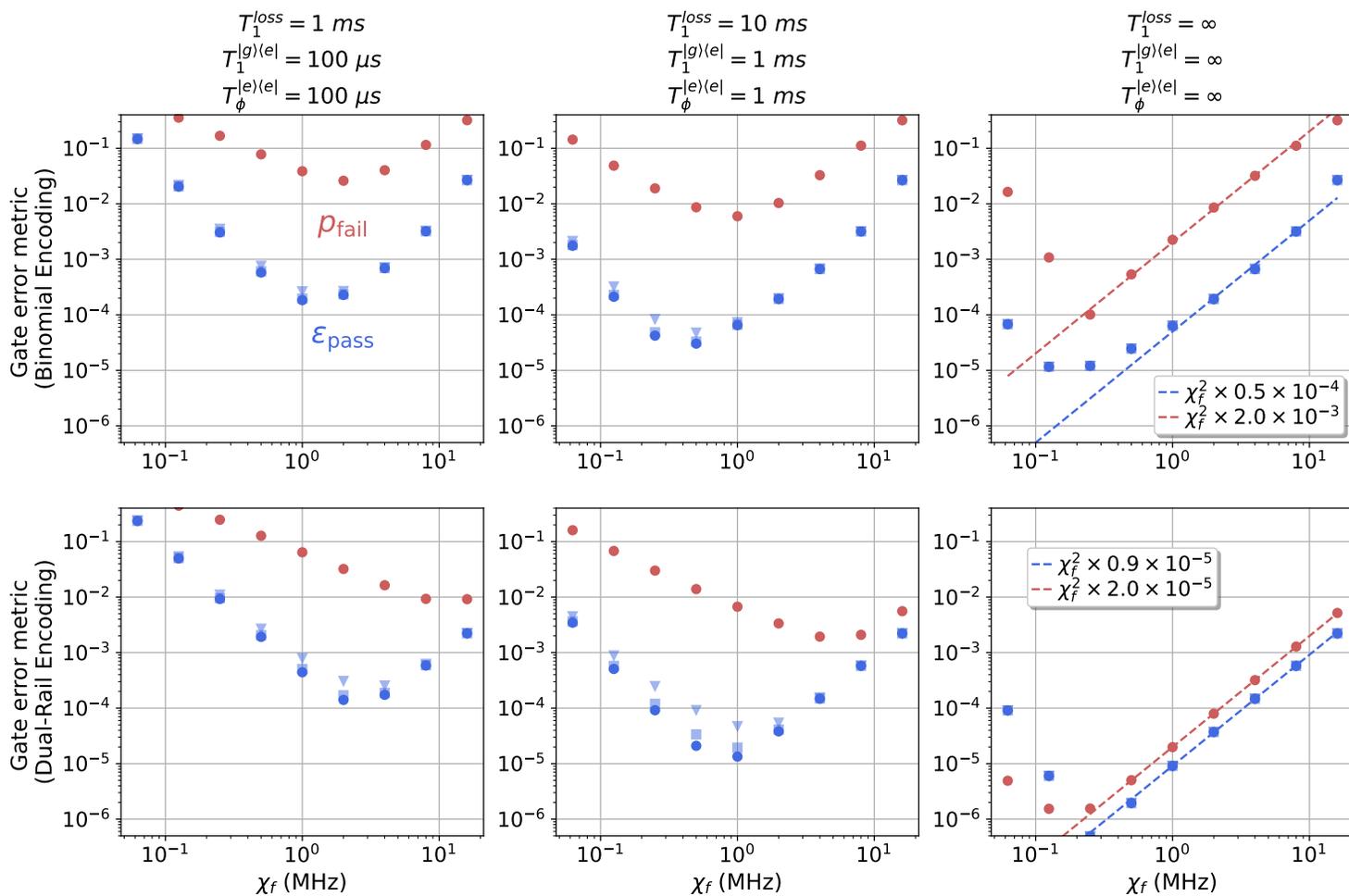


Figure 5.15: **Effects of Kerr and χ'** p_{fail} (red circles) and ϵ_{pass} (blue circles) for three sets of coherence times as functions of χ_{gf} . We investigate both the kitten code and dual-rail code and find optimal values of $\chi_{gf} \sim 1$ MHz for realistic (left column) and optimistic (middle column) coherence times.

herence by increasing τ_{gate} . These competing effects mean there is an optimal value of χ_{gf} to engineer.

Simulations (performed by **William Kalfus**) displayed in Fig. 5.3.4 highlight the competition between nonlinearity and decoherence. From typical values we measure in experiment [Reinhold et al., 2020], we assume an initial operating point where $\chi'_{gf}/2\pi = 2$ kHz, $\chi'_{ge}/2\pi = 1.125$ kHz and $K_a/2\pi = K_b/2\pi = 2$ kHz when $\chi_{gf}/2\pi = -1$ MHz, and that these quantities scale quadratically if we were to vary χ_{gf} . We also assume $\chi_{ab}/2\pi = 100$ Hz for all χ_{gf} , chosen to reflect values attainable in modern hardware. For the transmon pulse we again neglect the dispersive interaction and its higher-order corrections.

The rightmost column shows that in the absence of decoherence, the error associated with nonlinearity scales quadratically in χ_{gf} as expected. The leftmost and center columns highlight that when decoherence is introduced, the benefits of a weaker coupling to the nonlinear ancilla are overridden by the increased incoherent error experienced by the slower gate. This leads to an optimum which may be found for different bosonic codes and decoherence rates. For the codes and parameters explored here, we find the optimal χ_{gf} to be ~ 1 MHz and observe persistent error-detected gate infidelities $\sim 0.1\%$ and below.

This simulations show that designing a HFT two-qubit gate can have big implications for reducing hardware error rates without needing to improve coherence times. In the context of an erasure qubit, $\varepsilon_{\text{pass}}$ represents the residual Pauli error when no errors are detected. These simulations suggest we can achieve $\varepsilon_{\text{pass}} \lesssim 0.1\%$, far below the Pauli threshold of $\sim 1\%$ for circuit-level noise in a surface code. Even more impressive is we can be below threshold when our hardware error rate is at the level of $\sim 1\%$ per gate!

The other requirement is that p_{fail} , which would set the erasure rate per gate must also be sufficiently below the erasure threshold of $\sim 5\%$ for Yue Wu's circuit-level

noise model for erasure errors [Wu et al., 2022]. Our simulations show we can reach $p_{\text{fail}} < 5\%$, $\varepsilon_{\text{pass}} \ll 1\%$ for the dual-rail encoding which makes it a promising candidate for an erasure qubit, especially since the two-qubit entangling gate is often a highly error-prone operation in cQED.

Experimental implementation should be imminent, especially given that all these building blocks (dispersive interaction, g - f control and parametric beamsplitters) have all been experimentally realized individually to extremely high fidelity. [Wallraff et al., 2004, Rosenblum et al., 2018b, Lu et al., 2023, Chapman et al., 2022]

5.3.5 Realistic pulse shaping

The last ‘added realism’ we will investigate for $ZZ(\theta)$ gates is the effect of experimentally realistic pulse shaping. This addresses the practical constraint of finite pulse bandwidth. We find pulse shaping can fix two unwanted effects that arise from finite pulse bandwidth: unwanted dynamics during the gradual ramp-up of our beamsplitter amplitude, g_{BS} and unwanted dynamics during the short but finite-duration transmon rotations. These unwanted dynamics are due to the ‘always-on’ nature of the dispersive interaction between the transmon and cavity mode in the cQED systems we consider. Crucially, we show that simple pulse shaping solutions do not dramatically increase the duration or complexity of the gate. The full pulse sequence example is illustrated in Fig. 5.16.

To engineer our example pulse, we split the $ZZ_L(\theta)$ gate into 5 parts which are individually optimized: 3 ancilla rotations and 2 identical \mathbf{cZZ}_L unitaries. For simplicity, we examine the $ZZ_L(\pi/2)$ gate in the 0-1 encoding. We use optimal control pulses found with Gradient Ascent Pulse Engineering (GRAPE) and simple pulse detunings to compose our example pulse sequence. By analyzing the subsequent dynamics with QuTiP master equation simulations we verify we can perform the gate with coherent errors $\sim 0.01\%$ and with expected error-detected gate infidelities $\sim 0.01\%$, showing

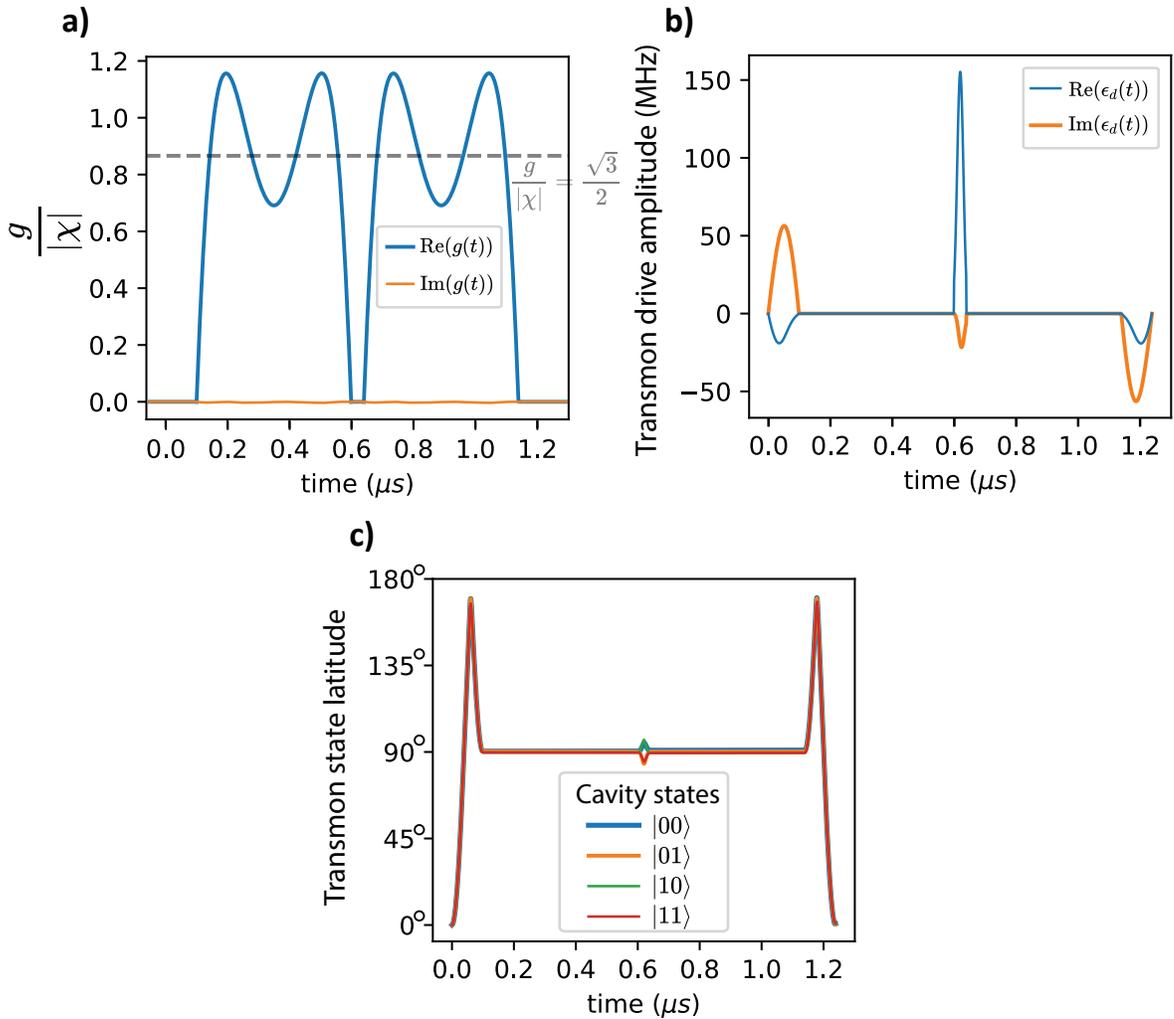


Figure 5.16: Example of pulse shaping for a $ZZ(\pi/2)$ gate in the 0-1 encoding. **(a)** Pulse for the beamsplitter drive. Dashed line indicates the beamsplitter strength required if we were to use rectangular (instantaneous ramp-up) pulses. **(b)** Pulse shaping for the three ancilla rotations during the $ZZ(\pi/2)$ gate. All pulses closely resemble Gaussian $3\pi/2$ pulses. **(c)** Latitude of the transmon on its g - f Bloch sphere, showing that we respect the equal latitude condition for the four different possible cavity states.

how we can overcome finite-bandwidth effects without compromising the key benefits of our gate.

Transmon rotations within the g - f manifold are expected to take 50–100 ns with current hardware, since we must be careful to avoid unwanted leakage to the $|e\rangle$ level (although this error is also detectable). We use GRAPE to find pulse $\epsilon_d(t) = \epsilon_x(t) + i\epsilon_y(t)$ with the Hamiltonian

$$\mathcal{H} = \epsilon_d(t) |f\rangle \langle g| + \epsilon_d^*(t) |g\rangle \langle f| + \frac{\chi}{2} (|g\rangle \langle g| - |f\rangle \langle f|) \hat{a}^\dagger \hat{a},$$

working with the constraint that the transmon reaches the equator state $(|g\rangle + |f\rangle)/\sqrt{2} = |+\rangle_{gf}$ after 100 ns, regardless of whether there are one or zero photons in cavity mode \hat{a} . Future work can also explicitly model the transmon's anharmonicity in the Hamiltonian to minimize leakage to $|e\rangle$ in the pulse. Here we assume our transmon drive can directly drive transitions in the g - f manifold.

We find our numerically optimized pulses closely resemble Gaussian $3\pi/2$ pulses (rather than a $\pi/2$ pulses) to reach $|+\rangle_{gf}$. A similar pulse is also found for the final transmon rotation, which takes $|+\rangle_{gf}$ to $|g\rangle$. For the intermediate transmon rotation, we use a detuned Gaussian pulse of duration 40 ns, also with an amplitude close to what would be required for a $3\pi/2$ pulse. We chose the detuning and phase such that the transmon state stays as close as possible to the equator plane of the g - f Bloch sphere.

In order to preserve error-detection to transmon dephasing errors, we should adhere to what we call the *equal latitude condition* as closely as possible throughout the gate. That is, the latitude of the transmon on its $|g\rangle$ - $|f\rangle$ Bloch sphere should remain independent of the state in the cavities at all times during the gate. We can see in Fig 5.16 c. this is mostly the case for our numerically optimized pulse, except for the intermediate transmon rotation where there is a slight deviation. Whilst we can no

longer error-detect 100% of transmon dephasing events during this rotation, as far as we can tell, this does not noticeably limit our error-detected gate fidelity in this particular pulse. Future pulse optimization efforts can take this equal latitude condition into account in the cost function, which seems to be a more general requirement for being able to detect ancilla dephasing errors.

Pulses for the \mathbf{cZZ}_L unitary are found from the Hamiltonian

$$\mathcal{H} = \frac{1}{2} (g(t)a^\dagger b + g^*(t)ab^\dagger) + \frac{\chi}{2} a^\dagger a (|g\rangle\langle g| - |f\rangle\langle f|),$$

with the constraint $g(0) = g(2\pi/\chi) = 0$. The solutions we find generally require g_{max} slightly larger than $\frac{\sqrt{3}}{2}\chi$.

The total gate duration for $\chi/2\pi = 2$ MHz is 1240 ns, which is close to the $4\pi/\chi = 1000$ ns limit for the fastest gate time possible with our construction. The actual unitary we implement is $ZZ_L(\pi/2)Z_1(\theta_1)$ where the additional single-qubit Z rotation is an artifact of the static dispersive interaction. For this particular pulse, $\theta_1 = 0.03$, an almost negligible effect. In the absence of decoherence, we obtain an average gate infidelity of 0.015% and average failure probability of 0.039% (both due to coherent control errors).

We expect to be only quadratically sensitive to pulse shaping errors. When we model pulse amplitude error by increasing all pulse amplitudes by 0.5%, we obtain an average gate infidelity of 0.030% and average failure probability of 0.065%. Similarly, when we detune all pulses by +10 kHz, we obtain an average gate infidelity of 0.030% and average failure probability of 1.8%, mainly from detuned ancilla pulses.

When we set $T_1^{|g\rangle\langle e|} = 100 \mu\text{s}$, $T_\phi^{|e\rangle\langle e|} = 100 \mu\text{s}$, $T_1^{\text{loss}} = \infty$ we find an average gate failure probability of 2.4% and an average error-detected gate infidelity of 0.019%, demonstrating that we maintain low error rates. (Including photon loss, we estimate an additional 0.2% failure probability if we were to use this pulse sequence for the

Dual-Rail encoding with $T_1^{\text{loss}}=1$ ms per cavity).

5.4 Pole-to-Pole: other useful trajectories on the Operator Bloch Sphere

Unlike a qubit Bloch sphere, we must be careful to only take trajectories that leave us back in the codespace. For this to be the case, the trajectory must end at one of the poles. At intermediate times, the evolution takes us out of the codespace but thanks to HFT, we can detect for errors that happen whilst we are out of the codespace after the gate, once we are back in the codespace (or the photon loss error space). This means there are only two types of trajectories. Ones where we end at the same pole we started from (performing \hat{R}_ϕ for geometric phase ϕ enclosed, or trajectories where we end at the opposite pole corresponding to a \hat{R}_ϕ SWAP unitary applied to the cavities.

With the constraint that the cavities must return back to the codespace, this means we can only engineer entangling unitaries of the form $ZZ(\theta)$ with the control unitary \mathbf{cZZ}_L or $e\text{SWAP}(\theta)$ via the control unitary \mathbf{cSWAP} , or combinations of the two. This is not too surprising since our Hamiltonian restricts us to excitation-preserving two-qubit gates anyways. One curiosity is that the OBS gives us several ways to realize the \mathbf{cSWAP} control unitary, where one trajectory returns to the initial pole and the other returns to the opposite pole as shown in Fig. 5.17. The geometric phase enclosed along the paths are different so this does not do \mathbf{cSWAP} exactly but rather the unitary

$$\hat{R}_{\phi_g} |g\rangle \langle g| + \hat{R}_{\phi_f} \text{SWAP} |f\rangle \langle f| \quad (5.77)$$

Where unwanted R_ϕ control unitaries must be mitigated, e.g., by adding delays before and after the control unitary.

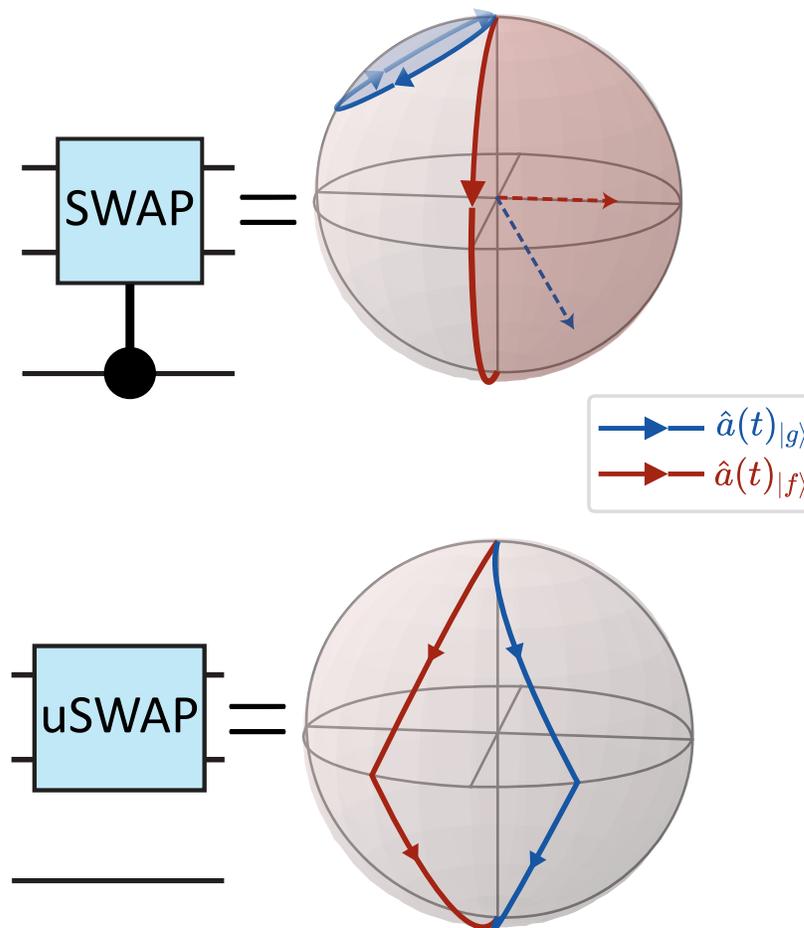


Figure 5.17: Trajectories on the operator Bloch sphere for **cSWAP** and unconditional SWAP (**uSWAP**) unitaries. The geometric phase enclosed by the trajectories is different, resulting in additional unwanted unitaries of the form $|g\rangle\langle g| + \hat{R}_{\phi_f - \phi_g} |f\rangle\langle f|$ which can be nulled by adding delays before and after the unitary.

The other trajectory of interest is what we call the ‘unconditional SWAP’ where the cavity states are swapped regardless of the state in the transmon. We can almost achieve this with trajectories on the Bloch sphere as shown in Fig. 5.17 which can be engineered by detuning the parametric beamsplitter by $\chi_{gf}/2$ and ensuring $g_{BS} \geq \chi_{gf}/2$. Once the trajectories reach the equator of the OBS we perform a π -pulse in the g - f manifold to effectively reverse the detunings. Similarly, this does not implement a true unconditional SWAP unitary but rather the unitary

$$\hat{R}_{\phi_g} \text{SWAP} |g\rangle \langle g| + \hat{R}_{\phi_f} \text{SWAP} |f\rangle \langle f| \quad (5.78)$$

$$= \hat{R}_{\phi_g} (\text{SWAP} |g\rangle \langle g| + \hat{R}_{\phi_f - \phi_g} \text{SWAP} |f\rangle \langle f|) \quad (5.79)$$

Again, the unwanted conditional rotations must be undone by adding delays before and after the unitary and using the dispersive interaction, or by finding trajectories such that $\hat{R}_{\phi_f - \phi_g} = \mathbb{1}_L$.

Another way to perform unconditional SWAP is to detune our beamsplitter coupling by $\chi/2$ as usual, and set $g_{BS} = |\chi|/2$ such that the polar angle of both precession vectors is 45° . After applying this Hamiltonian for time $t = \sqrt{2}\pi/\chi$, both trajectories reach the equator (and are antipodal). If we then flip the sign of the beamsplitter drive, such that $g_{BS} = -|\chi|/2$ then after the same duration both trajectories will reach the southpole at the same time. The area between these trajectories on the OBS is 2π steradians. This trajectory is shown in Fig. 5.18.

Since the parametric beamsplitter interaction on the Operator Bloch Sphere is analogous to a Rabi drive on the qubit Bloch Sphere, any pulse shaping techniques (such as composite pulses) are also transferable.

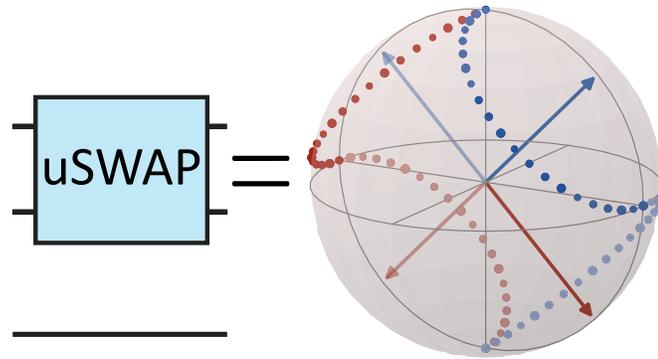
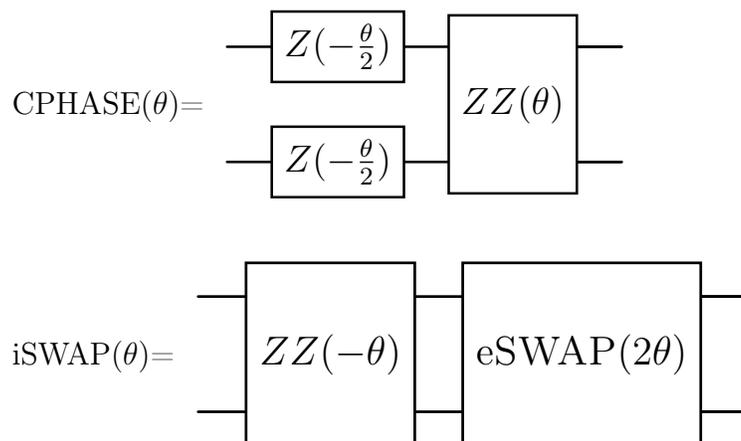
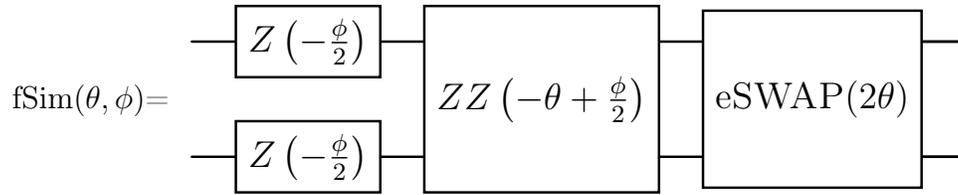


Figure 5.18: Trajectories for an unconditional SWAP. This time, we flip the sign of g_{BS} rather than χ at the halfway point in the sequence. The conditional trajectory is shown by blue (red) circles for transmon state $|g\rangle$ ($|f\rangle$)

5.4.1 A family of excitation-preserving two-qubit gates

As we explored back in Sec. 2.2.4, $Z(\theta)$, $ZZ(\theta)$ and $e\text{SWAP}(\theta)$ can construct any excitation-preserving gate. In this chapter, we have shown each of these constituent gates can be HFT when error-detected and realized with the dispersive Hamiltonian and a parametric beamsplitter (which themselves are also excitation-preserving). It then follows we can realize any excitation-preserving gate on rotationally symmetric bosonic codewords with this HFT error-detection property. Below we list recipes for constructing some more well-known gates from these building blocks:





where fSim is the fermionic simulation gate. (See [Google AI Quantum et al., 2020]) Are excitation-preserving gates sufficient to perform universal computation? Yes! But we must use a multimode encoding. The dual-rail code is the simplest multimode encoding for which all logical gate operations are excitation-preserving.

5.5 Conclusions

We began this chapter with the goal of finding an easy-to-engineer two-qubit entangling gate for some common bosonic qubits (the rotationally symmetric ones designed to correct for single photon loss). Most of the time on a quantum processor will be spent doing gates, rather than idling so it is important that this gate be able to correct for bosonic errors (i.e., photon loss) that occur before, after or during the gate.

This turned out to be non-trivial! There are only a few Hamiltonians we can engineer (namely linear and bilinear couplings as shown in Tab. 5.3) that preserve this property! We also need a control-ancilla (we opt for a transmon) to give us the non-linearity required for these gates³, but now we are even more susceptible to ancilla errors which corrupt our bosonic qubits!

The property we want our two-qubit gates to have is hardware fault-tolerance (HFT). That is, the dominant errors in the ancilla and cavities may occur during the gate, but are still correctable after the gate. Moreover, as we saw with $e\text{SWAP}$, just because our building blocks are HFT does not guarantee the overall gate construction is HFT.

3. For the GKP code, we can realize all Clifford gates with Gaussian operations which do not require a control-ancilla

We have found two easy-to-implement solutions to this problem that use only the dispersive interaction and a parametric beamsplitter interaction with hardware in Fig 5.6. Both are based on being able to engineer the \mathbf{cZZ}_L unitary and control an ancilla transmon in its g - f manifold. The first solution is an error-corrected QND ZZ measurement which in principle has the same ‘entangling power’ as a CNOT gate but in practice requires many repeated measurements to check for errors. The second solution is not to do error-correction, but error-detection, which allows us to use the exponentiation gadget with \mathbf{cZZ}_L to realize a $ZZ(\theta)$ gate. For $\theta = \pi/2$ this gate is locally equivalent to a CNOT.

The fact that hardware jump errors may be detected but not corrected is not as limiting as we might think. We do not have to postselect out attempts that detect errors, instead we can flag these as erasure errors and fully correct them in the next layer of the code. In fact, the error scaling ends up being the same as the case where we fully corrected these errors at the hardware level!

If we are now only *detecting* errors at the hardware level, it makes more sense to use a bosonic code that is designed to *detect* single photon loss rather than *correct* for it. Our choice of code for this is the dual-rail encoding. This will in fact make our lives a lot simpler! Having such simple codewords will vastly simplify the set of operations needed for full quantum computing. We now move onto the next chapter, where we lay out a new architecture for quantum computing based entirely on the dual-rail code for microwave cavities, where the goal is now to detect dominant hardware errors, convert them to erasure errors and correct them in the next level of encoding which will be the surface code.

The Dual-rail Code with Superconducting Cavities

6.1 Error-detection Enables Erasure Qubits

So far we have considered ways to realize hardware-fault tolerance in bosonic qubits in cQED. What we have found is error-detection of hardware errors is significantly easier than error-correction at the hardware-level, but provides much of the same benefits thanks to the possibility of using the physical qubits as erasure qubits. We saw back in Chapter 3 that if one can detect ‘first-order’ hardware errors and treat them as erasures, to leave a residual set of ‘second-order’ hardware errors that contribute to the Pauli noise, then our erasure qubit will have close to the optimal ratio of erasure to Pauli errors for a stabilizer code designed to correct for both Pauli and erasure errors.

Ultimately, what we want for practical error correction (with say a surface code or similar large-scale error correction code with a threshold) is for error rates on the physical qubits to be far-below their respective error thresholds. Erasure qubits pro-

vide a path to achieving this with a lot less stringent demands on the hardware error rates, the rate at which jump errors occur due to physical processes. In particular, 1% level hardware errors are now acceptable. In cQED experiments, achieving 1% error rates in all operations (gates, measurements and error-detection itself) is immminently doable, with logical measurement recently demonstrated [Chou et al., 2023]. Achieving 0.1% error rates to be sufficiently below threshold for a standard surface code (without erasure qubits) in all operations has been achieved for gates [Sung et al., 2021] on small scale processors but not state preparation and readout (and gate fidelity degrades in larger scale systems [Acharya et al., 2023].)

The dual-rail code with superconducting cavities is then the simplest way we can think of for realizing an erasure qubit in cQED. The dominant hardware jump errors are detected and treated as erasures to first-order, and ‘double-jump’ errors and much rarer hardware errors set the Pauli error rate. The dual-rail architecture we propose here provides recipes to combine the well-established building blocks of cQED in new ways to realize all operations needed for a good erasure qubit. These building blocks are transmons, dispersive coupling of transmons to microwave cavity modes and single-shot dispersive readout of the transmon. Aside from the recent realization of high-fidelity beamsplitters, there is nothing fundamentally new in the hardware. Rather, we are combining the building blocks we all know and love in a new way, within the paradigm of error-detection and erasure qubits.

It is no accident that the development of dual-rail cavity qubits coincided with the advent of high-fidelity cavity-cavity beamsplitter operations [Lu et al., 2023, Chapman et al., 2022]. Bosonic encodings such as GKP or cat qubits rely on parametric beamsplitter operations or related processes to do entangling gates. From an experimental perspective, the dual-rail encoding is attractive because it only requires a beamsplitter interaction for single qubit gates, and a beamsplitter interaction combined with a cavity-transmon dispersive interaction for entangling gates.

In this architecture we must delicately balance how we deal with hardware errors in not just the transmon ancilla but also the cavities, in all operations! In the previous chapter we already solved this problem with an HFT error-detectable \mathbf{cZZ}_L which can be used for both dual-rail entangling gate and mid-circuit erasure detection, arguably the most difficult parts of realizing an erasure qubit. As we unveil our architecture we will also detail schemes for state preparation, logical readout and single-qubit gates and take into account as many known experimental errors and imperfections as we can to make our proposal as water-tight as possible.

Compared to other bosonic codes, the dual-rail code is remarkably simple and has been largely overlooked until recently in cQED, because it was largely assumed that one had to correct bosonic errors such as photon loss at the hardware level, before concatenation. The concept of the erasure qubit alleviates this requirement, allowing one to detect errors at the hardware level and correct them at the level of the stabilizer code instead by treating them as erased physical qubits.

6.2 Engineering the Error Hierarchy

The concept of erasure qubits has seen a renaissance at the time of this writing in several qubit platforms. In neutral atoms/trapped ion platforms, erasure qubits are attractive because leakage errors during two-qubit gates can be the largest source of error, but converted to erasures if mid-circuit erasure conversion can be performed [Wu et al., 2022, Ma et al., 2023, Kang et al., 2023]. What sets cQED dual-rail proposals apart is that most aspects of the system can be specifically engineered to have detectable errors. In our case, the choice of bosonic codewords, the exact way the control ancillas interact with the cavities, and the choice of parametric interactions all ensures that the dominant first-order hardware errors are easily detectable via standard transmon measurements, and remaining errors contribute to a much smaller

Pauli error rate (and an even smaller leakage rate). We call this engineering the error hierarchy. Whilst an idle qubit may naturally exhibit a strong hierarchy of errors, it is not something that tends to occur by default in qubit operations and care needs to be taken when designing every operation to ensure we preserve the error hierarchy.

One way to think about what this means is to imagine there are three relevant ‘error hierarchies’. The first is the **Optimum error hierarchy**. Any stabilizer code can correct for double the number of erasure errors per round as Pauli errors, and it can correct for simultaneous erasure and Pauli errors (assuming both are well below their relevant thresholds. Thresholds can only really be numerically computed by holding one error rate fixed and varying the other). In the optimum error hierarchy we have

$$p^{\text{Erasure}} > p^{\text{Pauli}} > p^{\text{Leakage}} \quad (6.1)$$

$$p > p^2 \frac{p_{\text{th}}^{\text{Pauli}}}{(p_{\text{th}}^{\text{Erasure}})^2} > 0 \quad (6.2)$$

$$p > 4p^2 > 0 \quad (6.3)$$

Where the last line is when we assume the erasure model in [Wu et al., 2022] and p is the characteristic probability of a hardware jump error happening during an operation. All the optimum error hierarchy says is this is the ideal ratio of erasure to Pauli errors for both to contribute equally to the total logical error rate (assuming both are far-below threshold and large code distance).

Next we have the **Idling error hierarchy**. This is the hierarchy of error rates when we do no operations on our physical qubits, except those needed to preserve the qubit information as a quantum memory. For an erasure qubit, this just means we perform mid-circuit erasure detection and no other operations. This sets a lower bound for how good the **Operational error hierarchy** can be.

The operational error hierarchy is ultimately the one we care about and the one we

compare with the optimum error hierarchy. These are the average error rates whilst doing single-qubit gates, two-qubit gates, physical qubit state preparation and readout — everything else needed to run a quantum computer. Our choice of physical qubit sets the idling error hierarchy, and our design of the gate/measurement operations sets the operational error hierarchy, which is inevitably worse than the idling error hierarchy but still aims to adhere closely to the optimum error hierarchy.

If the operational error hierarchy does not match the optimum error hierarchy but error rates are still below threshold for our noise-model, this is still quite a good spot to be in! What this means is that either the erasure rate or the Pauli rate will become the limiting error rate at large code distances and the other error type will become negligible, e.g., if $p^{\text{Erasure}} = 2\%$ and $p^{\text{Pauli}} = 0.001\%$ then both could be below threshold in a surface code, but the logical error rate will quickly be dominated by the erasure errors as we scale to larger code distances.

6.3 Dual-rail in Microwave Cavities

Why do we choose to encode our information in microwave cavities? Three arguments are often given. For a *bare* cavity:

- Dominant error channel is single photon loss, dephasing is negligible
- Long T_1 times, exceeding the best J.J.-based qubits by 10× or more
- They have a large Hilbert space in which we can redundantly encode information

The dual-rail cavity qubit benefits greatly from the first two points. (As we have seen in Chapter 5, the last point is often more of a hindrance!). Two cavity modes (Alice and Bob) are required to realize the dual-rail qubits with a single photon shared between them. The $|0_L\rangle$ and $|1_L\rangle$ states are defined with the photon residing in Alice

or Bob as $|0\rangle_L = |01\rangle$ and $|1\rangle_L = |10\rangle$. We can quickly see why this makes a good erasure qubit.

Under only photon loss, we can decay to the leakage state $|00\rangle$. We call this a **cavity decay error** in a dual-rail qubit. To very good approximation, this is the *only* error that can occur in an idle dual-rail cavity qubit. It is not desirable to stay in this leakage state for long. Although it is a rather inert state, we saw in Sec 3.1.6 of Chapter 3 how damaging leakage errors can be if they are allowed to accumulate. For this to be converted to an erasure error, we must use mid-circuit erasure detection which distinguishes between the leakage space $\{|00\rangle\}$ and the dual-rail codespace $\{|01\rangle, |10\rangle\}$, without destroying the coherence of superpositions of states in the codespace! Later in Sec. 6.5.5 we give ways of realizing the mid-circuit erasure detection, e.g., via the QND joint-parity measurements introduced in Sec 5.2 of Chapter 5.

The dual-rail cavity qubit has multiple hardware components - two superconducting cavities (Alice and Bob), a single parametric coupler, and a single transmon ancilla with its own dispersive readout, as shown in Fig. 6.1.

6.4 Expected Idling Error Hierarchy

To estimate the physical error rate of the proposed dual-rail cavity qubit during idling time, and the logical errors in the encoded qubit they can cause, we assume typical performances achievable for the individual components today. We also assume that our qubits will not be idle for more than $\sim 1 \mu\text{s}$

By construction, the idling error hierarchy shown in Tab. 6.1 is well-aligned to the optimal error hierarchy of an erasure qubit: the dominant error is photon loss out of the computational space caused by the average cavity relaxation at rate $\bar{\kappa} = (\kappa_a + \kappa_b)/2$, which, by measuring the joint photon number parity, can be converted into

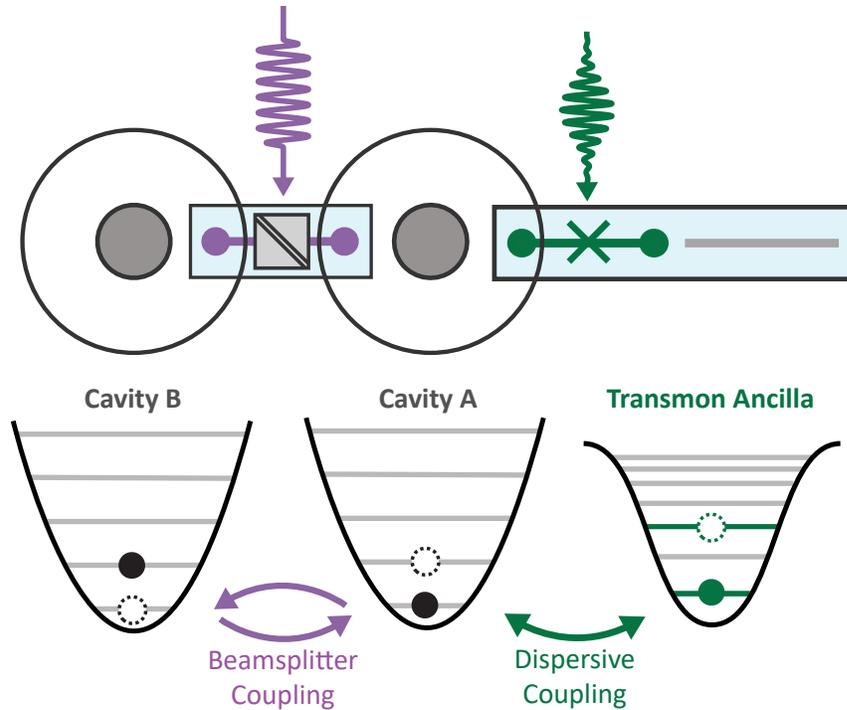


Figure 6.1: cQED hardware needed to operate one dual-rail cavity qubit. The parametric beamsplitter interaction is shown in purple whilst a transmon ancilla operated in its g - f manifold is shown in green.

erasure errors. During the idling time, phase flip events are predominantly caused by heating events in the transmon, where $\bar{n}_{\text{th},A}$ is the average thermal ancilla population and $\Gamma_{1,A}$ the energy relaxation rate. We also presume a phenomenological intrinsic cavity dephasing rate of 10 ms although the actual rate is thought to be much lower for typical 3D microwave cavities.

By measuring the transmon via dispersive readout and ensuring it remains in its ground state we can detect ancilla heating events, at least to first-order. Changes in population in the $|01\rangle \leftrightarrow |10\rangle$ states are a consequence of the no-jump backaction, a second-order effect, in the event that there is a finite difference $\Delta\kappa = \kappa_b - \kappa_a$ between the single-photon decay rates of the cavities, or by the unlikely combination of a cavity decay and heating event, resulting in extremely low rates. Leakage events in the cavities are in principle always detectable by performing the appropriate measurements of the cavity populations; if not by joint-parity measurements then by the

Error process	Scaling	Probability in 1 μs	Erasure noise bias	Effective lifetime	Error type
cavity photon loss	$\bar{\kappa}t$	10^{-3}	1	1 ms	Leakage*
cavity heating	$\bar{n}_{\text{th}}\bar{\kappa}t$	10^{-5}	10^2	100 ms	Leakage*
cavity dephasing	$\gamma_{\varphi}t$	10^{-4}	10^1	10 ms	Phase flip
ancilla heating	$\bar{n}_{\text{th},A}\Gamma_{1,A}t$	10^{-4}	10^1	10 ms	Phase flip*
no-jump backaction	$(\frac{1}{4}\Delta\kappa t)^2$	10^{-6}	10^3	1 s	Phase flip
cavity photon loss + heating	$\bar{n}_{\text{th}}(\bar{\kappa}t)^2$	10^{-8}	10^5	100 s	Bit flip
cavity heating $\times 2$	$3(\bar{n}_{\text{th}}\bar{\kappa}t)^2$	10^{-9}	10^6	~ 1 h	Leakage

*=can convert to **erasure error** via mid-circuit measurements.

Table 6.1: The expected idling error hierarchy for a dual-rail cavity qubit.

appropriate set of number-selective π -pulses or similar. However, joint-parity measurements alone should be sufficient to discern the vast majority of leakage events to $|00\rangle$, the most common leakage state and $|11\rangle$, the next most common leakage state (already rare with probability 10^{-5} per μs .)

For the calculation of the error probabilities, we assume single-photon decay rates of $\kappa_a = (1.5 \text{ ms})^{-1}$ and $\kappa_b = (0.5 \text{ ms})^{-1}$ for the cavities where we have chosen $\Delta\kappa \sim \bar{\kappa}$ to examine the ‘worse case scenario’ for no-jump backaction, $\Gamma_{1,A} = (100 \mu\text{s})^{-1}$ for the ancilla transmon, pure dephasing rates $\gamma_{\varphi,a} = \gamma_{\varphi,b} = (20 \text{ ms})^{-1}$, and a thermal population $\bar{n}_{\text{th}} = \bar{n}_{\text{th},A} = 0.01$ for both the cavities and transmon ancillas. We define the erasure noise bias to refer specifically to erasure errors happening much more frequently than all other types of errors such as Pauli and leakage errors. The erasure noise bias is quantified for each error process by comparing to the cavity photon loss error rate, the dominant source of erasure errors in the system when idling.

Some of the idling coherence times in Tab. 6.1. have been measured or bounded by recent experimental work by [Chou et al., 2023], on which I am a co-author.

6.4.1 No-jump backaction

No-jump backaction is a curious error in an idle dual-rail qubit. Although it is projected to be a small error at short times, and completely correctable by regularly applying SWAP operations to the cavity to echo-out the error, it is interesting from a theoretical perspective and is relevant for the long-time dynamics of an idling dual-rail qubit. This polarization error channel can be written with only $\mathbb{1}$ and Z Pauli operators making it a dephasing channel. Unlike ‘pure’ dephasing due to, e.g., fluctuations in a mode’s frequency which only act to destroy the coherence in a superposition, no-jump backaction affects the state populations in $|0_L\rangle$ and $|1_L\rangle$, increasing the probability of being in the less lossy cavity.

Within the short time interval δt we can model decay in a dual-rail cavity qubit with the set of Kraus operators

$$\begin{aligned}\hat{E}_0 &= \sqrt{1-p}|01\rangle\langle 01| + \sqrt{1-q}|10\rangle\langle 10| \\ \hat{E}_1 &= \sqrt{p}|00\rangle\langle 01| \\ \hat{E}_2 &= \sqrt{q}|00\rangle\langle 10|\end{aligned}\tag{6.4}$$

Where p is the probability of photon loss in cavity Bob with probability $p = \kappa_b \delta t$ and q is the probability of photon loss in cavity Alice with probability $q = \kappa_a \delta t$. The Kraus operator \hat{E}_0 represents no-jump backaction and can be written in the form of a dephasing channel as

$$\hat{E}_0 = \frac{\sqrt{1-p}}{2} (\mathbb{1} + Z_L) + \frac{\sqrt{1-q}}{2} (\mathbb{1} - Z_L)\tag{6.5}$$

Where $Z_L = |01\rangle\langle 01| - |10\rangle\langle 10|$ is the logical Z operator for the dual-rail cavity qubit and $\mathbb{1}_L = |01\rangle\langle 01| + |10\rangle\langle 10|$ is the projector onto the dual-rail codespace. Alternatively, we can rewrite this channel for any idling time t by combining the

no-jump backaction for two cavities as

$$\hat{E}_0 = e^{-\frac{1}{2}(\kappa_a a^\dagger a + \kappa_b b^\dagger b)t} \quad (6.6)$$

When projected onto the dual-rail codespace this becomes

$$\hat{E}_0 = e^{\frac{\Delta\kappa t}{4} Z_L} \quad (6.7)$$

where $\Delta\kappa = \kappa_a - \kappa_b$. The sign of $\Delta\kappa$ dictates which cavity we polarize our state populations towards.

6.5 Operating One Dual-rail Cavity Qubit

The hardware shown in Fig. 6.1 is capable of all operations needed for one dual-rail cavity qubit. This includes state preparation, readout, single qubit gates and mid-circuit erasure detection. The dual-rail Bloch sphere and control spaces are shown in Fig. 6.2

6.5.1 State preparation

Preparing a state in the logical Z basis is simply a matter of loading a single photon into one of the two cavities. This may be done via optimal control pulses [Heeres et al., 2017] or cavity-transmon sideband drives [Wallraff et al., 2007, Premaratne et al., 2017, Rosenblum et al., 2018a, Elder et al., 2020]. We can tolerate some errors better than others during state preparation. For instance, if the primary error during state preparation is unknowingly preparing $|00\rangle$ then this is an example of an inert leakage error which will be quickly detected in our mid-circuit erasure detection. However, if we use OCT pulses, we may populate higher states of the cavity and errors during OCT may leave us in harder to detect odd-parity leakage states such as

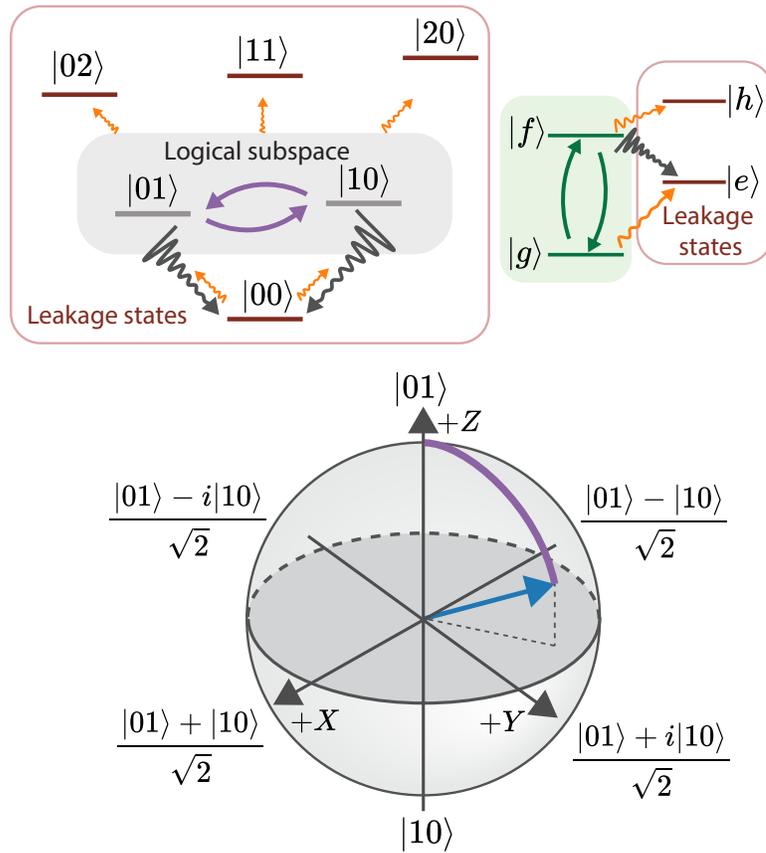


Figure 6.2: **(Top)** Codespace and control manifold for a dual-rail cavity qubit. The dual-rail qubit is operated within the $\{|01\rangle, |10\rangle\}$ subspace of the cavities. The transmon ancilla is operated in its g - f manifold (except for dual-rail logical measurements.) All other states are (detectable) leakage states. The most common leakage states are $|00\rangle$ due to cavity decay and $|e\rangle$ in the transmon due to transmon decay. **(Bottom)** The dual-rail Bloch sphere with the logical state represented by the blue vector. With only a parametric beamsplitter interaction (purple) we can perform arbitrary rotations in this Bloch sphere.

$|30\rangle$.

Regardless of our method of state preparation, we can always verify if our state preparation was successful by measuring if the cavity contains one photon (through selective π -pulses or photon number parity measurements and transmon readout). Single photon states cannot be dephased and the main error we incur from this check is a higher probability of preparing $|00\rangle$ due to undetected cavity decay during the verification measurement itself. With current state of the art, it should be doable to prepare single photon states with $> 99\%$ fidelity, with the dominant error being decay to $|00\rangle$.

6.5.2 Logical measurement with in-built erasure detection

Logical measurement of a dual-rail cavity qubit was recently demonstrated in [Chou et al., 2023]. The measurement scheme is similar to what we present in this section and this work confirms that low SPAM errors and logical assignment errors are indeed possible with these qubits.

Logical measurement in the Z basis of a dual-rail cavity qubit (or the physical qubit measurement if we are concatenating with an outer stabilizer code) is a bit different than usual logical measurement. Since we are using an erasure qubit, we can assign three possible outcomes from the measurement: ‘0’, ‘1’ or ‘erasure’. The type of erasure detection we do here is ‘end-of-the-line’ (EOTL) erasure detection, and we do not worry about whether our erasure detection dephases the states in the dual-rail codespace, since we are performing a logical measurement in the Z basis which dephases the dual-rail qubit anyways.

