# Device- and Application-Adapted Quantum Error Correction

by

David Layden

B.S. Mathematical Physics, University of Waterloo (2014)
MMath Applied Mathematics, University of Waterloo (2016)

Submitted to the Department of Nuclear Science and Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Quantum Science and Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2020



Author .....
Department of Nuclear Science and Engineering
May 20, 2020

Certified by .....
Paola Cappellaro
Professor of Nuclear Science and Engineering
Thesis Supervisor

Certified by .....
William D. Oliver
Associate Professor of Electrical Engineering and Computer Science
Thesis Reader

Accepted by .....
Ju Li
Battelle Energy Alliance Professor of Nuclear Science and Engineering
Professor of Materials Science and Engineering
Chair, Department Committee on Graduate Students

**Device- and Application-Adapted Quantum Error Correction**

by

David Layden

Submitted to the Department of Nuclear Science and Engineering
on May 20, 2020, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Quantum Science and Engineering

## Abstract

Precise control of coherent quantum systems could enable new generations of sensing, communication and computing technologies. Such systems, however, are typically noisy and difficult to stabilize. One promising technique to this end is called quantum error correction, which encodes quantum states in such a way that errors can be detected and corrected, much like in classical error-correcting codes.

Quantum error-correcting codes usually cast a wide net, in that they are designed to correct errors regardless of their physical origins. In large-scale devices, this is an essential feature. It comes at a cost, however: conventional quantum codes are typically resource-intensive in terms of both the system size and the control operations they require. Yet, in smaller-scale devices the main error sources are often well-understood. In the near term, it may therefore be advantageous to cast a more targeted net through specialized codes.

This thesis presents new families of such quantum error-correcting codes, which are adapted either for leading candidate devices, or for near-term applications. The device-adapted codes require exponentially less overhead than conventional codes to achieve the same level of protection, whereas the application-adapted codes can enhance quantum sensors, in which conventional codes cannot readily be used.

The new techniques presented in this thesis adapt cornerstones of conventional theory in light of key experimental challenges and opportunities. The ultimate goal of this research is to help bridge the gap between the exacting requirements of proposed quantum technologies and the realities of emerging quantum devices. Bridging this gap is critical, if quantum technologies are to realize their full potential.

Thesis Supervisor: Paola Cappellaro
Title: Professor of Nuclear Science and Engineering

# Acknowledgments

This thesis is the culmination of a long process with many critical junctures. It is humbling to think of how strongly this process was shaped by the people surrounding me. Their contributions—big and small—compounded and intertwined so thoroughly that this end result would be inconceivable without them. I would therefore like to express my gratitude to a number of people while also stressing that this list is necessarily incomplete, both in breadth and in depth.

First, I wish to thank my supervisor Paola Cappellaro. The careful balance of freedom and guidance she offered, as well as her tireless support for her students (often behind the scenes), were essential to the success of this undertaking. I can only aspire to match her remarkable ability to see the forest through the trees, and to move seamlessly between fields in response to evolving problems. Working with her these past four years has been a privilege, and has shaped my approach to science and engineering.

I am also grateful to Liang Jiang, Seth Lloyd and Will Oliver for their guidance throughout my time at MIT. Liang's ability to instantly isolate the core of a research problem, and Seth's ability to spot such problems long before the rest of us come around, have been inspiring. Will's technical expertise speaks for itself; I only hope I can be as kind and generous to those starting out as he is, should I ever find myself in his shoes.

I also wish to acknowledge my fellow Quantum Engineering Group members, past and present: Scott Alsid, Louisa Huang, J-C Jaskula, Changhao Li, Yixiang Liu, Dominika Lyzwa, Pai Peng, Akira Sone, Calvin Sun, Ken Wei, and Yuan Zhu for good times and useful discussions. Special thanks in particular to my friend and collaborator Mo Chen, whose experimental know-how, unbounded network, and general *joie de vivre* have been a constant help.

Many thanks are also due to members of the MIT community and beyond, including Victor Albert, Ike Chuang, Eddie Farhi, Aram Harrow, Wolfgang Ketterle, Iman and Milad Marvian, Florentin Reiter, Peter Shor and Sisi Zhou, who have all

impacted my work in one way or another. The same goes for Brandy Baker, Heather Barry, Pete Brenton, and Dianne Lior, without whom things would presumably have derailed long ago. On a personal level, I am grateful to Jochen Braumüller, Peter Johnson and the Zapatistas, Bharath Kannan, and Morten Kjærgaard, with whom I've tackled many of life's great questions while running along the Charles. Thank you also to Steve Girvin, Adrian Lupaşcu, Eduardo Martín-Martínez, and Michael Vanner for fruitful discussions and hospitality over these last few years, and to Michael Goldberg, Sally Gunz, Kris Prather and Alex Vitkin for helping me figure out where I was going and how to get there. Thank you finally to Achim Kempf (a diplomat in a former life, surely) for his unwavering support. Things would look quite different were it not for him.

Lastly, none of this would have been possible without the immeasurable help of my parents, Richard and Ruth. I am eternally grateful to them. Thank you also to my mother-in-law Linda, for enabling the juggling act of the last few years, and for everything else. Finally, I don't know where I would be without my wife. Sarah: thank you for making it all worthwhile.

# Contents

# List of Figures

# Citations to Previously Published Work

Most chapters of this thesis are based on material which has appeared in print elsewhere. By chapter number:

**Chapter 2** D. Layden, M. Chen, P. Cappellaro, Phys. Rev. Lett. **124**, 020504 (2020). © 2020 American Physical Society.

**Chapter 3** D. Layden, L. R. Huang, P. Cappellaro, to appear in Quantum Sci. Technol. (2020).

**Chapter 5** D. Layden, P. Cappellaro, npj Quantum Inf. **4**, 30 (2018).

**Chapter 6** D. Layden, S. Zhou (equal contributions), P. Cappellaro, L. Jiang, Phys. Rev. Lett. **122**, 040502 (2019). © 2019 American Physical Society.

# Preface

*"After growing wildly for years, the field . . . appears to be reaching its infancy."*
– John R. Pierce

Quantum mechanics has been revolutionized in recent decades by two complementary advances. The first advance is theoretical: a number of protocols have been proposed wherein a coherent quantum system, through careful control, could be made to output useful information that a classical system of equivalent size could not produce. Among these protocols are quantum algorithms with dramatically better scaling than their known classical counterparts on a number of important computational problems, as well as schemes for measuring physical quantities with unprecedented sensitivity and resolution. Other related protocols enable encrypted communication whose security relies on the laws of quantum physics rather than on computational complexity.

The second advance is experimental. The phenomena of quantum superposition and entanglement, which underlie the protocols above, were largely confined to *gedanken* experiments for much of the $20^{\text{th}}$ century. As a result of steady, interdisciplinary progress, however, researchers can now create, control, and measure coherent quantum systems of appreciable size with impressive precision. There are several different types of such quantum systems, based on superconducting circuits, trapped ions, neutral atoms, nuclear and electronic spins, and photons, to name a few, as well as combinations thereof. To some extent, these different platforms are equivalent; for instance, the same quantum algorithm run on any two of them should produce the same output, regardless of the underlying device and its physics. This has allowed

the two aforementioned advances to occur largely in parallel, requiring relatively little close interaction between the communities responsible for each.

As we progress along the path towards implementing these protocols and realizing true quantum-coherent technologies (henceforth simply "quantum technologies"), however, a new paradigm is emerging. It is not yet clear on which types of devices these technologies will ultimately rely; yet, some of the physics and engineering challenges with which they will likely contend have become apparent. These include the dominant error mechanisms in leading candidate systems, as well as challenges inherent in specific applications, such as quantum sensing. Incorporating these experimental concerns directly into theory—using a level of abstraction somewhere between those typical in the theoretical and experimental communities—is a promising way to shorten the path towards useful quantum technologies.

