

UC Berkeley

UC Berkeley Electronic Theses and Dissertations

Title

Measurement-based quantum feedback controls for quantum computation

Permalink

<https://escholarship.org/uc/item/6x53p38k>

Author

Zhang, Song

Publication Date

2020

Peer reviewed|Thesis/dissertation

Measurement-based quantum feedback controls for quantum computation

by

Song Zhang

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Physics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Birgitta Whaley, Co-chair

Professor Irfan Siddiqi, Co-chair

Professor Eric Neuscamman

Professor Holger Mueller

Fall 2020

Measurement-based quantum feedback controls for quantum computation

Copyright 2020
by
Song Zhang

Abstract

Measurement-based quantum feedback controls for quantum computation

by

Song Zhang

Doctor of Philosophy in Physics

University of California, Berkeley

Professor Birgitta Whaley, Co-chair

Professor Irfan Siddiqi, Co-chair

The advance of experimental techniques in circuit quantum electrodynamics over last decade has made it possible to monitor the evolution of qubits under continuous measurements in real time . This kind of measurements, as compared to the discrete ones, provides a stream of information for feedback to alter the system's quantum dynamics continuously in a desired manner, thus allowing broader possibilities. We obtain the information at a cost, the back-action on the states. However, the back-action turns out not to be just something disturbing and harmful to the control but instead can be harnessed, together with the control, to steer the system to our targets. This thesis discusses one such application, namely to use measurements to create entanglement from remote qubits under locally optimal control. It is illustrated that by carefully designing symmetric measurement and feedback operators, we can construct efficient Markovian or non-Markovian protocols which generate N -qubit W, general Dicke and Greenberger–Horne–Zeilinger states with high fidelity above 94% up to $N = 100$. In some cases, the fidelity even approaches unity. Following that, we describe how to take advantages of the continuous stream of information to protect the unitary dynamics of a quantum system subject to noises. We develop a practical protocol for continuous operation of a quantum error correcting code for protection of coherent evolution due to an encoded Hamiltonian against environmental errors, using the three qubit bit-flip code and bit-flip errors as a canonical example. We demonstrate, for both quantum memory and adiabatic Hamiltonians, that the protocol, once optimized, can significantly reduce logical errors with a performance which is close to the theoretical limit of the bit-flip code's capability. In the adiabatic Hamiltonian case, we also propose some novel correction operators instead of the conventional ones to further reduce the logical errors.

To my parents, for their unconditional and unlimited love

Contents

Contents	ii
List of Figures	iv
List of Tables	x
1 Introduction	1
1.1 Background and overview	1
1.2 Density operators	2
2 A short introduction on circuit quantum electrodynamics	6
2.1 Quantum LC resonator	7
2.2 Artificial atom	8
2.3 The Jaynes-Cummings interaction and its dispersive limit	11
2.4 Dispersive readout of the qubit	14
2.5 Controlling a single qubit	17
3 Continuous measurement theory and measurement-based control	18
3.1 Projective measurement	18
3.2 Generalized measurement	19
3.3 Formalism of continuous measurement	24
3.4 Implementation of continuous measurement on a computer	28
3.5 Measurement-based quantum feedback control	34
4 Entanglement generation for multi-qubit system	36
4.1 Introduction	36
4.2 Entanglement in a nutshell	38
4.3 Continuous measurement with locally optimal feedback	41
4.4 Generation of Dicke states	48
4.5 Generation of Greenberger–Horne–Zeilinger states	55
4.6 Conditions for deterministic evolution under measurement-based feedback control	63
4.7 Discussion and Conclusions	64

4.8	Appendix	65
5	Continuous error correction I: quantum memory	73
5.1	Introduction	73
5.2	Discrete error correction in a nutshell	75
5.3	Continuous operation of the three-qubit bit flip code.	79
5.4	The double threshold CQEC protocol	81
5.5	Effective logical error rate	84
5.6	Optimization and benchmarking	89
6	Continuous error correction II: time-dependent Hamiltonian system	91
6.1	Adiabatic quantum computation in a nutshell	91
6.2	Encoding the time-dependent Hamiltonian	93
6.3	Effective open-system model for the logical qubit	95
6.4	Optimization and benchmarking	100
6.5	Discussion and conclusions	115
7	Conclusion and outlook	119
	Bibliography	121

List of Figures

2.1	A lumped-element circuit diagram representing a capacitive coupling between an LC oscillator and a Cooper pair box. The circuit symbols have their standard meaning. In particular, the box with a cross in it denotes the Josephson junction, denoted JJ. The notation C with possibly a subscript denotes the capacitance of that specific device, L is the inductance and E_J is the Josephson energy.	7
2.2	Figure illustrates the first three energy levels of a Cooper pair box (CPB) as a function of n_g . We have chosen the ground energy as a reference point for each value of n_g , i.e., what we have plotted is $E_i - E_0$. Each panel has a different E_J/E_C ratio, as indicated on top. Figure adapted from [4].	10
2.3	An illustration of the first few energy levels of the dispersive Hamiltonian (Eq. (2.19)). We divide the energy levels into two groups, one containing the qubit ground state $ g\rangle$ and the other containing the excited state $ e\rangle$. $ n = 0, 1, 2\rangle$ is the photon number state.	15
2.4	Illustration of the photon's initial state, denoted by the purple circle, and the two steady states associated with the qubit ground (red) and excited (blue) states. .	16
3.1	Illustration of a quantum trajectory for a single qubit continuously measured along the $\hat{z}$ axis. The qubit is initialized in the completely mixed state. The measurement operator is $X = \hat{\sigma}_z$, and the measurement strength $\kappa = 1$. The notation x, y, z refers to the three Bloch components and $r = \sqrt{x^2 + y^2 + z^2}$ is the Bloch vector length. The x and y values stay at zero while z gradually approaches -1 , suggesting that the qubit is projectively measured into the $ 1\rangle$ state in the long run. The fact that r gradually approaches 1 from 0 is an indication that the qubit is being purified by the measurement, i.e., being driven by the measurement from a completely mixed state to a pure state.	27
3.2	Averaged relative error as a function of time step size. The left panel is for $\eta = 1$ and the right panel is for $\eta = 0.64$. Green dots in both panels is for Bayesian method, Blue dots in both panels is for Kraus decomposition method, and red dot in the left panel is for POVM method. Parameters: $\kappa = \frac{1}{2}$, total time $T = 10\Gamma_m^{-1}$, and number of trajectories averaged over is 10000.	33
4.1	A diagrammatic picture illustrating what happens during a measurement-feedback cycle over time dt	42

4.2	(a) Schematic of a non-Markovian feedback protocol. The number of possible states, and hence the number of potentially distinct feedback operations needed in response scale exponentially in time. We refer to this as the trajectory ensemble approach (TEA). (b) Schematic of the average state locally optimal (ASLO) feedback protocol, which applies controls that are determined using only the average time-evolved state and the most recent measurement outcome [26]. This is the behavior of a “forgetful” or Markovian controller.	47
4.3	Fidelity of controlled state with respect to the W state, $\mathcal{F}_W$ starting from the initial product state $ 000\rangle$. The solid red line shows results calculated with the trajectory ensemble approach (TEA) and the dashed red line shows results calculated with the ASLO approach. The dashed blue line shows the results obtained with measurement alone, <i>i.e.</i> , in absence of feedback, after an initial rotation to the state of maximum fidelity with the W state that is achievable by $U_F^W(\theta)$. Over 1000 trajectories were averaged for TEA, and over 10,000 trajectories for the ASLO and zero feedback calculations. The measurement strength is set to $k = 1$ MHz, corresponding to realistic values for superconducting qubits [17, 26], and the time step is set at $\Delta t \ll \frac{1}{k}$	51
4.4	Optimal angle coefficients A_1 and A_2 , evaluated as a function of time for the ASLO protocol for generation of the $N = 3$ W state from the initial product state $ 000\rangle$	52
4.5	Fidelity with respect to generalized Dicke states $ N, k\rangle$ under the ASLO protocol with locally optimal feedback, equations (4.31) - (4.32), with U_F determined by the PaQS approach of section 4.3. Panel <i>a</i>) shows fidelity as a function of number of qubits N (from 3 to 48, in increments of 3) and excitation number k . The two right panels show cuts through this two-dimensional plot for fixed values of N and k . Panel <i>b</i>): k -dependence of fidelity for $N = 48$ qubits. Panel <i>c</i>): N -dependence for states with $k = 1$ excitation (W states), extending to $N = 100$ qubits.	54
4.6	Fidelity $\mathcal{F}_G(t)$ of the controlled state with respect to GHZ state as a function of time, obtained from the locally optimal ASLO approach. The red line is the result obtained with the symmetric two-body measurement operator X_G^S . The blue line shows the result obtained with the non-symmetric one-body measurement operator X_G . All simulations were run with measurement strength $k = 1$ MHz and time step $\Delta t \ll \frac{1}{k}$, starting from the complete superposition state (see text).	57
4.7	Fidelity of four-qubit GHZ state, computed with the TEA protocol using the complete superposition as initial state.	58
4.8	Fidelity and purity of the controlled state under the multiple measurement and multiple feedback operator protocol for generation of N qubit GHZ states, in the ASLO implementation. For $N = 2$ and $N = 3$ the purity remains exactly 1 for all time.	62

- 4.9 Bloch sphere illustrating the geometric meaning of the locally optimal control protocol for the GHZ state. The blue vector $\hat{n} = (\frac{\sqrt{3}}{2}, 0, \frac{1}{2})$ is the rotation axis. Initially, our state is located exactly along this direction. Note that the measurement will always keep the state in the xoz plane, where o denotes the origin. Suppose that after a measurement, the state is one of the two red dashed vectors. The feedback operator will rotate this vector along the yellow circle around the $\hat{n}$ axis. To maximize the fidelity with respect to the $|\tilde{0}\rangle$ state, we should always rotate it to the *upper* position labeled by $|\psi\rangle$. This means that the rotation angle θ should be 0 or $\frac{\pi}{2}$, depending on whether the state is above or below $\hat{n}$ after the measurement. The green circle at constant z is drawn for reference as a visual aid. 67
- 4.10 Time dependence of the three-tangle τ and the fidelity $\mathcal{F}_G$ for three qubits with the TEA protocol using feedback based on direct optimization of the three-tangle cost function τ (Eq. (4.88)). The initial condition here is the complete superposition state, Eq. (4.49). Red solid curve: three-tangle for state measured with two-body observable, averaged over 1000 trajectories. Red dashed curve: fidelity with respect to GHZ state, same state evolution as the red solid curve. Blue solid curve : three-tangle of state measured with one-body observable, averaged over 100 trajectories. For all trajectories the measurement strength was $k = 1$ and time step $dt \ll \frac{1}{k}$ 70
- 4.11 Histograms of control angle distributions resulting from optimization of the three-tangle at different times. The total trajectory count is 1000. The y-axis shows the normalized count in each bin. 71
- 5.1 Example of filtered readout signals $\mathcal{I}_1(t)$ and $\mathcal{I}_2(t)$ when a bit-flip error X_2 occurs. This error is detected by the CQEC protocol (see main text) at the moment when both filtered readout signals have exited the “syndrome uncertainty region” below the lower error threshold Θ_1 . The filtered readout signals $\mathcal{I}_k(t)$ are discontinuous since the CQEC protocol reset them to the value +1 at the moment when the occurred error is diagnosed. 83
- 5.2 Probability $p_{\text{misdiag}}^{(X_q)}$ that the CQEC protocol misdiagnoses the bit-flip error X_q as function of the measurement averaging time parameter τ . The error thresholds are fixed at $\Theta_1 = -0.54$ and $\Theta_2 = 0.8$. Symbols indicate numerical results and the solid line depicts the analytical formula (6.22) of the main text with $c = 1.607$. The dashed lines are guides to the eye. 84

- 5.3 Time windows $\Delta t_{qq'}$ for two consecutive bit-flip errors X_q (first) and $X_{q'}$ (second) to be diagnosed as a single (false) error $X_{q''}$, in the noiseless approximation. This is scenario three of the main text. The first error (X_q) occurs at the moment $t = 0$. The solid lines depict the “signal part” of the filtered readout signals $\mathcal{I}_1(t)$ [blue lines] and $\mathcal{I}_2(t)$ [red lines]. In panels (a) and (b), the window times Δt_{12} and Δt_{23} are both equal to the time that the blue solid line [i.e., $\langle \mathcal{I}_1(t) \rangle$, given in Eq. (5.23) of the main text] takes to cross the lower error threshold Θ_1 . In panel (c), the window time Δt_{13} is given by the time that the blue solid line spends inside the “syndrome uncertainty region”. Explicit formulas for the windows times $\Delta t_{qq'}$ are given in Eq. (5.30) of the main text. If the second error occurs within the time window $t = \Delta t_{qq'}$, the depicted two-error combinations are diagnosed in the noiseless approximation as $X_{q''} = X_3$ [panel (a)], X_1 [panel (b)] and X_2 [panel (c)]. 87
- 5.4 Logical error rate Γ_L as function of the measurement averaging time parameter τ . The curves depict formula Eq. (5.31) of the main text, evaluated for bit-flip error rates $\gamma_q = \gamma = 1.25 \times 10^{-3} \Gamma_m$ ($q = 1, 2, 3$), error threshold parameters $\Theta_1 = -0.54$ and $\Theta_2 = 0.8$, and quantum efficiencies $\eta = 0.5$ (dashed curve) and $\eta = 1$ (solid curve). Γ_m denotes measurement strength from continuous measurement. 88
- 5.5 Optimized logical X error rate Γ_L^{opt} for quantum memory operation. Solid red and blue lines depict Γ_L^{opt} , obtained using our double threshold CQEC protocol (Eq. (5.31)), and the dashed red and blue lines depict the logical X error rate, $\Gamma_L^{\text{Wonham}} = 3\gamma^2 \tau_m \ln(2/\gamma \tau_m)$, for the linear variant of the optimal Wonham filter [134]. Γ_m and $\tau_m = (2\Gamma_m \eta)^{-1}$ denote the measurement strength and “measurement time” from continuous measurement, respectively. Red and blue lines depict results for quantum efficiencies $\eta = 0.5$ and 1, respectively. The dotted line depicts the quadratic scaling $\Gamma_L \sim \gamma^2$ 90
- 6.1 Reduction factor $R(\gamma, \Omega_0)$ of the final logical state infidelity due to continuous QEC, Eq. (6.47), for operation time $t_{\text{op}} = 500\Gamma_m$, shown as a function of the bit-flip error rate γ for a range of strengths Ω_0 of the annealing Hamiltonian. Lines present analytic estimates, symbols depict numerical results averaged over an ensemble of 60,000-80,000 realizations of the conditional quantum master equation Eq. (5.18) together with our CQEC protocol. Black lines depict results for the quantum memory (QM) limit, $\Omega_0 \rightarrow 0$, see Eq. (6.16) of the main text. The other lines depict the results for quantum annealing (QA) for $\Omega_0 = \Gamma_m/90$ (green lines), $\Gamma_m/30$ (blue lines), $0.1\Gamma_m$ (purple lines) and $0.3\Gamma_m$ (red lines). The plateau values R_{plateau} are evident on the left hand side of the plot. Solid and dashed lines depict the results for quantum efficiencies $\eta = 1$ and 0.5, respectively. Γ_m denotes the strength of the continuous measurement. 103

- 6.2 Optimal measurement averaging time parameter τ_{opt} . Black solid and dashed lines depict the results for quantum memory (QM), where the frequency parameter Ω_0 vanishes, see Eq. (6.16) of the main text. The other lines depict the results for quantum annealing (QA) with $\Omega_0 = \Gamma_m/90$ (green lines), $\Gamma_m/30$ (blue lines), $0.1\Gamma_m$ (purple lines) and $0.3\Gamma_m$ (red lines). Solid and dashed lines depict the results for quantum efficiencies $\eta = 1$ and 0.5 , respectively. Γ_m denotes the strength of the continuous measurement. 105
- 6.3 Accuracy of the effective evolution equation (6.38) for the logical state infidelity during quantum annealing. The frequency parameter Ω_0 determines the strength of the logical Hamiltonian (6.16), where the coefficients $a(t)$ and $b(t)$ correspond to quantum annealing with linear schedule, see Eq. (6.43). Solid lines: full numerical calculations averaged over an ensemble of 20,000 realizations. Dotted lines: effective model of Eq. (6.38). Dashed line: unencoded qubit infidelity [see Eq. (6.48)] for $\Omega_0 = 0.1\Gamma_m$. Parameters: $\Theta_1 = -0.54$, $\Theta_2 = 0.8$, $\tau = 2.5\Gamma_m^{-1}$, $\eta = 1$, $t_{\text{op}} = 500\Gamma_m^{-1}$ and $\gamma = 1.25 \times 10^{-3}\Gamma_m$ 106
- 6.4 Logical state infidelity for two logical qubits. Parameters: $\gamma = 1.25 \times 10^{-3}\Gamma_m$ (error rate of all qubits), $\Omega_0 = 0.1\Gamma_m$ (Hamiltonian strength, see Eq. (6.56)), $\Theta_1 = -0.54$ and $\Theta_2 = 0.8$, $\tau = 2.1\Gamma_m^{-1}$, $\eta = 1$, and Γ_m denotes the strength of continuous measurement. Blue dotted line: effective model of Eq. (6.57). Red solid line: full numerical calculations averaged over an ensemble of 5,000 realizations (see text). 107
- 6.5 Reduction factor R of the final logical state infidelity as a function of the operation time t_{op} for single logical qubit annealing under CQEC. Blue, orange and purple lines depict the reduction factors corresponding to the cases where the error correction operations are given by $\tilde{C}_{\text{op}}(t)$ [Eq. (6.58)], $\tilde{C}_{\text{op}}^{\text{approx}}(t)$ [Eq. (6.59)] and $C_{\text{op}} = X_q$, respectively. Parameters: $\gamma = 1.25 \times 10^{-4}\Gamma_m$, $\Omega_0 = 0.1\Gamma_m$, $\tau = 2.5\Gamma_m^{-1}$, $\Theta_1 = -0.54$, $\Theta_2 = 0.8$, and $\eta = 1$ 109
- 6.6 Final infidelity with respect to the ground state of the problem Hamiltonian for one qubit annealing as a function of operation time t_{op} . Panel (a): $\gamma = 1.25 \times 10^{-4}\Gamma_m$. Panel (b): $\gamma = 1.25 \times 10^{-3}\Gamma_m$. Red line: one unencoded physical qubit, no error correction. Green line: classic protection scheme of Ref. [153] (see text). Blue line: one logical qubit with CQEC and modified error correction, Eq. (6.58). Parameter values: $\eta = 1$, $\Omega = 0.1\Gamma_m$, $\theta_1 = -0.54$, $\theta_2 = 0.8$, and $\tau = 2.5\Gamma_m^{-1}$ 112

- 6.7 Ratio of the TTS value for the classical (C) strategy to that for CQEC, as a function of t_{op} . Panel (a): one logical qubit, $\gamma = 1.25 \times 10^{-4}\Gamma_{\text{m}}$. Panel (b): one and two logical qubits, $\gamma = 1.25 \times 10^{-3}\Gamma_{\text{m}}$. Green lines are for one logical qubit, with error correction operations in the CQEC calculations given by Eq. (6.58) (solid lines) and Eq. (6.59) (dashed lines). Blue lines are for two logical qubits, with error correction operations in the CQEC calculations given by Eq. (6.58) (solid line) and Eq. (6.59) (dashed line). Parameters: one logical qubit calculations use the same values as in Fig. 6.6, two logical qubit calculations employ $\tau = 2.1\Gamma_{\text{m}}^{-1}$, with all other parameter values the same as for one logical qubit. 113

List of Tables

4.1	Summary of some measurement-based feedback protocols that may be derived from Eq. (4.26) or its generalization to multiple measurement and feedback operators in section 4.5. [†] denotes this work. Protocols in which perfect noise cancellation (cancellation of dW terms, see discussion in section 4.6) can occur are marked with an asterisk. Note that for N -qubit GHZ states, noise cancellation occurs only for the multiple measurement operator protocol with $N = 2$ and $N = 3$	46
-----	--	----

Acknowledgments

This acknowledgements' part of the thesis to me is a place for recollecting the memories during my eight and half years in Berkeley. I first would like to thank my academic advisor, Birgitta Whaley, for introducing me to the field I am working in, her patient guidance and encouragement when I am stuck and struggling, her gentle critique when I am in the wrong path, and all the inspirations and ideas I got during our conversations. I am deeply grateful to my collaborator Leigh Martin, a former joint doctoral student between the Whaley group and the Siddiqi's lab, for helping me get into the field of quantum control and circuit quantum electrodynamics, giving detailed and tireless explanations of all the questions I got, inspiring me ideas and showing me the genuine enthusiasm toward research. I also wish to thank my collaborator Juan Atalaya, a former postdoc in our group, for his advice inside and outside research and the meticulous attitude toward details which I would otherwise neglected. I would also like to thank my friends residing in the same basement as me, Dip, Carlos, Norm, Jefferey, Rocío, Sudi and Katie for their support, the cheers, joys and food we shared together and oh, the great magic shows by Carlos. There are also all the other Whaley group members I would like to thank for all the wonderful conversations we had and inspirations I got from their talks. I also want to thank my friends Chin-shen and Avinash, who were my housemates when I first came to Berkeley and also pursuing Ph.D. at UC Berkeley, for all the moments we spent together. I am also very grateful to Irfan Siddiqi, my co-advisor in the physics department, for being my co-advisor and being considerate of me when I was coping with my inner problems. The gratitude also goes to the graduate student advisors in the physics department, Anne, Donna and Joelle, for the excellent services they provided. I also want to thank myself for constantly fighting the battles within my mind. Eight years and a half is long for completing a physics Ph.D. I wouldn't call the fights a victory, but I didn't give in and am not defeated. And therefore, I claim the win. Finally and above all, I give my greatest thanks to my parents Yonghua and Hong, for everything.

Chapter 1

Introduction

1.1 Background and overview

In pursuit of the holy grail of quantum computation, the research on superconducting qubits has rapidly progressed over the past two decades - ever since the first observation of their coherent time-domain oscillations in 1999 [1]. Compared to other candidates for quantum computation platforms such as trapped ions, nitrogen-vacancy (NV) centered diamonds, etc., superconducting qubits have considerable advantages in scalability and controllability, due to the well-developed technologies in circuit fabrication. The disadvantage of superconducting qubits is their short coherence time due to various decoherence channels like spontaneous emission and photon loss into uncontrolled environmental degrees of freedom, quasiparticle tunneling etc. In fact, the coherence time was only in the nanosecond time scale in [1]. Since then, many efforts have been dedicated to improve the coherence time, such as increasing the resonator-qubits coupling, going to the dispersive regime, and utilizing the transmon qubits [2–5]. Now the coherence time has been significantly extended to a scale of microseconds. Superconducting qubits have indeed become the leading architecture in studying and building a quantum computer in the noisy intermediate-scale quantum (NISQ) era [6].

The advances in measurement techniques of superconducting qubits make it possible to continuously monitor the system state. In 2013, by using a near quantum-limited amplifier and measuring the quadratures of the resonator field, Murch et al. [7] observed the real-time evolution of the qubit state, i.e., the quantum trajectory of the qubit, for the first time. From a theoretical point of view, this real-time monitoring, or continuous measurement serves as the natural playground for quantum feedback control, since it allows more controllability compared to feedback on discrete measurements. So this motivates us to think about the question of how we can use the continuous measurement-based control to assist quantum information processing. Exploring along this direction, we present two applications of quantum feedback control in this thesis.

The reminder of the thesis is organized as follows. In section 1.2, we give a quick introduction to the background knowledge about the density operator that is essential to understand

the rest of the thesis (which is intended to be accessible to anyone who has taken an undergraduate quantum mechanics class). In chapter 2, we give a short introduction to circuit QED, with a focus on the experimental aspects of superconducting qubits. The purpose of this chapter is to provide a bridge to the experimental side for the readers who have a theoretical background. Chapter 3 introduces the framework of generalized measurements and, in particular, continuous measurements. The core of this chapter is the stochastic master equation describing the diffusive evolution of conditioned states under real-time monitoring and its interpretations. The last section of this chapter is dedicated to a short discussion on measurement-based feedback control.

Chapter 4 is our research work [8] on how to use joint continuous measurement on remote qubits to efficiently generate states possessing multi-qubit entanglement. Section 4.3 gives the general formalism of our locally optimal protocols, which we refer as proportional and quantum state-based (PaQS) feedback and explains how we can use these for a Markovian feedback, which is much more feasible than a non-Markovian type of feedback that relies on knowledge of the complete history of the state under the continuous measurement. Sections 4.4 and 4.5 demonstrate how we apply PaQS to efficiently generate multi-qubit Dicke and Greenberger–Horne–Zeilinger (GHZ) states with high fidelity. It turns out that the symmetries of these entangled states play an important role in choosing the proper form of both the measurement and the feedback operators. The main results are summarized in Figures 4.5, 4.6, 4.7 & 4.8.

Chapters 5 & 6 present our work [9] on protecting coherent evolution from noise by developing a protocol to continuously correct errors on quantum memory and time-dependent Hamiltonian systems. Using bit-flip errors and the three qubit repetition code as a canonical example, the formalism of the protocol is explained in Sections 5.3 and 5.4 where we derive an effective master equation in the codespace for the logical qubits dynamics with continuous error correction applied. With this equation we can facilitate the optimization and benchmarking of our protocol. A specific example is presented in Section 5.6 for quantum memory and in Section 6.4 for systems driven by adiabatic Hamiltonians. Finally we conclude the thesis with an outlook in chapter 7.

One convention used in the thesis is that we set $\hbar = 1$ occasionally. It should be understood from the context whenever we use this convention.

1.2 Density operators

Considering a quantum system, if we know the system completely, its state can be represented by the state vector $|\psi\rangle$. An alternative way to describe the state is to rewrite it in this form

$$\rho = |\psi\rangle\langle\psi|. \quad (1.1)$$

We call ρ the *density operator* or *density matrix* of the system. If we expand these two mathematical objects in some basis, it is not hard to show that we can construct one from another up to a global phase term of the wavefunction. So these two representations of the

system's state are equivalent. But what if we don't have perfect knowledge of the system's state? Suppose all we know about the system is that it can be in a state $|\psi_i\rangle$ with probability p_i for $i = 1, \dots, n$. The mathematical object to describe the state in this situation is still the density operator, but now constructed as

$$\rho = \sum_i^n p_i |\psi_i\rangle \langle \psi_i|, \quad (1.2)$$

where $\sum_i p_i = 1$. We refer to the former state where the system can be described by a state vector as the *pure state* and to the latter case where the system is described by an ensemble of pure states ($n > 1$) as the *mixed state*. There are three characteristic properties of the density operator:

1. ρ is Hermitian.
2. $\text{Tr}(\rho) = 1$
3. ρ is a semi positive-definite operator.

The probability distribution $\{p_i | i = 1, \dots, n\}$ reveals the **classical** uncertainties in the system, e.g., we don't know which of the pure states the system is in, only the set of states and the probability of being in each of them. The degree of uncertainty can be very different for different density operators. One extreme case is when we have an equal mixture of the basis states. In this case, it is referred as *completely mixed state*. For example, suppose we have a spin system with basis vectors $|0\rangle$ and $|1\rangle$, the completely mixed state is

$$\rho = \frac{1}{2}(|0\rangle \langle 0| + |1\rangle \langle 1|) = \frac{I}{2} \quad (1.3)$$

where I is the identity operator. It means we have no knowledge of which pure state the system can be in. One metric to quantify how mixed or how pure the states are is the *purity*

$$\gamma = \text{Tr}(\rho^2) \quad (1.4)$$

This quantity must be positive and upper bounded by 1. For a pure state, $\gamma = 1$. For any mixed state $\gamma < 1$ and in particular, for a completely mixed state, $\gamma = I/d$, where d is the dimension of Hilbert space of the system and I is the identity operator. It can be shown that the completely mixed state has the lowest purity. We want to point out the pure state decomposition (Eq. (1.2)) of a density operator may not be unique. For example, the completely mixed state can be constructed from the any basis set of the system since it is just proportional to an identity.

Qubits and Bloch vectors

In this subsection, we consider a two-level quantum system, i.e., the qubit, which is the primary object we study in our research. Any density operator of a qubit can be written in

the so-called *Bloch vector form*

$$\rho = \frac{1}{2}(I + x\hat{\sigma}_x + y\hat{\sigma}_y + z\hat{\sigma}_z) \quad (1.5)$$

where the three matrices $\hat{\sigma}_{x/y/z}$ are the Pauli matrices. The three coordinates x, y, z are real numbers and termed as the Bloch components. The 3-tuple (x, y, z) is an equivalent representation of the most general qubit state. We can treat it as a vector in $\mathbb{R}^3$ and hence the name Bloch vector. Each individual coordinate is called a *Bloch component*. They can be computed from the density operator by

$$x = \text{Tr}(\hat{\sigma}_x \rho), \quad y = \text{Tr}(\hat{\sigma}_y \rho), \quad z = \text{Tr}(\hat{\sigma}_z \rho), \quad (1.6)$$

where $\text{Tr}(\cdot)$ denotes the trace over any eigenbasis of the qubit. Let $\vec{r} = (x, y, z)$ denote the Bloch vector. Due to the positivity of the density operator, the Bloch vector length $r \equiv \|\vec{r}\| = \sqrt{x^2 + y^2 + z^2} \leq 1$. The equal sign holds when the state is pure. Hence the Bloch vector is bounded inside a sphere with center at the origin and radius being 1. The origin denotes a completely mixed state, while the surface denotes pure states. The Bloch vector length can be related to the state purity by

$$r = 2\text{Tr}(\rho^2) - 1. \quad (1.7)$$

So Bloch vector length is a strict increasing function of the purity. There are some special points on the Bloch surface worth mentioning

$$\begin{aligned} \vec{r} = (0, 0, \pm 1) & \text{ denotes eigenstates of } \hat{\sigma}_z \\ \vec{r} = (\pm 1, 0, 0) & \text{ denotes eigenstates of } \hat{\sigma}_x \\ \vec{r} = (0, \pm 1, 0) & \text{ denotes eigenstates of } \hat{\sigma}_y. \end{aligned} \quad (1.8)$$

An example of the Bloch sphere is illustrated in Figure 4.9.

Composite system and reduced density operator

Consider a composite system with two subsystems A and B. Let $\{|\psi_i\rangle | i = 1, \dots\}$ be a basis set of A and $\{|\phi_i\rangle | i = 1, \dots\}$ be a basis set of B. We can define the partial trace on the subsystem B as

$$\rho_A = \text{Tr}_B(\rho) = \sum_m \langle \phi_m | \rho | \phi_m \rangle. \quad (1.9)$$

This partial trace gives us an operator acting on subsystem A. We can check it satisfies the three characteristic properties we introduced above and hence it is a density operator on the subsystem A. The physical meaning of ρ_A is that by tracing out the degree of freedoms of the subsystem B, we are left with the best state-of-knowledge of the subsystem A. This is the information available to us if we can only observe A instead of having full access to the composite system. We can also interpret this from a measurement point of view. Suppose

we want to measure an observable $\hat{O}$ defined on A. This means we want to know the trace average¹ of the observable $\hat{O}$

$$\langle \hat{O} \rangle = \text{Tr}(\hat{O}\rho) = \text{Tr}_A \text{Tr}_B(\hat{O}\rho) = \text{Tr}_A(\hat{O} \text{Tr}_B(\rho)) = \text{Tr}_A(\hat{O}\rho_A) \quad (1.10)$$

The last equality on the right hand side of this equation tells us to compute the trace average of the operator $\hat{A}$ on the subsystem A (we call the operator a local operator in this case), i.e., we don't have to know the entire state. All we need is the reduced state of the subsystem, i.e., ρ_A . Hence indeed, the quantity ρ_A captures all the information we need about the subsystem A if we can only perform local measurements. Let us give some specific examples of reduced states.

Product state: $\rho = |\psi\rangle\langle\psi| \otimes |\phi\rangle\langle\phi|$ with $|\psi\rangle$ being any state of A and $|\phi\rangle$ any state of B. Taking the partial trace on B, we have $\rho_A = |\psi\rangle\langle\psi|$, a pure state. So if A and B are not entangled, we would have a well-defined state of A.

Maximally entangled state: We consider two qubits. Let $|0\rangle, |1\rangle$ be the spin up and down states, respectively. Suppose the composite system is in a maximally entangled state (see Chapter 4), e.g.,

$$\rho = \frac{1}{2}(|01\rangle + |10\rangle)(\langle 01| + \langle 10|). \quad (1.11)$$

Then the reduced state obtained by tracing over the second qubit is $\rho_A = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|)$, the completely mixed state. This tells us that when A is maximally entangled with B, we know nothing about the subsystem A if we can only perform local measurements on A itself.

Unitary time evolution

The coherent evolution of a density operator caused by a Hamiltonian follows from the Schrödinger equation. We first consider its form for a pure state. Using the Schrödinger equation, we have

$$\begin{aligned} d\rho &= (d|\psi\rangle)\langle\psi| - |\psi\rangle d\langle\psi| \\ &= -\frac{i}{\hbar} H|\psi\rangle\langle\psi| dt + \frac{i}{\hbar} |\psi\rangle\langle\psi| H dt \\ &= -\frac{i}{\hbar} [H, \rho] dt. \end{aligned} \quad (1.12)$$

Hence we have the evolution equation

$$\dot{\rho} = -\frac{i}{\hbar} [H, \rho]. \quad (1.13)$$

The unitary time evolution for mixed states has the same expression due to linearity.

¹Or rather the trace averages of all the moments of $\hat{A}$. But the arguments follow from Eq. (1.10).

Chapter 2

A short introduction on circuit quantum electrodynamics

Generally speaking, our research work mainly focuses on measurement-based control of qubits. While theoretical, it is not a castle in the sky and has indeed solid connection with real world experiments. We feel it is necessary to spend a chapter discussing the relevant experimental topics and this serves as the purpose of this chapter. We will describe, although briefly, the experimental realization, measurement and control of superconducting qubits, which is nowadays a leading architecture for quantum computation and in for general for quantum information processing [5, 6]. Hopefully, this short introduction can partially fill in the gap between the theory and experiment side of the story.

Superconducting qubits are artificial atoms which are implemented through non-linear superconducting circuits, of which the central component is usually a Josephson junction (JJ). To probe and manipulate these qubits, the experimentalists couple them to linear resonators which contain quantized light in the microwave domain (around 1 GHz). Figure 2.1 is a schematic diagram illustrating the whole circuit. The study of the interaction between the non-linear superconducting circuits and the resonator is known as circuit quantum electrodynamics (c-QED). This is a field with interplay of quantum optics, superconducting condensates, decoherence, circuit signal processing, etc., which not only demonstrates rich physical phenomena but also has important applications like quantum information processing. It is fascinating to realize that all of these processes can happen on tiny and controllable chips. In the next two sections, we introduce the basic building blocks of circuit QED, namely the linear resonators and the non-linear artificial atoms. Then we discuss a common type of interaction between the two building blocks, the Jaynes-Cummings model. Following that, we turn to the standard approach to measure the qubits in circuit QED, the dispersive qubit readout. We conclude the chapter with a discussion on how to control the qubits.

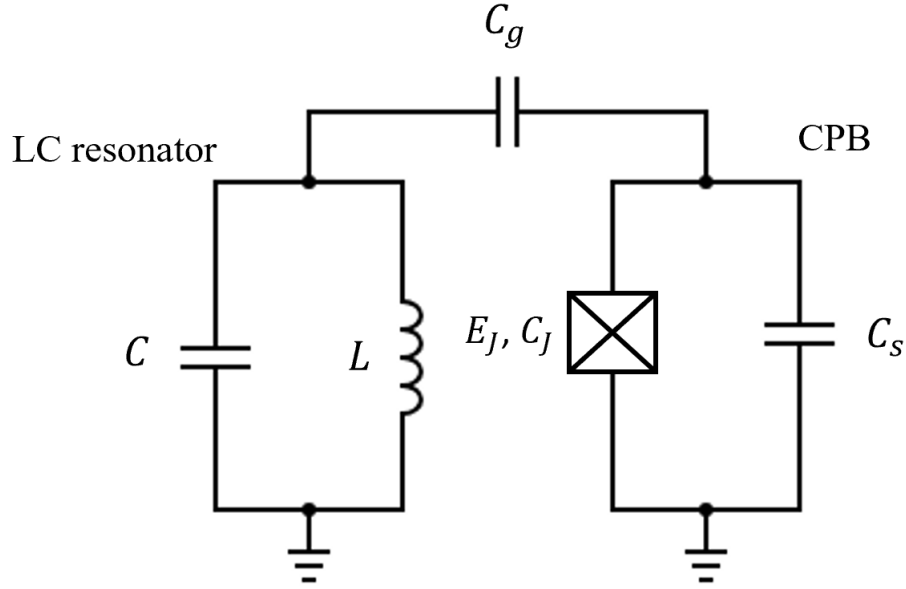


Figure 2.1: A lumped-element circuit diagram representing a capacitive coupling between an LC oscillator and a Cooper pair box. The circuit symbols have their standard meaning. In particular, the box with a cross in it denotes the Josephson junction, denoted JJ. The notation C with possibly a subscript denotes the capacitance of that specific device, L is the inductance and E_J is the Josephson energy.

2.1 Quantum LC resonator

A resonator is a device which stores photons and thus energy. A linear resonator can be considered as a harmonic oscillator and thus has constant separation between energy levels after quantization. It can come in many forms. The simplest one is perhaps a lumped-element¹ LC circuit as shown in Figure 2.1. The energy or Hamiltonian of the system can be written as

$$H_{\text{LC}} = \frac{Q^2}{2C} + \frac{\Phi^2}{2L}, \quad (2.1)$$

where Q is the charge on the capacitor plate, Φ is the magnetic flux through the inductor, C is the capacitance and L is the inductance. We can quantize this circuit by imposing the commutation relation [10]

$$[\Phi, Q] = i\hbar \quad (2.2)$$

¹Lumped-element circuits are circuits with small physical dimensions compared to signal wavelength such that currents or voltages do not vary through the circuits. So basically, they can be considered as point elements in the entire circuit.

This Hamiltonian is then identical to a quantum harmonic oscillator and produces a constant spacing of energy levels

$$\varepsilon_n = \hbar\omega_r(n + 1/2) \quad (2.3)$$

where $\omega_r = 1/\sqrt{LC}$ is the oscillator frequency. This resonator has thus only a single frequency mode and stores photons with frequency ω_r . We can also write the Hamiltonian in the ladder operator formalism of the harmonic oscillator. We introduce the resonator creation/annihilation operators a and $a^\dagger$ by the relations

$$\begin{aligned} \hat{\Phi} &= \sqrt{\frac{\hbar Z_r}{2}}(\hat{a}^\dagger + \hat{a}) \\ \hat{Q} &= i\sqrt{\frac{\hbar}{2Z_r}}(\hat{a}^\dagger - \hat{a}), \end{aligned} \quad (2.4)$$

where $Z_r = 1/\omega_r C = \sqrt{C/L}$ is the resonator impedance. The prefactors in front of the ladder operators carry the dimension and represent the zero-point fluctuation in the flux and charge. Note that the form of these relations indicates that the flux operator $\hat{\Phi}$ is the analog of position and the charge operator $\hat{Q}$ is the analog of momentum for the usual harmonic oscillator.

In practice, in order to operate the resonator in the quantum regime, at least two requirements must be met: low linewidth and low temperature. We will explain these one by one. The linewidth or photon loss rate of the resonator, denoted by κ , tells us how fast the photon is lost into the surrounding environment. The coherent lifetime of the photon stored in the resonator can be thought as $1/\kappa$. To observe quantum effects, we need the resonator to be weakly enough coupled to the environment such that its energy broadening must be less than the energy separation, i.e., the quality factor $Q = \omega_r/\kappa$ should be large [11].

The low temperature requirement means that we do not want too much thermal excitation of the photon mode, which will decohere the quantum nature as a pure quantum state can become mixed under thermal noise. A more precise condition is thus $\hbar\omega_r \gg k_B T$. For a photon in the microwave domain, $\omega_r \sim 1\text{GHz}$. This corresponds to a thermal temperature of $T \approx 50\text{mK}$.

Other typical types of linear resonator include two-dimensional planar forms (2D resonator) or a three-dimensional volume forms (3D resonator). These can be considered as a collection of quantum harmonic oscillators with different frequency modes. For example, a $\lambda/2$ coplanar waveguide consisting of a conductor, two ground planes and a substrate, is a 1D transmission resonator with open boundary conditions at two ends of the waveguide [3]. It has quantized frequency modes $f_m = (m + 1)f_0$, where f_0 is the fundamental frequency.

2.2 Artificial atom

We already saw that the resonators are linear components of the circuit. In order to have an artificial atom or qubit, we need non-linearity in the circuit. The reason is that a linear

device gives equal spacing energy levels, but in order to operate as a qubit, we only need two adjacent energy levels for which the transition frequency is well separated from the transition frequencies of any other two energy levels. Non-linearity achieves this for us.

The JJ is a non-linear superconducting component which can have a high quality factor and operate at low temperature, which makes it a perfect candidate as an artificial atom². It consists of two superconducting electrodes with some thin insulating material inbetween. A supercurrent can flow through the junction provided it is below the critical current. The supercurrent is related to the phase difference of the two superconductors in the following way

$$I = I_c \sin(\phi), \quad (2.5)$$

where I_c is the supercurrent and ϕ is the phase difference or the *Josephson phase*. The voltage V across the junction is related to the time derivative of ϕ

$$\dot{\phi} = 2\pi \frac{V}{\Phi_0}, \quad (2.6)$$

where $\Phi_0 = \hbar/2e$ is the magnetic flux quantum. The above two equations together are called the *Josephson equations* [12]. Introducing a new quantity which carries the unit of magnetic flux (note that there is no real magnetic field in the interior of a superconductor

$$\Phi = \Phi_0 \frac{\phi}{2\pi}, \quad (2.7)$$

allows us to rewrite Eq. (2.5) & (2.6) as $I = I_c \sin(2\pi\Phi/\Phi_0)$ and $\dot{\Phi} = V$. We can then make an analogue between the JJ and a linear inductor. The voltage equation is like Faraday's law and the current equation is a relation between the current and the flux. Except that in this case, unlike the linear inductor, the latter relation is non-linear. The energy stored in the JJ can be computed as

$$E = \int dt V(t)I(t) = I_c \int dt \dot{\Phi} \sin(2\pi \frac{\Phi}{\Phi_0}) = -E_J \cos(2\pi \frac{\Phi}{\Phi_0}), \quad (2.8)$$

where $E_J = \Phi_0 I_c / 2\pi$ is the Josephson energy. When the JJ is coupled to a capacitor, it is called the Cooper pair box (CPB), which together serves as the qubit. The Hamiltonian of the CPB is

$$\begin{aligned} H_q &= \frac{(\hat{Q} - Q_g)^2}{2C_q} - E_J \cos(2\pi \frac{\hat{\Phi}}{\Phi_0}) \\ &= 4E_C(\hat{n} - n_g)^2 - E_J \cos(\hat{\phi}), \end{aligned} \quad (2.9)$$

where C_q is the total capacitance of the circuit including the contribution from the JJ itself, $\hat{n} = \hat{Q}/2e$ is the charge number operator, $\hat{\phi} = (2\pi/\Phi_0)\Phi$ is the phase operator and

²To be more precise, we still need to connect it in parallel to a shunt capacitor to form the so called Cooper pair box or charge qubit (See Figure 2.1)

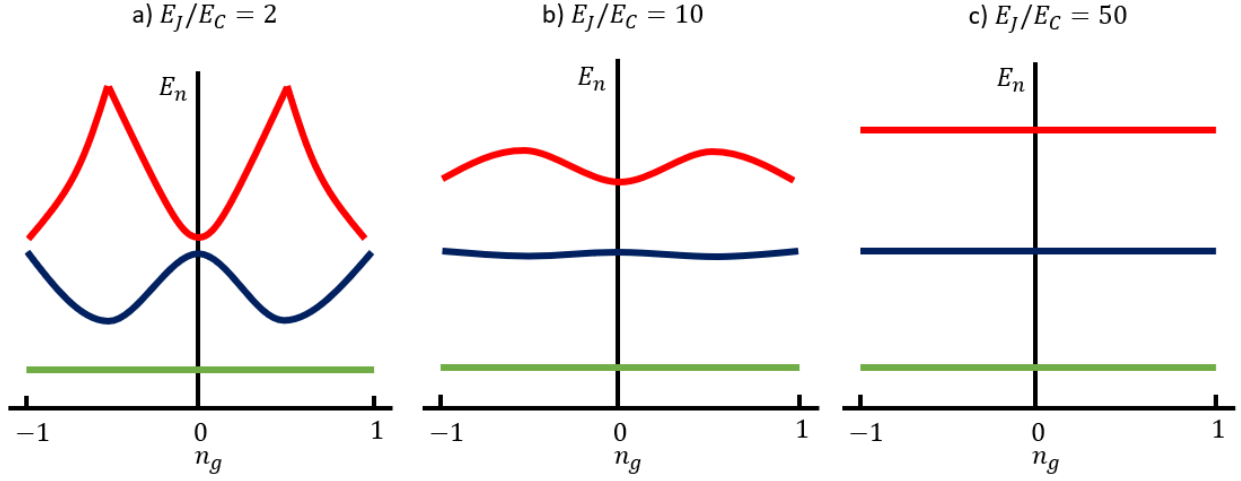


Figure 2.2: Figure illustrates the first three energy levels of a Cooper pair box (CPB) as a function of n_g . We have chosen the ground energy as a reference point for each value of n_g , i.e., what we have plotted is $E_i - E_0$. Each panel has a different E_J/E_C ratio, as indicated on top. Figure adapted from [4].

$E_C = e^2/2C$ is the charge energy. A possible offset charge Q_g or offset charge number $n_g = Q_g/2e$ is included to account for the fact that the CPB may be connected to an external voltage or coupled to uncontrolled environmental degrees of freedom. The latter results in fluctuation of the offset charge number and is a source of decoherence for the qubit. The non-linearity or anharmonicity of this Hamiltonian comes from the cosine term. This introduces non-equal spacing between the energy levels and enables the possibility of isolating two energy levels from others. Thus effectively, we can treat the CPB as a two-level system.

Transmon qubits

The spectrum of the CPB Hamiltonian (2.9) depends on the energy ratio E_J/E_C and the offset charge number n_g . By varying the ratio E_J/E_C , we have different types of qubit³. The first three energy levels are shown in Figure 2.2 as an illustration. We can see the dependence on n_g changes a lot as the ratio parameter E_J/E_C varies. When this ratio is small, like in panel a), we call it a charge qubit. The transition frequency is very sensitive to n_g for this type of qubit. We mentioned above that n_g can fluctuate due to environmental noise. This sensitive dependence will cause random disturbance in the transition frequency, thus leading to dephasing in the charge qubit. To fix this problem, Koch et al. [4] proposed a new

³The specific experimental realization of these can be quite different from the basic CPB circuit. For example, Koch et al. [4] implemented their transmon qubits by using two JJs connected in parallel to a shunt capacitor.

type of qubit called a transmon qubit which operates in the domain of large E_J/E_C ratio. As we see in Figure 2.2 panels b) and c), the dependence on n_g becomes less sensitive as the ratio increases. However, the anharmonicity in these energy levels (transition frequency difference) also decreases as the ratio increases, and in the limit of $E_J/E_C \rightarrow \infty$ we recover perfect harmonicity with transition frequency $\omega_q = \sqrt{8E_J E_C} - E_C$. Luckily, the drop in anharmonicity has a much weaker dependence on the ratio E_J/E_C [11]. So we can find a sweet spot of the ratio which gives us both improved coherence time compared to the charge qubit and also enough anharmonicity in the energy levels.

2.3 The Jaynes-Cummings interaction and its dispersive limit

The Jaynes-Cummings model (JCM) is an exactly solvable model in quantum optics describing the interaction of photons in a cavity and a two-level atom through electric dipole coupling. A classic review on the model itself is given by Shore and Knight [13]. The original form of the model is

$$H_{\text{JCM}} = \hbar\omega_c \hat{a}^\dagger \hat{a} + \hbar\omega_q \frac{\sigma_z}{2} + \hbar g(\hat{a} \hat{\sigma}_+ + \hat{a}^\dagger \hat{\sigma}_-). \quad (2.10)$$

Here $\hat{a}/\hat{a}^\dagger$ is the photon creation/annihilation operator, $\hat{\sigma}_\pm$ is the spin raising/lowering operator, ω_c is the single mode photon frequency or cavity frequency, ω_q is the atomic transition frequency, and g is the light-atom coupling strength. We have assumed the interaction is small and near resonant, i.e., $g \ll \omega_c, \omega_q$, $|\omega_c - \omega_q| \ll \omega_c + \omega_q$, in order to utilize the rotating-wave approximation⁴. When $g = 0$, the eigenstates of the system are called the *bare* states and denoted as $|s, n\rangle$, where $s = g$ or e means the atom is in the ground state or excited state, respectively, and n refers to the photon occupation number. These states form a basis and we can pair the bare states by their total excitation number. The system ground state $|g, 0\rangle$ has excitation number 0, while the two states $|g, n\rangle, |e, n-1\rangle$ have n excitations. After we turn on the interaction (i.e., make $g \neq 0$), we notice that the Hamiltonian is decoupled between subspaces with different excitation numbers. So in the bare state basis, H_{JCM} is a direct sum of 2 by 2 submatrices. We can diagonalize these submatrices easily. The resulting eigenstates become *dressed*

$$\begin{aligned} |\overline{g, 0}\rangle &= |g, 0\rangle \\ |\overline{g, n}\rangle &= \cos\left(\frac{\theta_n}{2}\right)|g, n\rangle - \sin\left(\frac{\theta_n}{2}\right)|e, n-1\rangle \\ |\overline{e, n-1}\rangle &= \sin\left(\frac{\theta_n}{2}\right)|g, n\rangle + \cos\left(\frac{\theta_n}{2}\right)|e, n-1\rangle, \end{aligned} \quad (2.11)$$

⁴In the rotating-wave approximation, the fast oscillation terms $\omega_c + \omega_q$ can be neglected since they are far off-resonant [14]. This means we drop the double excitation/de-excitation terms $\hat{a}^\dagger \hat{\sigma}_+$ and $\hat{a} \hat{\sigma}_-$ in the JCM Hamiltonian.

where $\theta_n = \arctan(2g\sqrt{n}/\Delta)$, with $\Delta = \omega_q - \omega_c$ being the qubit-cavity detuning and $n > 1$. The energy spectrum is given by

$$\begin{aligned} E_{\overline{g,0}} &= E_{g,0} = -\frac{\hbar\omega_q}{2} \\ E_{\overline{g,n}} &= n\hbar\omega_c - \frac{\hbar\omega_c}{2} - \frac{\hbar}{2}\sqrt{\Delta^2 + 4g^2n} \\ E_{\overline{e,n-1}} &= n\hbar\omega_c - \frac{\hbar\omega_c}{2} + \frac{\hbar}{2}\sqrt{\Delta^2 + 4g^2n} \end{aligned} \quad (2.12)$$

The energy splitting between the two levels with excitation number n is $\hbar\sqrt{\Delta^2 + 4g^2n}$ which increases with n , in contrast to the bare states whose energy separation is independent of n .