We walk through a scheme whereby repeated photon number measurements on each cavity in the dual-rail qubit are used to perform a logical measurement with in-built erasure detection. The circuit for this measurement is shown in Fig. 6.3 along with the interpretation of the assigned measurement outcomes.

input state	input probability	measurement outcome		
		$ 01\rangle$	$ 10\rangle$	erasure
$ 01\rangle$	$\frac{1-p}{2}$	✓	✗	\mathcal{E}
$ 10\rangle$	$\frac{1-p}{2}$	✗	✓	\mathcal{E}
$ 00\rangle$	p	✗	✗	✓

✓ - correct outcome

\mathcal{E} - incorrect outcome causes erasure

✗ - incorrect outcome causes Pauli error

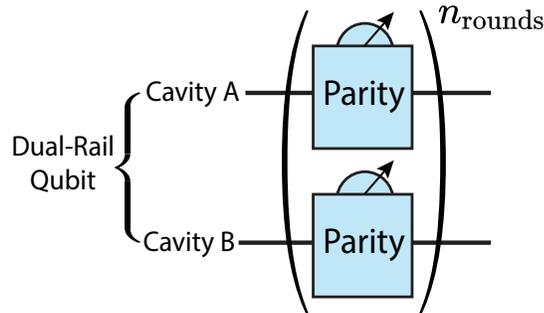


Figure 6.3: The table shows the three most likely input states $|01\rangle$, $|10\rangle$ and $|00\rangle$ together with their input probability, where p is the probability of a cavity decay event prior to the measurement, leaving the system in $|00\rangle$. We set $p = 1\%$ which is the erasure probability due to photon loss in the cavities after approximately ten two-qubit gates. Depending on the input state, the outcome of the measurement is either correctly assigned to the actual input state (green ticks), incorrectly flagged as an erasure (orange), or incorrectly assigned to the logical states (red cross) resulting in a Pauli error. The circuit shows one way of performing this measurement with n_{rounds} of repeated parity measurements on each cavity.

	Declared Outcome		
	$ 0\rangle_L$	$ 1\rangle_L$	Erasure
1R	(e,g)	(g,e)	(g,g),(e,e)
2R	(e,g,-,-)	(g,e,-,-)	(g,g,g,g),(e,e,e,e)
	(g,g,e,g),(e,e,e,g)	(g,g,g,e),(e,e,g,e)	(g,g,e,e),(e,e,g,g)
2R*	(e,g,e,g)	(g,e,g,e)	all other strings
3R	(\tilde{e}, \tilde{g})	(\tilde{g}, \tilde{e})	$(\tilde{g}, \tilde{g}), (\tilde{e}, \tilde{e})$

Table 6.2: Outcome binning for different logical measurement strategies. 1R — 1-round measurement, 2R—2-round measurement favoring less erasure, 2R*—2-round measurement (strict), 3R—3-round measurement with majority voting. Outcomes are listed chronologically left to right with two outcomes per measurement round. ‘-’ denotes the value can be either g or e, $\tilde{g} = \{(g, g, g), (g, g, e), (g, e, g), (e, g, g), \}$, the majority vote in favor of g (for measurements on the same cavity). $\tilde{e} = \{(e, e, e), (e, e, g), (e, g, e), (g, e, e), \}$, majority vote in favor of e .

One round of measurement constitutes a parity measurement on both cavities. This can be achieved with the hardware in Fig. 6.1 by using SWAP operations in between subsequent parity measurements. We also do not need to use the g - f manifold of the transmon in this measurement. The important feature is that the mapping $|0g\rangle \rightarrow |0g\rangle, |1g\rangle \rightarrow |1e\rangle$ is realizable with the measurement. We label the transmon measurement outcomes with the string format (a, b) for 1-round measurements, and $(a1, b1, a2, b2)$ for 2-round measurements, $(a1, b1, a2, b2, a3, b3)$ for 3-round measurements etc.. This is an example of a simple measurement decoder, in which we take the raw string of measured transmon outcomes and assign one of three outcomes based on a lookup table as shown in Tab. 6.2. By changing the number of rounds and our decoding ‘strategy’ we can tune our measurement to trade-off lower logical assignment error rates for higher probabilities of declaring erasure and vice versa.

A single round of measurement is already expected to be sufficient to reach the target error hierarchy of 1% erasure probability and 0.01% Pauli error per operation. The dominant cause of declaring erasure is not expected to be photon loss, but rather transmon errors and readout errors. These errors can be suppressed exponentially

by increasing n_{rounds} , which works until the measurement takes so long that cavity decay $|00\rangle$ eventually sets the probability of declaring erasure. In other words, we expect a higher false erasure assignment error with the single round strategy but a shorter duration logical measurement. Transmon readout fidelity can also be boosted by means of a ‘shelving’ pulse [Elder et al., 2020] that transfers population from $|e\rangle$ to $|f\rangle$ prior to readout. When performing a string of repeated transmon readout operations we can either reset the transmon to $|g\rangle$ between each operation or not. Without reset, we need to update the lookup table for the measurement decoder to assign alternating strings of outcomes to the presence of a photon in the cavity [Sun et al., 2014].

6.5.3 Simulating logical measurement

Modelling transmon readout

Logical measurement is one aspect of the proposal we can readily simulate with QuTiP master equation simulations. For this, we need a model for errors that occur during transmon readout. We consider decay of the transmon during the readout and readout misassignment due to the finite signal-to-noise ratio (SNR) of the readout signal. For numerical simulations of logical measurement, we model transmon readout in the g - e manifold as follows: First we calculate the (unnormalized) density matrices for the cavities, ρ_g^c and ρ_e^c that would result from perfect transmon measurement. We obtain these from

$$\begin{aligned}\rho_g &= |g\rangle \langle g| \rho |g\rangle \langle g| \\ \rho_e &= |e\rangle \langle e| \rho |e\rangle \langle e|,\end{aligned}$$

where ρ is the combined state of the transmon and cavities before the measurement. We obtain ρ_g^c (ρ_e^c) by tracing over the ancilla state in ρ_g (ρ_e). Given a particular

transmon measurement outcome, $M = \{g, e\}$, we model the state of the system to be

$$\rho_{M=g} = \left((1 - P_o)\rho_g^c + \frac{P_d}{2}\rho_e^c \right) \otimes |g\rangle\langle g| + P_o\rho_e^c \otimes |e\rangle\langle e| \quad (6.8)$$

$$\rho_{M=e} = (1 - P_o - P_d)\rho_e^c \otimes |e\rangle\langle e| + \left(\frac{P_d}{2}\rho_e^c + P_o\rho_g^c \right) \otimes |g\rangle\langle g|. \quad (6.9)$$

We set $P_d = 0.01$ to model a 1% chance of transmon decay at a random time during the readout. When this happens, half the time we will record the transmon in $|g\rangle$ and half the time we record $|e\rangle$. The overlap error $P_o = 10^{-4}$ is the probability of misassigning $|g\rangle$ as $|e\rangle$ and vice versa due to noise in our readout signal and is representative of current state-of-the-art transmon readout [Walter et al., 2017]. Additionally, cavity decay during the readout is modelled by including a 1 μ s idle time for a typical readout duration. This model can easily be extended to include the $|f\rangle$ level but is not necessary for logical readout simulations. Another advantage when using a three-level transmon for the erasure check and entangling gate is that transmon decay from $|f\rangle$ to $|e\rangle$ during transmon readout will be flagged as erasure whereas two decay errors are necessary to cause a Pauli error.

Master equation simulations

Here we describe how the QuTiP Lindblad master equation simulations were performed to estimate logical measurement fidelity for a dual-rail cavity qubit under realistic hardware coherence times listed in Tab. 6.3. For each round of measurement we first prepare the cavities in either $|01\rangle$, $|10\rangle$ or $|00\rangle$ and the transmon in $|g\rangle$. We then perform the standard parity mapping sequence, $\pi/2$ — wait time π/χ — $\pi/2$, in order to map parity information to the state of the transmon. We then simulate imperfect transmon measurement according to the readout model. For $M = e$ outcomes, we perform a perfect transmon π -pulse to reset the transmon. We then apply

Collapse operator	Coherence time
$ g\rangle\langle e $	100 μ s
$ e\rangle\langle e $	100 μ s
a, b	1 ms
a^\dagger, b^\dagger	100 ms
$ e\rangle\langle g $	10 ms

Table 6.3: Collapse operators and corresponding decoherence times used in Linblad master equation simulations for logical measurement. In all simulations, $\chi_{ge}/2\pi = -2$ MHz. Cavity dephasing is not included since this does not affect logical readout.

a perfect SWAP between the cavity modes and repeat the whole parity measurement sequence once more to obtain another transmon measurement outcome. The measurement round is concluded by another perfect SWAP operation to return the states to their initial cavities.

In one round of measurement we may obtain one of four outcomes (g,g), (g,e), (e,g) or (e,e). We store the probability and density matrix for each of these outcomes. Every time we increase the number of measurement rounds, the total number of outcomes and stored density matrices grows by a factor of four. We group all possible outcome strings into three bins: declare $|0\rangle_L$, declare $|1\rangle_L$ or declare erasure. By summing the probabilities of each of the outcomes in a given bin, we calculate the probability of declaring a logical outcome for a given input state.

The results of this simulated logical measurement, for the four decoding strategies in Tab. 6.2 are shown in Fig. 6.4. Even with a single round of measurement and imperfect readout, we should be able to achieve $\sim 10^{-4}$ logical assignment error and $\sim 1\%$ erasure errors, even though most of these erasures are false erasure assignments due to transmon and readout errors. This is why 3-round majority voting decreases both the additional erasure error and the Pauli error due to logical misassignment errors. Although we incur a factor of three increase in the probability of a cavity decay event

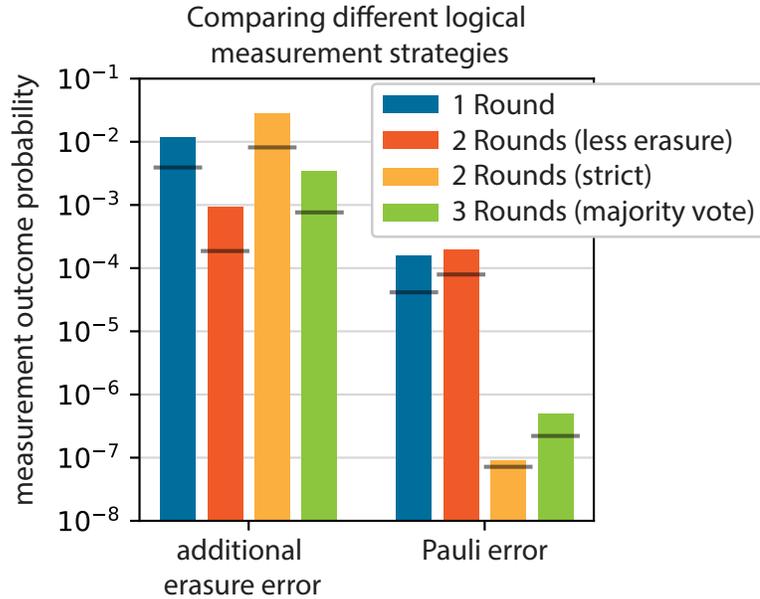


Figure 6.4: Comparison of 4 possible dual-rail logical measurement schemes listed in Tab. 6.2 for $p = 0.01$. The horizontal grey bars indicate the logical measurement performance if transmon readout were perfect. By changing our strategy, we can minimize additional erasure errors from the measurement or Pauli error. By majority voting, we can in principle reach exceedingly low Pauli rates. The additional erasure rate is the fraction of events we declare erasure, due to the measurement process itself either by cavity decay or readout errors. (The input states begin in the dual-rail codespace.)

due to a longer measurement process, this is greatly outweighed by the suppression of transmon and readout errors in the repeated rounds.

6.5.4 Single qubit gates

Single-qubit gates in the dual-rail subspace are realized by a parametric beamsplitter interactions using a dedicated coupling element [Lu et al., 2023, Chapman et al., 2022]. The combination of beamsplitter strength and duration, which are controlled by microwave pump signals in an experiment, define the polar angle of the rotation, while arbitrary rotations around the Z -axis can be implemented in software by changing the phase of the pump(s). This parameterizable control is in many ways analogous to typical use of RF signals for transmon single qubit gates via Rabi drives.

A SWAP gate between the cavities is equivalent to a π -pulse and a 50-50 beamsplitter corresponds to a $\pi/2$ -pulse. Changing the pump phase changes the equatorial axis we rotate around on the Bloch sphere. Techniques for suppressing control errors, such as composite pulses, can be directly transferred to our beamsplitter control [Brown et al., 2004, Alway and Jones, 2007]. Moreover, we can also make use of dynamical decoupling [Bylander et al., 2011, Yan et al., 2013] to further enhance the dephasing time of our dual-rail qubit and mitigate cross-talk between neighbouring dual-rail qubits.

High speed beamsplitters [Chapman et al., 2022] and randomized benchmarking [Lu et al., 2023] of single qubit gates in the dual-rail subspace have recently been experimentally demonstrated that approach 99.98% fidelity when combined with end-of-the-line erasure detection, in times of $\sim 50-100$ ns per gate. This is on par with single qubit gates in transmons [Chow et al., 2010, Sung et al., 2021]. Unlike a transmon qubit, there is no inherent speed limit set by the finite anharmonicity of the energy spectrum [Motzoi et al., 2009, Gambetta et al., 2011], which would increase the leakage out of the computational space as we decrease gate times. It is also observed that the beamsplitter interaction does not induce any additional leakage errors out of the dual-rail subspace, aside from the expected (and detectable) decay to the $|00\rangle$ state set by the bare cavity single-photon decay rates.

Using a parametric beamsplitter coupler allows our cavity modes to have large detunings, with vanishingly small swapping rates between the cavities in the absence of the pumps. Moreover, any static coupling between the cavities (within or between dual-rail qubits) simply leads to a renormalization of the cavity frequencies via mode hybridization, and these hybridized modes are the modes we use in any experiment. The ease of arbitrary single-qubit rotations in the dual-rail encoding is a huge advan-

tage compared to other bosonic encodings.

6.5.5 Mid-circuit erasure detection

Mid-circuit erasure detection refers to error detection measurements that occur in the middle of a gate sequence, e.g., during a stabilizer measurement circuit. Unlike EOTL erasure detection, mid-circuit erasure detection must preserve coherence in the dual-rail subspace in the event that no error has occurred, making their implementation more demanding. Any backaction from the measurement on the dual-rail subspace will contribute to the Pauli error rate.

When this measurement is done to detect leakage errors, we call this erasure conversion. We can now treat the leakage error as if it were an erasure channel and we can correct for the erasure error using the outer stabilizer code. How often should we perform mid-circuit erasure detection? The error model studied by [Wu et al., 2022] considers mid-circuit erasure detection sandwiched between every two-qubit entangling gate. This means we only need to erase two qubits (one data and one ancilla) when an erasure is declared on one of the qubits. One thing we also worry about are the errors introduced by the mid-circuit measurement itself. If we measure this frequently then most of the erasure and Pauli errors in the system may be induced by the measurement itself, rather than by the errors during the gates we wished to address.

We discuss two proposed methods for implementing the mid-circuit measurement which checks for cavity decay (and heating) error, primarily in the form of decay to the $|00\rangle$ leakage state.

Joint-parity measurements

This approach is based on the observation that the dual-rail codewords have odd total photon number whilst the four most common leakage states $\{|00\rangle, |11\rangle, |20\rangle, |02\rangle\}$

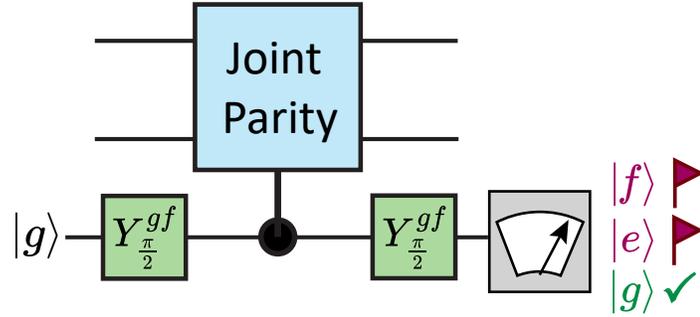


Figure 6.5: The circuit for a joint-parity measurement, now used for mid-circuit erasure detection in the dual-rail code. The phase of the last $\pi/2$ pulse is no longer flipped, such that we now map odd joint parity to $|g\rangle$

all have even total photon number. Thus a QND joint-photon number parity measurement can be used to distinguish between this set of leakage state and the codespace with a single measurement, ideally without affecting the coherence of states within the dual-rail codespace.

We used the same approach in Sec. 5.2 of Chapter 5 to realize a QND and hardware fault tolerant ZZ measurement for the 0-1 encoding. For this, we engineer the controlled joint-parity unitary by combining the dispersive interaction with a parametric beamsplitter interaction applied for the appropriate duration of time. (First row of Tab. 5.4.)

We use the same circuit, now in a different context, to realize mid-circuit erasure detection as shown in Fig. 6.5 In the absence of all transmon ancilla errors, if we detect $|g\rangle$, then we are still in the dual-rail codespace. If we detect $|f\rangle$, we are in $|00\rangle$ or $|11\rangle$ and we declare an erasure error (and perform erasure conversion).

It is important that the joint-parity measurement conforms to the error hierarchy when we introduce transmon errors as well. As we saw from its Hardware Fault Tolerance (HFT) properties, first-order transmon decay and dephasing events can be detected and contribute to the overall erasure rate, whilst double-jump errors involving the transmon are needed for these errors to pass through undetected by our measurement and contribute to the Pauli and/or leakage error rates. A less than

ideal consequence of using transmon ancillas this way is that the erasure rate tends to be set by the rate of first-order transmon errors, rather than the rate of cavity decay. Similarly, the Pauli error rate will be set by double-jump errors involving the transmon rather than the vanishingly small double-jump errors involving the cavity. This is the main reason why the operational error hierarchy is expected to be much worse than the idling error hierarchy.

Nonetheless, given that transmon jump errors are expected to occur with probability $\sim 1\%$ per gate and double jump errors with probability $\sim 0.01\%$ per gate, this still places us far below threshold for both erasure and Pauli error rates. These scalings are only possible because of the HFT properties specific to our control joint-parity implementation.

In the specific case of a joint-parity measurement, transmon decay events to first-order lead us to measure $|e\rangle$ at the end of the sequence. With high likelihood, the pair of cavities that make up the dual-rail qubit still contain a single photon¹, but the unknown time of the decay event constitutes complete dephasing of the dual-rail qubit. We thus do not need to reset the dual-rail qubit back into the codespace, since it should already be in some unknown state within the codespace. So when we measure $|e\rangle$ we should reset the transmon ancilla to $|g\rangle$ and remeasure stabilizers within the outer code to correct this erasure error. This form of erasure error is different from erasure conversion. The $|e\rangle$ state of the transmon ancilla is being used as a flag state to tell us when a dephasing error (a.k.a. a Z error) occurs on the dual-rail qubit due to transmon decay.

Transmon dephasing events cause us to readout $|f\rangle$ instead of $|g\rangle$ and vice-versa; they invert the measurement outcomes. Since transmon hardware errors are expected to occur roughly ten times more frequently than cavity decay errors, most of the times we measure $|f\rangle$ will not be due to actual leakage events in the dual-rail cavities but

1. Both the parametric beamsplitter and dispersive interaction conserve total photon number.

rather dephasing events in the transmon ancilla, which contribute to false erasure assignments and make up most of the fraction of events we declare as erasures. We can mitigate this error by performing two or three joint-parity measurements in a row and majority voting to decide if we declare an erasure outcome or not. These same dephasing events in the transmon ancilla can cause us to miss leakage errors. This is a double-jump error (cavity leakage + transmon dephasing) and so is expected to be rare at the 10^{-5} level per gate.

Y-mon

An alternate scheme for mid-circuit erasure detection can be implemented with the ‘Y-mon’ hardware in Fig. 5.1. This requires a transmon dispersively coupled to both cavities in the dual-rail qubit and the ability to readout this transmon. The idea is to play a selective π -pulse at the transmon ω_{ge} frequency. This technique shares more similarity with the fluorescence measurements used for mid-circuit measurements in neutral atoms — we only see a signal if the qubit is in a leakage state, in this case we record $|e\rangle$ only if the cavities are in the state $|00\rangle$. Unlike the joint-parity measurement, this measurement only excites the transmon ancilla out of the ground state if the cavity is in $|00\rangle$ and as such, we should be robust to transmon errors if the cavities are still within the dual-rail codespace. If we are in a leakage state, then we are now susceptible to transmon errors such as decay and dephasing, but these cannot further corrupt the cavity state since we are already in $|00\rangle$. At worst, these can result in leakage detection errors and are unlikely since they are all double-jump errors. Also, for this reason we can get away with using the g - e manifold of the transmon rather than the g - f manifold.

The drawback of this approach is an increase in hardware requirements: each dual-rail cavity qubit needs an ancilla transmon dispersively coupled to both cavities, as well as the ability to perform parametric beamsplitter interactions. This could

in principle be achieved with the same non-linear element functioning as the coupler and the ancilla. (The important thing is the element needs a dispersive shift to both cavities, it does not have to be a transmon). However, hardware for an additional readout would certainly be required. A dispersively coupled element also introduces a source of cavity dephasing via unwanted heating, although this in principle can also be detected. If the element is χ -matched to both cavities, then heating backaction only causes common dephasing in the dual-rail subspace, and does not present any source of error other than false erasure assignments, which are a relatively minor error.

6.6 Entangling Gates

Likely the hardest operation to implement is the two-qubit entangling gate. Fortunately, in a dual-rail qubit, we do not need all four cavity modes to interact to perform a two-qubit operation. It suffices for a cavity in one dual-rail qubit to interact with another cavity in the other dual-rail qubit. If we label the cavity modes as (\hat{a}_1, \hat{b}_1) for dual-rail qubit 1 and (\hat{a}_2, \hat{b}_2) for dual-rail qubit 2 then only modes, e.g., \hat{b}_1 and \hat{a}_2 need to directly interact.

One proposal for dual-rail entangling gates is to realize $ZZ(\theta)$ gates. The procedure is the exact same as in Sec. 5.3 of Chapter 5 for the 0-1 encoding for the cQED hardware shown in Fig. 6.6. By treating modes \hat{b}_1 and \hat{a}_2 as if they were in the 0-1 encoding and performing the corresponding $ZZ(\theta)$, we end up realizing the same $ZZ(\theta)$ for the two dual-rail qubits, even though \hat{a}_1 and \hat{b}_2 never interacted! This implementation is particularly hardware efficient with only one additional beamsplitter coupler required to perform the gate. A more detailed circuit showing the control joint-parity unitaries in an exponentiation gadget is shown in Fig. 6.7. In a stabilizer measurement code, we would almost exclusively use the $ZZ(\theta)$ gate at $\theta = \pi/2$, for

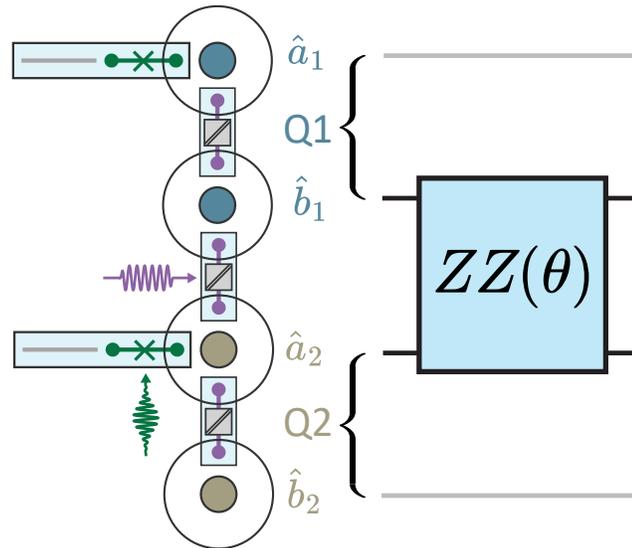


Figure 6.6: The two-qubit entangling gate can be realized by a $ZZ(\theta)$ gate performed between the central rails only, for which only an additional beamsplitter coupler is required.

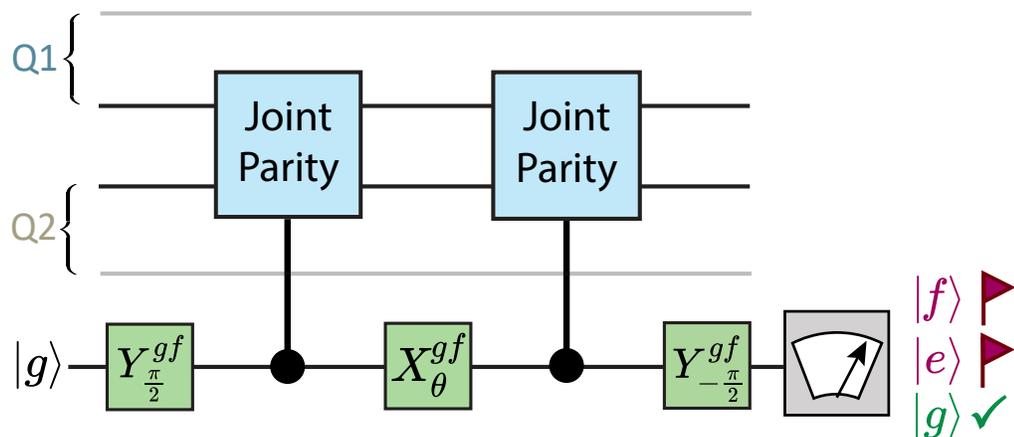


Figure 6.7: With control joint-parity unitaries that now act *between* rails of two dual-rail qubits, we can do the $ZZ(\theta)$ entangling gate for two dual-rail qubits. A successful gate is signalled by measuring the transmon in $|g\rangle$ after the gate.

which this gate is locally equivalent to a CZ gate. Once again, owing to our HFT gate construction, we have desirable error detection properties. Transmon errors during the gate result in uncorrectable unknown backaction on the cavities and as such, these errors can be detected and converted to erasure errors as well. This erasure detection is ‘built-in’ to the measurement. We just need to measure the transmon state after the measurement.

Transmon phase flips result in us measuring $|f\rangle$ after the gate. This is unrecoverable since we do not know whether $\mathbb{1}_L \otimes \mathbb{1}_L$ or $Z_L \otimes Z_L$ has been applied to the two dual-rail qubits. (See Sec. 5.3.1 of Chapter 5). Once again, a single transmon decay event during the measurement leaves us in $|e\rangle$ however this time, we may have evolved outside the two-qubit dual-rail codespace and need a more involved reset protocol to reset both dual-rail qubits back into the codespace. We can see why in Fig. 6.8, which shows how photon number population evolves during the control joint-parity unitary. With equal probabilities, the four cavity modes $(\hat{a}_1, \hat{b}_1, \hat{a}_2, \hat{b}_2)$ may be in the states $\{|1001\rangle, |1010\rangle, |1010\rangle, |0110\rangle\}$. If we start in any state other than $|1001\rangle$, then transmon decay during the $ZZ(\theta)$ can leave us in the leakage states $\{|1100\rangle, |0011\rangle, |0200\rangle, |0020\rangle\}$ which requires a more involved reset protocol. Like the joint-parity measurement, the errors in this gate are expected to be transmon-limited in the same way but still well below threshold.

6.6.1 With in-built erasure detection

A modified version of the $ZZ(\theta)$ dual-rail gate sequence shown in Fig. 6.9 has the capability to map the dominant cavity leakage states to the $|f\rangle$ state at the end of the gate. In particular, we will detect the leakage states where one of the dual-rail cavity qubits has leaked to $|00\rangle$ and the other remains in the dual-rail codespace. We would still measure $|g\rangle$ if both dual-rail qubits had lost their photons. This has the potential to remove the need for dedicated mid-circuit erasure detection operations.

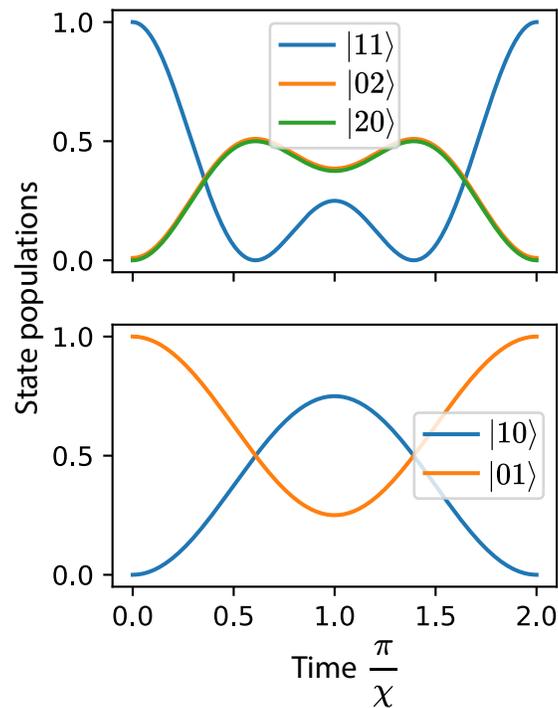


Figure 6.8: State populations of two cavities during the control joint-parity unitary. **(Top)** Population dynamics if the cavities are prepared in $|11\rangle$. **(Bottom)** Population dynamics if the cavities are prepared in $|01\rangle$. The case where we prepare $|10\rangle$ is symmetric and not shown, if we prepare $|00\rangle$ we remain in this state throughout the operation.

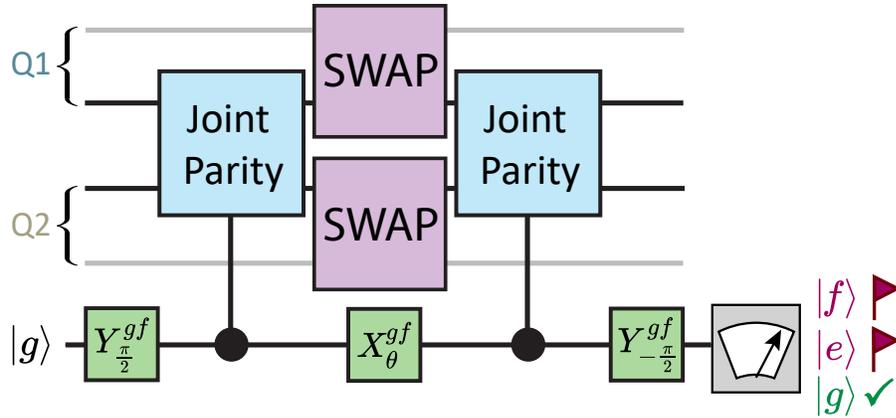


Figure 6.9: Variation of the $ZZ(\theta)$ gate construction to detect cavity decay events that occurred prior to the gate.

The idea is to perform the first control joint-parity unitary on cavity modes (\hat{b}_1, \hat{a}_2) and the second on (\hat{a}_1, \hat{b}_2) . Without additional hardware, this can be realized by performing swaps within each dual-rail qubit halfway through the gate sequence. This construction has the additional property that if the total number of photons amongst all four cavities is odd, we will also readout $|f\rangle$ at the end of the sequence. Using the hardware in Fig. 6.6, these SWAP operations must be unconditional on the state of the transmon dispersively coupled to mode \hat{a}_2 , which is possible to achieve as was discussed in Sec. 5.4 of Chapter 5, but takes a longer duration to cancel out the unwanted effects of the dispersive interaction. Whilst we can detect cavity decay events prior to the gate, they will be detected on average only half the time if they occur during the $ZZ(\theta)$, by dephasing the transmon ancilla (and randomizing the measured transmon state between $|g\rangle$ and $|f\rangle$). However, these decay events will be then be detected with high probability in the next entangling gate, since the total number of photons in the four cavities will likely still be odd when we start the next entangling gate.

6.7 Operational Error Hierarchy

Having gone through the required operations for dual-rail cavity qubits we can now estimate what the error hierarchy would be during operations, shown in Tab. 6.4. Since operation times are set by $\sim 1/\chi$, we often take $T_{\text{gate}} = 1 \mu\text{s}$.

Error process	Erasure errors		Pauli errors	
Single-photon loss	$\bar{\kappa}T_{\text{gate}}$	10^{-3}	-	-
no-jump backaction	-	-	$(\frac{1}{4}\Delta\kappa T_{\text{gate}})^2$	10^{-7}
cavity dephasing	-	-	$\gamma_{\varphi}T_{\text{gate}}$	10^{-4}
ancilla decay	$\Gamma_1^{ef}T_{\text{gate}}$	10^{-2}	$\Gamma_1^{ef}\Gamma_1^{ge}T_{\text{gate}}^2$	10^{-4}
undetected ancilla decay	-	-	$\eta_{ge}\Gamma_{1,ef}T_{\text{gate}}$	10^{-4}
ancilla dephasing	$\Gamma_{\varphi}^{gf}T_{\text{gate}}$	10^{-2}	$(\Gamma_{\varphi}^{gf}T_{\text{gate}})^2$	10^{-4}
photon loss + ancilla dephasing	-	-	$\bar{\kappa}\Gamma_{\varphi}^{gf}T_{\text{gate}}^2$	10^{-5}
undetected ancilla dephasing	-	-	$\eta_{gf}\Gamma_{\varphi}^{gf}T_{\text{gate}}$	10^{-6}
measurement infidelity	η_{gf}	10^{-4}	-	-
photon loss + msmt. infid.	-	-	$\eta_{gf}\bar{\kappa}T_{\text{gate}}$	10^{-7}

Table 6.4: Operational error hierarchy for dual-rail cavity qubits.

Erasure errors are caused both by single-photon loss events in the cavities happening at an average rate $\bar{\kappa}$ in the cavity system, but more frequently due to decay and dephasing events in the transmon ancilla and the finite measurement infidelity η_{ij} for distinguishing the ancilla states i and j . In general, the measurement infidelities to distinguish different pairs of states are not identical [Elder et al., 2020], which is why we consider two separate infidelities $\eta_{ge} = 10^{-2}$ and $\eta_{gf} = 10^{-4}$. A difference in the single-photon decay rates $\Delta\kappa = \kappa_b - \kappa_a$ causes a no-jump backaction slowly polarizing the dual-rail towards the cavity with the longer lifetime contributes only to second-order to the Pauli error rate. The pure dephasing of the cavities at rate $\gamma_{\varphi} = \gamma_{\varphi,a} + \gamma_{\varphi,b}$ is included to represent some upper limit to the intrinsic cavity dephasing time. For the calculation of the error probabilities we assume $\kappa_a = (0.5 \text{ ms})^{-1}$, $\kappa_b = (1 \text{ ms})^{-1}$, $\gamma_{\varphi,a} = (20 \text{ ms})^{-1}$, $\gamma_{\varphi,b} = (20 \text{ ms})^{-1}$, $\Gamma_1^{ge} = (100 \mu\text{s})^{-1}$, $\Gamma_1^{ef} = (100 \mu\text{s})^{-1}$,

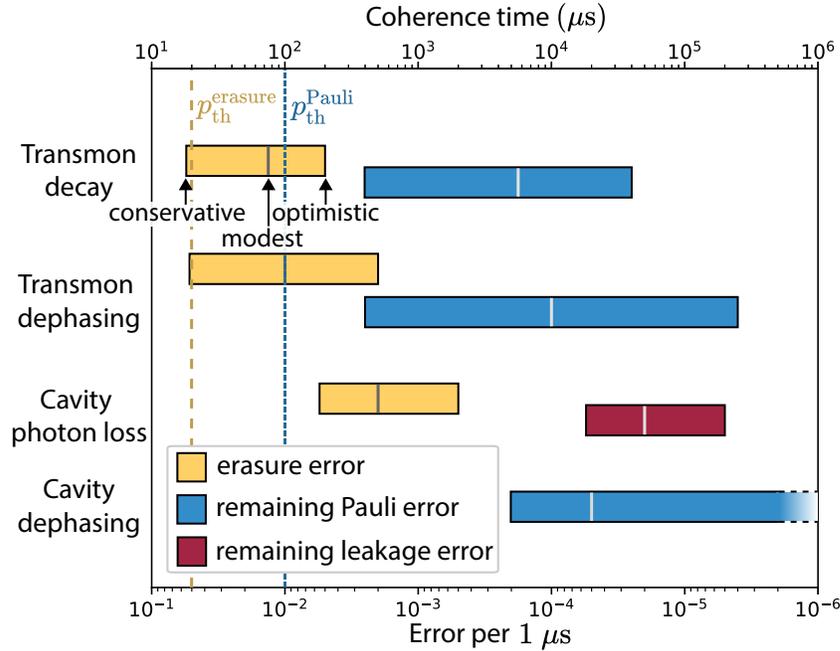


Figure 6.10: Graphical depiction of the operational error hierarchy. For modest coherence times and above, the erasure errors and Pauli errors are well below their respective thresholds.

$$\Gamma_{\varphi}^{gf} = (100 \mu\text{s})^{-1}.$$

We can represent the operational error hierarchy visually as shown in Fig. 6.10. Here, we only show the dominant sources of erasure, Pauli and leakage errors which are cavity photon loss and transmon errors for the erasures, transmon double-jump errors and cavity dephasing for the Paulis and undetected cavity photon loss for the leakage errors. On the plot we also mark the thresholds for our error model for erasure and Pauli errors. For each error mechanism we also include conservative, modest and optimistic error rates and a 99% conversion efficiency. These choices of error rates are shown in Tab. 6.5. By plotting these three coherence times, we can see that error-detected physical qubits such as ours greatly benefit from any improvements in coherence time or equivalently shortening of the gate times. If we were to double our coherence times, the erasure errors halve and the Pauli error rates are reduced by a factor of four. Ultimately the logical error rate of the outer code would drop

Coherence time	Conservative	Modest	Optimistic
T_1^{ef}	20 μ s	75 μ s	200 μ s
T_φ^{gf}	50 μ s	100 μ s	500 μ s
T_1^{cav}	200 μ s	500 μ s	2 ms
T_φ^{cav}	5 ms	20 ms	??

Table 6.5: Conservative, modest and optimistic coherence times for the four dominant types of error when operating dual-rail cavity qubits

by approximately $(\frac{1}{2})^d$. For comparison, without error detection the Pauli error rate would half and the logical error rate would drop by $(\frac{1}{2})^{\frac{d}{2}}$. The importance of the scaling in the exponent cannot be overstated, at large code distances we can benefit from increased coherence times exponentially faster with error-detected qubits vs non-error detected qubits.

6.8 Scaling Up

For erasure qubits to be useful, we must concatenate them with an outer stabilizer code. The surface code is a natural choice as we saw in Sec. 3.2.3 of Chapter 3 because it is adept at handling both erasures and Pauli errors simultaneously and has been shown to have high thresholds to both Pauli and erasure errors (compared with other large-scale stabilizer codes such as color codes.) One hypothetical cQED hardware layout for tiling dual-rail cavity qubits is shown in Fig. 6.11. With this layout, the transmons are arranged in such a way that we can perform joint-parity measurements on each dual-rail qubit as well as $ZZ(\theta)$ gates between each dual-rail ancilla qubit and its four surrounding data qubits. This is just one possible layout. One modification is to rearrange the transmon ancilla qubits such that only the dual-rail ancilla qubits are coupled to transmons in a bid to improve the idling errors in the dual-rail data qubits, although mid-circuit erasure detection via joint-parity measurements would then require additional SWAP operations.

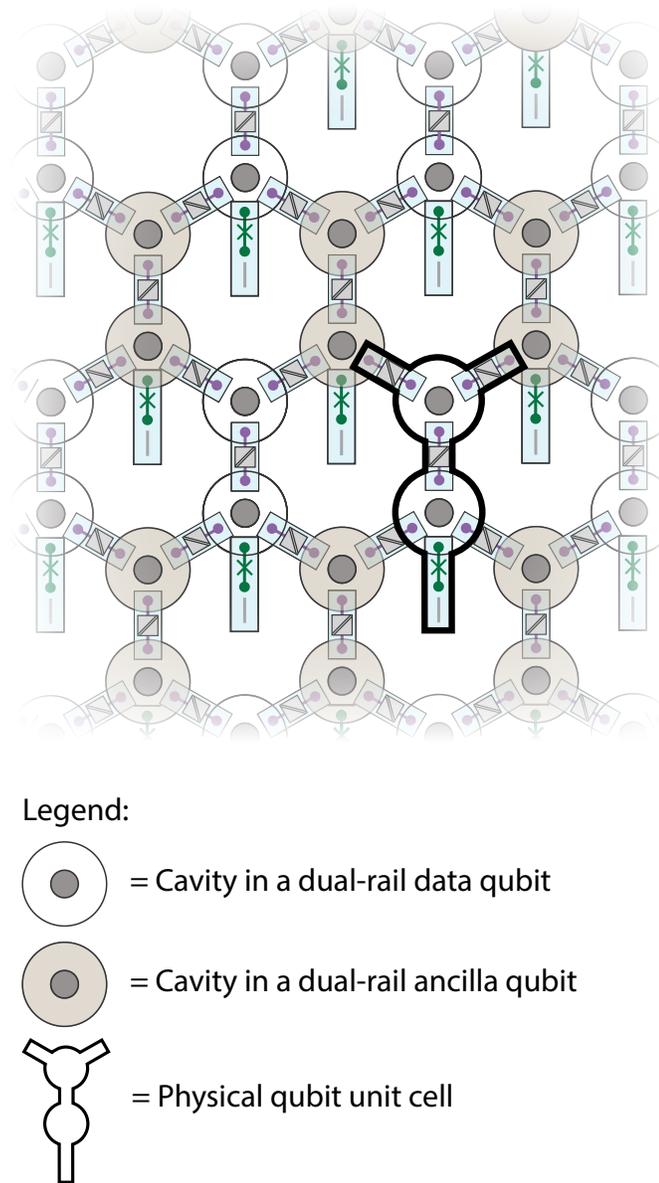


Figure 6.11: Example cQED hardware used for tiling dual-rail qubits together. Although the cavities are connected in a hexagonal lattice, this represents a square lattice of data qubits with ancilla qubits at the centre of each plaquette as needed for a surface code. In this case, the physical qubit unit cell is identical for ancilla and data qubits.

6.8.1 The physical qubit unit cell

Many different qubit platforms eventually aim to implement large-scale quantum error correction using an outer stabilizer code such as the surface code (or the cluster state equivalent in linear optics). Comparing across different platforms is non-trivial but one thing we can compare is the physical unit cell. This is the set of hardware we must ‘tile’ for large scale error correction. The unit cell includes not just a physical qubit, but also the required connections needed to tile it. A possible unit cell is outlined in Fig. 6.11 and consists of a single transmon (with readout), two cavity modes and a beamsplitter coupler for the dual-rail qubit, and two additional beamsplitter couplers to connect to neighbouring dual-rail qubits.

We can compare this to the unit cell of a transmon-based surface code, such as the one presented in [Acharya et al., 2023], where the unit cell consists of a single transmon and two couplers needed for the two-qubit gates. In this sense, the number of nonlinear elements per unit cell is similar in both cases (four vs three) but dual-rail cavity qubits require the two additional cavity modes.

Although the unit cell may be more complicated, if it can achieve a much lower physical error rate, then the total number of hardware components per logical qubit can be less. We can do some rough accounting by taking 10^{-14} to be a target logical error rate as taken in [Gidney and Ekerå, 2021]. Suppose for the sake of argument a transmon surface code can reach 0.1% error rates in all operations, then we would require $d = 27$ or 1,457 physical qubits per logical qubit, or 4,371 nonlinear elements.

How many nonlinear elements and cavity modes would be needed for a dual-rail cavity qubit? If we assume 1% erasure errors and 0.01% Pauli errors and the error scalings in Fig. 3.5 then we would require $d = 19$ or 721 physical qubits per logical qubit. This would then require a total of 2,884 nonlinear elements and 1,442 cavity modes for a total of 4,326 — coincidentally similar to our estimated number of modes for a transmon based approach.

All of this is to say that the hardware overhead in terms of the number of modes (not the physical footprint of the modes!) on the order of a few-thousand for current estimated dual-rail error rates, and this is not dissimilar from the number of modes needed in any platform targeting a 0.1% physical error rate. What is more important for us is whether we can reliably achieve the physical error rates required for below threshold operation, which is the main benefit of going down the erasure qubit route (rather than hardware efficiency).

6.8.2 Comparison to transmon-based approaches

From our analysis of the dual-rail unit cell, it is apparent that each dual-rail qubit will likely require one transmon control ancilla on average. Transmon errors are also expected to be the dominant errors in the system. We can then ask the following question:

Given that we have $2d^2 - 1$ transmons, is it better to correct their errors with an all-transmon-based surface code? or a dual-rail based surface code? Both implementations will have the same code distance, d . Answering this question depends on many specifics of the hardware implementation but we can make the following high-level observations.

We can define the ‘clock speed’ of the surface code to be the time to perform a complete cycle of stabilizer measurements. This is set by both the gate time and the time to measure the ancilla qubits. In a transmon-based surface code, this is usually limit by the transmon readout time which is typically longer than transmon gate times, which are set by the transmon anharmonicity. For the dual-rail surface code, the gate time is typically longer, set by the strength of the transmon-cavity dispersive interaction. The ancilla qubit measurement time is also set by transmon readout times. Overall, we expect the clock speed of a dual-rail surface code to be around $\sim 10 \mu\text{s}$ whereas the clock speed of a transmon-based surface code has been

shown to be $\sim 1 \mu\text{s}$ [Acharya et al., 2023]. As such, we expect transmon errors to be $\times 10$ more frequent in a dual-rail based surface code.

Nevertheless, there are good reasons to believe a dual-rail based surface code can out-perform a transmon-based one. This is primarily because the transmons in a dual-rail surface code are used exclusively as control ancillas - they do not need to store any qubit information as they do with a transmon-based surface code. And as such, we do not need to worry about preserving any physical qubit information when we detect errors in the transmons. In comparison to a transmon-based surface code where transmon errors can only be detected and corrected through stabilizer measurements, in a dual-rail surface code, we detect transmon errors instead by measuring them directly and using the $|e\rangle$ and $|f\rangle$ states as flag states for transmon errors that are ultimately converted to erasure errors. By using transmons not as qubits but instead as control ancillas we can evacuate the entropy due to transmon errors more efficiently. The dual-rail encoding with its extra cavities and couplers can be thought of as the scaffolding required to relegate transmons to control ancillas whilst still being able to encode and control qubits in the system.

6.8.3 Higher-order nonlinearities and cross-talk

A nice feature of the dual-rail code is that it is largely unsusceptible to unwanted higher order Hamiltonian terms such as self-Kerr, cross-Kerr and χ' . This is because there is only one photon in a dual-rail qubit. There are however a few exceptions where these terms may have a noticeable effect which involve wanted (and unwanted) interactions between neighbouring dual-rail qubits. The Cross-Kerr interaction between two adjacent cavities each in different dual-rail qubits is analagous to ZZ -crosstalk in transmon qubits. It essentially presents as an unwanted ZZ interaction that slowly implements a $ZZ(\theta)$ gate when we want the dual-rail qubits to be idle. Like ZZ -crosstalk, this effect can be mitigated with dynamical decoupling techniques.

During the $ZZ(\theta)$ gates, we are susceptible to all three of these effects since the two interacting cavities can be in the states $\{|11\rangle, |02\rangle, |20\rangle\}$ and the transmon ancilla is in a superposition. The contributions to the gate fidelity scale as $(K'T_{\text{gate}})^2$ for $K' \sim K, \chi_{ab}, \chi'$ and can be largely calibrated out by adjusting the angle θ of the middle transmon ancilla rotations and performing additional Z rotations $Z_1(\theta_1), Z_2(\theta_2)$ after the entangling gate.

Crosstalk such as ZZ crosstalk can be damaging for a stabilizer code, since it can be a source of correlated errors. Aside from small cross-Kerr couplings, we expect other sources of cross-talk to be small if we use a 3D microwave cavity architecture. This is because the cavity modes can be detuned from each other by as much as several GHz and are only coupled to each other indirectly through couplers in tunnels that also function as waveguides above cutoff. We also do not need to worry about unwanted package modes that plague 2D cQED systems on a single, large chip (The 3D cavity *is* the package mode!). We also do not expect any ‘global’ noise sources that affect all the physical qubits like has been observed for dephasing noise in neutral atoms platforms [Singh et al., 2023]. Also worth noting is that we expect our architecture with 3D cavities to be naturally robust to cosmic rays, which are known to affect multi-qubit processors [McEwen et al., 2022]. Although rare, a single cosmic ray event can corrupt multiple J.J.-based qubits on the same chip.

6.8.4 4-qubit erasure code

What is the minimum hardware we would need to realize a 4-qubit erasure code with dual-rail cavity qubits? This can be achieved with four data qubits and one ancilla qubit which can measure stabilizers on each of the four data qubits. A potential cQED hardware layout is shown in Fig. 6.12 To correct for erasures in this code we need to both be able to measure the stabilizers $ZZZZ$ and $XXXX$ and perform mid-circuit erasure detection for cavity leakage. This code can correct for a single

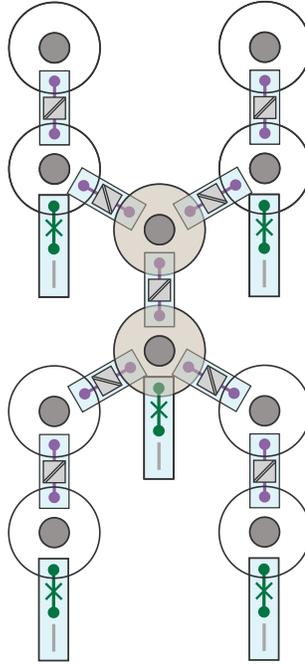


Figure 6.12: Potential cQED hardware for realizing a 4-qubit erasure code.

cavity leakage event during a round of stabilizer measurements but not cavity decay in two cavities or any dual-rail dephasing errors.

How would we operate such a code in practice? First we would use the transmons to prepare a single photon in each of the dual-rail data qubits. Then we would measure the stabilizers $ZZZZ$ and $XXXX$ using the circuits in Figs. 3.2 and 3.3 compiled for native $ZZ(\pi/2)$ gates with the central dual-rail ancilla qubits to project us into the codespace. Then, in intervals of δt we perform mid-circuit erasure detection, e.g., via joint parity measurements on each dual-rail data qubit. If we detect any of the four surrounding transmons to be in a state other than $|g\rangle$ after the joint-parity measurement we declare erasure on that qubit and reset that dual-rail cavity qubit, e.g., by Q-switching (temporarily increasing the loss rate of the cavity by swapping photons to an additional lossy mode, quickly resetting the state to vacuum) and repreparing a photon. To correct the erasure we remeasure the two stabilizers again and update the Pauli frame as necessary.

We can estimate what it would take for this code to reach breakeven as a memory compared to say the 0-1 encoding in a single cavity, the usual benchmark for memory breakeven experiments with bosonic qubits. The total number of photons is four, a factor of 8 overhead in \bar{n} . The probability of double photon loss in short time window δt is $\frac{1}{2}(4\kappa\delta t)(3\kappa\delta t) = 6(\kappa\delta t)^2$. The probability of any of the four dual-rail qubits dephasing is $4\gamma_\phi\delta t$. We reach breakeven when the logical error rate of our 4-qubit code is less than the error rate of the 0-1 code, due to single photon loss which is satisfied when

$$6(\kappa\delta t)^2 + 4\gamma_\phi\delta t < \frac{1}{2}\kappa\delta t \quad (6.10)$$

Taking $\kappa^{-1} = 1$ ms and $\gamma_\phi^{-1} = 1$ ms this requires $\delta t \approx 17$ μ s and corresponds to a logical error rate of 1% per stabilizer cycle. It is important to remember this is a back-of-the-envelope calculation that assumes perfect mid-circuit erasure detection and stabilizer measurements. Although not expected to be a practically useful encoding, the 4-qubit erasure code would serve as a useful intermediary on the way to realizing a $d = 3$ surface code as it requires the same hardware capabilities but can be tested on a smaller-scale system.