This device- and application-adapted approach is timely, as many quantum devices are entering a gray area wherein they are neither obviously useful, nor obviously useless. That is, they are too small and too noisy for most proposed protocols, and yet, they are increasingly hard to mimic classically. For instance, a fledgling quantum computer recently outperformed a supercomputer on an *ad hoc* problem (with few clear applications) [1]. Theory closely informed by experiment can provide a powerful means for getting the most out of these limited—yet increasingly substantive—devices. Therefore, barring an unexpected hardware breakthrough, it will likely be a critical ingredient in moving from proof-of-principle demonstrations to useful early applications in the coming years.

An important area where this approach could have a substantial impact is in noise suppression, and in quantum error correction (QEC) more specifically. As mentioned above, noise sets an important—if not the main—limit on current quantum devices. Not only are quantum systems often highly susceptible to small disturbances, but their manifestly quantum nature prohibits straightforward feedback stabilization, as this would collapse their state. QEC encompasses a family of noise-suppression techniques which use encoding and (typically) feedback based on partial measurements to stabilize quantum systems without completely collapsing their states. In effect, it is

a method of making a noisy quantum device behave like a smaller, but less noisy one; that is, of trading off size for reduced noise. It is a remarkably powerful tool, which conventionally requires little physical knowledge of the noise afflicting a device. In fact, this feature underpins QEC's envisioned role as the main tool to reduce noise to tolerable levels for large-scale quantum technologies. Characterizing the noise mechanisms in a quantum device can be exponentially hard in the device size; conventional QEC casts a wide enough net to avoid this eventual bottleneck.

Of course, casting such a wide net comes at a price: conventional QEC is very resource-intensive. Stated differently, it trades device size for reduced noise at a rate too exacting to be useful in most current and emerging devices [2]. This thesis instead takes a more targeted approach to QEC, which is directly informed by current experiments. Part I deals with device-adapted QEC, and shows that one can dramatically improve QEC's efficiency in leading candidate devices by incorporating their underlying physics from the start. Part II focuses on the specific QEC challenges posed by quantum sensing, rather than on a particular physical device, and develops application-adapted QEC schemes where conventional ones do not work. Taken together, both parts aim to help bridge the gap between the long-term plans for handling noise in large quantum devices, and the reality of current experiments. This is a critical task, if the long-term plans for quantum technologies are to be realized at all.

# Part I

# Device-Adapted Quantum Error Correction

*"All models are wrong, but some are useful."* – George E. P. Box

# Chapter 1

# Introduction

## 1.1 Formalism

This thesis deals exclusively with finite-dimensional quantum systems; that is, those with a finitely many energy eigenstates. These ubiquitous systems underlie many current candidate realizations of quantum technologies. Their states are most simply represented by a vector $|\psi\rangle$ in a complex vector space $\mathscr{H} = \mathbb{C}^d$ ($d < \infty$), called the system's Hilbert space. Since most envisioned quantum technologies involve many interacting subsystems, we will often be concerned with the aggregate Hilbert space arising from those of various subsystems. If subsystems 1 and 2 have Hilbert spaces $\mathscr{H}_1$ and $\mathscr{H}_2$ of dimensions $d_1$ and $d_2$ respectively, their combined Hilbert space is the tensor product of $\mathscr{H}_1$ and $\mathscr{H}_2$, written $\mathscr{H} = \mathscr{H}_1 \otimes \mathscr{H}_2$, of dimension $d = d_1 d_2$. For our purposes, states in $\mathscr{H}_1 \otimes \mathscr{H}_2$ and matrices on it can be constructed in any basis by using as a definition

$$A \otimes B = \begin{pmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{pmatrix}, \tag{1.1}$$

for arbitrary matrices/vectors $A = (a_{ij})$ (taken here to be $m$ by $n$) and $B$. For instance, if the subsystems are in states $|\psi_1\rangle \in \mathscr{H}_1$ and $|\psi_2\rangle \in \mathscr{H}_2$, their aggregate

state $|\psi_1\rangle \otimes |\psi_2\rangle \in \mathscr{H}$, often written as $|\psi_1\rangle\,|\psi_2\rangle$ or $|\psi_1\,\psi_2\rangle$, can be constructed using Eq. (1.1). We will largely be concerned with two-level systems, whose basis states are often denoted as $|0\rangle$ and $|1\rangle \in \mathbb{C}^2$. As such systems are the quantum analogs of classical bits of information, they are often called qubits (pronounced "Q-bits"). The aggregate Hilbert space for $n$ qubits is $\mathscr{H} = \mathbb{C}^{2^n}$, a convenient basis for which is the set of states corresponding to classical $n$-bit strings, e.g., $|010\ldots0\rangle$. This is often called the computational basis. Note that not all states in $\mathbb{C}^{2^n}$ can be factored into an $n$-fold tensor product; those that can are said to be separable, and those that cannot are said to be entangled. The latter are distinctly non-classical, and form a key ingredient for most quantum technologies.

Calculating quantum measurement outcomes in this description often involves taking inner products—here, simply the dot product on $\mathbb{C}^d$. It is useful to define the dual space $\mathscr{H}^*$ to $\mathscr{H}$, comprising the set of linear maps $\langle\psi|$ from $\mathscr{H}$ to $\mathbb{C}$. $\mathscr{H}^*$ is also a vector space, and it has the same structure as $\mathscr{H}$ (more precisely, they are isomorphic). In fact, if $|\psi\rangle$ is written out as a $d$-dimensional column vector, $\langle\psi|$ is simply the corresponding row vector with each element replaced by its complex conjugate. In this notation, the inner product between $|\psi\rangle$ and $|\phi\rangle$ is simply $\langle\psi|\phi\rangle$. For $|\psi\rangle$ to encode a valid quantum state, we will demand that it have unit length, i.e., that $\langle\psi|\psi\rangle = 1$. Moreover, states which represent mutually-exclusive classical outcomes, such as computational basis states, are taken to be orthogonal.

Quantum measurements are described by a set of linear operators (for $d < \infty$, simply matrices) $\{M_j\}$, where $M_j$ encodes the $j^{\text{th}}$ potential measurement outcome. The probability of getting this outcome for a system in state $|\psi\rangle$ is

$$p_j = \langle\psi|\, M_j^\dagger M_j\, |\psi\rangle = ||M_j\,|\psi\rangle\,||^2, \tag{1.2}$$

where $M_j^\dagger$ is called the adjoint of $M_j$ and denotes its conjugate transpose, and $||\,|x\rangle\,|| = \sqrt{\langle x|x\rangle}$. To ensure that these probabilities sum to unity for all $|\psi\rangle$, the measurement operators must satisfy $\sum_j M_j^\dagger M_j = I$. If a measurement returns the $j^{\text{th}}$ outcome, the system will be left in the post-measurement state $M_j\,|\psi\rangle\,/\sqrt{p_j}$ (provided $p_j \neq 0$; when

$p_j = 0$ the corresponding post-measurement state is irrelevant). Physically, this means that measurements are typically destructive in quantum mechanics. For instance, the most common measurements involve $M_j = |j\rangle\langle j|$ for some orthogonal basis of states $\{|j\rangle\}$. Before the measurement, the system's state can be a general superposition of these $|j\rangle$'s, i.e., $|\psi\rangle = \sum_j c_j |j\rangle$, but interaction with the measurement device causes the system to collapse to a particular state $|j\rangle$ with probability $p_j = |c_j|^2$. This collapse makes straightforward feedback stabilization impossible; if a generic state $|\psi\rangle$ has been subject to an unknown disturbance, there is no general way to measure the impact on the system and correct accordingly without collapse.