We have learned the eigenstate structure of the JCM Hamiltonian. Now let's come back to the superconducting circuit. The circuit structure is as shown in Figure 2.1. Assuming the coupling capacitance C_g is weak compared to that of the resonator and CPB, the circuit Hamiltonian is [3]

$$\hat{H} = 4E_C(\hat{n} + \hat{n}_r)^2 - E_J \cos(\hat{\phi}) + \sum_m \hbar\omega_m \hat{a}_m^\dagger \hat{a}_m \quad (2.13)$$

where the resonator is now multi-mode with frequencies ω_m and creation/annihilation operator $\hat{a}_m/\hat{a}_m^\dagger$ (Eq. (2.4)). The offset charge in Eq. (2.9) is replaced by the resonator charge number operator ($n_g \rightarrow -\hat{n}_r$) since this bias is now caused by the quantized resonator mode. The bias charge number operator is given by

$$\hat{n}_r = \sum_m \hat{n}_m = \sum_m \frac{C_g}{C_m} \frac{\hat{Q}_m}{2e}, \quad (2.14)$$

where $\hat{n}$, as introduced before, is the charge number of the CPB, C_m is the resonator capacitance associated with each resonator mode and $\hat{n}_m$ denotes the bias charge contribution from the m th mode of the resonator. As mentioned above, we assumed $C_g \ll C_m, C_q$. The quadratic term $(\hat{n} + \hat{n}_r)^2$ in Eq. (2.13) represents the interaction between the resonator and the CPB. If we assume the transmon frequency ω_q of interest is much closer to one resonator mode ω_r than to all the other modes, effectively the qubit is only coupled to a single mode. It can be shown, in the transmon qubit case ($E_J \gg E_C$), that under these conditions the Hamiltonian Eq. (2.13) reduces to the pre-rotating-wave form of the Jaynes Cummings Hamiltonian, i.e., to the parent form of Eq. (2.10) before applying the rotating wave approximation [11]:

$$\hat{H} = \hbar\omega_r \hat{a}^\dagger \hat{a} + \hbar\omega_q \hat{b}^\dagger \hat{b} - \frac{E_C}{2} \hat{b}^\dagger \hat{b}^\dagger \hat{b} \hat{b} - \hbar g (\hat{b}^\dagger - \hat{b})(\hat{a}^\dagger - \hat{a}). \quad (2.15)$$

In Eq. (2.15) $\hat{b}^\dagger/\hat{b}$ is the creation/annihilation operator for the qubit. The third term denotes virtual transitions to other levels in the artificial atom and will disappear when we restrict

the Hamiltonian to the two levels of interest. The coupling strength g between the resonator mode and the qubit is given by

$$g = \omega_r \frac{C_g}{C_q} \left(\frac{E_J}{2E_C} \right)^{1/4} \sqrt{\frac{Z_r}{Z_{\text{vac}}}} \sqrt{2\pi\alpha} \quad (2.16)$$

where ω_r is the relevant resonator mode frequency, $Z_{\text{vac}} = \sqrt{\mu_0/\epsilon_0}$ is the vacuum impedance with μ_0 and ϵ_0 being the vacuum permittivity and permeability, respectively, and α is the fine structure constant. Compared to the charge qubit, the coupling strength can now be very large due to the factor E_J/E_C . However, in common laboratory settings, g is still much smaller than microwave frequencies, i.e., $g \ll \omega_r, \omega_q$. This makes it possible to apply the rotating-wave approximation and after also restricting the Hamiltonian to the two levels, we then arrive finally at the JCM Hamiltonian [3, 15]

$$\hat{H}_{\text{JCM}} = \hbar\omega_r \hat{a}^\dagger \hat{a} + \frac{\hbar\omega_q}{2} \hat{\sigma}_z + \hbar g (\hat{a}^\dagger \hat{\sigma}_- + \hat{a} \hat{\sigma}_+), \quad (2.17)$$

where we have made the replacement $\hat{b}^\dagger \rightarrow \hat{\sigma}_+$ and $\hat{b} \rightarrow \hat{\sigma}_-$ and we now have the explicit form of both the qubit frequency $\omega_q = \sqrt{8E_J E_C} - E_C$ and the coupling g in terms of the resonator and qubit parameters, Eq. (2.16).

Dispersive limit

A widely used regime that the experimentalists work in to probe and detect the qubit is the so-called dispersive limit, $g \ll \Delta$ with $\Delta = |\omega_r - \omega_q|$ being the resonator qubit detuning. We will give two motivations below on why working in this regime is beneficial. First, in this regime, direct exchange of energy between the resonator and transmon qubit is greatly limited, yet at the same time the photons still carry information about the state of the qubit. This allows us to learn much about the qubit state using quantum non-demolition measurement (QND)⁵. We will come back to this later in the next section. For now, we can get some ideas by looking at the dressed states, Eq. (2.11). In the opposite regime, i.e., on resonance, $\Delta \rightarrow 0$, the dressed states become maximally entangled between the photon number state and the qubit state. If we only look at the qubit, its state (the reduced density matrix) becomes mixed. For example, if the qubit is initially in the excited state and the resonator is in the vacuum state, then it will not stay that way. Instead, the entire system state will evolve along the maximally entangled dressed states. This results in a mixed state of the qubit during most time of the evolution and thus causes additional relaxation. On the other hand, in the dispersive regime, the dressed state is only minimally coupled between the bare states. For example,

$$|\overline{e}, 0\rangle \sim |e, 0\rangle + \frac{g}{\Delta} |g, n\rangle, \quad (2.18)$$

⁵In reality, a measurement takes finite time to accomplish. QND is a type of quantum measurement whose only back-action is to project the system state into the observable eigenstate without disturbing its evolution during the measurement time. See more discussions in Section 3.2.

where we have ignored the normalization factor. This is very close to the product state $|e, 0\rangle$. So a state initialized in $|e, 0\rangle$ can more or less stay in that state.

Second, in the dispersive regime, the Purcell effect can be suppressed. When we put an atom in a cavity (or connect the qubit to a resonator in our case), the atom decay and the photon loss processes would actually enhance each other. The origin of this phenomenon lies in the fact that the qubit state is entangled with a photon state. Photon loss results in qubit decay and vice versa. It may not be surprising that compared to the on-resonance case, this effect can be less harmful at large detuning, since in the latter regime the qubit-photon entanglement is not as significant. It has been calculated [4, 16] that when on resonance the spontaneous emission rate is enhanced by $\frac{\kappa}{2}$, where κ is the bare photon loss rate when there is no atom. At large detuning, the spontaneous emission rate is enhanced by $(g^2/\Delta^2)\kappa \ll \kappa/2$.

An effective Hamiltonian can be derived in the dispersive limit as [3, 4]

$$H_{\text{disp}} = \hbar\omega_r\hat{a}^\dagger\hat{a} + \frac{\hbar\omega'_q}{2}\hat{\sigma}_z + \hbar\chi\hat{a}^\dagger\hat{a}\hat{\sigma}_z, \quad (2.19)$$

where χ is the dispersive coupling strength, and ω'_r and ω'_q are the dressed frequencies. These parameters are given by

$$\begin{aligned} \chi &= -\frac{g^2 E_C/\hbar}{\Delta(\Delta - E_C/\hbar)} \\ \omega'_r &= \omega_r - \frac{g^2 E_C/\hbar}{\Delta - E_C/\hbar} \\ \omega'_q &= \omega_q + \frac{g^2}{\Delta}. \end{aligned} \quad (2.20)$$

Clearly the eigenstates are the bare ones but there is a crucial difference from the no interaction case. Now the energy splitting between the photon levels depends on the qubit state, as illustrated in Figure 2.3. The dispersive Hamiltonian is actually obtained in a perturbative manner with an interaction term that depends on the photon number and on the excitation of the artificial atom. A more precise condition for Eq. (2.19) to be valid is that the average photon number has to be much smaller than a critical number [11]

$$\bar{n} \ll n_{\text{crit}} = \frac{1}{2j+1} \left(\frac{|\Delta - jE_C/\hbar|^2}{4g^2} - j \right), \quad (2.21)$$

where $j = 0, 1, \dots$ refers to the artificial atom's states that we are using as the qubit levels.

2.4 Dispersive readout of the qubit

In the last section, we mentioned that the dispersive regime is excellent for doing qubit measurement without unnecessary disturbance to its state. This can also be understood from the dispersive Hamiltonian. Eq. (2.19). The interaction term $\chi\hat{a}^\dagger\hat{a}\hat{\sigma}_z$ does not cause any transition between qubit states. Instead, it introduces a Stark shift to the qubit frequencies

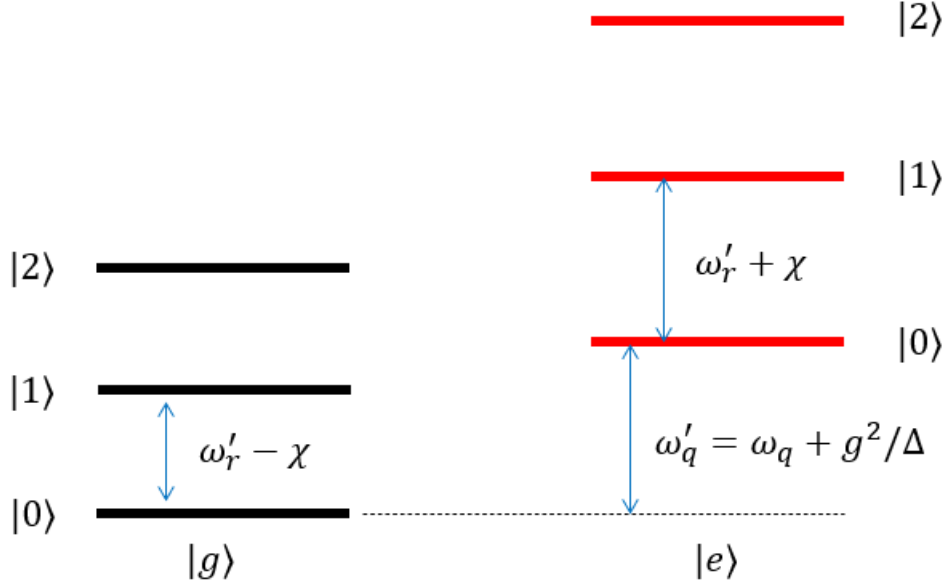


Figure 2.3: An illustration of the first few energy levels of the dispersive Hamiltonian (Eq. (2.19)). We divide the energy levels into two groups, one containing the qubit ground state $|g\rangle$ and the other containing the excited state $|e\rangle$. $|n = 0, 1, 2\rangle$ is the photon number state.

which depends on the dispersive coupling strength and the photon number. If we only focus on the qubit itself, this will result in some unavoidable measurement-induced dephasing as a cost of observing σ_z . From the resonator's point of view, the interaction term causes a photon frequency shift depending on the qubit state, as mentioned previously. So the photon state evolution is different for different qubit states. Suppose we prepare the photon in a coherent state $|\alpha(0)\rangle$ and the qubit is in a superposition state $c_0|g\rangle + c_1|e\rangle$. Then as the system evolves, it is not hard to show that due to the dispersive coupling, its state will become an entangled state of the photon coherent states and the qubit states

$$|\psi(t)\rangle = c_0|g\rangle|\alpha_g(t)\rangle + c_1|e\rangle|\alpha_e(t)\rangle, \quad (2.22)$$

where the two coherent states $|\alpha_{g/e}(t)\rangle$ rotate at a qubit-state dependent angular speed in the phase space. The subscript is to emphasize the dependence of the coherent states on the qubit state. The former behave like pointer states for the qubit and can be measured by a homodyne or heterodyne measurement [17]. In the lab, due to the presence of photon loss, the story is slightly more complicated. We need to drive the resonator a little bit to compensate the loss. In practice, the experimentalists send microwave probe photons to the resonator-qubit system through a transmission line for this purpose. Formally, this means

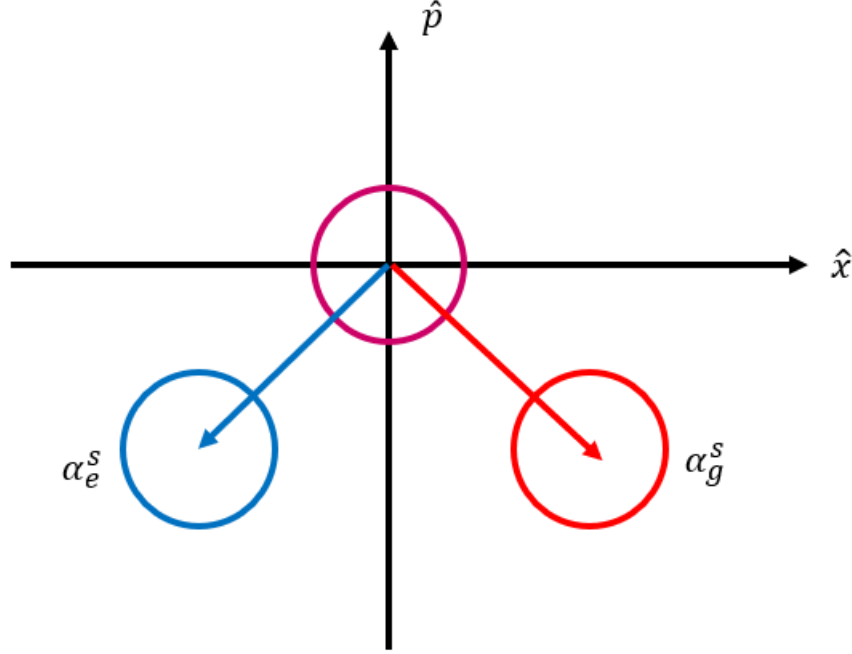


Figure 2.4: Illustration of the photon's initial state, denoted by the purple circle, and the two steady states associated with the qubit ground (red) and excited (blue) states.

we add a driving field to the dispersive Hamiltonian [3]

$$\hat{H}_{\text{disp}}^d = \frac{\hbar\omega'_q}{2} + \hbar\delta_r\hat{a}^\dagger\hat{a} + \hbar\chi\hat{a}^\dagger\hat{a}\hat{\sigma}_z + \hbar[\varepsilon_d(t)\hat{a}^\dagger e^{-i\omega_d t} + \varepsilon_d^*(t)\hat{a}e^{i\omega_d t}] \quad (2.23)$$

where ω_d is the driving frequency, $\varepsilon(t)$ is the time-dependent driving amplitude, $\delta_r = \omega_r - \omega_d$ is the detuning from the bare resonator frequency to the driving frequency. We have assumed the driving is far off-resonance from the qubit frequency. Also the Hamiltonian is written in the driving frame, i.e., we transformed to the rotating frame of the resonator at frequency ω_d . The evolution of the coherent state amplitudes can be obtained by solving the master equation [18] which treats the photon loss or gain from input-out theory applied at the boundary between the transmission line and the resonator [11]. We omit the calculation here and only quote the results

$$\begin{aligned} \dot{\alpha}_e(t) &= -i\varepsilon_d(t) - i(\delta_r + \chi)\alpha_e(t) - \kappa\alpha_e(t)/2 \\ \dot{\alpha}_g(t) &= -i\varepsilon_d(t) - i(\delta_r - \chi)\alpha_g(t) - \kappa\alpha_g(t)/2. \end{aligned} \quad (2.24)$$

The $\pm$ sign difference in the equations (second term on the RHS) is the signature of the dependence on the qubit state. We are interested in their steady state values which can be obtained by setting the derivative to zero,

$$\alpha_{e/g}^s = -\frac{\varepsilon}{(\delta_r \pm \chi) - i\kappa/2}. \quad (2.25)$$

where the driving amplitude is taken to be a constant, ε . When $\delta_r = 0$, the difference is only in the x axis of the phase space⁶. Figure 2.4 illustrates a initial vacuum state of the photon and its two steady states in the phase space. By measuring the x quadrature, we can know which state the qubit is in.

2.5 Controlling a single qubit

In this section, we will focus on how to control a single qubit, i.e., to make a single qubit gate, which is what our research work presented later involves (e.g., the quantum feedback operations employed in Chapter 4). Implementations of two qubit gates or multi-qubit gates can be found in [11, 19]. In addition to facilitating qubit readout, a driving field can also be used to control the qubit if we choose a proper phase and stay in the right frequency regime. If we connect the transmon qubit to a transmission line with a driving tone at frequency ω_d , in the rotating frame, the qubit Hamiltonian comes with an additional driving term similar to the one in Eq. (2.23) [11]:

$$\hat{H}' = \hat{H}'_q + \hbar\varepsilon(t)(\hat{b}^\dagger e^{-i\phi_d} + \hat{b}e^{i\phi_d}), \quad (2.26)$$

where $\hat{H}'_q$ is the qubit Hamiltonian in the rotating frame, $\hat{b}^\dagger/\hat{b}$ is the qubit creation/annihilation operator, $\varepsilon(t)$ is the driving amplitude as before, and ϕ_d is the phase of the driving tone. Rewriting $\hat{b}^\dagger/\hat{b}$ in terms of Pauli operators, we have

$$\hat{H}' = \frac{\hbar\delta_q}{2}\hat{\sigma}_z + \frac{\hbar\Omega_R(t)}{2}[\cos(\phi_d)\hat{\sigma}_x + \sin(\phi_d)\hat{\sigma}_y] \quad (2.27)$$

where $\delta_q = \omega_q - \omega_d$ is the detuning from the qubit frequency to the driving frequency, and $\Omega_R(t) = 2\varepsilon(t)$ is the Rabi frequency. At resonance, we are left with an arbitrary linear combination of spin x and spin z . The Hamiltonian can then generate rotations around any axis in the xy plane by choosing a proper phase. In particular, $\phi_d = 0$ gives us an $\hat{x}$ rotation and $\phi_d = \pi/2$ gives us a $\hat{y}$ rotation. Due to the fact that any rotation in the Bloch sphere can be decomposed into $\hat{x}$ and $\hat{y}$ rotations, we can then perform an arbitrary rotation. There is also a direct way of doing a $\hat{z}$ rotation. If instead of staying on-resonance, we make the detuning large, then the dominant term in Eq. (2.27) is $\hat{\sigma}_z$. Effectively, we are then performing a $\hat{z}$ rotation the angle of which is determined by the time over which the Hamiltonian is applied.

⁶In the phase space, the two axes or quadratures are defined as $\hat{x} = \text{Re}[\hat{a}]$ and $\hat{p} = \text{Im}[\hat{a}]$

Chapter 3

Continuous measurement theory and measurement-based control

3.1 Projective measurement

We start the discussion of this section with the standard projective measurement in quantum mechanics. Say we have some initial state ρ and a Hermitian observable $\hat{A}$ which we would like to measure. The operator $\hat{A}$ gives us a spectrum of real eigenvalues a_i and the corresponding eigenstates $|a_i, j\rangle$, where the additional label j denotes the possible degeneracy of the spectrum. From these eigenvectors, we can construct a set of projection operators¹ $\{\Pi_i | i = 1, \dots\}$

$$\Pi_i = \sum_j |a_i, j\rangle \langle a_i, j| \quad (3.1)$$

These operators are mutually orthogonal and complete

$$\Pi_i \Pi_j = \delta_{ij} \quad (3.2)$$

$$\sum_{i=1} \Pi_i \Pi_j = I \quad (3.3)$$

where I is the identity operator. Each projector correspond to a measurement outcome a_i . The state ρ collapses into one of the eigenstates of $\hat{A}$ after the measurement. The resulting conditioned state given the outcome a_i is

$$\rho \rightarrow \rho^{(i)} = \frac{\Pi_i \rho \Pi_i}{\text{Tr}(\Pi_i \rho \Pi_i)} \quad (3.4)$$

and this happens with probability

$$P(i) = \text{Tr}(\Pi_i \rho \Pi_i) = \text{Tr}((\Pi_i)^2 \rho) = \text{Tr}(\Pi_i \rho) \quad (3.5)$$

An interesting observation here is that the conditioned state will not necessarily be pure when the initial state is mixed and the outcome eigenspace is degenerate.

¹A projection operator is Hermitian and the square of it equals to itself.

3.2 Generalized measurement

Von Neumann extended the concept of measurement beyond the projective case by formulating it as a dynamical procedure [20]. To motivate the formulation, let's think about how we measure an object in real world. For example, when we measure the current through a resistor, we connect it in parallel to an ammeter. The pointer of the ammeter will then move from its 0 initial value to the value of the current we are measuring at the present time (the initial value is zero if we calibrated it carefully, but the specific value doesn't matter here). By reading out the value at which the pointer is, we know how large the current is. Now to retell the same story in a quantum setting, this means that to measure an object, referred as the *target*, we couple it to a measurement device, referred as the *probe*. Let's say the target is in an initial state $\rho_s(0)$ and the probe is in a default state $|0\rangle$ which is independent of the target state. Then we let the probe interact with the target, which leads to a unitary evolution of the combined target/probe state. Let's denote the unitary operator by $U(t, 0)$. At a later time t , we perform an **instantaneous** projective measurement on the **probe**. This means we project the combined system onto the probe's basis $\{|n\rangle | n = 0, 1, \dots\}$, referred as the *pointer states*. The label of the pointer states can be interpreted as the outcome of the measurement, or the pointer value ². Rewriting the unitary operator in the basis $\{|n, \alpha\rangle \equiv |n\rangle|\alpha\rangle\}$ where $\{|\alpha\rangle\}$ are the basis states of the target, we have

$$\begin{aligned} U(t, 0) &= \sum_{m,n,\alpha,\beta} |m\rangle\langle m| \otimes |\alpha\rangle\langle\alpha| U(t, 0) |n\rangle\langle n| \otimes |\beta\rangle\langle\beta| \\ &= \sum_{m,n} |m\rangle \left(\sum_{\alpha,\beta} \langle m, \alpha | U(t, 0) | n, \beta \rangle |\alpha\rangle\langle\beta| \right) \langle n| \\ &= \sum_{m,n} \Omega_{mn} |m\rangle\langle n|, \end{aligned} \tag{3.6}$$

where we have defined the set of operators acting on the target state

$$\Omega_{mn} \equiv \langle m | U(t, 0) | n \rangle = \sum_{\alpha,\beta} \langle m, \alpha | U(t, 0) | n, \beta \rangle |\alpha\rangle\langle\beta|. \tag{3.7}$$

Since for any unitary operator we have $U^\dagger U = U U^\dagger = I$, the Ω_{mn} operators must satisfy some completeness relation

$$\begin{aligned} \sum_i \Omega_{mi} \Omega_{ni}^\dagger &= \delta_{mn} I \\ \sum_i \Omega_{im}^\dagger \Omega_{in} &= \delta_{mn} I. \end{aligned} \tag{3.8}$$

²We assume no degeneracy in the pointer states for simplicity.

Using Eq. (3.4) & (3.5), the whole system state at time t after the measurement conditioned on the pointer value n will be

$$\begin{aligned}\rho_{\text{tot}}^{(n)}(t) &= \frac{1}{p_n} |n\rangle\langle n| [U(t, 0) (|0\rangle\langle 0| \otimes \rho_s(0)) U(t, 0)^\dagger] |n\rangle\langle n| \\ &= \frac{1}{p_n} |n\rangle\langle n| \left(\langle n| U(t, 0) |0\rangle \right) \rho_s(0) \left(\langle 0| U(t, 0)^\dagger |n\rangle \right) \langle n| \\ &= \frac{1}{p_n} |n\rangle\langle n| \otimes (\Omega_{n0} \rho_s(0) \Omega_{n0}^\dagger).\end{aligned}\tag{3.9}$$

Note that the subscript $n0$ indicates the change in the pointer value from 0 to n induced by the measurement. Here the denominator p_n is the trace of the nominator and also turns out to be the probability

$$\begin{aligned}p_n &= \text{Tr}(|n\rangle\langle n| U(t, 0) \left(|0\rangle\langle 0| \otimes \rho_s(0) \right) U(t, 0)^\dagger |n\rangle\langle n|) \\ &= \text{Tr}_s(\Omega_{n0}^\dagger \Omega_{n0} \rho_s(0))\end{aligned}\tag{3.10}$$

where the notation Tr_s denotes that the trace is taken over the target system. While this equation seems simple and intuitive, an important observation derives from it. The left hand side (LHS) represents the statistics we can obtain by measuring the probe. The right hand side (RHS) has the target's initial state, $\rho_s(0)$. This equation thus tells us how the measurement statistics of the probe contains the target's initial state information.

We are interested in the target state after the measurement, which can be computed by tracing over the probe degree of freedom on $\rho_{\text{tot}}^{(n)}(t)$:

$$\rho_s^{(n)}(t) = \text{Tr}_p(\rho_{\text{tot}}^{(n)}(t)) = \frac{\Omega_{n0} \rho_s(0) \Omega_{n0}^\dagger}{\text{Tr}_s(\Omega_{n0}^\dagger \Omega_{n0} \rho_s(0))},\tag{3.11}$$

where the notation Tr_p means tracing over the probe states. Comparing Eq. (3.9)-(3.11) with Eq. (3.4) & (3.5), we see a similarity, i.e., the projective operator Π_i and probability p_i in the latter have been replaced by the operator Ω_{n0} and probability p_n , respectively, in the former. In addition, the operator Ω_{n0} also satisfies the completeness relation to preserve the trace (this is equivalent to setting $m = n = 0$ in Eq. (3.8))

$$\sum_n \Omega_{n0}^\dagger \Omega_{n0} = I.\tag{3.12}$$

The difference is that now the operators don't need to be mutually orthogonal and there can be non-zero overlaps between the conditioned states $\rho_s^{(n)}(t)$ of different measurement outcomes. Thus the set $\{\Omega_{n0} | n = 0, \dots\}$ describes a measurement process just as well as the set of projection operators Π_i . In fact, it can be shown that any set of operators $\{\Omega_i | i = 0, \dots, n\}$ subject to the completeness condition

$$\sum_n \Omega_n^\dagger \Omega_n = I\tag{3.13}$$

describes a measurement procedure in the von Neumann sense that we have discussed above (See [21] Chapter 2). We call this kind of generalized measurement a *positive operator-valued measure* (POVM) and Ω_n the *measurement operators*. The term POVM comes from the fact that the operator $\Omega_{n0}^\dagger \Omega_{n0}$ appearing in the conditioned probability (Eq. (3.10)) is positive and thus guarantees that the probability is positive. We also define the *unconditioned state* of the POVM measurement as

$$\rho_s^{\text{uncon}}(t) = \sum_n p_n \rho_s^{(n)}(t) = \sum_n \Omega_{n0} \rho_s(0) \Omega_{n0}^\dagger \quad (3.14)$$

Conceptually, the unconditioned state means we have done the POVM measurements but do not look at the individual outcomes, only at the averaged result. It is easy to show this unconditioned state is nothing but the reduced state (Eq. (1.9) of the target

$$\begin{aligned} \rho_s(t) &\equiv \text{Tr}_p(\rho_{\text{tot}}(t)) = \sum_n \langle n | U(t, 0) | 0 \rangle \rho_s(0) \langle 0 | U(t, 0)^\dagger | n \rangle \\ &= \sum_n \Omega_{n0} \rho_s(0) \Omega_{n0}^\dagger = \rho_s^{\text{uncon}}(t), \end{aligned} \quad (3.15)$$

where we have traced out over the probe's degrees of freedom.

Using POVMs for probing quantum states

It seems that we have already understood what a generalized quantum measurement is. But wait, other than giving a bunch of definitions and terminologies, did we really explain what we can get from the target/probe measurement? Let's remind ourselves what the goal of the experiment is — this is to **learn the target** by reading out the value of the probe. Suppose the quantity that we actually want to know is the target observable $\hat{A}$, which has eigenstates $|\alpha\rangle$ with corresponding eigenvalue α . We want to learn which eigenstate the target is in. To gain this information, we must have strong correlations between the basis states $|\alpha\rangle$ of the observable $\hat{A}$ and the pointer states $|n\rangle$ of the probe. Using the resistor-ammeter measurement as an analogue, we have some current going I through the resistor: after we couple the ammeter to the resistor the pointer always moves from the default state to a **unique** state which is labelled by the same I each time that we make the measurement. It will not end up at any other state that is labelled by a different current value. To put it in other words, the current has a very strong correlation with the pointer position. In this case for the ammeter, the value of the current passing through it has a one-to-one correspondence to the pointer position.

Let us come back to the quantum target/probe system. To make the dynamics during the measurement more transparent, let us write down the Hamiltonian of the whole system

$$H_{\text{tot}} = H_s + H_p + H_I, \quad (3.16)$$

where H_s , H_p is the Hamiltonian of the target and probe, and H_I is the interaction. Say the target is initially in the eigenstate $|\alpha\rangle$ of $\hat{A}$ and the probe is in the default state $|0\rangle$. Assume the combined state will still be a product state at time t ,

$$|0, \alpha\rangle \rightarrow |\psi_\alpha(t)\rangle|\alpha\rangle, \quad (3.17)$$

where $|\psi_\alpha\rangle$ is the probe state which depends on the target eigenvalue α . This can be realized if $[H_s, H_I] = [H_s, \hat{A}] = [H_I, \hat{A}] = 0$. If we prepare the target at a different state $|\beta\rangle$ in the beginning, the probe state will be $|\psi_\beta(t)\rangle$ at time t . We can then learn well from the probe state which state the target is in, provided that the two probe states $|\psi_\alpha(t)\rangle$, $|\psi_\beta(t)\rangle$ are well separated. In this case, the two states are close to the basis states of some **probe observable** $\hat{P}$. We can measure $\hat{P}$ and its basis states serve as the pointer states. Equivalently, another good candidate for a probe observable $\hat{P}$ is one such that its expectation values with respect to the states $|\psi_\alpha(t)\rangle$ and $|\psi_\beta(t)\rangle$ are far apart enough compared to their corresponding variances. So by measuring the expectation values of $\hat{P}$, we can then distinguish the target state $|\alpha\rangle$ and $|\beta\rangle$. Hence, effectively, we have implemented the projective measurement of $\hat{A}$ through a combination of the target-probe interaction and projective measurements on $\hat{P}$.

The dispersive readout described in Sec. 2.4 is one such example. The target (transmon qubit) observable that we want to learn is $\hat{\sigma}_z$, while the probe (photon) observable that we actually measure is the quadrature $\hat{x}$. Also notice, for the dispersive Hamiltonian (2.23), we do have $[H_s, H_I] = [\hat{\sigma}_z, H_s] = [\hat{\sigma}_z, H_I] = 0$. We prepare the probe in the vacuum state which eventually evolves into the qubit-state-dependent steady state $|\alpha_{e/g}^s\rangle$ (Eq. (2.25)). For a strong enough dispersive coupling strength χ , the two averages $\langle\alpha_e^s|\hat{x}|\alpha_e^s\rangle = \text{Re}(\alpha_e^s)$, $\langle\alpha_g^s|\hat{x}|\alpha_g^s\rangle = \text{Re}(\alpha_g^s)$ are well separated as in Figure 2.4. Then by measuring the quadrature $\hat{x}$, we can learn the qubit state information.

Quantum non-demolition measurement

Using the generalized measurement scheme discussed above, such a POVM measurement is defined as a quantum non-demolition measurement (QND measurement) for some observable $\hat{A}$ if the probability distribution of the outcomes for a projective measurement of $\hat{A}$ remains unchanged before and after the generalized measurement period [22, 23]. The motivation of the QND concept is that we don't want to affect the state in terms of the measurement statistics. A sufficient condition for this is clearly

$$[\hat{A}, H_{\text{tot}}] = [\hat{A}, H_s + H_I] = 0. \quad (3.18)$$

The definition of QND measurement can also be restated as that all the moments of $\hat{A}$ do not change, i.e., for $k = 1, 2, \dots$, we have

$$\text{Tr}(\hat{A}^k \rho_s(0)) = \text{Tr}(\hat{A}^k \rho_s(t)) \quad (3.19)$$

where $\rho_s(t)$ is the reduced state of the target (see Eq. (1.9) & (1.10)). Using Eq. (3.14) & (3.15), Eq. (3.19) can be written as

$$\text{Tr}(\hat{A}^k \rho_s(t)) = \sum_n \text{Tr}(\Omega_n^\dagger \hat{A}^k \Omega_n \rho_s(0)) = \text{Tr}(\hat{A}^k \rho_s(0)), \quad (3.20)$$

where we have dropped the subscript 0 corresponding to the initial pointer state value in the measurement operators for notational simplicity. Inserting the completeness condition (Eq. (3.8)) to the RHS, the above equation can be rewritten as

$$\sum_n \text{Tr}(\Omega_n^\dagger [\hat{A}^k, \Omega_n] \rho_s(0)) = 0 \quad (3.21)$$

If this equation is true for all initial states $\rho_s(0)$, we have another sufficient condition for a QND measurement, namely

$$[\hat{A}, \Omega_n] = 0 \quad \text{for all } n. \quad (3.22)$$

In particular, the projective measurement of $\hat{A}$ itself is a QND measurement. The second sufficient condition, Eq. (3.22), is weaker than the first one, Eq. (3.18), since we do not impose any condition on other Ω_{mn} operators (Eq. (3.7)) and in Eq. (3.22) we are only considering the measurement operators defined at the end time of the whole process.

An interesting observation follows from the QND condition Eq. (3.22). We assume further that there is no degeneracy in the eigenbasis of $\hat{A}$. Due to the commutativity of $\hat{A}$ and Ω_n , we can then write the measurement operators in the common eigenbasis of A as

$$\Omega_n = \sum_\alpha f(n, \alpha) |\alpha\rangle \langle \alpha|, \quad (3.23)$$

where $f(n, \alpha) = \langle \alpha | \Omega_n | \alpha \rangle$ is in general a complex number. Let s_n be the eigenvalue of the probe observable $\hat{P}$ corresponding to the state $|n\rangle$, i.e., s_n is the pointer value. Denoting the reduced density matrix for the probe as $\rho_p(t)$ and making use of Eq. (3.10), we find after some algebra that

$$\begin{aligned} \text{Tr}(\hat{P}^k \rho_p(t)) &= \sum_n (s_n)^k p_n = \sum_n (s_n)^k \sum_\alpha \langle \alpha | \Omega_n^\dagger \Omega_n \rho_s(0) | \alpha \rangle \\ &= \sum_\alpha \langle \alpha | \rho_s(0) | \alpha \rangle \sum_n (s_n)^k |f(n, \alpha)|^2 \\ &= \sum_\alpha g_k(\alpha) \langle \alpha | \rho_s(0) | \alpha \rangle \\ &= \text{Tr}(g_k(\hat{A}) \rho(0)), \end{aligned} \quad (3.24)$$

where we have defined

$$g_k(\alpha) = \sum_n (s_n)^k |f(n, \alpha)|^2. \quad (3.25)$$

Notice that $|f(n, \alpha)|^2$ is in fact a probability distribution function for n due to the completeness condition (Eq. (3.8)). The information we can obtain by directly and repeatedly measuring the probe observable $\hat{P}$ is $\text{Tr}(\hat{P}^k \rho_p(t))$ for $k = 1, 2, \dots$. Eq. (3.24) tells us how this is related to the expectation values of a set of functions of the target observable $\hat{A}$.

3.3 Formalism of continuous measurement

In this section, we follow the approach in [24] to introduce continuous measurements. Say we want to measure an observable X of a quantum system using the POVM approach. We start the whole process at time 0 and stop at T . Let us divide the whole period into small intervals with length Δt . On each of the intervals, we repeatedly perform the POVM measurement described by a set of measurement operators $\{\Omega(\alpha) | \alpha \in \mathbb{R}\}$ ³. In general $\Omega(\alpha)$ depends on the interval length Δt . This is called a continuous measurement if as $\Delta t \rightarrow 0$, $\Omega(\alpha) \rightarrow I$ for all α . Without loss of generality, assume the operator X has a continuous spectrum

$$X|x\rangle = x|x\rangle. \quad (3.26)$$

Let us further assume the spectrum is non-degenerate and the measurement operators of the POVM commute with X :

$$[X, \Omega(\alpha)] = 0 \quad \text{for all } \alpha. \quad (3.27)$$

This is indeed a QND measurement. These two additional assumptions imply i) that $\Omega(\alpha)$ is diagonal in the $|x\rangle$ basis, and ii), mutual commutativity between the measurement operators. Finally, we also assume the measurement operators have a Gaussian form in the basis $|x\rangle$:

$$\begin{aligned} \Omega(\alpha) &= \left(\frac{2\kappa\Delta t}{\pi}\right)^{1/4} e^{-\kappa\Delta t(X-\alpha)^2} \\ &= \left(\frac{2\kappa\Delta t}{\pi}\right)^{1/4} \int_{-\infty}^{\infty} dx e^{-\kappa\Delta t(x-\alpha)^2} |x\rangle\langle x|. \end{aligned} \quad (3.28)$$

Here κ is the measurement strength. The prefactor in front is to make the measurement operators satisfy the completeness relation (Eq. (3.8)). Consider the time interval $[t, t + \Delta t]$. Given the system state $\rho(t)$ at time t , the probability to have an outcome α at $t + \Delta t$ is (Eq. (3.10))

$$p(\alpha) = \text{Tr}(\Omega(\alpha)^\dagger \Omega(\alpha) \rho(t)). \quad (3.29)$$

³It doesn't matter whether α is continuous or discrete. Unlike in the previous section, we choose it to be continuous here to facilitate our discuss below.

Its expectation value can be computed as

$$\begin{aligned}
 \langle \alpha \rangle(t + \Delta t) &= \int_{-\infty}^{\infty} d\alpha \alpha p(\alpha) \\
 &= \int_{-\infty}^{\infty} d\alpha \alpha \text{Tr}(\Omega(\alpha)^\dagger \Omega(\alpha) \rho(t)) \\
 &= \sqrt{\frac{2\kappa\Delta t}{\pi}} \int_{-\infty}^{\infty} dx \langle x | \rho(t) | x \rangle \int_{-\infty}^{\infty} d\alpha \alpha e^{-2\kappa\Delta t(\alpha-x)^2} \\
 &= \int_{-\infty}^{\infty} dx x \langle x | \rho(t) | x \rangle \\
 &= \langle X \rangle(t),
 \end{aligned} \tag{3.30}$$

where $\langle X \rangle$ denotes the trace average. We would like to make some comments about this expression. First, the LHS of the equation is a quantity which we can get by directly measuring the probe. So this equation is a demonstration of how we get to learn the target observable X by measuring the probe. Second, this equation does not hold in general. We have used here the fact that $\Omega(\alpha)$ commutes with X and also the special properties of the Gaussian function.

The probability $p(\alpha)$ can also be computed

$$\begin{aligned}
 p(\alpha) &= \text{Tr}(\Omega(\alpha)^\dagger \Omega(\alpha) \rho) \\
 &= \sqrt{\frac{2\kappa\Delta t}{\pi}} \int_{-\infty}^{\infty} dx \langle x | \rho | x \rangle e^{-2\kappa\Delta t(\alpha-x)^2}.
 \end{aligned} \tag{3.31}$$

When $\kappa\Delta t \ll 1$, the Gaussian weight is much wider than the function $\langle x | \rho | x \rangle$. We can thus replace $\langle x | \rho | x \rangle$ by a delta function centered at the expectation value $\langle X \rangle$

$$\begin{aligned}
 p(\alpha) &\approx \sqrt{\frac{2\kappa\Delta t}{\pi}} \int_{-\infty}^{\infty} dx \delta(x - \langle X \rangle) e^{-2\kappa\Delta t(\alpha-x)^2} \\
 &= \sqrt{\frac{2\kappa\Delta t}{\pi}} e^{-2\kappa\Delta t(\alpha-\langle X \rangle)^2}.
 \end{aligned} \tag{3.32}$$

Thus α is a Gaussian random variable with mean $\langle X \rangle$ and variance $1/4\kappa\Delta t$. We can write it in an alternative way

$$\alpha(t + \Delta t) = \langle X \rangle(t) + \frac{\Delta W_t}{2\sqrt{\kappa\Delta t}}, \tag{3.33}$$

where ΔW_t is a Gaussian random variable with zero mean and variance Δt . We have denoted the time-dependence of W_t explicitly with a subscript, to stress the point that this is a time-dependent process. Also notice that the values of ΔW_t on different time intervals are independent, since the measurements that we are performing are not correlated. To make things simpler, we assume that the initial state at $t = 0$ is pure. Then the conditioned state

would remain pure during the entire measurement procedure by construction (assuming that the measurement is the only source of dynamics). We then can safely replace $\rho(t)$ by $|\psi(t)\rangle$. The **conditioned** state after the measurement is (see Eq. (3.11)⁴)

$$\begin{aligned} |\psi(t + \Delta t)\rangle &\propto \Omega(\alpha)|\psi(t)\rangle \\ &\propto e^{-\kappa\Delta t(\alpha - X)^2}|\psi(t)\rangle \\ &\propto e^{-\kappa\Delta tX^2 + X(2\kappa\langle X\rangle\Delta t + \sqrt{\kappa}\Delta W)}|\psi(t)\rangle \end{aligned} \quad (3.34)$$

where we have ignored the normalization factor and dropped all the scalar terms in the exponential. The **continuous** limit is obtained by making $\Delta t \rightarrow 0$. Normalizing and expanding the above equation, keeping everything up to order $\mathcal{O}(\Delta t)$ ⁵ and replacing Δt by dt , and ΔW_t by dW_t , we have the normalized *stochastic Schrödinger equation* (SSE)

$$d|\psi\rangle = \left[-\frac{\kappa}{2}(X - \langle X\rangle)^2 dt + \sqrt{\kappa}(X - \langle X\rangle)dW_t\right]|\psi(t)\rangle. \quad (3.35)$$

Notice this equation is non-linear, due to the presence of the trace term $\langle X\rangle = \text{Tr}(X\rho)$. In the continuous limit, the random variable ΔW_t goes to its infinitesimal version dW_t which has the properties

$$\overline{dW_t} = 0 \quad (3.36)$$

$$(\overline{dW_t})^2 = dt, \quad (3.37)$$

and at different times t , the values of dW_t are independent. In fact, dW_t is the infinitesimal increment of a Brownian motion [25]. Dividing dW_t by dt , we have the so called *white noise* $\xi(t) = dW_t/dt$. We rewrite the measurement outcome (3.33) in an increment form

$$dy \equiv \alpha dt = \langle X\rangle dt + \frac{1}{2\sqrt{\kappa}}dW. \quad (3.38)$$

This quantity is called the *measurement record* and is the data we continuously collect during the continuous measurement. The SSE (3.35) describes the evolution of a state conditioned on a stream of measurement outcomes $\{\alpha(t)|t : 0 \rightarrow T\}$. The solution $|\psi(t)\rangle$ of the equation given a particular realization of the outcomes $\{\alpha(t)|t : 0 \rightarrow T\}$ is called a *quantum trajectory*. Figure 3.1 shows how the quantum trajectory looks like for a continuous measurement on a single qubit. To allow the situation for a mixed initial state, we can also derive the *stochastic master equation* (SSE) in the same way as for the pure state from Eq. (3.11). The final result is

$$d\rho = \kappa\mathcal{D}[X]\rho dt + \sqrt{\kappa}\mathcal{H}[X]\rho dW_t \quad (3.39)$$

where, in the general case when X is not Hermitian, the two superoperators are given by

$$\begin{aligned} \mathcal{D}[X]\rho &= X\rho X^\dagger - \frac{1}{2}(X^\dagger X\rho + \rho X^\dagger X) \\ \mathcal{H}[X]\rho &= X\rho + \rho X - \rho\text{Tr}(X\rho + X^\dagger\rho) \end{aligned} \quad (3.40)$$

⁴The pure state version trivially derives from this equation.

⁵Note in the infinitesimal limit, $(\Delta W_t)^2 = \Delta t$.

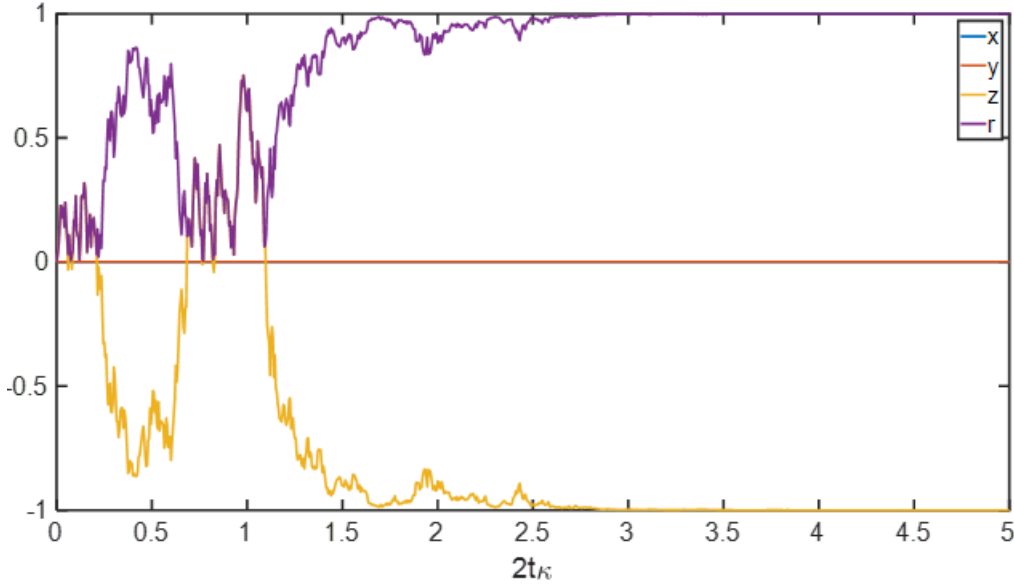


Figure 3.1: Illustration of a quantum trajectory for a single qubit continuously measured along the $\hat{z}$ axis. The qubit is initialized in the completely mixed state. The measurement operator is $X = \hat{\sigma}_z$, and the measurement strength $\kappa = 1$. The notation x, y, z refers to the three Bloch components and $r = \sqrt{x^2 + y^2 + z^2}$ is the Bloch vector length. The x and y values stay at zero while z gradually approaches -1 , suggesting that the qubit is projectively measured into the $|1\rangle$ state in the long run. The fact that r gradually approaches 1 from 0 is an indication that the qubit is being purified by the measurement, i.e., being driven by the measurement from a completely mixed state to a pure state.

Now let us pause a little and try to understand this SME a bit more. We can get the unconditioned state by averaging out the second term. So then we are only left with the first term, which denotes pure dephasing in the eigenbasis of the observable X . In the long time limit, it results in complete dephasing from any superposition state, the same effect that we would observe from a projective measurement on X . The second term denotes the continuous information gain from measurements. It makes the eigenstates of X the only possible steady states and rules out the mixed steady states as allowed by the first term⁶. Hence, in the long time limit, the continuous measurement results in a projective measurement on X (See Figure 3.1). For the sake of completeness, we also show below the continuous measurement SME for a system with coherent evolution and an inefficient detector:

$$d\rho = -i[H, \rho] + \kappa \mathcal{D}[X]\rho dt + \sqrt{\eta\kappa} \mathcal{H}[X]\rho dW_t. \quad (3.41)$$

Here the Hamiltonian term can be simply included as a commutator term, due to the fact that the unitary evolution operator commutes with the POVM operator (Eq. (3.28)) in the

⁶This is a consequence of the non-linear term $\text{Tr}(X\rho)\rho$ in the second term.

infinitesimal time interval limit. The quantity $\eta \in [0, 1]$ is the *measurement efficiency*, which describes how effectively the detector can collect data. Qualitatively, when $\eta < 1$, we only have partial access to the measurement records. We will omit its precise definition and the derivation to include it, which can be found in [24]. In particular, we want to point out the detector efficiency does not affect the evolution of the unconditioned state, i.e. if we average the SME over all trajectories, we are only left with the first two terms and η is absent.

3.4 Implementation of continuous measurement on a computer

In this section, we would like to discuss the various numerical approaches we used for our research work to simulate continuous measurement processes described by the following stochastic master equation (SME)

$$d\rho = \kappa \mathcal{D}[X]\rho dt + \sqrt{\kappa\eta} \mathcal{H}[X]\rho dW_t \quad (3.42)$$

where dW_t is the white Gaussian noise process with zero mean and variance dt , κ is the measurement strength, and η is the detector efficiency. The two superoperators $\mathcal{D}[\cdot]$ and $\mathcal{H}[\cdot]$ acting on ρ are as in Eq. (3.40). At each time step, the incremental measurement signal is

$$dy = \frac{\langle X + X^\dagger \rangle}{2} dt + \frac{dW}{\sqrt{4\kappa\eta}} \quad (3.43)$$

where the bracket notation denotes the trace average $\langle A \rangle = \text{Tr}(A\rho)$ for an arbitrary operator A . In the following subsections, we first summarize the different approaches to integrate Eq. (3.42) and then finally we compare the advantages and disadvantages of each approach.

POVM method

In Jacob and Steck's paper [24], they formalized the outcome of an **ideal** ($\eta = 1$) continuous measurement as a Gaussian process, which is valid for ideal quantum measurements by a coherent state probe field, and derived the corresponding positive-operator valued measure (POVM) when X is **Hermitian**

$$M_{dy} \propto e^{-\kappa dt (\frac{dy}{dt} - X)^2} \quad (3.44)$$

where we dropped the normalization factor (hence we use the notation $\propto$). The resulting conditioned state given the output dy is then

$$\rho(t + dt) = \frac{M_{dy}\rho(t)M_{dy}^\dagger}{\text{Tr}(M_{dy}\rho(t)M_{dy}^\dagger)} \quad (3.45)$$

Expanding the above equation up to order dt , we will recover the master equation (3.42). I will briefly describe the steps here. First, we expand the POVM operator up to order dt

$$\begin{aligned} M_{dy} &\propto e^{-\kappa dt((X - \langle X \rangle)^2 - \frac{(X - \langle X \rangle)dW}{\sqrt{\kappa dt}} + \frac{1}{4\kappa dt})} \\ &\propto e^{-\kappa dt((X - \langle X \rangle)^2 - \frac{(X - \langle X \rangle)dW}{\sqrt{\kappa dt}})} \\ &\approx 1 - \frac{\kappa}{2}dt(X - \langle X \rangle)^2 + \sqrt{\kappa}(X - \langle X \rangle)dW \end{aligned} \quad (3.46)$$

To obtain the second line, we have dropped the scalar term in the exponent since it will only contribute to the norm. Then using Eq. (3.45), we arrive at the master equation

$$\begin{aligned} \rho(t + dt) &\propto M_{dy}\rho(t)M_{dy}^\dagger \\ &= \rho + (-k\{(X - \langle X \rangle)^2, \rho\} + 2k(X - \langle X \rangle)\rho(X - \langle X \rangle))dt + \sqrt{2k}dW\{(X - \langle X \rangle), \rho\} \\ &= \rho + \kappa\mathcal{D}[X]\rho dt + \sqrt{\kappa}\mathcal{H}[X]\rho dW \end{aligned}$$

Notice this form is already normalized so we can ignore the denominator in Eq. (3.45) safely. Then we can use the POVM operator (3.44) to numerically simulate a continuous measurement provided $\eta = 1$ and X is Hermitian. We just discretize the process and at each time step and replace the infinitesimal dt with a finite step size Δt . Then we update the density matrix with Eq. (3.44) and (3.45). In some works from the Whaley group, e.g., [26], and our work [8], we have used this exponential-form POVM directly.

Kraus decomposition method

Eq. (3.42) can be rewritten in a Kraus decomposition form since the measurement is actually a completely positive map

$$\rho(t + dt) = \sum_n A_n \rho(t) A_n^\dagger \quad (3.47)$$

This motivates our second method to simulate a continuous measurement. Starting from the smallest number of Kraus operators and by doing a term matching, we can find out the density matrix can be updated in the following way

$$\rho(t + dt) = \frac{\Omega_{dy}\rho(t)\Omega_{dy}^\dagger + (1 - \eta)\kappa X\rho X^\dagger dt}{\text{Tr}(\Omega_{dy}\rho(t)\Omega_{dy}^\dagger + (1 - \eta)\kappa X\rho X^\dagger dt)} \quad (3.48)$$

where

$$\Omega_{dy} = I - \frac{\kappa}{2}X^\dagger X dt + 2\kappa\sqrt{\eta}dyX \quad (3.49)$$

and dy is as defined in Eq. (3.43). Notice in the special case when X is Hermitian and $\eta = 1$, Ω_{dy} is just the linear expansion of the POVM M_{dy} . This approach has been proposed by Rouchon and Ralph [27]. We often use this on account of its good stability.