6.9 The Dual-rail Glossary

Dual-rail cavity qubit:

A dual-rail qubit where two superconducting microwave cavity modes are used to encode a qubit in the dual-rail subspace. May be two separate cavities are two modes of a multimode cavity.

Logical dual-rail qubit:

If a dual-rail qubit is the highest level of the encoding, it is the logical qubit too.

Logical state:

Any state in the dual-rail codespace.

Physical dual-rail qubit:

When concatenating with an outer stabilizer code, e.g., the surface code, dual-rail qubits make up the physical qubits.

Logical measurement (of a dual-rail qubit):

The sequence of operations that performs a logical Z measurement, which usually has end-of-the-line erasure detection built-in to the measurement. Used to determine if the dual-rail is in $|01\rangle$ or $|10\rangle$ or if it is assigned an erasure outcome. Natively measurements are in the Z basis.

Assigned outcome:

The outcome we declare from a logical measurement, e.g., ‘0’, ‘1’ or ‘erasure’.

Logical assignment error:

when we declare ‘0’ when we should have declared ‘1’ and vice versa.

Erasure detection:

The act of detecting an error on a particular qubit (or qubits), which can then

be treated as an erasure error. The qubit may have suffered from a leakage or Pauli error.

Erasure fraction:

The fraction of outcomes assigned erasure when averaging over many erasure detection measurements.

End-of-the-line (EOTL) erasure detection:

Detecting erasure in a dual-rail qubit while performing a logical Z measurement, or during the last measurement on the dual-rail qubit before reset. Does not need to preserve coherence in the dual-rail subspace!

Mid-circuit erasure detection:

Detecting errors in such a way that coherence of the dual-rail qubit is preserved if no error occurred, a property necessary for erasure conversion.

Erasure conversion:

The act of detecting a leakage error, via mid-circuit measurements, such that the leakage error channel may now be modelled as an erasure channel.

Correcting for erasure errors:

After erasure detection (either EOTL or midcircuit depending on the context), we reset the dual-rail qubit to the codespace and measure the stabilizers and decode in the outer stabilization code.

Leakage error:

Errors that take us out of the dual-rail codespace.

Decay error:

A single photon loss error in a dual-rail qubit, which usually leaves us in $|00\rangle$.

Leakage detection:

A measurement that tells us if we have leaked out of the dual-rail codespace. If measurement is part of the procedure for correcting an erasure error it is an example of erasure detection.

Leakage detection error:

Given that we have leaked out of the codespace, the probability a leakage detection measurement does not flag this leakage. (This is the ‘false negative’ rate that in part sets the total fraction of leaked qubits.)

False erasure assignment:

When we assign a state with no error as an erasure, the false positive rate associated with an erasure detection.

Readout error:

An error that occurs during the dispersive readout of a transmon, typically contributes to the false positive/negative rate.

Transmon error:

A phase-flip or decay error in a transmon.

Double-jump error:

An error that is described by two jump errors from $\{\hat{\epsilon}\}_{\text{hardware}}$. We also typically include readout misassignment as a jump error.

No-jump backaction:

A type of dephasing error that polarizes a dual-rail photon towards the less-lossy cavity mode. Occurs whenever cavity loss rates are unequal.

Dual-rail dephasing:

Dephasing within the dual-rail subspace which destroys coherence in superpositions of the codewords.

Common dephasing:

Dephasing noise experienced equally by both cavity modes, such that it does not cause dephasing on the dual-rail subspace.

Complete dephasing:

A type of cavity dephasing that is an unknown rotation in phase space, i.e., $e^{i\theta\hat{a}^\dagger\hat{a}}$ where angle θ is unknown and uniformly distributed for $\theta \in [0, 2\pi]$. This error is the backaction on the cavity when a dispersively coupled element such as an ancilla transmon undergoes a decay or heating event.

Post-selection:

Removing experimental shots from data analysis.

Nonlinear element:

Devices that contain at least one J.J., transmon (ancilla and couplers).

Intrinsic cavity dephasing:

Dephasing in a bare cavity mode in the absence of coupled non-linear elements.

Extrinsic cavity dephasing:

Dephasing in a cavity that are caused by the coupled nonlinear elements.

Beamsplitter:

A parametric beamsplitter interaction, used to perform single qubit gates. Also used in the $ZZ(\theta)$ entangling gate.

ZZ crosstalk:

ZZ crosstalk between two dual-rail qubits, due to cross-Kerr between the two neighbouring cavities (each cavity belonging to a separate dual-rail qubit). Analogous to ZZ crosstalk in, e.g., transmon qubits.

Data dual-rail qubit:

A dual-rail qubit that stores logical information within the outer stabilizer code.

Ancilla dual-rail qubit:

A dual-rail qubit used to measure stabilizers within the outer stabilizer code.

6.10 The Dual-rail FAQ

1. Do erasures errors only help with the surface code? How does it work for other codes?

Having erasures as the dominant error channel at the physical qubit level helps for *any* stabilizer code, not just the surface code. This could be a color code or Bacon-Shor code etc. The general property is that any stabilizer code that can correct for N Pauli errors can correct for $2N$ erasure errors [Grassl et al., 1997]. For codes with a threshold, the threshold tends to be higher for erasure errors than Pauli errors.

2. Don't jumps in the transmon/couplers ruin the 'intrinsic' noise bias of the cavities?

It is undoubtedly true that unwanted energy jumps in dispersively coupled modes dephase the cavity modes. And this is large angle dephasing — once a jump happens, the cavity is completely dephased (see glossary). The idea is that most jumps are detectable by regularly measuring these modes, or otherwise unlikely because the modes are cold and well-initialized in their ground state. Undetectable errors manifest from double jump errors, e.g., a transmon heating even followed quickly by a transmon decay event and are expected to be rare, contributing to the background rate of Pauli errors.

3. Why not do dual-rail with transmons?

dual-rail has been proposed in transmons [Kubica et al., 2022] and has recently been experimentally demonstrated [Levine et al., 2023]. The key difference here is that transmons do not intrinsically have a low dephasing errors and this must be engineered for dual-rail to be viable. This is possible by tuning two transmons on resonance and using the new energy eigenstates $|ge\rangle \pm |eg\rangle$ to define the codewords. The energy splitting and hence T_ϕ of the dual-rail

subspace is only quadratically sensitive to shifts in either transmon's frequency. It is worth noting that this approach to dual-rail could work for any pair of two-level systems tuned on resonance with each other.

4. **Why not do dual-rail with Fluxonium?**

Why not? We would need to operate the fluxoniums at the 'sweet spot', where the T_ϕ time is long compared to its T_1 time. High fidelity two-qubit gates could be implemented with transmon couplers (See the FTF scheme [Ding et al., 2023]). As with any erasure qubit, the main questions are: can we do good mid-circuit erasure detection? And is the dominant error during entangling gates an erasure error?

5. **The cavity modes participate in the Josephson junctions (if only weakly), do the cavities inherit dephasing?**

They inherit some dephasing, expected to be $\Gamma_\phi^{\text{cav}} \approx \left(\frac{g_c}{\Delta}\right)^2 \Gamma_\phi^J$ where Γ_ϕ^J is the dephasing rate of the junction mode. If the junction mode only suffers from $1/f$ -type noise then this relation is $\Gamma_\phi^{\text{cav}} \approx \left(\frac{g_c}{\Delta}\right)^4 \Gamma_\phi^J$. Given that transmon T_ϕ s can be as long as 1 ms, we do not expect this to limit our cavity dephasing rate — more likely unwanted jumps in the transmon state will be the dominant source of cavity dephasing!

6. **How many error hierarchies are there? Is it the same as noise-bias?**

The error hierarchy refers to the ratios between different types of error a physical qubit can experience. Since some types of error are much easier to correct than others, it can be advantageous to realize a hierarchy of error rates. The main error hierarchy we consider is erasure > Pauli > leakage. The term Noise-bias was introduced to quantify physical qubits on which one type of Pauli error (e.g., Z) occurs much more frequently than the other two. Under this definition, noise-bias is a specific example of a desirable error hierarchy.

We consider three error hierarchies: The **optimal error hierarchy**, which is determined by the choice of outer stabilizer code and the error model. This is the ratio of error rates for each type of error such that each error contributes ‘equally’ to the overall logical error rate. The **idling error hierarchy** is the error hierarchy we achieve when performing no operations on the physical qubits. The **operational error hierarchy** is the error hierarchy we realize during all required operations.

7. Do we expect any noise-bias in the sense that Z errors occur more often than X errors?

Most sources of error are either leakage to $|00\rangle$ or dephasing of the individual microwave cavities, which constitutes dephasing of the dual-rail qubit. The dominant Pauli errors are expected to be Z errors. The $XZZX$ surface code can take advantage of this, even without the need for bias-preserving gates. The set of gates we present here do not preserve this noise bias.

8. Do I care about no-jump back-action in a dual-rail qubit?

No-jump backaction arises whenever the two cavity modes of a dual-rail qubit have different loss rates where $\Delta\kappa = |\kappa_1 - \kappa_2|$. No-jump backaction is small for short times, polarizing at rate $(\Delta\kappa t)^2/16$. At long times we can compensate this effect by ensuring the photon spends equal time in both cavities, regardless of the qubit state, e.g., by regularly performing SWAPs.

9. How often do we have to do erasure detection measurements?

In the main error model, erasure detection is performed after every entangling gate on the two qubits that participated in the gate. This is an ongoing area of research — can we move all erasure detection to the end of the stabilizer measurement cycle? Can we build-in error detection into the entangling gate?

10. If the cavities have a natural noise bias, why not use the XZZX surface code for biased Pauli noise?

Microwave cavities have an excellent noise bias when idling, but bias-preserving gates have not been realized for dual-rail cavity qubits, e.g., single qubit gates via beamsplitters such as Hadamards transform Z errors to X errors and vice versa, and so are not bias preserving.

11. Why the XZZX surface code? Why not concatenate with something else?

We like the XZZX surface because the ancilla dual-rail qubit is always used as the control in the required CZ or CX gates, which is better suited for our native $ZZ(\pi/2)$ gates (fewer beamsplitter gates needed on the ancilla dual-rail qubit). It also benefits from a slightly higher threshold when there is biased Pauli noise but no bias-preserving gates [Puri et al., 2020]. Concatenation with codes other than the surface code has not yet been explored in detail for erasure qubits.

12. Is this the same dual-rail code from optics? How is it different?

The dual-rail cavity qubit does share some similarities with the dual-rail code from linear quantum optics. Two photonic modes are used to encode the dual-rail qubit. In our case, these modes are standing modes in cavities at microwave frequencies. For the linear optics case, the two modes are propagating modes at optical frequencies, e.g., two different polarizations or two different spatial modes in an optical fibre or waveguide. The other similarities are that photon loss is the main source of error which forms an erasure error in both platforms, and single photon sources and single photon detection constitutes state preparation and measurement. We use parametric beamsplitter interactions in place of actual beamsplitters and phase shifters used for the single qubit gates in

dual-rail optical qubits.

The most significant difference is that the main source of nonlinearity in the optics case is single photon and detection while in circuit QED we can use the dispersive interaction to do direct entangling gates between dual-rail qubits — something that is not typically considered in optics. This has big implications. Optical dual-rail schemes propose measurement-based quantum computation and probabilistic state generation because of this.

13. Why not a measurement-based scheme like they do with the dual-rail code in optics? KLM protocol?

As shown by the KLM protocol, universal quantum computing is possible with just single photon sources, single photon detection and linear optics operations (beam splitters and phase shifters). We could do the same thing with dual-rail cavity qubits, and these operations can be made very robust to transmon ancilla errors. However, we incur a significant overhead in two major ways.

Firstly, in a measurement based scheme we generally need more dual-rail physical qubits per logical qubit, by roughly a factor of 4 (for 4-star resource states investigated in [Bartolucci et al., 2023]). We need double the number of physical qubits in a given layer compared to a surface code lattice, and then we need at least two layers to physically exist in our hardware). A factor of 4 is a big deal! For context, with four times the number of physical qubits, we could double our code distance, resulting in many orders of magnitude reduction in the logical error rates. Alternatively, if we have satisfactory error correction, with four times the number of physical qubits, we could quadruple the total number of logical qubits!

Secondly, we also need Bell measurements for measurement-based computation. These can be realized in linear optics with just beamsplitters and single photon detection but are probabilistic — the simplest Bell measurements only work 50% of the time! The other half of the time the outcome is inconclusive [Browne and Rudolph, 2005]. In linear optics, this can be boosted by introducing more modes [Grice, 2011] but increases the hardware overhead. In circuit-QED we can also perform deterministic Bell measurements on dual-rail cavity qubits by using QND ZZ measurements from chapter 5.2 in the Bell measurement circuit in 2.4. But now we are starting to use a transmon control ancilla which pushes us towards a gate-based implementation.

14. Does the dual-rail have biased erasure errors?

Not in its current form. Biased erasures arise when we also know which state the physical qubit was in before the erasure error. When leakage to $|00\rangle$ occurs via photon loss, the environment knows which cavity lost the photon (and so measures the dual-rail qubit in the Z basis) but this information is lost to us when we detect the cavity in $|00\rangle$. Related encodings such as the 0-2 encoding or 2-photon dual rail $|0\rangle_L = |02\rangle$, $|1\rangle_L = |20\rangle$ can convert single photon loss to a biased erasure error but are more complicated encodings in general. Biased erasures benefit from roughly double the threshold of regular erasure errors [Sahay et al., 2023]. When we detect errors in the ancilla transmon, they do not generally tell us information about the dual-rail qubit other than it has dephased or is in a leakage state, so these are also not biased erasures.

15. Why not use the 0-2 encoding and reduce the number of cavities?

We need extremely good single qubit gates (at the 99.99% level). This is achievable with parametric beamsplitter interactions. For the 0-2 encoding, we would need Rabi-oscillations between $|0\rangle \leftrightarrow |2\rangle$ cavity states, which appears harder to

engineer to high fidelity.

16. I'm confused, what are the logical qubits? the data qubits? the ancilla qubits?

When we are talking about dual-rail qubits as the highest level of the encoding, i.e., no concatenation with a stabilizer code/surface code, then dual-rail qubits are the logical qubits. If we are concatenating, then dual-rail qubits are the physical qubits and many physical qubits comprise the logical qubit in the next level of the encoding. In this scenario, data qubits refer to physical qubits that encode part of the logical information. Ancilla qubits refer to physical qubits are used solely to extract information from the data qubits, usually in the form of stabilizer measurements.

17. I'm confused how you actually correct erasure errors in the stabilizer code, can you go through it again?

Reset the affected qubit(s) back to the codespace, remeasure the usual stabilizers of the code, use the knowledge of which physical qubits had the erasures in the decoding step. See Appendix A for an example with the 4-qubit erasure code.

18. How do you do the non-Clifford gate?

The parametric beamsplitter interaction is analogous to a qubit Rabi drive, and can readily implement arbitrary rotations on the dual-rail Bloch sphere, including T gates at the physical qubit level which could then be used in a magic state factory.

19. How fast does the reset need to be? How do you reset? The timescale of both dual-rail measurement and gate operations are $\sim \frac{1}{\chi}$. So ideally we want to reset back into the codespace on this timescale. Reset is an open problem for dual-rail cavity qubits. If we are sure we are in $|00\rangle$, then reset is fairly simple,

requiring us to reprepare one of the cavities in $|1\rangle$ which can be achieved via sideband loading and SWAP operations. Reset from other leakage states such as $|20\rangle$ or $|11\rangle$ is less straightforward and may require Qswitching the cavities to a lossy mode, or more detailed measurement-based reset.

20. How bad is it if we fail to detect a dual-rail qubit in $|00\rangle$?

This would represent an ‘inert’ leakage error during the second round of stabilizer measurements where it would be detected with high probability. Whenever we try to do an entangling gate with an ‘inert’ leakage qubit, we generally incur Pauli errors on the un-leaked qubits — it is as though the entangling gate never occurred. Once we detect leakage, we can backtrack to reconstruct what errors happened on the un-leaked qubit to mitigate some of the effects of leakage errors. When our native entangling gates are $ZZ(\pi/2)$, inert leakage presents a source of Z error on the un-leaked qubits. General treatment of leakage errors in the surface code can be found in [Fowler, 2013]

21. Does transmon readout dephase the cavity?

Transmon readout can dephase the cavity in two ways. First, if the transmon is in an excited state, then it may decay during readout. This happens with a higher probability than when we are not reading out due to the ‘ T_1 vs \bar{n} ’ effect. When the jump happens, the cavity mode completely dephases. Most of the time, this error is detected — we require two decay errors (double jump errors) to end up in $|g\rangle$. Most of the time, we also expect to readout the transmon in $|g\rangle$ for joint parity measurements (with the appropriate choice of phase in the final transmon $\pi/2$ pulse) and $ZZ(\theta)$ gates, so undetected jump errors are generally second-order. Secondly, there exists a small cross-Kerr [Reinhold et al., 2020] between the readout resonator and the cavity. If the readout resonator is populated with photons, this can dephase the cavity. This error

can be mitigated (at least when reading out $|g\rangle$) by reducing $\kappa_r < \chi_r$ and playing our readout pulse close to $\omega_r + \chi_r$ such that the readout resonator is only populated with a large number of photons if the transmon is (unexpectedly) in $|e\rangle$ or $|f\rangle$ due to an error.

22. What happens if the transmon gets stuck in higher levels?

As observed in [Elder et al., 2020] and [Sivak et al., 2023], when performing many repeated measurements of the transmon, there is a chance we get ‘stuck’ in states higher than the $|f\rangle$ level, and we must wait until the transmon decays to the $|f\rangle$ level or lower until it is usable again. How do we cope with this error? First we must identify when this happens, and then we can omit all entangling gates involving this transmon from our rounds of stabilizer measurements. This is analogous to a ‘fabrication error’ that is considered in [Auger et al., 2017] and [Nagayama et al., 2017], although eventually the transmon will fix itself. In a surface code, or related codes [Aasen et al., 2023], we can temporarily extend the size of our plaquettes to avoid using the faulty transmon.

23. Hey there are a lot of bosonic codes out there, GKP, 4-cats, stabilized cats... what makes you think dual-rail is the best?

Because it’s the simplest bosonic qubit to operate. (Gates are even simpler than the 0-1 encoding!)

24. How good do my cavity T_1 times need to be? What about the trans-

mons? We have been benchmarking our proposed dual-rail architecture assuming cavity lifetimes ~ 1 ms and transmon coherence times ~ 100 μ s. First and foremost, the dual-rail code is designed to detect errors in the transmon control ancilla. Cavity errors are much rarer but it is important that we can detect those as well to make sure most cavities remain in the dual-rail codespace. We estimate cavity lifetimes on the order ~ 1 ms are sufficient. The longer the cavity

T_1 , the lower the probability of leakage to $|00\rangle$ which is important for keeping the fraction of ‘inert’ leaked dual-rail qubits small and for low idling errors.

25. What about noise in the coupler?

The effect of noise in the coupler is highly dependent on the choice of coupler. In general, if noise from the environment can activate the parametric beamsplitter interaction this will be a source of bit-flip errors. The cavity modes can inherit dephasing from the coupler mode as it would from say a transmon. This is still an area of investigation. It has been shown that at least for single-qubit gates, noise in the coupler does not matter on the 10^{-4} level per gate [Lu et al., 2023]. This kind of performance is also required for the entangling gate.

26. Are erasure errors always leakage errors?

No. One of the most frequent errors, ancilla transmon dephasing during two qubit gates is detectable (and mapped to the $|f\rangle$ state of the transmon), and tells us a dephasing error of the form ZZ has occurred on the two dual-rail qubits that participated in the gate. This is a Pauli-type error!

27. If most of the erasure errors are due to transmon errors, isn’t that a bit annoying?

Yes, but this allows us to use the non-linearity (the dispersive interaction) to do the entangling gates. We think it would be more annoying to do a measurement-based approach!

28. So I don’t need χ -matching but do I still need a three-level ancilla?

A three-level ancilla is being used here solely to detect decay from $|f\rangle$ to $|e\rangle$ which results in an erasure error. It is a hardware-efficient way of detecting transmon decay errors. χ -matching is not needed in any operation but is still beneficial in principle. χ -matching reduces the probability of leaking into the states $|11\rangle, |20\rangle, |02\rangle$ during joint-parity measurements and $ZZ(\theta)$ gates.

29. How do you do gates in a dual-rail surface code?

The native single-qubit gates are arbitrary rotations about an axis on the equator (e.g., X or Y axis) of the dual-rail Bloch sphere. This is analogous to an on-resonance Rabi drive for a qubit and is achieved via a parametric beam-splitter interaction. The native entangling gate we have proposed here is the $ZZ(\theta)$ gate. We can make use of the parameterization this gate affords but for universality, $ZZ(\pi/2)$ suffices. The entangling gate may be realized in other ways, e.g., two-mode SNAP. In general, these are expected to make use of the dispersive interaction between a cavity and an ancilla transmon.

30. 1 transmon, 3 couplers and 2 cavities per physical qubit doesn't seem very hardware efficient to me!

The primary focus of the dual-rail architecture is to make it easier to achieve below-threshold operation, and by aiming to be further below threshold than other approaches ($\times 100$ rather than $\times 10$ for Pauli errors), the total number of hardware elements per outer code logical qubit is reduced. For comparison, a transmon architecture may use 1 transmon and 2 couplers per physical qubit, so the total number of J.J.-based elements is similar.

31. Hey, I'm confused about the noise-model you've been using, can you run through it again?

Yue Wu's error model is described in [Wu et al., 2022] and shows that it is possible to run a surface code with both erasure and Pauli errors, provided we have a strong error hierarchy. It is an error model originally designed for Rydberg atom entangling gates, but is fairly general for other platforms. Here we point out which aspects are not experimentally realistic:

- Leakage detection is presumed instantaneous and occurs after every entangling gate.

- Undetected leakage errors are instantly converted to Pauli errors on the two qubits that participated in the entangling gate, rather than persisting as leakage errors.
- Reset of an erased qubit is perfect and instantaneous.

6.11 Conclusions

Erasure qubits present a whole new paradigm that is only recently being studied for specific hardware platforms, and carries great potential for simplifying practical error correction. The core idea is that we can design our hardware such that the most likely errors are the ones that are the easiest to correct in the outer layers of our error correcting code. This is the motivation behind engineering an error hierarchy and for our case, the three types of errors in our hierarchy are erasures, Paulis and leakage errors.

To this end, we have introduced the dual-rail cavity qubit. Especially pertinent in the cQED platform is the ability to design all operations and choose an encoding such that the qubit has the ideal properties needed for an erasure qubit, while remaining relatively simple to realize. We have provided recipes for hardware implementation of all the key operations that should be imminently realizable with today's hardware. These are:

- State preparation and readout of dual-rail cavity qubits
- Single qubit gates (parametric beamsplitter interaction)
- Two-qubit gates ($ZZ(\theta)$)
- mid-circuit erasure detection (via joint parity or Y-mon measurements)
- dual-rail reset (cavity measurements and photon loading via sidebands)

Each building block has been designed so it conforms closely to the optimal error hierarchy, with erasure errors set by cavity decay and transmon ancilla errors at the 1% per gate level, Pauli errors by double jump errors in the transmon ancilla/transmon readout at 0.01% per gate level and a low background of undetected leakage at the 0.001% level per gate. With these building blocks, it should be possible to realize a good erasure qubit by concatenating with a surface code such as the $XZZX$ surface code. The set of operations we introduce are by no means final, and can likely be further optimized. Rather they are an example of a complete set of operations that adhere to the desired error hierarchy.

Future work will likely focus on optimizing dual-rail cavity qubits for specific use in circuits that measure stabilizers — the circuit in Fig. 3.4 and we could further design operations to respect the different classes of error in the surface code. For example, the surface code is relatively robust to stabilizer measurement errors and this could be factored in to the design of operations.

The concept of dual-rail cavity qubits is still in its infancy, and while we have outlined one way to realize them, this is not set in stone. Major hardware changes could involve replacing all 3D cavity modes with on-chip stripline resonators, and using multimode cavities to make the footprint of these qubits more compact. Other areas to investigate are the many possible ways of implementing entangling gates other than via $ZZ(\theta)$ gates, and the many possibilities for what to use as the ancilla physical qubit. Here the prospect exists for a ‘mixed species’ code, where the ancilla physical qubits are different from the hardware used for the data qubits. In our case, the data qubits would be dual-rail cavity qubits but really, a whole range of cQED devices could be used to measure the stabilizers, provided they rarely cause undetectable backaction on the data qubits when hardware errors occur.

What is equally pressing is the development of accurate, hardware-specific error models and decoders dedicated to dual-rail cavity qubits. Certain details, such as ‘inert’ leakage errors, reset errors, and accurate circuit level noise (Z errors are the most common Pauli error) are not yet accounted for in surface code simulations for these codes. This knowledge is needed to tell us precisely what error rates we need to target in operations and what kind of logical qubit errors we could expect in a real surface code. A framework where we can make changes at the hardware level and observe their direct effect on the surface code logical error rate would be a powerful design tool.

For near-term characterization, dual-rail cavity qubits and erasure qubits in general present new challenges. Techniques such as randomized benchmarking (RB) and interleaved randomized benchmarking (IRB) work well when we want to characterize gate fidelities with a single number, by measuring out to long times after the application of thousands of gates. With erasure qubits, we now also need to characterize p_{fail} and $\varepsilon_{\text{pass}}$ for all the gate operations. $\varepsilon_{\text{pass}}$, the gate fidelity when no errors are detected is challenging because randomized benchmarking techniques will be less effective. This is because double-jump errors are expected to limit fidelities at long times, leading to RB dynamics that cannot be described by a simple exponential decay. We will be most interested in ‘short time’ fidelities, where the errors here represent the errors we would experience in a round of stabilizer measurements in say a surface code. To extract these numbers we will instead need to rely on the high-fidelity SPAM expected for dual-rail qubits and careful characterization of systematic errors such as measurement bias. Full process tomography for the entangling gates may also be useful for constructing accurate error models for our circuit-level noise. It is likely that a whole suit of tools will need to be developed to characterize the relevant coherence times and relevant operation fidelities in dual-rail cavity qubits.

Error Detection for Quantum Networks

Now we move onto the final major part of this thesis, where we apply the principles of error detection to the problem of quantum communication. In general, the goal of quantum communication is to transfer quantum information across some appreciable distance, via a communication channel. Ultimately, the goal is to send qubits from point A to point B with a high rate of data transfer and a high fidelity. The use context does not have to be quantum computing but could for example be for quantum cryptography instead (See BB84 protocol [[Bennett and Brassard, 2014](#)]). Quantum communication complements quantum computing in the sense that multiple (hypothetical) quantum computers can be networked together [[Kimble, 2008](#)] over long distances with quantum communication protocols to build a quantum network where logical qubits are exchanged between nodes in the network.

This quantum internet is a long way away from reality! In the meantime, we will be primarily interested in using quantum communication to ease the task of building a quantum computer, through the paradigm of modularity. It is no secret

that multi-qubit processors are harder to build and are more error-prone than few-qubit processors. This is due to fabrication reliability, crosstalk and wiring issues. The principle of modularity says we should build and independently test small-scale qubit processors, and then network them together with quantum communication techniques to build a more powerful processor [Jiang et al., 2007]. The challenge of building a monolithic processor is now replaced with the challenge of building good quantum communication links between modules - which we hope is easier!

Most quantum industry players have some sense of modularity built into their road maps [Bravyi et al., 2022, Gold et al., 2021]. It is probably unreasonable (though perhaps with CMOS-based qubits) to build 20 million physical qubits in a single device. The first step however is just to link modules together over short distances, much less than 1m. Here the problem is still a quantum communication problem but the context and scales are quite different from what the phrase quantum communication usually invokes — fiber optic channels and free-space (sometimes outer space!) transmission of qubits over 10s or 100s of km.

For microwave frequency platforms, the quantum communication schemes developed for optics are not the optimum (although we often find optical quantum communication proposals are sometimes first implemented in superconducting qubits!). One thing we will rigorously explore are the consequences of being able to treat our communication channel not as a continuum of modes, but a single standing wave mode. The experiment we will discuss hinges heavily on this fact.

What are the figures of merit for quantum communication? We have already mentioned the fidelity and rate of qubit transfer. In a computing context, we want both to be high. Ideally the fidelity of the communication should be similar to the fidelity of the local gate operations on the physical qubits, so we do not further complicate the task of quantum error correction. The rate of qubit transfer should be similar to the rate local gates are performed, otherwise we reduce the ‘clock speed’ of

our quantum processor and incur more errors from idling.

We often focus on qubit transfer as the goal of quantum communication in a modularity setting but this is not strictly necessary. We may for instance perform an entangling gate directly across the network between two distant qubits and that would also suffice for quantum computing. We simplify the task to qubit transfer because this enables all the other quantum processing we could possibly want to do across our network (e.g., we could transfer the qubit from module A to module B, perform the entangling gate locally with qubits in module B and then transfer back).

Other important metrics are the success probability. In quantum communication, this can sometimes be shockingly low, often requiring thousands or millions of attempts until a qubit is transferred successfully. The trick to get away with low success rates is to instead try many times to herald entanglement and then teleport the qubit state, as was discussed in Sec. 2.4.2 of Chapter 2. We do pay a penalty, in the form of a reduced and probabilistic transfer rate. The average transfer rate will be the rate of attempts multiplied by the success probability of attempts, but we will be waiting around an uncertain amount of time for each qubit transfer. This does not seem ideal for a modular quantum computer where we might occasionally be waiting a long time until we can proceed with the next gate in the circuit. However, this can be mitigated by parallelizing the heralding of entangled states with multiple links, such that with high probability there is always an entangled state ready to teleport a qubit ‘on-demand’.

This is where error-detection fits in. Error-detection allows us to reduce rate of transfer in exchange for increasing the fidelity of qubit transfer. Depending on the communication context, this can be a worthwhile trade-off. Quantum communication has more leniency for success probability than quantum computing. For instance, if I detect an error half the time on average while trying to send an entangled state then this reduces my average transfer rate by a factor of 2. In a quantum computing

context, detecting an error in a gate 50% of the time is much too high to do anything useful like high-depth circuits or erasure qubits.

The main challenge to the modularity approach is that the links themselves tend to be the most lossy and error-prone parts of the system, and so quickly set the bottleneck of our modular processor's performance. Although optical fibre can have attenuation lengths of 10s of km and similarly for microwave superconducting coaxial cable, this is still the major source of loss compared to the physical qubits themselves in both platforms.

The problem our experiment tackles will be centimeter-meter scale communication between superconducting cavity modules such as those introduced in Chapter 4. Over this length scale (and for our experimentally available coupling strengths), our communication channel ought to be treated as a single standing wave mode. We will also be interested in the regime where the loss in our communication channel is significantly greater than our coupling rate. At first glance, it looks like we can *never* succeed in sending anything through this channel, although this is certainly the regime most easiest to realize in experiment!

We will show that by combining three main ingredients: a standing wave mode link, bosonic cat qubits and error-detection via heralding, it is in fact possible to herald Bell states with a remarkably high success probability $\sim 50\%$ per attempt, and such Bell states should be practically free of all errors due to loss in the link!

Our scheme is, to our knowledge, the first example of a quantum communication scheme where the heralding relies specifically on interference effects in a standing wave mode link. To better understand the context and use cases for this scheme, it is helpful to first review the 'standard' regimes of quantum communication that have been implemented in various qubit platforms.

7.1 Regimes of Quantum Communication

We simplify the problem of quantum communication to focus on qubit transfer between two nodes, Alice and Bob. The qubit to be transferred may be a two-level system or a bosonic qubit. We also take our communication channel to be a section of fibre or microwave transmission line/coax/waveguide of distance L , which naturally supports multiple standing wave modes that can be excited by coupling to them and serve as the medium for the transfer of energy from node 1 to node 2. These standing wave modes are spaced in frequency by Δ_{FSR} , the Free Spectral Range where

$$\Delta_{\text{FSR}} = \frac{c}{2L} \quad (7.1)$$

Where c is the group velocity in the channel, typically the speed of light (or a significant fraction of) in a photonic link or the speed of sound in a phononic link.

We presume all our couplings between the qubit and the link to be beamsplitter couplings of the form

$$\hat{\mathcal{H}}_{\text{comm}} = \sum_j g_{1,j}(t) \hat{a}_1^\dagger \hat{b}_j + g_{2,j}(t) \hat{a}_2^\dagger \hat{b}_j + \text{h.c.} + \sum_j \Delta_j \hat{b}_j^\dagger \hat{b}_j \quad (7.2)$$

Where

$$\Delta_j = j \Delta_{\text{FSR}} \quad (7.3)$$

For the j^{th} harmonic where we designate $j = 1$ to be the fundamental mode. Importantly, we can vary the couplings $g_{1,j}(t)$ and $g_{2,j}(t)$ in time. We need to be able to turn our communication on and off! At the very least we need to be able to set both couplings to zero in the ‘off’ state.

When we couple resonantly to a single standing wave mode in the link, \hat{b} our

Hamiltonian becomes

$$\hat{\mathcal{H}}_{\text{disc}} = g_1(t)\hat{a}_1^\dagger\hat{b} + g_2(t)\hat{a}_2^\dagger\hat{b} + \text{h.c.} \quad (7.4)$$

Finally there is the characteristic loss rate, κ of the link. For the discrete case, this is the loss rate of the standing wave mode we couple to. For the continuum case, this is the typical loss rate of the modes within the frequency bandwidth of the wavepacket we excite which is typically taken to be a constant.

These three parameters, g , κ and Δ_{FSR} largely govern the physical regime of our quantum communication and which quantum communication protocols are suitable. As shown later in Fig. 7.1, we can re-express these terms in the time and spatial domains, where they correspond to different physical aspects of our communication protocol.

An important figure of merit for any deterministic state transfer scheme is the energy transfer efficiency, η or more conveniently the energy transfer inefficiency $\bar{\eta} = 1 - \eta$. Simply put, this is the probability that a single photon from node 1 does not make it to node 2 at the end of the protocol or in other words, the fraction of energy lost during the state transfer. For a two-level qubit this sets an upper bound $\sim \eta$ on the transfer fidelity. For a bosonic qubit, this sets an upper bound on the average energy transfer but not the fidelity, since we may encode in a bosonic code designed to correct for photon loss (which also corrects for loss that happens in the link since all the link modes are bosonic and we are assuming beamsplitter couplings). For example, by error-correcting loss in the link using a code designed for single-photon loss [Burkhart et al., 2021], we now have a infidelity $\sim \bar{\eta}^2$.

There is a fundamental limit to how large $\bar{\eta}$ can be. We call this limit the ‘single pass loss’ and it is set by

$$\bar{\eta} = \frac{\kappa L}{c} = \frac{\kappa}{2\Delta_{\text{FSR}}} \quad (7.5)$$

Simply put, the fastest possible time we can transfer energy through the link is given by the speed of light/sound in the link and the distance traversed. During this time we experience energy loss at rate κ . Often we find adiabatic schemes that use discrete modes appear perfect, until we consider the effects of other, off-resonance modes in the link. Although the effect of these modes is typically small, we always end up recovering the single pass loss as the upper bound on how well an adiabatic scheme could theoretically perform.

Before we move on to describe briefly some of the many qubit transfer protocols that have been realized, we first present a mini-glossary that explains the various terms used to qualify a particular scheme

Continuum vs Discrete

Whether the communication channel can be treated as a continuum of (bosonic) modes or a single standing wave mode. At optical frequencies in the THz regime, optical communication links are always treated as a continuum of modes. This is also often the case for phononic microwave links due to their much slower wavespeed, which is given by the speed of sound rather than the speed of light. In these links the coupling strength g is larger than Δ_{FSR} . In microwave photonic links which operate at frequencies of only a few GHz, over distances of up to 10s of metres, the link can be well-treated as a single standing wave mode, since we typically can only engineer $g < \Delta_{\text{FSR}}$. There are some exceptions to this, either by engineering $g > \Delta_{\text{FSR}}$ with certain couplers [Zhong et al., 2019], or alternatively, we can set Δ_{FSR} to zero by introducing lossy directional elements, although there is arguably little benefit in doing so since $\bar{\eta}$ will still be limited by κ/g and the circulator presents an additional source of loss.

Deterministic vs Probabilistic

Deterministic or ‘on-demand’ communication means the communication protocol succeeds 100% of the time. Probabilistic means that success is not guar-

anteed on each attempt, but is instead heralded by some measured signal that tells us if a particular attempt was successful. For this reason, probabilistic schemes often generate entanglement first for teleportation.

Direct transfer vs Teleported

Deterministic schemes usually opt for direct state transfer, where a qubit is literally transported through the communication channel and received on the other side. Teleported schemes must first generate an entangled pair between the two nodes, and then perform local Bell measurements and classically communicate the results to teleport the qubit state.

Microwave vs Optical

Microwave schemes use microwave links made out of coaxial transmission line or waveguides (photonic or phononic) and typically interface with superconducting qubits (or other qubits with accessible transitions in the microwave regime which we can couple to a microwave link). Optical schemes work at optical frequencies and use fiber optics, optical waveguides or free space as the communication link.

Adiabatic

Adiabatic schemes are sometimes used for discrete mode communication. By changing parameters slowly compared to g , we can achieve state transfer with inefficiencies $\bar{\eta} < \kappa/g$ whilst incurring a penalty in an increase in the overall transfer time.

Intra-fridge vs Inter-fridge

For superconducting qubits (or other qubit platforms that require dilution refrigerators), intra-fridge communication refers to communication within a fridge, e.g., when building a modular quantum processor by joining chips/modules together over the scale of a few cm at most. Inter-fridge communication is link-

ing processors between fridges and requires either microwave-to-optical transduction [Lauk et al., 2020] or a long cryogenic link [Magnard et al., 2020]

Directional vs Bidirectional

A directional scheme can only transfer from node 1 to node 2 but not the other way around, without significant hardware changes. This is the case when we insert circulators into microwave links. Here we define bidirectional to mean we can transfer a qubit from node 1 to node 2 or from node 2 to node 1, without making any hardware changes. We might need to still change the details of the protocol, e.g., the exact pulse shaping if pitching and catching wavepackets. Schemes that are the ‘most’ bidirectional implement a SWAP gate between the qubits at node 1 and node 2.

7.2 Examples of Deterministic State Transfer Schemes

The generalized framework for analyzing deterministic qubit transfer protocols is shown in Fig. 7.1 We first describe the continuum protocols, since these were first developed for optics. Note that with many of these transfer schemes, the underlying physics principles are classical. They are compatible with quantum information processing because they preserve qubit coherence if the state we are trying to transfer is in a quantum superposition.

7.2.1 Pitch and catch

Owing to low fiber coupling efficiencies and photon collection efficiency, optical protocols typically rely on heralding and are probabilistic. There is however a deterministic protocol that we will call ‘Pitch and Catch’ which is based off of the proposals in [Cirac et al., 1997] in cavity QED and then adapted for microwaves in [Korotkov, 2011]. This scheme is also referred to as the CZKM scheme after the authors in [Cirac

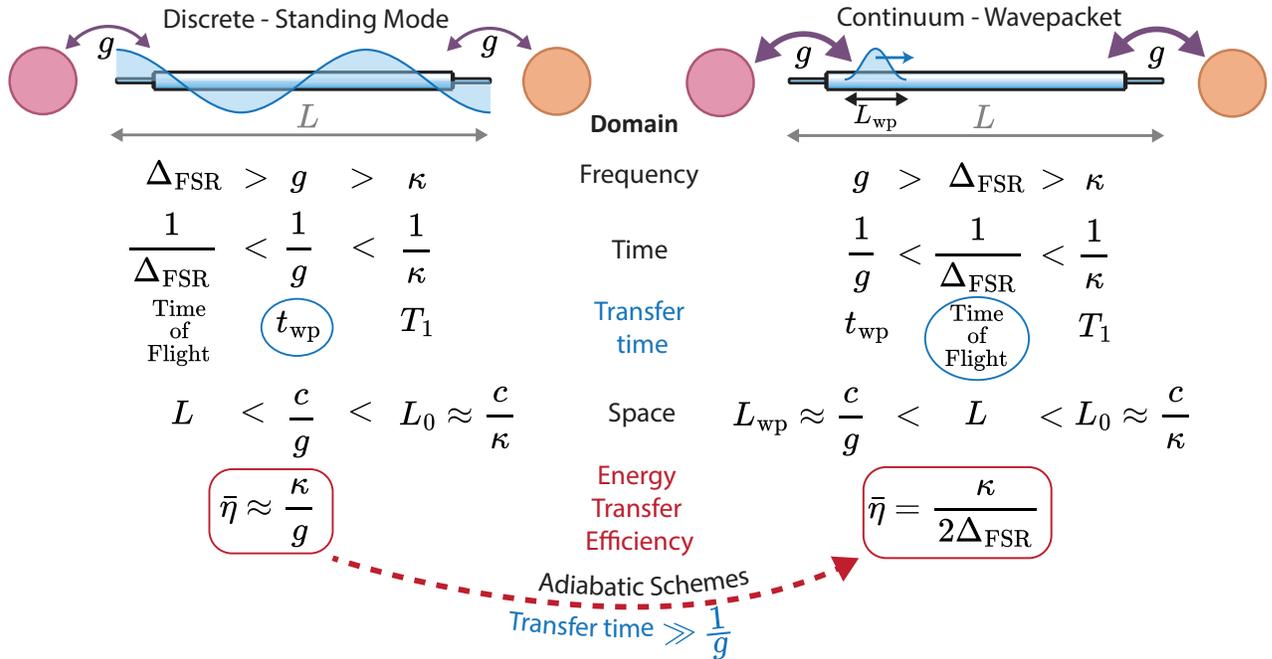


Figure 7.1: General properties and limiting factors for discrete and continuous deterministic quantum communication schemes. We can analyze the physics of a scheme in the frequency, time or spatial domain.

et al., 1997]. If we were to engineer a constant beamsplitter coupling g between a qubit and some intermediary mode that has a strong coupling to the continuum of modes on the communication channel, then the qubit state would ultimately ‘leak out’ into the channel as a wavepacket with an exponentially decaying envelope. It turns out this wavepacket cannot be perfectly absorbed with only a constant coupling at the receiver node, due to its discontinuous envelope. The ‘Pitch and Catch’ scheme states that if we can instead modulate the amplitude of the beamsplitter couplings $g_1(t)$ and $g_2(t)$ in time then we can shape the envelope of the wavepacket. By choosing $g_1(t)$ such that the envelope of the wavepacket is time symmetric¹, then in theory it can be perfectly absorbed at node 2 with a $g_2(t)$ that satisfies $g_1(t) = g_2(\tau - t)$ where τ is the small delay that accounts for the time of flight of the wavepacket through the link.

1. Strictly the envelope does not need not to be time symmetric, so long as the catch pulse is the time reverse of the pitch pulse.

Aside from technical difficulties such as experimental pulse shaping errors and efficiencies in coupling to the link, the energy transfer efficiency of this scheme is limited by the single pass loss, i.e., $\bar{\eta} = \kappa/2\Delta_{\text{FSR}}$

With a circulator

Because it is easier to engineer the necessary couplings to the communication channel, this scheme was first realized not in optics but in microwaves [Kurpiers et al., 2018, Axline et al., 2018, Campagne-Ibarcq et al., 2018]. Herein lies an experimental challenge: it is hard to engineer $g > \Delta_{\text{FSR}}$ unless we use an impractically long microwave link that would struggle to fit inside a single dilution refrigerator. The experimental solution was to add a circulator/isolator, which acts like a one-way valve for microwave signals, effectively setting the Δ_{FSR} to zero and attenuating the signal typically by a few decibels.

A curious side effect is the time to release the wavepacket is now longer than the time-of-flight through the system, and so the overall transfer time is set by $1/g$, not $1/\Delta_{\text{FSR}}$ as we would expect for the optical case. This also means the energy transfer efficiency is limited by $\bar{\eta} \sim \kappa/g$, plus the loss from the circulator.

Despite its shortcomings, this scheme remains popular for microwave communication, which aside from a few exceptions still requires a circulator, even when traversing distances of up to 30 m. An advantage of this scheme is that it can truly be described as remote. The protocol does not change significantly if we extend the length of our link to arbitrarily long distances, aside from an increase in the energy loss due to a longer time of flight. From a practical perspective, loss in the link is typically one of the main limits on the transfer fidelity, but inefficiencies due to the ‘pitching’ and ‘catching’ of the wavepacket also contribute to the transfer infidelity.

Time-binned

The pitch and catch protocol can be modified with a Time-binned protocol. Usually, if the qubit at node 1 is in $|1_L\rangle$ we pitch the wavepacket and if it is in $|0_L\rangle$ we do not. With time-binning, we always pitch a wavepacket but the time delay at which we pitch the wavepacket is dependent on the qubit state. The catch procedure at node 2 is sensitive to the time at which the wavepacket arrives. In addition, we can then use a ‘witness’ at node 2 which detects if a wavepacket arrives, but knows nothing about the time of arrival. This allows us to error-detect the transfer. Success is only declared if the wavepacket arrives. As a heralded protocol, this scheme has the potential to overcome all loss in the link, where loss acts only to reduce the success probability. This scheme has also been realized in superconducting qubits [Kurpiers et al., 2019, Zalus-Geller, 2022].

7.2.2 The quantum bus

The quantum bus is the name we give to any scheme that uses on-resonant couplings to a single standing wave mode in the communication channel, which we call the bus mode. These schemes can all be described by the Hamiltonian in Eq. 7.4 and this was the first way superconducting qubits were coupled together, albeit on the same chip [Majer et al., 2007]. As long as this Hamiltonian is a valid model for our system, we can use these schemes for qubit transfer over appreciable distances. For a typical coupling rate on the order of $g/2\pi \sim 0.1\text{-}1$ MHz, this corresponds to link lengths up to 10-100 m to satisfy $g < \Delta_{\text{FSR}}$.

With our knowledge of beamsplitters from Chapter 4, a simple way we can engineer a (directional) qubit transfer is to first turn on only $g_1(t)$ at constant amplitude $g_0/2$ for time $t = \pi/g_0$, which initiates a SWAP between \hat{a}_1 , the mode used to store the qubit in node 1 to the bus mode \hat{b} . We then set $g_1(t) = 0$ and $g_2(t) = g_0/2$. After time $t = \pi/g_0$ we will have swapped the state to node 2, completing our transfer.

The total time of the transfer is $t = 2\pi/g_0$ giving us an energy transfer inefficiency of $\bar{\eta} = 2\pi\kappa_b/g_0$, where κ_b is the loss rate of the bus mode. This scheme can generate entanglement if using the 0-1 encoding by performing a 50-50 beamsplitter or ‘half swap’ with $g_1(t)$ followed by a full swap with $g_2(t)$.

We can improve on this scheme in what has been called the ‘hybrid’ scheme [Wang and Clerk, 2012] whereby we set $g_1 = g_2 = g_0/2$ during the full duration of the transfer, with the Hamiltonian

$$\hat{\mathcal{H}}_c = \frac{g_0}{2}(\hat{a}_1^\dagger + \hat{a}_2^\dagger)\hat{b} + \text{h.c.} \quad (7.6)$$

After time $t = \sqrt{2}\pi/g_0 = \tau_{\text{SWAP}}$, we will have implemented a SWAP gate between modes \hat{a}_1 and \hat{a}_2 with a much improved energy transfer inefficiency of $\bar{\eta} = \frac{1}{\sqrt{32}}(2\pi\kappa_b/g_0)$. These improvements come from the faster transfer time and the fact that we can begin ‘unloading’ our state from the bus mode from the other side as soon as we start ‘loading’ by turning on both couplings simultaneously, which reduces the average photon number of the bus mode (full details worked out in [Burkhart et al., 2021]). This improvement comes from the fact that our transfer is now faster and the population of the bus mode is reduced (see Fig. 7.2).

This is another popular scheme for microwave communication and forgoes both the circulator and complicated pulse shaping needed for the Pitch and Catch protocol. The highest fidelity microwave communication in a modular superconducting processor uses this scheme.

A variation of the hybrid scheme can also be used to generate deterministic entanglement. We instead detune both our couplings from resonance by Δ_{BS} to realize the Hamiltonian

$$\hat{\mathcal{H}} = \frac{g_0}{2}(\hat{a}_1^\dagger + \hat{a}_2^\dagger)\hat{b} + \text{h.c.} + \Delta_{BS}\hat{b}^\dagger\hat{b} \quad (7.7)$$

where we choose $\Delta_{BS} = g_0\sqrt{\frac{2}{3}}$. After time $t = \sqrt{\frac{3}{4}}\tau_{\text{SWAP}}$, we will have enacted

a 50-50 beamsplitter between \hat{a}_1 and \hat{a}_2 . The reasons these particular values work are discussed in [Burkhart et al., 2021]. By tuning our beamsplitter couplings from resonance we can find points in time where the bus mode is empty and a beamsplitter transformation is enacted between modes \hat{a}_1 and \hat{a}_2 . A 50-50 beamsplitter is sufficient to generate entanglement provided we start with the appropriate input states. By preparing a single photon in one or both of the qubits, we can use this beamsplitter to generate a Bell state in the 0-1 or 0-2 encoding.

In Figs. 7.2 and 7.3 we compare the state transfer dynamics for various states. Note, that since we are describing state transfer between different modes, the overall transfer operation is state independent. In these QuTiP simulations we always begin in the initial state $|1\rangle_{a_1} \otimes |0\rangle_{a_2} \otimes |0\rangle_b$ and set $g_0/2\pi = 1$ MHz for a fair comparison. The dynamics we plot are also reminiscent of the evolution of the mode operators, $\hat{a}_1(t)$, $\hat{a}_2(t)$ and $\hat{b}(t)$.

7.2.3 Adiabatic schemes

We define an adiabatic scheme for discrete transfer to be qubit transfer when the bus mode can be adiabatically eliminated from the dynamics. If some condition for adiabaticity is reached (in a limiting case) then the population of the bus mode tends to zero at all times during the protocol, and hence we are no longer sensitive to loss from the bus mode. Invariably these schemes come with an ‘adiabatic slowdown’ and it is impossible to transfer better than the single pass loss limit.