The dynamics of a closed quantum system is described by the Schrödinger equation:

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = H(t) |\psi(t)\rangle , \tag{1.3}$$

where $H(t)$ is the system's Hamiltonian, the matrix encoding its energy levels. This dynamics can equivalently be written as $|\psi(t)\rangle = U(t) |\psi(0)\rangle$, where the matrix-valued function $U(t)$ satisfies the same differential equation as $|\psi(t)\rangle$:

$$i\hbar \frac{d}{dt} U(t) = H(t)U(t). \tag{1.4}$$

To simplify the notation, we will often leave the time dependence implicit in this and similar equations. Moreover, throughout this thesis we will use units in which $\hbar = 1$, unless otherwise stated. Since $H$ is Hermitian/self-adjoint ($H = H^\dagger$), $U$ is unitary ($U^\dagger = U^{-1}$). This ensures that $U$ preserves the length of state vectors, and therefore represents a valid evolution.

The Hamiltonian of a single qubit can always be written as a real linear combination of $I$ (the identity matrix) and the "Pauli matrices"

$$X = |0\rangle\langle 1| + |1\rangle\langle 0| \qquad Y = i\big( |1\rangle\langle 0| - |0\rangle\langle 1| \big) \qquad Z = |0\rangle\langle 0| - |1\rangle\langle 1| , \tag{1.5}$$

which have the form

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{1.6}$$

in the computational basis. These matrices are sometimes denoted as $\sigma_x, \sigma_y$ and $\sigma_z$, although in this thesis our notation will prove more convenient. The reason $I$, $X$, $Y$ and $Z$ span all possible qubit Hamiltonians is that they form a basis for matrices on $\mathbb{C}^2$. More precisely, any $2 \times 2$ Hermitian matrix can be expressed as as a linear combination of $I, X, Y$ and $Z$ with real coefficients, and any general $2 \times 2$ matrix can be expressed as a complex combination thereof. Furthermore, Pauli matrices have a natural geometric interpretation: the eigenstates of $X$, $Y$, and $Z$ encode the possible outcomes of measuring a spin-$\frac{1}{2}$ system in the spatial $x$, $y$, and $z$ directions respectively. More generally, a generic Hamiltonian for $n$ non-interacting qubits (which does not produce entanglement) is in $\text{span}\{I, X_j, Y_j, Z_j\}$, where for instance

$$X_3 = I \otimes I \otimes X \otimes \underbrace{I \otimes \cdots \otimes I}_{n-3} \tag{1.7}$$

acts as $X$ on qubit 3 and trivially on the others. A key property of tensor products is

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD), \tag{1.8}$$

for matrices/vectors $A$, $B$, $C$ and $D$ of compatible sizes. This means that we can write more general Hamiltonians describing coupled subsystems (which produce entanglement) compactly, e.g.:

$$\begin{aligned} H &= \frac{\omega_1}{2} Z \otimes I + \frac{\omega_2}{2} I \otimes Z + JX \otimes X \\ &= \frac{\omega_1}{2} Z_1 + \frac{\omega_2}{2} Z_2 + JX_1X_2. \end{aligned} \tag{1.9}$$

Note that it is common to drop certain $I$'s and $\otimes$ symbols to simplify the notation, when such shorthand introduces no ambiguity.

We will make occasional use of "rotating frames" in this thesis. That is, when

$H = H_0 + H_1$ has well-understood part $H_0$, and a part $H_1$ to be analyzed more closely, one can absorb the effects of $H_0$ into the frame of reference, therefore isolating those of $H_1$. More precisely, a state $|\psi\rangle$ in the "lab frame" becomes $e^{iH_0 t}|\psi\rangle$ in this rotating frame. One can easily show that this rotating frame state satisfies the Schrödinger equation, but with $H$ replaced by $\tilde{H}_1(t) = e^{iH_0 t} H_1 e^{-iH_0 t}$. Similarly, an operator $M$ in the lab frame, such as a measurement operator, becomes $\tilde{M}(t) = e^{iH_0 t} M e^{-iH_0 t}$ in the rotating frame. When $[H_0, M] := H_0 M - M H_0 = 0$, however, $M$ is not affected by the change of frame; a common reason for choosing $H_0$ to commute with the measurement operators one plans to implement. Note that we have implicitly assumed $H_0$ to be time-independent here. This is not strictly necessary, although it is the relevant case for this thesis.

It is often convenient to coarse-grain quantum dynamics into lumped unitaries $U$, describing a system's net evolution over some time interval, and thus abstracting away the fine-grained description of how this evolution occurred. This is an efficient way to describe quantum control sequences and algorithms; one can specify a desired sequence of lumped operations (often in a rotating frame) whose effect is readily understood, which can then be translated into a more opaque time-dependent Hamiltonian to be implemented. Common examples of such $U$'s acting on single qubits are Pauli matrices (which are both unitary and Hermitian) and Hadamard matrices:

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{1.10}$$

Unfortunately Hadamard matrices and Hamiltonians share the same symbol; however, the intended meaning of $H$ is usually clear from context. Another important type of lumped $U$ is that of controlled operations, which act on a qubit (called the target) only if another (called the control) is in the state $|1\rangle$. One common example is the controlled-$Z$ or controlled-phase operation, denoted $cZ$, which acts as

$$cZ = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes Z, \tag{1.11}$$

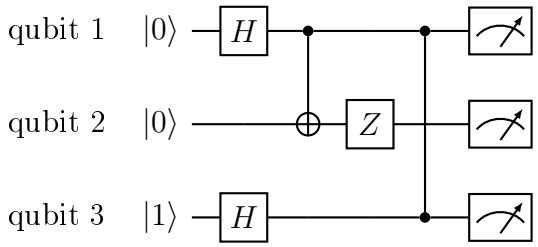



Figure 1-1: An example of a quantum circuit, showing three qubits initially prepared in the state $|001\rangle$, which are then subject to: (i) $H \otimes I \otimes H$ where $H$ is the Hadamard operation, (ii) a $c$NOT with qubits 1 and 2 as the control and target respectively, (iii) $Z_2$, and (iv) a controlled-$Z$, whose action is symmetric under the exchange of control and target qubits, which are therefore denoted with the same symbol. Finally, the qubits are measured in the computational basis, producing various 3-bit strings with known probabilities. The circuits in this thesis were made using Ref. [3].

where the first qubit is the control and the second is the target. Note that since both qubits may start in superposition states, this is an entangling operation that can produce non-classical effects; in particular, it can modify the control qubit's state. A closely related operation is the controlled-$X$ or controlled-NOT, whose name derives from the fact that the Pauli $X$ acts like a NOT gate in classical logic. It has the effect

$$c\text{NOT} = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes X. \tag{1.12}$$

It is often useful to represent these and other operations using circuit diagrams, inspired by classical logic circuits, as illustrated in Fig. 1-1. These circuits can be used to represent quantum control sequences (at a high level) and algorithms.

The framework introduced so far is that of pure states, which have zero entropy. We now introduce the more general formalism of density matrices, which can describe both pure and mixed states, the latter having non-zero entropy. This thesis will use both formalisms; while that of density matrices is more powerful in principle, that of pure states provides an easier means of understanding aspects of quantum error correction. To motivate the definition of density matrices, we note that the dynamics of a quantum system may not be identical in every run of an experiment, e.g., due to random variations in the system's environment. For instance, rather than

evolving each time by the same unitary $U$, a system may evolve by different $U(\theta)$'s, each occurring with probability $p(\theta)$. One might expect the average dynamics to be described by $\int p(\theta)U(\theta)d\theta$, but upon substituting this into Eq. (1.2), one sees that it does not give the correct averaged probabilities of measurement outcomes.