Bayesian method

The Bayesian method has root in experiments and many experimentalists use this to update the state. It provides an exact update when we continuously measure a diagonal observable in the absence of Rabi oscillation. This is the method we used in our work [9]. The physical motivation is explained in [28–31]. Supposing X is **Hermitian** and **diagonal**, the Bayesian method updates the density matrix in the following way:

$$\rho_{ij}(t + dt) = \frac{\sqrt{p_i(\dot{y}(t)) p_j(\dot{y}(t))}}{p(\dot{y}(t))} e^{-\gamma_{ij} dt} \times \rho_{ij}(t), \quad (3.50)$$

where $p_i(\dot{y}) = \sqrt{\frac{2\eta\kappa dt}{\pi}} \exp[-2\eta\kappa dt(\dot{y}(t) - \langle i|X|i\rangle)^2]$ is the conditional probability density for the output signal $\dot{y}(t)$ given that the system is in the state $|i\rangle$, $p(\dot{y}(t)) = \sum_i \rho_{ii}(t) p_i(\dot{y})$ is the unconditioned probability of the signal $\dot{y}(t)$, and $\gamma_{ij} = \kappa(1 - \eta)(\langle i|X|i\rangle - \langle j|X|j\rangle)^2/2$. The reason for our notation $\dot{y}$ will become clear below. In the case of ideal measurement, $\gamma_{ij} = 0$ and it seems that Eq. (3.50) is no different from Eq. (3.45) but now just written in components form. But there is a crucial difference when numerically implementing these two methods. In a Bayesian update the output signal is generated as

$$\dot{y} = x(t) + \frac{1}{\sqrt{4\eta\kappa}} \frac{dW}{dt}, \quad (3.51)$$

where $x(t)$ is a random variable taking the eigenvalues of the observable X and the probability to get $x(t)$ is $\langle x|\rho(t)|x\rangle$ ($|x\rangle$ is eigenstate of X corresponding to the eigenvalue x). We make $x(t)$ explicitly dependent on t , in order to emphasize the fact that its probability is changing with time although the value it can possibly take is always one of the eigenvalues. Note the difference from Eq. (3.43), in which dy is not divided by dt .

When we do a real measurement in the lab, the infinitesimal dt should be replaced by a finite step size Δt , corresponding to the discretized time interval over which we record the signal or the finite time-resolution of measurement devices. Eq. (3.51) is valid over a wider range of step size than Eq. (3.43) in the sense that the latter is only true when $\Delta t \rightarrow 0$, while the former always holds. This can be seen by looking at the probability distribution of the signal $\dot{y}$ (see Eq. (24) in [24])

$$P(\dot{y}) = \sqrt{\frac{2\kappa\Delta t}{\pi}} \sum_x \langle x|\rho|x\rangle e^{-2\kappa\Delta t(x-\dot{y})^2}, \quad (3.52)$$

where we have assumed an ideal measurement for now, i.e., $\eta = 1$. Using the conditional probability formula $P(\dot{y}) = \sum_x P(\dot{y}|x)P(x)$, this probability suggests that we can actually write the random variable $\dot{y}$ as a summation of two random variables which is exactly given by Eq. (3.51).

We can see that Eq. (3.51) is true no matter how large Δt is. An important approximation now enters here (the same approximation is made in [24]). In the limit $\Delta t \rightarrow 0$, the

Gaussian function in Eq. (3.52) has a much broader width than $\langle x|\rho|x\rangle$. So the latter can be approximated by a delta function which must be centered at $\langle X\rangle = \text{Tr}(X\rho)$. Then the equation reduces to

$$\begin{aligned} P(\dot{y}) &= \sqrt{\frac{2\kappa\Delta t}{\pi}} \sum_x \delta(x - \langle X\rangle) e^{-2\kappa\Delta t(x-\dot{y})^2} \\ &= \sqrt{\frac{2\kappa\Delta t}{\pi}} e^{-2\kappa\Delta t(\langle X\rangle - \dot{y})^2}. \end{aligned} \quad (3.53)$$

Now the random variable $\dot{y}$ is a Gaussian distribution centered at $\langle X\rangle$. Its form will reduce to Eq. (3.43) divided by dt .

With this in mind,⁷ it is now straightforward to show the equivalence between Eq. (3.50) and (3.42). For $\eta = 1$, Eq. (3.50) is just Eq. (3.45) but written in components form. When $\eta < 1$, expanding everything up to order dt , we have from the Bayesian method

$$\begin{aligned} \sqrt{p_i(\dot{y}(t))} &\propto e^{-\eta\kappa dt(\dot{y}-x_i)^2} = e^{-\eta\kappa dt[(\langle X\rangle - x_i) + \frac{1}{\sqrt{4\eta\kappa}} \frac{dW}{dt}]^2} \\ &\propto e^{-\eta\kappa[(\langle X\rangle - x_i)^2 dt + (\langle X\rangle - x_i) \frac{dW}{\sqrt{\eta\kappa}}]} \rho_{ii}(t) \\ &\approx [1 + \sqrt{\eta\kappa}(x_i - \langle X\rangle)dW - \frac{\eta\kappa}{2}(\langle X\rangle - x_i)^2 dt] \end{aligned} \quad (3.54)$$

$$e^{-\gamma_{ij}dt} \approx 1 - \kappa(1 - \eta)(x_i - x_j)^2 dt/2 \quad (3.55)$$

$$\begin{aligned} \rho_{ij}(t + dt) &\propto \sqrt{p_i(\dot{y}(t))p_j(\dot{y}(t))} e^{-\gamma_{ij}dt} \rho_{ij}(t) \\ &\approx [1 - \frac{\kappa}{2}(x_i - x_j)^2 dt + 2\sqrt{\eta\kappa}(\frac{x_i + x_j}{2} - \langle x\rangle)dW] \rho_{ij}(t) \end{aligned} \quad (3.56)$$

where x_i is a short-hand notation for $\langle i|X|i\rangle$ and we have dropped all scalar terms as before. This is exactly a restatement of Eq. (3.42) written in the eigenbasis of the operator X .

Taylor expansion method

The usual way of simulating ordinary differential equation (ODE) is based on Taylor expansion and its generalizations. In this subsection, we would discuss briefly how to use this kind of method to simulate stochastic differential equation (SDE). In general, when it comes to SDE, these traditional methods do not work as well as for simulating ODEs. The main reference for this subsection is Chapter 10 in [32].

Consider a general type of SDE

$$dY_t = a(t, Y_t)dt + \sum_{j=1}^m b^j(t, Y_t)dW_t^j \quad (3.57)$$

⁷Technically we have only proved the $\eta = 1$ case here, but it is not hard to generalize this to the non-ideal case with $\eta < 1$. See section VII in [24] on how to treat inefficient measurement as limited access to multiple ideal measurement channels.

where Y , a and b^j are in general d -dimensional vectors. We can also write it in an integral form

$$Y(t + \Delta t) = Y(t) + \int_t^{t+\Delta t} a(s, Y_s) ds + \sum_{j=1}^m \int_t^{t+\Delta t} b^j(s, Y_s) dW_s^j \quad (3.58)$$

By substituting in Y_s recursively and truncating it, we can have approximations at different orders. This is called Ito-Taylor expansion (see Eq. (5.3) in Chapter 5 of [32]). The detailed form can be very tedious to write down. We omit it here) and gives us a way to numerically simulate Eq. (3.57).

Taking $s = t$ and $Y_s = Y_t$, we have the **Euler-Maruyama** method, which is order 0.5 of strong convergence, or weakly convergent to first order⁸

$$Y(t + \Delta t) = Y(t) + a(t, Y_t)\Delta t + \sum_j b^j(t, Y_t)\Delta W_t^j \quad (3.59)$$

Substituting Y_s in once, truncating it and then Taylor expanding the function a and b , we will have the **Milstein** method, which is order 1.0 of both strong and weak convergence,

$$Y(t + \Delta t) = Y(t) + a(t, Y_t)\Delta t + b(t, Y_t)\Delta W_t + b(t, Y_t)\frac{\partial b}{\partial Y}(t, Y_t)(\Delta W^2 - \Delta t) \quad (3.60)$$

where we assumed $d = 1$ and $m = 1$ for simplicity. For the general case when $m \neq 1$ and $d \neq 1$, please refer to Eq. (3.2) and (3.3) of Chapter 10 in [32].

The Euler-Maruyama and Milstein method is widely used in simulating general type of SDEs. Going to higher orders is possible but the analytical form is more tedious and is more computationally expensive. We will not discuss them here. Interested readers can find the details in Chapter 10 of [32].

Advantages and disadvantages of each approach

In this subsection we discuss the advantages and disadvantages of each method to simulate Eq. (3.42). First of all, we would like to point out that for numerical simulations using the first three approaches, a good (empirical) rule of thumb is that the time step should be such that $dt\kappa \leq 1/20$. The Taylor-expansion-based methods are not very suitable to simulate continuous measurement, which is a physical process in the sense that the property of density matrix must be preserved during the evolution. These properties include:

1. ρ has unit trace,
2. ρ is semi-positive definite,
3. for $\eta = 1$, an initial pure state must remain pure throughout the evolution.

⁸Simply put, strong convergence characterizes how the numerical errors behave per trajectory and weak convergence characterizes how the errors behave on average. For exact definition, see page xxv and xxvi in [32].

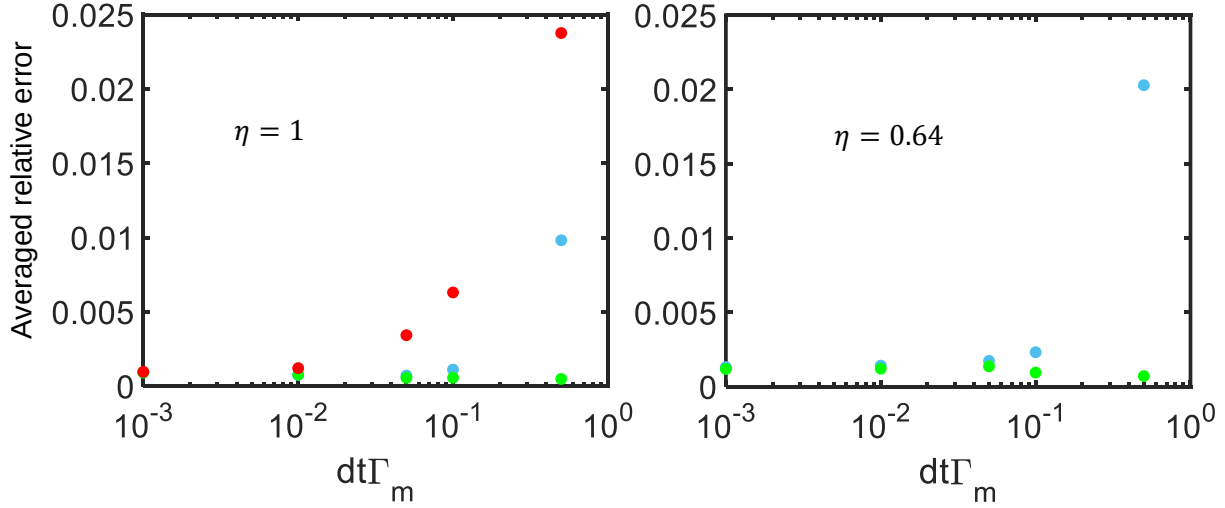


Figure 3.2: Averaged relative error as a function of time step size. The left panel is for $\eta = 1$ and the right panel is for $\eta = 0.64$. Green dots in both panels is for Bayesian method, Blue dots in both panels is for Kraus decomposition method, and red dot in the left panel is for POVM method. Parameters: $\kappa = \frac{1}{2}$, total time $T = 10\Gamma_m^{-1}$, and number of trajectories averaged over is 10000.

Taylor expansion methods do not guarantee these properties, no matter how high the order is. In contrast, it should be clear that both the POVM and Kraus decomposition method automatically preserve these properties by design. Also it is easy to see the Bayesian method guarantees the first and third property. For the second property, we have checked that the Bayesian method preserves it when ρ is a 2 by 2 matrix. When the dimension goes higher, we are currently not sure whether it still holds. But in principle, by checking the principal minors of ρ , we can find out whether it remains as a semi-positive definite matrix. So the conclusion is that we better avoid using Taylor expansion methods when simulating the continuous measurement process.

Now we compare the first three methods in terms of accuracy and computational time. First we should be clear about the conditions when each method can or should be used. The POVM method only works when X is Hermitian and $\eta = 1$.⁹ The Bayesian method only works when X is Hermitian. The Kraus decomposition method doesn't have either of these restrictions.

From the discussion in previous sections, we know Bayesian method provides a more accurate description of the output signal. For a finite step size Δt , usually the Bayesian method gives the smallest error. For POVM and Kraus decomposition method, we currently

⁹It can be extended to $\eta < 1$ by including multiple measurement channels, as noted above. But it is best to avoid using the POVM method when $\eta < 1$ since this involves sampling additional noises and taking averages.

don't have any theoretical argument which one performs better in general. Instead, we numerically simulate a continuous measurement process on a single qubit with all the three methods and computed the accuracy for each method. We choose $X = \sigma_z$ at all times and start from a completely mixed state. The analytical solution for the averaged Bloch vector length is available for this process and it turns out to be

$$\bar{r}_{\text{dW}}^{\text{analytic}}(t) = \text{erf}(\sqrt{2\kappa\eta t}) \quad (3.61)$$

where r is the Bloch vector length, the notation $(\bar{\cdot})_{\text{dW}}$ denotes average over all trajectories, and $\text{erf}(\cdot)$ is the error function. We use the time averaged relative error as the metric to determine the accuracy, which is defined as

$$\varepsilon = \frac{1}{N} \sum_{i=1}^N \frac{|\bar{r}_{\text{dW}}^{\text{analytic}}(t_i) - \bar{r}_{\text{dW}}^{\text{numeric}}(t_i)|}{\bar{r}_{\text{dW}}^{\text{analytic}}(t_i)} \quad (3.62)$$

where $\bar{r}_{\text{dW}}^{\text{numeric}}$ is the averaged Bloch vector length obtained numerically, t_i is the discretized time for each step, and N is the number of time steps in a single trajectory. We summarize the result in Fig. 3.2. At least for this particular example, it seems Bayesian method is the most accurate while POVM is the worst. In terms of computational time, one thing we need to pay attention to is that the Bayesian method involves sampling extra random variables $x(t)$ (See Eq. (3.51)). This can be very expensive if the process goes too long and if we need to generate very many trajectories. Another thing that we need to take into account is whether we should diagonalize X very frequently. When X is diagonal like the Pauli σ_z operator, of course we don't need to worry about that. When it is not diagonal but still a constant operator, for the exponential POVM and Bayesian method, we can update ρ in the eigenbasis of X and do a transformation at the end of the calculation. This is still not harmful to the computational time. The real cost comes when we change the measurement axis continuously, i.e., the operator X now depends on time. If we insist on using exponential POVM and Bayesian method in this situation, we have to diagonalize the matrix at every step and constantly transform ρ back and forth. This can significantly increase the computational time. In this situation, the Kraus decomposition method may be preferred since it doesn't involve diagonalizing X .

3.5 Measurement-based quantum feedback control

In a measurement-based quantum feedback control (MQFC), we measure the quantum system which is our control target, collect the measurement signal, Eq. (3.38) as an example, and use that classical information to decide how we make changes to the quantum system. This process involves explicit measurements on the quantum system and thus is sometimes also called incoherent control. The complete state-of-knowledge that we can have about a quantum system is represented by its density matrix. One way to do MQFC is to estimate the density matrix from the measurement records we have collected up to that point. This

estimation step, based on the classical data, of the true quantum state is referred as *quantum filtering* [33–35]. Controls based on the resulting estimations are often termed as *Bayesian* feedback control [36, 37]¹⁰. Our research work on continuous quantum error correction in Chapters 5 & 6 utilizes the idea of filtering the measurement records to estimate the state and then use a unitary feedback operator to correct any error that occurred.

Another type of MQFC is to directly feed back the measurement signals into the system without any filtering, first introduced by Wiseman and Milburn [38–40]. This approach is referred as *Markovian feedback* due to the fact that the averaged state evolution can be described by a Markovian master equation. One of the implementations of the feedback schemes we propose in Chapter 4, the ASLO implementation, can be considered as a non-trivial generalization of this approach and is also of Markovian type. Below we illustrate the Markovian control of Wiseman and Milburn, since it naturally fits into the continuous measurement setting and is commonly used. It is obtained by simply adding a Hamiltonian term in Eq. (3.39) that is proportional to the measurement signal, i.e.,

$$d\rho_c = -i[H + a(t)F, \rho_c]dt + \kappa\mathcal{D}[X]\rho_c + \sqrt{\eta\kappa}\mathcal{H}[X]\rho_c dW, \quad (3.63)$$

where $a(t)$ is the measurement signal given by $a(t) = \frac{dy}{dt}$ for the dy in Eq. (3.38) and F is some Hermitian operator which determines the form of a unitary feedback. Since the control Hamiltonian $a(t)F$ is proportional to the signal, this approach is often referred to as *proportional feedback*. Averaging over the noise term, we have the *Wiseman-Milburn master equation* for the unconditioned state

$$\dot{\rho}_c = -i[H, \rho_c] + \kappa\mathcal{D}[X]\rho_c - i\sqrt{\kappa}[F, \{X, \rho_c\}] + \frac{1}{\eta}\mathcal{D}[F]\rho_c, \quad (3.64)$$

where $\{\cdot, \cdot\}$ is the anti-commutator. The third term in the equation coherently steers the system to the desired target state and represents the goal of our control. In particular, in Chapter 4 we seek to find the locally optimal form of this term at each time t in the evolution. The fourth term causes extra decoherence due to the control, which is something we usually don't want. A fine-tuned control should also suppress the effect of this term. This decoherence term actually decays to zero in the locally optimal protocol we propose in Chapter 4.

¹⁰If we consider the continuous measurement introduced earlier, a full integration of the SME (3.39), given the collected measurement records up to that time and the initial state, can be done in principle to obtain the true density matrix. But in practice, this is computationally expensive and much slower than the time scale of the system evolution itself. The quantum filtering involves a much simplified integration of the records and thus provides only an approximation of the true density matrix.

Chapter 4

Entanglement generation for multi-qubit system

4.1 Introduction

Entanglement is a crucial resource for quantum information science, with applications in secure cryptography [41] long-range quantum state transfer [42], quantum computation [43], quantum-enhanced sensors [44–46] and quantum simulation [47]. Many applications, particularly large-scale quantum information processing, will require modular quantum devices that are able to talk to each other [48, 49]. However, the objects that we would like to entangle often have little or no direct interaction to enable entanglement generation, either because they are separated by significant distances, or because of intrinsically weak physical interactions. The former is the case in most quantum networking applications, as well as in nitrogen vacancy centers, for which large inter-qubit spacing is necessary to maintain coherence properties [50]. The latter is the case for dilute neutral atoms, which have received considerable interest for generation of spin-squeezed states involving large numbers of entangled atoms [51, 52].

When no direct interaction is available, joint measurement on all parties offers a practical method to project the total system into a desired entangled state. This has been demonstrated by interfering spontaneously emitted optical photons emitted from (artificial) atomic systems [53–56], and also by performing cavity-QED-based joint measurements on superconducting qubits [57] or cold neutral atoms [58].

However, often the stochastic nature of a quantum measurement prevents such an approach from succeeding with unit probability target [59]. In order to improve the probability of reaching an entangled state, or to obtain it deterministically, one can then add control via feedback unitary operations. Several experimental and theoretical works have shown that this approach can yield deterministic remote entanglement generation between two qubits [26, 60, 61]. Multi-partite entanglement generation has been experimentally demonstrated in Ref. [52], with a feedback strategy for deterministically preparing an ensemble of

atoms near the maximally spin-squeezed subspace. Generation of multi-spin states has also been discussed in the theoretical control literature in the context of state stabilization [62–64]. Refs. [65], [51], and [66] have proposed state-based *i.e.*, non-Markovian protocols for deterministic preparation of Dicke states. In the conclusion of Ref. [51], the authors emphasize both the experimental difficulty of implementing non-Markovian protocols and the importance of constructive methods to derive feedback control laws, but leave these issues as open problems for future research.

In this chapter we address both of these issues with a general method for constructing locally optimal measurement-based feedback protocols. Measurement-based feedback protocols rely on the fact that realistic measurements acquire only partial information over a finite time interval. Such protocols can in principle be applied to any system in which the measurement signal is collected with high efficiency, and have proven particularly useful in recent years for superconducting qubits [17, 57, 61, 67–69]. In these systems, the finite time interval is taken to be small enough that the measurements may be regarded as continuous and feedback consists of applying additional control unitary operations conditional on the outcome of the continuous measurement.

Our approach to design locally optimal protocols is based on the introduction of time-dependent feedback unitaries composed of two independently varying components, one of which depends linearly on the measurement outcome with a state-dependent coefficient, while the other depends only on the quantum state. We term this approach “Proportional and Quantum State” (PaQS) feedback, to emphasize the increased flexibility offered by these two independent feedback components. We also show how this construction can be modified to guarantee Markovianity, meaning that each feedback operation only depends on the immediately proceeding measurement outcome, which greatly simplifies experimental implementation. We demonstrate this formalism by first using it to reproduce several well-established measurement-based feedback protocols, and then giving a systematic treatment of entanglement generation in three-qubit systems and beyond, emphasizing the crucial role that symmetry plays in the protocols in order to achieve high fidelities. We consider generation of both the N -qubit GHZ states and the full range of N -qubit symmetric Dicke states, from W states to the half-filled states, which are maximally spin-squeezed. Our protocols are characterized by a high degree of symmetry and are derived using general analytic techniques that allow us study up to 100 qubits. These advantages also confer remarkably simple experimental implementations of the feedback controller.

The remainder of the chapter is organized as follows. Section 4.2 will discuss basic concepts and background knowledge of entanglement which will be involved in our work. In section 4.3, we first summarize the evolution of a general quantum state under continuous measurement and feedback as described by the stochastic master equations, and then present our PaQS formalism for efficiently computing locally optimal protocols in the context of a general measurement-based feedback system.

In sections 4.4 and 4.5 we apply the PaQS formalism to the task of generating entangled many-qubit states, considering specifically Dicke states and GHZ states. In section 4.5 we then show that the PaQS formulation of section 4.3 can be extended to systems undergo-

ing multiple distinct measurements and commuting feedback Hamiltonians, and use this to obtain a second protocol for GHZ states. A third protocol for GHZ states that is based on optimizing an entanglement measure rather than a target state fidelity is presented in appendix 4.8. Section 4.6, which is mainly Martin Leigh's (Co-author of the paper [8]) contribution, shows how under certain conditions the feedback can be chosen to eliminate the randomness introduced by the "quantum noise" inherent to the measurement process. We introduce a strict condition for deterministic evolution, given fixed measurement and feedback operators and show that this provides a simple tool for finding solutions to proportional feedback master equations in general. Section 4.7 provides a summary of the results and an outlook for future work. Key calculational details and supplementary materials are presented in the Appendix.

4.2 Entanglement in a nutshell

Simply put, entanglement is a situation where the state of a composite system cannot be represented as product states of the individual entities. It may sound simple, yet it has profound consequences and a complicated structure of its own. One telling observation was made by Schrödinger, which perhaps best explains why entanglement phenomenon is bizarre and intriguing: "The best possible knowledge of a whole does not include the best possible knowledge of its parts". In this section, we will only mention a few basic concepts and techniques that will be helpful in understanding our work. Interested readers should definitely take a look at the classic review on this topic [70].

Basic concepts

As mentioned above, entanglement can be easily defined as the negative of what is not entanglement. Consider a composite system which consists of n bodies. Restricting to pure states first, the global state $|\psi\rangle$ is *entangled* if it cannot be written in the following way

$$|\psi\rangle \neq \otimes_{i=1}^n |\psi_i\rangle, \quad (4.1)$$

for any local state $|\psi_i\rangle$ of the i th body. For a mixed state ρ , the definition is a bit more complicated. We can always write the mixed state as a convex linear combination (or an ensemble) of pure states

$$\rho = \sum_{\alpha} p_{\alpha} |\psi^{(\alpha)}\rangle \langle \psi^{(\alpha)}|. \quad (4.2)$$

Notice this pure state decomposition is not unique. The mixed state ρ is *entangled* if we cannot find a decomposition of which every pure state is not entangled. No matter whether pure or mixed, when a state is not entangled, we call it *separable*. A special case is when we can write the global mixed state as product of local mixed states

$$\rho = \otimes_{i=1}^n \rho_i, \quad (4.3)$$

where ρ_i is a possibly mixed state for the i th object. Rewriting each of the ρ_i as an ensemble of pure states

$$\rho_i = \sum_{\alpha=1} p_i^{(\alpha)} |\psi_i^{(\alpha)}\rangle \langle \psi_i^{(\alpha)}|, \quad (4.4)$$

then the global state has the form

$$\rho = \sum_{\alpha} (\prod_{i=1}^n p_i^{(\alpha)}) (\otimes_{i=1}^n |\psi_i^{(\alpha)}\rangle) (\otimes_{i=1}^n \langle \psi_i^{(\alpha)}|). \quad (4.5)$$

It is clear that this form satisfies the definition of a separable state and has one additional property—the classical ensemble probability distribution is a product of individual probabilities $p_i^{(\alpha)}$. This reminds us of the fact that for a mixed separable state, the classical joint ensemble probability distribution p_{α} can still have correlations among the individual components.

Now we have clarified the basic notion of what is an entangled state. Let's look at some concrete examples.

Bell state The four Bell states of a two-qubit system are defined as

$$|\Psi_{\pm}\rangle = \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle) \quad (4.6)$$

$$|\Phi_{\pm}\rangle = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle). \quad (4.7)$$

W state The W state is a tri-partite entangled state

$$|W\rangle = \frac{1}{\sqrt{3}}(|100\rangle + |010\rangle + |001\rangle). \quad (4.8)$$

GHZ state There is another distinct type of tri-partite entangled state called the GHZ state [71]

$$|GHZ\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle) \quad (4.9)$$

Through proof by contradiction, all three states can be verified easily to be entangled from the definition. However, in order not to mislead the readers, we would like to point out in general it can be very hard to determine whether a given state is entangled or not just from the definition itself.

Entanglement measures

We have seen that there are four Bell states. They share the common name for a reason. Every Bell state can be transformed into every other Bell state through some **local** unitary rotation acting on each qubit. A local unitary should not change the nature of the entangled state since entanglement is a global property. In this sense, all the four states are equivalent.

So we may determine whether two entangled states are the same by checking whether they are connected by some local unitary operations. But we may ask whether there is a simpler way of doing this. Can we just use some numbers to describe different entangled states?

Furthermore, the Bell states are certainly not the only entangled states of a bipartite system. We can easily construct an entangled state of the form

$$|\varepsilon\rangle \propto |01\rangle + \varepsilon|10\rangle \quad (4.10)$$

where ε is a very small positive number. Let us ignore the normalization factor for a moment. This state is entangled but very close to the product state $|01\rangle$. We may feel that this state is not as entangled as the Bell state and this feeling is indeed correct. A legitimate question to ask here is how we can describe the degree of entanglement or, more specifically, how we can quantify the entanglement.

Researchers have proposed quantities called entanglement measures to characterize and quantify entanglement. There are many of these. We will only mention a few basic ones here.¹

Schmidt rank The Schmidt theorem tells us for any bipartite system consisting of two subsystems A and B , the global state $|\psi\rangle$ can be decomposed in the following way

$$|\psi\rangle = \sum_{i=1}^n c_i |a_i\rangle |b_i\rangle, \quad (4.11)$$

where $|a_i\rangle$ and $|b_i\rangle$ are the basis states for the subsystem A and B respectively, the c_i , called Schmidt coefficients, are real non-negative numbers, and $n = \min\{\dim(A), \dim(B)\}$. The Schmidt rank is defined as the number of non-zero terms in the expression. When the Schmidt rank is 1, the state is separable. Otherwise, it is entangled. The Schmidt theorem is very useful to study bipartite systems but unfortunately the Schmidt rank doesn't tell us how much the system is entangled. For example, both the Bell states and the $|\varepsilon\rangle$ state have Schmidt rank 2.

An interesting consequence of the Schmidt theorem is that the reduced density matrix for subsystem A and B is the same for both subsystem up to a basis change, i.e. $\rho_A = \text{Tr}_B \rho$ is identical to $\rho_B = \text{Tr}_A \rho$ if we equate their basis.

Concurrence Concurrence is a measure defined for two-qubit systems. For pure states, it has the form

$$c(\rho) = \sqrt{2(1 - \text{Tr} \rho_A^2)}. \quad (4.12)$$

Expressed in terms of Schmidt coefficients, the concurrence is $c(\rho) = 2c_1c_2$. A larger concurrence means that we have more entanglement in the state. For example, the Bell states have concurrence 1 while the $|\varepsilon\rangle$ state, Eq. (4.10), has concurrence $c \propto \varepsilon$. Concurrence can fully characterize the entanglement for all pure two-qubit states [72].

¹We only discuss pure states here. These measures can be extended to mixed states through a convex roof construction [70].

Three-tangle The three-tangle is a genuine tri-partite entanglement measure for three-qubit system, in the sense that it vanishes for some tri-partite entangled states as well as for any bi-separable state. For instance, it is zero on the W state, although we cannot find a separable bipartition for it. Considering three qubits A, B and C, the definition is as follows [73, 74]:

$$\tau(A : B : C) = \tau(A : BC) - \tau(AB) - \tau(AC). \quad (4.13)$$

In contrast to the function $\tau(A : B : C)$ on the left hand side, here the functions $\tau(AB)$, $\tau(BC)$ on the right hand side act on only two qubits. They are referred to as 2-tangles and can be computed as the square of the concurrence for the two-qubit system. Note that now the reduced density matrices for both subsystem AB and subsystem AC are mixed. We need to use the convex roof extension of Eq. (4.12) to compute the 2-tangles (see [75]). The symbol A:BC means that we treat BC as a single system and consider the total system as a bipartite system between A and BC, although in this case the two-qubit system BC has dimension 4. The state of the entire system is still pure by assumption. So we can formally use Eq. (4.12) to compute $\tau(A : BC)$, being careful not to forget the square when computing the concurrence).

The three-tangle takes values ranging from 0 to 1. The GHZ state happens to have the maximum value of 1. It can also be written as a hyperdeterminant [76]. In that way, it is clear that the three-tangle has S_3 permutation symmetry as it should be.

4.3 Continuous measurement with locally optimal feedback

Continuous measurement and feedback

We consider a joint measurement that is realized by an indirect simultaneous measurement on multiple qubits. Such measurements are routinely made for superconducting qubits in the dispersive regime, using homodyne detection of cavity transmission [17, 18, 57, 61]. The readout voltage of the signal is given by

$$dV = \langle X \rangle(t)dt + \frac{dW}{\sqrt{8\eta k}} \quad (4.14)$$

where X is the measurement operator which will be specified in detail below, $\langle \cdot \rangle$ denotes the trace average of this, dW is a Wiener increment satisfying $dW(t)dW(t') = \delta(t - t')dt$ that represents the quantum uncertainty in the homodyne detection (represented by white Gaussian noise in the continuous limit [77, 78]), η is the measurement efficiency, and k is the measurement strength.

The evolution of the quantum state conditioned on this measurement signal is given by a stochastic master equation (SME),

$$d\rho = -\frac{i}{\hbar}[H_S, \rho]dt + 2k\mathcal{D}[X]\rho(t)dt + \sqrt{2k\eta}\mathcal{H}[X]\rho dW. \quad (4.15)$$

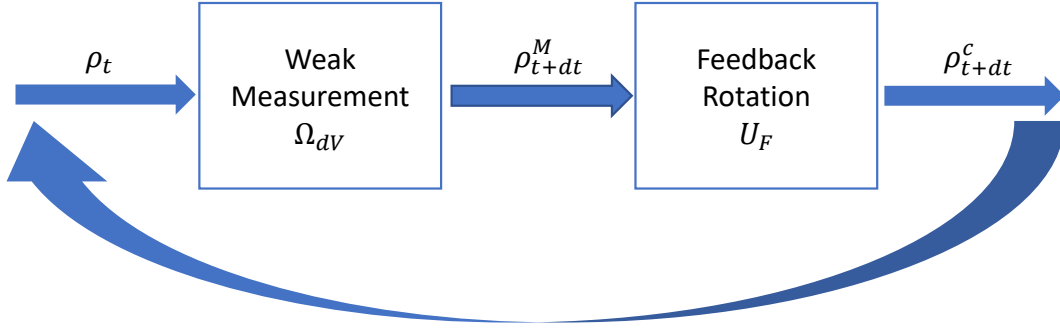


Figure 4.1: A diagrammatic picture illustrating what happens during a measurement-feedback cycle over time dt .

Here H_S is the system Hamiltonian which is not relevant to this analysis and will henceforth be set to 0, $\mathcal{D}[X]\rho = X\rho X^\dagger - \frac{1}{2}(X^\dagger X\rho + \rho X^\dagger X)$, and $\mathcal{H}[X]\rho = X\rho + \rho X^\dagger - \langle X + X^\dagger \rangle \rho$. The second, deterministic, term of order dt describes the dephasing effect of the measurement and is just the usual dissipator term in the Markovian master equation. It describes the unconditioned dynamics of an open system coupled to an external bath, *i.e.*, the dynamics after averaging over all possible measurement records. The third, stochastic, term of order dW continuously updates ρ based on knowledge gained from the measurement.

As the qubits are assumed to be too remote to allow direct interactions, we construct a unitary feedback operator in the form of local operations on each of the N qubits,

$$U_F = \bigotimes_{i=1}^N U_i. \quad (4.16)$$

In the first instance we shall employ

$$U_i = e^{-i\theta_i H_i/\hbar}, \quad (4.17)$$

with $H_i = \frac{1}{2}\hat{n} \cdot \vec{\sigma}$, where $(\theta_i, \hat{n}_i)$ are rotation angle parameters for the i th qubit. The single qubit unitaries are then simply rotation operators on a single qubit, $U_i = e^{-i\frac{\theta}{2}\hat{n} \cdot \vec{\sigma}/\hbar}$, with $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ denoting the usual Pauli operators. These rotation angles will be determined based on the measurement result. For simplicity, we assume that these rotations are realized instantaneously.

After adding the feedback, the complete evolution of the qubits over a single infinitesimal cycle of measurement and feedback (see Fig. 4.1) is described by the following:

$$\rho_{t+dt}^c = U_F \rho_{t+dt}^M U_F^\dagger = U_F(\rho_t + d\rho) U_F^\dagger \quad (4.18)$$

where $d\rho$ is taken from Eq. (4.15), the superscript M indicates it is the conditioned state resulting from the measurement, and the superscript c indicates that this state is obtained after the feedback control.

General locally optimal control protocol with PaQS

Now we are ready to state the locally optimal control problem in complete generality and to find the solution. Suppose we have a state ρ_t at time t , and that we ultimately wish to reach some target state $|\psi_T\rangle$. In this work we shall focus primarily on the fidelity with respect to the target state as the cost function, whose optimization determines the parameter of the feedback operator. (Appendix 4.8 presents optimization under an alternative, non-linear cost function given by an entanglement measure.) For the entangled states of interest in later sections, i.e., Dicke and GHZ states, we shall take all values of θ_i to be equal, to ensure that the permutation symmetry of the state is maintained by the feedback operations (sections 4.4 and 4.5) although we shall relax this assumption when considering feedback with multiple, simultaneous measurements (section 4.5). Our feedback unitary is then

$$U_F(\theta) = e^{-i\theta H_F/\hbar}. \quad (4.19)$$

For the time being, we place no constraints on H_F , although for practical settings it will often be useful to take this to be separable. The fidelity after a feedback operation is given by

$$\begin{aligned} \mathcal{F}_{t+dt}(\theta) &= \langle \psi_T | \rho^c(\theta) | \psi_T \rangle, \\ \rho^c(\theta) &\equiv U(\theta)(\rho + d\rho)U^\dagger(\theta). \end{aligned} \quad (4.20)$$

For convenience, we shall set $\hbar = 1$ from now on. Our local optimality condition is given by

$$\begin{aligned} \mathcal{G} &\equiv \frac{\partial \mathcal{F}_{t+dt}(\theta)}{\partial \theta} = \langle \psi_T | \left[U_F'(\theta)(\rho_t + d\rho)U_F^\dagger(\theta) + h.c. \right] | \psi_T \rangle \\ &= -i \langle \psi_T | [H_F, \rho^c] | \psi_T \rangle = 0 \end{aligned} \quad (4.21)$$

with U_F' the derivative of U_F .

Using the fact that $d\rho$ is infinitesimal, we can derive an analytical expression for the optimal rotation angle θ , which we denote θ^* . Assuming that ρ_t is already optimized from the previous time step, then typically θ^* will also be $\mathcal{O}(d\rho)$ (we deal with possible exceptions below). We can therefore parameterize the rotation angle as

$$\theta^* = A_1(t)dW + A_2(t)dt. \quad (4.22)$$

Expanding U_F to second order in dW and making use of Ito's lemma [78] yields

$$U \equiv I - iA_1H_FdW - (iA_2H_F + \frac{1}{2}A_1^2H_F^2)dt. \quad (4.23)$$

Expanding Eq. (4.18) and using the measurement stochastic master equation Eq. (4.15) for $d\rho$, together with this second order expansion, yields the master equation

$$\begin{aligned} d\rho^c &= \mathcal{D}[Y]\rho dt + \mathcal{H}[Y]\rho dW - i(A_1dW + A_2dt)[H_F, \rho] \\ &+ A_1^2\mathcal{D}[H_F]\rho dt - i\sqrt{\eta}A_1[H_F, Y\rho + \rho Y^\dagger]dt, \end{aligned} \quad (4.24)$$

where we have defined $Y = \sqrt{2k}X$ to suppress k in the result. Employing ρ^c in Eq. (4.21) leads to the following explicit form for $\mathcal{G}$:

$$\begin{aligned} \mathcal{G} &= -i\langle\psi_T|[H_F, \rho_t]|\psi_T\rangle \\ &- i\langle\psi_T|[H_F, \mathcal{D}[Y]\rho_t dt + \sqrt{\eta}\mathcal{H}[Y]\rho_t dW]|\psi_T\rangle - A_1\langle\psi_T|[H_F, [H_F, \rho_t]]|\psi_T\rangle dW \\ &- iA_1^2\langle\psi_T|[H_F, \mathcal{D}[H_F]\rho_t]|\psi_T\rangle dt - \langle\psi_T|[H_F, [H_F, \sqrt{\eta}A_1\mathcal{H}[Y]\rho_t + A_2\rho_t]]|\psi_T\rangle dt, \end{aligned} \quad (4.25)$$

We now solve Eq. (4.25) order by order in dW . Despite the large number of terms, it is nevertheless possible to solve for A_1 and A_2 in complete generality. The assumption that the optimal rotation was applied at the immediately preceding time step implies that $\partial\mathcal{F}_t(\theta)/\partial\theta|_{\theta=0} = -i\langle\psi_T|[H_F, \rho_t]|\psi_T\rangle = 0$, so that the first term in Eq. (4.25) may be dropped. Terms proportional to dW yield a linear equation in A_1 , which is easily solved. Once A_1 is known, terms proportional to $dW^2 = dt$ are gathered to yield another linear equation, this time for A_2 . The final result in full form is

$$\begin{aligned} A_1 &= \frac{-i\langle\psi_T|[H_F, \sqrt{\eta}(Y\rho_t + \rho_t Y^\dagger)]|\psi_T\rangle}{\langle\psi_T|[H_F, [H_F, \rho_t]]|\psi_T\rangle} \\ A_2 &= \frac{-\langle\psi_T|[H_F, i\mathcal{D}[Y]\rho_t + \sqrt{\eta}A_1[H_F, Y\rho_t + \rho_t Y^\dagger] + iA_1^2\mathcal{D}[H_F]\rho_t]|\psi_T\rangle}{\langle\psi_T|[H_F, [H_F, \rho_t]]|\psi_T\rangle}. \end{aligned} \quad (4.26)$$

Using Eq. (4.14) and (4.22), the locally optimal feedback rotation can also be written as

$$\theta^* = \sqrt{8\eta k}A_1dV + (A_2 - 2A_1\langle Y \rangle)dt, \quad (4.27)$$

from which we see that A_1 can be identified with a proportional feedback term, while the second term - dependent on A_1 , A_2 , and the state at time t - can be identified with an additional

time dependent effective Hamiltonian drive. These two terms motivate the designation of “PaQS”.

Note that as A_2 only appears in Eq. (4.24) to first order, $\partial(d\mathcal{F})/\partial A_2$ is not a function of A_2 , so then the naive method of computing the locally optimal protocol by maximizing $d\mathcal{F} \equiv \langle \psi_T | d\rho | \psi_T \rangle$ with respect to A_2 would not yield Eq. (4.26).

So far, we have assumed that the optimal angle is infinitesimal. However, Eq. (4.26) only guarantees that the solution θ^* is a local extremum and does not guarantee that it is necessarily a maximum. A sufficient condition for θ^* to be a local maximum is that the second derivative of the fidelity function evaluated at θ^* be negative:

$$\left. \frac{\partial^2 \mathcal{F}_{t+dt}(\theta)}{\partial \theta^2} \right|_{\theta=\theta^*} = -\langle \psi_T | [H_F, [H_F, \rho^c]] | \psi_T \rangle \Big|_{\theta=\theta^*} < 0. \quad (4.28)$$

Failure of this test, *i.e.*, when the second derivative is positive, suggests the presence of a local minimum from the infinitesimal solution. Then we will need a large (*i.e.*, non-infinitesimal) rotation, which we compute by maximizing the fidelity over the entire angular range.

It should be noted that A_1 and A_2 can in principle become singular. However as $\rho_{t-dt}^c|_{\theta^*} = \rho_t$, the denominator diverges only when the second derivative test failed at the previous time step (compare Eq. (4.28) to the denominator of Eq. (4.26)). Thus this divergence is typically prevented by the global search described above. A special case is $[H_F, \rho_t] = 0$, in which case feedback has no effect on the state, so that we may simply set $\theta^* = 0$.

To simulate this form of feedback in practice, we assume that the controller chooses the rotation angle θ^* that ensures a global maximum of $\mathcal{F}_{t=0}$ at the initial time step. During evolution of the state, the above protocol typically continues to pick θ^* as the global maximum of $\mathcal{F}_t$ and thus maintains the system on a locally (time-)optimal trajectory. However even if Eq. (4.28) remains negative, it is possible that the nearest local maximum of $\mathcal{F}_t(\theta)$ can fail to be the global maximum. The only way to catch such instances is to occasionally undertake a brute force maximization of $\mathcal{F}$ over the full range of θ and to thereby check whether the local maximum identified by Eq. (4.26) is also a global maximum. In practice, such global maximization procedures are often unnecessary. Table 4.1 shows many combinations of measurement operators X and feedback Hamiltonians H_F that allow Eqs. (4.26) and (4.28) to reproduce the indicated feedback protocols established in the control literature. Global searches are only required when indicated by Eq. (4.28). We shall see examples of this in the following section.

The functions A_1 and A_2 are dependent on both the initial and target states, as well as on the state at time t . Dependence on the current state implies implicit dependence on the full measurement record, yielding a potentially non-Markovian feedback protocol in general. In experimental situations, non-Markovian protocols are significantly more difficult to implement than Markovian protocols. The feedback controller must either calculate $\rho(t)$ in real time, which is both challenging and time-consuming, or else perform an exponential amount of precomputation to determine the optimal action for all possible measurement records. This is illustrated schematically in Fig. 4.2a. However we can also use the feedback

Feedback protocol	Measurement operator (X)	Feedback Hamiltonian (H_F)	Target state ($ \psi_T\rangle$)
Adaptive phase measurement [79]*	σ	σ_z (Heisenberg picture)	$(0\rangle + i 1\rangle)/\sqrt{2}$
Rapid qubit purification [80]*	σ_z	σ_y	$(g\rangle + e\rangle)/\sqrt{2}$
Half-parity Bell state preparation [26]*	$\sigma_{z,1} + \sigma_{z,2}$	$\sigma_{y,1} + \sigma_{y,2}$	$(eg\rangle + ge\rangle)/\sqrt{2}$
Full-parity Bell state preparation [60, 81]*	$\sigma_{z,1}\sigma_{z,2}$	$\sigma_{x,1}$	$(gg\rangle + i eg\rangle + i eg\rangle + ee\rangle)/2$
N-qubit Dicke states [51, 65, 66] [†]	$\sum_i \sigma_{z,i}$	$\sum_i \sigma_{y,i}$	N-qubit Dicke state with n excitations
N-qubit GHZ states [†]	$\sum_{i \neq j} \sigma_{z,i}\sigma_{z,j}$	$\sum_i \sigma_{x,i}$	N-qubit GHZ state
N-qubit GHZ states* [†]	$\sigma_{z,i} - \sigma_{z,j}$ for all i, j	$\sigma_{y,i} - \sigma_{y,j}$ for all i, j	N-qubit GHZ state
Hong-Ou-Mandel Bell state [82]*	$i(\sigma_1 + \sigma_2),$ $\sigma_1 - \sigma_2$	$\sigma_{x,1} + \sigma_{x,2}$ $\sigma_{y,1} - \sigma_{y,2}$	$(gg\rangle + ee\rangle)/\sqrt{2}$

Table 4.1: Summary of some measurement-based feedback protocols that may be derived from Eq. (4.26) or its generalization to multiple measurement and feedback operators in section 4.5. [†] denotes this work. Protocols in which perfect noise cancellation (cancellation of dW terms, see discussion in section 4.6) can occur are marked with an asterisk. Note that for N -qubit GHZ states, noise cancellation occurs only for the multiple measurement operator protocol with $N = 2$ and $N = 3$.

master equation to simulate a memoryless controller, which ensures that all realizations have the same feedback, regardless of the individual measurement records. This is the average state locally optimal (ASLO) protocol of Ref. [26]. The basic concept is illustrated in Fig. 4.2b.

In an ASLO protocol, instead of using the controlled state ρ_{t+dt}^c of Eq. (4.18) to determine the feedback operation for the next time step, we use an unconditioned state that is obtained by averaging over the entire measurement record at time t . We express the evolution of the conditioned state as

$$\rho_{t+dt}^M = \frac{\Omega_{dV} \rho_t \Omega_{dV}^\dagger}{\text{Tr}(\Omega_{dV} \rho_t \Omega_{dV}^\dagger)}, \quad (4.29)$$

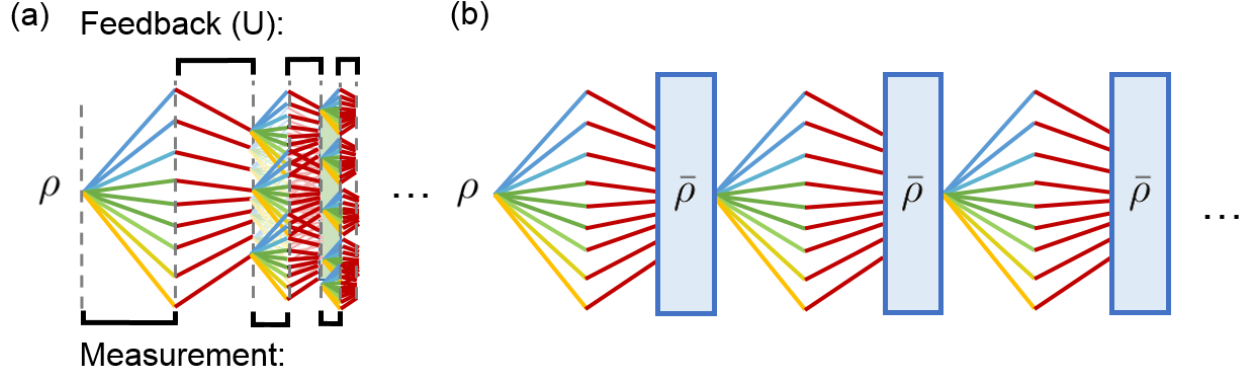


Figure 4.2: (a) Schematic of a non-Markovian feedback protocol. The number of possible states, and hence the number of potentially distinct feedback operations needed in response scale exponentially in time. We refer to this as the trajectory ensemble approach (TEA). (b) Schematic of the average state locally optimal (ASLO) feedback protocol, which applies controls that are determined using only the average time-evolved state and the most recent measurement outcome [26]. This is the behavior of a “forgetful” or Markovian controller.

where Ω_{dV} is the POVM for continuous measurement with $I = \int dV \Omega_{dV}^\dagger \Omega_{dV}$ (Eq. (3.44)):²

$$\Omega_{dV} = \left(\frac{4k}{\pi dt} \right)^{\frac{1}{4}} \exp \left[-2kdt \left(\frac{dV}{dt} - X \right)^2 \right]. \quad (4.30)$$

Eqs. (4.29) and (4.30) are equivalent to the stochastic master equation of Eq. 4.15 with $H_S = 0$. Averaging over the measurement record at time t then gives the unconditioned state

$$\bar{\rho}_{t+dt} \equiv \int_{-\infty}^{\infty} d(dV_t) P(dV_t) \rho_{t+dt}^c, \quad (4.31)$$

where $P(dV_t) = \text{Tr}(\Omega_{dV} \bar{\rho}_t \Omega_{dV}^\dagger)$ and ρ_{t+dt}^c is implicitly dependent on dV_t (Eq. (4.18)). Inspection of Eq. (4.18) shows that the averaged state $\bar{\rho}_{t+dt}$ is equal to

$$\bar{\rho}_{t+dt} = \int_{-\infty}^{\infty} d(dV) U_F \Omega_{dV} \bar{\rho}_t \Omega_{dV}^\dagger U_F^\dagger. \quad (4.32)$$

At the end of the next time step $t + 2dt$, the feedback controller then computes θ^* based on this unconditioned averaged state $\bar{\rho}_{t+dt}$ instead of based on ρ_{t+dt} . Iterating this procedure results in replacing ρ by $\bar{\rho}$ in Eq. (4.26). The state $\bar{\rho}_t$ along the evolution is then understood

²Technically, this POVM is applicable only to an ideal measurement. However, inefficient measurement can be treated as limited access to multiple ideal measurement channels. The analysis below can be thereby readily extended to inefficient measurements.

as an averaged state. It has deterministic dynamics, as can be seen by averaging over, (*i.e.*, dropping, since $\langle dW \rangle = 0$) the terms proportional to dW in the second order expansion Eq. (4.24). This allows efficient simulation of an arbitrary ASLO feedback protocol by a single trajectory. The feedback protocol is now essentially Markovian, since the dependence on the previous measurement history has been removed by the averaging.

The averaging procedure in Eq. (4.31) is a mathematical step that corresponds exactly to the averaging over trajectories with different measurement records and hence over quantum noise histories that is done in an experiment. The feedback unitary characterized by $A_1(t)$ and $A_2(t)$ for the averaged state can then be applied to any individual realization of the state at each instant. This provides significant advantages for experimental implementation, since the functions $A_1(t)$ and $A_2(t)$ can be pre-calculated efficiently with the same procedure as for the TEA approach above but using a single calculated trajectory for $\bar{\rho}(t)$, and the resulting feedback operation applied to each experimental trajectory without the need for real-time state estimation.

Finally, we note that although the general optimization formalism above has been illustrated using fidelity with respect to a desired target state as the cost function, it can straight-forwardly be applied to other cost functions that are linear in ρ , such as the expectation value of an operator. In addition, one can also go beyond linear protocols to consider alternative cost functions based on entanglement measures.