STIRAP

STIRAP stands for STimulated Raman Adiabatic Passage. Originally developed for state transfer between energy levels in various quantum systems, this approach works equally well for distinct quantum modes in a quantum communication context. We use on-resonance couplings but the ‘counter-intuitive’ pulse sequence. To transfer

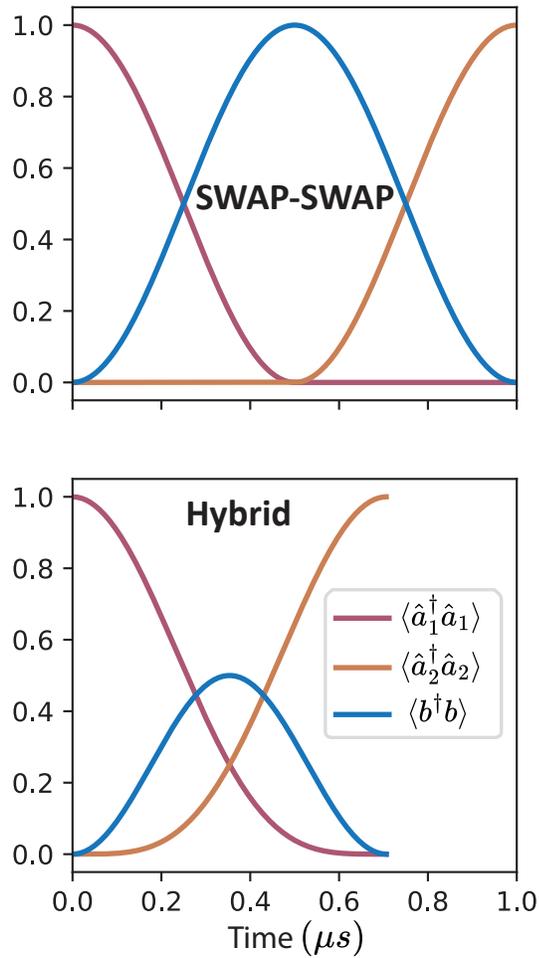


Figure 7.2: QuTiP simulations of population dynamics during schemes for on-resonance deterministic state transfer. **(Top)** The Swap-Swap scheme. The qubit at node 1 is first swapped to the bus mode, and then to node 2. **(Bottom)** The ‘hybrid’ scheme where $g_1 = g_2 = \frac{g_0}{2}$ for the entire protocol, which performs a SWAP operation on modes \hat{a}_1 and \hat{a}_2

from node 1 to node 2, we must first turn on $g_2(t) = g_0$ and set $g_1(t) = 0$. This is not intuitive at all! We then slowly ramp down $g_2(t) \rightarrow 0$ and ramp up $g_1(t) \rightarrow g_0$. If we do this slowly enough then at the end of the pulse sequence we will have performed a directional transfer from node 1 to node 2. We are making use of the adiabatic approximation: if the Hamiltonian changes slowly enough, we will remain in the eigenstate we started in, even if that eigenstate is slowly being redefined. The same applies to the eigenmodes of a multimode Hamiltonian which is our case. The instantaneous eigenmode of this system (sometimes called the ‘dark mode’) is

$$\hat{a}_d = \cos \theta \hat{a}_1 - \sin \theta \hat{a}_2 \quad (7.8)$$

where $\tan \theta = g_1/g_2$. By starting with $g_2 = g_0$ and $g_1 = 0$, we begin in eigenmode \hat{a}_1 which becomes \hat{a}_2 at the end of the sequence. This was experimentally demonstrated in superconducting qubits on a single chip in [Chang et al., 2020].

Formally, the condition for adiabaticity is

$$\dot{\theta}(t) \ll g_0 \quad (7.9)$$

which states the eigenmodes of the Hamiltonian must change slowly with respect to g_0 .

Virtual Raman transitions

Virtual Raman transitions were also developed for state transfer within a single quantum mode but the same principle also extends to multimode systems. This time, we detune both couplings from resonance by Δ_{VR} to engineer the Hamiltonian

$$\hat{\mathcal{H}} = \frac{g_0}{2}(\hat{a}_1^\dagger + \hat{a}_2^\dagger)\hat{b} + \text{h.c.} + \Delta_{VR}\hat{b}^\dagger\hat{b} \quad (7.10)$$

The key difference with Eq. 7.7 is that we now choose $\Delta_{VR} \gg g_0$. In this limit, it is as if we are doing a parametric beamsplitter interaction directly between modes \hat{a}_1 and \hat{a}_2 at the reduced effective beamsplitter rate $g_{BS} = g_0^2/2\Delta_{VR}$. At the same time, the population of the bus mode is reduced to sinusoidal oscillations at frequency Δ_{VR} and amplitude g_0/Δ_{VR} with suppressed average population $\frac{1}{2}(g_0/\Delta_{VR})^2$. When we calculate the energy transfer inefficiency, we find these two effects do not cancel out! Overall, we find the fraction of the energy lost is $\bar{\eta} = \pi\kappa_b/\Delta_{VR}$, which decreases as we increase Δ_{VR} , meaning this scheme can be advantageous if we are willing to increase the overall transfer time. We can also realize an arbitrary angle beamsplitter transformation with this approach.

In an elegant way we can recover the single pass loss limit by noticing that there is a *maximum* detuning from a mode. If we detune by $\Delta_{VR} > \Delta_{FSR}/2$ then we will be closer to being on resonance with the adjacent harmonic! At this maximum detuning $\Delta_{VR} = \Delta_{FSR}/2$ we calculate the energy loss in the bus to be *exactly* the single pass loss! (See [Leung et al., 2019], the same goes for STIRAP too — see [Vogell et al., 2017])

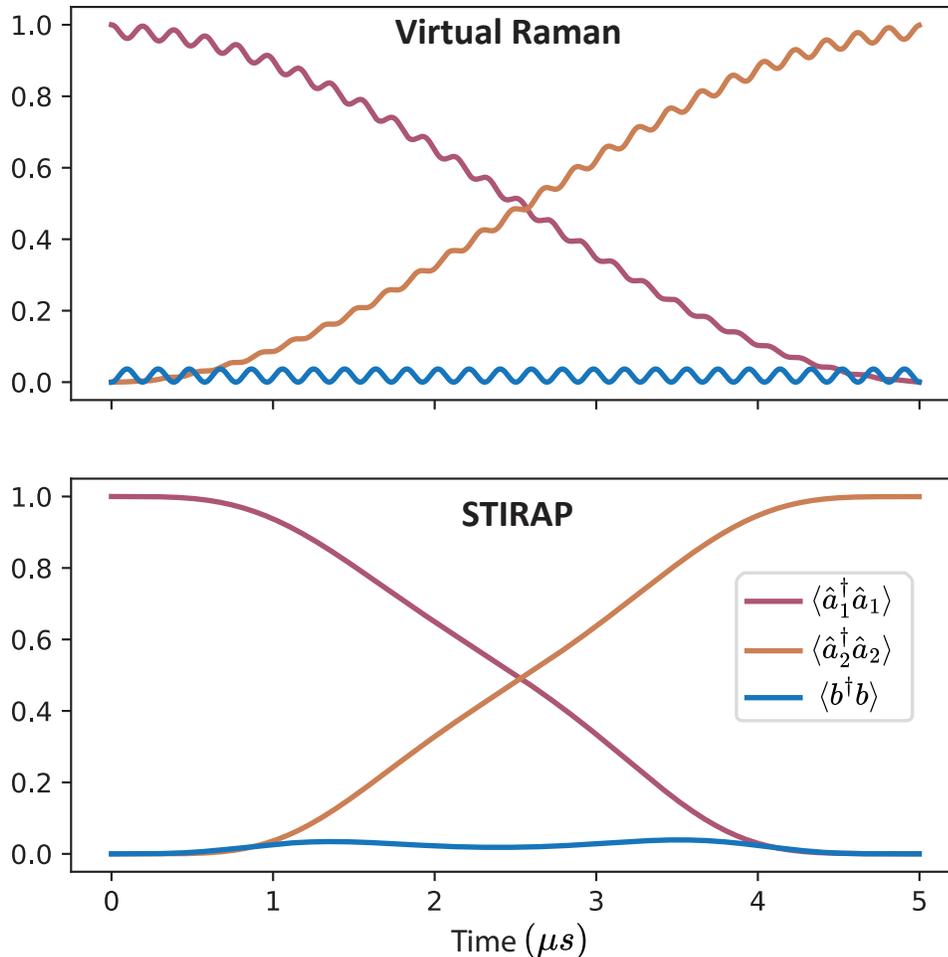


Figure 7.3: Simulated population dynamics for the two adiabatic transfer protocols. Here we have chosen parameters so we are not entirely in the adiabatic regime, to hint at the bus population dynamics. **(Top)** Virtual Raman scheme where we set $g_1 = g_2 = g_0/2$ and set $\Delta_{VR} = 5g_0$ with $g_0/2\pi = 1$ MHz for a fair comparison with the on-resonance schemes. Dynamics approximate a parametric beamsplitter interaction between \hat{a}_1 and \hat{a}_2 **(Bottom)** STIRAP for state transfer between modes. We chose a simple pulse shape where $g_1(t) = g_0 \frac{t}{t_{\text{ramp}}}$ where the beamsplitter amplitude linearly increases in time $t_{\text{ramp}} = 5 \mu\text{s}$ to its maximum value g_0 . Similarly, $g_2(t) = g_0(1 - t/t_{\text{ramp}})$, linearly ramps down. The condition for adiabaticity for this pulse is $t_{\text{ramp}} \gg 1/g_0$.

7.3 The Crossover from Discrete to Continuum

All the deterministic state transfer schemes we described can be modelled with the Hamiltonian in Eq. 7.2, even in the continuum limit (although this is not the most efficient way to simulate these dynamics.). This model is useful for investigating dynamics due to our imperfect assumption that we can only couple to exactly one mode in the link when implementing a discrete qubit transfer scheme, e.g., we can see the effects of off-resonant coupling to the other modes.

We use this model to examine the crossover from the discrete mode to the continuum of modes regime if we were to increase g_0 , keeping all other parameters fixed. In this model we assume only one node couples with the link with a static, always on coupling, to simplify the dynamics. We consider the first seven harmonics of the link, denoted by \hat{b}_j for $j = 0, 1, 2, 3, 4, 5, 6$ and our beamsplitter coupling is initially on-resonance only with the $j = 3$ harmonic. This gives us the Hamiltonian

$$\hat{\mathcal{H}} = g_0 \sum_j (-1)^{j+1} \left(\hat{a}_1 \hat{b}_j^\dagger + \hat{a}_1^\dagger \hat{b}_j \right) + \sum_j (j-3) \Delta_{\text{FSR}} \hat{b}_j^\dagger \hat{b}_j \quad (7.11)$$

Note the alternation of the signs in the coupling to the j^{th} harmonic. For the link we consider, the parity of each harmonic alternates between even (cosine) and odd (sine) as we increment j . In Fig. 7.4 we plot the dynamics of this Hamiltonian when we initially prepare mode \hat{a}_1 in $|1\rangle$, for various ratios of g_0/Δ_{FSR} and observe an abrupt crossover with an intermediate regime with messy dynamics and properties of both discrete and continuous dynamics. From the population of the j^{th} harmonic, $n_{b_j}(t) = \langle \hat{b}_j^\dagger \hat{b}_j(t) \rangle$, we can construct the spatial energy density distribution for the electric field, $u(x, t)$ in our link using the equation

$$u(x, t) = \frac{2}{L} \cos^2((j+1)\Delta_{\text{FSR}}t) \cos^2(kx) \quad (7.12)$$

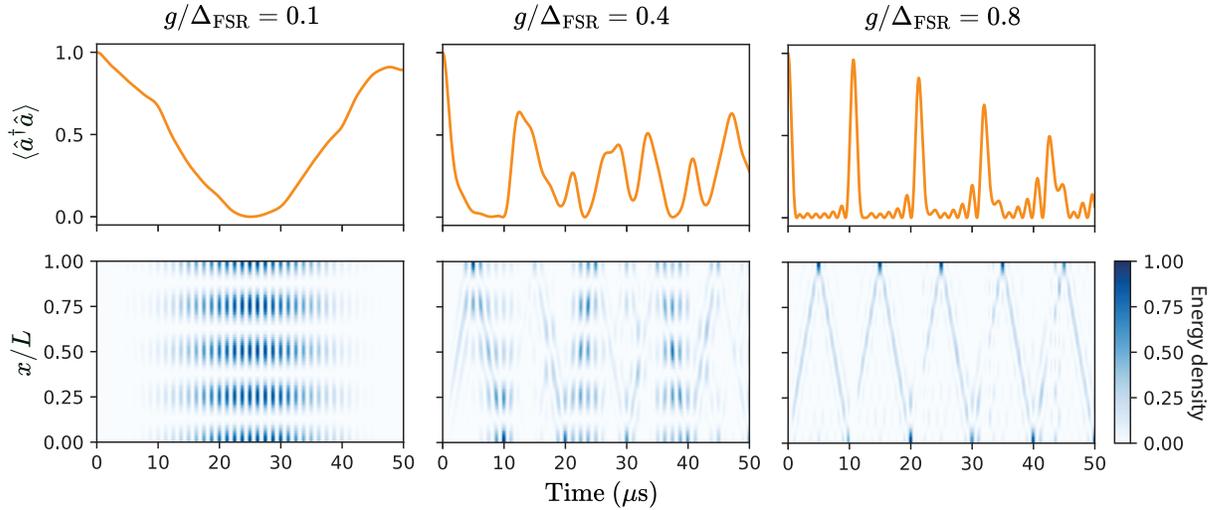


Figure 7.4: Simulating the crossover from a discrete mode to a continuum of modes. Here we fix $\Delta_{\text{FSR}}/2\pi = 100$ kHz, corresponding to a link length of ≈ 500 m. Top row - population of the qubit at node 1, Bottom row, electrical energy density in the link as a function of position along the link and time. Left column - in the discrete mode regime we mostly realize a parametric beamsplitter interaction between qubit 1 and the ‘bus’ mode, with distortion due to off-resonant couplings. At the halfway point, we have excited mainly the $j = 3$ harmonic only and swapped the qubit into the bus mode. At longer times we swap the energy from the bus mode back into the qubit. Right column - when g approaches Δ_{FSR} we begin launching wavepackets which travel down the link and reflect off the other end, travel back towards the start and are partially reabsorbed (since we never turn off g). Here we see the wavepackets travel much faster, at the speed of light in the link. Dynamics become messy since we never ‘turn off’ g and no pulse shaping is used. Middle column - intermediate regime. Dynamics are messy here and not useful for communication. We observe energy distributions in the link reminiscent of both cases.

Where $k = (j + 1)\pi/L$ is the wavevector for the j^{th} harmonic and L is the link length. We are also assuming ‘open’ terminations at either end of the link. This allows us to track the position of wavepackets in the link, particularly in the continuous regime where we observe ping-pong dynamics (as experimentally observed in [Zhong et al., 2019]) In our experiment, we are firmly in the single mode regime with $g/\Delta_{\text{FSR}} \sim 10^{-4}$

7.4 Heralded Schemes

Qubit platforms such as photonics, trapped ions, neutral atoms, and NV centers often couple to an optical fiber for long distance quantum communication. Traditionally, the energy transfer efficiency is low. Sometimes this is due to energy absorption in the fiber itself, although the attenuation length can be 10s of km for ‘telecom wavelength’ photons close to 1500 nm in wavelength. Unlike microwave platforms, the coupling efficiency to the fiber link is also low, due to the need to collect photons which may be emitted in all directions in 3D space (we can boost this by emitting photons into a photonic waveguide) and couple these photons into an optical fiber.

Heralded schemes overcome the fidelity limitation imposed by deterministic qubit transfer — a high single pass loss due to coupling inefficiencies (and often fiber absorption too). The goal in heralded schemes is to first herald an entangled state, and then use this as a resource for something else, e.g., qubit teleportation. Heralding relies on a measurement, typically a Bell measurement and the act of measuring is integral to producing the entangled state — even if there were no errors, the scheme could not produce entanglement without this measurement. We briefly walk through how this is typically achieved in optics, before we subvert the idea of heralding when we apply it to a microwave link.

We first prepare two qubits at nodes 1 and 2 in the product state $(|0\rangle + |1\rangle)_1 \otimes (|0\rangle + |1\rangle)_2$ where we are omitting normalization factors for brevity. We then engineer our systems so each qubit emits a photon, with a polarization that depends on the state of the qubit, e.g.,

$$(|0V\rangle + |1H\rangle)_1 \otimes (|0V\rangle + |1H\rangle)_2 \quad (7.13)$$

We now have two local Bell states. The idea is to perform a destructive Bell measurement on the photons, in order to entangle the qubits with each other. This is an

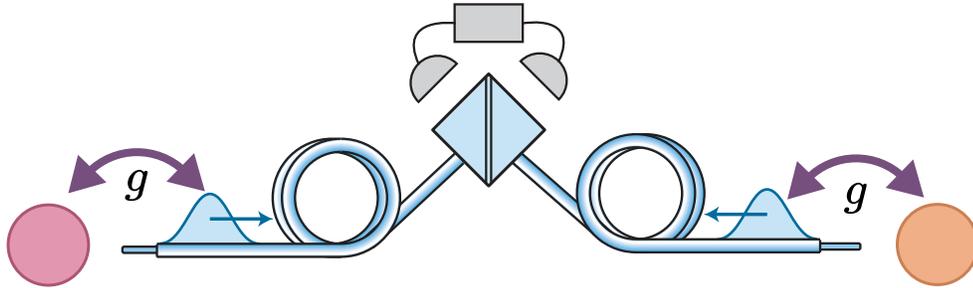


Figure 7.5: Typical setup for heralding entanglement in an optical quantum network. Here an optical fiber typically functions as the communication channel and the protocol begins by launching propagating wavepackets into the fiber from either end, where each wavepacket is entangled with the state of the qubit at each node. The wavepackets meet at some intermediate point along the link where a Bell measurement is performed, typically by interfering them on a beamsplitter and measuring them with photodetectors shown in grey. A particular measurement outcome heralds entanglement.

example of entanglement swapping.

As shown in Fig. 7.5, this can be achieved by sending both photons down an optical fiber and interfering them on an actual beamsplitter. There are variations on how this measurement is specifically implemented [Browne and Rudolph, 2005] but this beamsplitter can essentially act as a which-path information eraser, e.g., if we detect a photon of a particular polarization, we have no idea which qubit it came from. A certain pattern of clicks on the photodetectors heralds successful entanglement, leaving the qubits in a known Bell state. By ensuring a signal that is unlikely due to environmental noise (e.g., dark counts), such as two simultaneous clicks on two photodetectors, the entanglement fidelity between the qubits can be high fidelity, with entanglement fidelities in excess of 90% reported in trapped ion and NV center platforms [Hensen et al., 2015, Stephenson et al., 2020], even when the success probability per attempt is small. ($\sim 10^{-4}$ or below.)

7.5 Review of Previous Experiments

In Tab. 7.1 we summarize the experimental implementations of these various communication schemes over the years, restricting ourselves to microwave links between superconducting qubits. To compare both deterministic and probabilistic schemes we use the metric of entanglement fidelity — the fidelity of the Bell state these schemes can prepare between two nodes. In addition, we also compare them based on their success probability, energy transfer efficiency and repetition rate.

Work	Coupling to link (MHz)	Link loss rate (MHz)	Physical link	Channel type	Name of scheme	Transfer efficiency η	Success prob	Scheme duration	Bell state fidelity \mathcal{F}
Narla (2016)	-	-	copper coax cable, 180° hybrid, w/ circulators	Continuum	'optical' heralding	-	0.004	2.5 μ s	57 \pm 1%
Kurpiers (2018)	10.5	-	0.8 m NbTi Coax w/circulator	Continuum	CZKM	0.675	1.0	180 ns	78.9 \pm 0.1%
Campagne-Ibarcq (2018)	< 0.7	-	1 m NbTi Coax w/circulator	Continuum	CZKM	0.7	1.0	2.5 μ s	73%
Axline (2018)	< 0.4	-	0.6 m NbTi Coax w/circulator	Continuum	CZKM	0.74	0.78 (1.0)	6 μ s	77 \pm 2% (61 \pm 2%)
Kurpiers (2019)	5.6	-	0.9 m NbTi Coax w/isolator	Continuum	CZKM Time-Binned	0.72	0.615	480 ns	87%
Leung, Lu (2019)	1.3	0.29	1 m copper Coax hybridized w/on-chip resonators	Single mode	Half-swap, swap	0.61	1.0	265 ns	79.3 \pm 0.3%
Burkhart (2020)	0.56	0.1	6.6 cm NbTi Coax	Single mode	Detuned beamsplitter	0.84	0.79 (1.0)	520 ns	94 \pm 1% (88 \pm 1%)
Magnard (2020)	20	-	5 m WR90 Al waveguide w/circulator+ 0.8 m copper Coax	Continuum	CZKM	0.675	1.0	311 ns	79.5%
Zalys-Geller thesis (2021)	< 0.2	-	1 m NbTi Coax w/circulator	Continuum	CZKM Time-Binned	0.64	0.27	3 μ s	89 \pm 1%
Zhong (2021)	2.9,6.1	0.34	1 m NbTi Coax (wirebond)	Single mode	ST/2	0.881	1.0	62.8 ns	90.8 \pm 1.2%
Liu, Zhong (2023)	5	0.006	25 cm Al Coax (wirebond)	Single mode	Half-swap, swap	0.99	1.0	25 ns	98.9 \pm 0.6%
Storz (2023)	20	-	30 m WR90 Al waveguide w/circulator+ 0.8 m copper Coax	Continuum	CZKM	0.669	1.0	107 ns	80.4%
Teoh (2023)	0.16	0.6	6.7 cm NbTi Coax	Single mode	Dark Mode Msmt	-	0.35	4.8 μ s	92 \pm 1%

Table 7.1: Comparison of deterministic and probabilistic entanglement generation experiments in superconducting qubits across a quantum network. For our scheme (bottom row) we take the scheme duration to be 4.8 μ s for comparison, which excludes the time for state preparation.

7.6 Qubit Transfer in the ‘Impossible’ Regime

Is it possible to transfer a qubit in the regime $g < \kappa < \Delta_{\text{FSR}}$? In a way, this is the easiest regime to realize in an intrafridge microwave link. The link is lossy and the coupling rate is small. However, since $\bar{\eta} \sim \kappa/g > 1$, it seems that at first glance, all the energy will be lost in the link, and we will not be able to transfer any energy from node 1 to node 2 and we must resort to a slower adiabatic scheme or a heralded scheme with a vanishingly small success rate that scales as $e^{-\kappa/g}$. The goal of our experiment is to show that it is in fact possible to transfer a qubit in this regime without an adiabatic slowdown, by developing a new heralded scheme + teleportation. Even if we are not in the $\kappa > g$ limit, this scheme is still able to get around the $\bar{\eta} = \kappa/g$ ‘fidelity barrier’ that non-adiabatic deterministic schemes face. The price we pay is that our scheme to generate entanglement will succeed with at most 50% probability per attempt

To build some intuition, it is helpful to first consider the case where we try to transfer a single photon with the hybrid scheme in the case where $\kappa_b > g_0$. Something curious happens in the system — we do not decay to vacuum but instead reach a steady state which contains half the initial energy. Using the notation $|n_1\rangle_1 \otimes |n_2\rangle_2 \otimes |n_b\rangle_b = |n_1, n_2, n_b\rangle$, this state can be written as

$$\rho_{\text{steady}} = \frac{1}{2} |0, 0, 0\rangle \langle 0, 0, 0| + \frac{1}{2} \left(\frac{|1, 0, 0\rangle - |0, 1, 0\rangle}{\sqrt{2}} \right) \left(\frac{\langle 1, 0, 0| - \langle 0, 1, 0|}{\sqrt{2}} \right) \quad (7.14)$$

which curiously is a 50-50 incoherent mixture of vacuum and a Bell state, so we can in fact retain up to half the energy if we arrange our couplings right, even though $\kappa/g > 1$. If we had some way of distinguishing between these states, we could obtain a Bell state with 50% probability. However, in this case, there are no *local* measurements that can achieve this — we would have to do some non-local measurement between node 1 and node 2, but if we can do this, then we could just use this measurement

instead to make our entangled state! This state has a 50% fidelity to a Bell state, and so it has no useful entanglement. Nonetheless it is instructive to see why we reach this steady state.

We can rewrite the Hamiltonian from Eq. 7.6 as

$$\hat{\mathcal{H}}_c = g_{\text{BS}} (\hat{a}_1 + \hat{a}_2) \hat{b}^\dagger + \text{h.c.} = \sqrt{2} g_{\text{BS}} (\hat{a}_b \hat{b}^\dagger + \hat{a}_b^\dagger \hat{b}) \quad (7.15)$$

Where we have redefined g_{BS} to the ‘experimentalist’ definition, where $g_{\text{BS}} = g_0/2$ where we define the bright mode to be $\hat{a}_b = (\hat{a}_1 + \hat{a}_2)/\sqrt{2}$ which couples to the bus and the ‘dark mode’ $\hat{a}_d = (\hat{a}_1 - \hat{a}_2)/\sqrt{2}$, which does not.

We define ‘bright states’ to be excitations of the bright mode, i.e., they can be written as

$$|\psi\rangle_{\text{bright}} = \sum_n c_n (\hat{a}_b^\dagger)^n |\text{vac}\rangle \quad (7.16)$$

for scalar coefficients c_n and similarly for ‘dark states.’

Dissipation on the bus mode given by the collapse operator \hat{b} at rate κ_b for evolution under the Lindblad master equation. Only the bright state feels the effects of this dissipation.

We can now rewrite our initial single photon state as a superposition of a bright and dark state.

$$|\psi_{\text{init}}\rangle = |1, 0, 0\rangle = \frac{|1, 0, 0\rangle + |0, 1, 0\rangle}{\sqrt{2}} + \frac{|1, 0, 0\rangle - |0, 1, 0\rangle}{\sqrt{2}} \quad (7.17)$$

In the steady state, the bright state has decayed to vacuum and the dark state remains unaffected by the dissipation, leaving us in the state in Eq. 7.14.

7.6.1 Classical bright and dark states

It is possible to find a local measurement that can distinguish between vacuum and a dark state but to do this, we must prepare our bright and dark states within a larger Hilbert space. We can examine what happens when we prepare coherent states, which have a small overlap with the vacuum state provided α is large enough². ‘Classical’ bright states take the form $|\beta, \beta, 0\rangle$ we call this a classical state because it is classically analogous to two harmonic oscillators displaced in phase with the same amplitude. It is a bright state of the Hamiltonian since we can write this state as

$$|\beta, \beta, 0\rangle = \widehat{\mathcal{D}}_{\hat{a}_b}(\sqrt{2}\beta) |\text{vac}\rangle \quad (7.18)$$

Where $\widehat{\mathcal{D}}_{\hat{a}_j}(\alpha)$ is the displacement operator acting on mode \hat{a}_j with a displacement α .

Under the evolution of the Lindblad master equation, we will eventually lose all the energy in the bus mode and reach the vacuum state (See Fig. 7.8). This is true for any value of β . We can also prepare the corresponding dark state

$$|\beta, -\beta, 0\rangle = \widehat{\mathcal{D}}_{\hat{a}_d}(\sqrt{2}\beta) |\text{vac}\rangle \quad (7.19)$$

Regardless of our choice of β , this state never participates in the dynamics of Eq. 7.6, and it never loses any energy to the bus mode.

So far all that we have discussed is a classical interference effect that arises in coupled harmonic oscillators. The classical analog for this system is three coupled pendulums as shown in Fig. 7.6, where only the middle pendulum is lossy.

2. The overlap, $|\langle \alpha|0\rangle|^2 = e^{-|\alpha|^2}$ decreases exponentially to zero as $|\alpha|^2$ increases

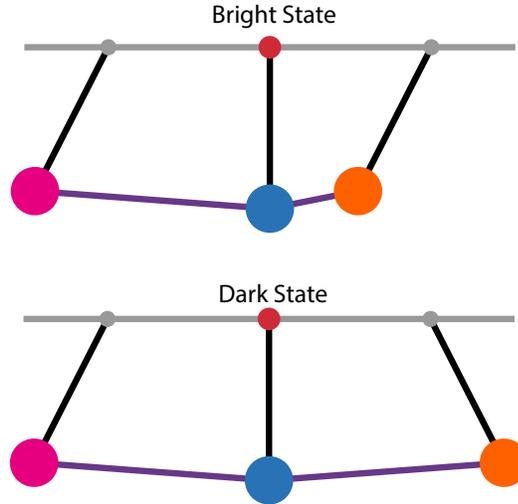


Figure 7.6: A classical pendulum analog for our system. Three coupled pendulums coupled together by elastic links (purple). Bright states correspond to the left and right pendulums displaced in phase. They excite the middle pendulum which has a lossy pivot shown in red, where eventually all the energy in the system is dissipated. The dark state is the left and right pendulum displaced in antiphase. The forces on the central pendulum from the purple links always cancels out so it never moves and no energy is dissipated in the system.

7.6.2 The Dark Mode Measurement scheme

By combining this classical interference effect with an initial state that is a quantum superposition of these bright and dark coherent states, we can reach a steady state from which we can herald a Bell state with local measurements with a 50% success probability. We call our scheme the ‘Dark mode measurement’ (DMM) scheme since ultimately we are performing a measurement which distinguishes between dark modes and bright modes of the system.

The scheme is simple to explain for the lossless case, where $\kappa_b = 0$. We start with two local ‘cat states’ in qubit 1 and qubit 2, which are now taken to be bosonic modes (cavity qubits). The initial state is

$$(|\alpha\rangle + |-\alpha\rangle)_1 \otimes (|\alpha\rangle + |-\alpha\rangle)_2 \otimes |0\rangle_b = |\alpha, \alpha, 0\rangle + |-\alpha, -\alpha, 0\rangle + |\alpha, -\alpha, 0\rangle + |-\alpha, \alpha, 0\rangle \quad (7.20)$$

Where we have omitted the normalization factors. After applying the Hamiltonian $\hat{\mathcal{H}}_c$ for time $t_{\text{SWAP}} = \pi/2\sqrt{2}g_{\text{BS}}$ we end up with the state

$$|0, 0, \sqrt{2}\alpha\rangle + |0, 0, -\sqrt{2}\alpha\rangle + |\alpha, -\alpha, 0\rangle + |-\alpha, \alpha, 0\rangle \quad (7.21)$$

Unlike the state in Eq. 7.14, this state is a pure state (no dissipation yet) but crucially, we can now distinguish the dark state from vacuum by local measurements. We can perform a QND measurement that tells us whether one of the qubits is in vacuum or not, which can be achieved with a number-selective π -pulse on a dispersively coupled transmon ancilla. We call this measurement the ‘vacuum check’. 50% of the time, we find the qubits to not be in vacuum and project into the state

$$|\Psi_+^{2\text{cat}}\rangle \approx \frac{1}{\sqrt{2}} |\alpha, -\alpha, 0\rangle + |-\alpha, \alpha, 0\rangle \quad (7.22)$$

which is a Bell state in the two-legged cat code provided $\alpha \gg 0$. This Bell state is also sometimes referred to as the ‘cat in two boxes’ state [Wang et al., 2016], although now these boxes can live in separate modules connected by a link. When we obtain the other outcome, we know the cavities are both empty and have transferred all their energy to the bus mode. Before retrying the scheme, we must be sure the bus mode is in vacuum.

The fact that this scheme only succeeds 50% of the time in the lossless case is reminiscent of some heralded entangling schemes in optics, where the Bell measurements can only succeed at most 50% of the time as well. Unlike continuum schemes in optics, there is no interference of wavepackets in the middle of the link, the interference happens immediately due to the standing wave nature of our microwave link. Dissipation in the bus mode is not strictly necessary for this scheme to work, although it is helpful for bus reset in between attempts.

7.6.3 The DMM scheme is robust to bus loss

The remarkable feature of this scheme is that it proceeds in more or less the same way even when there is significant energy loss in the bus. Regardless of the value of κ_b , the dark modes are always static under $\hat{\mathcal{H}}_c$ and hence loss in the bus does not degrade entanglement fidelity. If anything, a larger κ_b helps to quickly reset the bus mode after the failed attempts. However, t_{SWAP} may become prohibitively long if we are in the regime where $\kappa_b \gg g_{\text{BS}}$.

We can ask the question how long does it take for a bright mode to reach vacuum in the cavities? We must apply $\hat{\mathcal{H}}_c$ for this time before proceeding with the vacuum check measurement. This time depends on the ratio of g_{BS} to κ_b . As shown in Fig. 7.8, the functional form of the decay depends on whether we are in the ‘underdamped’ regime, where $g_{\text{BS}} > \kappa_b$, the overdamped regime where $g_{\text{BS}} < \kappa_b$ or the critically damped regime where $g_{\text{BS}} = \kappa_b/4\sqrt{2}$ which is when the energy decays to vacuum the fastest.

We can perform the vacuum check before all the energy has decayed if we started in a bright state. Instead, we only require the cavities to reach vacuum. In the underdamped regime this happens in time t_{SWAP} , which is the shortest possible duration for the DMM scheme and is independent of κ_b . If we are in the critically damped or overdamped regime, we will reach the steady state

$$\rho_{\text{steady}} = \frac{1}{2} |000\rangle \langle 000| + \frac{1}{2} |\Psi_+^{2\text{cat}}\rangle \langle \Psi_+^{2\text{cat}}| \quad (7.23)$$

from which we can succeed in creating the pure Bell state $|\Psi_+^{2\text{cat}}\rangle$ with 50% probability using the same vacuum check measurement. The DMM scheme is summarized in Fig. 7.7

The decay of bright states can again be analyzed in the Heisenberg picture for the

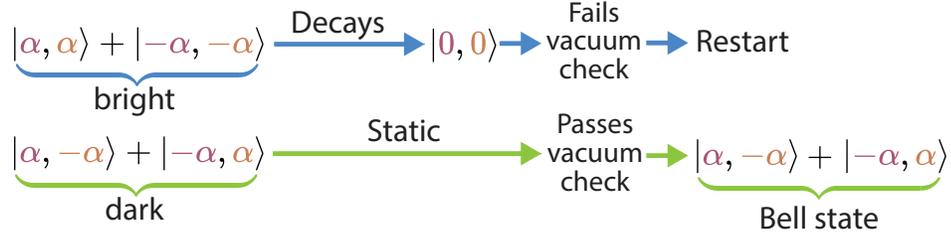


Figure 7.7: The Dark Mode Measurement (DMM) scheme for a lossy bus. We start off with a superposition of bright and dark coherent states that initially have no entanglement. After turning on our couplings to the bus for long enough, the bright state decays to the vacuum but the dark states are untouched. We can perform a QND measurement called the vacuum check that determines if the cavities are in vacuum or not. We declare success (with up to 50% probability) if we detect the cavities are not in vacuum. We have chosen our initial states carefully such that the dark states are also a Bell state. Thus, when we pass the vacuum check we herald a Bell state.

three field operators $\hat{a}_d(t)$, $\hat{a}_b(t)$ and $\hat{b}(t)$

$$\begin{aligned}
 \dot{a}_d(t) &= 0 \\
 \dot{a}_b(t) &= i\sqrt{2} g_{\text{BS}} b(t) \\
 \dot{b}(t) &= -i\sqrt{2} g_{\text{BS}} a_b(t) - \frac{\kappa_b}{2} b(t)
 \end{aligned} \tag{7.24}$$

We numerically solve for the dynamics in Fig. 7.8.

7.6.4 Fundamental loss limits to the DMM scheme

At first glance, it appears the DMM scheme is completely robust to bus loss, but is this actually true? Since this is a heralded scheme no single pass loss limit applies but we can identify two limitations that affect any experimental implementation, even if our cavities were completely error-free.

The first is the fact that any physical link must support multiple standing wave modes and it is impossible to only couple to one of these modes. We rewrite Eq. 7.15 to include the first two adjacent harmonics \hat{b}_n and \hat{b}_m , neglecting other harmonics

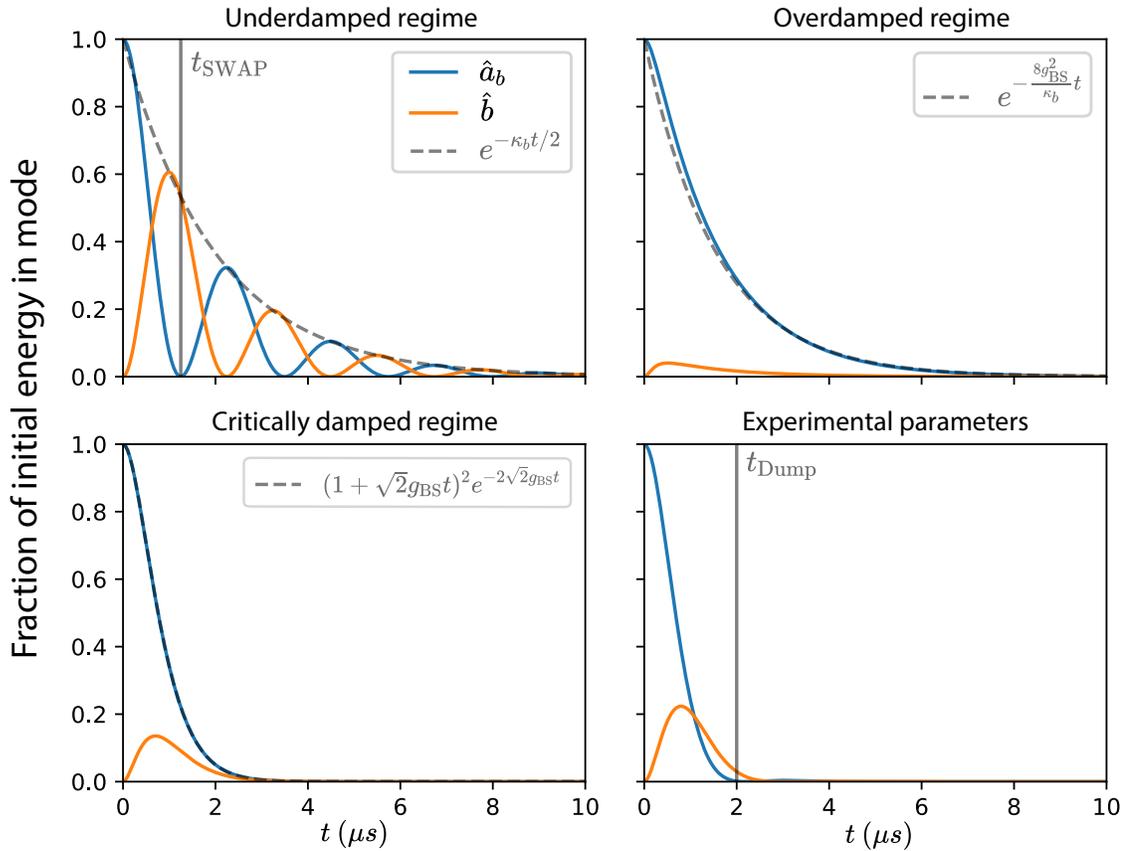


Figure 7.8: The DMM scheme in different bus loss regimes. The ratio of κ_b to g_{BS} determines how long we must apply until we can perform the vacuum check, and also how long we must wait for the bus mode to reset between failed attempts. We generate dynamics from Eq. 7.24 with initial conditions $a_d(0) = 0$, $a_b(0) = 1$, $b(0) = 0$. We fix $g_{BS}/2\pi = 160$ kHz, the value we will use in the experiment and set κ_b to be 160 kHz (underdamped), 2000 kHz (overdamped), 905 kHz (critically damped) or 600 kHz, as used in the experiment where we wait time $t_{\text{dump}} = 2$ μs until the vacuum check. Dashed lines are expressions for the energy decay rate of the bright mode, given by the analytical solutions to Eq. 7.24. Critical damping exhibits the fastest energy decay from the system. The shortest wait time, t_{SWAP} , is achieved in the underdamped regime.

which are further off-resonance.

$$\hat{\mathcal{H}}_c = \sqrt{2}g_{\text{BS}} \left(\hat{a}_b \hat{b}^\dagger + \hat{a}_d \hat{b}_n^\dagger + \hat{a}_d \hat{b}_m^\dagger + \text{h.c.} \right) + \Delta_{\text{FSR}} \left(\hat{b}_n^\dagger \hat{b}_n - \hat{b}_m^\dagger \hat{b}_m \right) \quad (7.25)$$

Although the dark mode, \hat{a}_d does not couple to the mode \hat{b} at all, it *does* couple to the adjacent harmonics albeit off-resonantly since they are detuned by $\pm\Delta_{\text{FSR}}$. This coupling happens because the parity of each consecutive harmonic alternates signs.

The limit on $\bar{\eta}$ due to this effect will scale as the time of the protocol, which is at minimum t_{SWAP} , the loss rate in these adjacent harmonics which we take to also be κ_b and the population in these modes, which scales as $\left(\frac{g_{\text{BS}}}{\Delta_{\text{FSR}}}\right)^2$ due to the off-resonant coupling. From this we find

$$\bar{\eta} \sim \frac{\kappa_b}{g_{\text{BS}}} \left(\frac{g_{\text{BS}}}{\Delta_{\text{FSR}}} \right)^2 \sim \left(\frac{g_{\text{BS}}}{\Delta_{\text{FSR}}} \right) \left(\frac{\kappa_b}{\Delta_{\text{FSR}}} \right) \quad (7.26)$$

which is a factor $\left(\frac{g_{\text{BS}}}{\Delta_{\text{FSR}}}\right)$ smaller than the single pass loss limit. For our experimental parameters, we will expect this contribution to be negligible at the 10^{-6} level.

Is this still true in the large loss limit? Do we do better than adiabatic schemes? Energy now decays from the system at the slower rate $8g_{\text{BS}}^2/\kappa_b$. Suppose we had to wait 8 time constants³ until performing the vacuum check, giving our scheme a duration $\sim \kappa_b/g_{\text{BS}}^2$. This would set

$$\bar{\eta} \sim \frac{\kappa_b^2}{g_{\text{BS}}^2} \left(\frac{g_{\text{BS}}}{\Delta_{\text{FSR}}} \right)^2 \sim \left(\frac{\kappa_b}{\Delta_{\text{FSR}}} \right)^2 \quad (7.27)$$

which still beats the single pass loss!

We can simulate the effects of off-resonant couplings with Eq. 7.11 in the regime where Δ_{FSR} and g_{BS} are within an order of magnitude of each other to magnify this effect. The energy distribution along the link shows small perturbations due to these

3. Chosen to simplify the maths.

off-resonant couplings as shown in Fig. 7.9

The other effect we consider is the Purcell limit of the cable on the cavity qubits. In the dressed mode basis, the mode we label the cavity mode (at each node) will have a small participation in the bus mode, mediated by the coupler. In the case where the coupler is dispersively coupled to both a cavity mode and the bus mode, as is the case for our experiment, we will experience an additional loss rate on each cavity mode given by

$$\kappa_{\text{cav}}^{\text{Purcell}} = \frac{\chi_{a_i t} \chi_{b t}}{\alpha_t \alpha_t} \kappa_b \quad (7.28)$$

Where α_t is the anharmonicity of our transmon couplers, $\chi_{a_i t}$ is the dispersive coupling between the transmon and the i^{th} cavity and $\chi_{b t}$ is the dispersive coupling between the bus mode and the transmon. For our couplers, $\chi/\alpha_t \sim 0.01$ and results in a minimum $\bar{\eta}$ of

$$\bar{\eta} = \frac{\kappa_{\text{cav}}^{\text{Purcell}}}{g_{\text{BS}}} \quad (7.29)$$

For our experimental parameters, we estimate this to contribute at the 10^{-4} level, larger than the off-resonant coupling contribution but still too small to be measurable. However, they illustrate that the intrinsic limit to the DMM scheme is small enough such that in reality, it is the local errors in the cavities that limits the scheme.

This includes things such as SPAM errors, photon loss in the cavities, errors in engineering g_{BS} coupling and in performing the vacuum check. The important thing to note is that we do not run into the κ/g barrier that typically limits deterministic qubit transfer schemes in microwaves. The rest of the chapter will be concerned with the details of experimental implementation of the DMM scheme, and how the Bell state can then be used to transfer a qubit via teleportation.

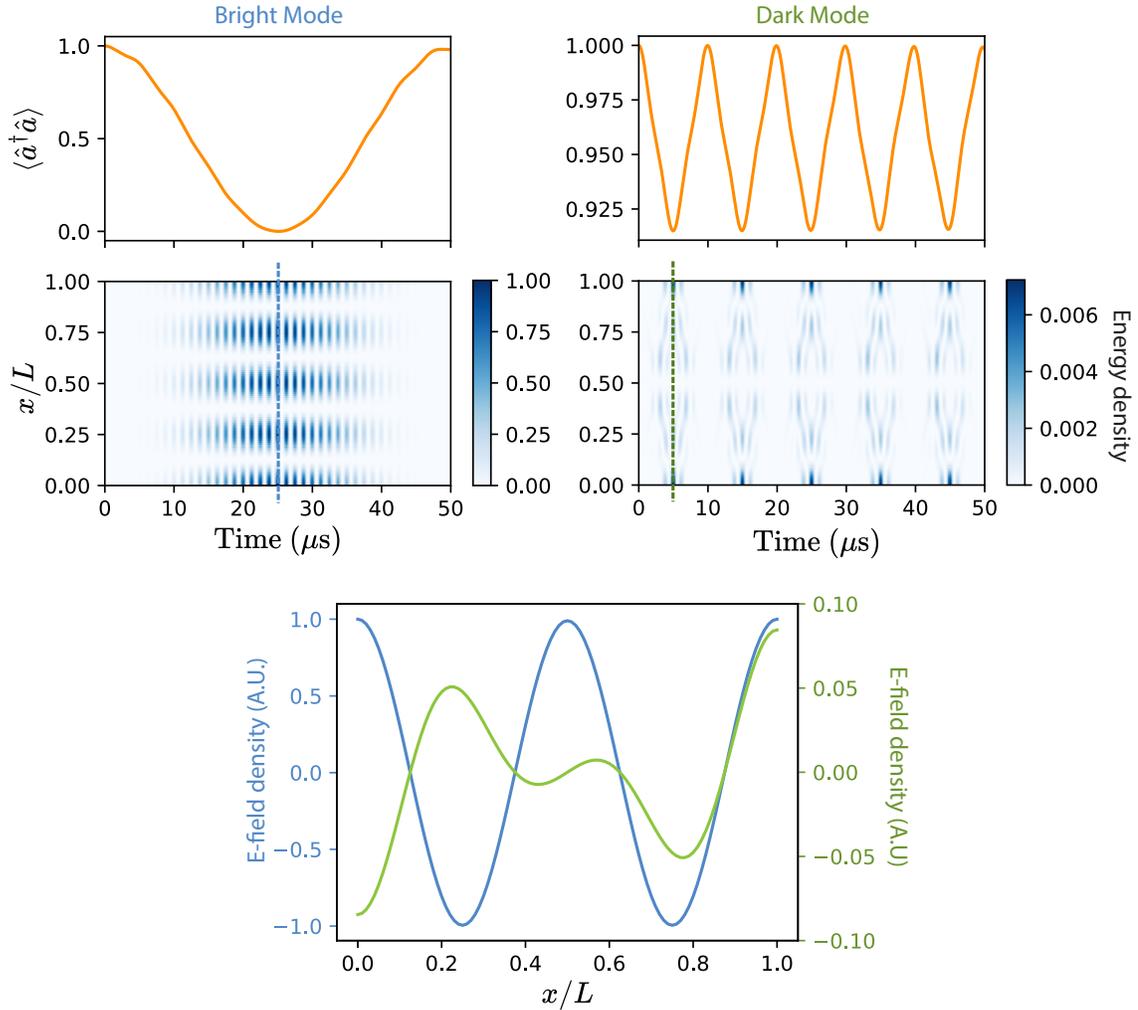


Figure 7.9: The effect coupling to off-resonant harmonics in the DMM scheme. Here we take $\Delta_{\text{FSR}}/2\pi = 100$ kHz and $g_{\text{BS}}/2\pi = 10$ kHz to intentionally exaggerate these effects. We initially prepare a single excitation in the bright or dark mode and model seven harmonic modes with the Hamiltonian in 7.11. We observe in the left column that the bright state more or less fully swaps to the bus mode in time t_{SWAP} , with the off-resonant couplings limiting the fraction of energy transferred. In the right column we observe faster dynamics and are able to excite the off-resonant modes in the bus, which would result in a loss of energy in the system. Note the different energy density scales for the bright and dark modes. At the bottom we plot the slices to show the E-field amplitude at points in time where the link contains the most energy. This shows how even if we start in a ‘dark state’ off-resonant couplings mean we can excite anti-symmetric modes in the link.

7.7 Making the Scheme Experimentally-friendly

One modification we make to the scheme to make our lives as experimentalists a little easier is to reduce the value of α . This is for three reasons. Firstly, the probability of photon loss in either of the cavities is

$$p_{\text{loss}} = 2|\alpha|^2\kappa\delta t \quad (7.30)$$

for cavity loss rate κ and short time interval δt . If we lose a single photon from either cavity, even if we began in the dark state, we will end up in the error state

$$|\Psi_-^{\text{2cat}}\rangle \approx \frac{1}{\sqrt{2}}|\alpha, -\alpha, 0\rangle - |-\alpha, \alpha, 0\rangle \quad (7.31)$$

which is the wrong Bell state. Reducing α reduces our susceptibility to this loss, essentially just by reducing the average number of photons in the system. The second reason is related to the self-Kerr of the cavities. We will control each cavity with its own dispersively coupled transmon and so we can expect self-Kerrs in the 1-10 kHz regime. This unwanted unitary evolution distorts our initial bright and dark states so they are not entirely bright or dark anymore. This happens at quite a slow rate, of order $K|\alpha|^2\delta t$ for typical cavity self-Kerr, K but can still be a source of error in our protocol.

Thirdly, we observe that the transmon couplers we use to engineer g_{BS} via a four-wave mixing process suffer from more anomalous heating when there are more photons in the cavities. This is thought to be due to unwanted multi-photon resonances [Xiao et al., 2023] which become harder to avoid with more photons in the system.

The DMM scheme can be modified to work at arbitrarily small α , and can produce *perfect* Bell states, provided we carefully update our logical qubit basis. But, as we decrease α , the success probability of the scheme also decreases towards zero. This

is because our dark states will now have a significant overlap with vacuum, which reduces the overall success probability of our scheme to

$$p_{\text{success}} = \frac{1}{2} (1 - 2|\langle 0|\alpha\rangle|^2 + |\langle 0|\alpha\rangle|^4) = \frac{1}{2} (1 - 2e^{-|\alpha|^2} + e^{-2|\alpha|^2}), \quad (7.32)$$

where we have used the formula $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ and we perform the vacuum check on both cavities (to reduce the effects of readout error).

The other consequence is that when we ‘pass’ the vacuum check the remaining dark state will have the vacuum ‘subtracted’ from it. If our target dark state would be $|\text{dark}\rangle$, then after passing the vacuum check we end up in the state

$$|\text{dark}'\rangle \propto \hat{\Pi}_{\bar{0}\bar{0}} |\text{dark}\rangle, \quad (7.33)$$

where $\hat{\Pi}_{\bar{0}\bar{0}}$ is the ‘not zero’ projector

$$\hat{\Pi}_{\bar{0}\bar{0}} = (\mathbb{1} - |0\rangle\langle 0|)_1 \otimes (\mathbb{1} - |0\rangle\langle 0|)_2 \otimes \mathbb{1}_b. \quad (7.34)$$

At first glance, removing vacuum would seem to reduce our entanglement fidelity, since it reduces our overlap with the target state, $|\Psi_+^{2\text{cat}}\rangle$. However, by redefining the basis of our bosonic encoding, we can show $|\text{dark}'\rangle$ is a perfect Bell state in this basis. The modified basis is

$$\begin{aligned} |+\rangle_L &\propto (\mathbb{1} - |0\rangle\langle 0|) (|\alpha\rangle + |-\alpha\rangle) \\ |-\rangle_L &\propto |\alpha\rangle - |-\alpha\rangle \end{aligned} \quad (7.35)$$

where the $|-\rangle_L$ state is unchanged from the usual two-legged cat code but we subtract vacuum from the state $|+\rangle_L$. It is easy to check these states are orthogonal since $|+\rangle_L$ contains only even photon numbers whilst $|-\rangle_L$ contains only odd. In Fig. 7.10. we show the Bloch sphere for the modified basis.