A more general description of a quantum state, which can encompass such uncertainty, is as a density matrix $\rho = \rho^\dagger$ with $\mathrm{tr}(\rho) = 1$ and with non-negative eigenvalues. In this description, a state vector $|\psi\rangle$ becomes $\rho = |\psi\rangle\langle\psi|$. If a density matrix $\rho$ can be written as $|\psi\rangle\langle\psi|$ (i.e., if it has a rank of 1), it is said to represent pure state; otherwise, it represents a mixed state with nonzero entropy. Any mixed state can be expressed as a probabilistic mixture of orthogonal pure states, $\rho = \sum_j \lambda_j |\psi_j\rangle\langle\psi_j|$, where the $\lambda_j$'s can be interpreted as probabilities. It is this decomposition that will allow us to move smoothly between the density matrix and state vector formalisms. For density matrices, Eq. (1.2) becomes instead $p_j = \mathrm{tr}(M_j^\dagger M_j \rho)$, and the post-measurement state for outcome $j$ is given by $M_j \rho M_j^\dagger / p_j$ (provided again $p_j \neq 0$; when $p_j = 0$ the corresponding post-measurement state is irrelevant). Notice that an average of density matrices (e.g., over noise realizations) gives the correct averaged measurement probabilities, as desired. Moreover, if a quantum system becomes entangled with its environment, the system's effective (mixed) state can always be expressed as a density matrix.

In the language of density matrices, the Schrödinger equation takes the form

$$\frac{d\rho}{dt} = -i[H, \rho], \tag{1.13}$$

which is sometimes called the quantum Liouville equation. It can be written more compactly by introducing the notion of a superoperator: a linear function on the space of matrices, which in turn act on $\mathscr{H}$. In effect, a superoperator is to a density matrix $\rho$ what a matrix is to a state vector $|\psi\rangle$. In particular, defining the Hamiltonian superoperator $\mathcal{H}$ by its action on a density matrix as $\mathcal{H}(\rho) := [H, \rho]$, Eq. (1.13) takes the simple form of $\dot{\rho} = -i\mathcal{H}(\rho)$. Just as a unitary matrix $U$ encoded the evolution of a state vector under the Schrödinger equation, the same evolution of a density

matrix is described by the superoperator $\mathcal{U}$, defined as $\mathcal{U}(\rho) := U\rho U^\dagger$, which can be found by integrating the quantum Liouville equation (or equivalently, the Schrödinger equation).

The Schrödinger and quantum Liouville equations describe unitary dynamics (so called because $U^\dagger = U^{-1}$). Such dynamics underpins most proposals for quantum technologies. In experiments, however, dynamics are seldom exactly unitary. Instead, they typically also have an irreversible character, in which information encoded in a quantum state is gradually lost, due to growing entanglement with an environment or classical noise processes affecting the system's Hamiltonian. We will refer to such imperfect systems as being open. The gradual loss of information in open quantum systems is broadly called decoherence, and it is arguably the central obstacle to developing useful quantum technologies. The aim of quantum error correction is to effectively suppress decoherence, so as to make real open quantum systems behave almost like ideal closed ones, in effect.

In principle it is possible to write an equation of motion for any open quantum system, analogous to the Schrödinger/quantum Liouville equations for closed systems [4, 5]. In practice this is rarely useful for two reasons: First, one often has little knowledge of the environment's internal dynamics, nor the exact nature of its coupling to the system. Second, even if one had this knowledge, the system's resulting equation of motion would almost surely be too complex to solve. Instead, one is typically forced to use a more empirically-motivated, effective description of open system dynamics in order to make progress. This can be as much an art as it is a science; there are many ways to model open quantum systems, and a model well-suited for one system and purpose can fail to capture important details of others.

Just as we can coarse-grain unitary quantum dynamics into a lumped matrix $U$, we can encode the net effect of both unitary and non-unitary quantum dynamics into a superoperator $\mathcal{K}$ called a quantum channel (under the reasonable assumption that the system was not initially correlated with its environment), whose name derives from classical information theory. Whereas $U$ must be unitary to describe a valid closed quantum dynamics, $\mathcal{K}$ must be completely positive and trace-preserving (CPTP, as

well as Hermiticity preserving). For our purposes, it suffices to say that any $\mathcal{K}$ admits a "Kraus decomposition," that is, it can be written as a sum

$$\mathcal{K}(\rho) = \sum_j K_j \rho K_j^\dagger, \tag{1.14}$$

where $\{K_j\}$ are matrices known as Kraus operators of $\mathcal{K}$, which have the property that $\sum_j K_j^\dagger K_j = I$. This expression generalizes the unitary superoperator $\mathcal{U}$ encoding the dynamics of the Schrödinger/Liouville equations, which has a single Kraus operator $U$. It is important to note that Kraus decompositions are not unique in general; rather, the same generic quantum channel $\mathcal{K}$ could be written in terms of different Kraus operators $\{K'_j\} \neq \{K_j\}$. In fact, quantum channels generally admit infinitely many Kraus decompositions, and we are free to pick the most convenient one—a fact that is important for quantum error correction.

## 1.2 Quantum Error Correction

The ubiquity of decoherence, together with the destructive nature of quantum measurements, would seem to prohibit the building of useful, controllable large-scale quantum devices. Mathematically, this difficulty is reflected in part by the fact that most quantum channels $\mathcal{K}$ describing non-unitary dynamics do not have an inverse channel. That is, there is generally no physically realizable $\mathcal{K}^{-1}$—even in principle—such that $\mathcal{K}^{-1}\mathcal{K} = \mathcal{I}$, where $\mathcal{I}$ denotes the identity channel (i.e., $\mathcal{I}(\rho) = \rho$ for all $\rho$). More broadly, there is generally no channel $\mathcal{G}$ such that $\mathcal{G}\mathcal{K}$ represents a unitary evolution.

Thankfully, there is a loophole which could allow decoherence to be effectively suppressed through quantum error correction (QEC). It relies on two key observations: First, we don't necessarily need to reverse decoherence on all quantum states; instead, we could realize quantum technologies by doing so only on a subset of them. Second, measurement need not completely collapse a quantum state. Rather than measuring whether a system is in some particular state $|j\rangle$, one can instead measure whether it

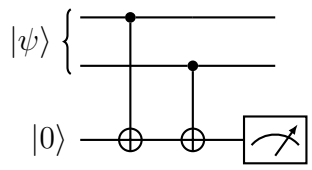



Figure 1-2: A parity measurement performed indirectly on the top two qubits, which realizes the measurement operators $M_0 = |00\rangle\langle 00| + |11\rangle\langle 11|$ and $M_1 = |01\rangle\langle 01| + |10\rangle\langle 10|$. For an initial state $|\psi\rangle = c_{00}|00\rangle + c_{01}|01\rangle + c_{10}|10\rangle + c_{11}|11\rangle$, the measurement returns 0 with probability $p_0 = |c_{00}|^2 + |c_{11}|^2$, and returns 1 with probability $p_1 = |c_{01}|^2 + |c_{10}|^2$. In either case, the state of the top two qubits is not totally collapsed, but instead projected into the even or odd parity subspaces respectively (namely, onto $\text{span}\{|00\rangle, |11\rangle\}$ or $\text{span}\{|01\rangle, |10\rangle\}$). In particular, these qubits remain entangled after the measurement for a generic $|\psi\rangle$.

is in some larger set of states. An example of such a measurement, which is usually performed indirectly with the help of an ancillary qubit (called an ancilla), is shown in Fig. 1-2.