In the next two sections we apply the PaQS protocol to N -qubit Dicke states (section 4.4) and to N -qubit GHZ states (section 4.5). These two canonical examples of entangled N -qubit states are simple generalizations of the 2-qubit Bell states. For simplicity, we shall employ perfect measurement efficiency, $\eta = 1$, in all simulations throughout the rest of this chapter, although the theory described in this section is valid also for inefficient measurements with $\eta < 1$.

4.4 Generation of Dicke states

The Dicke states are defined as

$$|N, k\rangle = \frac{1}{\sqrt{\binom{N}{k}}} \sum_{P \in S_N} P(|0\rangle^{\otimes(N-k)} \otimes |1\rangle^{\otimes k}) \quad (4.33)$$

where P is an operator belonging to the permutation group S_N on N qubits. When $k = 1$, we have the well-known W state. When $k = \frac{N}{2}$ (N even or $k = \frac{N+1}{2}$ when N is odd), we have the half-filled Dicke state. $|N, n\rangle$ is a uniform superposition over all states with the same number of excitations, and is also known as a spin squeezed state. The Dicke states have been proposed for applications in a wide range of sensing protocols, including very long baseline interferometry [83] and Heisenberg-limited measurement sensitivity [84].

We shall first consider the canonical W state with $N = 3$, before addressing the generation of general Dicke states for arbitrary N values.

Generation of the $N = 3$ W state

The $N = 3$ W state, given by

$$|W\rangle = \frac{1}{\sqrt{3}}(|001\rangle + |010\rangle + |100\rangle), \quad (4.34)$$

represents a generic type of weak three qubit entanglement that is characterized by zero three-way entanglement but maximal retention of bipartite entanglement on loss of a qubit [71]. For simplicity, in the rest of this section we shall refer to the $N = 3$ state just as the W state. To prepare this state, we first choose a measurement observable

$$X_W = \sigma_{z,1} + \sigma_{z,2} + \sigma_{z,3}, \quad (4.35)$$

where $\sigma_{z,i}$ is the Pauli operator along the $\hat{z}_i$ axis for qubit i . Note that this measurement operator has the same permutation symmetry as the W state. It is also a linear combination of single-body observables, which is the easiest form to implement on spatially separated qubits, in particular, on remote qubits, such as superconducting qubits in different microwave cavities.

The W state is an eigenstate of X_W , with eigenvalue $+1$. This implies that for any initial state ρ_0 , the long time limit of evolution under continuous measurement of X_W in the absence of feedback will result in projection onto the W state with probability $p_W = \langle W|\rho_0|W\rangle$.

To increase this success probability, we apply the fidelity-optimized protocol of the previous section, choosing the target state as $|\psi_T\rangle = |W\rangle$. To choose a proper feedback rotation operator, we generalize the locally optimal two-qubit feedback unitary used to generate the 2-qubit W state $[|10\rangle + |01\rangle]/\sqrt{2}$ in Ref. [26], which was shown to provide a globally optimal protocol for both maximal fidelity and concurrence [60]. As in that work, we restrict the evolution of each qubit to lie in the xz plane ($\phi_i = 0, i = \{1, 2, 3\}$), with local rotations around the y axis having equal angles for each of the three qubits, *i.e.*, $\theta_1 = \theta_2 = \theta_3 = \theta$. The latter condition is consistent with the fact that both the target W state and the observable X_W are symmetric with respect to any permutation of the three qubits. For the W state, the rotation operator in Eq. (4.19) then takes the following form

$$U_F^W(\theta) = e^{-i\frac{\theta}{2}(\sigma_{y,1} + \sigma_{y,2} + \sigma_{y,3})}, \quad (4.36)$$

where $\sigma_{y,i}$ are the Pauli operators along the $\hat{y}_i$ axis for qubit i .

Now the permutation symmetry of both measurement X_W and feedback $U_F^W(\theta)$ operators induces a symmetric subspace $\mathcal{H}_W^S$ that is spanned by

$$\begin{aligned} |\phi_W^1\rangle &= |000\rangle \\ |\phi_W^2\rangle &= \frac{1}{\sqrt{3}}(|001\rangle + |010\rangle + |100\rangle) \\ |\phi_W^3\rangle &= \frac{1}{\sqrt{3}}(|011\rangle + |101\rangle + |110\rangle) \\ |\phi_W^4\rangle &= |111\rangle \end{aligned} \quad (4.37)$$

Thus, if our initial state is also symmetric to all permutations of the qubits, then the state will only evolve in the subspace $\mathcal{H}_W^S$ under the action of X_W and $U_F^W(\theta)$. To see this, consider an arbitrary operator P from the permutation group S_3 . This clearly commutes with both X_W and $U_F^W(\theta)$ in Eqs. (4.35) and (4.36). Then we can use Eq. (4.18) to obtain

$$\begin{aligned}
 P\rho_{t+dt}^c P &= P U_F^W \rho_{t+dt}^M U_F^{W\dagger} P = U_F^W P \rho_{t+dt}^M P U_F^{W\dagger} \\
 &= U_F^W (P \rho_t P + P d\rho P) U_F^{W\dagger} \\
 &= U_F^W (\rho_t + d\rho) U_F^{W\dagger} \\
 &= \rho_{t+dt}^c,
 \end{aligned} \tag{4.38}$$

where to arrive at the final equality, we have used the assumption that ρ_t is in the subspace $\mathcal{H}_W^S$ and the fact that X_W commutes with P . Provided that the initial state $\rho(0)$ is in $\mathcal{H}_W^S$, this assumption will hold at any time t . Now we wish to control the evolution so that the quantum state evolves from the initial state to the W state, which is $|\phi_W^2\rangle$. Noting that the basis vectors in Eq. (4.37) constitute non-degenerate eigenspaces of the measurement operator X_W in the symmetric subspace, we can then achieve the desired evolution by implementing a feedback protocol in which the control angle θ is determined by a cost function that specifically targets the state $|\phi_W^2\rangle$ within $\mathcal{H}_W^S$. The fidelity $\mathcal{F}_W = \langle \phi_W^2 | \rho_t | \psi_W^2 \rangle$ is a suitable cost function for this.

Experimentally it is generally easy to prepare qubits in a product state, so we shall assume without loss of generality that the initial state is $|000\rangle$. The advantage of keeping state evolution in the symmetric subspace is that this will help us rule out many unwanted states. This makes reaching the target state in the end more likely as well as finding the optimal angle numerically more efficient. We shall discuss this role of symmetry further when developing feedback protocols for the GHZ state, where it places important physical constraints on the measurement observable. To numerically test the PaQS protocol, Eqs. (4.22) and (4.26), we start from the product state $|000\rangle$. Before evolving this state, we first locally rotate it with $U_F^W(\theta)$ to a state with maximum fidelity. Following this initial rotation, we evolve under the local optimization protocol, computing the optimal feedback coefficient at every time step using Eq. (4.26).

Fig. 4.3 shows the resulting trajectory ensemble average (TEA) fidelity (red solid line) with respect to the W state, obtained by averaging over 1000 trajectories. The TEA results are seen to saturate at fidelity ~ 1 at a time of $\sim 1.5\mu\text{s}$. We note that in order to ensure that the fidelity saturates at 1, it is necessary to occasionally apply non-infinitesimal rotations when the second derivative test fails (Eq. (4.28)). Approximately 100 out of 10^4 trajectories, i.e., $\sim 1\%$ required at least one non-infinitesimal rotation angle (for a total of $\sim 10^5$ time steps). Without feedback, the fidelity remains constant at the value 0.44 that is obtained after an optimal rotation of $\theta = 2\arcsin(\frac{1}{\sqrt{3}})$, which we plot for comparison with a blue dashed line.

We now compare the TEA-based protocol, which implicitly assumes knowledge of the entire measurement record for ρ and requires real-time state estimation, with the more computationally and experimentally efficient ASLO protocol in which the feedback controls

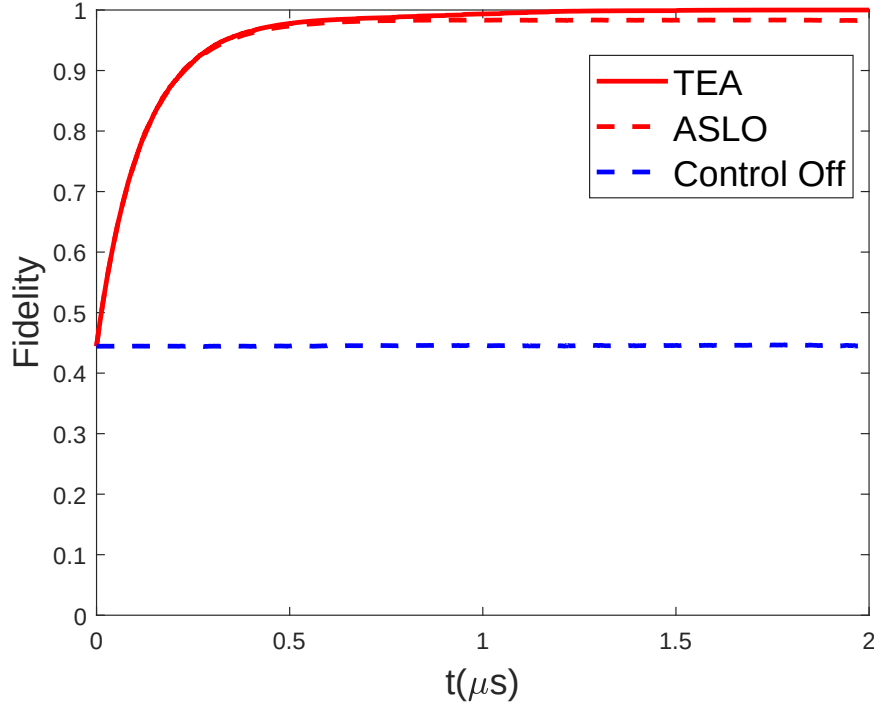


Figure 4.3: Fidelity of controlled state with respect to the W state, $\mathcal{F}_W$ starting from the initial product state $|000\rangle$. The solid red line shows results calculated with the trajectory ensemble approach (TEA) and the dashed red line shows results calculated with the ASLO approach. The dashed blue line shows the results obtained with measurement alone, *i.e.*, in absence of feedback, after an initial rotation to the state of maximum fidelity with the W state that is achievable by $U_F^W(\theta)$. Over 1000 trajectories were averaged for TEA, and over 10,000 trajectories for the ASLO and zero feedback calculations. The measurement strength is set to $k = 1$ MHz, corresponding to realistic values for superconducting qubits [17, 26], and the time step is set at $\Delta t \ll \frac{1}{k}$.

are optimized based on the average state $\bar{\rho}(t)$, Eq. (4.31). Fig. 4.4 shows the evolution of the optimal feedback angle coefficients $A_1(t)$ and $A_2(t)$ evaluated by the ASLO protocol. The coefficient $A_2(t)$ determines the average value of θ^* , while $A_1(t)$ determines its variance. It is evident that the average value is considerably smaller than the variance at all except the earliest times in the evolution.

Using the optimized ASLO feedback angles, Eq. (4.27), with ρ_t replaced by the unconditioned state $\bar{\rho}_t$ to control the averaged state dynamics gives rise to the fidelity shown in Fig. 4.3 by the dashed red line. The same initial condition is used here as for the TEA protocol, *i.e.*, the 3-qubit product state $|000\rangle$. It is apparent that while the W state is reached with high fidelity within a comparable time of approximately 600 ns, the ASLO protocol neverthe-

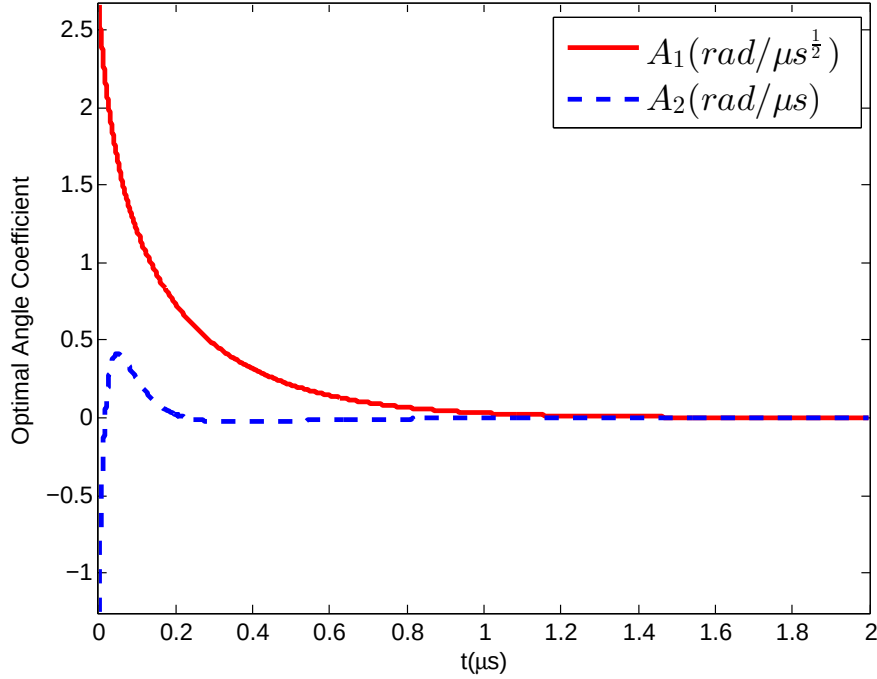


Figure 4.4: Optimal angle coefficients A_1 and A_2 , evaluated as a function of time for the ASLO protocol for generation of the $N = 3$ W state from the initial product state $|000\rangle$.

less saturates slightly below unity, at ~ 0.98 . The origin of this difference lies in the different sampling of the density matrix that is enabled by the TEA and ASLO approaches. Note that while the procedure for generating the coefficients $A_1(t)$ and $A_2(t)$ is identical, Eq. (4.31), the input density matrices are different, with the TEA approach sampling these from many trajectories while the ASLO approach takes just one averaged trajectory. However, these two approaches can be identical in the situation where feedback cancels measurement noise exactly, as is the case for two-qubit optimal control [60].

It should be noted that although the ASLO protocol does not achieve unit fidelity, one can nevertheless still produce unit-fidelity states by adding a final projective measurement. As the symmetry reduction has removed all degeneracy from the measurement operator X , the resulting measurement outcome now uniquely determines the state. Although the success probability under the ASLO protocol is less than one, it has been significantly enhanced by feedback. This argument applies to all protocols in which the symmetry reduction is made, so that one may thereby interpret the final ASLO fidelity as a success probability for perfect state preparation.

Generation of general Dicke states

We now consider the general Dicke states of Eq. (4.33) with $k > 1$ and arbitrary N . Several previous works have demonstrated deterministic Dicke state preparation with feedback, but these protocols were state-based and hence non-Markovian [51, 65, 66]. It is interesting to ask how well a Markovian protocol performs for the same task, particularly since the dynamical state estimation required for non-Markovian protocols becomes exponentially more challenging for larger systems. The most straightforward generalization of our $N = 3$ qubit protocol is to simply add more operator components to X_W and H_F

$$\begin{aligned} X_D^N &= \sigma_{z,1} + \sigma_{z,2} + \cdots + \sigma_{z,N} \\ H_F^N &= \sigma_{y,1} + \sigma_{y,2} + \cdots + \sigma_{y,N}. \end{aligned} \quad (4.39)$$

H_F^N is still local as required. X_D^N can be measured by placing each qubit in a separate cavity in an extension of the scheme demonstrated in [57], or by coupling many qubits dispersively to the same cavity. The latter method (without continuous feedback) has been applied to generate spin squeezing in cold neutral atoms coupled to an optical cavity [52].

Imposing the permutation group symmetry on N qubits allows the Dicke states to be represented within the symmetric subspace of dimension $N + 1$ [85]. Each state in this subspace is just a Dicke state with k excitations, which is also associated with a non-degenerate eigenspace of the observable X_D^N . In particular, the W state for N qubits, defined as

$$|W\rangle = \frac{1}{\sqrt{N}}(|10 \cdots 0\rangle_N + |01 \cdots 0\rangle_N + \cdots + |00 \cdots 1\rangle_N) \quad (4.40)$$

belongs to the eigenspace of X_D with eigenvalue $(N - 1) - 1 = N - 2$. This implies that the computational resources required to compute the feedback protocol scale only polynomially with the number of qubits, which is a huge improvement compared to the exponential scaling of the full Hilbert space dimension. We note that other than for the smallest case of $N = 2$ [26, 60], the average state $\bar{\rho}$ does not remain pure under ASLO feedback, so the stochastic terms do not cancel in Eq. (4.24) and the unaveraged dynamics (*i.e.* conditioned on the entire measurement record) depends on the entire measurement record.

The results of ASLO calculations with the locally optimal protocol of section 4.3 (see Eq. (4.24)) for Dicke states with variable excitation number k for up to $N = 48$ qubits, are shown in Fig. 4.5. We see that the final fidelities in Fig. 4.5 are all above 94% despite the fact that the system does not take a predictable path through Hilbert space and hence an ASLO feedback controller does not know the true state. This is exemplified in the two inserts. The lower inset in Fig. 4.5 shows that the maximum fidelity for the W state ($k = 1$) decreases with the number of qubits N . This is not surprising, given the increased dimensionality of the system. However it is remarkable that the fidelity decreases only to $\simeq 0.94$ for $N = 100$, indicating a strong robustness of the ASLO protocol for these target states. The upper inset shows how the maximum fidelity depends on the excitation number k for Dicke states of $N = 48$ qubits. Here the fidelity increases with k , indicating that the more symmetric

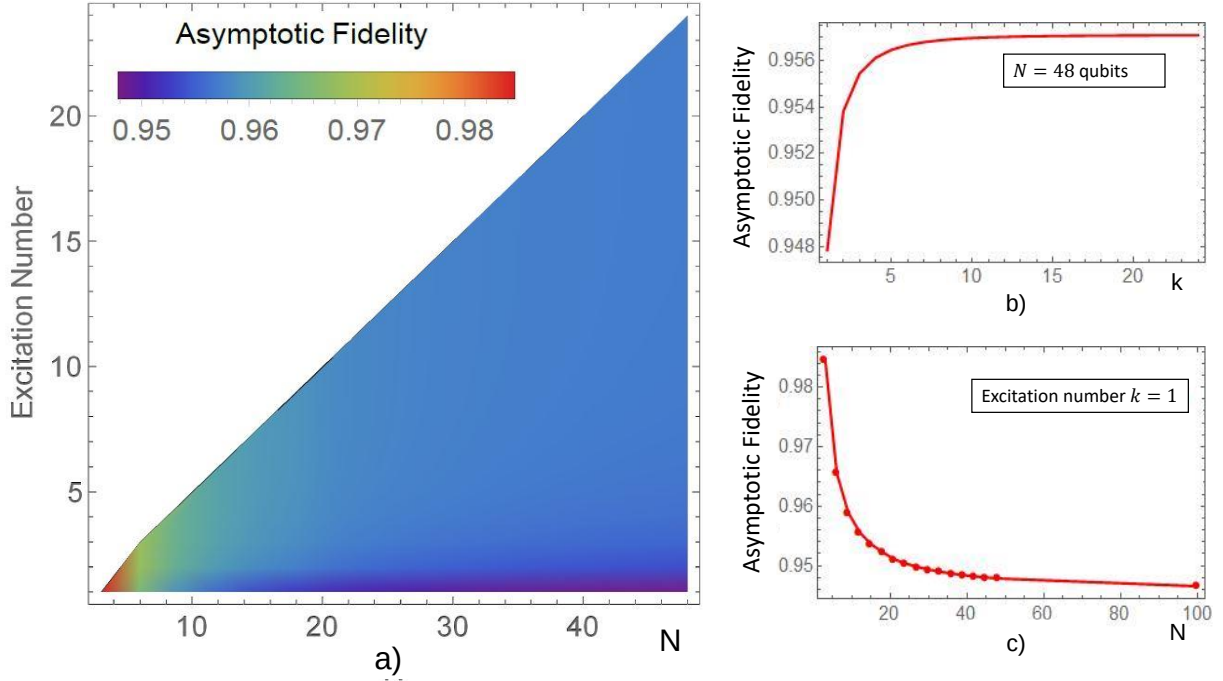


Figure 4.5: Fidelity with respect to generalized Dicke states $|N, k\rangle$ under the ASLO protocol with locally optimal feedback, equations (4.31) - (4.32), with U_F determined by the PaQS approach of section 4.3. Panel a) shows fidelity as a function of number of qubits N (from 3 to 48, in increments of 3) and excitation number k . The two right panels show cuts through this two-dimensional plot for fixed values of N and k . Panel b): k -dependence of fidelity for $N = 48$ qubits. Panel c): N -dependence for states with $k = 1$ excitation (W states), extending to $N = 100$ qubits.

states, i.e., states that have closer to half of the qubits in $|0\rangle$ and half in $|1\rangle$, are more efficiently prepared. Indeed we find empirically that $A_2 = 0$ for half-filled Dicke states with even N . The high performance thus appears to be due in part to the highly symmetric nature of the problem. The symmetry reduction to $|N, n\rangle$ has removed all degeneracy from the measurement operator X_D^N , so that the measurement outcome now uniquely determines the state.

4.5 Generation of Greenberger–Horne–Zeilinger states

The Greenberger–Horne–Zeilinger (GHZ) state

$$|\text{GHZ}\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle) \quad (4.41)$$

is the three-qubit state with both maximal multi-particle entanglement and maximal fragility of three-way entanglement [71, 86]. The state

$$|\text{GHZ}\rangle_N = \frac{1}{\sqrt{2}}(|00 \cdots 0\rangle_N + |11 \cdots 1\rangle_N). \quad (4.42)$$

is the extension to N qubits. We shall use the notation $|\text{GHZ}\rangle$ to refer specifically to the $N = 3$ state.

Generation of GHZ states using the above method poses new challenges. As noted above, measurement operators that are linear combinations of single-body observables like Z are the easiest to implement in remote and weakly interacting settings. The GHZ state is an eigenvector of such operators, with the general form

$$X_G = (a + b)Z_1 - aZ_2 - bZ_3. \quad (4.43)$$

However, such an operator does not satisfy the full permutation symmetry of the target state. Furthermore, the target state $|\text{GHZ}\rangle$ is now one of a degenerate set of eigenstates of X_G . As we saw for the W state above, it is important that the target state constitute a non-degenerate eigenspace within a subspace that is determined by a symmetry imposed by the measurement operator. In order to achieve this, we first need to impose the correct symmetry requirement on the measurement operator.

We consider here two forms of measurement operators for the GHZ state. Both of these possess the full S_3 permutation symmetry and the additional bit-flip symmetry that characterizes the GHZ state. The first approach uses a symmetrized sum of two-qubit operators that imposes the additional bit flip symmetry of the GHZ state. The second approach achieves the same symmetries by the action of multiple measurement operators with single-qubit components.

Two-qubit measurement operators

To prepare the GHZ state we can use the following symmetric two-qubit measurement observable

$$X_G^S = \sigma_{z,1}\sigma_{z,2} + \sigma_{z,2}\sigma_{z,3} + \sigma_{z,3}\sigma_{z,1}. \quad (4.44)$$

X_G^S is the lowest order n -qubit observable that possesses both full permutation symmetry and bit-flip symmetry. Together, these two symmetries induce a symmetric subspace $\mathcal{H}_G^S$ of

the full Hilbert space. The basis vectors of $\mathcal{H}_G^S$ are expressed in the computational basis as

$$|\phi_G^1\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle) \quad (4.45)$$

$$|\phi_G^2\rangle = \frac{1}{\sqrt{6}}(|011\rangle + |101\rangle + |110\rangle + |100\rangle + |010\rangle + |001\rangle). \quad (4.46)$$

$|\phi_G^i\rangle, i = 1, 2$ are eigenstates of X_G^S with eigenvalues $e_1 = +3$ and $e_2 = -1$. Within this subspace, the GHZ state $|\phi_G^1\rangle$ is then a non-degenerate eigenstate of the measurement operator X_G^S . Thus, restricting the dynamics governed by Eq. (4.18) to this subspace and employing a cost function that is biased towards the state $|\phi_G^1\rangle$, will allow continuous measurement of X_G^S together with the feedback controls to extract the GHZ state. Similarly to the procedure for the Dicke states in the previous section, we achieve this by using the fidelity with the target state as the cost function, i.e., $\mathcal{F}_G = \langle\phi_G^1|\rho_t|\psi_G^1\rangle$. But in contrast to the imposition of only permutation symmetry on the feedback operator in section 4.4, we need to now impose the full permutation and bit-flip symmetry of X_G^S on the feedback rotation operator U_F , to ensure that the action of this does not take the state out of $\mathcal{H}_G^S$. Therefore we define U_F here by a single rotation around the x axis, which is consistent with the bit-flip symmetry, and set the rotation angles to be equal for all three qubits, to be consistent with the permutation symmetry. This yields the rotation operator

$$U_F^G(\theta) = e^{-i\frac{\theta}{2}(\sigma_{x,1} + \sigma_{x,2} + \sigma_{x,3})}. \quad (4.47)$$

If the initial state is within $\mathcal{H}_G^S$, then the combined action of $U_F^G(\theta)$ and X_G^S will ensure that the subsequent evolution always remains in $\mathcal{H}_G^S$.

Similar to the analysis above for the W state, we then determine the locally optimal angle θ^* by maximizing the fidelity at each time step

$$\mathcal{F}_G(t + dt) = \langle\text{GHZ}|\rho_{t+dt}^c|\text{GHZ}\rangle, \quad (4.48)$$

where ρ_{t+dt}^c is again the output state given by Eq. (4.18). When the symmetric feedback operator (4.47) is used with the GHZ state target, we can obtain an exact solution for $\partial\mathcal{F}/\partial\theta$ and $\partial^2\mathcal{F}/\partial\theta^2$ and hence find the explicit values for the optimal angle θ^* . Full details of the solution are given in appendix 4.8, where we show that the optimal angle is always either 0 or $\pi/2$, with the specific choice given by the sign of a state-dependent term. Thus there is no need to use a PaQS protocol based on infinitesimal rotation angles in this situation. This also means that we cannot use the SME Eq. (4.24), which assumes infinitesimal rotations, to simulate the dynamics and instead we must use the full POVM equation, Eq. (4.18).

The GHZ state also allows an interesting alternative optimization approach, deriving from the fact that after the dynamics are constrained to the symmetric subspace, the dimension of the Hilbert space is reduced from eight to two. Consequently, under these constrained dynamics the three-qubit problem can be mapped to a single qubit problem. This mapping is described explicitly in appendix 4.5, where it is shown that this allows an alternative

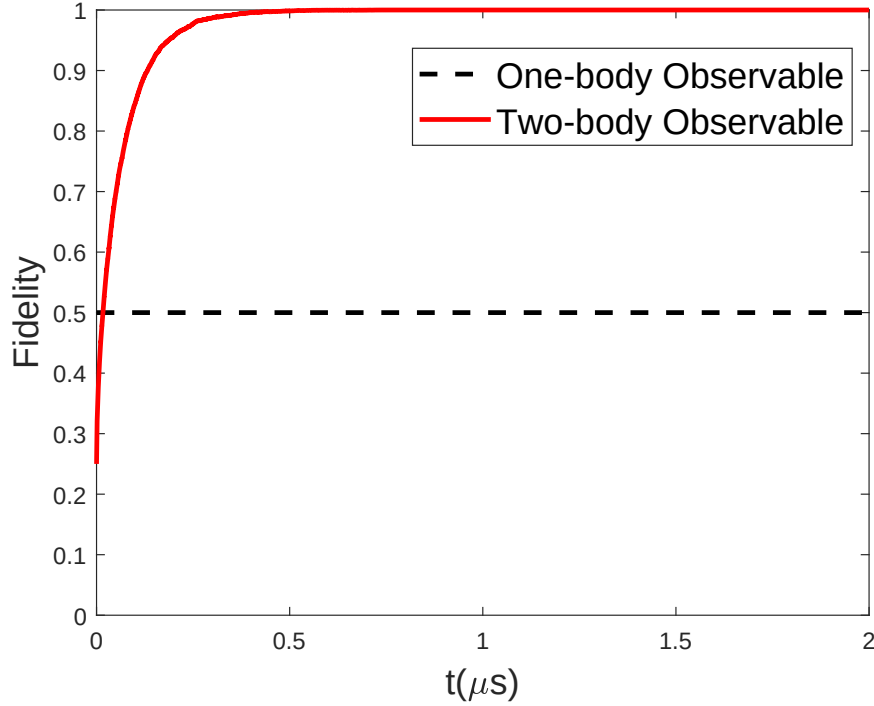


Figure 4.6: Fidelity $\mathcal{F}_G(t)$ of the controlled state with respect to GHZ state as a function of time, obtained from the locally optimal ASLO approach. The red line is the result obtained with the symmetric two-body measurement operator X_G^S . The blue line shows the result obtained with the non-symmetric one-body measurement operator X_G . All simulations were run with measurement strength $k = 1\text{MHz}$ and time step $\Delta t \ll \frac{1}{k}$, starting from the complete superposition state (see text).

approach to determination of the optimal angle that also results in an optimal angle of either 0 or $\frac{\pi}{2}$.

Fig. 4.6 shows the time dependence of the fidelity of formation of the GHZ state obtained using the locally optimal protocol within the ASLO approach, with the symmetrized measurement X_G^S , and using as initial condition the complete superposition state

$$\frac{1}{(\sqrt{2})^3} (|0\rangle + |1\rangle) \otimes (|0\rangle + |1\rangle) \otimes (|0\rangle + |1\rangle). \quad (4.49)$$

We see that in this situation the fidelity under the ASLO protocol asymptotically approaches unity.

It is instructive to compare the performance of this protocol with a symmetrized two-body measurement operator, to that obtained from feedback control based on measurements of an observable not respecting this permutation symmetry. We consider here the one-body

observable of Eq. (4.43) with $a = b = 1$, which is permutation symmetric only with respect to exchange of the last two qubits. The target state is then contained in a degenerate eigenspace with eigenvalue zero that is spanned by $\frac{1}{\sqrt{2}}(|000\rangle \pm |111\rangle)$. While we can still impose the symmetry condition on the feedback rotation operator, the measurement is now unable to distinguish the target GHZ state from another state ($\frac{1}{\sqrt{2}}(|000\rangle - |111\rangle)$) within the degenerate eigenspace. We therefore expect that any feedback protocol based on this measurement will be unlikely to reach unit fidelity. In fact, the results achieved with this protocol, shown as the blue line in Fig. 4.6, are of very low quality.

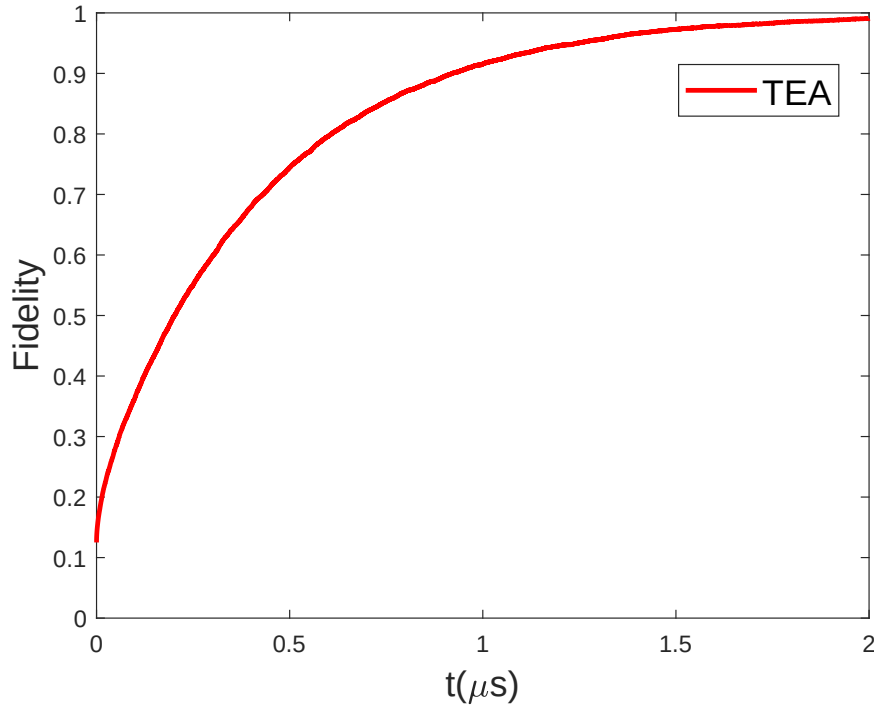


Figure 4.7: Fidelity of four-qubit GHZ state, computed with the TEA protocol using the complete superposition as initial state.

We can also generalize our protocol to the GHZ state of N qubits, which is defined as

$$|\text{GHZ}\rangle_N = \frac{1}{\sqrt{2}}(|00 \cdots 0\rangle_N + |11 \cdots 1\rangle_N). \quad (4.50)$$

The observable we use is still the two-body symmetrized observable

$$X_G^S = \sum_{i < j} \sigma_{z,i} \sigma_{z,j}. \quad (4.51)$$

After imposing both the permutation and bit flip symmetry, the corresponding symmetric subspace $\mathcal{H}_G^S$ now has dimension $\frac{N+1}{2}$ when N is odd, and dimension $\frac{N}{2} + 1$ when N is even.

Explicitly, the symmetrized basis set is given by

$$|\phi_G^1\rangle = \frac{1}{\sqrt{2}}(|00\cdots 0\rangle_N + |11\cdots 1\rangle_N) \quad (4.52)$$

$$|\phi_G^2\rangle = \frac{1}{\sqrt{2N}} \sum_{P \in S_N} P(1 + \Pi) |10\cdots 0\rangle_N \quad (4.53)$$

$\vdots$

$$|\phi_G^m\rangle = \frac{1}{\sqrt{2\binom{N}{m}}} \sum_{P \in S_N} P(1 + \Pi) |\underbrace{11\cdots 1}_m 0\cdots 0\rangle_N \quad (4.54)$$

where $\Pi = X_1 X_2 \cdots X_N$ is the N -qubit bit flip operator. The index m indicates how many spins are pointing downward and runs from 0 to either $\frac{N-1}{2}$ (N odd) or $\frac{N}{2}$ (N even). Note that every basis vector $|\phi_G^m\rangle$ is a non-degenerate eigenstate of the observable X_G^S . In particular, the GHZ state is just $|\phi_G^1\rangle$, with eigenvalue $\binom{N}{2}$. Following the same procedure as for $N = 3$, we choose the feedback rotation operator to be of the form $U_F(\theta) = \bigotimes_{i=1}^N e^{-i\frac{\theta}{2}X_i}$ and start

from the symmetric initial state $|\psi(0)\rangle = \bigotimes_{i=1}^N \frac{|0\rangle + |1\rangle}{\sqrt{2}}$. The locally optimal rotation angles are calculated using the same local expansion approach employed for $N = 3$ qubits (see appendix 4.8).

Results obtained with the trajectory ensemble approach for generation of the $N = 4$ GHZ state using this locally optimal GHZ protocol are shown in Fig. 4.7. Undertaking these non-Markovian locally optimal calculations for the GHZ state is significantly more expensive than the corresponding calculations for the W state in section 4.4. While possible in principle, this again motivates the use of the ASLO approach for larger N values, as was done for the generalized Dicke states for large N in section 4.4. In the following subsection we address this, together with the goal of finding a feedback protocol for GHZ states that also uses only single qubit measurement and feedback Hamiltonian operators. To do this, we have to first generalize our PaQS protocol to allow multiple operators acting on each qubit.

Multiple measurement and multiple feedback Hamiltonian operators

The extension of PaQS to handle multiple simultaneous measurement operators and feedback Hamiltonians is straightforward so long as the feedback Hamiltonians commute. In what follows, we extend PaQS to cover these cases. The result also turns out to hold for non-commuting Hamiltonians under certain conditions [82]. We then show that with this generalization it is possible satisfy the permutation symmetry of the GHZ state, while also restricting the measurements to the experimentally more accessible linear combinations of single-body observations. We start by noting that the GHZ state is also an eigenstate of the

following operators

$$M_{ij} = \frac{\sigma_{z,i} - \sigma_{z,j}}{2}, \quad i, j = 1 \cdots N, \quad i \neq j. \quad (4.55)$$

These resemble the half-parity measurement operators of Table 4.1, with the crucial difference that M_{ij} now give identical outcomes on the states $|00\rangle$, $|11\rangle$ rather than on the states $|01\rangle$, $|10\rangle$ of each i, j pair of qubits. The action of this set of measurement operators imposes both permutation and bit flip symmetry on the corresponding set of terms in the SME, Eq. (4.15).

To find the locally optimal protocol, we need to generalize the expression for θ_{opt} to handle multiple simultaneous measurements M_i and multiple feedback operations H_j . In general there is no preferred pairing between the measurement and feedback operations in these groups, so that the i th measurement outcome may affect how we apply the j th feedback Hamiltonian. This forces us to rederive the feedback master equation with a more general feedback unitary,

$$\begin{aligned} U &= \exp \left(-i \sum_j H_j \theta_j \right) \\ &= I - i \sum_{ij} A_{ij} H_j dW_i - \left[i \sum_j B_j H_j + \frac{1}{2} \sum_{ijk} A_{ij} A_{ik} H_j H_k \right] dt \\ \theta_j &= B_j dt + \sum_i A_{ij} dW_i, \end{aligned} \quad (4.56)$$

where A_{ij} and B_i are analogous to A_1 and A_2 respectively, and our goal is now to find the locally optimal values of all $\{A_{ij}\}$ and $\{B_i\}$. The presence of cross terms in Eq. (4.56) modifies the resulting feedback master equation

$$\begin{aligned} \rho(t+dt) &= \rho - i \sum_j B_j [H_j, \rho] dt + \sum_i \left[2k \mathcal{D}[M_i] \rho dt + \sqrt{\eta_i} 2k \mathcal{H}[M_i] \rho dW_i \right. \\ &\quad - i \sum_j A_{ij} [H_j, \rho] dW_i - i \sum_j \sqrt{\eta_i} A_{ij} [H_j, \sqrt{2k} (M_i \rho + \rho M_i^\dagger)] dt \\ &\quad \left. + \sum_{jk} A_{ij} A_{ik} [H_k \rho H_j - \frac{1}{2} (H_j H_k \rho + \rho H_j H_k)] dt \right] \end{aligned} \quad (4.57)$$

In contrast to Eq. (4.24), the last term is not in Lindblad form due to the presence of cross terms, though it may be cast into Lindblad form by defining the modified feedback operators $\tilde{H}_i = \sum_j A_{ij} H_j$. For simplicity, we assume that the control Hamiltonians commute pairwise, although the end result is essentially unmodified if one relaxes this assumption [82].³ The

³In general, one must have that the H_i form a Lie algebra *i.e.*, that the vector space formed by H_i is closed under commutation.

locally optimal feedback coefficients must satisfy $\langle \psi_T | [H_\alpha, d\rho] | \psi_T \rangle = 0$, this time for all α . Dealing first with the dW terms, we find

$$\sum_i \left[\sum_j -iA_{ij} \underbrace{\langle \psi_T | [H_\alpha, [H_j, \rho]] | \psi_T \rangle}_{c_{j\alpha}} + \underbrace{\sqrt{\eta_i} \langle \psi_T | [H_\alpha, \sqrt{2k}(M_i \rho + \rho M_i^\dagger)] | \psi_T \rangle}_{a_{i\alpha}} \right] dW_i = 0 \quad (4.58)$$

The solution is evident if we rewrite the expression in matrix form with the help of a and c , which are state-dependent

$$-iAc + a = 0 \implies A = -iac^{-1}. \quad (4.59)$$

In general, one should use the Moore-Penrose pseudoinverse to handle the case in which c is not invertible[82]. With a solution for A in hand, we can solve for the B coefficients by collecting $\mathcal{O}(dt)$ terms from $\langle \psi_T | [H_\alpha, d\rho] | \psi_T \rangle = 0$

$$\begin{aligned} b_\alpha \equiv \langle \psi_T | [H_\alpha, \sum_i \left[\mathcal{D}[M_i] \rho - i \sum_j \sqrt{\eta_i} A_{ij} [H_j, \sqrt{2k}(M_i \rho + \rho M_i^\dagger)] \right. \\ \left. + \sum_{jk} A_{ij} A_{ik} [H_k \rho H_j - \frac{1}{2}(H_j H_k \rho + \rho H_j H_k)] \right] | \psi_T \rangle \\ \left. - i \sum_i B_i c_{i\alpha} + b_\alpha = 0 \implies \vec{B} = -i\vec{b}c^{-1} \end{aligned} \quad (4.60)$$

treating $\vec{b}$ and $\vec{B}$ as row vectors. Eqs. (4.59) and (4.60) generalize Eq. (4.26).

Now we have a locally optimal solution for general measurements and feedback. To generate the GHZ states, we will set the measurement operators M_i to be M_{ij} as given in Eq. (4.55), which ensures that the measurement term in the SME Eq. (4.57) satisfies both permutation and bit-flip symmetries. Then we still need to choose a set of feedback operators. For simplicity, we limit ourselves to local σ_y rotations. c can fail to be invertible if there are too many H_i s, and we have learned the importance of symmetry, so some care is required in choosing them. One obvious choice of Hamiltonian basis is $H_i = \sigma_{y,i}$, one for each qubit. However, this basis does not enforce the $|11\dots 1\rangle \leftrightarrow |00\dots 0\rangle$ bit flip symmetry of the GHZ state and the M_{ij} operators. A better choice would be a set orthogonal to $\sum_i \sigma_{y,i}$, such as

$$H_i = \frac{\sigma_{y,i} - \sigma_{y,i+1}}{2}, \quad i = 1 \dots N-1 \quad (4.61)$$

where N is the number of qubits. Although this basis appears to break permutation symmetry, any Hamiltonian of the form

$$H_{ij} = \frac{\sigma_{y,i} - \sigma_{y,j}}{2} \quad (4.62)$$

can be written as a linear combination of the set of H_i . Thus the locally optimal solution can retain permutation symmetry.

The time dependence of GHZ state generation under ASLO feedback is plotted in Fig. 4.8. The 2-qubit case reproduces the half-parity measurement of Refs. [57] and [26] under the transformation $\sigma_{z,2} \leftrightarrow -\sigma_{z,2}$, while the remaining results $N = 3$ through $N = 8$ are novel. Although we use Eq. (4.61) as our feedback Hamiltonians, the locally optimal protocol takes linear combinations of the form Eq. (4.62), so that each H_{ij} is paired with its corresponding M_{ij} . We speed up the numerics by taking this observation into account, so that many terms in Eq. (4.57) may be set to zero. Perhaps most interestingly, the $N = 3$ case reaches exactly unit fidelity, and the purity of the average state remains 1 throughout the feedback process, as in the example at the beginning of this section. One can derive an analytic solution in this case.

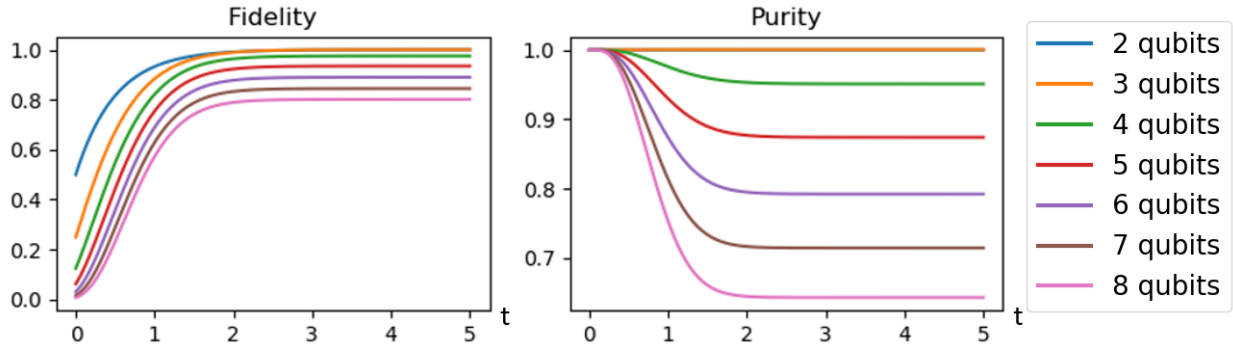


Figure 4.8: Fidelity and purity of the controlled state under the multiple measurement and multiple feedback operator protocol for generation of N qubit GHZ states, in the ASLO implementation. For $N = 2$ and $N = 3$ the purity remains exactly 1 for all time.

Feedback protocols for GHZ states based on non-fidelity measures

As noted in section 4.3 above, it is also possible to construct feedback protocols based on cost functions providing a direct measure of entanglement, rather using as cost function the fidelity with a specific target state. While, as seen above, the latter choice simplifies many calculations, it does however require that a single specific target state is singled out. This ignores the possibility that one might be able to do better by targeting a different target state that is nevertheless locally equivalent to the desired target state. In appendix 4.8 we present a method for deriving locally optimal protocols based on entanglement measures, and apply it to directly optimize the three-tangle, a measure of entanglement for three qubits that achieves its maximal unit value for a GHZ state and all locally equivalent states.

4.6 Conditions for deterministic evolution under measurement-based feedback control

This work has revealed a number of intriguing examples in which feedback perfectly eliminates the randomness of the measurement process. These cases are highlighted with asterisks in table 4.1. All of these specific protocols can be shown to be globally optimal for the task at hand, except for the 3-qubit GHZ protocol that we conjecture to also be optimal. These examples also yield analytic solutions in all cases, further indicating that they belong to a natural class of protocols in continuous measurement. Ref. [60] provides detailed analysis of this for the half- and full-parity Bell state measurements in table 4.1.

It is straightforward to introduce a strict necessary condition for whether or not deterministic evolution is possible, given fixed measurement and feedback operators. This condition also provides a simple tool for finding solutions to the general proportional feedback master equation, supposing that we apply proportional feedback of the form $U = \exp(-iH(Pdr + cdt))$, where H is the feedback operator, P and c are time-dependent scalars, and $dr = \sqrt{\eta}\langle M + M^\dagger \rangle + dW$ is the incremental output signal for the measurement operator M .⁴ Applying U to the stochastic Schrödinger equation yields a pure-state equation of motion for $|\psi\rangle$ valid for $\eta = 1$

$$|\psi(t + dt)\rangle = \left[I + (M - m - iPH)dW - \frac{1}{2}(M^\dagger M - 2Mm + m^2 + P^2 H^2)dt - iH(P(M + m) + c)dt \right] |\psi(t)\rangle \quad (4.63)$$

where $m \equiv \langle M + M^\dagger \rangle/2$. We should allow for the solution to be deterministic up to a global phase, so we make the substitution

$$H \rightarrow H + V_0. \quad (4.64)$$

For a deterministic solution, the dW terms of the stochastic Schrödinger equation must cancel, *i.e.*,

$$[M - m - iPH - iPV_0]|\psi\rangle = 0. \quad (4.65)$$

This is a nonlinear constraint on $|\psi\rangle$, since the expectation value m depends on the state. It can only be satisfied if $|\psi\rangle$ is a right eigenvector of $M - iPH$. To see that this necessary condition is also sufficient, compute the corresponding eigenvalue λ by taking an expectation value

$$\langle \psi | [M - iPH] | \psi \rangle = \lambda = \langle M \rangle - iP\langle H \rangle \quad (4.66)$$

⁴Note that we have absorbed the measurement strength into M here.

and then compute the left-hand side of (4.65)

$$\begin{aligned} [M - m - iPH - iPV_0]|\psi\rangle &= \left[\langle M \rangle - iP\langle H \rangle - \frac{\langle M + M^\dagger \rangle}{2} - iPV_0 \right] |\psi\rangle \\ &= \left[\frac{\langle M^\dagger - M \rangle}{2} - iP\langle H \rangle - iPV_0 \right] |\psi\rangle \end{aligned} \quad (4.67)$$

As $\langle M^\dagger - M \rangle$ is purely imaginary, we can always find a V_0 such that the above expression is zero. Thus feedback can exactly cancel the measurement-induced noise on $|\psi\rangle$ if and only if $|\psi(t)\rangle$ is a right eigenvector of $M - iPH$. As P is a free parameter, we can check if this condition is possible for any value of it.

With the stochastic terms cancelled, the next step is to see if the deterministic part of the equations of motion can maintain the above condition for a finite time interval. Given the above result, we can search for such solutions within each eigenspace of $M - iPH$ separately (the eigenspaces will typically evolve continuously as a function of P). Let $\{|v_i(P(t))\rangle\}$ be a set of orthonormal eigenvectors associated with a chosen eigenspace. To obtain a deterministic solution, we must have

$$|\psi(t)\rangle = \sum_i c_i(t) |v_i\rangle. \quad (4.68)$$

Now consider the deterministic part of the equation of motion. Using the above expansion, we have

$$\begin{aligned} \frac{d|\psi\rangle}{dt} &= \sum_i \dot{c}_i(t) |v_i\rangle + c_i(t) \frac{dP}{dt} \frac{d|v_i\rangle}{dP} \\ &= \left[\frac{1}{2} (M^\dagger M - 2Mm + m^2 + P^2 H^2) dt - iH(P(M + m) + c) dt \right] |\psi(t)\rangle. \end{aligned} \quad (4.69)$$

As $m = \text{Re}(\lambda)$ and $\lambda(t)$ is fixed by working within a fixed eigenspace, the above equation of motion is linear in $|\psi\rangle$, which significantly simplifies the finding of a solution. One can use the above procedure to identify candidate solutions, which appear to be plentiful and relatively unexplored. For example, the multiple measurement, multiple feedback operator GHZ state protocol in section 4.5 yields an analytic solution for $N = 2$ and $N = 3$ under the above procedure.

4.7 Discussion and Conclusions

We have presented a general, analytic construction of locally optimal measurement-based feedback protocols, termed PaQS, in which the feedback unitary has two independently varying components, one having dependence on the measurement outcome and the other having dependence only on the quantum state. We showed that the resulting protocol can

be modified to generate an average state locally optimal (ASLO), or Markovian protocol that can be efficiently implemented.

In this work we demonstrated the effectiveness of PaQS feedback by applying it to the generation of W, general Dicke and GHZ states for $N = 3 - 100$ qubits, highlighting the utility of symmetry constraints. The ASLO versions of these protocols reach target state fidelities above 94%, despite the significant constraints imposed by Markovianity. This shows that a simple Markovian-type feedback strategy exists to prepare any of these maximally spin-squeezed states deterministically.

While we have focused on application to entanglement generalization, the full scope of Eqs. (4.57)-(4.60) extends to virtually any application of measurement-based feedback. The ability of these equations to reproduce all existing feedback protocols that we have tried them on (see table 4.1) and the ease with which they generate new protocols suggests that there is much more to be gained from them. One immediate extension of this work would be to study the behavior under non-ideal measurement conditions, i.e., for $\eta < 1$.

Several interesting challenges remain regarding systematization of this PaQS feedback approach. Although we have so far only considered specific target states, one may ask more generally which classes of quantum states may be generated under constraints on the measurement and feedback operators. Similar questions have been successfully answered in the context of Markovian stabilization [87, 88] and it would be natural to apply and extend these questions to non-Markovian situations. Another important direction for future applied work would be to classify what states can be prepared deterministically (under a suitable definition) using local feedback and measurement operators that are linear combinations of local observables. Alternatively, one could ask the converse question of how to efficiently determine suitable measurement and feedback operators given a desired target state. Finally, questions of state stabilization and robustness leave important directions for future work. Recent work has shown that global exponential stabilization of two-level systems is possible using a quantum state and proportional control method similar to that introduced here [89]. Whether PaQS yields global or exponential stabilization for higher dimensional quantum systems is an important question for further work.