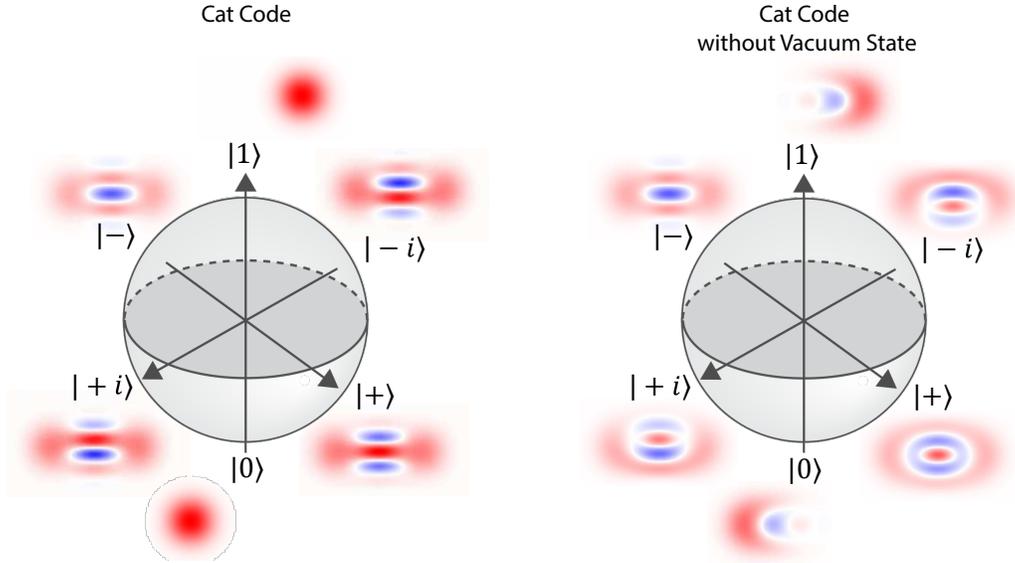


Figure 7.10: Modifying the bosonic code basis for small α . On the left we show the logical Bloch sphere with Wigner functions for the standard two-legged code and compare it to the Bloch sphere on the right which is for the modified basis. Here we set $\alpha = \sqrt{2}$. Note we now define this Bloch sphere such that the cat states with well-defined parity now lie on the equator.

Now we show how we end up in a perfect Bell state if our target state was initially the dark state

$$|\text{dark}\rangle \propto |\alpha\rangle_1 |-\alpha\rangle_2 - |-\alpha\rangle_1 |\alpha\rangle_2 \quad (7.36)$$

which we can rewrite as

$$|\text{dark}\rangle \propto (|\alpha\rangle + |-\alpha\rangle)_1 (|\alpha\rangle - |-\alpha\rangle)_2 \quad (7.37)$$

$$- (|\alpha\rangle - |-\alpha\rangle)_1 (|\alpha\rangle + |-\alpha\rangle)_2 \quad (7.38)$$

If we use the standard definition of the cat code

$$|\pm\rangle_{\text{cat}} = \frac{|\alpha\rangle \pm |-\alpha\rangle}{\sqrt{N_{\pm}}}, \quad (7.39)$$

where $N_{\pm} = 2(1 \pm e^{-2|\alpha|^2})$, then our initial dark state can be rewritten exactly as

$$|\text{dark}\rangle = \frac{|+\rangle_{\text{cat},1} |-\rangle_{\text{cat},2} - |-\rangle_{\text{cat},1} |+\rangle_{\text{cat},2}}{\sqrt{2}}. \quad (7.40)$$

It then follows that after applying the projector, $\hat{\Pi}_{\bar{0}\bar{0}}$ to this state we end up in the final dark state

$$|\text{dark}'\rangle = \frac{|+\rangle_{L,1} |-\rangle_{L,2} - |-\rangle_{L,1} |+\rangle_{L,2}}{\sqrt{2}} \quad (7.41)$$

which is indeed a perfect Bell state in the new basis. This is not true for all target dark states. For instance, if our target dark state were instead

$$|\text{dark}\rangle \propto |\alpha\rangle_1 |-\alpha\rangle_2 + |-\alpha\rangle_1 |\alpha\rangle_2, \quad (7.42)$$

which has a subtle sign flip in the superposition then

$$|\text{dark}\rangle \approx \frac{|+\rangle_{\text{cat},1} |+\rangle_{\text{cat},2} - |-\rangle_{\text{cat},1} |-\rangle_{\text{cat},2}}{\sqrt{2}}, \quad (7.43)$$

is only an approximate Bell state, owing to $N_+ \neq N_-$. This cannot be fixed by a simple change of basis. In the experiment we will prepare the initial state

$$(|\alpha\rangle + i|-\alpha\rangle)_1 \otimes (|\alpha\rangle - i|-\alpha\rangle)_2 \otimes |0\rangle_b \quad (7.44)$$

which is another possible starting state which gives us the same desired dark state as in Eq. 7.36. Experimentally, we are willing to trade the success probability of our scheme in order to work with smaller α entangled states, which ultimately leads to higher entanglement fidelities.

What happens as $\alpha \rightarrow 0$? Although the success probability drops dramatically to $p_{\text{success}} \rightarrow |\alpha|^4/2$, the modified basis DMM scheme still produces perfect entanglement even as $\alpha \rightarrow 0$, provided that there is non-zero population in the $|2\rangle$ state initially.

(The DMM scheme does not work for two-level systems, but it does for some three-level systems.) In this limit, $|+\rangle_L \rightarrow |2\rangle$ and $|-\rangle_L \rightarrow |1\rangle$ which begins to be apparent if one squints at the Wigner functions in Fig. 7.10.

7.8 cQED Hardware for the DMM Scheme

Now we can finally get stuck into a *real* experiment! The minimalist diagram for the cQED hardware we use is shown in Fig. 7.11 and 7.13 and consists of two modules joined together by a short, 6.7 cm long Niobium Titanium (NbTi) superconducting coaxial cable with a Teflon dielectric. (SC-086/50-NbTi-NbTi From Coax Co, Ltd., Japan).

Each module consists of a 3D stub cavity where the bosonic qubit is encoded in the $\lambda/4$ fundamental mode. Additionally, a single chip contains both a transmon and readout resonator. The transmon is dispersively coupled to the cavity mode and serves a multitude of purposes.

Firstly, it provides the requisite non-linearity via the dispersive coupling with the cavity which enables the preparation of cat states in the cavity mode. Secondly, by applying two off-resonant microwave pump tones we can actuate tunable beamsplitter couplings between the cavity and the bus mode via a four-wave mixing process [Zhang et al., 2019, Gao et al., 2018, Burkhardt et al., 2021] to engineer the Hamiltonian $\hat{\mathcal{H}}_c$. Thirdly, number selective π -pulses applied to the transmon suffice as vacuum check measurement and finally, the transmons can be used to perform two-qubit quantum state tomography to measure the Bell-state fidelity and transfer fidelity.

These modules are the same as physical modules used in [Burkhardt et al., 2021]. The cavities remain the same but the chips and cable are new. The cavities, which we refer to as cavity 1 and cavity 2 are etched 99.999% purity Aluminum stub cavities with lifetimes of 385 μs and 520 μs respectively, two orders of magnitude longer than

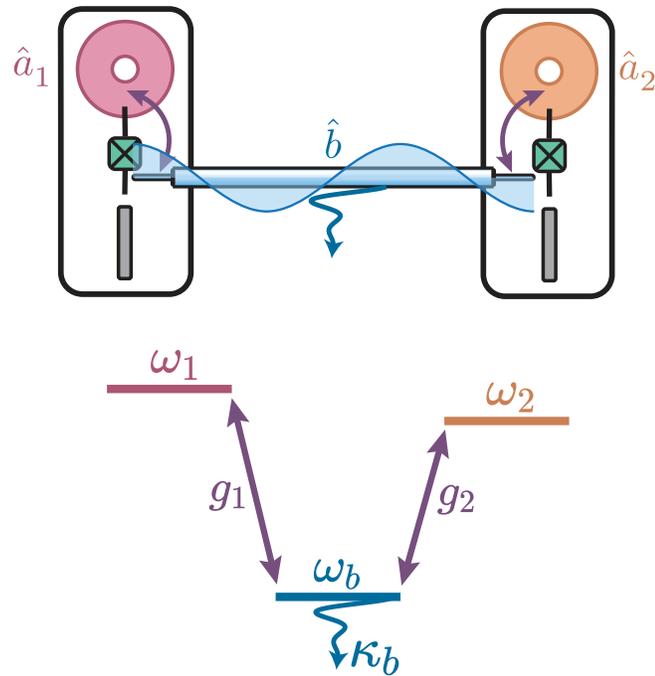


Figure 7.11: Minimalist layout of the cQED hardware needed for the DMM scheme. The hardware must be capable of four main tasks: preparation of cat states, actuating the Hamiltonian $\hat{\mathcal{H}}_c$, performing the vacuum check measurement and joint-cavity tomography to verify entanglement. **(Top)** cavity modules consist of a 3D stub cavity (pink or orange), a transmon (green) and a readout resonator (grey). With 4-wave mixing, we can actuate parametric beamsplitter couplings (purple), g_1 and g_2 between each cavities and a cable harmonic we designate the bus mode. **(Bottom)** Energy level diagram for the three coupled bosonic modes.

the duration of our DMM scheme. This ensures errors due to cavity photon loss during the scheme is small (at the few percent level). Their mode frequencies are 6509 MHz and 6487 MHz respectively and no direct coupling between the cavities is observed either when idling the system or when applying microwave drives.

Each transmon has its own readout resonator and Purcell filter resonator (both meandered) fabricated on the same chip. Each transmon has a ‘standard’ dispersive readout where single shot readout of the transmon state can be achieved within 1.5 μ s.

The cable is modified in two ways. Firstly, we strip back the outer conductor and dielectric by around 8 mm such that only the inner conductor protrudes far into the tunnel. The capacitive coupling between the inner conductor and the transmon pads allows us to engineer a dispersive coupling between the standing modes of the cable and the transmon mode. The outer conductor is bolted into the outer conductor of the package with a single brass screw (See [Burkhart et al., 2021]) tipped with indium in an attempt to reduce the joint-resistance. It is thought that this joint resistance is what currently limits the T_1 times of the cable harmonics. This simple clamping technique is easily reversible — helpful for exchanging parts in a modular device.

We use the $3\lambda/2$ harmonic of the cable as our bus mode, with a frequency of 5658 MHz. The loss rate of these superconducting cable modes is typically ~ 100 kHz with Quality factors between 10,000 – 100,000 having been reported [Kurpiers et al., 2017, Burkhart et al., 2021]. Recently, Quality factors on the order of 1 million have been reported in Al-Al coaxial cable [Niu et al., 2023], although requires careful engineering to minimize joint resistance.

7.8.1 Purcell filter for the bus mode

For purely technical reasons, we must do some more engineering to further increase the loss rate of the bus mode. The reason for this stems from our design choice to minimize hardware components and use a single transmon in each module to

do four different tasks. Three of these tasks could be in principle outsourced to another transmon, in the currently unused chip tunnel (See Fig. 7.13). These are state preparation, the vacuum check and cavity tomography. For all these tasks, we ideally want a high coherence transmon to maximize the fidelity of the vacuum check and minimize SPAM errors. The fourth task, engineering the beamsplitter coupling, would have its own dedicated coupling element.

One subtlety of having each transmon be dispersively coupled to its cavity mode *and* the bus mode is that we must wait for any bright states to fully decay to vacuum in both the cavities and the bus mode — we do not want to operate in the underdamped regime of Fig. 7.8. After time t_{SWAP} , the cavities will be empty but the bus mode will not, and our selective π -pulse will fail to excite the transmon to $|e\rangle$. Waiting for the bus mode to decay to vacuum increases the duration of our scheme, which increases our susceptibility to Kerr and photon loss in the cavities. Instead, we opt to increase the loss rate of the bus mode, aiming for the critical damping regime.

Now, the simplest thing would seem to be to use a more lossy material for the cable, such as copper which typically has a Q factor of around 5,000. However, this has a detrimental effect on the transmon (and possibly the cavity) via the inverse Purcell effect. We would observe a drop in the transmon T_1 which cause problems for all three tasks.

Instead our solution is a little contrived but does the job. We increase the loss rate of the bus mode by creating a 3D Purcell filter (to my knowledge the first of its kind) which functions to increase the loss rate of the bus mode, but protects the transmon from the Purcell effect. This functions in the exact same way as a Purcell filter for the readout resonator [Reed et al., 2010], except now instead of the readout resonator mode it is the bus mode.

We need to create another lossy mode, which we call the Purcell mode, that is strongly coupled to the bus mode with a small detuning. We make this Purcell mode

out of a 6061 Al stub cavity that has its center ‘stub’ galvanically coupled to a 50 Ω terminated copper coax cable. This stub is also made of copper and is what we typically use for the ‘drive pins’. The construction of this Purcell filter is shown in Fig. 7.12

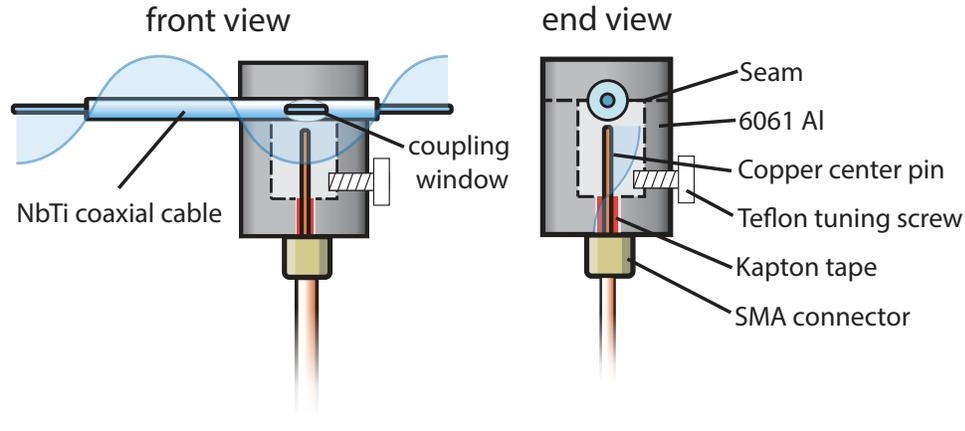


Figure 7.12: Design of a 3D Purcell filter for the bus mode. A window is drilled through the outer conductor to expose the inner conductor. The Purcell filter mode is the $\lambda/4$ mode of a microwave stub cavity resonator, which capacitively couples to the bus mode at an E-field anti-node. The frequency of this mode is set by the length of the copper center pin but can be fine-tuned via the Teflon tuning screw. The copper center pin directly protrudes from an SMA flange connector, meaning the purcell filter mode is galvanically coupled to the 50 Ω copper coax transmission line. However, the copper pin passes through a narrow tunnel in between the stub cavity and the SMA flange. The sudden change in tunnel width creates an impedance mismatch that boosts the Quality factor of the Purcell filter mode. Kapton tape prevents the copper center pin from shorting with the 6061 Aluminum housing.

Despite the galvanic coupling to ground, the large impedance mismatch means that the $\lambda/4$ mode of this stub cavity has a quality factor of $Q = 170 \pm 10$ when measured at room temperature on a VNA. By adjusting the Teflon tuning screw we decrease the frequency of this mode by increasing the mode participation in the Teflon, which has $\epsilon_r \approx 1.7$. This lets us tune the frequency of this mode to $\omega_{\text{filter}} = 5688$ MHz, 30 MHz above the bus mode frequency (although we cannot measure the Purcell mode frequency at cryogenic temperatures and it likely increases slightly due to the Teflon shrinking in size).

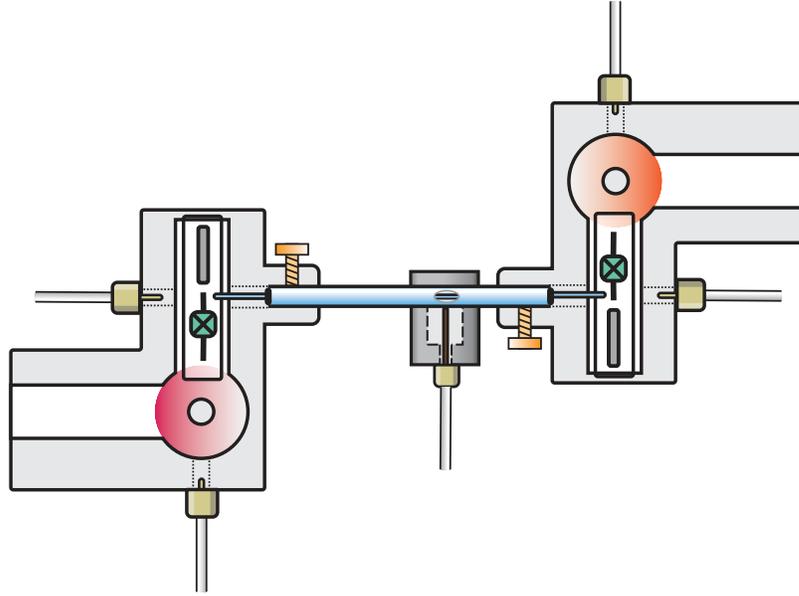


Figure 7.13: A more detailed look at the cQED hardware used in the DMM experiment. Each module contains two tunnels (white) in which chips can be placed (only one is used) and each module has two drive lines. One drive line is used exclusively to displace the cavity mode. The other is used to displace the readout resonator, drive the transmon qubit and deliver the pumps for four-wave mixing processes. Transmons are shown in green with a grey readout resonator. The readout Purcell filters are omitted. The cable's purcell filter is also shown, along with the brass screw clamping system. Dotted lines indicate ports drilled into the package.

Overall, we are able to reduce the loss rate of our bus mode to $\kappa_b/2\pi = 600 \pm 10$ kHz, which makes it easier to realize g_{BS} close to the critical damping regime. The complication of the Purcell filter can be avoided by using two transmons per module instead, where one is a dedicated coupling transmon, a la [Burkhart et al., 2021].

7.8.2 Detailed cQED hardware

In Fig. 7.13 we show the detailed cQED hardware for the DMM scheme. In Tab. 7.8.2, we show all the measured values of the relevant mode frequencies, static couplings and coherence times used in the experiment.

Mode	Parameter	Frequency (MHz)
Cavity 1	$\omega_{a_1}/2\pi$	6509
Cavity 2	$\omega_{a_2}/2\pi$	6487
Bus mode	$\omega_b/2\pi$	5658
Bus purcell filter	$\omega_{\text{filter}}/2\pi$	5688
Transmon 1	$\omega_{t_1}/2\pi$	5358
Transmon 2	$\omega_{t_2}/2\pi$	4759
Pump x_1	$\omega_{x_1}/2\pi$	5425.0
Pump y_1	$\omega_{y_1}/2\pi$	6271.76
Pump x_2	$\omega_{x_2}/2\pi$	4950.381
Pump y_2	$\omega_{y_2}/2\pi$	5775.11
Readout resonator 1	ω_{RO1}	7981
Readout resonator 2	ω_{RO2}	7895

Coupling	Parameter	Frequency (MHz)
Cavity 1 - Transmon 1 dispersive	$\chi_{a_1 t_1}$	-3.75
Cavity 2 - Transmon 1 dispersive	$\chi_{a_2 t_2}$	-2.2
Bus mode - Transmon 1 dispersive	$\chi_{b t_1}$	-2.1
Bus mode - Transmon 2 dispersive	$\chi_{b t_2}$	-2.5
Cavity 1 self-Kerr	K_1	-0.023
Cavity 2 self-Kerr	K_2	-0.007
Cavity 1 - Transmon 1 χ'	$\chi'_{a_1 t_1}$	+0.017
Cavity 2 - Transmon 2 χ'	$\chi'_{a_2 t_2}$	+0.004
Transmon 1 anharmonicity	α_{t_1}	-182
Transmon 2 anharmonicity	α_{t_2}	-187
Transmon 1 - Transmon 2 dispersive	$\chi_{t_1 t_2}$	0.05

Mode	T_1 or T_2 (μs)	κ (kHz)	Quality Factor
Bus mode	0.27	600 ± 10	9400
Purcell filter mode			170 ± 10
Cavity 1 (T_1)	385 ± 5		
Cavity 1 (T_2^R)	235 ± 10		
Cavity 2 (T_1)	520 ± 15		
Cavity 2 (T_2^R)	240 ± 10		
Transmon 1 (T_1)	31 ± 4		
Transmon 1 (T_2^R)	27 ± 1		
Transmon 1 (T_2^E)	34 ± 1		
Transmon 2 (T_1)	52 ± 1		
Transmon 2 (T_2^R)	21 ± 1		
Transmon 2 (T_2^E)	36 ± 1		
Readout resonator 1		360 ± 5	
Readout resonator 2		730 ± 5	

Table 7.2: Experimental parameters used for the DMM scheme

7.8.3 Tuning up simultaneous parametric beamsplitters

Being able to engineer the Hamiltonian $\hat{\mathcal{H}}_c$ is crucial for the DMM scheme to work. To actuate $\hat{\mathcal{H}}_c$, we tune $g_1 = g_2 = 160 \pm 5$ kHz in-situ but the tune-up process is non-trivial! When activating a beamsplitter process via four-wave mixing, we must satisfy the frequency matching condition $|\omega_{a_i} - \omega_b| = |\omega_{y,i} - \omega_{x,i}|$. We find the engineered beamsplitter is the ‘cleanest’ (lowest transmon heating, few spurious parametric resonances) when one of the pumps, pump x is placed around 100 MHz to 200 MHz above the transmon ω_{ge} frequency and the other pump, pump y is placed around 900 MHz to 1000 MHz above. In system 1, pump x_1 is designated the ‘strong pump’, since it induces the most stark shift on transmon 1. In system 2, both pumps induce similar amounts of stark shift on transmon 2. When we actuate both beamsplitters simultaneously, the frequency matching conditions change for both beamsplitter processes, since every pump tone induces a stark shift on the bus mode.

The first step in tune-up is to separately tune up the g_1 and g_2 beamsplitter processes and to match their beamsplitter rates as closely as possible. This is achieved by preparing a coherent state, applying the beamsplitter for a variable time and then measuring the probability the cavity is in the vacuum state via selective π -pulses on the transmon. By fitting the time dynamics, we can tune both g_1 and g_2 to be close to 160 kHz, although since we are in the slightly overdamped regime, the lack of oscillations gives an uncertainty of ~ 5 kHz. These pump conditions are also used to Q-switch cavity states to the bus mode to reset them to vacuum at the end of an experimental shot, by first Q-switching cavity 1 to the bus mode for 4 μ s, waiting 5 μ s, and then Q-switching cavity 2 to the bus mode for another 4 μ s.

After g_1 and g_2 have been tuned up individually, we must tune them up simultaneously. This mostly amounts to sweeping the pump y frequencies, ω_{y_1} and ω_{y_2} until both processes are on resonance when activated simultaneously. This calibration procedure is shown in Fig. 7.14.

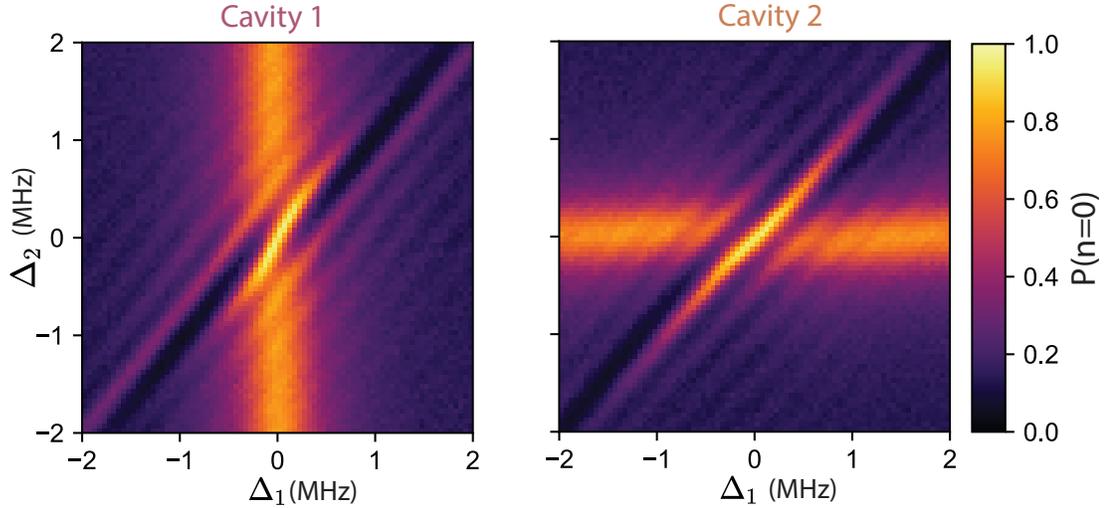


Figure 7.14: Pump calibration to account for stark shift. We sweep the difference in the pump frequencies, $\Delta_i = \omega_{y_i} - \omega_{x_i}$ by sweeping the value of ω_{y_i} . We choose to sweep ω_{y_i} because this pump is far detuned from the transmon and so this does not drastically change the normalized pump amplitude. To perform this calibration we prepare the coherent state $|\alpha\rangle_1 |\alpha\rangle_2$ (here $\alpha = \sqrt{2}$), apply all four pump tones for $2\ \mu\text{s}$, and then measure the cavity states with selective π -pulses to see if it is in the vacuum state. When Δ_1 is close to zero, but Δ_2 is far from zero, we see a vertical bar showing the g_1 process on resonance, but the g_2 process off-resonance. The same thing happens on the right plot but with Δ_1 and Δ_2 reversed, which gives a horizontal bar. The ideal pump detunings are found from where the vertical and horizontal bar intersect. As we approach $\Delta_1 = \Delta_2 = 0$, we see the appearance of fringes, a result of the cavity dynamics now becoming dependent on the relative phases between both cavities. This signature also verifies our pumps are phase-locked correctly.

Next we verify these pump frequencies are close to optimum by looking at time evolution vs relative cavity phase (See Fig. 7.18) and looking for any asymmetric or skewed dynamics. The final part of the calibration is a direct optimization of the Bell state fidelity where we fine tune the pump frequencies and amplitudes.

7.8.4 Pump phase-locking

The four pump tones must be appropriately phase-locked with each other, to ensure we enact the $\hat{\mathcal{H}}_c$ Hamiltonian on each shot with the correct relative phase between the g_1 and g_2 beamsplitter tones. Each tone is synthesized by sending an LO tone through an IQ mixer where it is combined with a single-sideband (SSB) tone from the FPGA controller. All SSBs remain phase-locked with each other over the course of a single shot, since they are sourced from the same FPGA controller, but drift apart on the timescale of seconds, and so their frequencies are reset after every experimental shot. Note that the two cavity drives at ω_{a_1} and ω_{a_2} must be phase locked with each other (and hence share the same LO generator) but do not need to be phase locked with any of the pumps that actuate $\hat{\mathcal{H}}_c$.

We enact $\hat{\mathcal{H}}_c$ with fixed relative phase when pump x_1 is phase-locked with pump x_2 , and pump y_1 is phase locked with pump y_2 , the same phase-locking condition required in [Burkhart et al., 2021]. This is typically achieved by sourcing both pump x_1 and x_2 LO tones from the same signal generator, and using SSB tones to synthesize the final pump tones. Achieving this phase-locking is more involved than merely sourcing pump tones from the same LO. The ω_{ge} transitions of transmon 1 and transmon 2 are detuned by ~ 500 MHz. This is by design, since we can have an appreciable ZZ coupling between the transmons mediated by the bus mode if they are too close in frequency. With this detuning, we observe a ZZ cross talk of 50 kHz (which is still quite large!). For phase locking, this means we must phase-lock pump tones that differ in frequency by ~ 500 MHz (in order to place each pump at the desired

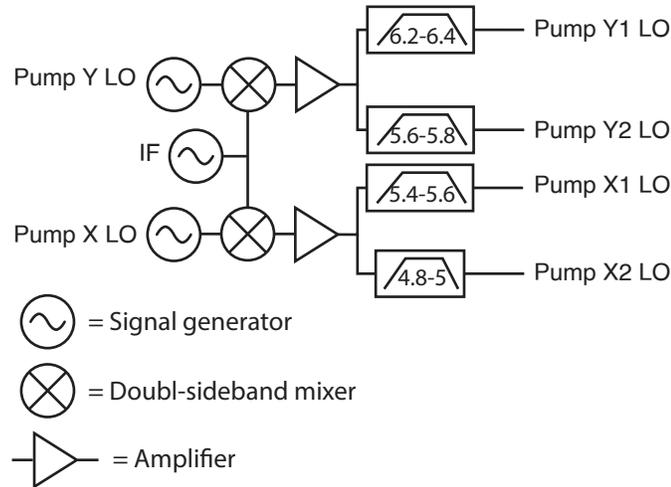


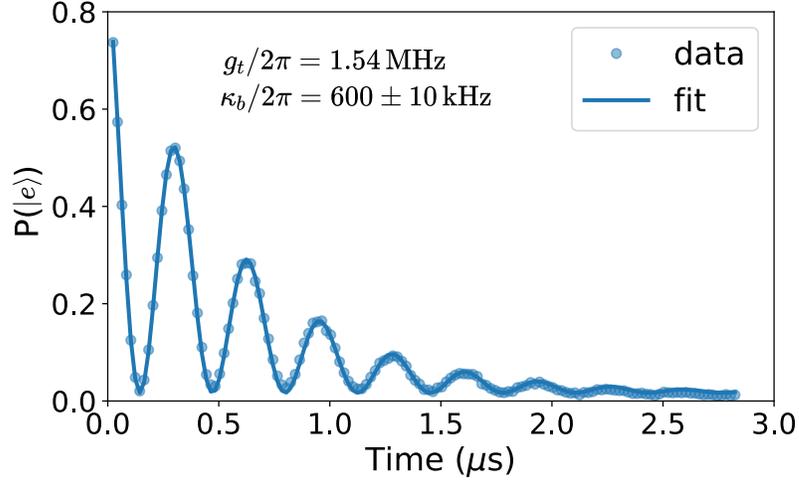
Figure 7.15: Wiring diagram for generating pairs of phase-locked pumps. Pump x_1 and x_2 are phase-locked with each other, as are Pump y_1 and y_2 . There is no phase-locking between the pairs but this is unimportant for us.

frequency above the transmons.) However, our FPGA controller can only synthesize SSB tones in the range ± 125 MHz.

To solve this problem, we use another signal generator, which we call the ‘IF’ generator. We set its frequency to 300 MHz, whilst the pump X LO generator is set to 5200 MHz and the pump Y LO generator set to 6000 MHz. With the use of additional mixers, we can generate phase-locked tones at 5200 ± 300 MHz and 6000 ± 300 MHz. By using four tunable narrow-band cavity filters, we can pick off just one of these tones to source the IQ mixers. The pump x_1 IQ mixer LO is at 5500 MHz, the pump x_2 IQ mixer LO is at 4900 MHz, the pump y_1 IQ mixer LO is at 6300 MHz and the pump y_2 IQ mixer LO is at 5700 MHz. This solution, drawn in Fig. 7.15 is very general and can be used to produce phase-locked sets of pump tones when their frequencies are inconvenient.

7.8.5 Measuring loss in the bus mode

The bus mode is lossy with a T_1 decay close to 300 ns. How do we measure this loss rate accurately when most operations in our system take on order of ~ 1 μ s?

Figure 7.16: Measuring κ_b in-situ.

Previously, this has been measured by swapping an excitation back and forth between one of the cavities and the bus mode by turning on just one of the beamsplitter processes [Burkhart et al., 2021]. In the critical coupling regime, we must opt for another approach. Instead we prepare transmon 2 in $|e\rangle$ and then actuate a sideband transition with the bus mode. This is another four-wave mixing process with the matching condition $|\omega_{t_2} - \omega_b| = |\omega_{y,2} - \omega_{x,2}|$. Since the transmon mode participates much more strongly in the J.J. than a cavity mode, with the same pump strengths we can engineer the Hamiltonian

$$\hat{\mathcal{H}}_t = g_t (\hat{t}^\dagger \hat{b} + \hat{t} \hat{b}^\dagger) \quad (7.45)$$

with $g_t \gg g_{BS}$ by roughly an order of magnitude in our case. With these oscillating dynamics, as shown in Fig. 7.16, it is easy to extract both κ_b and g_t from an exponentially decaying sinusoidal fit.

7.8.6 Kerr-cancellation

We measure the self-Kerrs of the cavities to be $K_1/2\pi = -23$ kHz and $K_2/2\pi = -7$ kHz respectively, distorting the dark states with inconvenient unitary dynamics during the $5.6\ \mu\text{s}$ duration of the measurement. Most of the time, we can compensate for the effects of Kerr by further modifying our logical basis. However, it is convenient to cancel out the effects of Kerr. We observe that the pumps used to actuate hf_c also happen to suppress the self-Kerr in the cavities. Most unwanted Kerr evolution occurs after $\hat{\mathcal{H}}_c$, during the vacuum check.

The most direct way of doing this is to apply an additional pump tone to cancel or change the sign of the self-Kerr in the cavity. This can be achieved by applying a blue-detuned pump to the transmon at $\omega_{ge} + 60$ MHz during the selective π -pulse, as used for the data in Fig. 7.23. Alternatively, we can undo the effects of Kerr with a SNAP gate combined with a selective π -pulse that implements the unitary which is used in the teleportation data in Fig. 7.29.

$$|0\rangle\langle 0| \otimes |e\rangle\langle g| + \hat{\mathcal{U}}_{\text{Kerr},1}^\dagger (\mathbb{1} - |0\rangle\langle 0|) \otimes |g\rangle\langle g|. \quad (7.46)$$

7.9 Observing Bright and Dark State Dynamics

We begin by first checking that we observe the expected dynamics of the ‘classical’ bright and dark states $|\alpha, \alpha, 0\rangle$ and $|\alpha, -\alpha, 0\rangle$. This experiment requires actuating the Hamiltonian $\hat{\mathcal{H}}_c$ and performing a vacuum check measurement. If this works as expected then we can use the same settings to generate entanglements merely by preparing cat states instead of coherent states.

Cavities are initially displaced via the cavity drives to the coherent states $|\alpha, \alpha e^{i\phi}\rangle$ where the phase, ϕ , is swept, where $\alpha = 1.414 \approx \sqrt{2}$. This is the value of α for which we will measure the highest Bell state fidelity. We apply the coupling Hamiltonian, $\hat{\mathcal{H}}_c$

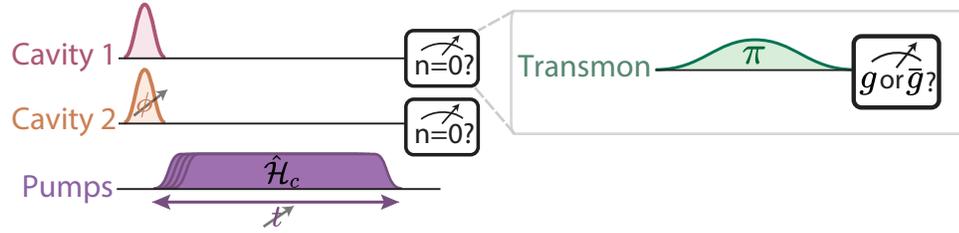


Figure 7.17: Pulse sequence for characterizing the DMM scheme with classical bright and dark states. Initially the cavities are displaced to the state $|\alpha, \alpha e^{i\phi}\rangle$

for variable pump time t , and then perform the vacuum check. These are implemented via number selective π -pulses on each transmon, followed by standard transmon readout from which we obtain four possible outcomes (g, g) , (g, e) , (e, g) , (e, e) . The pulse sequence is shown in Fig. 7.17

We only pass the vacuum check if we obtain the outcome (g, g) , where we then presume we are in a dark state. All of outcomes, denoted by $(\overline{g\overline{g}})$, fail, and we presume we started in a bright state which has since decayed to vacuum. Measuring both cavities is not strictly necessary but makes us robust to first-order transmon measurement errors — we require two errors to incorrectly declare a dark state. These errors are around 5% for each transmon and arise from transmon decoherence during the selective π -pulse, transmon decay during the readout and readout overlap error (in the histograms). In this experiment, we plot $p(\overline{g\overline{g}})$ as a function of pump time and ϕ as shown in Fig. 7.18

What is the figure of merit for the vacuum check? How well can we distinguish dark states from bright states? The success probability of our scheme will be

$$p_{\text{success}} = \frac{1}{2}p(\overline{g\overline{g}}|\text{dark}) \quad (7.47)$$

Because of the finite overlap with vacuum at $\alpha = \sqrt{2}$, we expect this to be limited to 0.374 which is close to what we will measure. Applying the pumps to actuate \hat{H}_c further reduces this probability by a few percent, not because these pumps induce

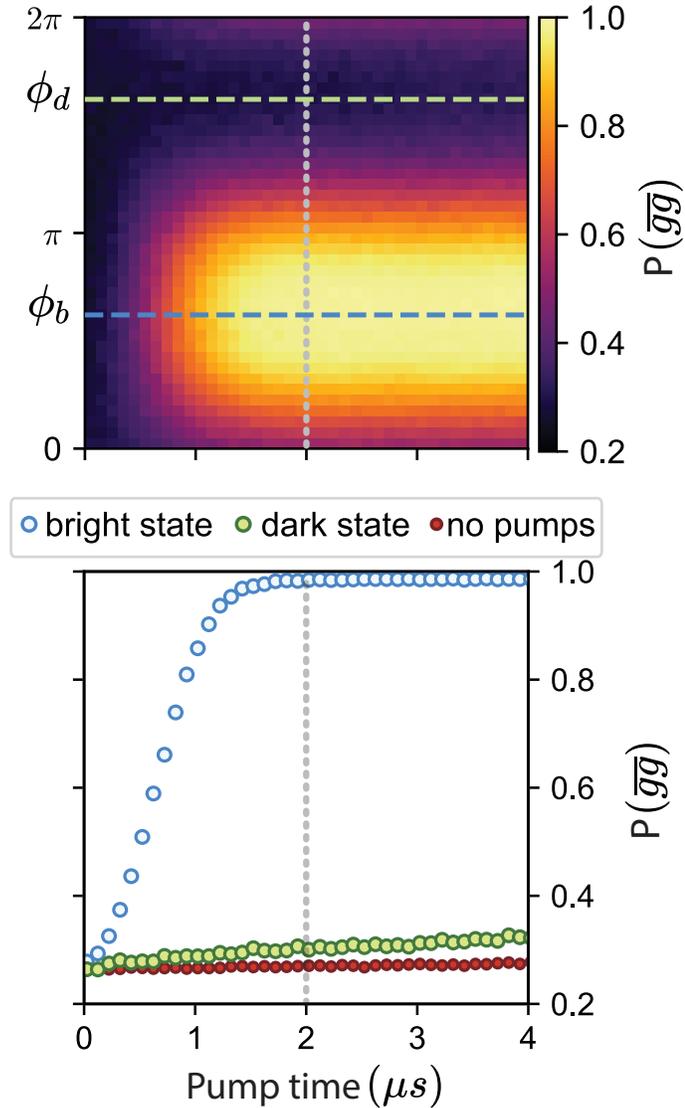


Figure 7.18: **(Top)** Dynamics of $|\alpha, e^{i\phi}\alpha\rangle$ under , for $\alpha = 1.414$. We identify a particular phase, ϕ_d for which we have prepared the dark state, where no dynamics should occur. At phase $\phi_b = \phi_d - \pi$ we have the bright state, which decays to vacuum quickly within 2 μs . **(Bottom)** Line cuts of the above plot, along with the dynamics when all pumps are off (red circles).

additional photon loss on the cavities, but rather because they heat the transmons out of their ground state, and so we may then record (g, e) or (e, g) after the vacuum check. For this reason it is important to make sure we can engineer our required g_{BS} with minimal pump strength (in terms of its Stark shift on the transmon). However, our scheme is naturally robust to pump heating, due to the double transmon measurement. Ultimately transmon heating from the pumps should only reduce the success probability, and not the Bell state fidelity.

The figure of merit we care about for the Bell state fidelity is the ‘false positive rate’: the probability we declare success when we actually started in a bright state. This is defined as

$$\mathcal{F}_{\text{DMM}} = \frac{p(gg|\text{bright})}{p(gg|\text{dark}) + p(gg|\text{bright})}. \quad (7.48)$$

Where \mathcal{F}_{DMM} is the fidelity of the dark mode measurement, which represents an upper bound on how well we would expect this scheme to work in the absence of SPAM errors. From this experiment, we find $\mathcal{F}_{\text{DMM}} = 2.1\%$, a far cry from the fundamental limits of this scheme! Is this because of transmon decoherence during the vacuum check? We know this cannot be the whole picture. This is because we can record the probability an individual transmon reads out $|g\rangle$ when we prepare a bright state, which we find to be 5% and 7% respectively. From this we would expect 0.5% for the dark mode measurement fidelity, if each transmon outcome is *independent* of the other. The missing 1.5% must then be due to *correlated* errors between the transmon outcomes, where a single error mechanism causes both transmons to readout g instead of e . At the moment, the source of these correlated errors is unknown.

Improvements to \mathcal{F}_{DMM} in any future implementations could result from using ‘cleaner’ parametric beamsplitter processes [Chapman et al., 2022, Lu et al., 2023] and having a dedicated coupling element and a separate transmon ancilla in each module.

7.9.1 High-fidelity cat state preparation

The ‘only’ modification needed to create entanglement is that we must run the same experiment but starting with a cat state in each cavity. Displaced states are trivial to prepare to high fidelity in a cavity mode but what is the best way to prepare cats? We tried OCT pulses for even, odd and parity-less cats which take the form $|\alpha\rangle \pm i|-\alpha\rangle$ which do not have a well-defined photon number parity. Parityless cats are the easiest to prepare with an estimated state fidelity $\sim 99\%$ via the following protocol.

First we use simple displacement pulses to prepare $|\alpha\rangle$ in both cavities. We then apply an OCT pulse to each transmon that functions similarly to a SNAP pulse. For example, the unitary that takes $|\alpha\rangle \rightarrow |\alpha\rangle \pm i|-\alpha\rangle$ can be written as a SNAP unitary, i.e., is of the form

$$\hat{U}_{\text{SNAP}} = \sum_{n=0}^{\infty} e^{i\theta_n} |n\rangle \langle n|, \quad (7.49)$$

where $\theta_n = \pi/4$ for $n = \text{even}$ and $\theta_n = -\pi/4$ for $n = \text{odd}$. Whilst we can construct this SNAP pulse by hand, we find OCT pulses that are constrained to only drive the transmon perform better, since effects such as the finite selectivity of π -pulses are automatically accounted for. As for a regular SNAP pulse, the OCT pulse leaves the transmon back in its ground state at the end of the pulse. If we were initialize in $|-\alpha\rangle$ instead, this same OCT pulse would prepare the state $\propto |\alpha\rangle - i|-\alpha\rangle$.

When we implement these ‘OCT-SNAP’ pulses in experiment, we see visible distortions in the Wigner functions, typically in the fringes or ‘whiskers’ of the cat. We attribute this to frequency dispersion in our control lines which distorts our OCT pulse. This means the actual state we prepare is of the form.

$$|\tilde{\psi}\rangle = \tilde{U}(|\alpha\rangle \pm |-\alpha\rangle)/\sqrt{2}, \quad (7.50)$$

where \tilde{U} is an unwanted SNAP unitary, also of the form $\tilde{U} = \sum_{n=0}^{\infty} e^{i\tilde{\theta}_n} |n\rangle \langle n|$ where

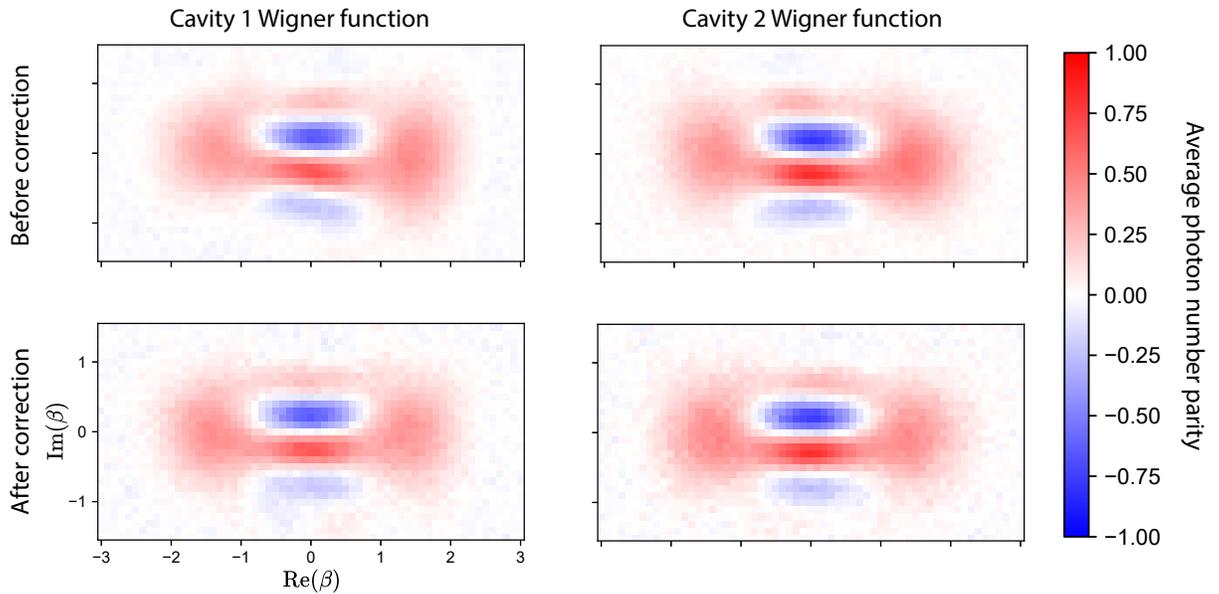


Figure 7.19: Correcting for pulse distortions in the control lines during cat state preparation. Visually from the measured Wigner functions we see these states are initially distorted (top row). In the bottom row once we have tuned up the corrected OCT pulse, we see less distortion, especially for cavity 1.

$\tilde{\theta}_n$ are unwanted phase distortions due to the unknown transfer function of the control lines⁴. However, since we can reconstruct the density matrix from the Wigner function, we can find the values of $\tilde{\theta}_n$ from the density matrix. To correct for this effect, we simply ask our OCT optimization to create a pulse that implements the unitary $\tilde{\mathcal{U}}^\dagger \hat{\mathcal{U}}_{\text{SNAP}}$ instead, using the previous pulse as the seed waveform. We find visually this improves the appearance of the fringes and increases state fidelity by 1–2%. Fig. 7.19. shows the Wigner functions of the prepared cat states before and after this correction.

4. Maybe this technique could be used to calibrate dispersion in the control lines in general?

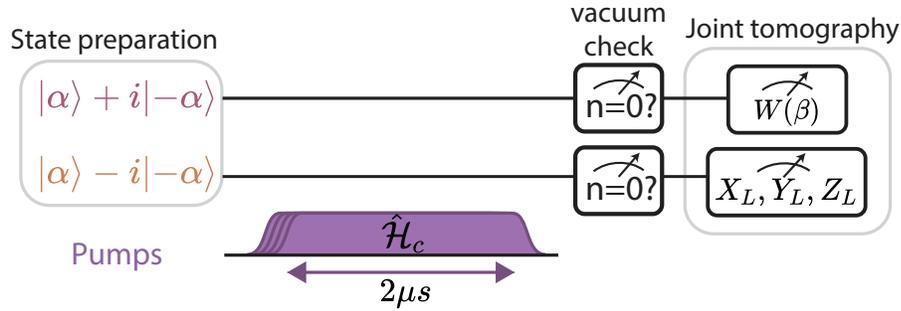


Figure 7.20: Pulse sequence for entanglement generation. We declare success when we measure (g, g) in the vacuum check. The duration of the scheme, including step preparation to the end of the vacuum check measurement is $5.6 \mu\text{s}$

7.9.2 Procedure for entanglement generation

The pulse sequence for generating entanglement with the DMM scheme is shown in Fig. 7.20. State preparation is achieved by displacing the cavities to $|\alpha, -\alpha\rangle$ and then applying the respective OCT SNAP pulses to each transmon. We then apply $\hat{\mathcal{H}}_c$ for $2 \mu\text{s}$ so the bright states decay in the bus. It takes an additional $48 \mu\text{s}$ to ‘ring up’ the four pumps used for g_1 and g_2 . We add a phase offset of ϕ_b to one of these pumps, which is chosen to define the bright states of our system as $|\alpha, \alpha\rangle$ on every attempt. Pump ‘ring-down’ takes the same duration. We then perform the vacuum check with simultaneous selective π -pulses on the transmons centered at the ω_{ge} frequency. These are Gaussian pulses with duration $4\sigma = 800$ (1200) ns for transmon 1 (2). Immediately after measuring the transmons in (g, g) , we proceed with joint-cavity tomography. Otherwise we attempt the experiment again after performing the following reset protocol.

7.9.3 System reset

We use a reset protocol that is a simpler version of the feedback cooling scheme in [Burkhart et al., 2021]. To reset the system, we first ensure the cavities are empty (we may have been in a dark state but declared failure due to readout error, and need

to empty the cavities). We Qswitch each cavity by actuating g_1 and g_2 sequentially. This ensures that the cavities lose their energy to the bus regardless of the cavity state. We apply g_1 for $5\ \mu\text{s}$, wait $2\ \mu\text{s}$ and then repeat with g_2 . We then use selective π -pulses to confirm the cavities are indeed in vacuum, restarting the entire protocol if required. Finally, we use feedback cooling on the transmons by repeating dispersive readout. If the transmon is not in its ground state, we apply a π -pulse and repeat the measurement. The cooling sequence terminates once we measure both transmons to be in the ground state.

Feedback cooling also helps cool the transmons to their ground state. We find their thermal populations to be quite high: 4% (6%) for transmon 1 (2). After cooling, we measure 0.9% (0.5%) to be the probability the transmon is not in the ground state. Our scheme will not work if the transmons are not in their ground state, adversely affecting the state preparation, $\hat{\mathcal{H}}_c$ and joint-tomography. Thankfully, this will be a source of error that is well-detected by default in our vacuum check measurement.

7.10 Wigner Functions and Parity Measurements

Measuring a single cavity's Wigner function is not enough to verify entanglement by itself, but it is incredibly useful in calibrating the experiment and can be incorporated into joint-cavity tomography which can verify entanglement.