To see how QEC exploits these observations, we will start with an abstract picture, and then gradually build up a more concrete description until we can finally meld in the physics of certain quantum devices. QEC makes use of states within a special subspace $\mathscr{C}_0$ of $\mathscr{H}$, called the codespace, over which $\mathcal{K}$ can be reversed. (At least approximately.) Typically, on a system of $n$ qubits the codespace has dimension $\dim(\mathscr{C}_0) = 2^k$ for $k < n$, meaning it has the same structure as (i.e., is isomorphic to) the Hilbert space of a smaller, $k$-qubit system[1]. We can therefore define a basis of orthogonal codeword states $\{|0\ldots00_{\text{L}}\rangle, |0\ldots01_{\text{L}}\rangle, \ldots, |1\ldots11_{\text{L}}\rangle\}$ for $\mathscr{C}_0$, labeled by $k$-bit strings with a subscript L for "logical." Ultimately, QEC will provide a method by which $n$ noisy qubits can be made to behave like $k$ less noisy ones, at the cost of an $n-k$ qubit overhead, as well as additional control operations. The $k$ effective qubits encoded in $\mathscr{C}_0 \subset \mathscr{H}$ are called logical qubits, in contrast to the $n$ physical qubits.

The idea is to prepare some initial logical state $|\psi_{\text{L}}\rangle \in \mathscr{C}_0$, and to perform operations and measurements on the $n$ physical qubits which effectively enact desired ones at the logical level, i.e., on the $k$ logical qubits. The typical QEC strategy is to peri-

[1]Of course, $\dim(\mathscr{H})$ and $\dim(\mathscr{C}_0)$ need not be powers of 2—nor even finite—in general [6]. However, we will focus on the qubit case here for concreteness, as it will be the most relevant for this thesis.

odically detect and correct errors from decoherence (and perhaps also from imperfect control and measurements) all the while. For this latter part to be possible, we must choose an appropriate codespace $\mathscr{C}_0$ for the channel $\mathcal{K}$ representing the decoherence we aim to reverse. Ideally, it should be chosen such that each of the Kraus operators $K_j$ of $\mathcal{K}$ maps states in $\mathscr{C}_0$ to mutually orthogonal subspaces $\mathscr{C}_j$ without distortion. Let's unpack this statement and make it more precise.

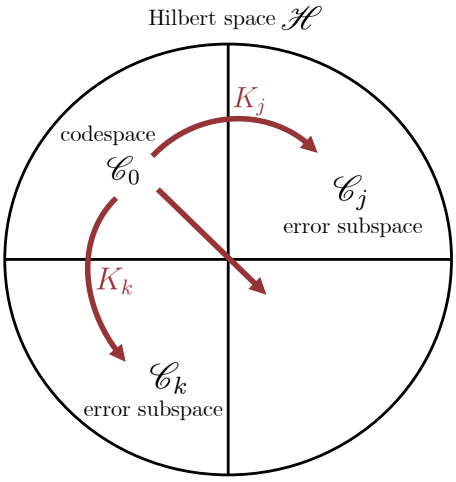
### 1.2.1 Knill-Laflamme Condition

Much like classical error correction, QEC requires one to identify "which error occurred" in a quantum system (or whether an error occurred at all), in the typical parlance. The meaning of this phrase will be explained shortly. Unlike in classical error correction, however, in QEC one must take care to identify errors without destroying the encoded state. To see how this is possible, consider the action of a channel $\mathcal{K}$ describing decoherence on an initial logical state $\rho_{\text{L}} = |\psi_{\text{L}}\rangle\langle\psi_{\text{L}}|$:

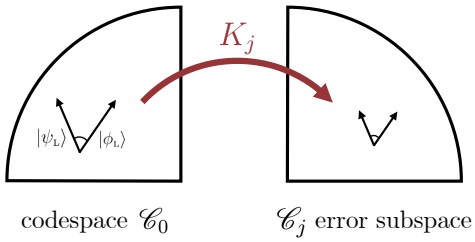
$$\mathcal{K}(\rho_{\text{L}}) = K_1 |\psi_{\text{L}}\rangle\langle\psi_{\text{L}}| K_1^\dagger + K_2 |\psi_{\text{L}}\rangle\langle\psi_{\text{L}}| K_2^\dagger + \dots . \tag{1.15}$$

Suppose that each $K_j$ mapped a generic $|\psi_{\text{L}}\rangle$ out of $\mathscr{C}_0$ and into some new subspace $\mathscr{C}_j$ (independent of $|\psi_{\text{L}}\rangle$), and that all of these subspaces were mutually orthogonal, as illustrated in Fig. 1-3a. Then, measuring which subspace the system was in would produce a pure post-measurement state $K_j |\psi_{\text{L}}\rangle\langle\psi_{\text{L}}| K_j^\dagger / p_j$—or, written as a state vector, $K_j |\psi_{\text{L}}\rangle / \sqrt{p_j} \in \mathscr{C}_j$—with probability $p_j = \langle\psi_{\text{L}}| K_j^\dagger K_j |\psi_{\text{L}}\rangle$. (If $p_j = 0$ the corresponding post-measurement state is irrelevant and undefined.) Therefore, even though $\mathcal{K}$ encompassed all of the Kraus operators $\{K_j\}$ at once, this choice of codespace would allow us to only worry about correcting one at a time. In fact, it would allow us to think of the $K_j$'s as mutually-exclusive errors, each occurring with some probability $p_j$; and of the measurement as revealing only the "error syndrome" $j$, telling us which error occurred without fully collapsing the system's state.

Mathematically, to find such a codespace we demand that any two logical states $|\psi_{\text{L}}\rangle$ and $|\phi_{\text{L}}\rangle \in \mathscr{C}_0$ be mapped to different orthogonal subspaces $\mathscr{C}_j$ and $\mathscr{C}_k$ by different




(a) Orthogonal subspaces




(b) No distortion

Figure 1-3: Our two requirements for QEC codes. Left: Distinct Kraus operators $K_j$ and $K_k$ $(j \neq k)$ must map all states in the codespace $\mathscr{C}_0$ into orthogonal subspaces $\mathscr{C}_j$ and $\mathscr{C}_k$ respectively. Right: Each $K_j$ must preserve the angle between any two logical states $|\psi_L\rangle, |\phi_L\rangle \in \mathscr{C}_0$ (or cause both to vanish).

errors $K_j$ and $K_k$:

$$\langle\psi_L| K_j^\dagger K_k |\phi_L\rangle = 0 \qquad (j \neq k), \tag{1.16}$$

as in Fig. 1-3a. This will allow us to identify which error occurred through measurement. However, simply identifying an error is not enough to reverse its effect. For instance, knowing that a qubit emitted a photon and decayed to state $|0\rangle$ does not enable one to restore its full initial state $\alpha|0\rangle + \beta|1\rangle$. Additionally, we must insist that $K_j$ not distort states as it maps them from $\mathscr{C}_0$ to $\mathscr{C}_j$. For starters, this means that if $p_j = 0$ for some $|\psi_L\rangle \in \mathscr{C}_0$ (that is, if $K_j|\psi_L\rangle = 0$), $p_j$ must be identically zero for all logical states. More broadly, when $p_j \neq 0$, we must demand that the angle between arbitrary initial states and that between the corresponding post-measurement states be the same, as illustrated in Fig. 1-3b, that is:

$$\langle\psi_L|\phi_L\rangle = \frac{1}{||K_j|\psi_L\rangle|| \, ||K_j|\phi_L\rangle||} \langle\psi_L| K_j^\dagger K_j |\phi_L\rangle, \tag{1.17}$$

for all $|\psi_L\rangle, |\phi_L\rangle \in \mathscr{C}_0$. This will allow us to correct the post-measurement states through a unitary operation, even though $K_j$ need not be unitary, nor even invert-

ible. In particular, if two logical states are orthogonal, their corresponding post-measurement states must be too. It follows immediately that $p_j$ must be independent of the initial state. This was to be expected; were it not the case, the error syndrome measurement would reveal information about the encoded state, thus damaging it. In aggregate, our requirement that $K_j$ cause no distortion amounts to demanding

$$\langle\psi_{\text{L}}|\, K_j^\dagger K_j\, |\phi_{\text{L}}\rangle = \lambda_j \langle\psi_{\text{L}}|\phi_{\text{L}}\rangle, \tag{1.18}$$

for all $|\psi_{\text{L}}\rangle\,, |\phi_{\text{L}}\rangle \in \mathscr{C}_0$, where $\lambda_j$ is some constant that depends only on $j$.