4.8 Appendix

Second Derivative Test for Local Optimal Control

In this appendix we discuss the situation when the infinitesimal solution Eq. (4.22) with Eq. (4.26) fails to pass the second derivative condition, Eq. (4.28). Using the same expansion

as in section 4.3, we have

$$\begin{aligned}
 \langle \psi_T | [H, [H, \rho^c]] | \psi_T \rangle &= \langle \psi_T | [H, [H, \rho]] | \psi_T \rangle \\
 &+ \langle \psi_T | [H, [H, \sqrt{\eta}(Y\rho + \rho Y^\dagger)]] - iA_1[H, [H, [H, \rho]]] | \psi_T \rangle dW \\
 &+ \langle \psi_T | [H, [H, \mathcal{D}[Y]\rho]] + A_1^2[H, [H, \mathcal{D}[H]\rho]] | \psi_T \rangle dt \\
 &- i\langle \psi_T | [H, [H, [H, \sqrt{\eta}A_1(Y\rho + \rho Y^\dagger) + A_2\rho]] | \psi_T \rangle dt > 0
 \end{aligned} \tag{4.70}$$

We can see that due to the stochastic nature of dW , there will always be some non-zero probability that this condition is violated. Specifically, when

$$\begin{aligned}
 |dW| &> \left| \frac{1}{\langle \psi_T | [H, [H, \sqrt{\eta}(Y\rho + \rho Y^\dagger)]] - iA_1[H, [H, [H, \rho]]] | \psi_T \rangle} \times \right. \\
 &\left. \langle \psi_T | [H, [H, \rho]] | \psi_T \rangle + \right. \\
 &\left. \langle \psi_T | [H, [H, \mathcal{D}[Y]\rho]] + A_1^2[H, [H, \mathcal{D}[H]\rho]] - i[H, [H, [H, \sqrt{\eta}A_1(Y\rho + \rho Y^\dagger) + A_2\rho]]] | \psi_T \rangle dt \right|
 \end{aligned} \tag{4.71}$$

the infinitesimal solution will give a local minimum instead of a maximum. As long as the term $\langle \psi_T | [H, [H, \rho]] | \psi_T \rangle$ is not zero (note that this term is always non-negative if the fidelity of ρ is assumed to be locally maximal), in the limit that dt goes to zero, the violation probability will be small for sufficiently small dt .

Optimal Angle for GHZ State with Two-body Observable

Local Expansion Method We can also use the calculus method of local expansion employed in section 4.3 in order to find the optimal angle for the GHZ state. Using Eq. (4.21), we have an explicit expression of $\mathcal{G}$ proportional to $\sin(2\theta)$ provided the input state is an extremal state. So the optimal angle would be either 0 or $\frac{\pi}{2}$. However, an extremal state can be either a local maximum or minimum. To determine which of these will be generated at time $t + dt$, we look again at the second derivative of the fidelity function (Eq. (4.28)):

$$\begin{aligned}
 \frac{d^2 \mathcal{F}}{d\theta^2} &= \cos(2\theta) \left[\frac{3 - 6\rho_{11} - 2\sqrt{3}\rho_{12}}{2} \right. \\
 &+ 4\sqrt{2}kdW(6\rho_{11}^2 - \sqrt{3}\rho_{12} + 2\rho_{11}(-3 + \sqrt{3}\rho_{12})) \\
 &\left. + 16\sqrt{3}k\rho_{12}dt \right].
 \end{aligned} \tag{4.72}$$

It is then evident that the optimal choices of angle to ensure that the sign of the second derivative is negative are $\theta^* = 0$ or $\theta^* = \pi/2$, depending on the sign of the state-dependent term in parentheses.

Mapping of $N = 3$ GHZ state to a single qubit state Since the symmetric subspace obtained by imposing the S_3 permutation symmetry on three qubits, has dimension two, we

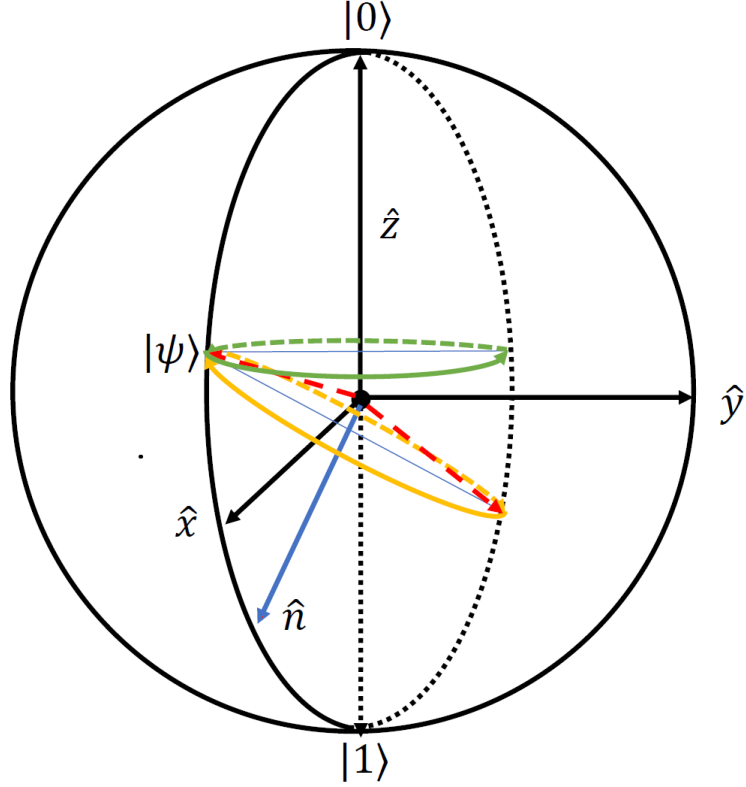


Figure 4.9: Bloch sphere illustrating the geometric meaning of the locally optimal control protocol for the GHZ state. The blue vector $\hat{n} = (\frac{\sqrt{3}}{2}, 0, \frac{1}{2})$ is the rotation axis. Initially, our state is located exactly along this direction. Note that the measurement will always keep the state in the xoz plane, where o denotes the origin. Suppose that after a measurement, the state is one of the two red dashed vectors. The feedback operator will rotate this vector along the yellow circle around the $\hat{n}$ axis. To maximize the fidelity with respect to the $|\tilde{0}\rangle$ state, we should always rotate it to the *upper* position labeled by $|\psi\rangle$. This means that the rotation angle θ should be 0 or $\frac{\pi}{2}$, depending on whether the state is above or below $\hat{n}$ after the measurement. The green circle at constant z is drawn for reference as a visual aid.

can map the three-qubit problem into an effective single-qubit representation. The geometric meaning of the resulting control protocol will become clear below.

In order to ensure that the state stays in the symmetric subspace, we have to apply a symmetric rotation around the qubit x axes:

$$U_F = e^{-i\frac{\theta}{2}(\sigma_{x,1} + \sigma_{x,2} + \sigma_{x,3})}. \quad (4.73)$$

Then the symmetric subspace within which the dynamics takes place is spanned by

$$|\tilde{0}\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle) \quad (4.74)$$

$$|\tilde{1}\rangle = \frac{1}{\sqrt{6}}(|100\rangle + |010\rangle + |001\rangle + |011\rangle + |101\rangle + |110\rangle) \quad (4.75)$$

The $|\tilde{0}\rangle$ basis is just the GHZ state we are trying to generate. Let us see how encoded operations are realized in this subspace. First, the rotation axis $\Sigma \equiv \sigma_{x,1} + \sigma_{x,2} + \sigma_{x,3}$ is equivalent to

$$\begin{aligned} \tilde{\Sigma} &= \begin{pmatrix} \langle \tilde{0} | \Sigma | \tilde{0} \rangle & \langle \tilde{0} | \Sigma | \tilde{1} \rangle \\ \langle \tilde{1} | \Sigma | \tilde{0} \rangle & \langle \tilde{1} | \Sigma | \tilde{1} \rangle \end{pmatrix} = \begin{pmatrix} 0 & \sqrt{3} \\ \sqrt{3} & 2 \end{pmatrix} \\ &= \tilde{I} + \sqrt{3}\tilde{X} - \tilde{Z} \end{aligned} \quad (4.76)$$

where $\tilde{I}$, $\tilde{X}$, and $\tilde{Z}$ are the 2 by 2 identity, Pauli X and Z matrix. We can drop the identity here since it only gives a global phase. Then we have

$$\tilde{\Sigma} = \sqrt{3}\tilde{X} - \tilde{Z} = 2\hat{n} \cdot \vec{\sigma}, \quad (4.77)$$

where $\hat{n} = (\frac{\sqrt{3}}{2}, 0, \frac{1}{2})$ is a unit vector lying in the $x - z$ plane which gives us the rotation axis, and $\vec{\sigma}$ is just the vector of the effective Pauli matrices. So in the effective single qubit picture, the rotation operator $\tilde{\sigma}$ and the corresponding rotation angle $\tilde{\theta}$ are given by

$$\tilde{\sigma} = \frac{\sqrt{3}}{2}\tilde{X} - \frac{1}{2}\tilde{Z} \quad (4.78)$$

$$\tilde{\theta} = 2\theta \quad (4.79)$$

Note that the rotation angle in the effective single qubit space is equal to twice that in the original space.

Now let us look at the measurement process, which is given by the encoded operator

$$\tilde{X}_{obs} \equiv \begin{pmatrix} \langle \tilde{0} | X_{obs} | \tilde{0} \rangle & \langle \tilde{0} | X_{obs} | \tilde{1} \rangle \\ \langle \tilde{1} | X_{obs} | \tilde{0} \rangle & \langle \tilde{1} | X_{obs} | \tilde{1} \rangle \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix} = \tilde{I} + 2\tilde{Z} \quad (4.80)$$

The evolution equation is given by Eq. (4.15) of the main text. In our case, the Hamiltonian is zero, and for the terms deriving from the measurement, we have

$$\mathcal{D}[\tilde{I} + 2\tilde{Z}] = \mathcal{D}[2\tilde{Z}] \quad (4.81)$$

$$\mathcal{H}[\tilde{I} + 2\tilde{Z}] = \mathcal{H}[2\tilde{Z}]. \quad (4.82)$$

Dropping the identity in the measured observable, we find that the measurement in the effective single qubit subspace is a measurement along the encoded z axis, with a four-fold increase in the measurement strength, *i.e.*,

$$\tilde{X}_{obs} = Z \quad (4.83)$$

$$\tilde{k} = 4k. \quad (4.84)$$

where $\tilde{k}$ is the effective measurement strength in the single qubit space.

It is easy to show that the initial state $|\psi\rangle_0 = (\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle))^{\otimes 3}$ becomes

$$|\tilde{\psi}\rangle_0 = \frac{1}{2}|\tilde{0}\rangle + \frac{\sqrt{3}}{2}|\tilde{1}\rangle \quad (4.85)$$

in the effective single qubit space, or in Bloch vector form,

$$\tilde{\rho}_0 = |\tilde{\psi}\rangle_0\langle\tilde{\psi}|_0 \sim (\frac{\sqrt{3}}{2}, 0, -\frac{1}{2}) \quad (4.86)$$

The fidelity with respect to the GHZ state is then given in terms of this Bloch vector by

$$f \equiv \langle\tilde{0}|\rho|\tilde{0}\rangle = \frac{1+\tilde{z}}{2}, \quad (4.87)$$

where $\tilde{z}$ is the z component of the Bloch vector in the effective single qubit space. Clearly the rotation angle has to be equal to either 0 or $\frac{\pi}{2}$, in order to ensure that this fidelity is optimal at each time step (See Fig. 4.9).

Tangle-based protocols

Here we present an alternative locally optimal protocol for generation of a three-qubit GHZ state that is based on optimization of the three-tangle τ , which provides a measure for tri-partite entanglement [70, 90]. This protocol has the advantage that it not only differentiates between the two distinct types of tri-partite entanglement of three-qubit states, but is also invariant under local rotations of the state. Under this measure the GHZ state reaches the upper bound, with value $\tau = 1$, while all two-particle entanglements are zero [91, 92]. In the following we shall maximize the tangle under measurements of the symmetrized two-body operator X_G^S , Eq. (4.44).

We note first that starting from a pure state, the conditioned state after the weak measurement will still be pure. So if we avoid averaging the state along the evolution over measurement outcomes, the state will remain pure at all times. This allows us to use the pure state definition of the three-tangle [70]

$$\tau = \tau(ABC) = \tau(A|BC) - \tau(A|B) - \tau(B|C), \quad (4.88)$$

where the quantities on the right hand side (referred to as “two-tangles”) are given by squares of the relevant concurrences [70]. This considerably simplifies the determination of the feedback angles, since computing the tangle for a mixed state can be very difficult, involving determination of a convex roof extension [93].

Since the tangle is invariant under our feedback unitary operations, the maximization procedure used for the fidelity cost function in the main text of the chapter does not work here. Instead, we determine the optimal angle by maximizing the expected increase in tangle

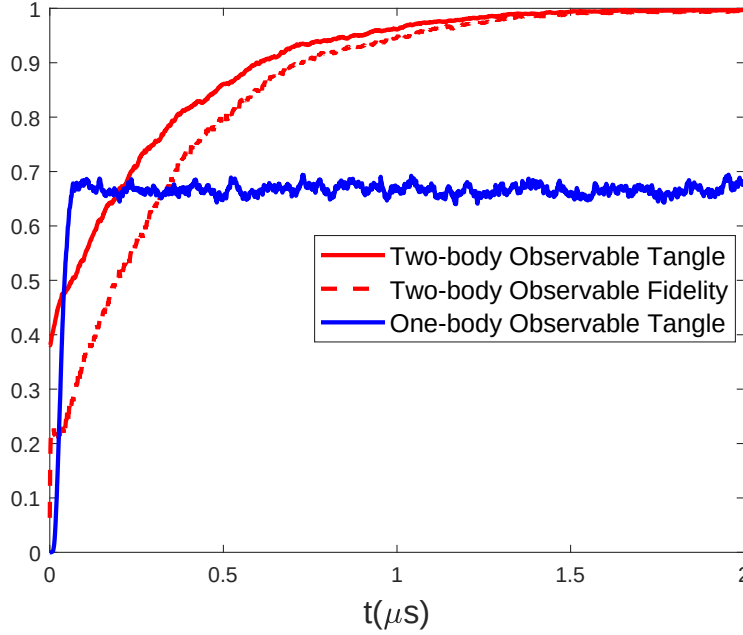


Figure 4.10: Time dependence of the three-tangle τ and the fidelity $\mathcal{F}_G$ for three qubits with the TEA protocol using feedback based on direct optimization of the three-tangle cost function τ (Eq. (4.88)). The initial condition here is the complete superposition state, Eq. (4.49). Red solid curve: three-tangle for state measured with two-body observable, averaged over 1000 trajectories. Red dashed curve: fidelity with respect to GHZ state, same state evolution as the red solid curve. Blue solid curve : three-tangle of state measured with one-body observable, averaged over 100 trajectories. For all trajectories the measurement strength was $k = 1$ and time step $dt \ll \frac{1}{k}$.

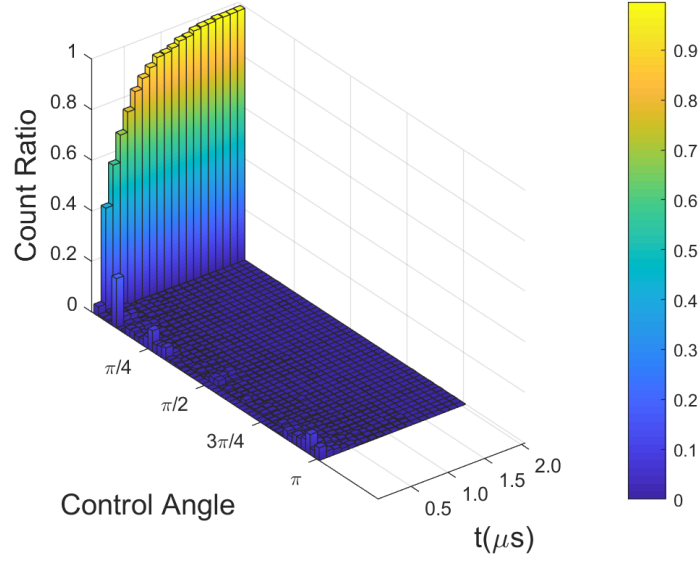
after measurement. The feedback angle at each infinitesimal time step is now computed as follows. At time t , the (pure) input state $|\psi\rangle_t$ is rotated using the feedback control operator $U_F^G(t)$ (Eq. (4.47))

$$|\psi\rangle_t^c = U_F^G(t)|\psi\rangle_t, \quad (4.89)$$

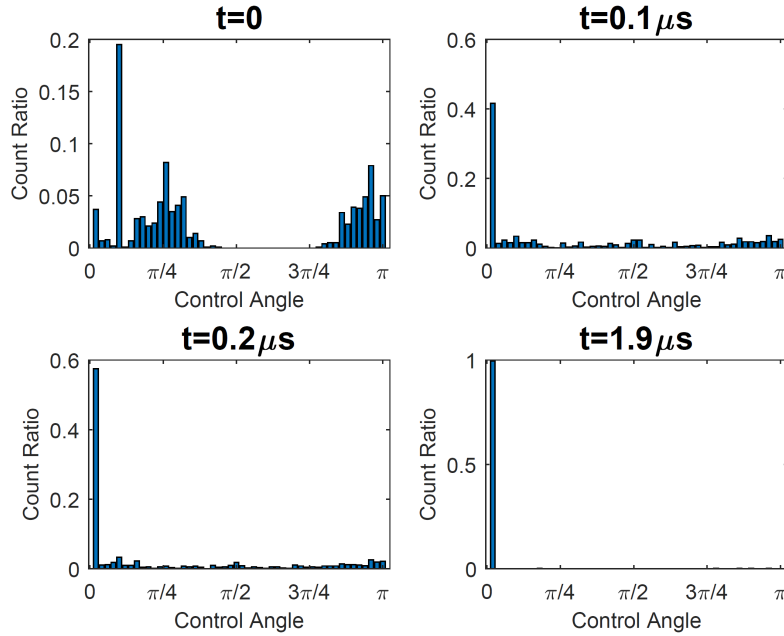
with the rotation angle parameter determined as described below. We then make a weak measurement on the controlled state:

$$\begin{aligned} |\psi\rangle_{t+dt}^c &= \frac{\Omega_{dV}|\psi\rangle_t^c}{\|\Omega_{dV}|\psi\rangle_t^c\|} \\ &= \frac{\Omega_{dV}U_F^G(t)|\psi\rangle_t}{\|\Omega_{dV}U_F^G(t)|\psi\rangle_t\|}. \end{aligned} \quad (4.90)$$

Note that we are now controlling the state *before* the measurement instead of *after* measurement: we choose to do this because a local rotation on a pure state will not change the



(a) Histograms of control angle distribution as a function of time. 20 snapshots of the angle distribution are taken between the initial and final times, with the control angle space $[0, \pi]$ divided into 50 bins in each snapshot. The count ratio for the 0 angle bin gradually increases to 1, indicating approach to a steady state.



(b) Four snapshots of the angular distribution, at times $t = 0, 0.1 \mu s, 0.2 \mu s$, and $1.9 \mu s$. The distribution rapidly becomes dominant in the first bin, and subsequent rotations of the state are very small.

Figure 4.11: Histograms of control angle distributions resulting from optimization of the three-angle at different times. The total trajectory count is 1000. The y-axis shows the normalized count in each bin.

value of the tangle for the state, so the tangle is not affected by the control.

Now the choice of rotation should not be determined by a particular measurement outcome that occurs *after* imposition of the control. Therefore in order to obtain the optimal rotation angle while avoiding issues of causality, we may simply average the tangle over all possible measurement outcomes and choose the control rotation as the value maximizing this average, i.e.,

$$U_F^*(t) = \operatorname{argmax}_{U_F(t)} \int dV \tau(|\psi\rangle_{t+dt}^c). \quad (4.91)$$

This requires sampling values of rotation angle and evaluating the average over measurement outcomes for each case. The state is then evolved forward by acting with the measurement after the optimal rotation, yielding the evolution described by Eq. (4.91) with $U_F(t)$ replaced by $U_F^*(t)$.

Fig. 4.10 shows that when this tangle-based protocol is implemented using the two-body measurement observable X_G^S , both the value of the three-tangle τ and the corresponding fidelity $\mathcal{F}_G$ appear to asymptotically reach a value of one, although on a slower timescale than the corresponding fidelity under the fidelity based approach (compare with Fig. 4.6). We also demonstrate how the control angle distributions vary with time for this situation in Fig. 4.11. In contrast, when the tangle-based optimization is used with the non-symmetrized one-body observable X_G for measurement, a significantly lower value of the tangle is obtained (not shown), with an asymptotic value of approximately 0.7 being reached. It is thus evident again that a protocol based on symmetrized two-body observable measurements significantly outperforms a protocol based on measurement with a non-fully symmetrized observable.

The three-tangle τ is one of five non-trivial polynomial invariants that characterize normalized three-qubit states [91, 92]. Our work suggests that optimization of multiple invariants might be useful for construction of feedback protocols to systematically generate arbitrarily entangled three-qubit states. For three-qubit states, τ achieves its maximal value for the GHZ state and all other invariants automatically reach the boundary value. In this case, optimizing the tangle alone then guarantees that the other invariants reach the correct values for the state. This is not the case for other states in general. One alternative choice of cost function in more general situations is to use the sum of the squared differences between the invariants of the current state and the target state, which can act as a measure of the distance between the two states.

Chapter 5

Continuous error correction I: quantum memory

5.1 Introduction

Quantum error correction (QEC) is an essential component of quantum information processing. The need to either avoid or correct errors on quantum states due to imperfect quantum operations or decohering interactions with the environment places stringent requirements on realization of the promise of quantum computation and quantum simulations. Various tools have been developed to mitigate the effect of such errors, including encoding into decoherence free subspaces or subsystems [94, 95], addition of penalty Hamiltonians [96–98], dynamical decoupling methods [99–101] and other applications of pulse sequences [102], as well as the use of quantum error correcting codes (QECC) that delocalize the errors over multiple physical qubits, combined with error recovery operations [103–106]. The latter provides a powerful approach to systematically correct errors that can also be made fault tolerant [107].

The canonical operation mode for quantum error correction codes [21, 108–110] employ projective measurements and discrete recovery operations to provide reduction of errors that are treated as discrete events occurring at a specified rate. The formalism of QEC has been developed to provide firm guarantees of protection in terms of reduced scaling of the logical error rate for an encoded state. However, in practice, few measurements can be described as projective, and are instead better described as finite strength weak measurements that are characterized by a gradual collapse of the measured system wavefunction [7, 69, 111–121]. A continuous quantum error correction code, i.e., a CQEC, is based on the continuous quantum measurement of the error syndrome operators of the conventional QEC code. Another reason to utilize CQEC instead of traditional discrete QEC lies in the fact that with continuous measurement we only collect partial information but in a much shorter time period and thus the error can be detected and corrected faster than that with discrete QEC [122]. For reasons which will be clear later, this is especially more important in the presence of a

time-depending Hamiltonian. Previous theoretical work on such continuous quantum error correction has been devoted primarily to analysis of the continuous operation performance of stabilizer [123–135] and subsystem [136, 137] QEC codes for quantum memory, where the Hamiltonian of the encoding physical qubits is disregarded in the analysis. In contrast, in this work we focus on protecting the coherent evolution of an encoded qubit system evolving under a time-dependent Hamiltonian, against environmental decoherence. This problem is particularly important for the development of quantum error correction for a broad range of quantum information applications employing continuously varying Hamiltonians. These include quantum annealing and adiabatic quantum computation [138], and quantum simulation [139].

A major challenge for application of either discrete or continuous QEC to protect coherent evolution of an encoded qubit system is that perfect identification and correction of errors (in the example studied here, these will be bit-flip errors) does not imply absence of logical errors [140]. We can understand this difficulty by thinking of the action of errors on the Hamiltonian instead of on the quantum state—a perspective somewhat similar to the Heisenberg picture. In this picture, an error causes the Hamiltonian to effectively change from $H(t)$ to $EH(t)E$, where E is the operator associated to the error that occurred and is assumed to be a single-qubit Pauli operator. Subsequent coherent evolution is due to the new Hamiltonian $EH(t)E$, until the moment when the error that occurred is detected and corrected. During this period of error diagnosis and correction, logical errors will accrue if the original Hamiltonian does not commute with the error operators, i.e., if $H(t) \neq EH(t)E$. Since Hamiltonians that commute with all error operators are difficult to implement [140], this problem has constituted a major stumbling block for the development of quantum error correction for quantum annealing and for analog quantum simulation in general. This is precisely the situation that we address in this work.

We consider here the continuous operation of a quantum code that is designed to protect the coherent evolution of the encoded qubit system. As a specific example we consider a single logical qubit and encode it with the three-qubit bit flip code [21], which is a stabilizer code [110] with two commuting stabilizer operators that constitute the measurement operators. We will divide the discussion of CQEC into two chapters. In the current chapter, we will focus on the case of quantum memory while in the next chapter, we will consider the case in the presence of a time-dependent Hamiltonian. In Section 5.2, we briefly introduce what the discrete error correction is about. Section 5.3 describes the continuous operation of this code. In Section 5.4 we propose and analyze an error detection protocol based on time-averaging (filtering) of the bare readout signals from simultaneous continuous monitoring of the error syndrome operators, together with a double error thresholding scheme that is applied to the filtered readout signals in order to explicitly diagnose errors. Unlike previous schemes [124, 127], partial errors are not acted on—the error diagnosis is acted on only when occurrence of a complete, i.e., discrete, error has been diagnosed with high probability. Filtering is necessary in the protocol to reduce (but not eliminate) the amount of noise in the filtered readout signals, while double error thresholding is essential to reduce the probability of mis-identification of single bit-flip errors that affect several readout signals

at the same time [137].

As for the next chapter, in Section 6.1, we give a short introduction on quantum annealing, which is the primary context of the time-dependent Hamiltonian we considered. In Section 6.3 we develop an effective open quantum system model to describe the Hamiltonian-driven evolution of the encoded qubit system in the presence of both bit-flip errors and CQEC. In Section 6.4 we use this model to optimize the performance of the proposed continuous QEC protocol for operation under quantum annealing. In this case, the performance of our protocol depends on the relative strength of three parameters; namely, the error rate γ , the Hamiltonian strength parameter Ω_0 , and the strength Γ_m of the continuous measurements. We find that our CQEC protocol yields a significant reduction of the final logical state infidelity when the measurements are sufficiently strong relative to the Hamiltonian. We show that further improvements are possible by using modified error correction operators. To demonstrate the capability of our proposed CQEC approach, we present detailed results for one logical qubit and then show that the strategy can be readily generalized to quantum simulation with many encoded qubits. We show explicitly that a high level of protection can be obtained for quantum annealing of two logical qubits. In addition, we analyse the performance of the proposed continuous QEC protocol for quantum memory operation, in the absence of a Hamiltonian. We minimize the logical error rate and find that its optimal value is slightly larger than the logical error rate obtained from using the linear variant [134] of the optimal Wonham filter [131, 132] for error diagnosis. The advantage of our continuous QEC protocol is that it can be simpler to implement. We also show that the resulting optimized double thresholding error diagnosis scheme is very effectively combined with discrete recovery operations to obtain the reduced scaling of the logical error rate that is necessary for a valid quantum error correcting code. Finally, in Section 6.5, we provide a summary and discussion, with additional analysis of the time-to-solution metric for estimating the cost of applying CQEC to quantum annealing problems. We conclude with a prognosis for future applications and useful directions for further work. The main reference of these two chapters is our paper [9].

5.2 Discrete error correction in a nutshell

In this section we give a brief introduction to discrete error correction and stabilizer codes. More detailed discussion of quantum error correction can be found in [141, 142]. In real world quantum computation, qubits are subject to noise. The evolution of the qubits can then be described by quantum noise channels [23]. These quantum noise channels are completely positive trace-preserving (CPTP) maps which have the general form

$$\rho \rightarrow \sum_i E_i \rho E_i^\dagger. \quad (5.1)$$

The subscript i indicates the i th noise channel and E_i is the error operator, which can be represented as linear combinations of tensor products of Pauli operators acting on different

qubits. The error operators must satisfy the completeness condition

$$\sum_i E_i^\dagger E_i = I. \quad (5.2)$$

Three-qubit code

We will begin the introduction with a gentle example, the three-qubit repetition code. Considering a computational single qubit subject to pure bit-flip errors X , we would like to encode this logical qubit by three physical qubits to protect it from the error. This means that instead of using the single qubit basis $|0\rangle$ and $|1\rangle$ as the logical basis, we would rather choose the logical basis as

$$|0_L\rangle = |000\rangle \quad (5.3)$$

$$|1_L\rangle = |111\rangle. \quad (5.4)$$

The subspace spanned by these two logical basis states is called the *codespace*. Any vector inside this codespace is called a *codeword*. In this simple example, there is only one single error type - the bitflip error. Assume for now only a single qubit error has happened. In order to correct the error, we only need to locate the error position. We can obtain this information by measuring two observables, $Z_1 Z_2$ and $Z_2 Z_3$. The good thing about this measurement is that it does not change the state if it is in the codespace, since any codeword is a **eigenvector** of the two observables, with eigenvalue labels $(+1, +1)$. If an error happened on the first qubit, then the state will be in the *error subspace* that is spanned by $|100\rangle$ and $|011\rangle$. This in turn is an eigenspace of the observables with the eigenvalue labels $(-1, +1)$. Similarly, an error on the second qubit results in a transition to the subspace with labels $(-1, -1)$ and an error on the third qubit results in transition to the subspace with labels $(+1, -1)$. Notice that each subspace has a different pair of labels. A measurement on the state (made via an ancilla) can then give us this information about the labels and we can correct the state accordingly by applying a bitflip at the location of the error. These labels are called *error syndromes* and the operators corresponding to the two observables are called *stabilizer operators*.

General properties of quantum error correcting codes

Now we can discuss error correction in a more general setting. We mentioned that a general error can be a superposition of tensor products of Pauli operators. But let us assume first that errors are pure tensor products of Pauli operators. Consider two errors E_a and E_b of this type that act on two different basis codewords $|\psi_i\rangle$ and $|\psi_j\rangle$, respectively. If we want to correct these two errors, then a necessary condition is that the two resulting states must be distinguishable. This implies

$$\langle \psi_i | E_a^\dagger E_b | \psi_j \rangle = 0 \text{ for } i \neq j. \quad (5.5)$$

This condition alone is not enough to characterize the correctable errors. We also need to guarantee that the measurement doesn't disturb the codeword itself. This gives us the following condition:

$$\langle \psi_i | E_a^\dagger E_b | \psi_i \rangle \text{ does not depend on } i, \quad (5.6)$$

since otherwise we are able to learn what the original state is from the outcome of the measurement. Combining these two equations, we have

$$\langle \psi_i | E_a^\dagger E_b | \psi_j \rangle = C_{ab} \delta_{ij}. \quad (5.7)$$

In fact this is also a sufficient condition for correctable errors [143, 144].

If the two errors E_a and E_b are correctable through some measurements, their linear combinations must also be correctable. To see this, say an error type of $aE_a + bE_b$ has happened. Then we can just do the same set of measurements. This will project the system to $E_a|\psi\rangle$ with probability $\langle \psi | E_a^\dagger (aE_a + bE_b) | \psi \rangle^2 = |a|^2$ and to $E_b|\psi\rangle$ with probability $\langle \psi | E_b^\dagger (aE_a + bE_b) | \psi \rangle^2 = |b|^2$. No matter which case happens, we can always detect the error type and correct the state accordingly. This fact means we can focus on correcting errors which are pure tensor products of Pauli operators and not worry about a separate protocol for linear combinations of these.

The *weight* of an error (considering now only tensor products of Pauli operators) is the number of qubits on which this error acts non-trivially. The *distance* of the code is the smallest weight of $E_a^\dagger E_b$ such that Eq. (5.7) does not hold. To correct weight t errors, we at least need a distance $2t+1$ code. We use the notation $[[n, k, d]]$ to denote a distance d quantum code which encodes k logical qubits by n physical qubits. For example, the three-qubit code is a $[[3, 1, 3]]$ code. With an encoding into the computational basis, $|0_L\rangle = |000\rangle, |1_L\rangle = |111\rangle$, this corrects only bitflip errors and not phase errors. Phase errors can be corrected by encoding into the Hadamard basis $|0_L\rangle = |+++ \rangle, |1_L\rangle = |--+\rangle$ with $|\pm\rangle = 1/\sqrt{2}(|0\rangle + |\pm\rangle)$. To be able to correct all type of single-qubit errors, we can concatenate these two codes to form the $[[9, 1, 3]]$ nine-qubit Shor code [108].

Stabilizer codes

Consider n physical qubits and the Pauli group $\mathcal{P}_n \equiv \{\pm 1, \pm i\} \otimes \{I, X, Y, Z\}^{\otimes n}$. A subgroup $\mathcal{S} \in \mathcal{P}_n$ which consists of 2^{n-k} mutually commuting operators without the negative identity $-I$ is called the *stabilizer group*. Ruling out the negative identity from the stabilizer group make all the operators in the group Hermitian with eigenvalues ± 1 . These stabilizer operators can be generated from a smaller subset of itself, of which the element is called the *stabilizer generator*. The size of the smallest generator set is $n - k$ [141] and we denote it by $\langle S \rangle$. The eigenspace $C(\mathcal{S})$ of the stabilizer group with simultaneous eigenvalue $+1$ for all stabilizer operators is the codespace of the stabilizer group. It can be shown that this codespace has dimension k [141].

Consider a set $\mathcal{E}$ of errors which contains all linear combinations of some basis errors.¹ This is a **correctable** error set if for any two different basis errors E_a and E_b , we have

$$\{E_a^\dagger E_b, S_i\} = 0 \text{ for } a \neq b \quad (5.8)$$

for some element S_i in the stabilizer group. Note that there can be multiple S_i satisfying this condition.

Proof: For any basis codeword $|\psi_i\rangle$ and $|\psi_j\rangle$ in the codespace, we have

$$\langle\psi_i|E_a^\dagger E_b|\psi_j\rangle = \langle\psi_i|S_i E_a^\dagger E_b|\psi_j\rangle = -\langle\psi_i|E_a^\dagger E_b S_i|\psi_j\rangle = -\langle\psi_i|E_a^\dagger E_b|\psi_j\rangle = 0 \quad (5.9)$$

where we have used the anticommutation relation Eq. (5.8) to obtain the second equality and made use of the simultaneous +1 eigenvalue for all codewords in the third equality. This satisfies Eq. (5.7). $\square$

By default, the identity I is always in the correctable error set $\mathcal{E}$. Let's illustrate how to correct the errors in this set. We measure all the operators in the generator set $\langle S \rangle$. The mutually commuting property means that we can do a simultaneous measurement on all of them. Assuming that the initial state $|\psi\rangle$ is in the codespace and no error has occurred, then this measurement does not disturb the state, since $|\psi\rangle$ is the simultaneous eigenstate of the stabilizers. The measurement produces all +1 eigenvalues. Notice any two operators from the Pauli group $\mathcal{P}_n$ either commute or anticommute. If an error E_a has happened, the state becomes $E_a|\psi\rangle$. An element M_i from the generator set $\langle S \rangle$ either commutes or anticommutes with E_a . The measurement of M_i on the state $E_a|\psi\rangle$ produces +1 eigenvalue if $[M_i, E_a] = 0$ and -1 eigenvalue if $\{M_i, E_a\} = 0$. The collection of ± 1 eigenvalues obtained by measuring the generator set $\langle S \rangle$ is called the *error syndrome*.

Condition (5.8) guarantees that every error in the set $\mathcal{E}$ has a unique error syndrome. To see this, consider two different errors E_a and E_b in $\mathcal{E}$. Then there exists M_i from the generator set which satisfies Eq. (5.8), i.e. $\{E_a^\dagger E_b, M_i\} = 0$, since otherwise, all stabilizer operators have to commute with the product $E_a^\dagger E_b$, which contradicts the condition (5.8). Assuming $[E_b, M_i] = 0$,² we have

$$E_a^\dagger E_b M_i = E_a^\dagger M_i E_b = -M_i E_a^\dagger E_b \quad (5.10)$$

The second equality implies that E_a anticommutes with M_i since E_b is invertible and $E_a^\dagger = E_a$ or $-E_a$. Thus, measuring M_i gives different eigenvalues for the two errors. Then by examining the error syndrome, we know exactly which error happened and can correct it by applying the Hermitian conjugate of that error to the state.

To make the discussion complete, let's consider other possible situations for the errors. Denote the product of two errors E_a and E_b as $E = E_a^\dagger E_b$. Condition (5.8) describes the situation when E anticommutes with some stabilizer operators. It is also possible that E commutes with all the operators in the stabilizer group, i.e., E belongs to the centralizer

¹We call an error from the Pauli group as a basis error.

²The argument would be similar if E_b anticommutes with M_i .

group $\text{Cen}(\mathcal{S})$ of $\mathcal{S}$.³ Note $\mathcal{S}$ itself is in $\text{Cen}(\mathcal{S})$. When $E \in \mathcal{S}$, following a similar argument leading to Eq. (5.10), we can show the two errors E_a and E_b must have the same error syndrome. These two errors are indeed effectively the same error. Once we detect the corresponding error syndrome, we can always correct the state back by applying $E_a^\dagger$ no matter which error actually happened, since $E_a^\dagger E_b$ is the identity operator when we restrict to the codespace. E_a and E_b are called the *degenerate errors* in this case. For the situation when $E \in \text{Cen}(\mathcal{S}) - \mathcal{S}$, any codeword transformed by E must remain as a codeword. Also there must be some codeword $|\psi\rangle$ such that $E|\psi\rangle \neq |\psi\rangle$. This implies Eq. (5.7) cannot be satisfied and thus E_a, E_b are not correctable errors [141].

5.3 Continuous operation of the three-qubit bit flip code.

In contrast to the discrete operation of the three-qubit bit flip code [21], in the continuous operation, the error syndrome operators (stabilizer generators),

$$S_1 = Z_{12} = Z_1 Z_2 \quad \text{and} \quad S_2 = Z_{23} = Z_2 Z_3, \quad (5.11)$$

are continuously measured at the same time. In Eq. (5.11), Z_1 represents the Pauli z operator that acts on the first physical qubit; that is, $Z_1 |q_1 q_2 q_3\rangle = (-1)^{q_1} |q_1 q_2 q_3\rangle$, where the set of states $|q_1 q_2 q_3\rangle$ with $q_1, q_2, q_3 = \{0, 1\}$ defines the computational basis. Similar definitions hold for the Pauli z operators Z_2 and Z_3 . The corresponding normalized readout signals are given by ($k = 1, 2$)

$$I_k(t) = \text{Tr}[S_k \rho(t)] + \sqrt{\tau_k} \xi_k(t), \quad (5.12)$$

where $\rho(t)$ is the 8×8 density matrix of the three physical qubits and τ_k is the so-called “measurement time” to distinguish between the ± 1 eigenvalues of the stabilizer generator S_k with a signal-to-noise ratio (SNR) of 1 [29]. Note that the detector readout signals $I_k(t)$ are given by the sum of the “signal part” $\text{Tr}[S_k \rho(t)]$ and the noise part $\xi_k(t)$, which has a vanishing mean. In the Markovian approximation, the noises $\xi_k(t)$ are assumed Gaussian and white with a two-time correlation function:

$$\langle \xi_k(t) \xi_{k'}(t') \rangle = \delta_{kk'} \delta(t - t'), \quad (5.13)$$

where $\langle \cdot \rangle$ denotes average over an ensemble of noise realizations.

The evolution of the three-qubit quantum state $\rho(t)$ in the absence of environmental decoherence is described by (in Itô interpretation [114])

$$\begin{aligned} \dot{\rho}(t) = & -i[H(t), \rho] + \\ & \sum_{k=1,2} \left[\frac{\Gamma_k}{2} (S_k \rho S_k - \rho) + \frac{\xi_k}{\sqrt{\tau_k}} \left(\frac{S_k \rho + \rho S_k}{2} - \rho \text{Tr}[S_k \rho] \right) \right]. \end{aligned} \quad (5.14)$$

³The set whose elements commute with all the elements in $\mathcal{S}$ is called the centralizer of $\mathcal{S}$ and the centralizer is a group.

The first line of Eq. (5.14) describes the coherent evolution of the three physical qubits due to a time dependent Hamiltonian and in this chapter it will be set to 0 as for quantum memory.

The second line of Eq. (5.14) describes the measurement-induced quantum back-action on the three-qubit quantum state that is due to simultaneous continuous measurement of the stabilizer generators Z_{12} and Z_{23} . Each measurement channel is characterized by the measurement time parameter τ_k and the measurement-induced ensemble dephasing rate Γ_k , which are related via the quantum efficiency η_k as follows $\tau_k = 1/(2\Gamma_k\eta_k)$ [29]. For ideal detectors, the quantum efficiency is unity, while for nonideal detectors the quantum efficiency is less than one. For simplicity of notation, we shall assume below that both detectors have identical parameters:

$$\Gamma_k = \Gamma_m, \quad \tau_k = \tau_m, \quad \text{and} \quad \eta_k = \eta \quad (k = 1, 2). \quad (5.15)$$

(This assumption can be readily removed and the analysis continued with different parameters for each detector.)

Encoding with the three-qubit bit flip code effectively divides the full eight-dimensional Hilbert space of the three physical qubits into four two-dimensional subspaces, where the stabilizer generators Z_{12} and Z_{23} have definite ± 1 values. As usual, the two-dimensional subspace where both stabilizer generators have values $+1$ is referred to as the code space, denoted as $\mathcal{Q}_0$, while the two-dimensional subspaces where (Z_{12}, Z_{23}) have values $(-1, +1)$, $(-1, -1)$ and $(+1, -1)$ are referred to as the error subspaces, denoted as $\mathcal{Q}_1$, $\mathcal{Q}_2$ and $\mathcal{Q}_3$, respectively. The code space is spanned by the zero and one logical states, which are expressed in the computational basis as

$$|0_L\rangle = |000\rangle \quad \text{and} \quad |1_L\rangle = |111\rangle, \quad (5.16)$$

respectively.

In the presence of bit-flip errors, the (mixed) three-qubit state $\rho(t)$ evolves according to the evolution equation that results from adding to the right-hand side of Eq. (5.14) the following decoherence terms

$$\dot{\rho}_{\text{decoh}}(t) = \sum_{q=1,2,3} \gamma_q [X_q \rho X_q - \rho], \quad (5.17)$$

where γ_q denotes the bit-flip error rate of the q th physical qubit. Thus in the presence of bit-flip errors, the full three-qubit state evolves as

$$\begin{aligned} \dot{\rho}(t) = & -i[H(t), \rho] + \\ & \sum_{k=1,2} \left[\frac{\Gamma_k}{2} (S_k \rho S_k - \rho) + \frac{\xi_k}{\sqrt{\tau_k}} \left(\frac{S_k \rho + \rho S_k}{2} - \rho \text{Tr}[S_k \rho] \right) \right] + \\ & \sum_{q=1,2,3} \gamma_q [X_q \rho X_q - \rho]. \end{aligned} \quad (5.18)$$

Our analysis of logical errors presented below is based on the jump/no-jump method [21] for bit-flip errors. In this method, gradual decoherence due to the terms (5.17) is described as the average effect of bit-flip errors X_1 , X_2 or X_3 that occur at random times, as follows. At the infinitesimal time interval $(t, t + \delta t)$, a bit-flip error X_q occurs with probability $\delta t \gamma_q$. If this error occurs, the system state “jumps” from $\rho(t)$ to $\rho(t + \delta t) = X_q \rho(t) X_q$; otherwise, the system state continuously evolves according to Eq. (5.14), without environmental decoherence. On averaging over many instances of the bit-flip errors, the jump/no-jump approach reduces to the open quantum system model (5.18), where errors continuously change the mixed system state $\rho(t)$.

The encoded logical state is obtained from the mixed three-qubit state $\rho(t)$ as follows

$$\varrho_L(t) = \left\langle \frac{1}{p_{\text{code-space}}(t)} \begin{bmatrix} \rho_{000,000}(t) & \rho_{000,111}(t) \\ \rho_{111,000}(t) & \rho_{111,111}(t) \end{bmatrix} \right\rangle, \quad (5.19)$$

where $p_{\text{code-space}}(t) = \rho_{000,000}(t) + \rho_{111,111}(t)$ is the probability of the system being in the code space.

Our goal is to maximize the fidelity between the target logical wavefunction (6.14) and the true (mixed) logical state (5.19), at some final time, where the evolution includes the decoherence effect of bit-flip errors as well as the effect of the spurious coherent evolution in the error subspaces due to an added Hamiltonian. To counteract the latter two effects, we introduce the double threshold CQEC protocol described in the following section.

5.4 The double threshold CQEC protocol

In the three-qubit bit flip code, the conventional error correction operations are (we will modify these operations to improve the continuous operation performance in Section 6.4):

$$C_{op} = X_1, X_2 \text{ or } X_3. \quad (5.20)$$

These operations are applied on the physical qubits when the error syndrome (defined as the values of the stabilizer generators Z_{12} and Z_{23} , in this order) is equal to $(-1, +1)$, $(-1, -1)$ or $(+1, -1)$, respectively. To apply these error correction operations in the continuous operation, we have to estimate the error syndrome from the noisy readout signals $I_k(t)$ given in Eq. (5.12). To do this, we filter the latter to obtain smoother signals $\mathcal{I}_k(t)$ that obey the following filter equation:

$$\dot{\mathcal{I}}_k(t) = -\frac{\mathcal{I}_k(t)}{\tau} + \frac{I_k(t)}{\tau}, \quad (5.21)$$

where τ plays the role of an averaging time parameter. The initial condition for Eq. (5.21) is discussed below. In practice, the filtered readout signals $\mathcal{I}_k(t)$ can be obtained, e.g., by passing the bare readout signals $I_k(t)$ through a resistor-capacitor circuit (RC lowpass filter [145]). Note that the SNRs of the filtered readout signals can be increased by choosing

a larger value of τ . For instance, in the absence of bit-flip errors, the filtered readout signals read as

$$\mathcal{I}_k(t) = \int_0^t dt' \frac{e^{-\frac{t-t'}{\tau}}}{\tau} I_k(t') \quad (5.22)$$

in the stationary regime ($t \gg \tau$) and their SNRs are equal to $2\tau/\tau_m$. The averaging time parameter τ , however, should not be chosen arbitrarily large; there is an optimal value that is obtained below.

To diagnose the error syndrome, we use a double thresholding scheme that is applied to the filtered readout signals $\mathcal{I}_1(t)$ and $\mathcal{I}_2(t)$. We introduce two error threshold parameters Θ_1 and Θ_2 ($\Theta_1 < \Theta_2$) that define the interval $[\Theta_1, \Theta_2]$, which is referred to as the “syndrome uncertainty region”, see Fig. 5.1. If at least one of the filtered readout signals lies within this interval, we say that we are not certain about the value of the error syndrome, and do nothing. More precisely, the double thresholding scheme works as follows. If $\mathcal{I}_1(t)$ and $\mathcal{I}_2(t)$ are both larger than Θ_2 , the diagnosed error syndrome is $(+1, +1)$ and no error correction operation is applied, since the system quantum state is most likely in the code space. If $\mathcal{I}_1(t) < \Theta_1$ and $\mathcal{I}_2(t) > \Theta_2$, the diagnosed error syndrome is $(-1, +1)$ and the error correction operation to be applied is $C_{op} = X_1$, since the system quantum state is most likely in the error subspace $\mathcal{Q}_1$. If $\mathcal{I}_1(t)$ and $\mathcal{I}_2(t)$ are both smaller than Θ_1 , the diagnosed error syndrome is $(-1, -1)$ and the error correction operation to be applied is $C_{op} = X_2$, since the system quantum state is most likely in the error subspace $\mathcal{Q}_2$. If $\mathcal{I}_1(t) > \Theta_2$ and $\mathcal{I}_2(t) < \Theta_1$, the diagnosed error syndrome is $(+1, -1)$ and the error correction operation to be applied is $C_{op} = X_3$, since the system quantum state is most likely in the error subspace $\mathcal{Q}_3$.

The error correction operations C_{op} must now be applied immediately after an error is detected. Note that this contrasts with the situation in operation of a quantum memory, where correction of errors can be delayed to the end of the continuous operation of the code [131, 132, 134, 137]. In the present analysis, we shall assume that the error correction operations are applied instantaneously on the physical qubits, changing the three-qubit state from $\rho(t)$ to $C_{op}\rho(t)C_{op}$ when the error correction operation C_{op} is applied.

Finally, the filtered readout signals $\mathcal{I}_k(t)$ are reset to the initial condition $+1$ at the moment when an error is diagnosed (see Fig. 5.1). Their subsequent values are dictated by the filter equation (5.21) until the next error is diagnosed, and so on.

Figure 5.1 depicts an example showing how the filtered readout signals $\mathcal{I}_1(t)$ and $\mathcal{I}_2(t)$ are affected by the occurrence of a bit-flip error X_2 at the moment $t_{err} = 162\Gamma_m^{-1}$. Before this error occurs, the system state is in the code space, so the filtered readout signals fluctuate around 1. After the occurrence of the error X_2 , the “signal part” of the filtered readout signals becomes (for $t \geq t_{err}$)

$$\langle \mathcal{I}_1(t) \rangle = \langle \mathcal{I}_2(t) \rangle = -1 + 2e^{-(t-t_{err})/\tau}. \quad (5.23)$$

Equation (5.23) is the solution of Eq. (5.21) with $I_k(t)$ replaced by -1 , which is the “signal part” of the bare readout signal Eq. (5.12) after the error X_2 occurs. Even though both

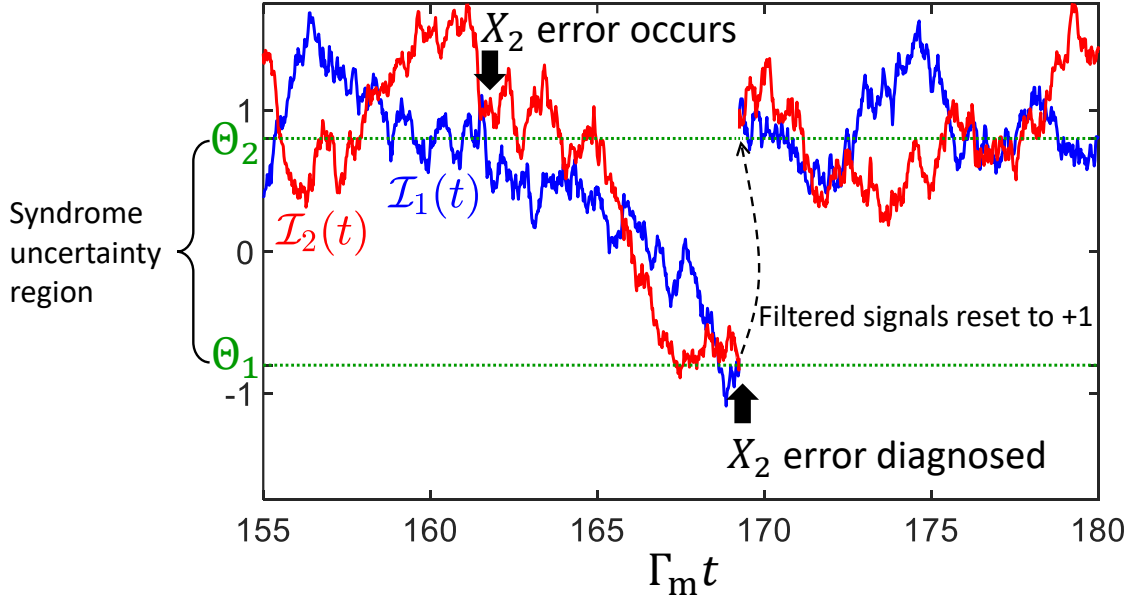


Figure 5.1: Example of filtered readout signals $\mathcal{I}_1(t)$ and $\mathcal{I}_2(t)$ when a bit-flip error X_2 occurs. This error is detected by the CQEC protocol (see main text) at the moment when both filtered readout signals have exited the “syndrome uncertainty region” below the lower error threshold Θ_1 . The filtered readout signals $\mathcal{I}_k(t)$ are discontinuous since the CQEC protocol reset them to the value $+1$ at the moment when the occurred error is diagnosed.

filtered readout signals have the same “signal part”, we see in Fig. 5.1 that these signals follow different paths due to noise. This indicates that if we had used a single error threshold to detect errors, the error X_2 would have been most likely misdiagnosed, because the filtered readout signals do not cross the given error threshold at the same time, see Fig. 5.1. In contrast, our double thresholding scheme performs well unless relatively large fluctuations occur in the filtered readout signals. For instance, in the example considered above, the error X_2 would be diagnosed as X_1 if a relatively large positive fluctuation (of magnitude of the order of $\Theta_2 - \Theta_1$) had made the filtered readout signal $\mathcal{I}_2(t)$ be above the upper error threshold Θ_2 at the moment when the other filtered readout signal $\mathcal{I}_1(t)$ is below the lower error threshold Θ_1 . We will show below that the probability to misdiagnose errors in our double threshold scheme can be made exponentially small by both increasing the length of the “syndrome uncertainty region” and increasing the averaging time parameter τ , see Fig. 5.2. Generally speaking, detecting errors that affect several error syndrome signals $\mathcal{I}_k(t)$ at the same time are the most difficult to detect under continuous monitoring (e.g., error X_2 in the three-qubit bit flip code), and the performance of the latter critically depends on suppressing misdiagnosis of such errors [137].