7.10.1 Measuring Wigner functions

Whether or not we can measure the Wigner function well largely hinges upon whether we can implement a good parity measurement. At large values of β or γ , our parity measurement will invariably get worse. The averaged parity will invariably decrease with the size of the displacement, resulting in a dimming or 'vignetting' towards the edges of the measured Wigner function. This effect is especially pronounced when

using the $\pi/2$ -wait $\pi/\chi - \pi/2$ pulse sequence to measure parity. The issue is that due to the dispersive interaction and the finite bandwidth of the $\pi/2$ pulses. At higher photon numbers, our $\pi/2$ pulses are effectively off-resonant, and so they rotate the transmon by some angle less than $\pi/2$. Ultimately this biases our parity measurement towards recording the g outcome, regardless of what the averaged parity should be. We can partially mitigate this effect by detuning our $\pi/2$ pulses to be centered around frequency $\omega_{ge} - |\beta|^2\chi$, but this is only a partial correction since there is a spread of photon numbers in the cavity.

Symmetrizing a single-mode Wigner function involves measuring the photon number parity in two ways: first by mapping even photon number of the $|g\rangle$ state, and then by mapping to the $|e\rangle$ state instead, which is typically achieved by inverting the amplitude of the last $\pi/2$ pulse. We denote these two averaged Wigner functions as W^g and W^e respectively. Taking the ‘symmetrized’ Wigner function as

$$W(\beta) = \frac{W^e(\beta) - W^g(\beta)}{2} \quad (7.51)$$

removes ‘vignetting’ but it does *not* remove all systematic errors. The ‘contrast’ of the Wigner function can be dependent on β , although this contrast reduction depends not only on β, χ and the pulse bandwidth but also the state being measured. A good rule of thumb is not to trust the fidelity extracted from a Wigner function to below 1% precision, due to systematic effects such as these.

7.10.2 Frequency combs for parity measurements

One alternate way of implementing a parity measurement is to replace the standard pulse sequence with a comb of selective π -pulses. This pulse sequence is slightly longer, which gives a larger reduction in contrast but we expect this contrast to be more uniform over a wider range of cavity photon numbers up to the maximum

frequency in the comb.

In Fig. 4.12 we find this approach is particularly useful when measuring the Wigner function of cavity 1, which has a larger than desired value of $\chi_{a_1 t_1}$. A comb is constructed by playing simultaneous selective π -pulses at frequencies $\omega_n = \chi n + \chi'(n^2 - n)$ for even n up to a maximum of $n = 14$. In all measured Wigner functions, we opt to use the frequency comb rather than the standard parity pulse sequence⁵.

7.10.3 Fine-tuning pump parameters with joint parity

We use the averaged joint parity to fine tune the pump calibration. The target dark state is always expected to have a well-defined joint parity: odd in our case. This is also a good proxy for the entanglement fidelity (although we need to verify correlations in phase and cavity photon number to verify entanglement). We finely adjust the pump frequencies ω_{y_1} and ω_{y_2} (sensitive to 10 kHz changes), as well as the pump y_2 DAC amplitude (sensitive to 2% changes) to minimize the joint parity.

We must repeat this pump tuneup procedure whenever we change the value of α we use for our entangled state generation. Typically, these adjustments are of order 10 kHz and are thought to be as a result of cavity self-Kerr, meaning the resonance condition is slightly dependent on the average number of photons in the cavities.

7.11 Joint-cavity Tomography

How do we verify entanglement? We must perform quantum state tomography on both of the cavities. Here we outline how this may be done either by measuring the Joint Wigner function, or by ‘decoding’ one of the cavities onto the state of its transmon. To extract a Bell state fidelity, we opt for the latter approach.

5. In the limit of large n , it can be shown that the frequency comb and the standard pulse sequence with instant $\pi/2$ pulses are mathematically the same pulse sequence!

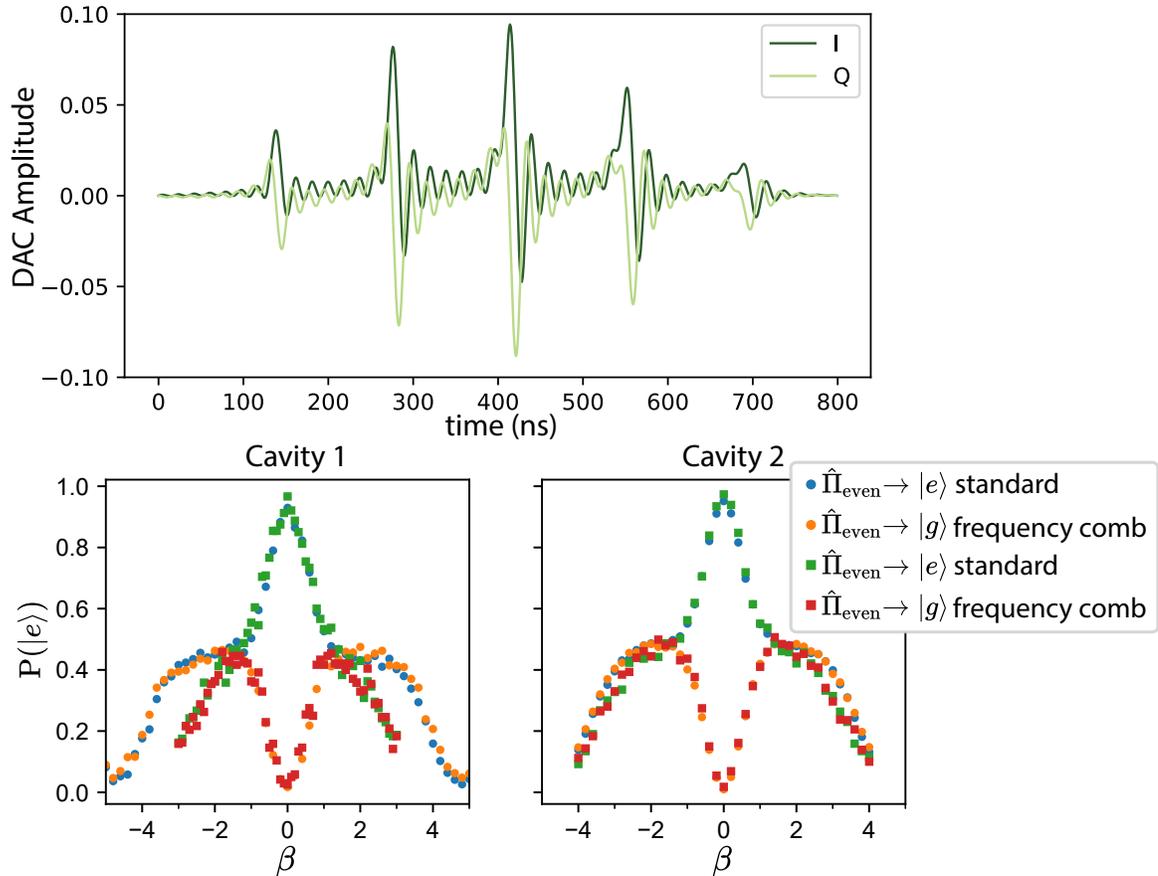


Figure 7.21: Improving parity measurements with combs of selective π -pulses. **(Top)** Example pulse constructed from a comb of eight selective Gaussian π -pulses with $\sigma = 200$ ns and 4σ duration. Each Gaussian is centered around frequency $\omega_n = \chi n + \chi'(n^2 - n)$ for $n = \{0, 2, 4, 6, 8, 10, 12, 14\}$. **(Bottom)** Measured 1D slices of the Wigner function of the vacuum state, for different parity measurement implementations. For a parity measurement that works well for larger photon numbers, we expect $P(|e\rangle) = 0.5$ for large β . Mainly due to finite pulse bandwidth, we observe a drop off towards $P(|e\rangle) = 0$ as β increases, regardless of which way we are mapping the parity to the transmon state. We see for cavity 1 that using a frequency comb instead results in a (steeper) drop-off that appears at almost twice the value of β (around $\beta = 3$ rather than $\beta = 1.5$). For cavity 2, we observe similar performance for both the frequency comb and the standard parity sequence.

7.11.1 Joint Wigner functions

One approach commonly used is measuring the Joint Wigner function, defined by extending Eq. 3.36 to two bosonic modes.

$$W_J(\beta, \gamma) = \left(\frac{2}{\pi}\right)^2 \text{Tr} \left(\hat{D}_\beta \hat{D}_\gamma \rho \hat{D}_\beta^\dagger \hat{D}_\gamma^\dagger \widehat{JP} \right) \quad (7.52)$$

This gives us a clear way to measure the Joint Wigner function, which constitutes a complete tomography of the two-cavity system. We must displace cavity 1 by β and cavity 2 by γ . Crucially, \widehat{JP} can be measured from two local parity measurements on either cavity (e.g., if on one experimental shot we measure cavity 1 to have odd parity and cavity 2 to have odd parity, we declare even joint parity).

The main drawback of measuring the Joint Wigner is that this 4D function can take an extremely long time to measure completely. For reference, measuring a single cavity's Wigner function with a 51×51 grid of points for β to cover area $\text{Re}(\beta), \text{Im}(\beta) \in [-2.5, 2.5]$ takes a few hours. Measuring the Joint Wigner function over the full 4D space would take several months!

There are a few solutions to this that sample more efficiently but we tend to be more susceptible to transmon readout errors with these approaches – transmon errors during the parity measurement tend to reduce the ‘contrast’ of a Wigner function, provided they result in unbiased errors. If we are confident our errors are unbiased (e.g., by symmetrizing the Wigner function), and if we are sure the reduction in contrast does not depend on α (See [Sivak et al., 2023] for when this is the case), then we can use the fact that all Wigner functions should be normalized to correct these errors. This works well for single-mode measured Wigner functions but not sparsely sampled Joint Wigner functions.

We can instead take 2D ‘slices’ of the Joint Wigner function to qualitatively examine if we indeed have entanglement, e.g., we can take points on a 2D grid where we

fix $\text{Im}(\beta), \text{Im}(\gamma) = 0$ (The Re-Re slice). and another slice when $\text{Re}(\beta), \text{Re} = 0$ (The Im-Im slice). These measured slices are shown in Fig. 7.22 and show clear features that can only be explained by an entangled state, but are not sufficient to reliably extract a Bell state fidelity. Slices of the Joint Wigner function were also used in [Wang et al., 2016] for the cat-in-two-boxes state within a single module, and indeed our slices share many of the same qualitative features.

7.11.2 Symmetrizing Joint Wigner functions

When measuring the Joint Wigner function from two local parity measurements, there are now four ways we can configure the two parity measurements, by choosing whether we map even parity to g or e . We denote these configurations as $W^{g_1, g_2}(\beta, \gamma)$, $W^{g_1, e_2}(\beta, \gamma)$, $W^{e_1, g_2}(\beta, \gamma)$ and $W^{e_1, e_2}(\beta, \gamma)$. The Joint Wigner function is ‘properly symmetrized’ when we take

$$W_J(\beta, \gamma) = \frac{W_J^{g_1, g_2}(\beta, \gamma) - W_J^{e_1, e_2}(\beta, \gamma) - W_J^{e_1, g_2}(\beta, \gamma) + W_J^{g_1, e_2}(\beta, \gamma)}{4} \quad (7.53)$$

7.11.3 OCT decode + a measured Wigner function

Characterizing entanglement in two-level systems is much easier, since the Hilbert space is smaller and it is easy to measure in different logical bases. We can do this in our system too with the help of OCT pulses.

The idea is to implement ‘decode’ pulses, which transfer the logical state of the cavity onto its transmon, i.e.,

$$|0_L\rangle_a \otimes |g\rangle \rightarrow |\text{vac}\rangle_a \otimes |g\rangle \quad (7.54)$$

$$|1_L\rangle_a \otimes |g\rangle \rightarrow |\text{vac}\rangle_a \otimes |e\rangle \quad (7.55)$$

Where $|0_L\rangle, |1_L\rangle$ are the logical basis of the cavities, the modified basis from Sec. 7.7,

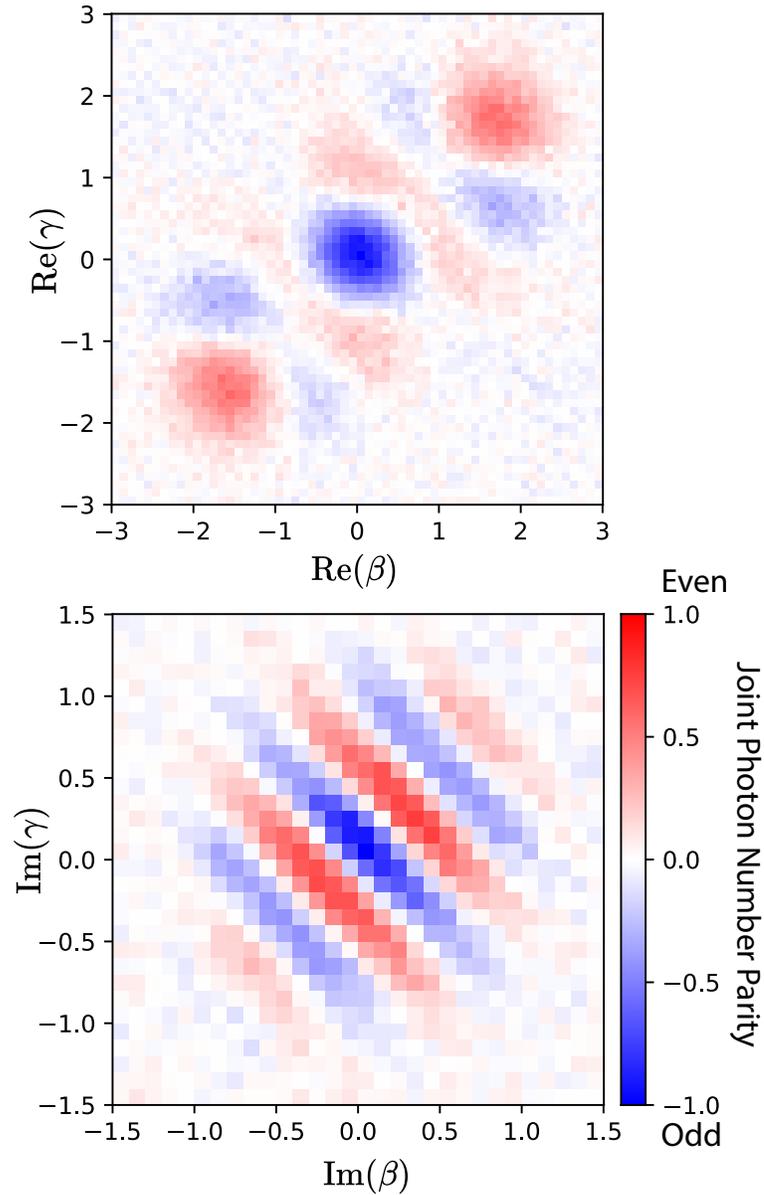


Figure 7.22: 2D slices of the Joint Wigner function for the DMM scheme at $\alpha = 1.6$, showing the signatures of entanglement. On the left we fix β and γ displacements to be real-valued. The principal features of this plot are correlations visible around $(\beta, \gamma) = (1.6, -1.6)$ and $(\beta, \gamma) = (-1.6, 1.6)$, which verifies we are confined to the $\{|\pm\alpha, \mp\alpha\rangle\}$ manifold. The negative joint parity centered around the origin verifies we are in an entangled state. Fainter features are artifacts from removing $|0\rangle$ from our final state via the vacuum check and are expected. On the right we fix β and γ to be imaginary, and see interference fringes, another signature of entanglement.

plus any Kerr evolution on the cavity which can be determined by measuring its Wigner function.

Naturally, with dispersive readout, we measure the transmon in its Z basis but we can readout the transmon (and hence the cavities) in any basis by applying the appropriate transmon rotation pulse before the measurement.

This technique certainly speeds up the measurement time: we only need to measure at 36 different pulse settings compared to 51^4 with Joint Wigner! However, the OCT decode pulses themselves are error prone, each one introducing a 2% error. In the experiment, we compromise between Joint Wigners and OCT decode by decoding cavity 2 onto transmon 2 (which has the better coherence time) whilst simultaneously measuring cavity 1's Wigner function. This is the same procedure used in [Burkhart et al., 2021]. The data processing is more involved to extract a Bell state fidelity, since we must process a mixture of measured single cavity Wigner functions and transmon 2 measurements, but it significantly reduces the SPAM error and provides a very striking confirmation of entanglement, where correlations in different bases becomes clearly apparent.

7.12 Extracting Bell State Fidelity

What do the tomography results look like and how do we extract a Bell state fidelity from them? With the aforementioned procedure for joint-cavity tomography, we expect a total of six separate (symmetrized) measured Wigner functions for cavity 1, because there are six possible outcomes from the logical measurement of cavity 2

(two outcomes, three different bases). The target Bell state takes the form

$$\frac{1}{\sqrt{2}} (|0\rangle |1\rangle - |1\rangle |0\rangle) \quad (7.56)$$

$$= \frac{1}{\sqrt{2}} (|+\rangle |-\rangle - |-\rangle |+\rangle) \quad (7.57)$$

$$= \frac{1}{\sqrt{2}} (|+i\rangle |-i\rangle - |-i\rangle |+i\rangle) \quad (7.58)$$

$$(7.59)$$

For this particular Bell state, the anti-correlation persists through each of the three measurement bases. We denote the six outcomes of cavity 2's measurement as $\{|+z\rangle, |-z\rangle, |\pm x\rangle, |\pm y\rangle\}$ which corresponds to the state of cavity 2 having been $\{|1_L\rangle, |0_L\rangle, |\pm_L\rangle, |\pm i_L\rangle\}$. These six measured Wigner functions are shown for $\alpha = 1.414$ in Fig. 7.23

The method we will use to extract a Bell state fidelity is the same as the method used in [Burkhart et al., 2021]. The goal is to reconstruct the expectation values of all sixteen joint Pauli correlations, $\langle \sigma_{i,1} \rangle \sigma_{j,2}$ as mentioned way back in Sec. 2.3.1 of Chapter 2.

The first step is to reconstruct cavity 1's density matrix for each measured Wigner function. Whilst it is possible to directly invert the Wigner function to obtain a density matrix, this method is very sensitive to experimental noise. Because of the finite measurement contrast, this often results in an 'unphysical' density matrix with a norm less than 1. To combat these problems, we use the well-established technique of maximum likelihood estimation (MLE) with a convex optimization method described in Section 5.5. of Kevin Chou's thesis [Chou, 2018] and used in the following works [Axline et al., 2018, Burkhart et al., 2021]. The optimization returns the closest physical density matrix (normalized, Hermitian etc.) that most likely produced the Wigner data. To avoid over-fitting to the noise we average each point in the Wigner func-

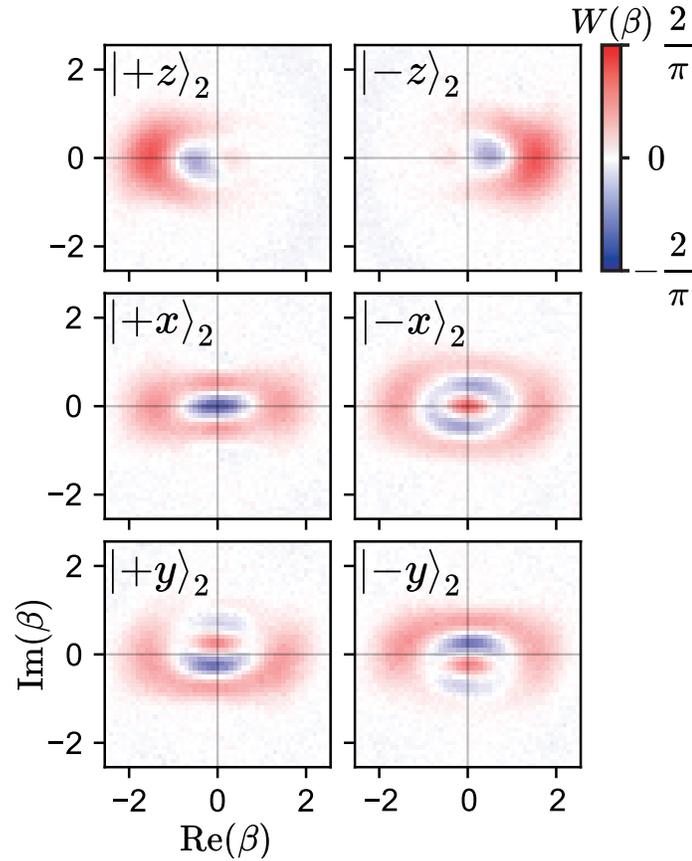


Figure 7.23: Measured Wigner functions of cavity 1, for $\alpha = 1.414$, conditioned on the outcome of cavity 2 measurements in each of the three logical bases (outcome indicated by the inset). The basis is the modified basis with vacuum removed as shown in the right panel of Fig. 7.10. A Kerr-cancelling pump was applied to transmon 1 during the vacuum check, which is why no Kerr distortions are apparent. The decode pulse used for cavity 2 takes into account the Kerr evolution of cavity 2 during the protocol in the definition of states $|0_L\rangle_2$ and $|1_L\rangle_2$

tion from 10,000 different experimental shots (before post-selection), which reduces the ‘shot noise’. We also truncate the maximum number state in our reconstructed density matrix to be $N_{\max} = 10$ (increased to $N_{\max} = 13$ for $\alpha = 1.8, 2$).

We test the robustness of the MLE reconstruction procedure by generating simulated Wigner function data, adding in randomly generated shot noise⁶ and trying to reconstruct the original state. For this test we used various cat states with $\alpha = \sqrt{2}$. We find that for $N_{\text{shots}} > 1000$, the reconstruction error is below 1%, a good rule of thumb when taking Wigner data. We estimate a $< 0.3\%$ reconstruction error due to shot noise. Most of the uncertainty in the reconstructed density matrix comes from systematic errors from measuring the Wigner function, which are much harder to characterize. For this reason, we quote an uncertainty of 1% in fidelities extracted from measured Wigner functions, which was also done in [Burkhart et al., 2021, Eickbusch et al., 2022].

We can calculate the Joint Pauli correlators, also called the ‘Pauli bars’ from these six reconstructed density matrices, designated $\rho_j^{(1)}$ and the probabilities for obtain the different measurement outcomes on cavity 2 denoted by $p_{\pm i}^{(2)}$ where $p_{+i}^{(2)} + p_{-i}^{(2)} = 1$. In both cases, $i, j = \{\mathbb{1}, x, y, z\}$.

When $j \neq \mathbb{1}$, the Pauli bar is given by

$$\langle \sigma_i \sigma_j \rangle = p_{+j}^{(2)} \text{Tr} \left(\rho_{+j}^{(1)} \sigma_i^{(1)} \right) - p_{-j}^{(2)} \text{Tr} \left(\rho_{-j}^{(1)} \sigma_i^{(1)} \right) \quad (7.60)$$

When $j = \mathbb{1}$ we use

$$\langle \sigma_i \sigma_{\mathbb{1}} \rangle = p_{+j}^{(2)} \text{Tr} \left(\rho_{+j}^{(1)} \sigma_i^{(1)} \right) + p_{-j}^{(2)} \text{Tr} \left(\rho_{-j}^{(1)} \sigma_i^{(1)} \right) \quad (7.61)$$

For this case it does not matter which basis, j we use to calculate these Pauli bars, any would suffice. We calculate these Pauli bars by averaging which basis we choose

6. Which follows a Binomial distribution not a Gaussian!

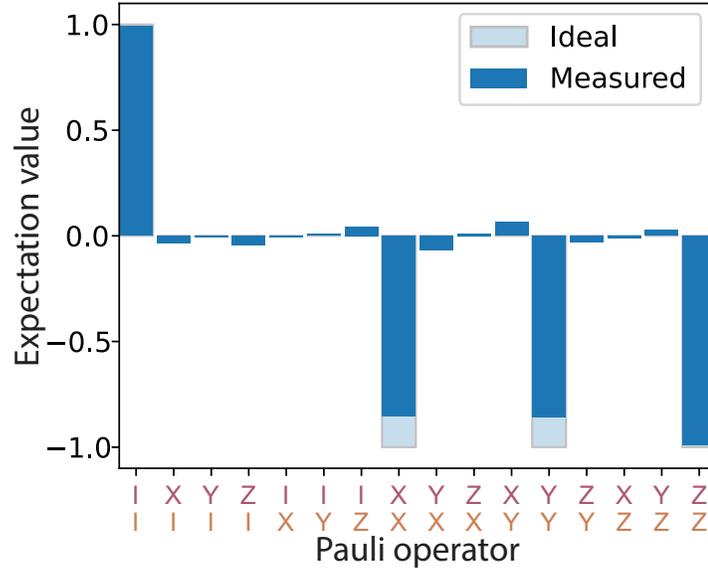


Figure 7.24: Reconstructed Pauli bars for $\alpha = 1.414$, where the highest Bell state fidelity is obtained at $92 \pm 1\%$

to calculate these Pauli bars. (We could obtain them from either of the three bases. We choose to average over all three.)

We can calculate terms such as $\text{Tr}(\rho_{+j}^{(1)} \sigma_i^{(1)})$ from our reconstructed density matrix and $\sigma_i^{(1)}$.

The form of $\sigma_i^{(1)}$ is the only part of the reconstruction process which is parameterized. Changes in our definition of the logical basis for cavity 1 will change the set of $\sigma_i^{(1)}$ operators, which is ultimately used to find the best basis for cavity 1 in the presence of Kerr, rotation offsets, etc. (The logical basis for cavity 2 is fixed by our choice of OCT decode pulses, and we must calibrate the Kerr, α and the rotation angle separately from measured Wigner functions of cavity 2). We show the reconstructed Pauli bars in Fig. 7.24 for $\alpha = \sqrt{2}$.

Having the expected Pauli bars is enough to reconstruct the full two-qubit density matrix in the logical subspace, by using the equation

$$\rho_{\text{recon}} = \frac{1}{2} \sum_{\substack{i,j=0,1, \\ 2,3}} \langle \sigma_i, \sigma_j \rangle \sigma_i \sigma_j \quad (7.62)$$

which is the two-qubit version of Eq. 7.62. As we saw with the Wigner function, whilst we can directly invert this equation, there is no guarantee the resulting density matrix is physical due to measurement errors, and so we can use Maximum Likelihood Estimation again to find the closest physical two-qubit density matrix, ρ_{final} which best matches the set of Pauli bars. This time, we remove the constraint that the reconstructed density must be normalized since our logical qubits are within the subspace of bosonic modes, and we can suffer leakage out of the codespace, e.g., to the vacuum state $|0\rangle$. Finally, from this density matrix we can quote the Bell state fidelity as

$$\mathcal{F}_{\text{Bell}} = \langle \Psi_- | \rho_{\text{final}} | \Psi_- \rangle \quad (7.63)$$

For our target Bell state, the only non-zero Pauli bars should be $\langle \mathbb{1}_1 \mathbb{1}_2 \rangle = 1$ and $\langle X_1 X_2 \rangle = \langle Y_1 Y_2 \rangle = \langle Z_1 Z_2 \rangle = -1$ and we see qualitatively we are close to this scenario.

We numerically optimize the definition of cavity 1's logical basis to maximize the Bell state fidelity. This basis is parameterized by three parameters: $\alpha_{\text{basis}}, \theta_K, \theta_r$ which respectfully account for shrinking α due to the amplitude damping back-action associated with cavity photon loss, Kerr distortion during the time of the protocol and any mis-calibration of drive phases that leads to a cavity rotation offset. The logical basis is defined from these parameters as:

$$\begin{aligned} |+\rangle_L &\propto e^{i\theta_r \hat{a}^\dagger \hat{a}} e^{\frac{i}{2}\theta_K \hat{a}^\dagger \hat{a} \hat{a}^\dagger \hat{a}} \hat{\Pi}_0 (|\alpha_{\text{basis}}\rangle + |-\alpha_{\text{basis}}\rangle) \\ |-\rangle_L &\propto e^{i\theta_r \hat{a}^\dagger \hat{a}} e^{\frac{i}{2}\theta_K \hat{a}^\dagger \hat{a} \hat{a}^\dagger \hat{a}} (|\alpha_{\text{basis}}\rangle - |-\alpha_{\text{basis}}\rangle). \end{aligned} \quad (7.64)$$

When we quote a Bell state fidelity we also quote the logical basis parameters for cavity 1. (Cavity 2 has been decoded onto transmon 2 so really we are measuring the Bell state fidelity between cavity 1 and transmon 2 as a lower bound for the Bell state fidelity between the cavities prior to tomography.)

We obtain $\mathcal{F}_{\text{Bell}} = 92 \pm 1\%$ for $\alpha = 1.414$ with the Kerr cancelling pump on, the highest Bell state fidelity we achieve, with a success probability of 0.349.

Examining the Pauli bars tells us that the main source of error appears to be photon loss. The $\langle \mathbb{1}_1 \mathbb{1}_2 \rangle$ bar is very close to 1, indicating our choice of basis is correct. The $\langle Z_1 Z_2 \rangle$ bar is also very close to -1, indicating that our scheme is very good at distinguishing between bright and dark states. The main source of error is apparent from the $\langle X_1 X_2 \rangle$ and $\langle Y_1 Y_2 \rangle$ Pauli bars. (The extra Pauli bars at XY and YX look bad but actually contribute less than 0.1% to the entanglement infidelity.) These both suggest that photon loss, which reduces the contrast of the fringes of the cats when measuring in the x or y basis is responsible for most of the error. We can check if this is as expected by making an error budget outlined below.

7.12.1 Error budget

We identify three dominant error sources that contribute to this infidelity. The largest contribution is from photon loss in the cavity modes, a local error. We can estimate this from the time of the protocol (from the initial cavity displacements until state tomography). This time is 5.6 μs (800 ns OCT SNAP pulse + 2092 ns pumps + 1200 ns selective π -pulses + 1500ns transmon readout = 5592 ns = t_{protocol}). During this time, a single photon loss event in either cavity would take us to the state $(|\alpha\rangle|-\alpha\rangle + |-\alpha\rangle|\alpha\rangle)/\sqrt{2}$, which orthogonal to our desired entangled state. The probability of this occurring is given by

$$p_{\text{loss}} = |\alpha|^2 t_{\text{protocol}} \left(\frac{1}{T_{1,a_1}} + \frac{1}{T_{1,a_2}} \right) \approx 5\%, \quad (7.65)$$

if we assume the cavity lifetimes remain unchanged throughout the protocol.

The next dominant source of error is a SPAM-type error that comes from our OCT

decode pulse used for mapping the logical state from cavity 2 onto transmon 2. We estimate this error to be $p_{\text{decode}} = 1.7\%$, estimated with the same method used in [Burkhart et al., 2021] by creating an ‘encode’ OCT pulse.

The last error we suspect is the ‘false positive’ rate of the vacuum check. This is defined to be

$$\mathcal{F}_{\text{DMM}} = \frac{p(gg|\text{bright})}{p(gg|\text{dark}) + p(gg|\text{bright})} \quad (7.66)$$

We can measure the numerator and denominator directly from Fig. ?? when we take a line-cut at $2\ \mu\text{s}$. We find $p(gg|\text{bright}) = 1\%$ and $p(gg|\text{dark}) = 70\%$, close to the theoretical value of $1 - 2e^{-|\alpha|^2} + e^{-2|\alpha|^2} \approx 73\%$ from the overlap with vacuum. The 3% difference is due to transmons being excited out of their ground states due to rethermalization and additional heating induced by the pump tones. This gives $\mathcal{F}_{\text{DMM}} = 2.1\%$.

Our total error budget is therefore

$$p_{\text{loss}} + p_{\text{decode}} + \mathcal{F}_{\text{DMM}} = 9\%, \quad (7.67)$$

in fairly good agreement with our measured fidelity at $\alpha = \sqrt{2}$.

This error model also informs us why we should expect an optimum value of α that maximizes the entanglement fidelity. At larger values of α , $p_{\text{false positive}}$ saturates to 1% whereas p_{loss} continues to increase quadratically. At smaller values of α , the success probability, set by $p(gg|\text{dark})$, drops rapidly, causing $p_{\text{false positive}}$ to increase rapidly as well.

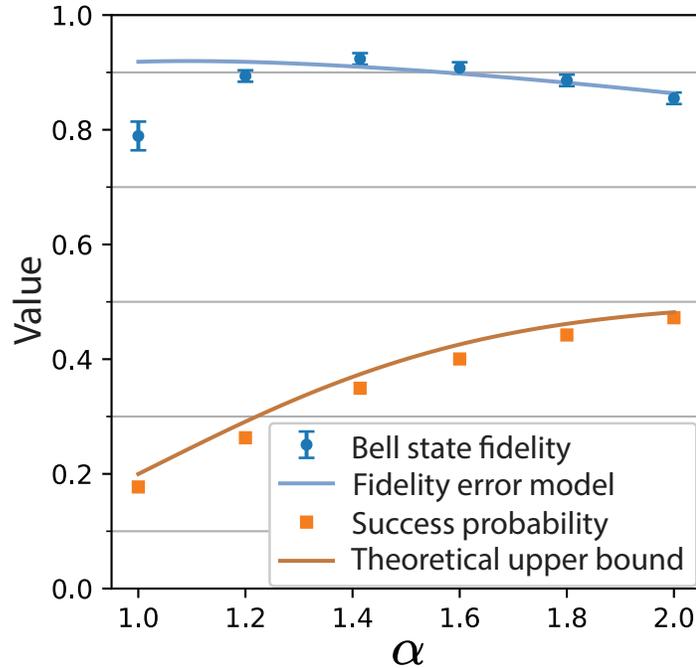


Figure 7.25: Measured Bell state fidelities and success probabilities for varying values of α . Solid blue line indicates expected Bell state fidelity from Eq. 7.67 Solid orange line is the success probability from Eq. 7.32. The fidelity error model does not account for the drop in fidelity observed at $\alpha = 1.2, 1.0$. An unknown process leads to higher than expected population of the vacuum state in cavity 1.

7.12.2 Bell state fidelities for different values of α

We repeat the DMM protocol for different values of α , which is achieved simply by adjusting the size of the initial cavity displacements and leaving the rest of the scheme unchanged. In Fig. 7.25 we show both the Bell state fidelity and success probability for the values $\alpha = 1.0, 1.2, 1.414, 1.6, 1.8, 2.0$, with no Kerr cancellation pump applied in all cases. The measured scheme success probability closely follows the theoretical upper limit, with an extra 2-3% additional failure probability due to imperfect preparation of the transmons in their ground state and heating of the transmons out of the ground state when the pumps are applied.

The measured Bell state fidelity is well-described by our model at most values of α , with the predicted drop-off at large α observed due to the increased probability of single photon loss in the cavities. At low values of α , we see a drop in measured

Basis parameters	$\alpha = 1.0$	$\alpha = 1.2$	$\alpha = 1.414$	$\alpha = 1.414$ Kerr-cancelled	$\alpha = 1.6$	$\alpha = 1.8$	$\alpha = 2.0$
α_{basis}	0.938	1.125	1.331	1.344	1.508	1.706	1.904
θ_K	0.467	0.473	0.493	0.001	0.502	0.502	0.502
θ_r	0.001	0.010	-0.749	0.009	0.119	-0.012	-0.014
Success probability	0.1772	0.263	0.347	0.349	0.400	0.442	0.472
Bell state fidelity	79.0%	89.5%	91.3%	92.3%	90.8%	88.6%	85.6%

Table 7.3: DMM scheme performance for different values of α .

Bell state fidelity that is not explained by the error model used to calculate the error-budget. Our reconstructed Wigner functions (and measurements of the cavity number population through transmon spectroscopy) suggest this is due to extra cavity population in the $|0\rangle$ state, but the exact mechanism for this is not yet known. This is not something we observe at values of $\alpha > 1.4$.

In Fig. 7.26 we show the measured Wigner functions (without Kerr correction) and Pauli bars. We see that the Wigner functions for $\alpha = 1.0$ look more noisy, due to a much lower success probability (and hence less shots averaged over which adds to a higher uncertainty in $\mathcal{F}_{\text{Bell}}$). We also see less Kerr distortion, due to the small \bar{n} . For $\alpha = 2.0$, this Kerr distortion is much more pronounced, but is compensated by updating the definition of cavity 1's logical basis.

In Tab. 7.3 we list the optimized logical basis parameters for cavity 1 (using the Nelder-Mead method), the success probability and the Bell state fidelity for all values of α used for the scheme. The value of α in the optimized basis is lower than the initial value of α , due to no-jump backaction in the $\sim 5 \mu\text{s}$ duration of the scheme. The value of θ_K is very similar for the different values of α , as expected since the cavity self-Kerrs, K_1 and K_2 , do not change with α .

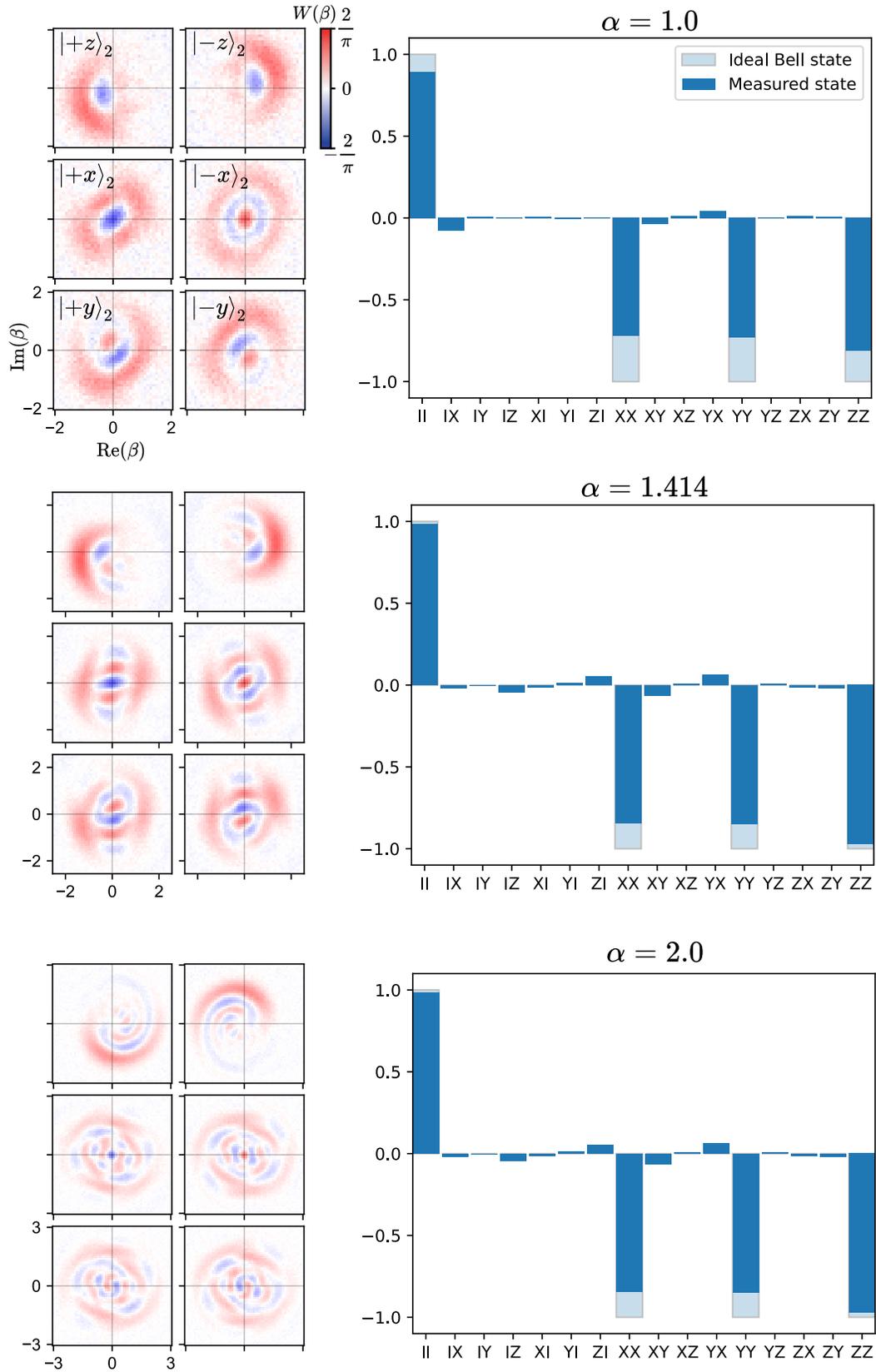


Figure 7.26: Measured Wigner functions of cavity 1 conditioned on cavity 2's logical measurement outcome and the reconstructed Pauli bars, for different values of α .

7.13 State Transfer via Teleportation

We can go one step further than entanglement generation, by using this Bell state as a resource in a quantum teleportation procedure. This is where the strength of quantum teleportation as a means of quantum state transfer becomes significant. In our modular network we have been confronted by a lossy link. We have developed a *probabilistic* scheme to generate a high-fidelity Bell state. Now we will use it to *deterministically* teleport a quantum state from one module to the other, avoiding the loss in the link as we do so.⁷

For this extension of the scheme we need a few more ingredients. Firstly, we need another quantum mode in which we prepare the qubit state we wish to teleport. For this we use the transmon in module 2 (which has better coherence than the transmon in module 1). After passing the vacuum check which heralds the Bell state, both transmons are in their ground state and so it is easy to initialize transmon 2 in an arbitrary qubit state $c_0 |g\rangle + c_1 |e\rangle$ with a single π -pulse.

As we saw from the ‘textbook’ quantum teleportation protocol discussed in Sec. 2.4.1 of Chapter 2, also what is needed is a Bell measurement between transmon 2 and cavity 2. The teleportation circuit we are implementing on our hardware is illustrated in Fig. 7.27. Ultimately we teleport the qubit state from transmon 2 to cavity 1, changing the logical basis as we do so. The state transfer fidelity is the bounded by the fidelity of our Bell state and the fidelity of our Bell measurement. We now describe how the Bell measurement can be implemented to high fidelity in our system.

7.13.1 Bell measurements in a module

In this realization of a Bell measurement, we first perform a CNOT gate between cavity 2 and transmon 2 and then measure transmon 2 in its X basis (m_1) and cavity

⁷ What remains probabilistic is how long we need to wait before the Bell state is successfully generated.

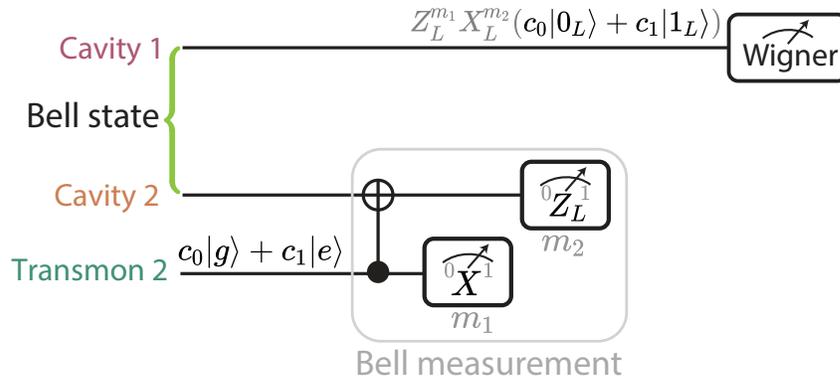


Figure 7.27: Circuit for quantum teleportation from transmon 2 to cavity 1.

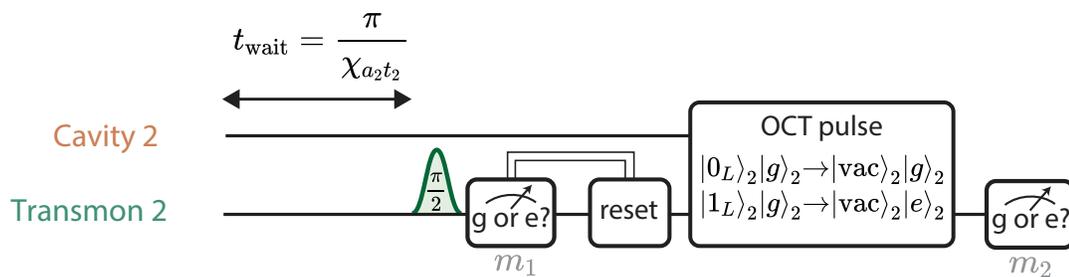


Figure 7.28: Circuit for implementing a Bell measurement between transmon 2 and cavity 2.

2 in its Z basis (m_2), no joint measurements are required.

Our strategy is to use the dispersive interaction to do the CNOT gate, measure transmon 2 in its X basis, then decode the logical state from cavity 2 onto the transmon and perform another transmon measurement. These two bits of information tell us the Pauli frame updates we need to apply to cavity 1, completing the teleportation. This is summarized by the circuit in Fig. 7.28. To see how this circuit works, it is helpful to consider what happens if cavity 2 is encoded in the cat code at large α , instead of our modified basis. The desired Bell measurement amounts to performing

a destructive measurement that distinguishes between the following four states:

$$\begin{aligned}
|\psi_1\rangle &\propto |\alpha\rangle_2 |g\rangle_2 + |-\alpha\rangle_2 |e\rangle_2 \\
|\psi_2\rangle &\propto |\alpha\rangle_2 |g\rangle_2 - |-\alpha\rangle_2 |e\rangle_2 \\
|\psi_3\rangle &\propto |\alpha\rangle_2 |e\rangle_2 + |-\alpha\rangle_2 |g\rangle_2 \\
|\psi_4\rangle &\propto |\alpha\rangle_2 |e\rangle_2 - |-\alpha\rangle_2 |g\rangle_2
\end{aligned} \tag{7.68}$$

The value of m_1 distinguishes between states $\{|\psi_1\rangle, |\psi_3\rangle\}$ and $\{|\psi_2\rangle, |\psi_4\rangle\}$ whereas the value of m_2 distinguishes between states $\{|\psi_1\rangle, |\psi_2\rangle\}$ and $\{|\psi_3\rangle, |\psi_4\rangle\}$. Once we know (m_1, m_2) , we know which state, $|\psi_i = 1, 2, 3, 4\rangle$ we started in, completing our Bell measurement. The CNOT gate is performed simply by waiting time $\pi/\chi_{a_2 t_2}$. Similar to a parity measurement sequence, this causes the states evolve as

$$\begin{aligned}
|\psi_1\rangle &\propto |\alpha\rangle_2 |g\rangle_2 + |-\alpha\rangle_2 |e\rangle_2 \rightarrow |\alpha\rangle_2 (|g\rangle + |e\rangle)_2 \\
|\psi_2\rangle &\propto |\alpha\rangle_2 |g\rangle_2 - |-\alpha\rangle_2 |e\rangle_2 \rightarrow |\alpha\rangle_2 (|g\rangle - |e\rangle)_2 \\
|\psi_3\rangle &\propto |\alpha\rangle_2 |e\rangle_2 + |-\alpha\rangle_2 |g\rangle_2 \rightarrow |-\alpha\rangle_2 (|g\rangle + |e\rangle)_2 \\
|\psi_4\rangle &\propto |\alpha\rangle_2 |e\rangle_2 - |-\alpha\rangle_2 |g\rangle_2 \rightarrow |-\alpha\rangle_2 (|g\rangle - |e\rangle)_2
\end{aligned} \tag{7.69}$$

We can see that we measure m_1 by measuring the transmon in the $|\pm\rangle$ basis after this wait time, which is achieved by a $\pi/2$ pulse about the Y axis followed by standard dispersive readout. We then conditionally reset the transmon to $|g\rangle$ if we measured $|e\rangle$ and measure cavity 2 in the $|\pm\alpha\rangle$ basis by using an OCT pulse to map these cavity states to the states of the transmon before another transmon readout. After successfully heralding a Bell state, the sequence takes an additional 6 μs to complete the local Bell measurement and teleport the state. To avoid idling errors during the rather long Bell measurement sequence, we begin measuring the Wigner function of cavity 1 as soon as the Bell state is heralded successfully (making this teleportation passive.) We take $\sim 2,800$ shots for each point in the Wigner function. (A 51×51

grid.) In our experiment, we also have to use the modified cavity 2 logical basis to define the OCT decode pulse (similar to the one used to measure entanglement fidelity but also compensating for additional Kerr evolution during the CNOT gate).

7.13.2 Average teleported state transfer fidelity

Average state transfer fidelity in the context of qubit transfer is defined to be the fidelity of the final transferred qubit to its initial state, averaged over randomly chosen qubit states. Fortunately, we do not need to average over all possible qubit states, but instead averaging over ‘cardinal’ states will suffice. As discussed in [Axline et al., 2018], a convenient choice of cardinal states is $\{|0\rangle, |1\rangle, |+\rangle, |-\rangle, |+i\rangle, |-i\rangle\}$. In our case, it suffices to average over the four states $\{|0\rangle, |1\rangle, |+\rangle, |+i\rangle\}$, provided we define our average to be

$$\mathcal{F}_{\text{QST}}^{\text{avg}} = \frac{\mathcal{F}_{\text{QST}}^{|0\rangle} + \mathcal{F}_{\text{QST}}^{|1\rangle} + 2\mathcal{F}_{\text{QST}}^{|+\rangle} + 2\mathcal{F}_{\text{QST}}^{|+i\rangle}}{6} \quad (7.70)$$

Determining the state transfer fidelities for the teleportation channel has the added complication that we need to average over the four different possible Bell state outcomes, which each have roughly 25% probability of occurring.

What is the correct order of the averaging, and how should we optimize the logical basis properly? First we note that we are allowed to choose one logical basis for cavity 1 for each measurement outcome but we must fix this logical basis across the qubit states being transferred. We denote the basis we are using with the subscript L_{00} for measurement outcome $(0, 0)$ etc. The state transfer fidelity averaged over all the Bell state outcomes is then given by

$$\begin{aligned}
\mathcal{F}_{\text{QST}}^{|\psi\rangle} &= P(0, 0) \langle \psi_{L_{00}} | \rho_{00}^{\text{meas}} | \psi_{L_{00}} \rangle \\
&+ P(0, 1) \langle \psi_{L_{01}} | X_{L_{01}} \rho_{01}^{\text{meas}} X_{L_{01}} | \psi_{L_{01}} \rangle \\
&+ P(1, 0) \langle \psi_{L_{10}} | Z_{L_{10}} \rho_{10}^{\text{meas}} Z_{L_{10}} | \psi_{L_{10}} \rangle \\
&+ P(1, 1) \langle \psi_{L_{11}} | Y_{L_{11}} \rho_{11}^{\text{meas}} Y_{L_{11}} | \psi_{L_{11}} \rangle
\end{aligned} \tag{7.71}$$

Where $P(m_1, m_2)$ is the probability of obtaining Bell measurement outcome (m_1, m_2) for the given state, $|\psi_{L(m_1, m_2)}\rangle$ is the state we were trying to transfer, given cavity 1's logical basis for measurement outcome (m_1, m_2) , $X_{L(m_1, m_2)}$ is the X logical Pauli operator on cavity 1 for the appropriate basis (and similar for the other Pauli operators), and $\rho_{(m_1, m_2)}^{\text{meas}}$ is the reconstructed density matrix from the measured Wigner function when we try to transfer state $|\psi\rangle$ and obtain Bell measurement outcome (m_1, m_2) .