Combined, the two requirements derived above become

$$\langle\psi_{\text{L}}|\, K_j^\dagger K_k\, |\phi_{\text{L}}\rangle = \lambda_j \delta_{jk} \langle\psi_{\text{L}}|\phi_{\text{L}}\rangle \qquad \text{for all } |\psi_{\text{L}}\rangle\,, |\phi_{\text{L}}\rangle \in \mathscr{C}_0. \tag{1.19}$$

We can re-write this expression in a more useful form by defining the orthogonal projector $P$ onto the codespace:

$$P = \sum_{i=0}^{2^k-1} |i_{\text{L}}\rangle\langle i_{\text{L}}|\,, \tag{1.20}$$

where $|i_{\text{L}}\rangle$ denotes the binary representation of $i$ (for instance $|2_{\text{L}}\rangle = |0\ldots010_{\text{L}}\rangle$). In terms of $P$, Eq. (1.19) becomes simply

$$P K_j^\dagger K_k P = \lambda_j \delta_{jk}\, P. \tag{1.21}$$

One step remains. We mentioned in Section 1.1 that a quantum channel generically admits infinitely many Kraus decompositions. More precisely, if $\{K_j\}$ are Kraus operators for $\mathcal{K}$, then so too are $\{K'_j = \sum_k v_{jk} K_k\}$ for any unitary matrix $V = (v_{jk})$. In this derivation we have implicitly chosen a convenient set of Kraus operators which highlighted the underlying structure of QEC. Of course, all Kraus decompositions for $\mathcal{K}$ are physically equivalent, so our result should not depend on having chosen a particular one. Rather, in terms of the generic Kraus operators $\{K'_j\}$, the previous

equation becomes

$$PK_j'^{\dagger}K_k'P = m_{jk}\,P. \tag{1.22}$$

where $M = (m_{jk})$ is known as the code matrix, and has eigenvalues[2] $\mathrm{spec}(M) = \{\lambda_j\}$. Eq. (1.22) is the famous Knill-Laflamme condition [7], which can equivalently be written as

$$PK_j'^{\dagger}K_k'P \propto P. \tag{1.23}$$

While we have derived it as a sufficient condition for QEC, it is in fact both sufficient and necessary for the existence of a quantum channel that reverses $\mathcal{K}$ over a subspace $\mathscr{C}_0 \subset \mathscr{H}$ [7]. To simplify the notation, we will henceforth drop the prime marks ($'$) on the $K_j'$'s, and use $\{K_j\}$ to denote a generic set of Kraus operators for the channel $\mathcal{K}$.

As presented here, QEC looks very much like classical error correction, in the sense that both use clever encodings to identify and reverse discrete errors. However, the state space of $n$ qubits is much broader than that of $n$ classical bits. Quantum systems can therefore be affected by noise in a wide range of ways which have no classical analogs. Fortunately, QEC is a remarkably powerful tool for reversing such a continuum of decoherence processes. Specifically, a QEC code that can reverse some channel $\mathcal{K}$ over a codespace $\mathscr{C}_0 \subset \mathscr{H}$ will do the same for any other channel $\tilde{\mathcal{K}} \neq \mathcal{K}$ whose Kraus operators are linear combinations of $K_j$'s (not necessarily related through a unitary). This follows immediately from the Knill-Laflamme condition. There is no reason to view the Kraus operators of $\mathcal{K}$ as being more natural or fundamental to the code than those of $\tilde{\mathcal{K}}$, even though they are not generally equivalent. This suggests that we ought not think of QEC as necessarily correcting a discrete set of physical errors, each of which afflicting the system with some probability per unit time, nor as being tied to any particular channel. Rather, we should think of it as casting a "net" which can catch a continuum of possible errors, and of this net's precise shape as being specified through a list of discrete errors. More specifically, imagine designing

[2] Choosing different Kraus decompositions for $\mathcal{K}$ amounts to expressing $M$ in different bases. Our initial choice of $\{K_j\}$ diagonalizes $M$.

a QEC code such that $PE_j^\dagger E_k P \propto P$ for some desired set of "error operators" $\{E_j\}$, which one is free to choose. Then, the effect of any matrix in $\mathscr{E} = \text{span}\{E_j\}$—the continuous "net" cast by the code in this metaphor—on a logical state can be reversed. One has wide discretion in specifying these error operators $\{E_j\}$ when designing a QEC code; for instance, they need not accurately describe the decoherence in a particular device, so long as the true Kraus operators for this decoherence are contained in $\mathscr{E}$. (Since we care only about the "net" $\mathscr{E}$ resulting from $\{E_j\}$, we can equivalently pick different error operators with the same span.) A central theme of this thesis, however, is that QEC codes designed around an $\mathscr{E}$ which closely reflects the actual decoherence mechanisms in a device—that is, casting a targeted net— can provide substantial benefits.

**Example 1: Idealized Bit Flips**

Consider for illustration $n = 3$ qubits subject to the decoherence channel

$$\mathcal{K}(\rho) = (1-3p)\rho + p(X_1\rho X_1 + X_2\rho X_2 + X_3\rho X_3); \tag{1.24}$$

a highly idealized model in which a bit-flip error ($|0\rangle \leftrightarrow |1\rangle$) occurs on at most one qubit with some probability $p$. While there is no inverse channel $\mathcal{K}^{-1}$, we can reverse $\mathcal{K}$ over a two-dimensional subspace, i.e., encode $k = 1$ logical qubit in this system using the code $\mathscr{C}_0 = \text{span}\{|0_{\text{L}}\rangle, |1_{\text{L}}\rangle\}$, where $|0_{\text{L}}\rangle = |000\rangle$ and $|1_{\text{L}}\rangle = |111\rangle$. The Kraus operators $K_j = \sqrt{p}X_j$ (for $1 \le j \le 3$) map states in the codespace to the error subspaces

$$\mathscr{C}_1 = \text{span}\{|100\rangle, |011\rangle\} \qquad \mathscr{C}_2 = \text{span}\{|010\rangle, |101\rangle\} \qquad \mathscr{C}_3 = \text{span}\{|001\rangle, |110\rangle\} \tag{1.25}$$

without distortion, and $K_0 = \sqrt{1-3p}\, I$ acts trivially. That is, $P = |000\rangle\langle 000| + |111\rangle\langle 111|$ satisfies the Knill-Laflamme condition as required, with a code matrix

$$M = \begin{pmatrix} 1-3p & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{pmatrix} \tag{1.26}$$

with respect to $\{K_j\}_{j=0}^{3}$. This code casts a net defined by the error operators $\{E_j\} = \{I, X_1, X_2, X_3\}$, whose span is $\mathscr{E}$. In this first example, the Kraus operators $K_j$ describing the decoherence are simply proportional to these abstract error operators around which the code is designed. We will see in later examples how this need not be the case.