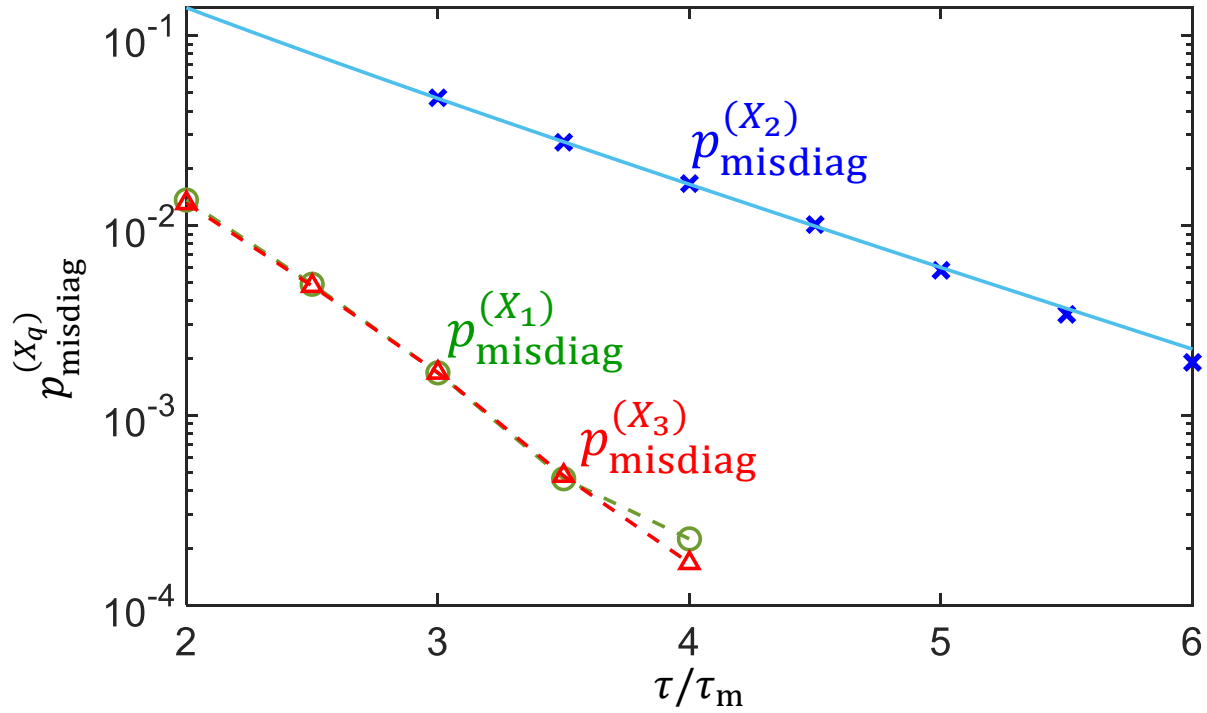


Figure 5.2: Probability $p_{\text{misdiag}}^{(X_q)}$ that the CQEC protocol misdiagnoses the bit-flip error X_q as function of the measurement averaging time parameter τ . The error thresholds are fixed at $\Theta_1 = -0.54$ and $\Theta_2 = 0.8$. Symbols indicate numerical results and the solid line depicts the analytical formula (6.22) of the main text with $c = 1.607$. The dashed lines are guides to the eye.

5.5 Effective logical error rate

With our double threshold protocol, a single-bit flip error could be detected and fixed correctly most of the time. However, there are situations our scheme could fail, the rate for which to happen is called *effective logical error rate*. In this section, we will derive a semi-analytical form for this quantity. There are two leading cases for which the protocol can diagnose the error wrongly. We will discuss them one by one.

Case I: mis-diagnosis of single qubit error

Considering a small time interval Δt , the probability for a X_q error to happen within it is $\gamma_q \Delta t$. Then given that a single error X_q has happened, let's call the probability to mis-diagnose it as $p_{\text{misdiag}}^{(X_q)}$. We will first derive the probability $p_{\text{misdiag}}^{(X_2)}$ to misdiagnose the bit-flip error X_2 . In the continuous operation, misdiagnosis of single bit-flip errors occurs when relatively large fluctuations affect one or both filtered readout signals $\mathcal{I}_k(t)$. The bit-flip error X_2 is misidentified as X_1 if, at the moment when the filtered readout signal $\mathcal{I}_1(t)$ exits

the "syndrome uncertainty region" by crossing the lower error threshold Θ_1 (see Fig. 5.1), the filtered readout signal $\mathcal{I}_2(t)$ is above the upper error threshold Θ_2 due to a unusually large positive fluctuation of size larger than $\Theta_2 - \Theta_1$. The probability that this situation occurs is given by the probability that $\Delta\mathcal{I}(t) \equiv \mathcal{I}_2(t) - \mathcal{I}_1(t) \geq \Theta_2 - \Theta_1$. From Eq. (5.21), we have

$$\frac{d\Delta\mathcal{I}(t)}{dt} = -\frac{\Delta\mathcal{I}(t)}{\tau} + \frac{\sqrt{\tau_m}}{\tau} (\xi_2(t) - \xi_1(t)), \quad (5.24)$$

where the noises $\xi_1(t)$ and $\xi_2(t)$ are the uncorrelated noises of the bare readout signals $I_k(t)$ (see Eqs. (5.12)–(5.13)). Note that Eq. (5.24) is valid both before and after the occurrence of the bit-flip error X_2 because the "signal parts" of the readout signals $I_1(t)$ and $I_2(t)$ cancel each other in $\Delta\mathcal{I}(t)$. Specifically, before (after) occurrence of the error X_2 , the "signals parts" of $I_1(t)$ and $I_2(t)$ are both equal to $+1$ (-1). This implies that the probability that $\Delta\mathcal{I}(t) \geq \Theta_2 - \Theta_1$ can be obtained from the stationary probability distribution, $p_{st}(\Delta\mathcal{I})$, of $\Delta\mathcal{I}(t)$. From Eq. (5.24), we obtain

$$p_{st}(\Delta\mathcal{I}) = \left[\frac{\tau}{2\pi\tau_m} \right]^{1/2} e^{-(\Delta\mathcal{I})^2\tau/2\tau_m}. \quad (5.25)$$

The probability that $\Delta\mathcal{I}(t)$ is larger than $\Theta_2 - \Theta_1$ is then equal to

$$\begin{aligned} p(\Delta\mathcal{I} \geq \Theta_2 - \Theta_1) &= \frac{1}{2} \left[1 - \operatorname{erf} \left(\sqrt{\frac{\tau}{2\tau_m}} (\Theta_2 - \Theta_1) \right) \right], \\ &\approx \frac{1}{\sqrt{2\pi}} \frac{e^{-(\Theta_2 - \Theta_1)^2\tau/2\tau_m}}{(\Theta_2 - \Theta_1)\sqrt{\tau/\tau_m}}, \end{aligned} \quad (5.26)$$

where $\operatorname{erf}(\cdot)$ is the error function and the approximation applies in the limit of large averaging time parameters τ . The result (5.26) is our estimation for the probability that the bit-flip error X_2 is misdiagnosed as the error X_1 . The same result can also obtained for the probability that the error X_2 is misdiagnosed as the error X_3 . Therefore, the probability that the error X_2 is misdiagnosed is given by

$$p_{\text{misdiag}}^{(X_2)} = c \frac{e^{-(\Theta_2 - \Theta_1)^2\tau/2\tau_m}}{(\Theta_2 - \Theta_1)\sqrt{\tau/\tau_m}} \quad (5.27)$$

The numerical coefficient c that follows from the above analysis is $\sqrt{2/\pi} \approx 0.7979$. By fitting our numerical results to Eq. (5.27), we obtain a larger coefficient c . Specifically, fitting to the data in Fig. 5.2 yields $c \approx 1.607$. Equation (5.27) with $c \approx 1.607$ has been successfully tested against numerical results for various values of the error threshold parameters Θ_1 and Θ_2 in addition to the values indicated in Fig. 5.2.

The other two mis-diagnosis probability $p_{\text{misdiag}}^{(X_1)}$ and $p_{\text{misdiag}}^{(X_3)}$ are actually equal to each other due to symmetry in the protocol. They can be derived in a similar argument and turns

out to be much smaller than $p_{\text{misdiag}}^{(X_2)}$ in the parameter regime we are interested in ($\tau \gtrsim 2\tau_m$). So practically we set them to 0. This is again verified in Fig. 5.2. So overall probability for this case to happen is

$$p_{\text{case-1}} = \sum_q \gamma_q \Delta t p_{\text{misdiag}}^{(X_q)} = 1.607 \gamma_2 \Delta t \frac{e^{-(\Theta_2 - \Theta_1)^2 \tau / 2\tau_m}}{(\Theta_2 - \Theta_1) \sqrt{\tau / \tau_m}} \quad (5.28)$$

Case II: mis-diagnosis of two consecutive single qubit errors

This case will happen if two single qubit errors occurred consecutively in time. Instead, the protocol diagnoses a different (false) error. It should be clear that if two consecutive errors occur sufficiently far apart in time, both errors will be correctly diagnosed. On the other hand, if these errors occur sufficiently close in time, the CQEC protocol can fail, since our protocol determines the error syndrome from the filtered readout signals $\mathcal{I}_k(t)$, which are slow and take some time (proportional to the averaging time parameter τ) to exit the “syndrome uncertainty region”, as evident in Fig. 5.1. Let us denote $\Delta t_{qq'}$ as the time window within which two consecutive errors, first X_q and then $X_{q'}$, are misdiagnosed as the false error $X_{q''}$ ($q \neq q' \neq q''$). The probability for this scenario is given by

$$p_{\text{case-2}} = 2 (\gamma_1 \gamma_2 \Delta t_{12} + \gamma_2 \gamma_3 \Delta t_{23} + \gamma_1 \gamma_3 \Delta t_{13}) \Delta t, \quad (5.29)$$

where the factor of 2 comes from the fact the order of the two errors does not matter. The time windows Δt_{12} , Δt_{23} and Δt_{13} can be easily evaluated in the noiseless approximation, which means we assume SNR is large and neglect the noisy part of the signal. As illustrated in Fig. (5.3) and by using Eq. (5.23), this yields

$$\Delta t_{12} = \Delta t_{23} = \tau \ln \left[\frac{2}{1 + \Theta_1} \right] \quad \text{and} \quad \Delta t_{13} = \tau \ln \left[\frac{1 + \Theta_2}{1 + \Theta_1} \right]. \quad (5.30)$$

The factor of 2 in Eq. (5.29) is due to the fact that the time window $\Delta t_{qq'}$ is the same as $\Delta t_{q'q}$, which is the corresponding time window for the case where the error $X_{q'}$ occurs before the error X_q .

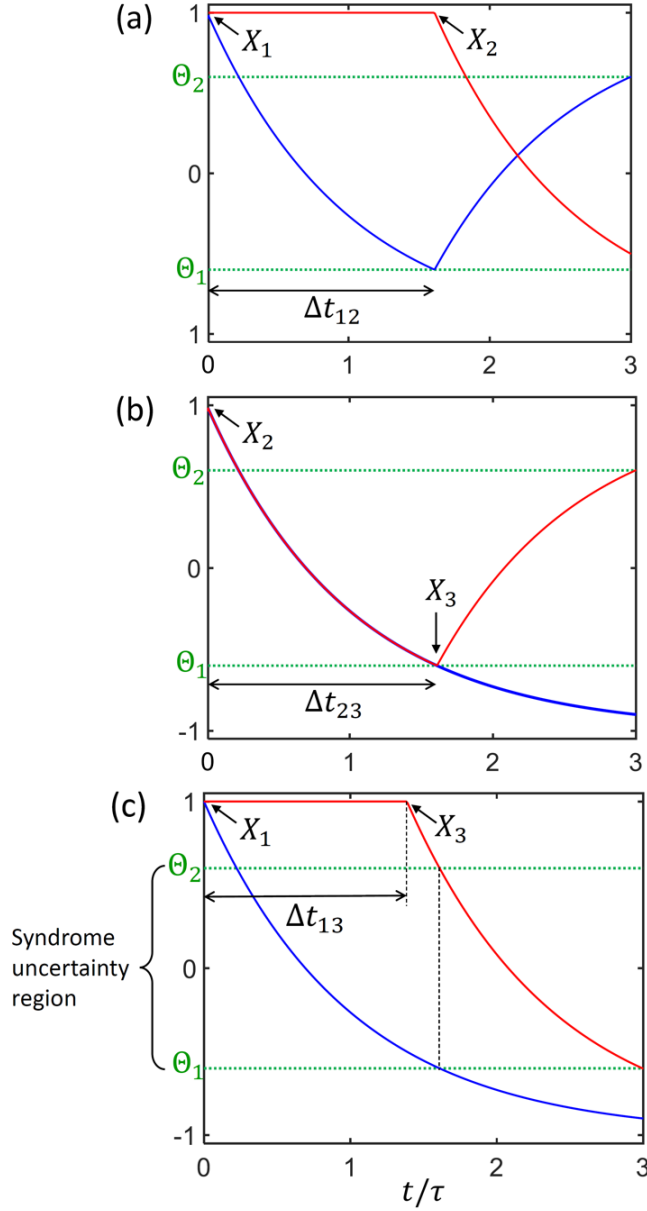


Figure 5.3: Time windows $\Delta t_{qq'}$ for two consecutive bit-flip errors X_q (first) and $X_{q'}$ (second) to be diagnosed as a single (false) error $X_{q''}$, in the noiseless approximation. This is scenario three of the main text. The first error (X_q) occurs at the moment $t = 0$. The solid lines depict the “signal part” of the filtered readout signals $\mathcal{I}_1(t)$ [blue lines] and $\mathcal{I}_2(t)$ [red lines]. In panels (a) and (b), the window times Δt_{12} and Δt_{23} are both equal to the time that the blue solid line [i.e., $\langle \mathcal{I}_1(t) \rangle$, given in Eq. (5.23) of the main text] takes to cross the lower error threshold Θ_1 . In panel (c), the window time Δt_{13} is given by the time that the blue solid line spends inside the “syndrome uncertainty region”. Explicit formulas for the windows times $\Delta t_{qq'}$ are given in Eq. (5.30) of the main text. If the second error occurs within the time window $t = \Delta t_{qq'}$, the depicted two-error combinations are diagnosed in the noiseless approximation as $X_{q''} = X_3$ [panel (a)], X_1 [panel (b)] and X_2 [panel (c)].

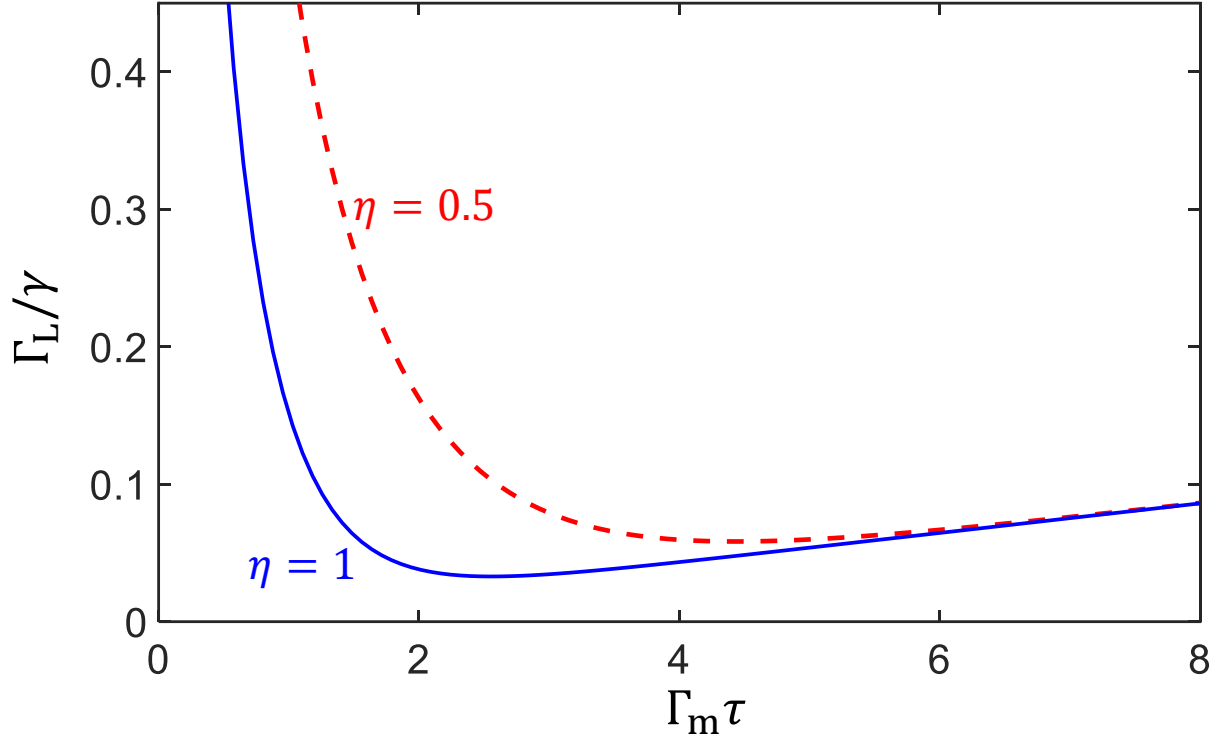


Figure 5.4: Logical error rate Γ_L as function of the measurement averaging time parameter τ . The curves depict formula Eq. (5.31) of the main text, evaluated for bit-flip error rates $\gamma_q = \gamma = 1.25 \times 10^{-3} \Gamma_m$ ($q = 1, 2, 3$), error threshold parameters $\Theta_1 = -0.54$ and $\Theta_2 = 0.8$, and quantum efficiencies $\eta = 0.5$ (dashed curve) and $\eta = 1$ (solid curve). Γ_m denotes measurement strength from continuous measurement.

Finally, we can sum up the probability for both cases and write down the formula for logical error rate

$$\begin{aligned}
 \Gamma_L &= \frac{p_{\text{case-1}} + p_{\text{case-2}}}{\Delta t} \\
 &= 1.607 \gamma_2 \frac{e^{-(\Theta_2 - \Theta_1)^2 \tau / 2\tau_m}}{(\Theta_2 - \Theta_1) \sqrt{\tau / \tau_m}} + 2(\gamma_1 \gamma_2 + \gamma_2 \gamma_3) \tau \times \\
 &\quad \ln \left[\frac{2}{1 + \Theta_1} \right] + 2\gamma_1 \gamma_3 \tau \ln \left[\frac{1 + \Theta_2}{1 + \Theta_1} \right].
 \end{aligned} \tag{5.31}$$

Note that Γ_L implicitly depends on the efficiency of the measurement, η , via the explicit dependence on measurement time $\tau_m = 1/(2\Gamma_m\eta)$. In a given experimental setup, the parameters τ, Θ_1, Θ_2 would constitute a minimal set of tunable parameters.

Figure 5.4 shows the non-monotonic dependence of Γ_L on the time-averaging parameter τ , for fixed values of the error threshold parameters $\Theta_1 = -0.54$ and $\Theta_2 = 0.8$, and equal

bit-flip error rates $\gamma_q = \gamma = 1.25 \times 10^{-3} \Gamma_m$. Note that, in the limit of relatively small τ , the logical X error rate increases exponentially because the SNR of the filtered readout signals decreases, leading to more frequent false diagnoses of X_2 errors. In this limit, the first term of Eq. (5.31) is dominant. In the opposite limit of relatively large τ , the logical X error rate increases linearly in τ , due to misdiagnosis of two errors that occur sufficiently close in time. We see that measurement inefficiency $\eta \leq 1$ affects the logical error rate only for small averaging times τ and has no effect at large τ . This reflects the fact that while the mis-diagnosis of single qubit errors that dominates Γ_L at small τ depends on measurement efficiency via τ_m (measurement time parameter), the mis-diagnosis of two errors occurring close in time was evaluated in the noiseless approximation and does not depend on η .

5.6 Optimization and benchmarking

As we pointed in previous section, there are three tunable parameters, the averaging filtering time τ , and the two threshold θ_1, θ_2 . We can adjust these quantities to minimize the logical error rate. Notice there are extra parameters γ_q and η (which is hidden in the measurement time τ_m) in Γ_L . These parameters come from the hardware setup and are not part of the control scheme. It is expected that the optimized control parameters should depend on the hardware parameters. In all the numerical discussion below in this and next chapter, we will assume the bit-flip is identical on each physical qubit, i.e. $\gamma_q = \gamma$ for all q . It turns out the two threshold parameters are insensitive to the bit-flip rate and measurement efficiency. Approximately, their optimal values are $\Theta_2 = 0.8$ and $\Theta_1 = -0.54$. On the other hand, the optimal value of τ varies a lot depending on the hardware parameters, as shown in figure 6.2. We will wait until next chapter to discuss its behavior together with the case when the system is also under coherent evolution.

Figure 5.5 depicts the optimized value of the logical X error rate Γ_L^{opt} for quantum memory. We find that this logical error rate scales approximately quadratically with the error rate γ :

$$\begin{aligned} \Gamma_L^{opt} &\approx 27.3208 (\gamma \Gamma_m^{-1})^{1.897} \Gamma_m \quad (\eta = 0.5), \\ &\approx 15.7580 (\gamma \Gamma_m^{-1})^{1.904} \Gamma_m \quad (\eta = 1). \end{aligned} \quad (5.32)$$

The numerical factors and exponents in the above equation are obtained from fitting for $\gamma \in [10^{-6} \Gamma_m, 10^{-4} \Gamma_m]$. The approximate quadratic scaling of Γ_L^{opt} with γ indicates that the double threshold CQEC protocol is both effective and accurate in diagnosing single bit-flip errors. Figure 5.5 also shows the logical X error rate for the linear variant of the optimal Wonham filter, $\Gamma_L^{Wonham} = 3\gamma^2 \tau_m \ln(2/\gamma \tau_m)$, that was obtained in Ref. [134]. We point out that our optimized logical error rate Γ_L^{opt} is very close to that of the linear variant of the optimal Wonham filter.

In addition, we find that the discrete and continuous operations can exhibit similar performance if the cycle time Δt_{cycle} from the discrete operation is related to the strength

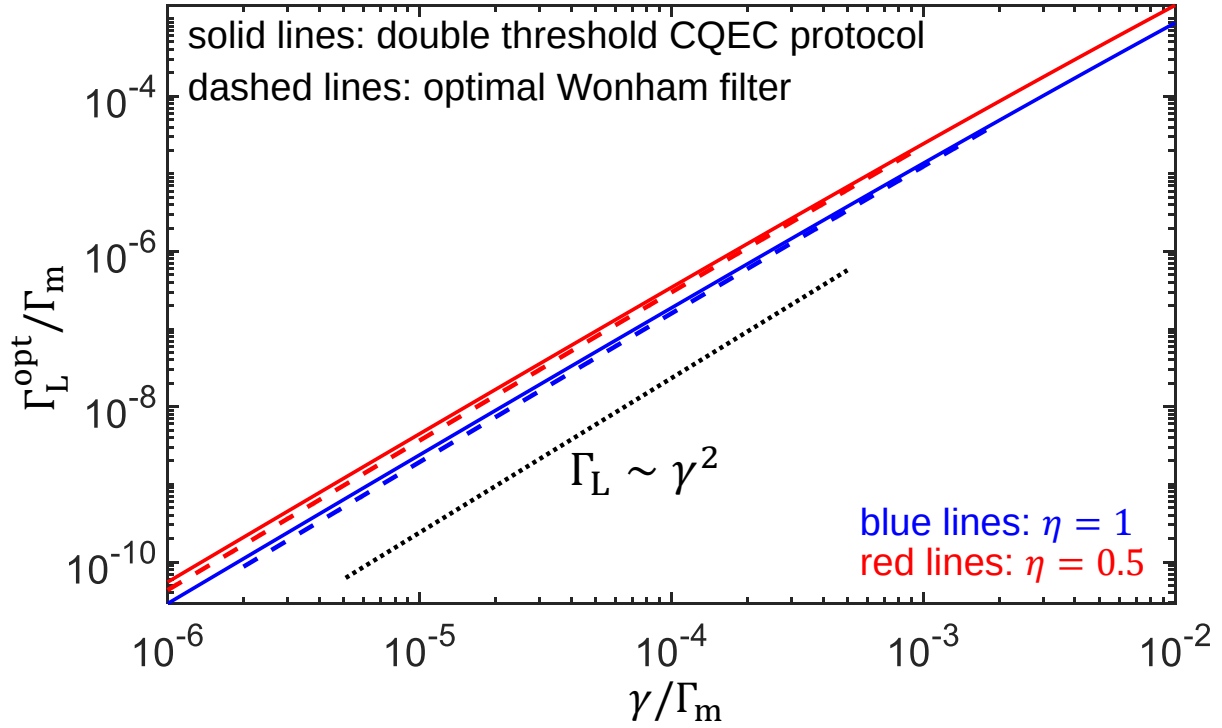


Figure 5.5: Optimized logical X error rate Γ_L^{opt} for quantum memory operation. Solid red and blue lines depict Γ_L^{opt} , obtained using our double threshold CQEC protocol (Eq. (5.31)), and the dashed red and blue lines depict the logical X error rate, $\Gamma_L^{\text{Wonham}} = 3\gamma^2\tau_m \ln(2/\gamma\tau_m)$, for the linear variant of the optimal Wonham filter [134]. Γ_m and $\tau_m = (2\Gamma_m\eta)^{-1}$ denote the measurement strength and “measurement time” from continuous measurement, respectively. Red and blue lines depict results for quantum efficiencies $\eta = 0.5$ and 1, respectively. The dotted line depicts the quadratic scaling $\Gamma_L \sim \gamma^2$.

Γ_m of the continuous measurements as follows:

$$\begin{aligned} t_{\text{cycle}} &\approx 9.1069\gamma^{-1} (\gamma\Gamma_m^{-1})^{0.897} & (\eta = 0.5), \\ &\approx 5.2527\gamma^{-1} (\gamma\Gamma_m^{-1})^{0.904} & (\eta = 1). \end{aligned} \quad (5.33)$$

The above results are obtained from the relation $\Gamma_L^{\text{opt}} = \Gamma_L^{\text{disc}}$, where $\Gamma_L^{\text{disc}} = 3\gamma^2\Delta t_{\text{cycle}}$ is the logical X error rate for the discrete operation.

Chapter 6

Continuous error correction II: time-dependent Hamiltonian system

6.1 Adiabatic quantum computation in a nutshell

Adiabatic quantum computation (AQC) is an alternative model for quantum computation in contrast to circuit gate model. It is a universal computation model which utilizes the dynamical evolution of a quantum system. In AQC, we encode the solution of a computation problem in the ground state of a constant Hamiltonian, which is usually called the problem Hamiltonian H_p . We start the evolution from an initial state which is experimentally easy to prepare. Usually it is the symmetric superposition state of the standard basis. The initial state is also the ground state of some initial Hamiltonian H_0 . The system Hamiltonian is controlled to transit from the initial Hamiltonian H_0 to the final problem Hamiltonian H_p . If this transition is sufficiently slow that we are in the adiabatic limit, the adiabatic theorem guarantees the system to evolve along the ground state. Then at the final moment of the whole evolution, the system would be in the ground state of H_p . For system consisting of qubits, which is the case we are interested in, local measurements on the final state along the z axis of each qubit can tell us the solution.¹

The system Hamiltonian is time-dependent and can be written as

$$H(t) = a(t)H_0 + b(t)H_p \quad (6.1)$$

where in the time interval $[0, t_{op}]$, $a(t)$ starts from 1 and ends at 0 while $b(t)$ starts from 0 and ends at 1. t_{op} is the operation time when we stop the evolution. The ground state of H_0 is usually required to be non-degenerate while H_p can have a degenerate ground state since the solution of a problem may not be unique. So over time, $H(t)$ will gradually transit from H_0 to H_p . How slow we should make the transition to be depends on the minimum gap Δ of between energy of the ground state and the first excited state. A sufficient condition is

¹The solution is represented by a binary string of 0s and 1s which correspond to product state of the standard basis of qubits.

that (up to some polylogarithmic factor in Δ) [146]

$$t_{op} \sim O\left(\frac{1}{\Delta^2}\right) \quad (6.2)$$

Adiabatic theorem

In this subsection, we briefly review the adiabatic theorem.² This theorem tells us if the Hamiltonian of the system varies slowly enough or adiabatically, then a state starting in an eigenstate will stay as the eigenstate of the same energy level. To prove this theorem, consider a time-dependent Hamiltonian $H(t)$. Let $|\epsilon_i(t)\rangle$ be the **instantaneous** eigenstate with energy $\epsilon_i(t)$. We assume the energy levels never touch each other throughout the evolution.

Let's introduce a unitary transformation which maps the instantaneous basis back to the fixed initial basis at $t = 0$

$$U(t, 0) = \sum_i e^{-i \int_0^t dt' \epsilon_i(t')} |\epsilon_i(0)\rangle \langle \epsilon_i(t)| \quad (6.3)$$

where the exponential prefactor is the dynamical phase term. Instead of looking at the original system state $|\psi(t)\rangle$, we transform it to a co-moving state

$$|\psi_c(t)\rangle = U(t, 0)|\psi(t)\rangle \quad (6.4)$$

The meaning of this state is clear if we do a basis expansion

$$|\psi_c(t)\rangle = \sum_i e^{-i \int_0^t dt' \epsilon_i(t')} |\epsilon_i(0)\rangle \langle \epsilon_i(t)| \psi(t)\rangle \quad (6.5)$$

It is a state with the same coefficients (up to a phase) of the instantaneous basis as $|\psi(t)\rangle$ but this time, the coefficients are along the initial basis. We would like to find the Hamiltonian which governs the dynamics of this new state. Taking time derivative of the state and applying Schrödinger equation to $|\psi\rangle$ (setting $\hbar = 1$)

$$\frac{d|\psi(t)\rangle}{dt} = -iH(t)|\psi(t)\rangle \quad (6.6)$$

we arrive at the equation of motion for $|\psi_c(t)\rangle$

$$\frac{d|\psi_c(t)\rangle}{dt} = -iH_c(t)|\psi_c(t)\rangle \quad (6.7)$$

where the co-moving Hamiltonian $H_c(t)$ is given by

$$\begin{aligned} H_c(t) &= i \sum_{ij} e^{\int_0^t dt' (\epsilon_i(t') - \epsilon_j(t'))} \langle \dot{\epsilon}_i(t) | \epsilon_j(t) \rangle |\epsilon_i(0)\rangle \langle \epsilon_j(0)| \\ &= i \sum_i \langle \dot{\epsilon}_i(t) | \epsilon_i(t) \rangle |\epsilon_i(0)\rangle \langle \epsilon_i(0)| \\ &\quad + i \sum_{i \neq j} e^{\int_0^t dt' (\epsilon_i(t') - \epsilon_j(t'))} \langle \dot{\epsilon}_i(t) | \epsilon_j(t) \rangle |\epsilon_i(0)\rangle \langle \epsilon_j(0)| \end{aligned} \quad (6.8)$$

²We only consider the simplest case here, neglecting the possibility of having multiple time scales in the Hamiltonian. So the only relevant time scale is t_{op} .

We can further evaluate the inner product term

$$\langle \dot{\epsilon}_i(t) | \epsilon_j(t) \rangle = \frac{\langle \epsilon_i(t) | \dot{H} | \epsilon_j(t) \rangle}{\epsilon_i(t) - \epsilon_j(t)} \text{ for } i \neq j \quad (6.9)$$

This Hamiltonian is not diagonal yet. Dropping the $i \neq j$ term in Eq. (6.8) gives us the diagonal adiabatic Hamiltonian. We can do this as long as the adiabatic condition is satisfied [147]

$$\frac{|\langle \dot{\epsilon}_i(t) | \epsilon_j(t) \rangle|}{|\epsilon_i(t) - \epsilon_j(t)|} = \frac{|\langle \epsilon_i(t) | \dot{H} | \epsilon_j(t) \rangle|}{|\epsilon_i(t) - \epsilon_j(t)|^2} \ll 1 \text{ for all } i \neq j \text{ and } t \in [0, t_{op}] \quad (6.10)$$

We would like to point out that there are cases that the evolution is not adiabatic even when this condition is satisfied. For a more careful condition when multiple time scales appear in the Hamiltonian, see [148, 149]. This concludes our minimal introduction to adiabatic quantum computation. The basic concepts in this section should be enough to help the readers understand the quantum annealing part for our research work, which will be discussed in the following sections. For a comprehensive review of AQC, see [138].

6.2 Encoding the time-dependent Hamiltonian

In this chapter we shall focus on CQEC with the encoded quantum annealing Hamiltonian ($\hbar = 1$)

$$H(t) = -\Omega_0 [a(t)X_L + b(t)Z_L], \quad (6.11)$$

where the frequency parameter Ω_0 sets the energy scale of the above Hamiltonian, and the coefficients $a(t)$ and $b(t)$ are functions of time with magnitudes smaller than 1. The operators X_L and Z_L denote the logical X and Z operators, given by

$$X_L = X_1 X_2 X_3 \quad \text{and} \quad Z_L = \frac{Z_1 + Z_2 + Z_3}{3}, \quad (6.12)$$

where X_q represents the Pauli x operator that acts on the q th physical qubit. Note that the system Hamiltonian (6.11) and the stabilizer generators (5.11) exhibit a block-diagonal matrix representation in the computational basis.

As in the previous chapter, we label the two logical basis states as

$$|0_L\rangle = |000\rangle \text{ and } |1_L\rangle = |111\rangle \quad (6.13)$$

In the absence of errors, the target logical wavefunction

$$|\psi_L(t)\rangle = \alpha_L(t)|0_L\rangle + \beta_L(t)|1_L\rangle \quad (6.14)$$

evolves according to the following Schrödinger equation for the probability amplitudes of the zero (α_L) and one (β_L) logical states:

$$\begin{bmatrix} \dot{\alpha}_L(t) \\ \dot{\beta}_L(t) \end{bmatrix} = -ih_L(t) \begin{bmatrix} \alpha_L(t) \\ \beta_L(t) \end{bmatrix}. \quad (6.15)$$

In the above equation, $h_L(t)$ represents the Hamiltonian of the logical qubit and is given by the 2×2 diagonal sub-matrix of $H(t)$ that corresponds to the code space, i.e.,

$$h_L(t) = -\Omega_0 [a(t)\sigma_x + b(t)\sigma_z]. \quad (6.16)$$

Here σ_x and σ_z denote the conventional Pauli x and z matrices of the logical qubit, and the coefficients $a(t)$ and $b(t)$ are the coefficients given in Eq. (6.11). In this chapter, we shall use the notation $|\psi_L(t)\rangle$ to denote the column matrix $[\alpha_L(t) \ \beta_L(t)]^T$.

We combine the evolution under the quantum annealing Hamiltonian Eq. (6.11) with the stochastic master equation describing the continuous measurement of the stabilizer generators S_k , (5.14). We emphasize that evolution of the target logical wavefunction (6.14) is not affected by measurement and on the other hand, the population in each subspace $\mathcal{Q}_i$ is not affected by the Hamiltonian. This is because the system Hamiltonian (6.11) and the three stabilizer generators ($S_k = Z_1 Z_2, Z_2 Z_3$, and $Z_1 Z_3$) commute with each other³. Thus there is no quantum Zeno effect (unlike the non-commuting situation, see e.g., [150]).

The error subspace $\mathcal{Q}_1$ is spanned by the computational states $|100\rangle = X_1|0_L\rangle$ and $|011\rangle = X_1|1_L\rangle$; the error subspace $\mathcal{Q}_2$ is spanned by the computational states $|010\rangle = X_2|0_L\rangle$ and $|101\rangle = X_2|1_L\rangle$; and the error subspace $\mathcal{Q}_3$ is spanned by the computational states $|001\rangle = X_3|0_L\rangle$ and $|110\rangle = X_3|1_L\rangle$. In addition, the 2×2 diagonal sub-matrices of $H(t)$ that correspond to these error subspaces are all identical and equal to

$$h_{\text{spurious}}(t) = -\Omega_0 \left[a(t)\sigma_x + \frac{1}{3}b(t)\sigma_z \right]. \quad (6.17)$$

Note the factor of $1/3$ in the above equation. This derives from the action of the system Hamiltonian $H(t)$, Eq. (6.11), on a state with support in one of the three error subspaces. For instance, for the system state, $|\psi(t)\rangle = \alpha X_1|0_L\rangle + \beta X_1|1_L\rangle$, which is in the error subspace $\mathcal{Q}_1$, we obtain $H(t)|\psi(t)\rangle = -\Omega_0 a(t)[\alpha X_1|1_L\rangle + \beta X_1|0_L\rangle] - (1/3)[\alpha X_1|0_L\rangle - \beta X_1|1_L\rangle]\Omega_0 b(t)$. In contrast, this factor of $1/3$ does not appear when the system Hamiltonian $H(t)$ acts on the (code space) logical states, Eq. (6.14). We can therefore say that when the system state is in the error subspaces, coherent evolution in those subspaces is due to the spurious Hamiltonian (6.17), instead of the intended logical Hamiltonian (6.16).

³For collecting the error information, we only need two of these. However in order for this statement to hold, we need all three.

6.3 Effective open-system model for the logical qubit

Quantum master equation for the logical qubit

In this subsection we develop an approximate evolution equation for the mixed logical state $\varrho_L(t)$ that describes the combined action of both bit-flip errors and the above CQEC protocol, and the action of an applied time-dependent Hamiltonian. We are particularly interested in the limit of sufficiently small bit-flip error rates γ_q , where single bit-flip errors are the most probable, followed by two bit-flip errors, and so on. In this regime there are three different scenarios that can give rise to logical errors during the time evolution—a single misdiagnosed bit-flip error, spurious coherent evolution in an error subspace prior to a correctly diagnosed bit-flip error, and two bit-flip errors that are misdiagnosed as one. Not surprisingly, two of the scenarios are the same as for the quantum memory case since our protocol has not been changed. We should expect to see a term with the logical error rate Γ_L (Eq. (5.31)) appearing in the final master equation. The only difference now is we will have spurious coherent evolution which leads to additional logical errors. We analyze each of these in turn below.

For the following analysis it is convenient to introduce a timestep Δt such that

$$t_{\text{det}} \ll \Delta t, \quad \Delta t \ll \gamma_q^{-1}, \quad \Delta t \ll \frac{\hbar}{|H_L(t)|}, \quad \text{and} \quad \Delta t \ll t_{\text{op}}, \quad (6.18)$$

where t_{det} denotes the characteristic time to detect a bit-flip error by our CQEC protocol, and t_{op} is the operation time of the continuous implementation. Because of the second inequality of Eq. (6.18), we assume below that at most two bit-flip errors occur within each timestep Δt . We shall eventually send Δt to zero, to obtain an effective evolution equation for the encoded density matrix $\varrho_L(t)$.

We consider first the scenario where a single bit-flip error that occurs in the time interval $(t, t + \Delta t)$ is misdiagnosed by the CQEC protocol. In this case, a wrong error correction operation is applied to one of the physical qubits: this incorrect operation transfers the system state to another error subspace, instead of back to the code space. For instance, if the actual error is X_2 but the diagnosed error syndrome is $(-1, +1)$ instead of $(-1, -1)$, the error correction operation that will be applied is $C_{\text{op}} = X_1$ instead of $C_{\text{op}} = X_2$. This will incorrectly transfer the system state from error subspace $\mathcal{Q}_2$ to error subspace $\mathcal{Q}_3$, resulting in a logical X error, since $X_1 X_2 = X_3 X_L$ and X_L is the logical X operator. The system state will be returned to the code space by the next iteration of the CQEC protocol if this iteration successfully diagnoses the new error syndrome. We shall assume that the probability to misdiagnose a bit-flip error is small enough that a series of two consecutive misdiagnoses is unlikely, and the next iteration does indeed return the system state to the code space. After completion of the next (successful) iteration of the CQEC protocol, the system state at the moment $t + \Delta t$ is equal to $X_L \rho(t) X_L$, which implies that the 2×2 logical density matrix at that moment is

$$\varrho_L^{\text{scn-1}}(t + \Delta t) = \sigma_x \varrho_L(t) \sigma_x. \quad (6.19)$$

Recall the probability of this scenario is given by

$$p_{\text{scn-1}} = \Delta t \left(\gamma_1 p_{\text{misdiag}}^{(X_1)} + \gamma_2 p_{\text{misdiag}}^{(X_2)} + \gamma_3 p_{\text{misdiag}}^{(X_3)} \right), \quad (6.20)$$

where the misdiagnosis probability $p_{\text{misdiag}}^{(X_q)}$ is as given in Chapter 5 and we restate them here

$$p_{\text{misdiag}}^{(X_1)} = p_{\text{misdiag}}^{(X_3)} \approx 0 \quad (6.21)$$

$$p_{\text{misdiag}}^{(X_2)} = 1.607 \frac{e^{-(\Theta_2 - \Theta_1)^2 \tau / 2\tau_m}}{(\Theta_2 - \Theta_1) \sqrt{\tau / \tau_m}} \quad (6.22)$$

This scenario results in a contribution $\Delta \varrho_L^{(1)}$ to the actual logical state $\varrho_L(t + \Delta t)$ at the moment $t + \Delta t$ (Eq. (6.37)), with

$$\Delta \varrho_L^{(1)} = p_{\text{scn-1}} \varrho_L^{\text{scn-1}}(t + \Delta t). \quad (6.23)$$

Note that in the argument leading to Eq. (6.19), we have disregarded the coherent evolution of the system state in the error subspaces because this leads to correction terms of the order of $(t_{\text{det}} \Omega_0)^2 \ll (\Delta t \Omega_0)^2$. These can be neglected since we shall take the limit $\Delta t \rightarrow 0$ below and thus only need to keep terms up to first order in Δt in Eq. (6.23).

The second scenario corresponds to the case of a single bit-flip error that is correctly diagnosed by the CQEC protocol. The probability for this scenario is

$$p_{\text{scn-2}}^{(q)} = \Delta t \gamma_q (1 - p_{\text{misdiag}}^{(X_q)}). \quad (6.24)$$

In contrast to the first scenario, logical errors are now due only to spurious coherent evolution in the corresponding error subspace during the time that it takes to diagnose and correct the occurred error. Let us assume that the bit-flip error X_q occurs at the instant $t' \in [t, t + \Delta t]$. We shall denote the time to detect such an error as $t_{\text{det}}^{(q)}$, where the upper index q indicates that in general the error detection time may depend on the bit-flip error type, $X_q = X_1, X_2$ or X_3 . The system density matrix at the moment $t + \Delta t$ is

$$\rho^{\text{scn-2}}(t + \Delta t) = \frac{\gamma_q}{\mathcal{N}} \int_t^{t + \Delta t - t_{\text{det}}^{(q)}} \mathcal{V}_q(t') \rho(t) \mathcal{V}_q^\dagger(t') dt', \quad (6.25)$$

where $\mathcal{N} = (\Delta t - t_{\text{det}}^{(q)}) \gamma_q$ is a normalization factor such that $\text{Tr}[\rho^{\text{scn-2}}(t + \Delta t)] = 1$. The integral in Eq. (6.25) evaluates the average over the error instant t' ,

$$\mathcal{V}_q(t') = \mathcal{U}(t' + t_{\text{det}}^{(q)}, t + \Delta t) X_q \mathcal{U}(t', t' + t_{\text{det}}^{(q)}) X_q \mathcal{U}(t, t'), \quad (6.26)$$

and $\mathcal{U}(t_1, t_2)$ with $t_1 \leq t_2$ denotes the unitary evolution operator associated to the system Hamiltonian (6.11). If we read the right-hand side of Eq. (6.26) from right to left, the first X_q operator accounts for the error that occurred, and the second X_q operator accounts

for the application of the error correction operation, which is $C_{\text{op}} = X_q$ since the occurred error is correctly diagnosed. We now seek to approximate $\rho^{\text{scn-2}}(t + \Delta t)$ to first order in Δt . Because the integral in Eq. (6.25) is over a time interval of duration approximately equal to Δt , we may write $\rho^{\text{scn-2}}(t + \Delta t) \approx \mathcal{V}_q(t) \rho(t) \mathcal{V}_q^\dagger(t)$, where the integrand of Eq. (6.25) has been evaluated at $t' = t$. In addition, the operator $\mathcal{V}_q(t)$ may be replaced by its zero-order approximation in Δt :

$$\begin{aligned} \mathcal{V}_q(t) &\approx e^{iH(t)t_{\text{det}}^{(q)}} C_{\text{op}} e^{-iH(t)t_{\text{det}}^{(q)}} X_q \\ &= e^{iH(t)t_{\text{det}}^{(q)}} X_q e^{-iH(t)t_{\text{det}}^{(q)}} X_q \\ &= e^{iH(t)t_{\text{det}}^{(q)}} e^{-iX_q H(t) X_q t_{\text{det}}^{(q)}}. \end{aligned} \quad (6.27)$$

Note that the 8×8 matrices $X_q H(t) X_q$ and $H(t)$ exhibit a similar block-diagonal matrix representation in the computational basis, since both commute with the stabilizer generators. This block-diagonal structure consists of 2×2 diagonal submatrices for each subspace $\mathcal{Q}_\ell$. In particular, the 2×2 diagonal submatrices of $X_q H(t) X_q$ and $H(t)$ corresponding to the code space are given by the spurious Hamiltonian $h_{\text{spurious}}(t)$ and the logical Hamiltonian $h_L(t)$ defined in Eqs. (6.16) and (6.17), respectively. This implies that the 2×2 diagonal submatrix of $\mathcal{V}_q(t)$ that corresponds to the code space can be approximated as

$$V_q(t) = \exp \left[i t_{\text{det}}^{(q)} h_L(t) \right] \exp \left[-i t_{\text{det}}^{(q)} h_{\text{spurious}}(t) \right]. \quad (6.28)$$

Up to first order in Δt , the logical state at the moment $t + \Delta t$ is then given by

$$[\varrho_L^{\text{scn-2}}]_q(t + \Delta t) = V_q(t) \varrho_L(t) V_q^\dagger(t). \quad (6.29)$$

Equation (6.28) provides an effective parameterization of the effective action of the logical error operation $V_q(t)$ due to spurious coherent evolution in an error subspace during detection of a single bit-flip error, in terms of the error detection time $t_{\text{det}}^{(q)}$. We can estimate this time from the “signal part” of the filtered readout signals $\mathcal{I}_k(t)$, i.e., disregarding the noise. In this noiseless approximation, the error-detection time is the same for all bit-flip errors; i.e., $t_{\text{det}}^{(q)} = t_{\text{det}}$, so we may consider a particular case. Let us consider the bit-flip error X_2 . If we apply the CQEC protocol to the “signal part” of the filtered readout signals, the error X_2 will be diagnosed when $\langle \mathcal{I}_k(t_{\text{err}} + t_{\text{det}}) \rangle = \Theta_1$ for $k = 1, 2$. From this condition and Eq. (5.23), we obtain the error-detection time

$$t_{\text{det}}^{(q)} = t_{\text{det}} = \tau \ln \left[\frac{2}{1 + \Theta_1} \right], \quad q = 1, 2, 3. \quad (6.30)$$

More generally, the presence of noise in the filtered readout signals will make the error-detection times random. For simplicity, and to obtain analytic estimates, we shall assume in this work that they are deterministic and given by Eq. (6.30).

The contribution of this scenario to the logical state $\varrho_L(t + \Delta t)$ at the moment $t + \Delta t$ is (see Eq. (6.37))

$$\Delta \varrho_L^{(2)} = \sum_{q=1,2,3} p_{\text{scn-2}}^{(q)} [\varrho_L^{\text{scn-2}}]_q(t + \Delta t). \quad (6.31)$$

The third scenario is the case of two errors that occur sufficiently close in time that they are not individually diagnosed by the CQEC protocol. This probability is as shown in Eq. (5.29) & (5.30)

$$p_{\text{scn-3}} = 2(\gamma_1 \gamma_2 \Delta t_{12} + \gamma_2 \gamma_3 \Delta t_{23} + \gamma_1 \gamma_3 \Delta t_{13}) \Delta t, \quad (6.32)$$

where the window times are

$$\Delta t_{12} = \Delta t_{23} = \tau \ln \left[\frac{2}{1 + \Theta_1} \right] \text{ and } \Delta t_{13} = \tau \ln \left[\frac{1 + \Theta_2}{1 + \Theta_1} \right]. \quad (6.33)$$

Neglecting spurious coherent evolution in the error subspaces, application of the wrong error correction operation $C_{\text{op}} = X_{q''}$ effectively induces a logical X operation on the system state $\rho(t)$ since $C_{\text{op}} X_{q'} X_q = X_{q''} X_{q'} X_q = X_L$, and then the logical density matrix changes from $\varrho_L(t)$ to

$$\varrho_L^{\text{scn-3}}(t + \Delta t) = \sigma_x \varrho_L(t) \sigma_x \quad (6.34)$$

at the moment $t + \Delta t$ (see also Eq. (6.19)). The contribution of this scenario to the logical state $\varrho_L(t + \Delta t)$ at the moment $t + \Delta t$ is (see Eq. (6.37))

$$\Delta \varrho_L^{(3)} = p_{\text{scn-3}} \varrho_L^{\text{scn-3}}(t + \Delta t). \quad (6.35)$$

Finally, if none of the above three scenarios occur, the logical state at the moment $t + \Delta t$ is given by the time evolved state under the logical Hamiltonian $h_L(t)$ of Eq. (6.16) and is equal to

$$\varrho_L^{\text{scn-0}}(t + \Delta t) = \varrho_L(t) - i[h_L(t), \varrho_L(t)]\Delta t, \quad (6.36)$$

where we have disregarded terms of order $(\Delta t)^2$.