From equations 7.71 and 7.70 we directly calculate the overall teleported average state transfer fidelity, reported in Tab. 7.4, with the measured Wigner functions in Fig. 7.30 and Fig. 7.29 for two teleportation configurations from which we quote $\mathcal{F}_{\text{QST}} = 89 \pm 1\%$ and $90 \pm 1\%$ respectively, indicating our Bell measurement contributes a $\sim 2\%$ error to the teleportation protocol and is still mostly limited by the Bell state fidelity. In both cases we begin with the Bell state for $\alpha = 1.414$. We find fidelity is persistently around 1% lower for all transferred states when we compensate for Kerr-distortion in cavity 1, since this pulse makes us slightly more susceptible to errors in transmon 1.

We can notice some trends from Tab. 7.4. Firstly, whenever we measure $m_1 = 1$ (transmon 2 in $|e\rangle$ in the first measurement) we obtain a lower transfer fidelity. This is likely due to errors in transmon 2 reset, due to transmon decay, which makes it more probable that m_2 is measured incorrectly resulting in an error. Transferring $|1\rangle$ also has a lower fidelity than $|0\rangle$ because initially transmon 2 is in $|e\rangle$ which has a

Kerr correction?	Metric	State prepared	Outcome (m_1, m_2)				Average	
			(0,0)	(0,1)	(1,0)	(1,1)		
No	Fidelity	$ 0\rangle$	0.9561	0.9644	0.9506	0.9142	0.9463	
		$ 1\rangle$	0.9000	0.9142	0.9206	0.9011	0.9090	
		$ +\rangle$	0.8915	0.8881	0.990	0.9190	0.9094	
		$ +i\rangle$	0.8991	0.8673	0.8693	0.8527	0.8721	
	$\mathcal{F}_{\text{QST}}^{\text{avg}}$					$90 \pm 1\%$		
	Probability	$ 0\rangle$	0.2485	0.2493	0.2516	0.2506		
		$ 1\rangle$	0.2508	0.2509	0.2496	0.2487		
		$ +\rangle$	0.2491	0.2501	0.2521	0.2487		
		$ +i\rangle$	0.2494	0.2497	0.2507	0.2501		
	Yes	Fidelity	$ 0\rangle$	0.9443	0.9681	0.9387	0.8981	0.9372
			$ 1\rangle$	0.8724	0.8990	0.8985	0.9102	0.8950
			$ +\rangle$	0.8915	0.8881	0.9274	0.9079	0.8998
$ +i\rangle$			0.8598	0.8608	0.8749	0.8472	0.8607	
$\mathcal{F}_{\text{QST}}^{\text{avg}}$						$89 \pm 1\%$		
Probability		$ 0\rangle$	0.2483	0.2495	0.2514	0.2507		
		$ 1\rangle$	0.2509	0.2506	0.2500	0.2485		
		$ +\rangle$	0.2491	0.2501	0.2521	0.2486		
		$ +i\rangle$	0.2494	0.2498	0.2508	0.2500		

Table 7.4: Teleportation fidelities for each cardinal state and the probability of obtaining each outcome.

probability of decaying during the CNOT gate.

In the first configuration, we cancel out the Kerr in cavity 1 by applying a SNAP pulse simultaneously during the vacuum check (See Eq. 7.49). This makes it easy to see the qualitative features of cavity 1’s Wigner functions after the teleportation, without the Kerr distortions but degrades the teleportation fidelity by around 1%, since we are prone to decoherence in transmon 1 during this pulse. In the second configuration, we omit the SNAP pulse.

The type of teleportation we do is deterministic and ‘passive’. This means that rather than actively applying the Pauli corrections to cavity 1, which are conditioned on the random measurement outcomes (m_1, m_2) , we instead track this in software.

Basis Parameters	Bell measurement outcome (m_1, m_2)							
	no Kerr-compensation				Kerr-compensation			
	(0, 0)	(0, 1)	(1, 0)	(1, 1)	(0, 0)	(0, 1)	(1, 0)	(1, 1)
α_{basis}	1.319	1.327	1.319	1.321	1.317	1.325	1.317	1.313
θ_K	0.515	0.509	0.474	0.471	-0.058	-0.072	-0.138	-0.129
θ_r	-0.012	0.000	0.113	0.123	0.198	0.226	0.381	0.366

Table 7.5: Optimized logical basis for cavity 1, for both teleportation configurations and for a given Bell measurement outcome. θ_K takes on a slightly lower value if $m_1 = 1$, due to Kerr evolution during the extra wait time needed for transmon 2 reset.

This also means the required update is very obvious from the measured Wigner functions, conditioned on each of the four possible (m_1, m_2) outcomes. This also means we do not have to wait for the Bell measurement to be completed before measuring the Wigner function of cavity 1. Instead, we measure it immediately proceeding successful heralding. This corresponds to the use case where we track the Pauli frame in software. We can proceed with the next step in our quantum algorithm provided this step does not need to know which ‘Pauli frame’ we are in. There are scenarios where we must update the Pauli frame in real time, which means actively applying Pauli operations to our qubits, e.g., before non-Clifford gates in a surface code.

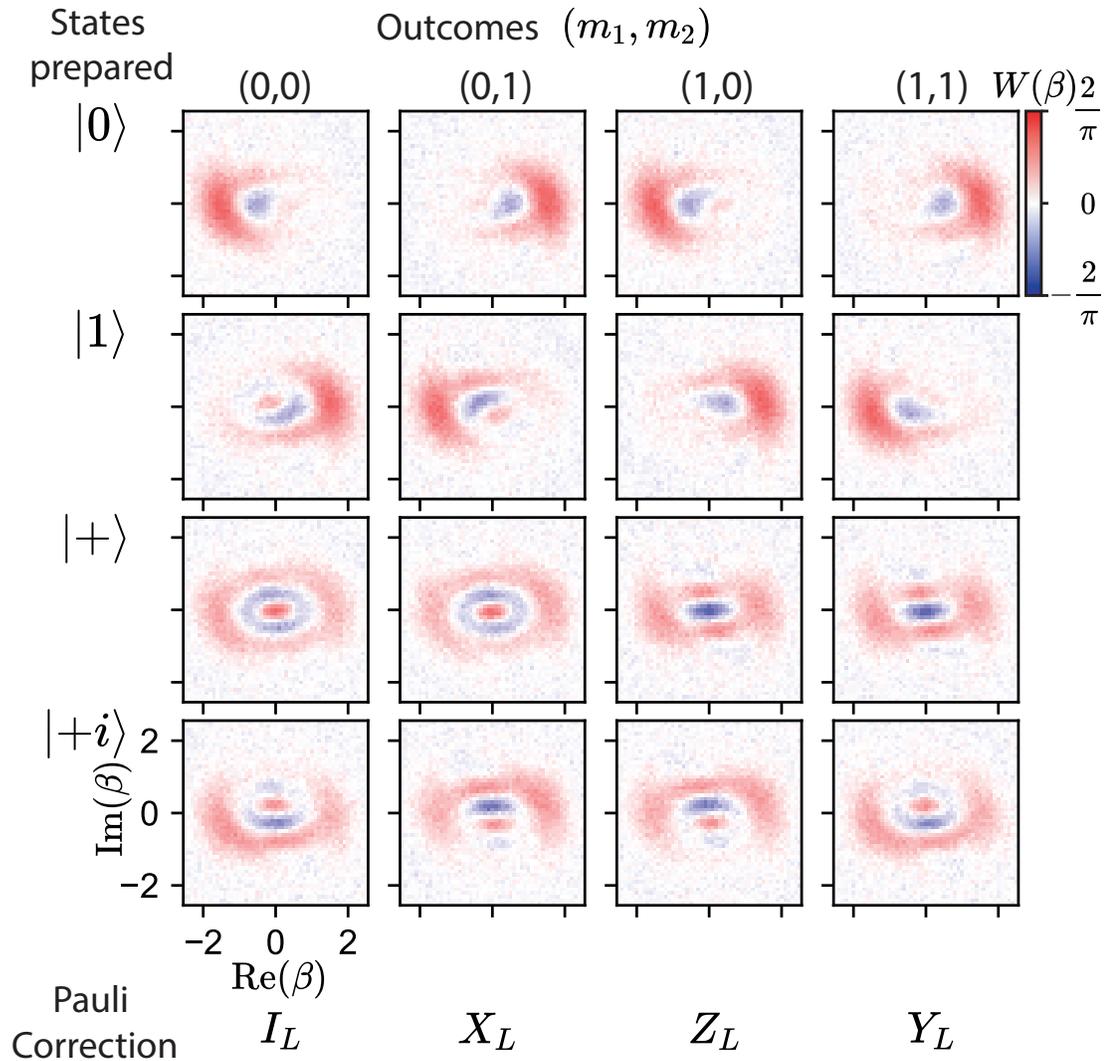


Figure 7.29: Measured Wigner functions after teleporting to cavity 1, with a Kerr cancellation pulse applied. Each row corresponds to a different cardinal state initially prepared in transmon 2. Each row corresponds to obtaining one of the four Bell state outcomes, (m_1, m_2) . Beneath each column we specify the Pauli correction required to ‘complete’ the teleportation. For ideal teleportation, applying the logical Pauli operator to each state should recover the leftmost column.

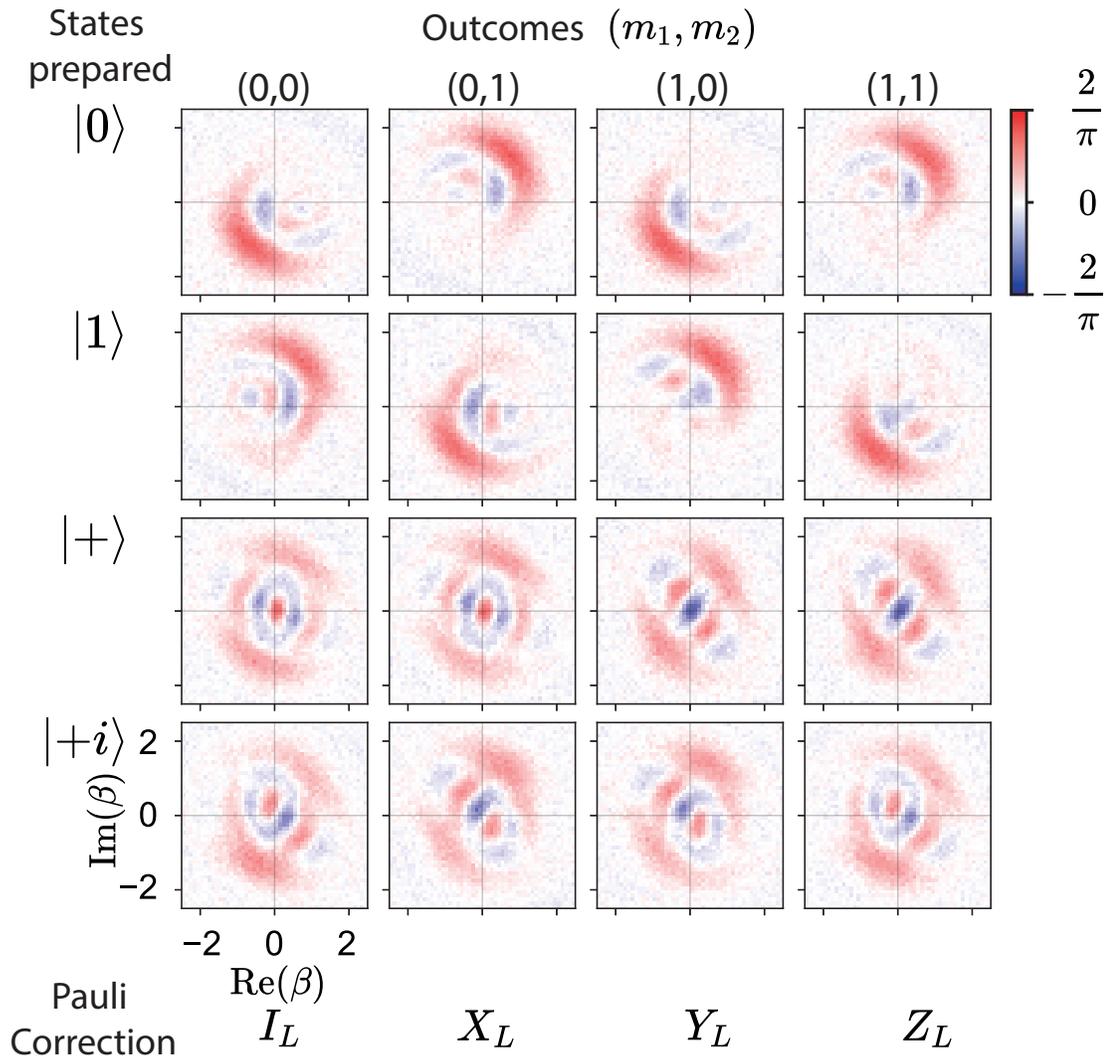


Figure 7.30: Measured Wigner functions after teleporting to cavity 1, without a Kerr cancellation pulse applied. This configuration results in a slightly improved \mathcal{F}_{QST} .

7.14 Conclusions

How should we interpret the experimental success in implementing the DMM scheme? At first glance, this seems like a specialized solution to a curiously hard and niche quantum communication problem; state transfer in the regime $g < \kappa < \Delta_{\text{FSR}}$. We point out this problem will likely become increasingly relevant as intrafridge microwave links become more commonplace, especially as qubit architectures move beyond a single module or chip that contains all the qubits.

Previous recent works [Burkhart et al., 2021, Zhong et al., 2019, Zhou et al., 2021] have shown that using a single-standing wave mode for inter-module quantum communication with superconducting qubits is a more natural choice than adapting schemes from quantum optics designed for ‘flying’ wavepackets. However, these schemes are still confronted with the usual fidelity ‘bottleneck’ when the link is lossy, set by $\bar{\eta} \sim \kappa/g$. Before the DMM scheme, the only way around this would be to resort to an adiabatic scheme — and the more lossy the bus mode, the longer the required scheme duration to mitigate this loss. The usual engineer’s solution would be to just maximize g and minimize the bus loss κ and indeed, this approach has seen remarkable success recently [Zhou et al., 2021].

The DMM scheme can be thought of a clear existence proof that we may overcome the ‘ κ/g ’ limit quite dramatically, without needing to resort to adiabatic schemes, if instead one is willing to use the trick of heralding from quantum optics. Using the paradigm of heralding for single-mode communication channels is a regime seldom explored in quantum communication, with this experiment being the first of its kind to our knowledge. But it solves the quantum communication problem in a regime we often find ourselves as experimentalists, when we wish our link had less loss. What is unique about the DMM scheme is how little the loss of the bus mode impacts the scheme (at least in the critically coupled and underdamped regime). In an optics scheme, we would at least expect κ to directly set the success probability but in the

DMM it remains 50% for any value of κ , thanks to the physics of a single-mode link which admits convenient interference effects. In the under-damped regime, the loss of the bus mode only affects the time to reset the scheme in the event that we are not successful in generating entanglement. In the overdamped regime, the scheme duration now scales as κ/g^2 but we are still not susceptible to loss of the bus mode. Interestingly, we are affected by loss in the bus mode due to the fact that a single-mode link is only an approximation, albeit a very good one for this experiment!

It is true that the DMM scheme requires some other specialist tools from bosonic QEC. It seems to only work well for the 2-legged cat code (and the modified basis) and requires QND ‘vacuum check’ measurements. Nevertheless, we presume other schemes exist and are yet to be found for the ‘heralded single-mode link’ regime of quantum communication.

How could the DMM scheme be improved in the future? At the moment, we are limited by ‘local loss’, that is, sources of error in the modules themselves rather than in the link. Many complications of the scheme can be circumvented by using a dedicated beamsplitter coupling element and a separate ancilla transmon in each module, the approach taken in [Burkhart et al., 2021]. This would allow us to operate in the ‘underdamped’ instead of ‘critically damped’ regime. With next-gen beamsplitter couplers [Lu et al., 2023, Chapman et al., 2022], we may only have to apply $\hat{\mathcal{H}}_c$ for a few hundred nanoseconds and most of the scheme duration would be cat state preparation and the vacuum check measurement!

Can the scheme be made deterministic? Here we briefly mention the adiabatic but deterministic version of the DMM scheme, which has since been proposed by others [Zapletal et al., 2022, Zhou et al., 2021]. In this version, we activate $\hat{\mathcal{H}}_c$ whilst also applying driven-dissipative stabilization of local cat states in each cavity. Local stabilization stabilizes the manifold of states $\{|\alpha, \alpha\rangle, |-\alpha, \alpha\rangle, |\alpha, -\alpha\rangle, |-\alpha, -\alpha\rangle\}$ and turning on $\hat{\mathcal{H}}_c$ makes the states $\{|\alpha, \alpha\rangle, |-\alpha, -\alpha\rangle\}$ ‘unstable’ causing them to

decay to vacuum. All these processes (local stabilization, $\hat{\mathcal{H}}_c$ and loss in the bus mode) also conserve the joint parity of the two cavities and so if we begin in the vacuum state (which has even joint parity) with $\hat{\mathcal{H}}_c$ activated, and slowly ramp up local cat stabilization in each cavity, we will stabilize the unique entangled state $|\Psi_+^{2\text{cat}}\rangle = \frac{1}{\sqrt{2}}(|\alpha, -\alpha\rangle + |-\alpha, \alpha\rangle)$, which is the only state within the manifold that has even joint parity.

For this scheme to work one must adhere to several rate hierarchies. Firstly, the bus mode must be ‘adiabatically eliminated’ such that it can be approximated as being in its vacuum state at all times. For this we need $\kappa_b \gtrsim g_0$. We also need the rate of local stabilization, $\kappa_{2\text{cat}} < g_0$ and to ramp up this stabilization slowly. This scheme was in fact the one we originally envisioned implementing in our hardware (hence empty tunnels in each module) but there are two major drawbacks that make experimental implementation difficult. Firstly, the number of pumps in the system to realize local cat stabilization and $\hat{\mathcal{H}}_c$ simultaneously requires a total of eight pump tones which is nontrivial to implement. Secondly, respecting adiabaticity conditions results in a scheme duration predicted to be 10-20 μs for our achievable g_0 rates, during which we are susceptible to single photon loss in both cavities. These schemes are the ‘automated versions’ of the DMM scheme and may be revisited in the future if faster beamsplitter rates (paired with a sufficiently lossy bus) can shorten the scheme duration. The DMM scheme takes a similar amount of time to ‘deterministically’ generate entanglement, requiring, e.g., 4 attempts on average to have a $> 90\%$ success rate.

What is the prognosis for near-term inter-module links? The DMM scheme is rather specialized to the two-legged cat code and its inherently probabilistic nature (even if its success probability is much higher than most optics realizations) is still inconvenient when used to link quantum processing together — syncing up gate operations becomes tedious when some steps take an unknown amount of time. In the

near-term, there is still more mileage that can be gained from improving the κ/g ratio, especially with the advent of faster beamsplitter couplers and low loss Al coaxial cable with carefully engineered joints to minimize resistance. Schemes such as the virtual Raman scheme and STIRAP become increasingly viable as we increase g_{BS} . Furthermore, from Chapter 6 we know that photon loss errors can merely be classed as erasure errors within the dual-rail code, for which we have a much higher error tolerance. Error-detected state transfer with these new qubits and mid-circuit erasure detection would be able to flag the vast majority of photon loss errors as erasures.

Finally, we will touch on one scenario where the DMM scheme may be useful. This is when our link has some additional desirable property, but also comes with a considerable amount of link loss. The cable may be particularly flexible, with the ability to support many separate links in parallel (i.e., a ribbon cable) or it may include on-chip elements such as a router which boost connectivity beyond two-nodes but presents an additional source of loss. So long as the link can be modelled well as a single standing wave mode, and we can actuate couplings to the two bosonic modes of interest, the DMM scheme is viable.

Outlook

We end this thesis with some of my speculative thoughts on the future of the work presented. At the core of this thesis is a new paradigm shift to orient around error-detection rather than error-correction at the small scale, which by all accounts should ease the task of quantum error correction at the large scale. It will naturally take a while to fully appreciate and discover how this changes and simplifies the hardware we now need to realize as experimentalists.

The proposed operations conform to the error-hierarchy as best they can, but they are not immutable. The requirement of error-detection offers a surprising amount of freedom in designing operations, which are now characterized by many metrics beyond just a single number for the fidelity. The best way for performing the two-qubit gate, mid-circuit erasure detection, and state preparation may look quite different than the initial set of operations we have presented.

The real test will of course lie in experimental realization. The proposed operations were designed to account for as many known errors as possible, and early implementations of the dual-rail proposal suggest our assumptions are not so far detached from reality. The spectre of transmon errors continues to haunt even dual-rail qubits and this is something that will have to be extensively managed through every operation. Nevertheless, I am confident each of the basic building blocks of error-detected operations, at least for dual-rail cavity qubits, will be demonstrated by the end of this year. The source of this optimism lies with the fact that little of the underlying hardware is novel. Aside from high-fidelity beamsplitter interactions, the rest of the required hardware has existed for close to a decade already and is well-studied.

Further down the line is the challenge of scaling up to large numbers of error-detected qubits. Dual-rail qubits are undoubtedly less hardware-efficient than previous approaches with bosonic qubits with one qubit per cavity mode, but I would still claim they are hardware-efficient compared to, e.g., an all-transmon architecture. Compared to this architecture, the number of modes per physical qubit has roughly doubled, but we are rewarded by efficiently correcting all first-order hardware errors by converting them to erasure errors. This is a good trade-off that I would claim is still hardware-efficient. Making the ~ 50 dual-rail cavity qubits needed to demonstrate practical error correction with a surface code will be a monumental challenge, but is no more difficult than the challenge of scaling any 3D-microwave cavity architecture, with the potential of minutarization with coaxial line architectures [Axline et al., 2016].

The area I believe urgently requires more attention (without needing to actually make anything) is how to implement large-scale error correction with dual-rail cavity qubits, and erasure qubits in general. There are several twists in the circuit-level error models that are seldom considered in analysis of the surface code (or similar stabilizer codes).

Some of the questions I would be interested to know the answers to are:

- What happens if I miss a leakage error but detect it later?
- What happens if ancilla physical qubit readout errors are the dominant error
- What happens with ‘inert’ leakage errors, when leaked qubits no longer interact with any other qubits. Can we detect them from stabilizer measurements alone?
- What happens if a detected error has affected multiple data qubits?
- Can we perform erasure detection at the end of a stabilizer measurement sequence?

Answering these questions involves not only simulating these kinds of errors, but developing new decoders tailored to these specific error models. The co-design of large-scale dual-rail hardware with tailored decoders is not something I expect to be easy, but is necessary if we are to take full advantage of the structured noise in these new qubits. Ultimately, we need to run the numbers. Given everything we know about the proposed building blocks, their duration, their errors, what is the circuit we run to measure stabilizers and detect errors? And what would the logical error rate be?

The next big milestone for error correction would be what I am calling Clifford Gate Break-even. (maybe CGB for short?) This is when the gate fidelity of *logical* Clifford gates is better than the fidelity of physical qubit Clifford gates. We deliberately omit non-Clifford gates from our definition since these are not expected to be implemented fault-tolerantly. To date, CGB has not been achieved in any device, which really speaks as to how difficult quantum error correction is. And the real goal is to go far beyond break-even to suppress logical gate errors by many orders of magnitude.

Beyond break-even has been conclusively demonstrated in bosonic qubits for a memory operation but CGB is much harder. In the surface code, the operations needed to preserve a quantum memory are also practically the same operations needed to also perform Clifford gates (through lattice surgery). Thus, memory break-even in a surface code also implies CGB should be possible in the same hardware. Because erasure qubits allow hardware errors at the 1% level and can still be well below threshold, I predict an erasure qubit architecture¹ will be the first to reach CGB. The task of building a quantum computer would still be far than over! We need to go far beyond break-even, ideally with an error-correction gain $\sim 10^{12}$, and then we would have succeeded in realizing just one logical qubit.

Where does modularity fit into all this? It is hard to imagine building a large-scale superconducting qubit processor without some degree of modularity built-in. The work of Chapter 7 has shown there is more to this problem than meets the eye, and it is indeed possible to circumvent loss in a microwave setting with heralding, just as it is possible in an optics setting. But the work of Chapter 6 also implies that loss errors in the link could be converted to erasure errors as well. Perhaps there is more leniency than we thought in how good our inter-module links need to be.

I like to view quantum error correction as an immense problem that can be tackled from two ends: by advancing theory and by advancing the experimental hardware. Often the largest advancements are made by taking some aspect from the theory side and applying it to the hardware, and vice versa. Error detection motivating the development of erasure qubits is a nice example of this – who knows what else can be found?

1. Or at least a qubit platform with highly structured noise.

Appendices

Correcting for erasure errors in the 4-qubit erasure code

We show how to fully correct an erasure error in the 4-qubit erasure code. We define a logical qubit with the codewords

$$|0_L\rangle = \frac{1}{\sqrt{2}} (|0000\rangle + |1111\rangle) \tag{A.1}$$

$$|1_L\rangle = \frac{1}{\sqrt{2}} (|1001\rangle + |0110\rangle) \tag{A.2}$$

$$\tag{A.3}$$

with the stabilizers

$$S_1 = Z_1 Z_2 Z_3 Z_4 \tag{A.4}$$

$$S_2 = X_1 X_2 X_3 X_4 \tag{A.5}$$

$$\tag{A.6}$$

We are also able to encode and correct another logical qubit in the manifold $\{\frac{1}{\sqrt{2}}(|0011\rangle + |1100\rangle), \frac{1}{\sqrt{2}}(|0101\rangle + |1010\rangle)\}$. The procedure we will show to correct an erasure will preserve this qubit as well, but in this example (and to simplify the maths) we leave this manifold unused.

Our initial qubit state is $|\psi_L\rangle = c_0 |0_L\rangle + c_1 |1_L\rangle$. Our initial state is then

$$|\psi_L\rangle = c_0 \frac{1}{\sqrt{2}} (|0000\rangle + |1111\rangle) + c_1 \frac{1}{\sqrt{2}} (|1001\rangle + |0110\rangle) \quad (\text{A.7})$$

It will be useful later to write out the full density matrix for this state

$$\begin{aligned} \rho_L &= \frac{1}{2} (c_0 |0000\rangle + c_0 |1111\rangle + c_1 |1001\rangle + c_1 |0110\rangle) \\ &\quad (c_0^* \langle 0000| + c_0^* \langle 1111| + c_1^* \langle 1001| + c_1^* \langle 0110|) \\ &= \frac{|c_0|^2}{2} (|0000\rangle \langle 0000| + |1111\rangle \langle 0000| + |0000\rangle \langle 1111| + |1111\rangle \langle 1111|) \\ &\quad + \frac{|c_1|^2}{2} (|1001\rangle \langle 1001| + |0110\rangle \langle 1001| + |1001\rangle \langle 0110| + |0110\rangle \langle 0110|) \\ &\quad + \frac{c_0 c_1^*}{2} (|0000\rangle \langle 1001| + |0000\rangle \langle 0110| + |1111\rangle \langle 1001| + |1111\rangle \langle 0110|) \\ &\quad + \frac{c_1 c_0^*}{2} (|1001\rangle \langle 0000| + |0110\rangle \langle 0000| + |1001\rangle \langle 1111| + |0110\rangle \langle 1111|) \end{aligned}$$

Our goal is to preserve the logical qubit state up to a Pauli frame update in the event that a single qubit undergoes an erasure error. Erasures are known errors, without loss of generality, let us suppose physical qubit 1 undergoes an erasure error. The error can be anything we want, as long as it only affects qubit 1 and we know that it has happened. We represent the most generic error by tracing out qubit 1:

$$\begin{aligned}
\rho_{\text{reduced}} &= \langle 0_1 | \rho_L | 0_1 \rangle + \langle 1_1 | \rho_L | 1_1 \rangle \\
&= |c_0|^2 (|000\rangle \langle 000| + |111\rangle \langle 111|) + |c_1|^2 (|110\rangle \langle 110| + |001\rangle \langle 001|) \\
&\quad + c_0 c_1^* (|000\rangle \langle 110| + |111\rangle \langle 001|) + c_1 c_0^* (|110\rangle \langle 000| + |001\rangle \langle 111|)
\end{aligned}$$

Now we reset qubit 1 to the codespace. In this example we use the mixed state $P_0 |0\rangle \langle 0| + P_1 |1\rangle \langle 1|$ where P_0, P_1 are positive and $P_0 + P_1 = 1$.

The state is now

$$\rho = P_0 |0\rangle \langle 0| + P_1 |1\rangle \langle 1| \otimes \rho_{\text{reduced}} \quad (\text{A.8})$$

By measuring the stabilizers S_1 and S_2 we re-entangle qubit 1 with the other qubits and restore $|\psi_L\rangle$ up to a Pauli frame update.

Let us consider the state after we measure S_1 and we obtain the outcome $S_1 = +1$ (with 50% probability). This means we only ‘keep’ the terms which have an even number of 1s and 0s in the kets. Now we are in the state

$$\rho = P_0 (|c_0|^2 |0000\rangle \langle 0000| + c_0 c_1^* |0000\rangle \langle 0110| + c_1 c_0^* |0110\rangle \langle 0000| + |c_1|^2 |0110\rangle \langle 0110|) \quad (\text{A.9})$$

$$= P_1 (|c_0|^2 |1111\rangle \langle 1111| + c_0 c_1^* |1111\rangle \langle 1001| + c_1 c_0^* |1001\rangle \langle 1111| + |c_1|^2 |1001\rangle \langle 1001|) \quad (\text{A.10})$$

This density matrix only contains eight terms and we need to get to sixteen to restore $|\psi_L\rangle$. Next we measure S_2 which increases the number of terms. If we obtain outcome $S_2 = +1$ then it is as though we applied the projector $(\mathbb{1} + S_2)/2$ to our

states. E.g.

$$\begin{aligned} & \left(\frac{\mathbb{1} + XXXX}{2} \right) |0000\rangle \langle 0110| \left(\frac{\mathbb{1} + XXXX}{2} \right) \\ &= \frac{1}{4} (|0000\rangle \langle 0110| + |1111\rangle \langle 1001| + |0000\rangle \langle 1001| + |1111\rangle \langle 0110|) \end{aligned}$$

If we do this for each term in Eq. A.9 and collect like terms we end up with the final density matrix

$$\begin{aligned} & \frac{|c_0|^2}{2} (P_0 + P_1)(|0000\rangle + |1111\rangle)(\langle 0000| + \langle 1111|) \\ &+ \frac{|c_1|^2}{2} (P_0 + P_1)(|0110\rangle + |1001\rangle)(\langle 0110| + \langle 1001|) \\ &+ \frac{c_0 c_1^*}{2} (P_0 + P_1)(|0000\rangle \langle 0110| + |0000\rangle \langle 1001| + |1111\rangle \langle 0110| + |1111\rangle \langle 1001|) \\ &+ \frac{c_1 c_0^*}{2} (P_0 + P_1)(|0110\rangle \langle 0000| + |1001\rangle \langle 0000| + |0110\rangle \langle 1111| + |1001\rangle \langle 1111|) \\ &= \rho_L \end{aligned}$$

and we have corrected the erasure error and recovered our initial state. If we obtained the stabilizer outcomes $m_1, m_2 = -1, +1$ our final state would be $Z_1 |\psi_L\rangle$. If we obtained the stabilizer outcomes $m_1, m_2 = +1, -1$ our final state would be $X_1 |\psi_L\rangle$. In general, if physical qubit i undergoes erasure then after we reinitialize and re-measure the stabilizers, the qubits will be in state $Z_i^{\left(\frac{1-m_1}{2}\right)} X_i^{\left(\frac{1-m_2}{2}\right)} |\psi_L\rangle$. We call $Z_i^{\left(\frac{1-m_1}{2}\right)} X_i^{\left(\frac{1-m_2}{2}\right)}$ the ‘Pauli frame update’ which can be tracked in software. This procedure generalizes to any stabilizer code used for the erasure channel. The reason why a stabilizer code can tolerate twice as many erasure errors as Pauli errors is because we know exactly which qubits require a Pauli-frame update after measuring the stabilizers. Without this knowledge (i.e., in the case of Pauli errors) we end up in scenarios where our stabilizer measurements tell us an error has occurred, but it is ambiguous as to which physical qubits we should apply the frame update to.

Magic values of α in the 4-legged cat code

The 4-legged cat code is made from superpositions of 4 coherent states instead of 2. Many unwanted effects that are noticeable at smaller values of α are bounded by $e^{-|\alpha|^2}$, rather than $e^{-2|\alpha|^2}$ as with the two-legged cat code. Although they both are suppressed exponentially with α , the extra factor of two means 4-legged cats become ‘ideal’ codewords at $\alpha \approx 3$ rather than $\alpha \approx 2$, giving this code a large \bar{n} overhead.

Correcting for no-jump backaction is non-trivial in the four-legged cat code, as it results in the gradual loss of energy in the system. In theory, this can be counteracted by stabilizing the manifold of four coherent states $\{|\alpha\rangle, |i\alpha\rangle, |-\alpha\rangle, |-i\alpha\rangle\}$. In practice, this is incredibly difficult to achieve with engineered dissipation [Mirrahimi et al., 2014] or Hamiltonians. Stabilization via ‘telecorrection’ [Hastrup and Andersen, 2022] is an alternative way of re-inflating cat states which may alleviate the demanding Hamiltonian engineering, but requires Bell states and Bell measurements in the cat-code basis.

The 4-legged cat code only satisfies the Knill-Laflamme conditions for error cor-

rection exactly in the limit of large α (typically $\alpha > 3$), in which $N_0, N_1, N_2, N_3 \rightarrow 2$. At smaller values of α , the value of \bar{n} is different for each codeword, and so no-jump backaction results not just in the usual reduction in $\alpha \rightarrow \alpha e^{-\kappa t}$ but also in polarization towards the codeword with the smaller \bar{n} . This rate of polarization depends on both the average number of photons in both the codewords, \bar{n}_{avg} and the difference in \bar{n} in each codeword, which we call $\Delta\bar{n}$. At short times, the logical error due to polarization is $O((\Delta\bar{n}\kappa\delta t)^2)$.

For certain values of α which are not in the large α limit, we can exactly satisfy $\Delta\bar{n} = 0$ [Li et al., 2017]. These values of α are termed ‘magic α s’ and are found by solving either $N_0 = N_2$ for even codewords or $N_1 = N_3$ for odd codewords. The transcendental equations that give the first set of magic α s are

$$\begin{aligned}\tan(|\alpha_{\text{even}}|^2) &= -\tanh(|\alpha_{\text{even}}|^2) \\ \cot(|\alpha_{\text{odd}}|^2) &= \tanh(|\alpha_{\text{odd}}|^2)\end{aligned}\tag{B.1}$$

with solutions $\alpha_{\text{even}} \approx 1.538, 2.345, 2.939, \dots$ and $\alpha_{\text{odd}} \approx 0.968, 1.982, 2.659$

Of course, no-jump backaction continuously shrinks the value of α , and values in the set of $\{\alpha_{\text{even}}\}$ is unfortunately not close $\{\alpha_{\text{odd}}\}$ at small values of α and so parity jumps, despite being trackable, still move us away from magic values of α .

We can define another set of ‘magic α s’ for which the $|\pm_L\rangle$ codewords are exactly two-legged cats. i.e.

$$\begin{aligned}|+_L\rangle &\propto |\alpha\rangle + |-\alpha\rangle \\ |-_L\rangle &\propto |i\alpha\rangle + |-i\alpha\rangle\end{aligned}\tag{B.2}$$

This is desirable since four ($\{|0\rangle, |1\rangle, |+\rangle, |-\rangle\}$) out of the six cardinal states can be prepared by first preparing coherent states such as $|\alpha\rangle$ or $|i\alpha\rangle$ and then performing QND photon number measurements. For example, $|\pm\rangle$ can be prepared from

QND photon number parity measurements exactly when $\alpha'_{\text{even}} = \sqrt{(n + 1/2)\pi} = 1.253, 2.171, 2.802\dots$ for integer n and for odd codewords when $\alpha'_{\text{odd}} = \sqrt{n\pi} = 1.772, 2.507, 2.802\dots$

To summarise, the 4-legged cat code is adept at tracking photon jumps but no-jump backaction and subtle effects at small α complicate its use as a fully error-correcting code. Many of these effects can be alleviated at $\alpha \approx 3$ or greater, but this results in a large \bar{n} overhead ($\bar{n} \approx 9$ vs $\bar{n} = 0.5$ for the 01 encoding.) and more demands on its physical implementation.

Bibliography

- [Aasen et al., 2023] Aasen, David, Haah, Jeongwan, Bonderson, Parsa, Wang, Zhenghan, and Hastings, Matthew (2023). Fault-Tolerant Hastings-Haah Codes in the Presence of Dead Qubits. arXiv:2307.03715 [quant-ph]. (Cited on page 260).
- [Acharya et al., 2023] Acharya, Rajeev, Aleiner, Igor, Allen, Richard, Andersen, Trond I., Ansmann, Markus, Arute, Frank, Arya, Kunal, Asfaw, Abraham, Atalaya, Juan, Babbush, Ryan, Bacon, Dave, Bardin, Joseph C., Basso, Joao, Bengtsson, Andreas, Boixo, Sergio, Bortoli, Gina, Bourassa, Alexandre, Bovaird, Jenna, Brill, Leon, Broughton, Michael, Buckley, Bob B., Buell, David A., Burger, Tim, Burkett, Brian, Bushnell, Nicholas, Chen, Yu, Chen, Zijun, Chiaro, Ben, Cogan, Josh, Collins, Roberto, Conner, Paul, Courtney, William, Crook, Alexander L., Curtin, Ben, Debroy, Dripto M., Del Toro Barba, Alexander, Demura, Sean, Dunsworth, Andrew, Eppens, Daniel, Erickson, Catherine, Faoro, Lara, Farhi, Edward, Fatemi, Reza, Flores Burgos, Leslie, Forati, Ebrahim, Fowler, Austin G., Foxen, Brooks, Giang, William, Gidney, Craig, Gilboa, Dar, Giustina, Marissa, Grajales Dau, Alejandro, Gross, Jonathan A., Habegger, Steve, Hamilton, Michael C., Harrigan, Matthew P., Harrington, Sean D., Higgott, Oscar, Hilton, Jeremy, Hoffmann, Markus, Hong, Sabrina, Huang, Trent, Huff, Ashley, Huggins, William J., Ioffe, Lev B., Isakov, Sergei V., Iveland, Justin, Jeffrey, Evan, Jiang, Zhang, Jones, Cody, Juhas, Pavol, Kafri, Dvir, Kechedzhi, Kostyantyn, Kelly, Julian, Khattar, Tanuj, Khezri, Mostafa, Kieferová, Mária, Kim, Seon, Kitaev, Alexei, Klimov, Paul V., Klots, Andrey R., Korotkov, Alexander N., Kostritsa, Fedor, Kreikebaum, John Mark, Landhuis, David, Laptev, Pavel, Lau, Kim-Ming, Laws, Lily, Lee, Joonho, Lee, Kenny, Lester, Brian J., Lill, Alexander, Liu, Wayne, Locharla, Aditya, Lucero, Erik, Malone, Fionn D., Marshall, Jeffrey, Martin, Orion, McClean, Jarrod R., McCourt, Trevor, McEwen, Matt, Megrant, Anthony, Meurer Costa, Bernardo, Mi, Xiao, Miao, Kevin C., Mohseni, Masoud, Montazeri, Shirin, Morvan, Alexis, Mount, Emily, Mruczkiewicz, Wojciech, Naaman, Ofer, Neeley, Matthew, Neill, Charles, Nersisyan, Ani, Neven, Hartmut, Newman, Michael, Ng, Jiun How, Nguyen, Anthony, Nguyen, Murray, Niu, Murphy Yuezhen, O'Brien, Thomas E., Opremcak, Alex, Platt, John, Petukhov, Andre, Potter, Rebecca, Pryadko, Leonid P., Quintana, Chris, Roushan, Pedram, Rubin, Nicholas C., Saei, Negar, Sank, Daniel, Sankaragomathi, Kannan, Satzinger, Kevin J., Schurkus, Henry F., Schuster, Christopher, Shearn, Michael J., Shorter,

- Aaron, Shvarts, Vladimir, Skruzny, Jindra, Smelyanskiy, Vadim, Smith, W. Clarke, Sterling, George, Strain, Doug, Szalay, Marco, Torres, Alfredo, Vidal, Guifre, Villalonga, Benjamin, Vollgraft Heidweiller, Catherine, White, Theodore, Xing, Cheng, Yao, Z. Jamie, Yeh, Ping, Yoo, Juhwan, Young, Grayson, Zalcman, Adam, Zhang, Yaxing, Zhu, Ningfeng, and Google Quantum AI. [Suppressing quantum errors by scaling a surface code logical qubit](#). *Nature*, 614(7949):676–681 (2023). (Cited on pages 210, 241, and 243).
- [Aharonov and Ben-Or, 1997] Aharonov, Dorit and Ben-Or, Michael. Fault-tolerant quantum computation with constant error. In *Proceedings of the twenty-ninth annual ACM symposium on Theory of computing*, pages 176–188 (1997). (Cited on pages 3 and 66).
- [Aliferis and Terhal, 2006] Aliferis, Panos and Terhal, Barbara M. (2006). Fault-Tolerant Quantum Computation for Local Leakage Faults. arXiv:quant-ph/0511065. (Cited on page 58).
- [Alway and Jones, 2007] Alway, William G. and Jones, Jonathan A. [Arbitrary precision composite pulses for NMR quantum computing](#). *Journal of Magnetic Resonance*, 189(1):114–120 (2007). (Cited on page 227).
- [Auger et al., 2017] Auger, James M., Anwar, Hussain, Gimeno-Segovia, Mercedes, Stace, Thomas M., and Browne, Dan E. [Fault-tolerance thresholds for the surface code with fabrication errors](#). *Physical Review A*, 96(4):042316 (2017). (Cited on page 260).
- [Axline et al., 2016] Axline, C., Reagor, M., Heeres, R., Reinhold, P., Wang, C., Shain, K., Pfaff, W., Chu, Y., Frunzio, L., and Schoelkopf, R. J. [An architecture for integrating planar and 3D cQED devices](#). *Applied Physics Letters*, 109(4):042601 (2016). (Cited on pages 112 and 356).
- [Axline et al., 2018] Axline, Christopher J., Burkhardt, Luke D., Pfaff, Wolfgang, Zhang, Mengzhen, Chou, Kevin, Campagne-Ibarcq, Philippe, Reinhold, Philip, Frunzio, Luigi, Girvin, S. M., Jiang, Liang, Devoret, M. H., and Schoelkopf, R. J. [On-demand quantum state transfer and entanglement between remote microwave cavity memories](#). *Nature Physics*, 14(7):705–710 (2018). (Cited on pages 276, 332, and 345).
- [Bartolucci et al., 2023] Bartolucci, Sara, Birchall, Patrick, Bombín, Hector, Cable, Hugo, Dawson, Chris, Gimeno-Segovia, Mercedes, Johnston, Eric, Kielsing, Konrad, Nickerson, Naomi, Pant, Mihir, Pastawski, Fernando, Rudolph, Terry, and Sparrow, Chris. [Fusion-based quantum computation](#). *Nature Communications*, 14(1):912 (2023). (Cited on pages 35 and 256).
- [Battistel et al., 2023] Battistel, Francesco, Chamberland, Christopher, Johar, Kauser, Overwater, Ramon W. J., Sebastiano, Fabio, Skoric, Luka, Ueno, Yosuke,

- and Usman, Muhammad (2023). Real-Time Decoding for Fault-Tolerant Quantum Computing: Progress, Challenges and Outlook. arXiv:2303.00054 [quant-ph]. (Cited on page 66).
- [Beenakker et al., 2004] Beenakker, C. W. J., DiVincenzo, D. P., Emary, C., and Kindermann, M. [Charge Detection Enables Free-Electron Quantum Computation](#). *Physical Review Letters*, 93(2):020501 (2004). (Cited on page 34).
- [BELL, 1966] BELL, JOHN S. [On the Problem of Hidden Variables in Quantum Mechanics](#). *Reviews of Modern Physics*, 38(3):447–452 (1966). (Cited on page 40).
- [Bennett and Brassard, 2014] Bennett, Charles H. and Brassard, Gilles. [Quantum cryptography: Public key distribution and coin tossing](#). *Theoretical Computer Science*, 560:7–11 (2014). (Cited on page 266).
- [Bennett et al., 1993] Bennett, Charles H., Brassard, Gilles, Crépeau, Claude, Jozsa, Richard, Peres, Asher, and Wootters, William K. [Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels](#). *Physical Review Letters*, 70(13):1895–1899 (1993). (Cited on page 40).
- [Bennett et al., 1996] Bennett, Charles H., Brassard, Gilles, Popescu, Sandu, Schumacher, Benjamin, Smolin, John A., and Wootters, William K. [Purification of Noisy Entanglement and Faithful Teleportation via Noisy Channels](#). *Physical Review Letters*, 76(5):722–725 (1996). (Cited on page 42).
- [Bonilla Ataides et al., 2021] Bonilla Ataides, J. Pablo, Tuckett, David K., Bartlett, Stephen D., Flammia, Steven T., and Brown, Benjamin J. [The XZZX surface code](#). *Nature Communications*, 12(1):2172 (2021). (Cited on page 62).
- [Braunstein and Mann, 1995] Braunstein, Samuel L. and Mann, A. [Measurement of the Bell operator and quantum teleportation](#). *Physical Review A*, 51(3):R1727–R1730 (1995). (Cited on page 40).
- [Bravyi et al., 2022] Bravyi, Sergey, Dial, Oliver, Gambetta, Jay M., Gil, Darío, and Nazario, Zaira. [The future of quantum computing with superconducting qubits](#). *Journal of Applied Physics*, 132(16):160902 (2022). (Cited on page 267).
- [Bravyi and Kitaev, 1998] Bravyi, S. B. and Kitaev, A. Yu (1998). Quantum codes on a lattice with boundary. arXiv:quant-ph/9811052. (Cited on page 62).
- [Briegel et al., 1998] Briegel, H.-J., Dür, W., Cirac, J. I., and Zoller, P. [Quantum Repeaters: The Role of Imperfect Local Operations in Quantum Communication](#). *Physical Review Letters*, 81(26):5932–5935 (1998). (Cited on page 42).
- [Brown et al., 2004] Brown, Kenneth R., Harrow, Aram W., and Chuang, Isaac L. [Arbitrarily accurate composite pulse sequences](#). *Physical Review A*, 70(5):052318 (2004). (Cited on page 227).

- [Browne and Rudolph, 2005] Browne, Daniel E. and Rudolph, Terry. [Resource-Efficient Linear Optical Quantum Computation](#). *Physical Review Letters*, 95(1):010501 (2005). (Cited on pages 257 and 287).
- [Burkhart et al., 2021] Burkhart, Luke D., **Teoh, James D.**, Zhang, Yaxing, Axline, Christopher J., Frunzio, Luigi, Devoret, M.H., Jiang, Liang, Girvin, S.M., and Schoelkopf, R.J. [Error-Detected State Transfer and Entanglement in a Superconducting Quantum Network](#). *PRX Quantum*, 2(3):030321 (2021). (Cited on pages 271, 278, 279, 305, 307, 310, 314, 316, 323, 331, 332, 334, 338, 351, and 352).
- [Burnett et al., 2014] Burnett, J., Faoro, L., Wisby, I., Gurtovoi, V. L., Chernykh, A. V., Mikhailov, G. M., Tulin, V. A., Shaikhaidarov, R., Antonov, V., Meeson, P. J., Tzalenchuk, A. Ya, and Lindström, T. [Evidence for interacting two-level systems from the 1/f noise of a superconducting resonator](#). *Nature Communications*, 5(1):4119 (2014). (Cited on page 17).
- [Bylander et al., 2011] Bylander, Jonas, Gustavsson, Simon, Yan, Fei, Yoshihara, Fumiki, Harrabi, Khalil, Fitch, George, Cory, David G., Nakamura, Yasunobu, Tsai, Jaw-Shen, and Oliver, William D. [Noise spectroscopy through dynamical decoupling with a superconducting flux qubit](#). *Nature Physics*, 7(7):565–570 (2011). (Cited on page 227).
- [Campagne-Ibarcq et al., 2020] Campagne-Ibarcq, P., Eickbusch, A., Touzard, S., Zalys-Geller, E., Frattini, N. E., Sivak, V. V., Reinhold, P., Puri, S., Shankar, S., Schoelkopf, R. J., Frunzio, L., Mirrahimi, M., and Devoret, M. H. [Quantum error correction of a qubit encoded in grid states of an oscillator](#). *Nature*, 584(7821):368–372 (2020). (Cited on pages 45 and 78).
- [Campagne-Ibarcq et al., 2018] Campagne-Ibarcq, P., Zalys-Geller, E., Narla, A., Shankar, S., Reinhold, P., Burkhart, L., Axline, C., Pfaff, W., Frunzio, L., Schoelkopf, R. J., and Devoret, M. H. [Deterministic Remote Entanglement of Superconducting Circuits through Microwave Two-Photon Transitions](#). *Physical Review Letters*, 120(20):200501 (2018). (Cited on page 276).
- [Chakram et al., 2021] Chakram, Srivatsan, Oriani, Andrew E., Naik, Ravi K., Dixit, Akash V., He, Kevin, Agrawal, Ankur, Kwon, Hyeokshin, and Schuster, David I. [Seamless High-Q Microwave Cavities for Multimode Circuit Quantum Electrodynamics](#). *Physical Review Letters*, 127(10):107701 (2021). (Cited on page 110).
- [Chang et al., 2020] Chang, H.-S., Zhong, Y. P., Bienfait, A., Chou, M.-H., Conner, C. R., Dumur, É., Grebel, J., Peairs, G. A., Povey, R. G., Satzinger, K. J., and Cleland, A. N. [Remote Entanglement via Adiabatic Passage Using a Tunably Dissipative Quantum Communication System](#). *Physical Review Letters*, 124(24):240502 (2020). (Cited on page 281).
- [Chapman et al., 2022] Chapman, Benjamin J., de Graaf, Stijn J., Xue, Sophia H., Zhang, Yaxing, **Teoh, James**, Curtis, Jacob C., Tsunoda, Takahiro, Eickbusch,

- Alec, Read, Alexander P., Koottandavida, Akshay, Mundhada, Shantanu O., Frunzio, Luigi, Devoret, M. H., Girvin, S. M., and Schoelkopf, R. J. (2022). A high on-off ratio beamsplitter interaction for gates on bosonically encoded qubits. arXiv:2212.11929 [quant-ph]. (Cited on pages [143](#), [145](#), [199](#), [210](#), [226](#), [227](#), [320](#), and [352](#)).
- [Chou, 2018] Chou, Kevin. *Teleported operations between logical qubits in circuit quantum electrodynamics*. PhD thesis, Yale University (2018). (Cited on page [332](#)).
- [Chou et al., 2018] Chou, Kevin S., Blumoff, Jacob Z., Wang, Christopher S., Reinhold, Philip C., Axline, Christopher J., Gao, Yvonne Y., Frunzio, L., Devoret, M. H., Jiang, Liang, and Schoelkopf, R. J. [Deterministic teleportation of a quantum gate between two logical qubits](#). *Nature*, 561(7723):368–373 (2018). (Cited on page [45](#)).
- [Chou et al., 2023] Chou, Kevin S., Shemma, Tali, McCarrick, Heather, Chien, Tzu-Chiao, **Teoh, James D.**, Winkel, Patrick, Anderson, Amos, Chen, Jonathan, Curtis, Jacob, de Graaf, Stijn J., Garmon, John W. O., Gudlewski, Benjamin, Kalfus, William D., Keen, Trevor, Khedkar, Nishaad, Lei, Chan U., Liu, Gangqiang, Lu, Pinlei, Lu, Yao, Maiti, Aniket, Mastalli-Kelly, Luke, Mehta, Nitish, Mundhada, Shantanu O., Narla, Anirudh, Noh, Taewan, Tsunoda, Takahiro, Xue, Sophia H., Yuan, Joseph O., Frunzio, Luigi, Aumentado, Jose, Puri, Shruti, Girvin, Steven M., Moseley, Jr, and Schoelkopf, Robert J. (2023). Demonstrating a superconducting dual-rail cavity qubit with erasure-detected logical measurements. arXiv:2307.03169 [quant-ph]. (Cited on pages [210](#), [216](#), and [220](#)).
- [Chow et al., 2010] Chow, J. M., DiCarlo, L., Gambetta, J. M., Motzoi, F., Frunzio, L., Girvin, S. M., and Schoelkopf, R. J. [Optimized driving of superconducting artificial atoms for improved single-qubit gates](#). *Physical Review A*, 82(4):040305 (2010). (Cited on page [227](#)).
- [Chuang et al., 1997] Chuang, Isaac L., Leung, Debbie W., and Yamamoto, Yoshihisa. [Bosonic quantum codes for amplitude damping](#). *Physical Review A*, 56(2):1114–1125 (1997). (Cited on page [84](#)).
- [Chuang and Yamamoto, 1995] Chuang, Isaac L. and Yamamoto, Yoshihisa. [Simple quantum computer](#). *Physical Review A*, 52(5):3489–3496 (1995). (Cited on page [84](#)).
- [Cirac et al., 1997] Cirac, J. I., Zoller, P., Kimble, H. J., and Mabuchi, H. [Quantum State Transfer and Entanglement Distribution among Distant Nodes in a Quantum Network](#). *Physical Review Letters*, 78(16):3221–3224 (1997). (Cited on pages [43](#) and [274](#)).
- [Cochrane et al., 1999] Cochrane, P. T., Milburn, G. J., and Munro, W. J. [Macroscopically distinct quantum-superposition states as a bosonic code for amplitude damping](#). *Physical Review A*, 59(4):2631–2634 (1999). (Cited on page [81](#)).