One can reverse the action of $\mathcal{K}$ over $\mathscr{C}_0$ by majority vote. As we have argued, however, it is imperative that the error detection process reveal only whether the state is in $\mathscr{C}_0$, $\mathscr{C}_1$, $\mathscr{C}_2$ or $\mathscr{C}_3$, but nothing about the encoded information. This can be done by measuring the parity between each pair of qubits, as in Fig. 1-2, rather than measuring any individual qubit's state. If each pair of qubits has even parity then no error occurred, and there is no need for feedback. Otherwise, one can infer which error occurred—or in this case, simply on which qubit it occurred—using the syndrome measurement outcomes for each of the three pairs, and by applying $X$ to the errant qubit. In fact, it is not necessary to measure the parity of all three qubit pairs; one gets the same information from measuring only two of them. The complete process of error detection and correction, which we will call the recovery, is shown in Fig. 1-4.

This code is called the bit-flip code. Notice that if the Kraus operators described phase-flip errors instead ($K_j \propto Z_j$, producing $|1\rangle \leftrightarrow -|1\rangle$ on qubit $j$), one could instead use $|0_{\text{L}}\rangle = |{+}{+}{+}\rangle$ and $|1_{\text{L}}\rangle = |{-}{-}{-}\rangle$, where $|\pm\rangle := \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$ and $Z|\pm\rangle = |\mp\rangle$. Errors could be detected by measuring parity in the $|\pm\rangle$ basis, and corrected by applying $Z_j$'s rather than $X_j$'s. This latter code is called the "phase flip code." Because both these codes are the same up to a change of basis, they are sometimes

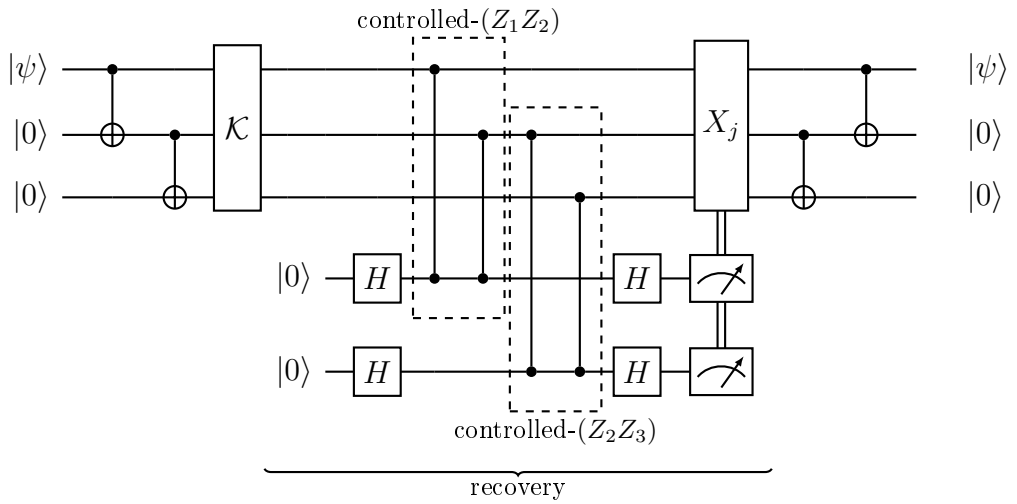



Figure 1-4: A circuit showing: (i) the encoding of a physical qubit into the bit-flip code, (ii) exposure to the channel $\mathcal{K}$, (iii) a recovery consisting of two parity measurements followed by the application of $I$, $X_1$, $X_2$ or $X_3$ as needed, and (iv) the mapping of the encoded state back to the first physical qubit. Note that double wires denote the transmission of classical information, and that we have expressed the same parity measurements found in Fig. 1-2 in terms of Hadamard and c$Z$ operations, for reasons that will be explained in Section 1.2.4. Finally, in practice one may want to avoid converting between physical and logical states as shown here, and instead work only with encoded states at every step.

both simply called repetition codes, after the family of classical error-correcting codes which encode each bit into $n \geq 3$ copies (i.e., $0 \mapsto 0\ldots0$ and $1 \mapsto 1\ldots1$). Like their classical counterparts, these quantum error-correcting codes are the smallest instances in a family of codes. The $n = 3$ repetition code can correct single $X_j$ errors (or linear combinations of these) at the cost of an $n - k = 2$ qubit overhead. One can instead encode a logical qubit ($k = 1$) into $n = 5$ physical qubits through $|i_{\rm L}\rangle = |iiiii\rangle$ for $i = 0/1$ and correct for arbitrary bit flips on up to two qubits through a majority vote (similarly for the phase-flip code). This increased protection, which leads to a wider "net" $\mathscr{E} = \text{span}\{I, X_j, X_jX_k\}$ comes at the expense of an increased overhead of $n - k = 4$. More generally, correcting for arbitrary bit or phase flips on $\leq w$ qubits with a repetition code requires an overhead of $n - k = 2w$ qubits.

### 1.2.2 Role of Quantum Error-Correcting Codes

Notice that the circuit in Fig. 1-4 represents a multi-step procedure, whereas the Knill-Laflamme condition specifies only an abstract codespace $\mathscr{C}_0 = \text{range}(P)$. There are two reasons why we will often treat such QEC codespaces as fundamental in this thesis, and focus on them accordingly. First, a QEC code satisfying the Knill-Laflamme condition automatically implies the form of an appropriate feedback correction scheme. Second, a QEC code can also be used to implement open-loop error suppression.

Let us briefly expand on the first point. Suppose $\{E_j\}$ are error operators for a QEC code that have been chosen to produce a diagonal code matrix, i.e., $PE_j^\dagger E_k P = \delta_{jk}\lambda_j P$. Such $E_j$'s exist for any appropriate $\mathscr{E}$. For each nonzero $\lambda_j$, one can define a unitary $U_j$ by performing a polar decomposition of $E_jP$ to give

$$E_jP = \sqrt{\lambda_j}U_jP. \tag{1.27}$$

We can then define orthogonal projectors $P_j := U_jPU_j^\dagger$ onto the error subspaces $\mathscr{C}_j$. Note that $P_jP_k = P_j\,\delta_{jk}$, as expected from our derivation of the Knill-Laflamme

condition. Finally, it is straightforward to show that the channel

$$\mathcal{R}(\rho) := \sum_{j\,|\lambda_j \neq 0} U_j^\dagger P_j \rho P_j U_j \tag{1.28}$$

reverses any channel $\mathcal{K}$ with Kraus operators in $\mathscr{E}$ on a logical state $\rho_{\text{L}}$:

$$\mathcal{R}\big[\mathcal{K}(\rho_{\text{L}})\big] = \rho_{\text{L}}, \tag{1.29}$$

Note that $\mathcal{R}$ describes the process of measuring operators $\{P_j\}$, then applying the unitary $U_j^\dagger$ in the event of outcome $j$. Therefore, as claimed above, a QEC code $\mathscr{C}_0$ and a set of error operators $\{E_j\}$ together imply a closed-loop correction scheme [7, 8].