The logical state at the moment $t + \Delta t$ that takes into account all of the above four scenarios is then given by

$$\begin{aligned} \varrho_L(t + \Delta t) = & \left(1 - p_{\text{scn-1}} - \sum_{q=1,2,3} p_{\text{scn-2}}^{(q)} - p_{\text{scn-3}} \right) \varrho_L^{\text{scn-0}}(t + \Delta t) + \\ & p_{\text{scn-1}} \varrho_L^{\text{scn-1}}(t + \Delta t) + \sum_{q=1,2,3} \left\{ p_{\text{scn-2}}^{(q)} [\varrho_L^{\text{scn-2}}]_q(t + \Delta t) \right\} + \\ & p_{\text{scn-3}} \varrho_L^{\text{scn-3}}(t + \Delta t). \end{aligned} \quad (6.37)$$

Inserting the approximations Eqs. (6.19)–(6.20), (6.29)–(6.24), (6.34)–(5.29) and (6.36) into Eq. (6.37) and then taking the limit $\Delta t \rightarrow 0$, we obtain the following effective evolution equation for the logical state $\varrho_L(t)$:

$$\begin{aligned} \dot{\varrho}_L = & -i[h_L(t), \varrho_L] + \Gamma_L [\sigma_x \varrho_L \sigma_x - \varrho_L] + \\ & \sum_{q=1,2,3} \gamma_q \left(1 - p_{\text{misdiag}}^{(X_q)}\right) [V_q \varrho_L V_q^\dagger - \varrho_L]. \end{aligned} \quad (6.38)$$

Here Γ_L is nothing but the logical error rate for quantum memory operations [134, 136]. The initial condition for Eq. (6.38) reads as

$$\varrho_L(0) = |\psi_L(0)\rangle\langle\psi_L(0)| = \begin{bmatrix} |\alpha_L^2(0)| & \alpha_L(0)\beta_L^*(0) \\ \alpha_L^*(0)\beta_L(0) & |\beta_L^2(0)| \end{bmatrix}. \quad (6.39)$$

Equation (6.38) is the main result of this subsection. To the best of our knowledge, the last term at the right-hand side of Eq. (6.38) has not been previously discussed in the context of QEC for quantum simulation or quantum annealing. This term quantifies the logical errors due to spurious coherent evolution in the error subspaces.

The numerical calculations presented at the end of this subsection show that the effective open-system model for the logical qubit [Eq. (6.38)] together with the estimates Eqs. (5.31), (6.28), (6.21) and (6.22) for the parameters Γ_L (logical X error rate), $V_q(t)$ (logical error operation parameterized in terms of error-detection times $t_{\text{det}}^{(q)}$, see Eq. (6.30)) and $p_{\text{misdiag}}^{(X_q)}$ (probability to misdiagnose bit-flip error X_q) provide a good description for the true evolution of the logical state $\varrho_L(t)$ that is encoded into the full system state $\rho(t)$, which evolves according to Eq. (5.18).

Final logical state fidelity

The figure of merit that we aim to maximize under evolution due to a time-dependent Hamiltonian is the final fidelity $\mathcal{F}$ between the target, Eq. (6.14), and the true, Eq. (5.19) logical states, defined as

$$\mathcal{F} = \langle\psi_L(t_{\text{op}})|\varrho_L(t_{\text{op}})|\psi_L(t_{\text{op}})\rangle. \quad (6.40)$$

Note that the target state $\psi_L(t_{\text{op}})$ is exactly equal to the ground state of the problem Hamiltonian only in the adiabatic limit as $t_{\text{op}} \rightarrow \infty$.

Using the effective evolution equation (6.38) for the logical state $\varrho_L(t)$, we can derive the following analytical expression for the final logical state infidelity

$$\begin{aligned} 1 - \mathcal{F} = & \Gamma_L \int_0^{t_{\text{op}}} \left(1 - |\langle\psi_L(t)|\sigma_x|\psi_L(t)\rangle|^2\right) dt + \\ & \sum_{q=1,2,3} \gamma_q (1 - p_{\text{mis}}^{(q)}) \int_0^{t_{\text{op}}} \left(1 - |\langle\psi_L(t)|V_q|\psi_L(t)\rangle|^2\right) dt, \end{aligned} \quad (6.41)$$

which is expressed in terms of the coherent evolution of the target logical state $|\psi_L(t)\rangle$. The first term on the right-hand side of Eq. (6.41) is the usual term in quantum memory, i.e., $\Gamma_L t_{\text{op}}$, generalized here to the case of a finite and time-dependent logical Hamiltonian (6.16). Note that the time integral accounts for the accumulated loss of fidelity due to logical X errors on the time-evolving logical state. The second term is due to the spurious coherent evolution in the error subspaces. Note that this term is positive, i.e., contributes a finite infidelity, because the operator V_q , given in Eq. (6.28), is unitary. Equation (6.41) is the main result of this subsection.

To obtain this result in Eq. (6.41), we have applied the jump/no-jump method in Eq. (6.38) to estimate $\varrho_L(t_{\text{op}})$ as follows:

$$\begin{aligned} \varrho_L(t_{\text{op}}) = & [1 - \Gamma_L t_{\text{op}} - \tilde{\gamma}_{\text{tot}} t_{\text{op}}] |\psi_L(t_{\text{op}})\rangle \langle \psi_L(t_{\text{op}})| + \\ & \Gamma_L \int_0^{t_{\text{op}}} dt' U_L(t', t_{\text{op}}) \sigma_x |\psi_L(t')\rangle \langle \psi_L(t')| \sigma_x U_L^\dagger(t', t_{\text{op}}) + \\ & \sum_q \tilde{\gamma}_q \int_0^{t_{\text{op}}} dt' U_L(t', t_{\text{op}}) V_q |\psi_L(t')\rangle \langle \psi_L(t')| V_q^\dagger U_L^\dagger(t', t_{\text{op}}), \end{aligned} \quad (6.42)$$

where $\tilde{\gamma}_q = \gamma_q(1 - p_{\text{mis}}^{(q)})$, $\tilde{\gamma}_{\text{tot}} = \tilde{\gamma}_1 + \tilde{\gamma}_2 + \tilde{\gamma}_3$ and $U_L(t_1, t_2)$ is the unitary evolution operator associated to the error free Schrödinger evolution equation (6.15). When the jump/no-jump approach is applied to Eq. (6.38), we see that logical errors come in two forms. First, the usual logical X errors that change the logical wavefunction from $|\psi_L(t)\rangle$ to $\sigma_x |\psi_L(t)\rangle$ (second term in Eq. (6.42)). These occur at the logical X error rate Γ_L given in Eq. (5.31). Second, logical errors that are characterized by the logical error operation V_q given in Eq. (6.28) (third term in Eq. (6.42)). This new type of logical errors is specifically due to spurious coherent evolution in the error subspaces. Such errors change the logical wavefunction from $|\psi_L(t)\rangle$ to $V_q |\psi_L(t)\rangle$ and occur at the rate $\tilde{\gamma}_q$. In addition, we also have the coherent no-jump evolution that is described by the unitary evolution operator $U_L(t_1, t_2)$ (first term in Eq. (6.42)). Note that in Eq. (6.42) we have disregarded cases where there are more than one logical error occurrences during the continuous operation duration t_{op} . This approximation is valid in the limit of small bit-flip error rates γ_q that we assume here.

6.4 Optimization and benchmarking

Optimization of the CQEC protocol

In this subsection we derive the optimal parameters (Θ_1^{opt} , Θ_2^{opt} and τ^{opt}) of the CQEC protocol that maximize the final logical state fidelity (6.40). The optimization will be specific to a particular choice of Hamiltonian evolution, i.e., to the choice of $h_L(t)$, since the temporal dependence of $|\psi_L(t)\rangle$ is determined by this. We shall consider here both quantum memory, $h_L(t) = 0$, and quantum annealing with the specific choice of a linear schedule. In the latter case the logical Hamiltonian $h_L(t)$ is given by Eq. (6.16) with the coefficients $a(t)$ and $b(t)$

equal to

$$a(t) = 1 - \frac{t}{t_{\text{op}}} \quad \text{and} \quad b(t) = \frac{t}{t_{\text{op}}}. \quad (6.43)$$

Following convention, we shall refer to the Hamiltonian component multiplied by $b(t)$ as the problem Hamiltonian, i.e., the final Hamiltonian at $t = t_{\text{op}}$. In the context of quantum annealing we shall assume that the adiabatic limit holds, $t_{\text{op}}\Omega_0 \gg 1$, so that we may approximate the target logical wavefunction as the instantaneous ground state, which reads as

$$|\psi_L(t)\rangle = \cos\left(\frac{\theta(t)}{2}\right) |0_L\rangle + \sin\left(\frac{\theta(t)}{2}\right) |1_L\rangle, \quad (6.44)$$

where $\theta(t) = \arctan(a(t)/b(t))$.

Inserting Eq. (6.44) into Eq. (6.41), we obtain for the final logical state infidelity

$$1 - \mathcal{F} = \frac{\Gamma_L t_{\text{op}}}{2} + \sum_{q=1,2,3} \frac{3\pi - 8}{54} \left(1 - p_{\text{mis}}^{(q)}\right) \left[\Omega_0 t_{\text{det}}^{(q)}\right]^2 \gamma_q t_{\text{op}}, \quad (6.45)$$

which is the cost function that we use in the optimization procedure. We emphasize that the result (6.45) applies to the special case of quantum annealing with a linear schedule, and note also that we have included terms up to second order in $\Omega_0 t_{\text{det}}^{(q)}$. The linear dependence of the infidelity Eq. (6.45) on the operation time t_{op} is valid for small infidelities, less than 0.1 for the parameters used in this work. More generally, the final infidelity $1 - \mathcal{F}$ for arbitrary annealing schedule parameters $a(t)$ and $b(t)$ can also be easily obtained, as long as these coefficients also satisfy the adiabatic condition $|\dot{a}(t)|, |\dot{b}(t)| \ll \Omega_0$. This can be accomplished by writing the first integrand of Eq. (6.41) as $1 - |\langle\psi_L(t)|\sigma_x|\psi_L(t)\rangle|^2 = \cos^2(\theta(t))$ and the second integrand as $1 - |\langle\psi_L(t)|V_q|\psi_L(t)\rangle|^2 = \sin^2(\theta(t) - \tilde{\theta}(t)) \sin^2(\tilde{\Omega}(t)t_{\text{det}}^{(q)})$, where $\tilde{\theta}(t) = \arctan(3a(t)/b(t))$ and $\tilde{\Omega}(t) = \Omega_0 \sqrt{a^2(t) + b^2(t)}/9$ is half the instantaneous energy gap of the spurious Hamiltonian (6.17). To obtain a final numerical value for the infidelity, the integrals of Eq. (6.41) would have to be evaluated numerically for evolution under a specific annealing Hamiltonian.

For simplicity, we discuss here the case of equal bit-flip error rates

$$\gamma_1 = \gamma_2 = \gamma_3 = \gamma. \quad (6.46)$$

When $\Omega_0 = 0$ the system operates as a quantum memory. The final logical state infidelity $1 - \mathcal{F}$ is given by the first term of Eq. (6.41) since the second term exactly vanishes. Assuming that the initial logical state is $|\psi_L(0)\rangle = |0_L\rangle$ or $|1_L\rangle$, we find that $1 - \mathcal{F}$ reduces to $\Gamma_L t_{\text{op}}$, because the target logical evolution is trivial in the quantum memory case ($|\psi_L(t)\rangle$ is constant). In addition, we may assume that the operation duration t_{op} is fixed. Then minimization of the final infidelity in quantum memory is equivalent to optimization of the logical X error rate Γ_L in Eq. (5.31), as what we did in Chapter 5.

When $\Omega \neq 0$, this is the case for quantum annealing. We now discuss the results of optimizing the double threshold error detection parameters in this context. To quantify the effectiveness of the CQEC protocol in correcting logical errors, we introduce here the ratio of the infidelity for an unencoded calculation, to the infidelity for an encoded calculation using the optimized double-threshold parameters. This ratio, $R(\gamma, \Omega_0)$, is defined for a given error rate and annealing Hamiltonian, which we shall denote here only by its strength Ω_0 . Specifically,

$$R(\gamma, \Omega_0) = \frac{1 - \mathcal{F}_{\text{unenc}}}{1 - \mathcal{F}_{\text{opt}}}, \quad (6.47)$$

where $\mathcal{F}_{\text{opt}}$ is the value of the final logical state fidelity Eq. (6.45), optimized with respect to τ, Θ_1, Θ_2 (see below), and

$$\mathcal{F}_{\text{unenc}} = \langle \psi_L(t_{\text{op}}) | \rho_{\text{unenc}}(t_{\text{op}}) | \psi_L(t_{\text{op}}) \rangle \quad (6.48)$$

is the fidelity between the final target logical state $|\psi_L(t_{\text{op}})\rangle$ and the final state $\rho_{\text{unenc}}(t_{\text{op}})$ of an unencoded qubit subject to bit-flip errors with rate γ and coherent evolution due to a Hamiltonian $h_L(t)$. We refer to $R(\gamma, \Omega_0)$ as the “reduction factor” of the final logical state infidelity, since by construction it shows by how much the infidelity is reduced by encoding together with optimization of the error detection.

It is easy to see, using the jump/no-jump method, that the unencoded final infidelity $1 - \mathcal{F}_{\text{unenc}}$ can be estimated by the first term of Eq. (6.41) with Γ_L replaced by γ . For the quantum annealing problem with a linear schedule, we obtain

$$1 - \mathcal{F}_{\text{unenc}} = \frac{\gamma t_{\text{op}}}{2}, \quad (6.49)$$

which is linear in the operation time t_{op} for sufficiently small physical error rate γ .

Figure 6.1 shows the dependence of the reduction factor (6.47) on the physical qubit error rate γ . We see that $R(\gamma, \Omega_0)$ increases as γ decreases, saturating at the value R_{plateau} in the limit of small bit-flip error rate γ . This behavior indicates our protocol works best in the case of small bit-flip error rate. This makes sense since now the errors occur further apart and the protocol has sufficient time to detect them out. This plateau value increases with decreasing Ω_0 as follows

$$\begin{aligned} R_{\text{plateau}} &\approx \frac{0.7174}{[\Omega_0 \Gamma_m^{-1} \ln(1.694 \Omega_0 \Gamma_m^{-1})]^2} \quad (\eta = 0.5) \\ &\approx \frac{2.512}{[\Omega_0 \Gamma_m^{-1} \ln(1.289 \Omega_0 \Gamma_m^{-1})]^2} \quad (\eta = 1). \end{aligned} \quad (6.50)$$

The numerical factors and exponents of Eq. (6.50) are obtained from fitting for $\Omega_0 = \Gamma_m/2430, \Gamma_m/810, \Gamma_m/270, \Gamma_m/90, \Gamma_m/30$ and $\Gamma_m/10$. Note that, in the quantum annealing operation considered here, the operation duration t_{op} and the frequency parameter Ω_0 have

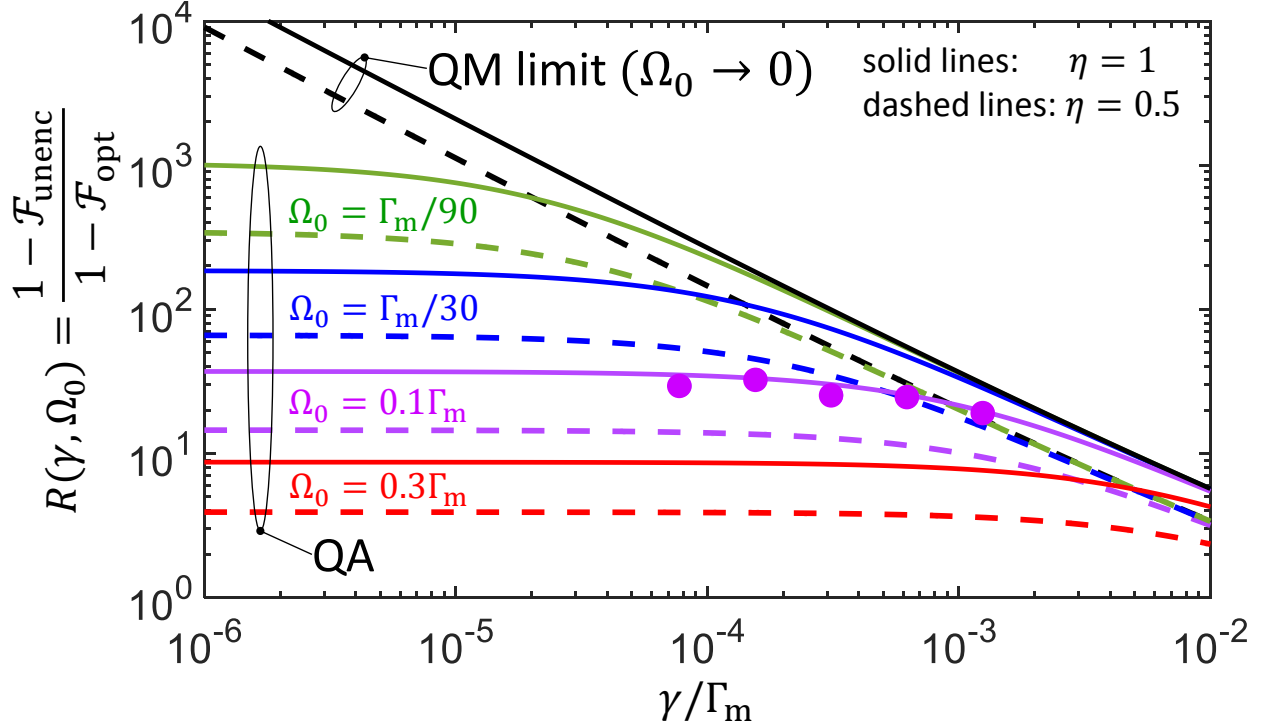


Figure 6.1: Reduction factor $R(\gamma, \Omega_0)$ of the final logical state infidelity due to continuous QEC, Eq. (6.47), for operation time $t_{\text{op}} = 500\Gamma_m$, shown as a function of the bit-flip error rate γ for a range of strengths Ω_0 of the annealing Hamiltonian. Lines present analytic estimates, symbols depict numerical results averaged over an ensemble of 60,000-80,000 realizations of the conditional quantum master equation Eq. (5.18) together with our CQEC protocol. Black lines depict results for the quantum memory (QM) limit, $\Omega_0 \rightarrow 0$, see Eq. (6.16) of the main text. The other lines depict the results for quantum annealing (QA) for $\Omega_0 = \Gamma_m/90$ (green lines), $\Gamma_m/30$ (blue lines), $0.1\Gamma_m$ (purple lines) and $0.3\Gamma_m$ (red lines). The plateau values R_{plateau} are evident on the left hand side of the plot. Solid and dashed lines depict the results for quantum efficiencies $\eta = 1$ and 0.5 , respectively. Γ_m denotes the strength of the continuous measurement.

to satisfy the adiabatic condition, $\Omega_0 t_{\text{op}} \gg 1$, which allowed us to use the instantaneous ground state (6.44) as the target logical state. Assuming that this condition is satisfied, the reduction factor (6.47) of the final logical state infidelity due to CQEC is independent of t_{op} . This monotonic decreasing is expected since smaller Ω_0 means a less harmful spurious evolution.

Finally, we summarize the optimized parameters $\tau_{\text{opt}}, \Theta_1^{\text{opt}}, \Theta_2^{\text{opt}}$ employed in Fig. 6.1. Figure 6.2 depicts the results for the optimal measurement averaging time τ_{opt} that minimizes the logical X error rate Γ_L in the case of quantum memory (black lines) and the final logical state infidelity (6.45) in the case of quantum annealing for $\Omega_0 = \Gamma_m/90$ (green lines),

$\Gamma_m/30$ (blue lines), $0.1\Gamma_m$ (purple lines) and $0.3\Gamma_m$ (red lines). We see that the optimal averaging time parameter τ_{opt} generally increases when the measurement quantum efficiency η decreases, due to the additional noise at the output of the readout signals, see Eq. (5.12) [29]. In the particular case of quantum memory ($\Omega_0 = 0$), we obtain

$$\begin{aligned}\tau_{\text{opt}}^{\text{QM}} &\approx -1.027 \Gamma_m^{-1} \log(9.6955 \gamma \Gamma_m^{-1}) \quad (\eta = 0.5), \\ &\approx -0.5192 \Gamma_m^{-1} \log(5.2891 \gamma \Gamma_m^{-1}) \quad (\eta = 1).\end{aligned}\tag{6.51}$$

The above results are obtained from fitting $\tau_{\text{opt}}^{\text{QM}}$ for the range of error rates indicated in Fig. 6.2. In the case of quantum annealing, for a fixed and finite Ω_0 , the optimal measurement averaging time τ_{opt} tends to increase as the error rate γ is decreased until it reaches a plateau level $\tau_{\text{opt}}^{\text{plateau}}$ that depends on Ω_0 as follows

$$\begin{aligned}\tau_{\text{opt}}^{\text{plateau}} &\approx -1.759 \Gamma_m^{-1} \log(1.3880 \Omega_0 \Gamma_m^{-1}) \quad (\eta = 0.5), \\ &\approx -0.9079 \Gamma_m^{-1} \log(1.2408 \Omega_0 \Gamma_m^{-1}) \quad (\eta = 1).\end{aligned}\tag{6.52}$$

The relations (6.52) are obtained from fitting for $\Omega_0 = \Gamma_m/810, \Gamma_m/270, \Gamma_m/90, \Gamma_m/30$ and $\Gamma_m/10$.

Qualitatively, τ_{opt} increases with decreasing bit-flip rate γ and Hamiltonian frequency Ω_0 . Keeping in mind this parameter is directly related to detection time (Eq. (6.30)), this suggests that when errors do not happen very often and the spurious evolution is less harmful, i.e. smaller Ω_0 , we can allow more detection time to collect more information and make more accurate decisions.

The optimal values for the error threshold parameters (Θ_1^{opt} and Θ_2^{opt}) are found to exhibit only a weak dependence on γ and Ω_0 parameters. For practical purposes these are given here for both the quantum memory and quantum annealing cases by

$$\Theta_1^{\text{opt}} \approx -0.54 \quad \text{and} \quad \Theta_2^{\text{opt}} = 0.8.\tag{6.53}$$

We point out that in our optimization procedure we have imposed two constraints: $-1 \leq \Theta_1 \leq 0$ and $0 \leq \Theta_2 \leq 0.8$. The reason for the constraint on Θ_2 is that our analytical estimates for the logical error rate Γ_L , see Eq. (5.31), and the error detection time t_{det} , see Eq. (6.30), are not accurate when Θ_1 approaches 1. The optimization finds that the optimal position of the upper error threshold should be as close to 1 as it is allowed. If we instead use the constraint $0 \leq \Theta_2 \leq 1$ (with the same previous constraint on Θ_1), the optimization finds that $\Theta_1^{\text{opt}} \approx -0.4$ and $\Theta_2^{\text{opt}} = 1.0$. This indicates that the optimal position of the lower error threshold is robustly located around -0.5 .

Overall performance of the double threshold CQEC protocol

To quantify the effectiveness of the double threshold CQEC protocol in correcting logical errors during the entire continuous operation, we introduce the time-dependent reduction

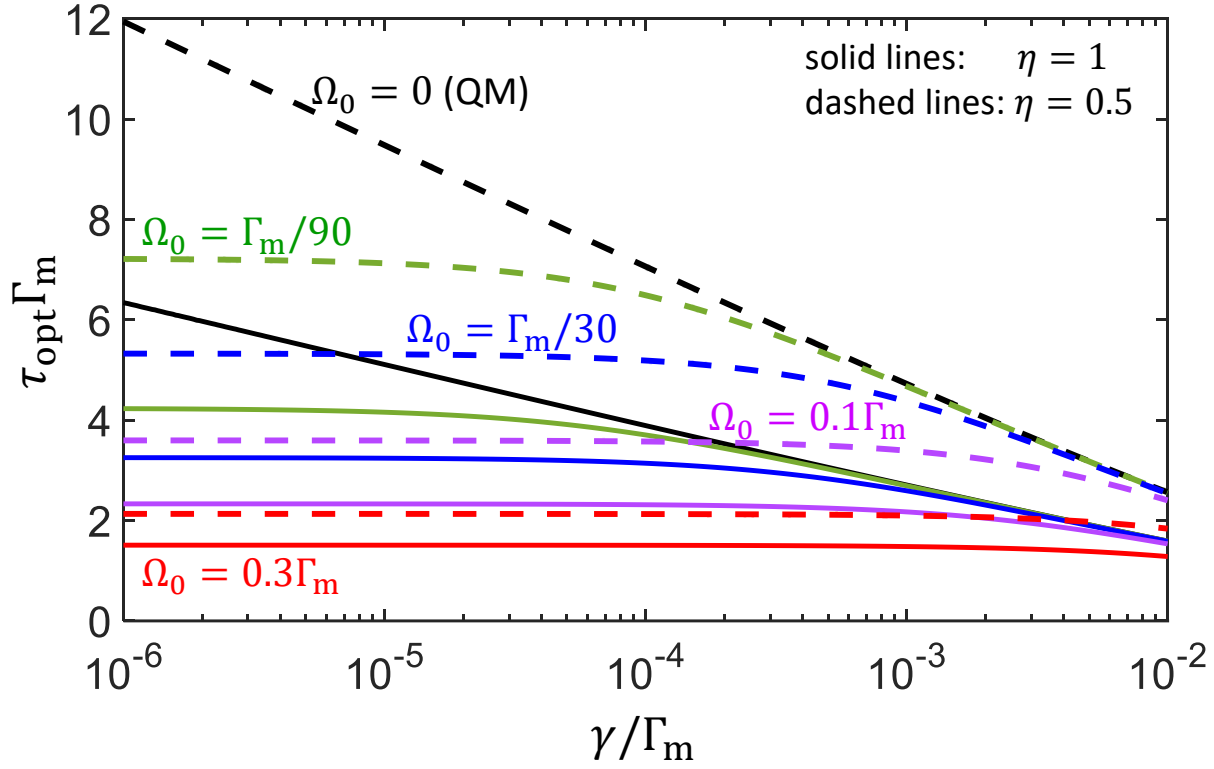


Figure 6.2: Optimal measurement averaging time parameter τ_{opt} . Black solid and dashed lines depict the results for quantum memory (QM), where the frequency parameter Ω_0 vanishes, see Eq. (6.16) of the main text. The other lines depict the results for quantum annealing (QA) with $\Omega_0 = \Gamma_m/90$ (green lines), $\Gamma_m/30$ (blue lines), $0.1\Gamma_m$ (purple lines) and $0.3\Gamma_m$ (red lines). Solid and dashed lines depict the results for quantum efficiencies $\eta = 1$ and 0.5 , respectively. Γ_m denotes the strength of the continuous measurement.

factor R_t of the logical state infidelity. This is defined analogously to Eq. (6.47);

$$R_t = \frac{1 - \mathcal{F}_{unenc.}(t)}{1 - \mathcal{F}_t}, \quad (6.54)$$

where $\mathcal{F}_{unenc.}(t)$ is now the time-dependent unencoded fidelity, defined as in Eq. (6.48) with the operation time t_{op} replaced by $t \in [0, t_{op}]$, and $\mathcal{F}_t$ is the time-dependent logical state fidelity

$$\mathcal{F}_t = \langle \psi_L(t) | \varrho_L(t) | \psi_L(t) \rangle. \quad (6.55)$$

Figure 6.3 shows the time dependence of the logical state infidelity, $1 - \mathcal{F}_t = 1 - \langle \psi_L(t) | \varrho_L(t) | \psi_L(t) \rangle$, obtained using two approaches: “full numerics” and “effective model”.

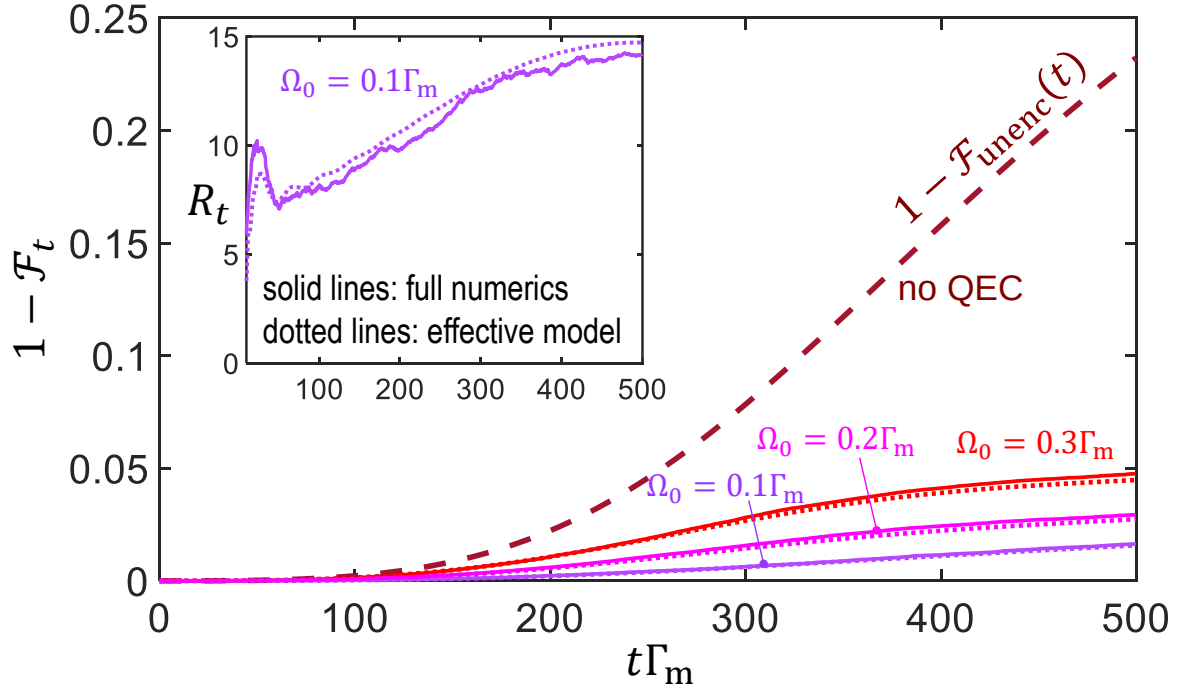


Figure 6.3: Accuracy of the effective evolution equation (6.38) for the logical state infidelity during quantum annealing. The frequency parameter Ω_0 determines the strength of the logical Hamiltonian (6.16), where the coefficients $a(t)$ and $b(t)$ correspond to quantum annealing with linear schedule, see Eq. (6.43). Solid lines: full numerical calculations averaged over an ensemble of 20,000 realizations. Dotted lines: effective model of Eq. (6.38). Dashed line: unencoded qubit infidelity [see Eq. (6.48)] for $\Omega_0 = 0.1\Gamma_m$. Parameters: $\Theta_1 = -0.54$, $\Theta_2 = 0.8$, $\tau = 2.5\Gamma_m^{-1}$, $\eta = 1$, $t_{\text{op}} = 500\Gamma_m^{-1}$ and $\gamma = 1.25 \times 10^{-3}\Gamma_m$.

In the first approach, the logical state $\varrho_L(t)$ is obtained by projecting out the code space components from the full system density matrix $\rho(t)$, where the latter evolves according to the evolution equation (5.18), together with the action of the instantaneous error-correction operations C_{op} [Eq. (5.20)] that are applied to the physical qubits whenever an error is diagnosed by the double threshold CQEC protocol. The ensemble average of Eq. (5.19) is generated over an ensemble of 20,000 realizations. The results of this approach are depicted in Fig. 6.3 by the solid lines, for Hamiltonian strength parameters $\Omega_0 = 0.1\Gamma_m$, $0.2\Gamma_m$ and $0.3\Gamma_m$. The second approach is that of our effective model derived in the previous section. Here the logical state infidelity is obtained from the numerical solution of the effective open-system model given by Eq. (6.38). The results of this approach are depicted in Fig. 6.3 by the dotted lines. The good agreement between the solid and dotted lines in Fig. 6.3 demonstrates that the effective open-system model accurately describes the evolution of the

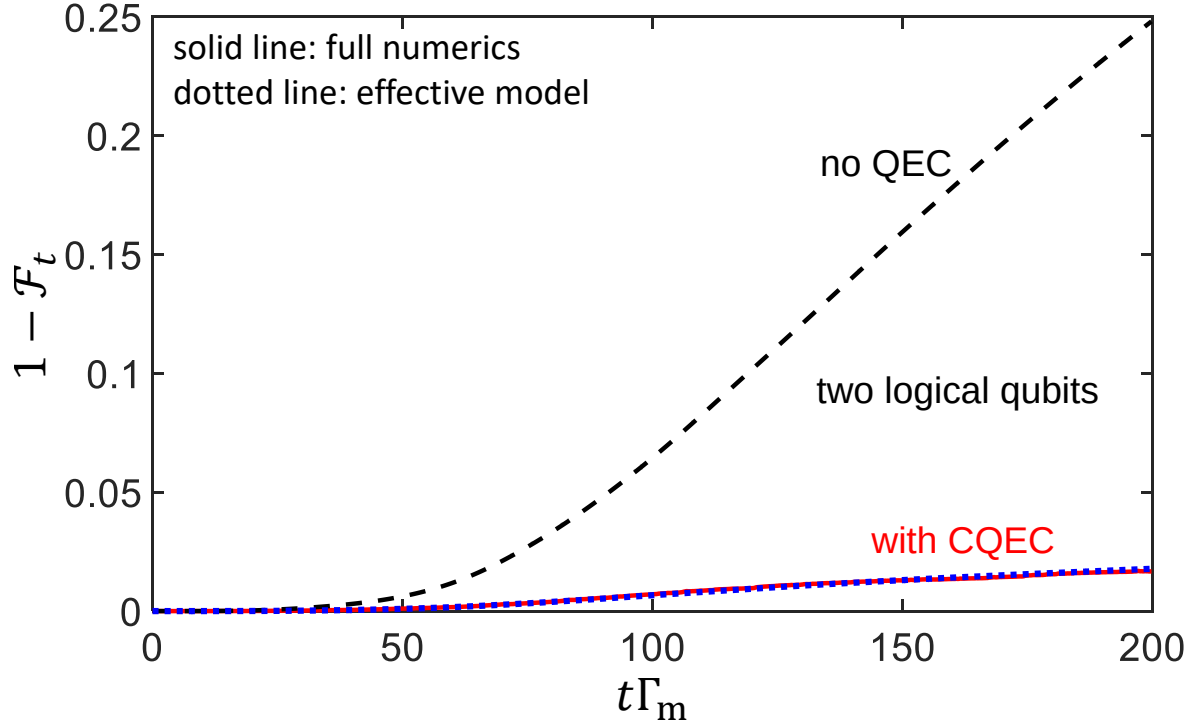


Figure 6.4: Logical state infidelity for two logical qubits. Parameters: $\gamma = 1.25 \times 10^{-3}\Gamma_m$ (error rate of all qubits), $\Omega_0 = 0.1\Gamma_m$ (Hamiltonian strength, see Eq. (6.56)), $\Theta_1 = -0.54$ and $\Theta_2 = 0.8$, $\tau = 2.1\Gamma_m^{-1}$, $\eta = 1$, and Γ_m denotes the strength of continuous measurement. Blue dotted line: effective model of Eq. (6.57). Red solid line: full numerical calculations averaged over an ensemble of 5,000 realizations (see text).

logical qubit during the entire continuous operation. This validates our analysis above for the optimized performance of the double threshold CQEC protocol. The inset of Fig. 6.3 shows the reduction factor R_t for the logical state infidelity during the entire duration of the continuous operation for $\Omega_0 = 0.1\Gamma_m$. Here also, good agreement is found between the full numerics and the effective model approaches. Although in this specific example the reduction factors of the logical state infidelity are modest (varying from 5 to 15), larger reduction factors can be readily achieved with stronger continuous measurements. This can be seen explicitly in Fig. 6.1, where the increase in R is evident for Γ_m larger than $10\Omega_0$.

We now discuss how to generalize the effective open-system model for one logical qubit, Eq. (6.38), to the general case of multiple logical qubits. In this general case, we again have logical errors that come in two forms: logical X errors, and logical errors that are characterized by a logical error operation $V_q^{(l)}$, where q now labels the three physical qubits that encode the l th logical qubit. The logical error operations $V_q^{(l)}$ are again given by Eq. (6.28), where $h_L(t)$ (logical Hamiltonian) and $h_{\text{spurious}}(t)$ (spurious Hamiltonian) are now

specified respectively by the code space diagonal submatrices of the system Hamiltonian $H(t)$ and $X_q H(t) X_q$. Logical X errors acting on the l th logical qubit occur at a rate $\Gamma_L^{(l)}$ that is also given by Eq. (5.31). Note that the set of parameters of the double threshold CQEC protocol (τ , Θ_1 and Θ_2) can differ for different logical qubits, so $\Gamma_L^{(l)}$ may not be the same for all logical qubits. The logical errors that are characterized by $V_q^{(l)}$ occur at the rate $\tilde{\gamma}_q = \gamma_q(1 - p_{\text{misdiag}}^{(X_q)})$ (Eq. (6.38)). This is approximately equal to the bit-flip error rate γ_q of the q th qubit, since the probability $p_{\text{misdiag}}^{(X_q)}$ to misdiagnose the error X_q is typically much smaller than one (see Fig. 5.2). As an example, we consider two logical qubits encoded by the physical qubits $q = 1, 2, 3$ (logical qubit with label $l = 1$) and $q = 4, 5, 6$ (logical qubit with label $l = 2$). Consider the two-qubit logical Hamiltonian

$$\tilde{h}_L(t) = -\Omega_0 \left[a(t) (\sigma_x^{(1)} + \sigma_x^{(2)}) + b(t) (\sigma_z^{(1)} + \sigma_z^{(2)} + \sigma_z^{(1)} \sigma_z^{(2)}) \right], \quad (6.56)$$

where $\sigma_x^{(l)}$ and $\sigma_z^{(l)}$ are the Pauli x and z operators corresponding to the l th logical qubit ($l = 1, 2$), and the quantum annealing coefficients $a(t)$ and $b(t)$ are given in this example by Eq. (6.43). We will assume that the initial condition for the target logical state evolution is $|\psi_L(0)\rangle = (|0_L\rangle + |1_L\rangle) \otimes (|0_L\rangle + |1_L\rangle)/2$. For this example of two logical qubits, the effective open-system model reads as

$$\begin{aligned} \dot{\varrho}_L = & -i[\tilde{h}_L(t), \varrho_L] + \sum_{l=1,2} \Gamma_L^{(l)} [\sigma_x^{(l)} \varrho_L \sigma_x^{(l)} - \varrho_L] + \\ & \sum_{q=1,2,3} \gamma_q [V_q^{(1)} \varrho_L V_q^{(1)\dagger} - \varrho_L] + \sum_{q=4,5,6} \gamma_q [V_q^{(2)} \varrho_L V_q^{(2)\dagger} - \varrho_L], \end{aligned} \quad (6.57)$$

where $V_q^{(1)}$ (for $q = 1, 2, 3$) is obtained from Eq. (6.28) with $h_L(t)$ replaced by $\tilde{h}_L(t)$ and $h_{\text{spurious}}(t)$ now given by Eq. (6.56) with $\sigma_z^{(1)}$ replaced by $\sigma_z^{(1)}/3$. Similarly, $V_q^{(2)}$ (for $q = 4, 5, 6$) is obtained from Eq. (6.28) with $h_L(t)$ also replaced by $\tilde{h}_L(t)$ and $h_{\text{spurious}}(t)$ given by Eq. (6.56) with $\sigma_z^{(2)}$ replaced by $\sigma_z^{(2)}/3$. Figure 6.4 shows that the logical state infidelity obtained from the effective open-system model for two logical qubits [Eq. (6.57), (dotted blue line)] agrees very well with the corresponding infidelity obtained from the full numerical calculations, (solid red line). This indicates that the effective model can be used to accurately estimate and optimize the performance of our CQEC protocol in order to protect the coherent evolution of several logical qubits. Most importantly, both effective model and full numerical calculations show that the CQEC protocol provides a significant reduction in the final state infidelity by a factor of ~ 14 relative to the value obtained without error correction. For the two-logical qubit Hamiltonian considered here, Eq. (6.56), this reduction is similar to that obtained for the corresponding single logical qubit in Fig. 6.3. However, in general, different values of $R(\gamma, \Omega_0)$ may be found, since the reduction also depends on the form of the coupling between the logical qubits. Thus if the coupling term is changed, the logical

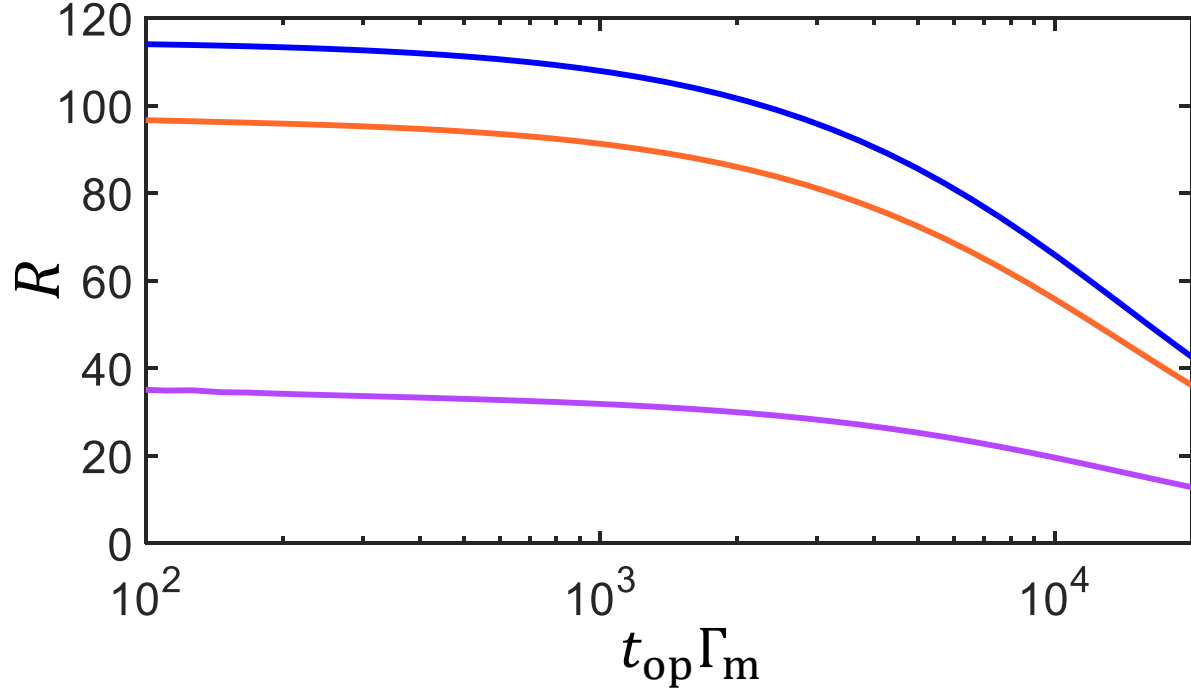


Figure 6.5: Reduction factor R of the final logical state infidelity as a function of the operation time t_{op} for single logical qubit annealing under CQEC. Blue, orange and purple lines depict the reduction factors corresponding to the cases where the error correction operations are given by $\tilde{C}_{\text{op}}(t)$ [Eq. (6.58)], $\tilde{C}_{\text{op}}^{\text{approx}}(t)$ [Eq. (6.59)] and $C_{\text{op}} = X_q$, respectively. Parameters: $\gamma = 1.25 \times 10^{-4} \Gamma_m$, $\Omega_0 = 0.1 \Gamma_m$, $\tau = 2.5 \Gamma_m^{-1}$, $\Theta_1 = -0.54$, $\Theta_2 = 0.8$, and $\eta = 1$.

state infidelity can change since both $\tilde{h}_L(t)$ and the $V_q^{(i)}$ terms in Eq. (6.57) are dependent on this coupling (see Eq. (6.28)). For example, if the sign of the $\sigma_z^{(1)} \sigma_z^{(2)}$ term in Eq. (6.56) is flipped, then the final time reduction factor R is close to 6 instead of 14.

Modified logical error operations

Our analysis of logical errors due to spurious coherent evolution in the error subspaces during error diagnosis (scenario 2 in Sec. 6.3) employed the conventional error correction operators of the three-qubit bit flip code as the correction operators, i.e., $C_{\text{op}} = X_q$ with $X_q = X_1, X_2$ or X_3 (see Eq. (6.26)). It is possible to further increase the reduction factors R of the final logical state infidelity by using the following modified error correction operations:

$$\tilde{C}_{\text{op}}(t) = \exp \left(-it_{\text{det}}^{(q)} H(t) \right) \exp \left(it_{\text{det}}^{(q)} X_q H(t) X_q \right) X_q, \quad (6.58)$$

where t is the time moment when the error X_q is detected, and $t_{\text{det}}^{(q)}$ is the time that it takes to diagnose such error. To analyze the performance of the modified error correction operations $\tilde{C}_{\text{op}}(t)$, we can use the effective evolution equation (6.38) for the logical state without the term describing the logical errors due to spurious evolution in the error subspaces; i.e., without the last term at the right-hand-side of Eq. (6.38). This can be shown by carrying out the analysis of Sec. 6.3 with error correction operations given by Eq. (6.58). We emphasize that spurious evolution is not fully compensated by using the modified error correction operations in an experimental setup because the error-detection times $t_{\text{det}}^{(q)}$ are random.

From the solution of Eq. (6.38) without the last term at its right-hand-side, we have estimated the reduction factor R for the final logical state infidelity that corresponds to our CQEC operation with error correction operations given by Eq. (6.58). The blue line in Fig. 6.5 shows our results for this situation and for a bit-flip error rate $\gamma = 1.25 \times 10^{-4} \Gamma_{\text{m}}$. By comparing the blue and the purple lines in Fig. 6.5, we see that, by using the modified error correction operations, R increases by a factor of $\simeq 3.3$ with respect to the case of CQEC with conventional error correction operations, $C_{\text{op}} = X_1, X_2$ or X_3 . This improvement factor in R depends on the physical error rate γ ; for instance, for $\gamma = 1.25 \times 10^{-3} \Gamma_{\text{m}}$ we obtain that R is increased by a smaller factor $\simeq 1.7$ by using the modified error correction operations. We have also calculated the reduction factor R from (computationally expensive) full quantum trajectory simulations of Eq. (5.18) with the double-thresholding error detection protocol and error correction operations given by Eq. (6.58), where $t_{\text{det}}^{(q)}$ is given by Eq. (6.30). Our results show that the reduction factors indicated by the blue line in Fig. 6.5 are somewhat larger than those from full trajectory simulations by $\simeq 15\%$, for operation times t_{op} in the range $100 \Gamma_{\text{m}}^{-1}$ to $1000 \Gamma_{\text{m}}^{-1}$, which is the region where the reduction factor is roughly independent of the operation time, see Fig. 6.5.

The modified error correction operations $\tilde{C}_{\text{op}}(t)$, Eq. (6.58), may be challenging to implement, so we have also analyzed the performance of the following error correction operations that can be regarded as an approximate version of $\tilde{C}_{\text{op}}(t)$:

$$\tilde{C}_{\text{op}}^{\text{approx}}(t) = \exp \left(i t_{\text{det}}^{(q)} (X_q H(t) X_q - H(t)) \right) X_q. \quad (6.59)$$

We note that the exponent in Eq. (6.59) is $2\Omega_0 b(t) Z_q / 3$ for the Hamiltonian given in Eq. (6.11). If we carry out the analysis of Sec. 6.3 with error correction operations given by Eq. (6.59), we arrive to an effective evolution equation for the logical state that is similar to Eq. (6.38) with the operator $V_q(t)$ replaced by

$$\begin{aligned} \tilde{V}_q^{\text{approx}}(t) = & \exp \left(i t_{\text{det}}^{(q)} h_{\text{L}}(t) \right) \times \\ & \exp \left(i t_{\text{det}}^{(q)} (h_{\text{spurious}}(t) - h_{\text{L}}(t)) \right) \exp \left(-i t_{\text{det}}^{(q)} h_{\text{spurious}}(t) \right). \end{aligned} \quad (6.60)$$

From this effective evolution equation for the logical state we have estimated the performance of our CQEC operation with error correction operations given by Eq. (6.59). The reduction factors in this case are somewhat smaller than those corresponding to the case of using error

correction operations given Eq. (6.58), see orange and blue lines in Fig. 6.5, respectively. The main advantage of using $\tilde{C}_{\text{op}}^{\text{approx}}(t)$ as the error corrections operations is that they can be simpler to implement.

Time-to-solution metric

The reduction factor expresses the decrease in error probability due to quantum error correction. This is the metric of success that is used for analysis of performance of error correcting codes in quantum computation, and can be also be used for quantum simulations. The cost of CQEC will depend on the specific application. Since quantum annealing is a heuristic analog approach that is not guaranteed to find the global minimum of the final Hamiltonian in a single run and instead finds it with a less than unit success probability p_s , it is conventional to measure the cost of quantum annealing on a specific device by measuring the time to achieve a desired success probability at least once during multiple runs, defined as the time-to-solution (TTS), often also taking the parallelization possible on a given device into account [151]. This TTS metric, which is applied to both algorithmic and error mitigation protocols, reflects a trade-off between a high single-run success probability with a long operation time, and a low single-run success probability with a short operation time [152]. It is generally defined for quantum annealing as

$$\text{TTS}(t_{\text{op}}) = t_{\text{op}} \mathcal{R}_s(t_{\text{op}}) \frac{N}{N_{\text{max}}} \quad (6.61)$$

with

$$\mathcal{R}_s(t_{\text{op}}) = \frac{\log(1 - p_d)}{\log(1 - p_s(t_{\text{op}}))}. \quad (6.62)$$

This refers just to the pure annealing time and not to the wall-clock time that also contains the setup, cooling and readout times on a real device. In Eq. (6.62), p_d is the desired success probability to get the solution at least once, $p_s(t_{\text{op}})$ is the success probability of a single run, N is the required number of qubits in a single run, and N_{max} is the total available number of qubits in the device. $\mathcal{R}_s$ is the number of total runs needed to get the solution at least once with desired probability p_d and the factor N/N_{max} takes into account the possible reduction by parallelization of runs. A separate analysis applies to the time for readout and state preparation. Comparing the TTS for our CQEC protocol with a classical protection strategy utilizing parallelization alone shows that CQEC provides a smaller TTS for the quantum annealing dynamics than does the classical strategy. This improvement over the classical strategy is noteworthy since at small encodings such as the three-qubit code employed here, proposals for suppression of errors during quantum annealing using stabilizer encodings with penalty Hamiltonians have shown that these perform worse than the classical parallelization scheme [153].