- [Darmawan et al., 2021] Darmawan, Andrew S., Brown, Benjamin J., Grimsmo, Arne L., Tuckett, David K., and Puri, Shruti. [Practical Quantum Error Correction with the XZZX Code and Kerr-Cat Qubits](#). *PRX Quantum*, 2(3):030345 (2021). (Cited on page 72).
- [Devoret et al., 2007] Devoret, M.h., Girvin, Steven, and Schoelkopf, Robert. [Circuit-QED: How strong can the coupling between a Josephson junction atom and a transmission line resonator be?*](#). *Annalen der Physik*, 519(10-11):767–779 (2007). (Cited on page 101).
- [Ding et al., 2023] Ding, Leon, Hays, Max, Sung, Youngkyu, Kannan, Bharath, An, Junyoung, Di Paolo, Agustin, Karamlou, Amir H., Hazard, Thomas M., Azar, Kate, Kim, David K., Niedzielski, Bethany M., Melville, Alexander, Schwartz, Mollie E., Yoder, Jonilyn L., Orlando, Terry P., Gustavsson, Simon, Grover, Jeffrey A., Serniak, Kyle, and Oliver, William D. (2023). High-Fidelity, Frequency-Flexible Two-Qubit Fluxonium Gates with a Transmon Coupler. arXiv:2304.06087 [quant-ph]. (Cited on page 253).
- [Eastin and Knill, 2009] Eastin, Bryan and Knill, Emanuel. [Restrictions on Transversal Encoded Quantum Gate Sets](#). *Physical Review Letters*, 102(11):110502 (2009). (Cited on page 65).
- [Eickbusch et al., 2022] Eickbusch, Alec, Sivak, Volodymyr, Ding, Andy Z., Elder, Salvatore S., Jha, Shantanu R., Venkatraman, Jayameenakshi, Royer, Baptiste, Girvin, S. M., Schoelkopf, Robert J., and Devoret, Michel H. [Fast universal control of an oscillator with weak dispersive coupling to a qubit](#). *Nature Physics*, 18(12):1464–1469 (2022). (Cited on pages 80 and 334).
- [Einstein et al., 1935] Einstein, A., Podolsky, B., and Rosen, N. [Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?](#) *Physical Review*, 47(10):777–780 (1935). (Cited on page 40).
- [Elder et al., 2020] Elder, Salvatore S., Wang, Christopher S., Reinhold, Philip, Hann, Connor T., Chou, Kevin S., Lester, Brian J., Rosenblum, Serge, Frunzio, Luigi, Jiang, Liang, and Schoelkopf, Robert J. [High-Fidelity Measurement of Qubits Encoded in Multilevel Superconducting Circuits](#). *Physical Review X*, 10(1):011001 (2020). (Cited on pages 128, 218, 223, 237, and 260).
- [Feynman, 1982] Feynman, Richard P. [Simulating physics with computers](#). *International Journal of Theoretical Physics*, 21(6):467–488 (1982). (Cited on page 1).
- [Flühmann et al., 2019] Flühmann, C., Nguyen, T. L., Marinelli, M., Negnevitsky, V., Mehta, K., and Home, J. P. [Encoding a qubit in a trapped-ion mechanical oscillator](#). *Nature*, 566(7745):513–517 (2019). (Cited on page 45).
- [Fowler, 2013] Fowler, Austin G. [Coping with qubit leakage in topological codes](#). *Physical Review A*, 88(4):042308 (2013). (Cited on page 259).

- [Fowler et al., 2012] Fowler, Austin G., Mariantoni, Matteo, Martinis, John M., and Cleland, Andrew N. [Surface codes: Towards practical large-scale quantum computation](#). *Physical Review A*, 86(3):032324 (2012). (Cited on pages 67 and 72).
- [Gambetta et al., 2011] Gambetta, J. M., Motzoi, F., Merkel, S. T., and Wilhelm, F. K. [Analytic control methods for high-fidelity unitary operations in a weakly nonlinear oscillator](#). *Physical Review A*, 83(1):012308 (2011). (Cited on page 227).
- [Gao et al., 2019] Gao, Yvonne Y., Lester, Brian J., Chou, Kevin S., Frunzio, Luigi, Devoret, Michel H., Jiang, Liang, Girvin, S. M., and Schoelkopf, Robert J. [Entanglement of bosonic modes through an engineered exchange interaction](#). *Nature*, 566(7745):509–512 (2019). (Cited on pages 27, 46, and 153).
- [Gao et al., 2018] Gao, Yvonne Y., Lester, Brian J., Zhang, Yaxing, Wang, Chen, Rosenblum, Serge, Frunzio, Luigi, Jiang, Liang, Girvin, S. M., and Schoelkopf, Robert J. [Programmable Interference between Two Microwave Quantum Memories](#). *Physical Review X*, 8(2):021073 (2018). (Cited on pages 145 and 305).
- [Gao et al., 2021] Gao, Yvonne Y., Rol, M. Adriaan, Touzard, Steven, and Wang, Chen. [Practical Guide for Building Superconducting Quantum Devices](#). *PRX Quantum*, 2(4):040202 (2021). (Cited on page 102).
- [Geller and Zhou, 2013] Geller, Michael R. and Zhou, Zhongyuan. [Efficient error models for fault-tolerant architectures and the Pauli twirling approximation](#). *Physical Review A*, 88(1):012314 (2013). (Cited on page 13).
- [Gertler et al., 2021] Gertler, Jeffrey M., Baker, Brian, Li, Juliang, Shirol, Shruti, Koch, Jens, and Wang, Chen. [Protecting a bosonic qubit with autonomous quantum error correction](#). *Nature*, 590(7845):243–248 (2021). (Cited on page 79).
- [Gidney and Ekerå, 2021] Gidney, Craig and Ekerå, Martin. [How to factor 2048 bit RSA integers in 8 hours using 20 million noisy qubits](#). *Quantum*, 5:433 (2021). (Cited on pages 2, 67, and 241).
- [Giovannetti et al., 2008] Giovannetti, Vittorio, Lloyd, Seth, and Maccone, Lorenzo. [Quantum Random Access Memory](#). *Physical Review Letters*, 100(16):160501 (2008). (Cited on page 29).
- [Girvin, 2014] Girvin, S. M. (2014). Circuit QED: superconducting qubits coupled to microwave photons. In Devoret, Michel, Huard, Benjamin, Schoelkopf, Robert, and Cugliandolo, Leticia F., editors, *Quantum Machines: Measurement and Control of Engineered Quantum Systems*, pages 113–256. Oxford University Press Oxford, 1 edition. (Cited on page 117).
- [Girvin, 2023] Girvin, Steven M. [Introduction to quantum error correction and fault tolerance](#). *SciPost Physics Lecture Notes*, page 070 (2023). (Cited on page 49).

- [Gold et al., 2021] Gold, Alysson, Paquette, J. P., Stockklauser, Anna, Reagor, Matthew J., Alam, M. Sohaib, Bestwick, Andrew, Didier, Nicolas, Nersisyan, Ani, Oruc, Feyza, Razavi, Armin, Scharmann, Ben, Sete, Eyob A., Sur, Biswajit, Venturelli, Davide, Winkleblack, Cody James, Wudarski, Filip, Harburn, Mike, and Rigetti, Chad. [Entanglement across separate silicon dies in a modular superconducting qubit device](#). *npj Quantum Information*, 7(1):1–10 (2021). (Cited on page 267).
- [Google AI Quantum et al., 2020] Google AI Quantum, Foxen, B., Neill, C., Dunsworth, A., Roushan, P., Chiaro, B., Megrant, A., Kelly, J., Chen, Zijun, Satzinger, K., Barends, R., Arute, F., Arya, K., Babbush, R., Bacon, D., Bardin, J. C., Boixo, S., Buell, D., Burkett, B., Chen, Yu, Collins, R., Farhi, E., Fowler, A., Gidney, C., Giustina, M., Graff, R., Harrigan, M., Huang, T., Isakov, S. V., Jeffrey, E., Jiang, Z., Kafri, D., Kechedzhi, K., Klimov, P., Korotkov, A., Kostritsa, F., Landhuis, D., Lucero, E., McClean, J., McEwen, M., Mi, X., Mohseni, M., Mutus, J. Y., Naaman, O., Neeley, M., Niu, M., Petukhov, A., Quintana, C., Rubin, N., Sank, D., Smelyanskiy, V., Vainsencher, A., White, T. C., Yao, Z., Yeh, P., Zalcman, A., Neven, H., and Martinis, J. M. [Demonstrating a Continuous Set of Two-Qubit Gates for Near-Term Quantum Algorithms](#). *Physical Review Letters*, 125(12):120504 (2020). (Cited on pages 29 and 207).
- [Gottesman and Chuang, 1999] Gottesman, Daniel and Chuang, Isaac L. [Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations](#). *Nature*, 402(6760):390–393 (1999). (Cited on page 44).
- [Gottesman et al., 2001] Gottesman, Daniel, Kitaev, Alexei, and Preskill, John. [Encoding a qubit in an oscillator](#). *Physical Review A*, 64(1):012310 (2001). (Cited on page 78).
- [Grassl et al., 1997] Grassl, M., Beth, Th., and Pellizzari, T. [Codes for the quantum erasure channel](#). *Physical Review A*, 56(1):33–38 (1997). (Cited on pages 55 and 252).
- [Grice, 2011] Grice, W. P. [Arbitrarily complete Bell-state measurement using only linear optical elements](#). *Physical Review A*, 84(4):042331 (2011). (Cited on page 257).
- [Grimm et al., 2020] Grimm, A., Frattini, N. E., Puri, S., Mundhada, S. O., Touzard, S., Mirrahimi, M., Girvin, S. M., Shankar, S., and Devoret, M. H. [Stabilization and operation of a Kerr-cat qubit](#). *Nature*, 584(7820):205–209 (2020). (Cited on page 81).
- [Guillaud and Mirrahimi, 2019] Guillaud, Jérémie and Mirrahimi, Mazyar. [Repetition Cat Qubits for Fault-Tolerant Quantum Computation](#). *Physical Review X*, 9(4):041053 (2019). (Cited on page 13).

- [Hann et al., 2019] Hann, Connor T., Zou, Chang-Ling, Zhang, Yaxing, Chu, Yiwen, Schoelkopf, Robert J., Girvin, S. M., and Jiang, Liang. [Hardware-Efficient Quantum Random Access Memory with Hybrid Quantum Acoustic Systems](#). *Physical Review Letters*, 123(25):250501 (2019). (Cited on page 29).
- [Hastrup and Andersen, 2022] Hastrup, Jacob and Andersen, Ulrik Lund. [All-optical cat-code quantum error correction](#). *Physical Review Research*, 4(4):043065 (2022). (Cited on page 364).
- [Heeres et al., 2017] Heeres, Reinier W., Reinhold, Philip, Ofek, Nissim, Frunzio, Luigi, Jiang, Liang, Devoret, Michel H., and Schoelkopf, Robert J. [Implementing a universal gate set on a logical qubit encoded in an oscillator](#). *Nature Communications*, 8(1):94 (2017). (Cited on pages 75, 136, 137, and 218).
- [Hensen et al., 2015] Hensen, B., Bernien, H., Dréau, A. E., Reiserer, A., Kalb, N., Blok, M. S., Ruitenberg, J., Vermeulen, R. F. L., Schouten, R. N., Abellán, C., Amaya, W., Pruneri, V., Mitchell, M. W., Markham, M., Twitchen, D. J., Elkouss, D., Wehner, S., Taminiau, T. H., and Hanson, R. [Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres](#). *Nature*, 526(7575):682–686 (2015). (Cited on page 287).
- [Hermans et al., 2022] Hermans, S. L. N., Pompili, M., Beukers, H. K. C., Baier, S., Borregaard, J., and Hanson, R. [Qubit teleportation between non-neighbouring nodes in a quantum network](#). *Nature*, 605(7911):663–668 (2022). (Cited on page 42).
- [Horsman et al., 2012] Horsman, Dominic, Fowler, Austin G., Devitt, Simon, and Meter, Rodney Van. [Surface code quantum computing by lattice surgery](#). *New Journal of Physics*, 14(12):123011 (2012). (Cited on page 65).
- [Jiang et al., 2007] Jiang, Liang, Taylor, Jacob M., Sørensen, Anders S., and Lukin, Mikhail D. [Distributed quantum computation based on small quantum registers](#). *Physical Review A*, 76(6):062323 (2007). (Cited on page 267).
- [Josephson, 1962] Josephson, B. D. [Possible new effects in superconductive tunnelling](#). *Physics Letters*, 1(7):251–253 (1962). (Cited on page 115).
- [Kang et al., 2023] Kang, Mingyu, Campbell, Wesley C., and Brown, Kenneth R. (2023). Quantum error correction with metastable states of trapped ions using erasure conversion. arXiv:2210.15024 [quant-ph]. (Cited on pages 58 and 211).
- [Kimble, 2008] Kimble, H. J. [The quantum internet](#). *Nature*, 453(7198):1023–1030 (2008). (Cited on page 266).
- [Knill, 2005] Knill, E. [Quantum computing with realistically noisy devices](#). *Nature*, 434(7029):39–44 (2005). (Cited on page 65).
- [Knill and Laflamme, 1997] Knill, Emanuel and Laflamme, Raymond. [Theory of quantum error-correcting codes](#). *Physical Review A*, 55(2):900–911 (1997). (Cited on page 48).

- [Knill et al., 2001] Knill, E., Laflamme, R., and Milburn, G. J. [A scheme for efficient quantum computation with linear optics](#). *Nature*, 409(6816):46–52 (2001). (Cited on page 84).
- [Koch et al., 2007] Koch, Jens, Yu, Terri M., Gambetta, Jay, Houck, A. A., Schuster, D. I., Majer, J., Blais, Alexandre, Devoret, M. H., Girvin, S. M., and Schoelkopf, R. J. [Charge-insensitive qubit design derived from the Cooper pair box](#). *Physical Review A*, 76(4):042319 (2007). (Cited on page 115).
- [Korotkov, 2011] Korotkov, Alexander N. [Flying microwave qubits with nearly perfect transfer efficiency](#). *Physical Review B*, 84(1):014510 (2011). (Cited on page 274).
- [Krantz et al., 2019] Krantz, P., Kjaergaard, M., Yan, F., Orlando, T. P., Gustavsson, S., and Oliver, W. D. [A quantum engineer’s guide to superconducting qubits](#). *Applied Physics Reviews*, 6(2):021318 (2019). (Cited on page 102).
- [Krastanov et al., 2015] Krastanov, Stefan, Albert, Victor V., Shen, Chao, Zou, Chang-Ling, Heeres, Reinier W., Vlastakis, Brian, Schoelkopf, Robert J., and Jiang, Liang. [Universal control of an oscillator with dispersive coupling to a qubit](#). *Physical Review A*, 92(4):040303 (2015). (Cited on page 75).
- [Kubica et al., 2022] Kubica, Aleksander, Haim, Arbel, Vaknin, Yotam, Brandão, Fernando, and Retzker, Alex (2022). Erasure qubits: Overcoming the \$T_1\$ limit in superconducting circuits. arXiv:2208.05461 [quant-ph]. (Cited on pages 7, 58, and 252).
- [Kurpiers et al., 2018] Kurpiers, P., Magnard, P., Walter, T., Royer, B., Pechal, M., Heinsoo, J., Salathé, Y., Akin, A., Storz, S., Besse, J.-C., Gasparinetti, S., Blais, A., and Wallraff, A. [Deterministic quantum state transfer and remote entanglement using microwave photons](#). *Nature*, 558(7709):264–267 (2018). (Cited on page 276).
- [Kurpiers et al., 2019] Kurpiers, P., Pechal, M., Royer, B., Magnard, P., Walter, T., Heinsoo, J., Salathé, Y., Akin, A., Storz, S., Besse, J.-C., Gasparinetti, S., Blais, A., and Wallraff, A. [Quantum Communication with Time-Bin Encoded Microwave Photons](#). *Physical Review Applied*, 12(4):044067 (2019). (Cited on page 277).
- [Kurpiers et al., 2017] Kurpiers, Philipp, Walter, Theodore, Magnard, Paul, Salathe, Yves, and Wallraff, Andreas. [Characterizing the attenuation of coaxial and rectangular microwave-frequency waveguides at cryogenic temperatures](#). *EPJ Quantum Technology*, 4(1):1–15 (2017). (Cited on page 307).
- [Laflamme et al., 1996] Laflamme, Raymond, Miquel, Cesar, Paz, Juan Pablo, and Zurek, Wojciech Hubert. [Perfect Quantum Error Correcting Code](#). *Physical Review Letters*, 77(1):198–201 (1996). (Cited on page 54).
- [Lauk et al., 2020] Lauk, Nikolai, Sinclair, Neil, Barzanjeh, Shabir, Covey, Jacob P., Saffman, Mark, Spiropulu, Maria, and Simon, Christoph. [Perspectives on quantum](#)

- [transduction](#). *Quantum Science and Technology*, 5(2):020501 (2020). (Cited on page [274](#)).
- [Leung et al., 2019] Leung, N., Lu, Y., Chakram, S., Naik, R. K., Earnest, N., Ma, R., Jacobs, K., Cleland, A. N., and Schuster, D. I. [Deterministic bidirectional communication and remote entanglement generation between superconducting qubits](#). *npj Quantum Information*, 5(1):1–5 (2019). (Cited on page [282](#)).
- [Levine et al., 2023] Levine, Harry, Haim, Arbel, Hung, Jimmy S. C., Alidoust, Nasser, Kalae, Mahmoud, DeLorenzo, Laura, Wollack, E. Alex, Arriola, Patricio Arrangoiz, Khalajhedayati, Amirhossein, Vaknin, Yotam, Kubica, Aleksander, Clerk, Aashish A., Hover, David, Brandão, Fernando, Retzker, Alex, and Painter, Oskar (2023). Demonstrating a long-coherence dual-rail erasure qubit using tunable transmons. arXiv:2307.08737 [quant-ph]. (Cited on page [252](#)).
- [Li et al., 2017] Li, Linshu, Zou, Chang-Ling, Albert, Victor V., Muralidharan, Sreraman, Girvin, S. M., and Jiang, Liang. [Cat Codes with Optimal Decoherence Suppression for a Lossy Bosonic Channel](#). *Physical Review Letters*, 119(3):030502 (2017). (Cited on page [365](#)).
- [Li et al., 2023] Li, Ziqian, Roy, Tanay, Perez, David Rodriguez, Lee, Kan-Heng, Kapit, Eliot, and Schuster, David I. (2023). Autonomous error correction of a single logical qubit using two transmons. arXiv:2302.06707 [quant-ph]. (Cited on page [79](#)).
- [Litinski, 2019a] Litinski, Daniel. [A Game of Surface Codes: Large-Scale Quantum Computing with Lattice Surgery](#). *Quantum*, 3:128 (2019a). (Cited on page [65](#)).
- [Litinski, 2019b] Litinski, Daniel. [Magic State Distillation: Not as Costly as You Think](#). *Quantum*, 3:205 (2019b). (Cited on page [29](#)).
- [Litinski and Oppen, 2018] Litinski, Daniel and Oppen, Felix von. [Lattice Surgery with a Twist: Simplifying Clifford Gates of Surface Codes](#). *Quantum*, 2:62 (2018). (Cited on page [34](#)).
- [Lu et al., 2023] Lu, Yao, Maiti, Aniket, Garmon, John W. O., Ganjam, Suhas, Zhang, Yaxing, Claes, Jahan, Frunzio, Luigi, Girvin, S. M., and Schoelkopf, Robert J. (2023). A high-fidelity microwave beamsplitter with a parity-protected converter. arXiv:2303.00959 [quant-ph]. (Cited on pages [140](#), [143](#), [145](#), [199](#), [210](#), [226](#), [227](#), [261](#), [320](#), and [352](#)).
- [Ma et al., 2023] Ma, Shuo, Liu, Genyue, Peng, Pai, Zhang, Bichen, Jandura, Sven, Claes, Jahan, Burgers, Alex P., Pupillo, Guido, Puri, Shruti, and Thompson, Jeff D. (2023). High-fidelity gates with mid-circuit erasure conversion in a metastable neutral atom qubit. arXiv:2305.05493 [physics, physics:quant-ph]. (Cited on page [211](#)).
- [Ma et al., 2020] Ma, Wen-Long, Zhang, Mengzhen, Wong, Yat, Noh, Kyungjoo, Rosenblum, Serge, Reinhold, Philip, Schoelkopf, Robert J., and Jiang, Liang.

- [Path-Independent Quantum Gates with Noisy Ancilla](#). *Physical Review Letters*, 125(11):110503 (2020). (Cited on page 6).
- [Magnard et al., 2020] Magnard, P., Storz, S., Kurpiers, P., Schär, J., Marxer, F., Lütolf, J., Walter, T., Besse, J.-C., Gabureac, M., Reuer, K., Akin, A., Royer, B., Blais, A., and Wallraff, A. [Microwave Quantum Link between Superconducting Circuits Housed in Spatially Separated Cryogenic Systems](#). *Physical Review Letters*, 125(26):260502 (2020). (Cited on page 274).
- [Majer et al., 2007] Majer, J., Chow, J. M., Gambetta, J. M., Koch, Jens, Johnson, B. R., Schreier, J. A., Frunzio, L., Schuster, D. I., Houck, A. A., Wallraff, A., Blais, A., Devoret, M. H., Girvin, S. M., and Schoelkopf, R. J. [Coupling superconducting qubits via a cavity bus](#). *Nature*, 449(7161):443–447 (2007). (Cited on page 277).
- [Martinis, 2009] Martinis, John M. [Superconducting phase qubits](#). *Quantum Information Processing*, 8(2):81–103 (2009). (Cited on page 31).
- [McEwen et al., 2022] McEwen, Matt, Faoro, Lara, Arya, Kunal, Dunsworth, Andrew, Huang, Trent, Kim, Seon, Burkett, Brian, Fowler, Austin, Arute, Frank, Bardin, Joseph C., Bengtsson, Andreas, Bilmes, Alexander, Buckley, Bob B., Bushnell, Nicholas, Chen, Zijun, Collins, Roberto, Demura, Sean, Derk, Alan R., Erickson, Catherine, Giustina, Marissa, Harrington, Sean D., Hong, Sabrina, Jeffrey, Evan, Kelly, Julian, Klimov, Paul V., Kostritsa, Fedor, Laptev, Pavel, Locharla, Aditya, Mi, Xiao, Miao, Kevin C., Montazeri, Shirin, Mutus, Josh, Naaman, Ofer, Neeley, Matthew, Neill, Charles, Opremcak, Alex, Quintana, Chris, Redd, Nicholas, Roushan, Pedram, Sank, Daniel, Satzinger, Kevin J., Shvarts, Vladimir, White, Theodore, Yao, Z. Jamie, Yeh, Ping, Yoo, Juhwan, Chen, Yu, Smelyanskiy, Vadim, Martinis, John M., Neven, Hartmut, Megrant, Anthony, Ioffe, Lev, and Barends, Rami. [Resolving catastrophic error bursts from cosmic rays in large arrays of superconducting qubits](#). *Nature Physics*, 18(1):107–111 (2022). (Cited on page 244).
- [Miao et al., 2022] Miao, Kevin C., McEwen, Matt, Atalaya, Juan, Kafri, Dvir, Pryadko, Leonid P., Bengtsson, Andreas, Opremcak, Alex, Satzinger, Kevin J., Chen, Zijun, Klimov, Paul V., Quintana, Chris, Acharya, Rajeev, Anderson, Kyle, Ansmann, Markus, Arute, Frank, Arya, Kunal, Asfaw, Abraham, Bardin, Joseph C., Bourassa, Alexandre, Bovaird, Jenna, Brill, Leon, Buckley, Bob B., Buell, David A., Burger, Tim, Burkett, Brian, Bushnell, Nicholas, Campero, Juan, Chiaro, Ben, Collins, Roberto, Conner, Paul, Crook, Alexander L., Curtin, Ben, Debroy, Dripto M., Demura, Sean, Dunsworth, Andrew, Erickson, Catherine, Fatemi, Reza, Ferreira, Vinicius S., Burgos, Leslie Flores, Forati, Ebrahim, Fowler, Austin G., Foxen, Brooks, Garcia, Gonzalo, Giang, William, Gidney, Craig, Giustina, Marissa, Gosula, Raja, Dau, Alejandro Grajales, Gross, Jonathan A., Hamilton, Michael C., Harrington, Sean D., Heu, Paula, Hilton, Jeremy, Hoffmann, Markus R., Hong, Sabrina, Huang, Trent, Huff, Ashley, Iveland, Justin, Jeffrey, Evan, Jiang, Zhang, Jones, Cody, Kelly, Julian, Kim, Seon, Kostritsa, Fedor, Kreikebaum, John Mark, Landhuis, David, Laptev, Pavel, Laws, Lily, Lee,

- Kenny, Lester, Brian J., Lill, Alexander T., Liu, Wayne, Locharla, Aditya, Lucero, Erik, Martin, Steven, Megrant, Anthony, Mi, Xiao, Montazeri, Shirin, Morvan, Alexis, Naaman, Ofer, Neeley, Matthew, Neill, Charles, Nersisyan, Ani, Newman, Michael, Ng, Jiun How, Nguyen, Anthony, Nguyen, Murray, Potter, Rebecca, Rocque, Charles, Roushan, Pedram, Sankaragomathi, Kannan, Schuster, Christopher, Shearn, Michael J., Shorter, Aaron, Shutty, Noah, Shvarts, Vladimir, Skruzny, Jindra, Smith, W. Clarke, Sterling, George, Szalay, Marco, Thor, Douglas, Torres, Alfredo, White, Theodore, Woo, Bryan W. K., Yao, Z. Jamie, Yeh, Ping, Yoo, Juhwan, Young, Grayson, Zalcman, Adam, Zhu, Ningfeng, Zobrist, Nicholas, Neven, Hartmut, Smelyanskiy, Vadim, Petukhov, Andre, Korotkov, Alexander N., Sank, Daniel, and Chen, Yu (2022). Overcoming leakage in scalable quantum error correction. arXiv:2211.04728 [quant-ph]. (Cited on page 58).
- [Michael et al., 2016] Michael, Marios H., Silveri, Matti, Brierley, R. T., Albert, Victor V., Salmilehto, Juha, Jiang, Liang, and Girvin, S. M. [New Class of Quantum Error-Correcting Codes for a Bosonic Mode](#). *Physical Review X*, 6(3):031006 (2016). (Cited on pages 77 and 83).
- [Mirrahimi et al., 2014] Mirrahimi, Mazhar, Leghtas, Zaki, Albert, Victor V., Touzard, Steven, Schoelkopf, Robert J., Jiang, Liang, and Devoret, Michel H. [Dynamically protected cat-qubits: a new paradigm for universal quantum computation](#). *New Journal of Physics*, 16(4):045014 (2014). (Cited on pages 81 and 364).
- [Motzoi et al., 2009] Motzoi, F., Gambetta, J. M., Rebentrost, P., and Wilhelm, F. K. [Simple Pulses for Elimination of Leakage in Weakly Nonlinear Qubits](#). *Physical Review Letters*, 103(11):110501 (2009). (Cited on pages 120 and 227).
- [Nagayama et al., 2017] Nagayama, Shota, Fowler, Austin G., Horsman, Dominic, Devitt, Simon J., and Meter, Rodney Van. [Surface code error correction on a defective lattice](#). *New Journal of Physics*, 19(2):023050 (2017). (Cited on page 260).
- [Ni et al., 2023] Ni, Zhongchu, Li, Sai, Deng, Xiaowei, Cai, Yanyan, Zhang, Libo, Wang, Weiting, Yang, Zhen-Biao, Yu, Haifeng, Yan, Fei, Liu, Song, Zou, Chang-Ling, Sun, Luyan, Zheng, Shi-Biao, Xu, Yuan, and Yu, Dapeng. [Beating the break-even point with a discrete-variable-encoded logical qubit](#). *Nature*, 616(7955):56–60 (2023). (Cited on page 51).
- [Nielsen and Chuang, 2010] Nielsen, Michael A. and Chuang, Isaac L. *Quantum computation and quantum information*. Cambridge University Press, Cambridge ; New York, 10th anniversary ed edition (2010). (Cited on pages 6, 12, 30, and 74).
- [Niu et al., 2023] Niu, Jingjing, Zhang, Libo, Liu, Yang, Qiu, Jiawei, Huang, Wenhui, Huang, Jiayang, Jia, Hao, Liu, Jiawei, Tao, Ziyu, Wei, Weiwei, Zhou, Yuxuan, Zou, Wanjing, Chen, Yuanzhen, Deng, Xiaowei, Deng, Xiuhao, Hu, Changkang, Hu, Ling, Li, Jian, Tan, Dian, Xu, Yuan, Yan, Fei, Yan, Tongxing, Liu, Song, Zhong, Youpeng, Cleland, Andrew N., and Yu, Dapeng. [Low-loss interconnects for modular superconducting quantum processors](#). *Nature Electronics*, 6(3):235–241 (2023). (Cited on pages 112 and 307).

- [Ofek et al., 2016] Ofek, Nissim, Petrenko, Andrei, Heeres, Reinier, Reinhold, Philip, Leghtas, Zaki, Vlastakis, Brian, Liu, Yehan, Frunzio, Luigi, Girvin, S. M., Jiang, L., Mirrahimi, Mazhar, Devoret, M. H., and Schoelkopf, R. J. [Extending the lifetime of a quantum bit with error correction in superconducting circuits](#). *Nature*, 536(7617):441–445 (2016). (Cited on page 51).
- [Olivieri et al., 2017] Olivieri, E., Billard, J., De Jesus, M., Juillard, A., and Leder, A. [Vibrations on pulse tube based Dry Dilution Refrigerators for low noise measurements](#). *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 858:73–79 (2017). (Cited on page 113).
- [Paladino et al., 2014] Paladino, E., Galperin, Y. M., Falci, G., and Altshuler, B. L. [1/f noise: Implications for solid-state quantum information](#). *Reviews of Modern Physics*, 86(2):361–418 (2014). (Cited on page 17).
- [Peterer et al., 2015] Peterer, Michael J., Bader, Samuel J., Jin, Xiaoyue, Yan, Fei, Kamal, Archana, Gudmundsen, Theodore J., Leek, Peter J., Orlando, Terry P., Oliver, William D., and Gustavsson, Simon. [Coherence and Decay of Higher Energy Levels of a Superconducting Transmon Qubit](#). *Physical Review Letters*, 114(1):010501 (2015). (Cited on page 118).
- [Pirandola et al., 2015] Pirandola, S., Eisert, J., Weedbrook, C., Furusawa, A., and Braunstein, S. L. [Advances in quantum teleportation](#). *Nature Photonics*, 9(10):641–652 (2015). (Cited on pages 41 and 42).
- [Place et al., 2021] Place, Alexander P. M., Rodgers, Lila V. H., Mundada, Pranav, Smitham, Basil M., Fitzpatrick, Mattias, Leng, Zhaoqi, Premkumar, Anjali, Bryon, Jacob, Vrajitoarea, Andrei, Sussman, Sara, Cheng, Guangming, Madhavan, Trisha, Babla, Harshvardhan K., Le, Xuan Hoang, Gang, Youqi, Jäck, Berthold, Gyenis, András, Yao, Nan, Cava, Robert J., de Leon, Nathalie P., and Houck, Andrew A. [New material platform for superconducting transmon qubits with coherence times exceeding 0.3 milliseconds](#). *Nature Communications*, 12(1):1779 (2021). (Cited on pages 112 and 119).
- [Pokharel et al., 2018] Pokharel, Bibek, Anand, Namit, Fortman, Benjamin, and Lidar, Daniel A. [Demonstration of Fidelity Improvement Using Dynamical Decoupling with Superconducting Qubits](#). *Physical Review Letters*, 121(22):220502 (2018). (Cited on page 17).
- [Premaratne et al., 2017] Premaratne, Shavindra P., Wellstood, F. C., and Palmer, B. S. [Microwave photon Fock state generation by stimulated Raman adiabatic passage](#). *Nature Communications*, 8(1):14148 (2017). (Cited on page 218).
- [Puri et al., 2017] Puri, Shruti, Boutin, Samuel, and Blais, Alexandre. [Engineering the quantum states of light in a Kerr-nonlinear resonator by two-photon driving](#). *npj Quantum Information*, 3(1):1–7 (2017). (Cited on page 81).

- [Puri et al., 2019] Puri, Shruti, Grimm, Alexander, Campagne-Ibarcq, Philippe, Eickbusch, Alec, Noh, Kyungjoo, Roberts, Gabrielle, Jiang, Liang, Mirrahimi, Mazyar, Devoret, Michel H., and Girvin, S. M. [Stabilized Cat in a Driven Non-linear Cavity: A Fault-Tolerant Error Syndrome Detector](#). *Physical Review X*, 9(4):041009 (2019). (Cited on pages 77 and 81).
- [Puri et al., 2020] Puri, Shruti, St-Jean, Lucas, Gross, Jonathan A., Grimm, Alexander, Frattini, Nicholas E., Iyer, Pavithran S., Krishna, Anirudh, Touzard, Steven, Jiang, Liang, Blais, Alexandre, Flammia, Steven T., and Girvin, S. M. [Bias-preserving gates with stabilized cat qubits](#). *Science Advances*, 6(34):eaay5901 (2020). (Cited on pages 13 and 255).
- [Raussendorf and Briegel, 2001] Raussendorf, Robert and Briegel, Hans J. [A One-Way Quantum Computer](#). *Physical Review Letters*, 86(22):5188–5191 (2001). (Cited on page 23).
- [Reagor et al., 2013] Reagor, Matthew, Paik, Hanhee, Catelani, Gianluigi, Sun, Luyan, Axline, Christopher, Holland, Eric, Pop, Ioan M., Masluk, Nicholas A., Brecht, Teresa, Frunzio, Luigi, Devoret, Michel H., Glazman, Leonid, and Schoelkopf, Robert J. [Reaching 10 ms single photon lifetimes for superconducting aluminum cavities](#). *Applied Physics Letters*, 102(19):192604 (2013). (Cited on page 110).
- [Reagor et al., 2016] Reagor, Matthew, Pfaff, Wolfgang, Axline, Christopher, Heeres, Reinier W., Ofek, Nissim, Sliwa, Katrina, Holland, Eric, Wang, Chen, Blumoff, Jacob, Chou, Kevin, Hatridge, Michael J., Frunzio, Luigi, Devoret, Michel H., Jiang, Liang, and Schoelkopf, Robert J. [Quantum memory with millisecond coherence in circuit QED](#). *Physical Review B*, 94(1):014506 (2016). (Cited on page 111).
- [Reed et al., 2010] Reed, M. D., DiCarlo, L., Johnson, B. R., Sun, L., Schuster, D. I., Frunzio, L., and Schoelkopf, R. J. [High-Fidelity Readout in Circuit Quantum Electrodynamics Using the Jaynes-Cummings Nonlinearity](#). *Physical Review Letters*, 105(17):173601 (2010). (Cited on page 308).
- [Reinhold, 2019] Reinhold, Philip. *Controlling Error-Correctable Bosonic Qubits*. PhD thesis, Yale University (2019). (Cited on page 137).
- [Reinhold et al., 2020] Reinhold, Philip, Rosenblum, Serge, Ma, Wen-Long, Frunzio, Luigi, Jiang, Liang, and Schoelkopf, Robert J. [Error-corrected gates on an encoded qubit](#). *Nature Physics*, 16(8):822–826 (2020). (Cited on pages 7, 77, 94, 198, and 259).
- [Rosenblum et al., 2018a] Rosenblum, S., Gao, Y. Y., Reinhold, P., Wang, C., Axline, C. J., Frunzio, L., Girvin, S. M., Jiang, Liang, Mirrahimi, M., Devoret, M. H., and Schoelkopf, R. J. [A CNOT gate between multiphoton qubits encoded in two cavities](#). *Nature Communications*, 9(1):652 (2018a). (Cited on page 218).

- [Rosenblum et al., 2018b] Rosenblum, S., Reinhold, P., Mirrahimi, M., Jiang, Liang, Frunzio, L., and Schoelkopf, R. J. [Fault-tolerant detection of a quantum error](#). *Science*, 361(6399):266–270 (2018b). (Cited on pages 6, 7, 77, 80, 124, 161, and 199).
- [Sahay et al., 2023] Sahay, Kaavya, Jin, Junlan, Claes, Jahan, Thompson, Jeff D., and Puri, Shruti (2023). High threshold codes for neutral atom qubits with biased erasure errors. arXiv:2302.03063 [quant-ph]. (Cited on pages 72, 86, and 257).
- [Schoelkopf et al., 2003] Schoelkopf, R. J., Clerk, A. A., Girvin, S. M., Lehnert, K. W., and Devoret, M. H. (2003). Qubits as Spectrometers of Quantum Noise. In Nazarov, Yuli V., editor, *Quantum Noise in Mesoscopic Physics*, NATO Science Series, pages 175–203. Springer Netherlands, Dordrecht. (Cited on page 16).
- [Scholl et al., 2023] Scholl, Pascal, Shaw, Adam L., Tsai, Richard Bing-Shiun, Finkelstein, Ran, Choi, Joonhee, and Endres, Manuel (2023). Erasure conversion in a high-fidelity Rydberg quantum simulator. arXiv:2305.03406 [cond-mat, physics:physics, physics:quant-ph]. (Cited on pages 7 and 58).
- [Schuster, 2007] Schuster, David Isaac. *Circuit Quantum Electrodynamics*. PhD thesis, Yale University (2007). (Cited on page 16).
- [Schwinger, 2001] Schwinger, Julian (2001). Angular Momentum. In Schwinger, Julian and Englert, Berthold-Georg, editors, *Quantum Mechanics: Symbolism of Atomic Measurements*, pages 149–181. Springer, Berlin, Heidelberg. (Cited on page 175).
- [Shillito et al., 2022] Shillito, Ross, Petrescu, Alexandru, Cohen, Joachim, Beall, Jackson, Hauru, Markus, Ganahl, Martin, Lewis, Adam G.M., Vidal, Guifre, and Blais, Alexandre. [Dynamics of Transmon Ionization](#). *Physical Review Applied*, 18(3):034031 (2022). (Cited on page 127).
- [Shor, 1994] Shor, P.W. Algorithms for quantum computation: discrete logarithms and factoring. In *Proceedings 35th Annual Symposium on Foundations of Computer Science*, pages 124–134 (1994). (Cited on page 1).
- [Shor, 1995] Shor, Peter W. [Scheme for reducing decoherence in quantum computer memory](#). *Physical Review A*, 52(4):R2493–R2496 (1995). (Cited on pages 3 and 53).
- [Singh et al., 2022] Singh, Kevin, Anand, Shraddha, Pocklington, Andrew, Kemp, Jordan T., and Bernien, Hannes. [Dual-Element, Two-Dimensional Atom Array with Continuous-Mode Operation](#). *Physical Review X*, 12(1):011040 (2022). (Cited on page 3).
- [Singh et al., 2023] Singh, K., Bradley, C. E., Anand, S., Ramesh, V., White, R., and Bernien, H. [Mid-circuit correction of correlated phase errors using an array of spectator qubits](#). *Science*, 380(6651):1265–1269 (2023). (Cited on page 244).

- [Sivak et al., 2023] Sivak, V. V., Eickbusch, A., Royer, B., Singh, S., Tsioutsios, I., Ganjam, S., Miano, A., Brock, B. L., Ding, A. Z., Frunzio, L., Girvin, S. M., Schoelkopf, R. J., and Devoret, M. H. [Real-time quantum error correction beyond break-even](#). *Nature*, 616(7955):50–55 (2023). (Cited on pages 51, 78, 80, 260, and 328).
- [Stace et al., 2009] Stace, Thomas M., Barrett, Sean D., and Doherty, Andrew C. [Thresholds for Topological Codes in the Presence of Loss](#). *Physical Review Letters*, 102(20):200501 (2009). (Cited on page 69).
- [Stephenson et al., 2020] Stephenson, L. J., Nadlinger, D. P., Nichol, B. C., An, S., Drmota, P., Ballance, T. G., Thirumalai, K., Goodwin, J. F., Lucas, D. M., and Ballance, C. J. [High-Rate, High-Fidelity Entanglement of Qubits Across an Elementary Quantum Network](#). *Physical Review Letters*, 124(11):110501 (2020). (Cited on page 287).
- [Sun et al., 2014] Sun, L., Petrenko, A., Leghtas, Z., Vlastakis, B., Kirchmair, G., Sliwa, K. M., Narla, A., Hatridge, M., Shankar, S., Blumoff, J., Frunzio, L., Mirrahimi, M., Devoret, M. H., and Schoelkopf, R. J. [Tracking photon jumps with repeated quantum non-demolition parity measurements](#). *Nature*, 511(7510):444–448 (2014). (Cited on pages 158 and 223).
- [Sung et al., 2021] Sung, Youngkyu, Ding, Leon, Braumüller, Jochen, Vepsäläinen, Antti, Kannan, Bharath, Kjaergaard, Morten, Greene, Ami, Samach, Gabriel O., McNally, Chris, and Kim, David. Realization of high-fidelity cz and z z-free iswap gates with a tunable coupler. *Physical Review X*, 11(2):021058 (2021). (Cited on pages 210 and 227).
- [Teoh et al., 2022] **Teoh, James D.**, Winkel, Patrick, Babla, Harshvardhan K., Chapman, Benjamin J., Claes, Jahan, de Graaf, Stijn J., Garmon, John W. O., Kalfus, William D., Lu, Yao, Maiti, Aniket, Sahay, Kaavya, Thakur, Neel, Tsunoda, Takahiro, Xue, Sophia H., Frunzio, Luigi, Girvin, Steven M., Puri, Shruti, and Schoelkopf, Robert J. (2022). Dual-rail encoding with superconducting cavities. arXiv:2212.12077 [quant-ph]. (Cited on pages 7, 58, and 85).
- [Terhal, 2015] Terhal, Barbara M. [Quantum error correction for quantum memories](#). *Reviews of Modern Physics*, 87(2):307–346 (2015). (Cited on page 66).
- [Tsunoda et al., 2022] Tsunoda, Takahiro, **Teoh, James D.**, Kalfus, William D., de Graaf, Stijn J., Chapman, Benjamin J., Curtis, Jacob C., Thakur, Neel, Girvin, Steven M., and Schoelkopf, Robert J. (2022). Error-detectable bosonic entangling gates with a noisy ancilla. arXiv:2212.11196 [quant-ph]. (Cited on pages 6, 7, and 148).
- [Vogell et al., 2017] Vogell, B., Vermersch, B., Northup, T. E., Lanyon, B. P., and Muschik, C. A. [Deterministic quantum state transfer between remote qubits in cavities](#). *Quantum Science and Technology*, 2(4):045003 (2017). (Cited on page 282).

- [Vool and Devoret, 2017] Vool, Uri and Devoret, Michel. [Introduction to quantum electromagnetic circuits](#). *International Journal of Circuit Theory and Applications*, 45(7):897–934 (2017). (Cited on page 108).
- [Wallraff et al., 2004] Wallraff, A., Schuster, D. I., Blais, A., Frunzio, L., Huang, R.-S., Majer, J., Kumar, S., Girvin, S. M., and Schoelkopf, R. J. [Strong coupling of a single photon to a superconducting qubit using circuit quantum electrodynamics](#). *Nature*, 431(7005):162–167 (2004). (Cited on page 199).
- [Wallraff et al., 2007] Wallraff, A., Schuster, D. I., Blais, A., Gambetta, J. M., Schreier, J., Frunzio, L., Devoret, M. H., Girvin, S. M., and Schoelkopf, R. J. [Sideband Transitions and Two-Tone Spectroscopy of a Superconducting Qubit Strongly Coupled to an On-Chip Cavity](#). *Physical Review Letters*, 99(5):050501 (2007). (Cited on page 218).
- [Walter et al., 2017] Walter, T., Kurpiers, P., Gasparinetti, S., Magnard, P., Potočnik, A., Salathé, Y., Pechal, M., Mondal, M., Oppliger, M., Eichler, C., and Wallraff, A. [Rapid High-Fidelity Single-Shot Dispersive Readout of Superconducting Qubits](#). *Physical Review Applied*, 7(5):054020 (2017). (Cited on page 224).
- [Wang et al., 2016] Wang, Chen, Gao, Yvonne Y., Reinhold, Philip, Heeres, R. W., Ofek, Nissim, Chou, Kevin, Axline, Christopher, Reagor, Matthew, Blumoff, Jacob, Sliwa, K. M., Frunzio, L., Girvin, S. M., Jiang, Liang, Mirrahimi, M., Devoret, M. H., and Schoelkopf, R. J. [A Schrödinger cat living in two boxes](#). *Science*, 352(6289):1087–1091 (2016). (Cited on pages 294 and 329).
- [Wang et al., 2011] Wang, David S., Fowler, Austin G., and Hollenberg, Lloyd C. L. [Surface code quantum computing with error rates over 1%](#). *Physical Review A*, 83(2):020302 (2011). (Cited on page 72).
- [Wang and Clerk, 2012] Wang, Ying-Dan and Clerk, Aashish A. [Using dark modes for high-fidelity optomechanical quantum state transfer](#). *New Journal of Physics*, 14(10):105010 (2012). (Cited on page 278).
- [Wood and Gambetta, 2018] Wood, Christopher J. and Gambetta, Jay M. [Quantification and characterization of leakage errors](#). *Physical Review A*, 97(3):032306 (2018). (Cited on page 58).
- [Wootters, 1998] Wootters, William K. [Entanglement of Formation of an Arbitrary State of Two Qubits](#). *Physical Review Letters*, 80(10):2245–2248 (1998). (Cited on page 26).
- [Wootters and Zurek, 1982] Wootters, W. K. and Zurek, W. H. [A single quantum cannot be cloned](#). *Nature*, 299(5886):802–803 (1982). (Cited on page 3).
- [Wu et al., 2022] Wu, Yue, Kolkowitz, Shimon, Puri, Shruti, and Thompson, Jeff D. [Erasure conversion for fault-tolerant quantum computing in alkaline earth Rydberg atom arrays](#). *Nature Communications*, 13(1):4657 (2022). (Cited on pages 7, 58, 60, 69, 70, 72, 199, 211, 212, 228, and 262).

- [Xia et al., 2017] Xia, Xiu-Xiu, Sun, Qi-Chao, Zhang, Qiang, and Pan, Jian-Wei. [Long distance quantum teleportation](#). *Quantum Science and Technology*, 3(1):014012 (2017). (Cited on page 42).
- [Xiao et al., 2023] Xiao, Xu, Venkatraman, Jayameenakshi, Cortiñas, Rodrigo G., Chowdhury, Shoumik, and Devoret, Michel H. (2023). A diagrammatic method to compute the effective Hamiltonian of driven nonlinear oscillators. arXiv:2304.13656 [quant-ph]. (Cited on pages 124, 127, and 301).
- [Yan et al., 2013] Yan, Fei, Gustavsson, Simon, Bylander, Jonas, Jin, Xiaoyue, Yoshihara, Fumiki, Cory, David G., Nakamura, Yasunobu, Orlando, Terry P., and Oliver, William D. [Rotating-frame relaxation as a noise spectrum analyser of a superconducting qubit undergoing driven evolution](#). *Nature Communications*, 4(1):2337 (2013). (Cited on page 227).
- [Zalys-Geller, 2022] Zalys-Geller, Evan. *Time-bin Entanglement Between Remote Superconducting Cavity Resonators*. PhD thesis, Yale University (2022). (Cited on page 277).
- [Zapletal et al., 2022] Zapletal, Petr, Nunnenkamp, Andreas, and Brunelli, Matteo. [Stabilization of Multimode Schrödinger Cat States Via Normal-Mode Dissipation Engineering](#). *PRX Quantum*, 3(1):010301 (2022). (Cited on page 352).
- [Zhang et al., 2019] Zhang, Yaxing, Lester, Brian J., Gao, Yvonne Y., Jiang, Liang, Schoelkopf, R. J., and Girvin, S. M. [Engineering bilinear mode coupling in circuit QED: Theory and experiment](#). *Physical Review A*, 99(1):012314 (2019). (Cited on pages 144 and 305).
- [Zhong et al., 2019] Zhong, Y. P., Chang, H.-S., Satzinger, K. J., Chou, M.-H., Bienfait, A., Conner, C. R., Dumur, É, Grebel, J., Pears, G. A., Povey, R. G., Schuster, D. I., and Cleland, A. N. [Violating Bell's inequality with remotely connected superconducting qubits](#). *Nature Physics*, 15(8):741–744 (2019). (Cited on pages 272, 285, and 351).
- [Zhou et al., 2021] Zhou, Zheng-Yang, Gneiting, Clemens, You, J. Q., and Nori, Franco. [Generating and detecting entangled cat states in dissipatively coupled degenerate optical parametric oscillators](#). *Physical Review A*, 104(1):013715 (2021). (Cited on pages 351 and 352).