We present here analysis of the application of our CQEC protocol to quantum annealing with the time-to-solution (TTS) metric. In particular, we show how the value of this metric

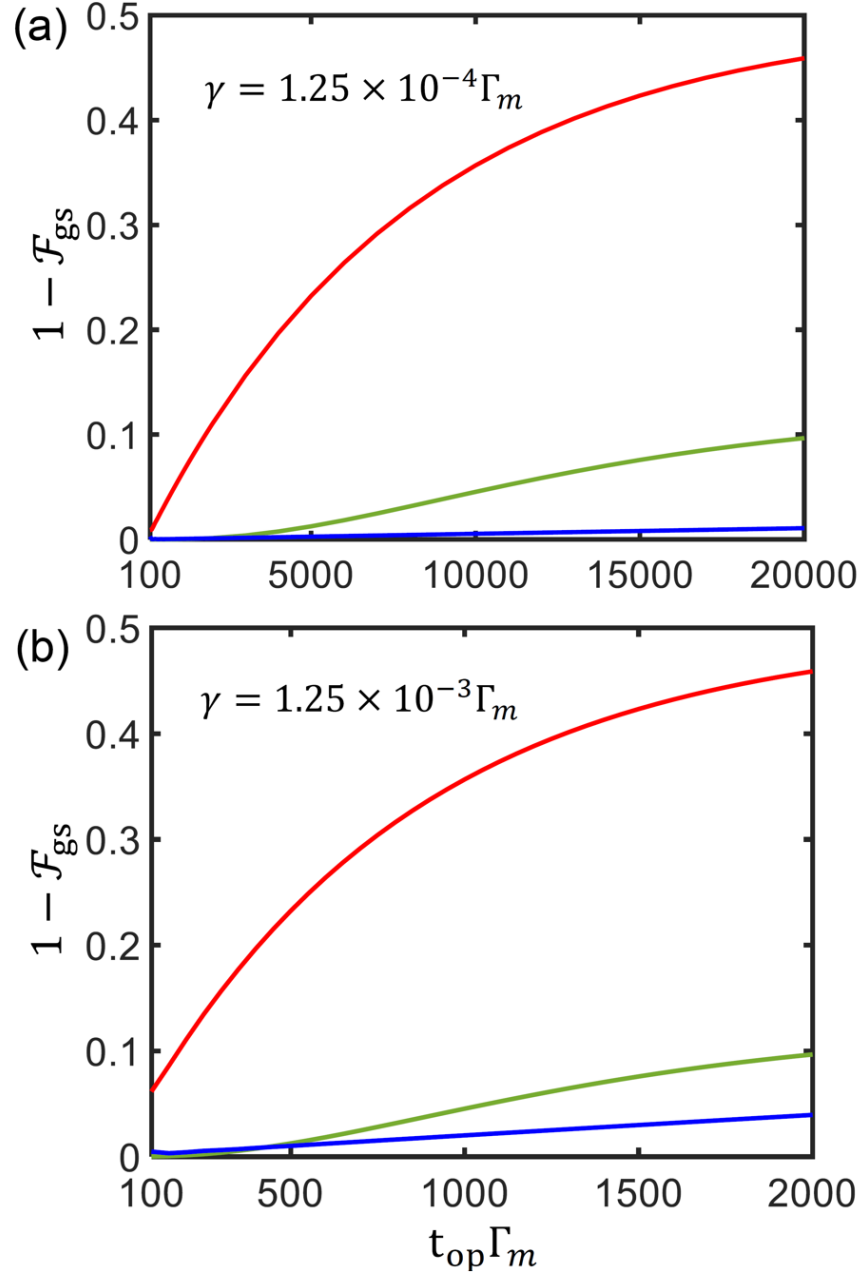


Figure 6.6: Final infidelity with respect to the ground state of the problem Hamiltonian for one qubit annealing as a function of operation time t_{op} . Panel (a): $\gamma = 1.25 \times 10^{-4} \Gamma_m$. Panel (b): $\gamma = 1.25 \times 10^{-3} \Gamma_m$. Red line: one unencoded physical qubit, no error correction. Green line: classic protection scheme of Ref. [153] (see text). Blue line: one logical qubit with CQEC and modified error correction, Eq. (6.58). Parameter values: $\eta = 1$, $\Omega = 0.1 \Gamma_m$, $\theta_1 = -0.54$, $\theta_2 = 0.8$, and $\tau = 2.5 \Gamma_m^{-1}$.

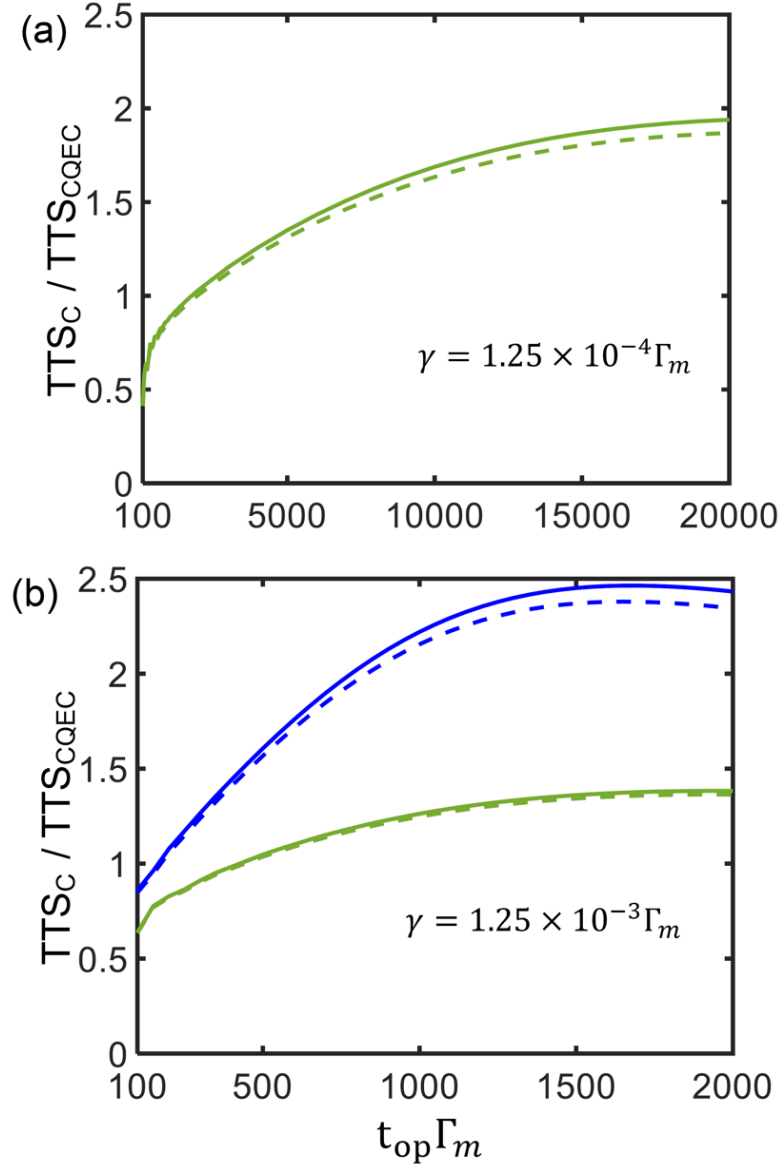


Figure 6.7: Ratio of the TTS value for the classical (C) strategy to that for CQEC, as a function of t_{op} . Panel (a): one logical qubit, $\gamma = 1.25 \times 10^{-4} \Gamma_m$. Panel (b): one and two logical qubits, $\gamma = 1.25 \times 10^{-3} \Gamma_m$. Green lines are for one logical qubit, with error correction operations in the CQEC calculations given by Eq. (6.58) (solid lines) and Eq. (6.59) (dashed lines). Blue lines are for two logical qubits, with error correction operations in the CQEC calculations given by Eq. (6.58) (solid line) and Eq. (6.59) (dashed line). Parameters: one logical qubit calculations use the same values as in Fig. 6.6, two logical qubit calculations employ $\tau = 2.1 \Gamma_m^{-1}$, with all other parameter values the same as for one logical qubit.

depends on the annealing operation time t_{op} . We compare the performance of our CQEC protocol with regard to TTS with that of the so-called classical (C) strategy of Ref. [153], in which annealing is carried out on m unencoded copies in parallel and the infidelity estimated from the binomial distribution function $(1 - p_s^{\text{unenc}}(t_{\text{op}}))^m$, with $p_s^{\text{unenc}}(t_{\text{op}})$ the probability of successfully reaching the ground state of the problem Hamiltonian, i.e., the final Hamiltonian at $t = t_{\text{op}}$, for a single unencoded qubit (see Ref. [153]). We use the example of single qubit quantum annealing from Section 6.4 for this detailed analysis, before summarizing the performance of both single qubit and two-qubit quantum annealing problem at a particular value of t_{op} .

Figure 6.6 shows how the final CQEC infidelity (blue lines, with modified logical error correction operations, see Eq. (6.58)) depends on the operation time t_{op} for single qubit quantum annealing, and compares this with the corresponding infidelities obtained for the unencoded dynamics (red lines) and the classical (C) strategy (green lines). Note that here we are plotting the final fidelity with respect to the ground state of the problem Hamiltonian. The two panels show results for two different physical error rates γ , with the value of the product γt_{op} for the maximal value of t_{op} held constant at 2.5. This corresponds to the same overall average number of errors at the largest shown value of t_{op} in both panels. The calculations related to the blue lines in Fig. 6.6 were carried out using the effective evolution equation Eq. (6.38) without the last term at its right-hand-side. The accuracy of this was confirmed by full quantum trajectory simulation of Eq. (5.18) together with the double-thresholding error detection protocol and error correction operations given by Eq. (6.58). For example, for $\gamma = 1.25 \times 10^{-4} \Gamma_m$ and $t_{\text{op}} = 10^4 \Gamma_m^{-1}$, we obtain infidelities of 6.6×10^{-3} from the quantum trajectory simulations (averaging over 1000 trajectories) and 5.14×10^{-3} from the effective evolution equation (6.38), respectively. The unencoded calculations (red lines) are simulated directly from Eq. (5.18) without the measurement terms, and the classical strategy values (green lines) are obtained as the third power of the unencoded values.

It is evident that CQEC performs better than both the unencoded annealing and the classical strategy at longer operation times. The final infidelity of unencoded annealing of the single qubit tends towards that of the completely mixed state on the time scales shown here, consistent with the lack of any error correction, while that of the classical strategy tends to the corresponding binomial function of this. In contrast, once the operation time is long enough that non-adiabatic errors are small, the CQEC strategy is successful in correcting the bit-flip errors due to coupling to the environment and can achieve significantly lower infidelities.

The data in Fig. 6.6 are then used to evaluate the TTS for single qubit annealing, according to Eqs. (6.61)–(6.62) in the main text. We have chosen the target success probability $p_d = 0.99$. Figure 6.7 shows how the ratio of the TTS value for the classical strategy to that for the CQEC strategy varies with operation time t_{op} , for a given value of physical error rate γ . At the longer operation times in the time scale range shown here, this TTS ratio is larger than 1, indicating that the classical strategy will require a longer time to reach the desired ground state solution for the annealing problem. In terms of this TTS metric, the CQEC strategy thus clearly outperforms the classical strategy at the longer operation times

t_{op} that suppress non-adiabatic errors and are preferred for quantum annealing.

To illustrate the performance on larger systems, we have also calculated the corresponding ratio of TTS values for quantum annealing of two logical qubits, using the effective two qubit model Eq. (6.57). The blue lines in Fig. 6.7 panel (b) show the resulting ratio $\text{TTS}_C/\text{TTS}_{\text{CQEC}}$ as a function of t_{op} . Comparing with the green lines for one logical qubit, it is evident that in terms of the TTS metric, the CQEC protocol performs increasingly better than the classic strategy when scaled up to two qubits. This is encouraging for the use of a modular form of CQEC for larger systems.

6.5 Discussion and conclusions

We have developed a continuous quantum error correction protocol for operation under both quantum memory and evolution under a time-dependent Hamiltonian. We illustrated the approach with a detailed analysis of the continuous operation performance of the three-qubit bit flip code, which is designed to preserve the coherent evolution of the logical qubits against decoherence from bit-flip errors. Error detection is carried out using a relatively simple and nearly optimal protocol that consists of filtering (time-averaging) the noisy bare readout signals and using a double thresholding scheme to diagnose the error syndrome in real time. In addition, immediately after diagnosing an error, discrete (i.e., instantaneous) error correction operations are applied to the physical qubits, as in the conventional code operation. We have shown that this combination of continuous detection of errors in real time with discrete correction of errors is very effective and yields, e.g., in the case of quantum memory operation, a logical X error rate that exhibits a nearly quadratic scaling on the physical qubit error rate and has a magnitude that is slightly larger than the logical X error rate of the linear variant of the optimal Wonham filter [134]. The advantage of our double threshold CQEC protocol is that it can be simpler to implement.

Spurious coherent evolution of the system state in the error subspaces [140], due to a (time-dependent) encoded Hamiltonian, leads to a new type of logical errors, for which we have found the corresponding effective Kraus logical error operators, $V_q(t)$, that act on the instantaneous logical state. The Kraus logical error operator $V_q(t)$ is parametrized by the time $t_{\text{det}}^{(q)}$ that the CQEC protocol takes to detect the error X_q , see Eq. (6.28). The time $t_{\text{det}}^{(q)}$ should be as small as possible in order to minimize the detrimental effect of logical errors due to spurious evolution on the performance of the double threshold CQEC protocol. For this protocol, $t_{\text{det}}^{(q)}$ is estimated to be proportional to the averaging time parameter τ (Eq. (6.30)), which, however, cannot be arbitrarily small without degrading the performance of the CQEC protocol to correctly diagnose single bit-flip errors X_1 , X_2 or X_3 .

We have developed an effective open-system model for the logical qubit state [see Eq. (6.38)] that accounts for the two types of logical errors that are relevant for, e.g., quantum simulation and quantum annealing applications: logical errors due to spurious coherent evolution in the error subspaces and the usual logical X errors of quantum memory operation. This effective model is very useful because it allows us to readily estimate and optimize the performance of

the double threshold CQEC protocol without performing computationally expensive numerical calculations on the full encoding qubit system (full numerics). We have shown that the effective model accurately describes the actual logical state during the continuous operation, see Fig. 6.3. In addition, we have discussed how to generalize the effective model for multiple logical qubits, where we have again found excellent agreement with the more cumbersome and computationally expensive full numerics approach, see Fig. 6.4.

Using the effective open-system model for one logical qubit, we have analyzed the performance of the double threshold CQEC protocol to preserve the coherent evolution of the logical qubit due to a quantum-annealing type Hamiltonian with a linear schedule. We have introduced the reduction factor R of the final logical state infidelity, see Eq. (6.47), as a measure of the performance of the CQEC protocol. The performance depends on the relative magnitudes of three problem-specific parameters; namely, the bit-flip error rate, γ , the strength of the logical Hamiltonian, Ω_0 , and the strength of the continuous measurements of the code stabilizer generators, characterized by the measurement strength parameter Γ_m .

For a given magnitude of the error rate γ , the reduction factor R increases as the magnitude of Ω_0 decreases, which our analysis shows is due to the fact that a reduction in Ω_0 causes the contribution of spurious coherent evolution to logical errors to decrease. In contrast, for a given ratio Ω_0/Γ_m , the reduction factor R increases as we decrease the physical qubit error rate γ , until R reaches a plateau level R_{plateau} . This reduction factor is the analog of the reduction of error probability obtained with discrete operation of quantum error correcting codes, which becomes increasingly effective as the physical qubit error rate decreases [154]. The plateau level depends on the relative magnitude of the logical Hamiltonian strength Ω_0 and the strength of continuous measurements Γ_m , as is evident in Fig. 6.1. For instance, we obtain $R_{\text{plateau}} \approx 37, 184$ or 1002 for measurement strengths $\Gamma_m = 10\Omega_0, 30\Omega_0$ or $90\Omega_0$, respectively, assuming that continuous measurements are performed by ideal detectors ($\eta = 1$). These reduction factors become $R_{\text{plateau}} \approx 15, 66$ or 340 , respectively, if the measurement efficiency is $\eta = 0.5$ (nonideal detectors).

Overall, our results reveal that the region of optimal performance is defined by the following hierarchy of timescales:

$$\frac{10}{\Gamma_m} \leq \Omega_0^{-1} \ll t_{\text{op}}$$

For example, for measurement rate $\Gamma_m = (10 \text{ ns})^{-1} \approx 2\pi \times 16 \text{ MHz}$, effective correction of bit flip errors can be achieved for annealing Hamiltonians of strength $\Omega_0 = 2\pi \times 1.6 \text{ MHz}$ and operation time $t_{\text{op}} \lesssim \gamma^{-1}$, with bit flip error rates $\gamma = (10 \mu\text{s})^{-1} \approx 2\pi \times 16 \text{ kHz}$. Such decoherence rates are the same order of magnitude as thermal relaxation rates of quantum annealers operating at temperatures of 8-20 mK [155–157]. In this case, the reduction factors R are approximately equal to 10 and 20 (see purple lines in Fig. 6.1) for measurement quantum efficiencies $\eta = 0.5$ and 1, respectively.

It is useful to consider here the feasibility of parity measurements on these timescales. Parity measurement with superconducting qubits [158] can be realized by dispersively coupling two qubits with a single resonator (cavity) such that their interaction Hamiltonian is

$H_{\text{int}} = \chi(Z_1 + Z_2)\hat{n}$, where χ is the dispersive coupling parameter and $\hat{n}$ is the intracavity photon number. In this setup, a superposition of odd-parity states $|01\rangle$ and $|10\rangle$ does not decohere due to measurement, because they do not shift the cavity resonance frequency. In contrast, a superposition of even-parity states $|00\rangle$ and $|11\rangle$ exhibits decoherence due to measurement at a rate that is proportional to $(\kappa/\chi)^2$ when $\chi \gg \kappa$. In this situation the odd parity states show an identical high response and the even parity states an identical low response to the cavity drive. To use this setup for parity measurement, it is required that $\chi \gg \kappa$ and the cavity is driven at a frequency equal to the average of the cavity resonance frequencies corresponding to the four two-qubit states: $|c, d\rangle$ with $c, d = \{0, 1\}$ [120]. The measurement rate is then $\Gamma_{\text{m}} \approx \kappa \bar{n}_{01}/2$, where $\bar{n}_{01}$ is the average number of intracavity photons when the two-qubit state is $|01\rangle$. This measurement rate dictates the exponential decay of the off-diagonal matrix elements of the two-qubit density matrix connecting states of even and odd parities; i.e., $\langle \rho_{00,01}(t) \rangle \sim e^{-\Gamma_{\text{m}} t}$. Now, assuming realistic parameter values of $\kappa/2\pi = 3 \text{ MHz}$ and $\bar{n}_{01} = 10$, we find that $\Gamma_{\text{m}} \approx (10.6 \text{ ns})^{-1}$.

The CQEC approach developed in this work is applicable to a wide range of analog quantum simulations, allowing the encoded logical quantum states to be efficiently corrected as they evolve under a general time-varying Hamiltonian and in the presence of environmental decoherence. The reduction factor R_t measures the overall effectiveness and success of the CQEC strategy under a given Hamiltonian evolution, while the quadratic scaling of the logical error rate Γ_{L} with physical qubit error rate γ guarantees that the strategy is correcting single physical qubit errors. As in discrete quantum error correction, this active correction of errors for greater fidelity of quantum simulation comes with an overhead of additional qubits for the logical encoding. However, unlike discrete protocols, we do not require additional ancilla qubits and entangling operations to transfer the logical state information for measurement.

Overall, the high level of reduction of infidelity achieved by this CQEC for a small three-qubit code requiring overhead of just two additional qubits in both quantum memory and quantum simulation under a time-evolving Hamiltonian is excellent. The reduction in time-to-solution relative to classical parallelization for quantum annealing shows the benefits of this in application to analog quantum simulations. In principle, one can expect even better performance with larger encodings. Indeed, the continuous time quantum error correction protocol presented in this work can be readily applied to any subspace stabilizer QEC code, such as the three-qubit repetition code studied here, and can also be extended to subsystem stabilizer codes [137]. In practice, the limitations will be achieving continuous measurement of multi-qubit stabilizer operators. For larger systems, subsystem codes [159] such as the Bacon-Shor codes [160, 161] or generalizations of this [162, 163] would be preferred since these require measurement of only two-body operators.

An important direction for further work is to apply the CQEC protocol to other error models. Clearly arbitrary single qubit errors can be corrected using this approach with larger stabilizer codes. Of particular interest for quantum annealing is correction of thermal errors. The bit-flip code alone is not effective here, since it cannot correct X and Y errors at the same time. However, active continuous correction of thermal errors can be achieved

by implementing the present CQEC protocol in a adiabatic frame and combining this with error suppression techniques in which an energy penalty consisting of the negative of the bit flip code stabilizer operators is added to the time-dependent Hamiltonian [164].

More generally, one would like to develop CQEC protocols for architecture-specific errors, such as biased noise. In the future, developing error correction diagnostics for physical errors encountered in realistic devices may be assisted by the use of machine learning techniques [165] or filters for non-Markovian noise [166]. In addition, experimental implementations frequently see drift of the key parameters such as the measurement rate Γ_m and efficiency η as well as slow temporal variations of the offset of the measurement signals. Exploring the use of machine learning techniques to track these parameters and adjust the CQEC protocol accordingly in real time during an experiment would be a useful direction for further work. Another important direction for further study based on CQEC is the extension of this approach to fault tolerant error correction [161, 167, 168].

The favorable performance of the CQEC protocol seen for the quantum annealing application presented here, in particular the lack of any significant decrease in performance going from one to two logical qubits, indicates the potential viability of modular approaches to quantum error correction for quantum simulation and for quantum annealing in particular. For quantum computation and simulation on near term quantum machines, it is advantageous to use encodings that generate only low weight logical operators, while also requiring only low weight measurement operators. Since the weight of the logical operators of stabilizer codes, whether subspace or subsystem, always grow with the number of encoding qubits, small codes are therefore highly attractive from this perspective. Indeed, quantum annealing Hamiltonians of the Ising spin glass form, i.e., containing only terms of the form $H_Z = \sum_i h_i^z \sigma_i^z$, $H_X = \sum_i h_i^x \sigma_i^x$, $H_{ZZ} = \sum_{ij} J_{ij} \sigma_i^z \sigma_j^z$, that are encoded with the three-qubit stabilizer code result in logical operator terms of only weight two and three. The three-qubit code thus presents an attractive modular option for implementing error correction of quantum annealing with large numbers of logical qubits.

Chapter 7

Conclusion and outlook

In this thesis, we have developed and demonstrated two applications of continuous measurement-based feedback control to assisting quantum computation. In chapter 4, we proposed the PaQS protocol for steering states of remote qubits and applied this to generate the highly symmetric multi-qubit Dicke and GHZ states. The symmetries appear to be essential to guarantee the performance. It is interesting to ask what other entangled states could be generated using PaQS? One good class of candidates is the cluster states. They have high symmetries and are widely used in measurement-based quantum computing [169]. Choosing the proper measurement operator with as low weight as possible may be a potential challenge for this problem. Looking at the protocol from another perspective, a broader question can also be asked—what are the possible states that can be generated from joint measurement and local controls without any non-local Hamiltonian interactions, and with what fidelity can they be generated? This is a hard theoretical question but it is very relevant if we want to achieve large scale distributed quantum computation. Another interesting question for PaQS is that while the PaQS protocol can be treated as a heuristic for finding pathways to entangled states and provides local optimality, we do not currently know whether the PaQS protocol is globally optimal in any situation. Motivating or showing global optimality from PaQS is thus another useful direction to explore.

In chapters 5 and 6, we developed a practical continuous error correction scheme to protect quantum computation from the noisy environment. Approaches to suppress the decoherence are essential and critical for realization of a large-scale fault tolerant quantum computer. In this thesis we demonstrated the effectiveness of our protocol against bit-flip errors. The next step is to consider more realistic error channels like thermal noise. In fact, this is a project we are currently working on. For this we again encode the qubits with the three-qubit repetition code to do continuous error correction, but now we also add a penalty Hamiltonian which separates the code spaces from the error subspaces. Promising preliminary results have been produced. Another problem we need to address is how to effectively protect the coherent evolution when we have strong noise. The bit-flip code works in the limit $t_{\text{det}}\gamma \ll 1$ (See Eq. (6.30)). This leads to another on-going project in the Whaley group, which is investigating how to use machine learning techniques to fully harness

the information in the continuous signal and thereby possibly shorten the detection time. Reducing the detection time is one way to raise the upper bound for the effectively correctable error rate. Another way to combat the effects of noise is to employ larger encodings such as the $[[11, 1, 5]]$ code [170] or surface code [171], which encode more qubits at the cost of measuring more stabilizers. Testing the performance of the double-threshold protocol using these encodings is a good next step for this work.

Bibliography

- [1] Y. Nakamura, Y. A. Pashkin, and J. S. Tsai. “Coherent control of macroscopic quantum states in a single-Cooper-pair box”. In: *Nature* 398 (1999).
- [2] D. Vion et al. “Manipulating the Quantum State of an Electrical Circuit”. In: *Science* 296 (2002), pp. 886–889.
- [3] A. Blais et al. “Cavity quantum electrodynamics for superconducting electrical circuits: An architecture for quantum computation”. In: *Physical Review A* 69 (2004), p. 062320.
- [4] J. Koch et al. “Charge-insensitive qubit design derived from the Cooper pair box”. In: *Physical Review A* 76 (2007), p. 042319.
- [5] M. Kjaergaard et al. “Superconducting Qubits: Current State of Play”. In: *Annual Reviews of Condensed Matter Physics* 11 (2020), pp. 369–395.
- [6] J. Preskill. “Quantum Computing in the NISQ era and beyond”. In: *Quantum* 2 (2018), p. 79.
- [7] K. W. Murch, a. C. M. S. J. Weber, and I. Siddiqi. “Observing single quantum trajectories of a superconducting quantum bit”. In: *Nature (London)* 502 (2013), p. 211.
- [8] S. Zhang, L. S. Martin, and K. B. Whaley. “Locally optimal measurement-based quantum feedback with application to multi-qubit entanglement generation”. In: *arXiv:1807.02029. To be published in Physical Review A* (2018).
- [9] J. Atalaya et al. “Continuous quantum error correction for evolution under time-dependent Hamiltonians”. In: *arXiv:2003.11248v2* (2020).
- [10] U. Vool and M. H. Devoret. “Introduction to Quantum Electromagnetic Circuits”. In: *arXiv:1610.03438v2* (2017).
- [11] A. Blais et al. “Circuit quantum electrodynamics”. In: *arXiv:2005.12667* (2020).
- [12] J. M. Martinis and K. Osborne. “Superconducting Qubits and the Physics of Josephson Junctions”. In: *arXiv:cond-mat/0402415* (2004).
- [13] B. W. Shore and P. L. Knight. “The Jaynes-Cummings Model”. In: *Journal of Modern Optics* 40 (1993), pp. 1195–1238.

- [14] M. S. Fox. *Quantum Optics: An Introduction*. Oxford University Press Inc., New York, 2006.
- [15] S. E. Nigg et al. “Black-box superconducting circuit quantization”. In: *Physical Review Letter* 108 (2012), p. 240502.
- [16] E. A. Sete, J. M. Gambetta, and A. N. Korotkov. “Purcell effect with microwave drive: Suppression of qubit relaxation rate”. In: *Physical Review B* 89 (2014), p. 104516.
- [17] K. W. Murch et al. “Observing single quantum trajectories of a superconducting quantum bit”. In: *Nature* 502 (2013), p. 211.
- [18] J. Gambetta et al. “Quantum trajectory approach to circuit QED: Quantum jumps and the Zeno effect”. In: *Physical Review A* 77.1 (2008), p. 012112.
- [19] J. M. Chow. “Quantum Information Processing with Superconducting Qubits”. In: *Ph.D. thesis, Yale University* (2010).
- [20] J. von Neumann. *Mathematical Foundations of Quantum Mechanics*. Princeton University Press, Princeton, New Jersey, 1955.
- [21] M. A. Nielsen and I. L. Chuang. “Quantum Computation and Quantum Information, (Cambridge University Press, Cambridge)”. In: (2000).
- [22] V. B. Braginsky and F. Y. Khalili. *Quantum Measurement*. Cambridge University Press, Cambridge, 1992.
- [23] H.-P. Breuer and F. Petruccione. *The theory of open quantum systems*. Oxford University Press, 2002.
- [24] K. Jacob and D. A. Steck. “A straightforward introduction to continuous quantum measurement”. In: *Contemporary Physics* 47.5 (2006), pp. 279–303.
- [25] C. W. Gardiner and P. Zoller. *Quantum Noise 3rd edn*. Springer Berlin Heidelberg New York, 2004.
- [26] L. Martin et al. “Deterministic generation of remote entanglement with active quantum feedback”. In: *Physical Review A* 92.6 (2015), p. 062321.
- [27] P. Rouchon and J. F. Ralph. “Efficient quantum filtering for quantum feedback control”. In: *Physical Review A* 91 (2015), p. 012118.
- [28] A. N. Korotkov. “Continuous quantum measurement of a double dot”. In: *Physical Review B* 60 (1999), p. 5737.
- [29] A. Korotkov. “Selective quantum evolution of a qubit state due to continuous measurement”. In: *Physical Review B* 63 (2001), p. 115403.
- [30] A. N. Korotkov. “Nonideal quantum detectors in Bayesian formalism”. In: *Physical Review B* 67 (2003), p. 235408.
- [31] A. N. Korotkov. “Quantum Bayesian approach to circuit QED measurement”. In: *arXiv:1111.4016* (2011), p. 11.

- [32] P. E. Kloeden and E. Platen. “Numerical Solution of Stochastic Differential Equations”. In: *Springer-Verlag Berlin, Heidelberg* (1992).
- [33] V. P. Belavkin. “Quantum continual measurements and a posteriori collapse on CCR”. In: *Communications in Mathematical Physics* 146 (1992), pp. 611–635.
- [34] V. P. Belavikin. “Quantum stochastic calculus and quantum nonlinear filtering”. In: *Journal of Multivariate Analysis* 42 (2002), pp. 171–202.
- [35] L. Bouten, R. van Handel, and M. R. James. “An introduction to quantum filtering”. In: *SIAM Journal on Control and Optimization* 46 (2007), pp. 2199–2241.
- [36] A. C. Doherty and K. Jacobs. “Feedback control of quantum systems using continuous state estimation”. In: *Physical Review A* 60 (1999), p. 2700.
- [37] H. M. Wiseman, S. Mancini, and J. Wang. “Bayesian feedback versus Markovian feedback in a two-level atom”. In: *Physical Review A* 66 (2002), p. 013807.
- [38] G. J. M. H. M. Wiseman. “Quantum theory of optical feedback via homodyne detection”. In: *Physical Review Letter* 70 (1993), pp. 548–551.
- [39] H. M. Wiseman and G. J. Milburn. “Quantum theory of field-quadrature measurements”. In: *Physical Review A* 47 (1993), pp. 642–662.
- [40] H. M. Wiseman. “Quantum theory of continuous feedback”. In: *Physical Review A* 49 (1994), pp. 2133–2150.
- [41] A. K. Ekert. “Quantum cryptography based on Bell’s theorem”. In: *Physical Review Letter* 67.6 (1991), pp. 661–663.
- [42] H. J. Kimble. “The quantum internet”. In: *Nature* 453.7198 (2008), pp. 1023–1030.
- [43] G. Wendin. “Quantum information processing with superconducting circuits: a Review”. In: *Reports on Progress in Physics* 80 (2016), p. 106001.
- [44] D. J. Wineland et al. “Squeezed atomic states and projection noise in spectroscopy”. In: *Physical Review A* 50.1 (1994), p. 67.
- [45] M. Kitagawa and M. Ueda. “Squeezed spin states”. In: *Physical Review A* 47.6 (1993), p. 5138.
- [46] O. Hosten et al. “Measurement noise 100 times lower than the quantum-projection limit using entangled atoms”. In: *Nature* 529.7587 (2016), p. 505.
- [47] I. Bloch, J. Dalibard, and S. Nascimbene. “Quantum simulations with ultracold quantum gases”. In: *Nature Physics* 8.4 (2012), p. 267.
- [48] M. H. Devoret and R. J. Schoelkopf. “Superconducting circuits for quantum information: an outlook”. In: *Science* 339.6124 (2013), pp. 1169–1174.
- [49] N. Nickerson, Y. Li, and S. Benjamin. “Topological quantum computing with a very noisy network and local error rates approaching one percent”. In: *Nature Communications*. 4 (2013), p. 1756.

- [50] R. Hanson et al. “Coherent dynamics of a single spin interacting with an adjustable spin bath”. In: *Science* 320.5874 (2008), pp. 352–355.
- [51] J. K. Stockton, R. van Handel, and H. Mabuchi. “Deterministic Dicke-state preparation with continuous measurement and control”. In: *Physical Review A* 70.2 (2004), p. 022106.
- [52] K. C. Cox et al. “Deterministic squeezed states with collective measurements and feedback”. In: *Physical Review Letter* 116.9 (2016), p. 093602.
- [53] C. W. Chou et al. “Measurement-induced entanglement for excitation stored in remote atomic ensembles”. In: *Nature* 438.7069 (2005), pp. 828–832.
- [54] D. L. Moehring et al. “Entanglement of single-atom quantum bits at a distance”. In: *Nature* 449.7158 (2007), pp. 68–71.
- [55] J. Hofmann et al. “Heralded Entanglement Between Widely Separated Atoms”. In: *Science* 337 (2012), pp. 72–75.
- [56] H. Bernien et al. “Heralded entanglement between solid-state qubits separated by three metres”. In: *Nature* 497 (2013), pp. 86–90.
- [57] N. Roch et al. “Observation of Measurement-Induced Entanglement and Quantum Trajectories of Remote Superconducting Qubits”. In: *Physical Review Letter* 112.17 (Apr. 2014), p. 170501.
- [58] T. Takano et al. “Spin squeezing of a cold atomic ensemble with the nuclear spin of one-half”. In: *Physical Review Letter* 102.3 (2009), p. 033601.
- [59] We note that for two qubits, entangled states can be produced with unit probability by measurements alone when full parity measurements are used. In this case, feedback provides additional speedup of the state generation [60, 81].
- [60] L. Martin, M. Sayrafi, and K. B. Whaley. “What is the optimal way to prepare a Bell state using measurement and feedback?”. In: *Quantum Science and Technology* 2.4 (2017), p. 044006.
- [61] D. Ristè et al. “Deterministic entanglement of superconducting qubits by parity measurement and feedback”. In: *Nature* 502 (2013), p. 350.
- [62] R. Van Handel, J. K. Stockton, and H. Mabuchi. “Feedback control of quantum state reduction”. In: *IEEE Transactions on Automatic Control* 50.6 (2005), pp. 768–780.
- [63] M. Mirrahimi and R. Van Handel. “Stabilizing feedback controls for quantum systems”. In: *SIAM Journal on Control and Optimization* 46.2 (2007), pp. 445–467.
- [64] K. Tsumura. “Global stabilization at arbitrary eigenstates of N-dimensional quantum spin systems via continuous feedback”. In: *American Control Conference, 2008. IEEE. 2008*, pp. 4148–4153.
- [65] L. Thomsen, S. Mancini, and H. Wiseman. “Continuous quantum nondemolition feedback and unconditional atomic spin squeezing”. In: *Journal of Physics B: Atomic, Molecular and Optical Physics* 35.23 (2002), p. 4937.

- [66] J. Wei et al. “Deterministic generation of symmetric multi-qubit Dicke states: an application of quantum feedback control”. In: *IET Control Theory & Applications* 9.17 (2015), pp. 2500–2505.
- [67] R. Vijay, D. H. Slichter, and I. Siddiqi. “Observation of quantum jumps in a superconducting artificial atom”. In: *Physical Review Letter* 106.11 (2011), p. 110502.
- [68] R. Vijay et al. “Stabilizing Rabi oscillations in a superconducting qubit using quantum feedback”. In: *Nature* 490 (2012), p. 77.
- [69] S. Hacothen-Gourgy et al. “Quantum dynamics of simultaneously measured non-commuting observables”. In: *Nature (London)* 538 (2016), p. 491.
- [70] R. Horodecki et al. “Quantum entanglement”. In: *Review of Modern Physics* 81.2 (2009), p. 865.
- [71] W. Dür, G. Vidal, and J. I. Cirac. “Three qubits can be entangled in two inequivalent ways”. In: *Physical Review A* 62 (2000), p. 062314.
- [72] S. Hill and W. K. Wootters. “Entanglement of a Pair of Quantum Bits”. In: *Physical Review Letter* 78.26 (1997), p. 5022.
- [73] V. Goffman, J. Kundu, and W. K. Wootters. “Distributed entanglement”. In: *Physical Review A* 61 (2000), p. 052306.
- [74] W. K. Wootters. “The rebit three-tangle and its relation to two-qubit entanglement”. In: *Journal of Physics A: Mathematical and Theoretical* 47 (2014), p. 424037.
- [75] C. H. Bennett et al. “Purification of Noisy Entanglement and Faithful Teleportation via Noisy Channels”. In: *Physical Review Letter* 76.722 (1996).
- [76] A. Miyake and M. Wadati. “Multipartite Entanglement and Hyperdeterminants”. In: *Quantum Information and Computation* 2 (special (2002), pp. 540–555.
- [77] H. M. Wiseman and G. J. Milburn. *Quantum measurement and control*. Cambridge University Press, 2009.
- [78] B. Øksendal. *Stochastic Differential Equations*. Springer-Verlag, 2003.
- [79] H. M. Wiseman. “Adaptive phase measurements of optical modes: going beyond the marginal Q distribution”. In: *Physical Review Letter* 75 (1995), p. 4587.
- [80] K. Jacobs. “How to project qubits faster using quantum feedback”. In: *Physical Review A* 67.3 (2003), 030301(R).
- [81] C. Hill and J. Ralph. “Weak measurement and control of entanglement generation”. In: *Physical Review A* 77 (2008), p. 014305.
- [82] L. S. Martin and K. B. Whaley. “Single-shot deterministic entanglement between non-interacting systems with linear optics”. In: *arXiv preprint arXiv:1912.00067* (2019).
- [83] D. Gottesman, T. Jennewein, and S. Croke. “Longer-baseline telescopes using quantum repeaters”. In: *Physical review letters* 109.7 (2012), p. 070503.

- [84] C. Gross. “Spin squeezing, entanglement and quantum metrology with Bose–Einstein condensates”. In: *Journal of Physics B: Atomic, Molecular and Optical Physics* 45.10 (2012), p. 103001.
- [85] J. K. Stockton et al. “Characterizing the entanglement of symmetric many-particle spin-1 2 systems”. In: *Physical Review A* 67.2 (2003), p. 022112.
- [86] N. Gisin and H. Bechmann-Pasquinucci. “Bell inequality, Bell states and maximally entangled states for n qubits”. In: *Physics Letter A* 246.1-2 (1998), pp. 1–6.
- [87] P. D. Johnson, F. Ticozzi, and L. Viola. “Exact stabilization of entangled states in finite time by dissipative quantum circuits”. In: *Physical Review A* 96.1 (2017), p. 012308.
- [88] F. Ticozzi et al. “Alternating Projections Methods for Discrete-Time Stabilization of Quantum States”. In: *IEEE Transactions on Automatic Control* 63.3 (2018), pp. 819–826.
- [89] G. Cardona, A. Sarlette, and P. Rouchon. “Exponential Stochastic Stabilization of a Two-Level Quantum System via Strict Lyapunov Control”. In: *IEEE Conference on Decision and Control (CDC), Miami Beach, FL* (2018), pp. 6591–6596.
- [90] V. Coffman, J. Kundu, and W. K. Wootters. “Distributed entanglement”. In: *Physical Review A* 61.5 (2000), p. 052306.
- [91] A. Sudbery. “On local invariants of pure three-qubit states”. In: *Journal of Physics A: Mathematical and General* 34.3 (2001), p. 643.
- [92] A. Acín et al. “Generalized Schmidt decomposition and classification of three-quantum-bit states”. In: *Physical Review Letter* 85.7 (2000), p. 1560.
- [93] K. Cao et al. “Efficient numerical method to calculate the three-tangle of mixed states”. In: *Physical Review A* 81.3 (2010), p. 034302.
- [94] D. A. Lidar, I. L. Chuang, and K. B. Whaley. “Decoherence-Free Subspaces for Quantum Computation”. In: *Physical Review Letter* 81 (1998), p. 2594.
- [95] D. A. Lidar and K. B. Whaley. “Decoherence-free Subspaces and Subsystems”. In: *Irreversible Quantum Dynamics, F. Benatti and R. Floreanini (Eds.), Springer Lecture Notes in Physics vol. 622, Berlin* (2003), pp. 83–120.
- [96] S. P. Jordan, E. Farhi, and P. W. Shor. “Error correcting codes for adiabatic quantum computation”. In: *Physical Review A* 74 (2006), p. 052322.
- [97] A. D. Bookatz, E. Farhi, and L. Zhou. “Error suppression in Hamiltonian-based quantum computation using energy penalties”. In: *Physical Review A* 92 (2015), p. 022317.
- [98] M. Marvian and D. A. Lidar. “Error suppression in Hamiltonian-based quantum computation using energy penalties”. In: *Physical Review A* 95 (2017), p. 032302.
- [99] L. Viola, E. Knill, and S. Lloyd. “Dynamical Decoupling of Open Quantum Systems”. In: *Physical Review Letter* 82 (1999), p. 2417.

- [100] S. Pasini and G. S. Uhrig. “Optimized dynamical decoupling for power-law noise spectra”. In: *Physical Review A* 81 (2010), p. 012309.
- [101] D. A. Lidar. “Review of Decoherence-Free Subspaces, Noiseless Subsystems, and Dynamical Decoupling”. In: *Advances in Chemical Physics* 154 (2014), p. 295.
- [102] D. J. Gorman, K. C. Young, and K. B. Whaley. “Overcoming dephasing noise with robust optimal control”. In: *Physical Review A* 86 (2012), p. 012317.
- [103] D. Gottesman. “An introduction to to quantum error correction”. In: *Proceedings of Symposia in Applied Mathematics* 58 (2002), p. 221.
- [104] S. J. Devitt, W. J. Munro, and K. Nemoto. “Quantum error correction for beginners”. In: *Rep. Prog. Physical* 76 (2013), p. 076001.
- [105] D. A. Lidar and T. A. Brun. “Quantum Error Correction”. In: *Cambridge University Press* (2013).
- [106] B. M. Terhal. “Quantum error correction for quantum memories”. In: *Review of Modern Physics* 87 (2015), p. 307.
- [107] P. W. Shor. “Fault-Tolerant Quantum Computation”. In: *Proceedings 37th Symposium on the Foundations of Computer Science, IEEE Computer Society Press* 56 (1996).
- [108] P. W. Shor. “Scheme for reducing decoherence in quantum computer memory”. In: *Physical Review A* 52 (1995), R2493(R).
- [109] A. M. Steane. “Error Correcting Codes in Quantum Theory”. In: *Physical Review Letter* 77 (1996), p. 793.
- [110] D. Gottesman. “Class of quantum error-correcting codes saturating the quantum Hamming bound”. In: *Physical Review A* 54 (1996), p. 1862.
- [111] K. Kraus. “States, Effects, and Operations: Fundamental Notions of Quantum Theory, (Springer, Berlin)”. In: (1983).
- [112] H. J. Carmichael. “An Open System Approach to Quantum Optics, Lecture Notes in Physics, (Springer, Berlin)”. In: (1993).
- [113] A. N. Korotkov. “Quantum Bayesian approach to circuit QED measurement with moderate bandwidth”. In: *Physical Review A* 94 (2016), p. 042326.
- [114] H. M. Wiseman and G. J. Milburn. “Quantum Measurement and Control, (Cambridge University Press, Cambridge, UK)”. In: (2010).
- [115] J. Dalibard, Y. Castin, and K. Mølmer. “Wave-function approach to dissipative processes in quantum optics”. In: *Physical Review Letter* 68 (1992), p. 580.
- [116] L. Diósi. “Continuous quantum measurement and itô formalism”. In: *Physical Letter A* 129 (1988), p. 419.
- [117] N. K. et al. “Coherent state evolution in a superconducting qubit from partial-collapse measurement”. In: *Science* 312 (2006), p. 1498.

- [118] A. Palacios-Laloy et al. “Experimental violation of a Bell’s inequality in time with weak measurement”. In: *Nature Physics* 6 (2010), p. 442.
- [119] M. H. et al. “Quantum Back-Action of an Individual Variable-Strength Measurement”. In: *Science* 339 (2013), p. 178.
- [120] D. Risté et al. “Deterministic entanglement of superconducting qubits by parity measurement and feedback”. In: *Nature (London)* 502 (2013), p. 350.
- [121] Q. Ficheux et al. “Dynamics of a qubit while simultaneously monitoring its relaxation and dephasing”. In: *Nature Communications*. 9 (2018), p. 1926.
- [122] C. K. Anderson et al. “Entanglement stabilization using ancilla-based parity detection and real-time feedback in superconducting circuits”. In: *npj Quantum Information* 5 (2019), p. 69.
- [123] J. P. Paz and W. H. Zurek. “Continuous error correction”. In: *Proceedings of the Royal Society of London A* 454 (1998), p. 355.
- [124] C. Ahn, A. C. Doherty, and A. J. Landahl. “Continuous quantum error correction via quantum feedback control”. In: *Physical Review A* 65 (2002), p. 042301.
- [125] C. Ahn, H. M. Wiseman, and G. J. Milburn. “Quantum error correction for continuously detected errors”. In: *Physical Review A* 67 (2003), p. 052310.
- [126] C. Ahn, H. M. Wiseman, and K. Jacobs. “Quantum error correction for continuously detected errors with any number of error channels per qubit”. In: *Physical Review A* 70 (2004), p. 024302.
- [127] M. Sarovar et al. “Practical scheme for error control using feedback”. In: *Physical Review A* 69 (2004), p. 052324.
- [128] M. Sarovar and G. J. Milburn. “Continuous quantum error correction by cooling”. In: *Physical Review A* 72 (2005), p. 012306.
- [129] O. Oreshkov and T. A. Brun. “Continuous quantum error correction for non-Markovian decoherence”. In: *Physical Review A* 76 (2007), p. 022318.
- [130] B. A. Chase, A. J. Landahl, and J. M. Geremia. “Efficient feedback controllers for continuous-time quantum error correction”. In: *Physical Review A* 77 (2008), p. 032304.
- [131] R. van Handel and H. Mabuchi. “Optimal error tracking via quantum coding and continuous syndrome measurement”. In: *arXiv:0511221* (2005).
- [132] H. Mabuchi. “Continuous quantum error correction as classical hybrid control”. In: *New Journal of Physics* 11 (2009), p. 105044.
- [133] K. C. Hsu and T. A. Brun. “Method for quantum-jump continuous-time quantum error correction”. In: *Physical Review A* 93 (2016), p. 022321.
- [134] R. Mohseninia et al. “Always-On Quantum Error Tracking with Continuous Parity Measurements”. In: *arXiv:1907.08882* (2019).

- [135] G. Cardona, A. Sarlette, and P. Rouchon. “Continuous-time quantum error correction with noise-assisted quantum feedback”. In: *IFAC-PapersOnLine*, 52 (2019), p. 198.
- [136] J. Atalaya et al. “Bacon-Shor code with continuous measurement of noncommuting operators”. In: *Physical Review A* 95 (2017), p. 032317.
- [137] A. N. K. J. Atalaya and K. B. Whaley. “Error correcting Bacon-Shor code with continuous measurement of noncommuting operators”. In: *arXiv:1910.08272* (2019).
- [138] T. Albash and D. A. Lidar. “Adiabatic quantum computation”. In: *Review of Modern Physics* 90 (2018), p. 015002.
- [139] I. M. Georgescu, S. Ashhab, and F. Nori. “Quantum simulation”. In: *Review of Modern Physics* 86 (2014), p. 153.
- [140] K. C. Young, M. Sarovar, and R. Blume-Kohout. “Error suppression and error correction in adiabatic quantum computation: Techniques and challenges”. In: *Physical Review X* 3 (2013), p. 041013.
- [141] D. Gottesman. “Stabilizer Codes and Quantum Error Correction”. In: *Ph.D. thesis, Caltech* (1997).
- [142] T. Brun. “Quantum Error Correction”. In: *arXiv:1910.03672* (2019).
- [143] B. C. H. et al. “Mixed-state entanglement and quantum error correction”. In: *Physical Review A* 54 (1996), p. 3824.
- [144] K. E. and L. R. “Theory of quantum error-correcting codes”. In: *Physical Review A* 55 (1997), pp. 900–911.
- [145] P. Horowitz and W. Hill. In: *The Art of Electronics (Cambridge University Press, Cambridge)* (2015).
- [146] A. Elgart and G. A. Hagedorn. “A note on the switching adiabatic theorem”. In: *Journal of Mathematical Physics* 53 (2012), p. 102202.
- [147] A. Messiah. *Quantum Mechanics, Vol. II*. North-Holland Publishing Company, Amsterdam, 1962.
- [148] D. Comparat. “General conditions for quantum adiabatic evolution”. In: *Physical Review A* 80 (2009), p. 012106.
- [149] M. H. S. Amin. “Consistency of the Adiabatic Theorem”. In: *Physical Review Letter* 102 (2009), p. 220401.
- [150] D. H. Slichter et al. “Quantum Zeno effect in the strong measurement regime of circuit quantum electrodynamics”. In: *New Journal of Physics* 18 (2016), p. 053031.
- [151] T. F. Rønnow et al. “Defining and detecting quantum speedup”. In: *Science* 345 (2014), p. 420.
- [152] T. Albash and D. A. Lidar. “Demonstration of a scaling advantage for a quantum annealer over simulated annealing”. In: *Physical Review X* 8 (2018), p. 031016.

- [153] K. L. Pudenz, T. Albash, and D. A. Lidar. “Error-corrected quantum annealing with hundreds of qubits”. In: *Nature Communications* 5 (2014), p. 1.
- [154] A. M. Steane. “A tutorial on quantum error correction”. In: *Proceedings of the International School of Physics Enrico Fermi, IOS press* 162 (2007), p. 1.
- [155] A. D. K. et al. “Observation of topological phenomena in a programmable lattice of 1,800 qubits”. In: *Nature* 560 (2018), p. 456.
- [156] S. Novikov et al. “Exploring More-Coherent Quantum Annealing”. In: *IEEE International Conference on Rebooting Computing (ICRC)* 1 (2018).
- [157] S. W. et al. “Coherent coupled qubits for quantum annealing”. In: *Physical Review Applied* 8 (2017), p. 014004.
- [158] D. P. DiVincenzo and F. Solgun. “Multi-qubit parity measurement in circuit quantum electrodynamics”. In: *New Journal of Physics* 15 (2013), p. 075001.
- [159] D. Poulin. “Stabilizer Formalism for Operator Quantum Error Correction”. In: *Physical Review Letter* 95 (2005), p. 230504.
- [160] D. Bacon. “Operator quantum error-correcting subsystems for self-correcting quantum memories”. In: *Physical Review A* 73 (2006), p. 012340.
- [161] P. Aliferis and A. W. Cross. “Subsystem fault tolerance with the Bacon-Shor code”. In: *Physical Review Letter* 98 (2007), p. 220502.
- [162] S. Bravyi. “Subsystem codes with spatially local generators”. In: *Physical Review A* 83 (2011), p. 012320.
- [163] M. Marvian and S. Lloyd. “Robust universal Hamiltonian quantum computing using two-body interactions”. In: *arXiv:1911.01354* (2019).
- [164] J. Atalaya and K. B. Whaley. In: *in preparation* (2020).
- [165] T. F. sel et al. “Reinforcement learning with neural networks for quantum feedback”. In: *Physical Review X* 8 (2018), p. 031084.
- [166] M. Tsang. “Volterra filters for quantum estimation and detection”. In: *Physical Review A* 92 (2015), p. 062119.
- [167] N. M. Linke et al. “Fault-tolerant quantum error detection”. In: *Science Advances*, 3 (2017), e1701074.
- [168] M. Li, D. Miller, and K. R. Brown. “Direct measurement of Bacon-Shor code stabilizers”. In: *Physical Review A* 98 (2018), 050301(R).
- [169] H. J. Briegel et al. “Measurement-based quantum computation”. In: *Nature Physics* 5 (2009), pp. 19–26.
- [170] A. R. Calderbank et al. “Quantum error correction via codes over $GF(4)$ ”. In: *arXiv:quant-ph/9608006* (1996).

- [171] A. G. Fowler et al. “Surface codes: Towards practical large-scale quantum computation”. In: *Physical Review A* 86 (2012), p. 032324.