Now the second point. Rather than correct errors explicitly, one could think of making them energetically unfavorable and thus unlikely to occur, instead of correcting them when they do [9]. This could be done by engineering a system Hamiltonian with low-energy eigenstates forming a desired error-correcting code[3], separated from higher-energy states by a large energy gap. Such a Hamiltonian could be realized through strong, carefully designed couplings between qubits, for instance. The result would be a system with a low-energy subspace in which decoherence is exponentially suppressed in the size of the energy gap (under quite general assumptions about the environment) [10–12]. For instance, a register of qubits with a Hamiltonian

$$H = \frac{1}{2}\sum_j \omega_j Z_j - \sum_{jk} J_{jk} Z_j Z_k, \tag{1.30}$$

with sufficiently strong couplings $J_{jk} > 0$ will have a protected subspace corresponding to the bit-flip code. We will not deal explicitly with such open-loop error suppression techniques in this thesis. However, it is natural to think of some of the results presented here in this context. Since both the closed- and open-loop control schemes discussed above are ultimately specified by a codespace $\mathscr{C}_0 \subset \mathscr{H}$, we will often treat

[3] In fact, an error-detecting code, which reveals when an error occurs though not necessarily which one occurred, is sufficient for this.

the latter as the core of QEC.

**Example 2: Local Phase Flips**

Let us build up towards more realistic decoherence models by considering the single-qubit channel

$$\mathcal{K}_1(\rho) = (1-p)\rho + pZ\rho Z, \tag{1.31}$$

representing a phase error with probability $p$. The aggregate channel[4] for $n$ qubits undergoing $\mathcal{K}_1$ simultaneously (and independently) is denoted $\mathcal{K}_1^{\otimes n} = \mathcal{K}_1 \otimes \cdots \otimes \mathcal{K}_1$.

Notice that for $n = 3$ this is not the same (contrived) channel as in Example 1 with $X \leftrightarrow Z$. Here we do not artificially impose that errors can occur only on one qubit at once; rather, since each qubit is subject to an independent decoherence process, errors can occur on $w$ qubits with probability $O(p^w)$. For $n = 3$, $\mathcal{K}_1^{\otimes n}$ has Kraus operators

$$\begin{array}{lll} & K_0 = (1-p)^{3/2} I & \\ K_1 = \sqrt{p}(1-p)Z_1 & K_2 = \sqrt{p}(1-p)Z_2 & K_3 = \sqrt{p}(1-p)Z_3 \\ K_4 = p\sqrt{1-p}Z_1Z_2 & K_5 = p\sqrt{1-p}Z_2Z_3 & K_6 = p\sqrt{1-p}Z_1Z_3 \\ & K_7 = p^{3/2}Z_1Z_2Z_3, & \end{array} \tag{1.32}$$

describing errors on 0 to 3 qubits by descending rows. We can still use the repetition code (this time for phase-flips) to suppress the decoherence described by this channel; however, we cannot hope to reverse it exactly, even over the codespace, as the number of Kraus operators grows too fast with $n$. This follows from a simple counting argument: $\mathcal{K}_1^{\otimes n}$ has $2^n$ distinct Kraus operators in general. To have one perfect logical qubit, we would need to decompose the total Hilbert space $\mathscr{H}$ into $2^n$ orthogonal 2-dimensional subspaces[5] $\mathscr{C}_0, \mathscr{C}_1, \ldots, \mathscr{C}_{2^n-1}$. Since $\dim(\mathscr{H}) = 2^n$, this is clearly impossible. Instead, we must make due with reversing the most damaging

[4] In general a density matrix for a bipartite system can be decomposed into a linear combination of separable matrices $\rho_i \otimes \rho_j$. The overall action of subsystem superoperators $\mathcal{A}$ and $\mathcal{B}$, described by the joint superoperator $\mathcal{A}\otimes\mathcal{B}$ can be understood through its action $(\mathcal{A}\otimes\mathcal{B})(\rho_i\otimes\rho_j) = \mathcal{A}(\rho_i)\otimes\mathcal{B}(\rho_j)$.

[5] Unless one can find a "degenerate code," in which distinct Kraus operators act identically on the codespace. We will seldom encounter such codes in this thesis.

Kraus operators; here, those of largest magnitude. Note that this means the error "net" $\mathscr{E}$ can no longer simply coincide with $\text{span}\{K_j\}_{j=0}^{7}$; this example therefore begins to illustrate the distinction between the Kraus operators of a physical decoherence process $\{K_j\}$, and the more abstract error operators $\{E_j\} = \{I, Z_1, Z_2, Z_3\}$ around which the code is built.

For $n = 3$, the effect of $\mathcal{K}_1^{\otimes n}$ followed by a recovery using the repetition code is

$$\rho_{\text{L}} \mapsto \mathcal{R}\big[\mathcal{K}_1^{\otimes 3}(\rho_{\text{L}})\big] = (1 - p_{\text{eff}})\rho + p_{\text{eff}}\, X_{\text{L}}\rho X_{\text{L}}, \tag{1.33}$$

where $p_{\text{eff}} = p^2(3 - 2p)$ is smaller than the physical error probability $p$ when $p < 1/2$, and $X_{\text{L}}$ maps $|0_{\text{L}}\rangle \leftrightarrow |1_{\text{L}}\rangle$. That is, at the logical level, the system behaves not like a noiseless qubit, but like a less noisy one, provided the physical noise strength is below a threshold value. While $\{K_j\}_{j\geq 1}$ can be viewed as describing physical errors, $X_{\text{L}}$ describes a "logical error" occurring with probability $p_{\text{eff}}$. For $n = 3$ $p_{\text{eff}} = O(p^2)$, and more generally $p_{\text{eff}} = O(p^{\frac{n+1}{2}})$ with these codes, reflecting the fundamental trade-off of QEC: decreased space for reduced noise.

Example 2 illustrates a common reality. A channel describing the open dynamics of a real quantum device is generally too complex (i.e., has too many Kraus operators, with too complicated a time dependence) to be exactly reversed over some codespace—even in principle. In this sense, all QEC is approximate QEC in practice. We will therefore aim to get the best logical error rates using the fewest possible resources in this thesis, by targeting the dominant decoherence mechanisms inherent in specific devices and applications.

### 1.2.3 Lindblad Equation

Moving closer yet to a picture of QEC in real devices, we introduce a simple model for the dynamics of open quantum systems called the Lindblad equation [13]:

$$\frac{d\rho}{dt} = -i[H, \rho] + \sum_{jk\geq 1} d_{jk}\Big(A_j\rho A_k^\dagger - \frac{1}{2}\{A_k^\dagger A_j, \rho\}\Big). \tag{1.34}$$

Here $H$ is the system's Hamiltonian, $D = (d_{jk})$ is a positive semidefinite matrix (i.e., Hermitian with non-negative eigenvalues), $\{A_j\}$ are arbitrary matrices on $\mathscr{H}$, and $\{A, B\} := AB + BA$ is an anti-commutator. There are numerous ways to derive the Lindblad equation. For instance, by postulating that the quantum channel which propagates a system from time $t$ to $t + \delta t$ depend only on $\delta t$, one arrives at Eq. (1.34) from the channel's Kraus decomposition in the $\delta t \to 0$ limit. More physically, one can arrive at the same equation by considering a system weakly coupled to an environment in which information about the system dissipates quickly compared to the system dynamics of interest [14, 15]. In Part II we will arrive at Eq. (1.34) differently still by analyzing the effect of a classical noise process on a quantum sensor. Suffice it to say that the Lindblad equation is an important tool for modeling open quantum systems. It therefore behooves us to connect it with the QEC formalism introduced thus far.

Eq. (1.34) can be written compactly as

$$\frac{d\rho}{dt} = -i\mathcal{H}(\rho) + \mathcal{D}(\rho) \tag{1.35}$$

in terms of the Hamiltonian superoperator $\mathcal{H}(\rho) := [H, \rho]$ and the "dissipator"

$$\mathcal{D}(\rho) = \sum_{jk \geq 1} d_{jk} \Big( A_j \rho A_k^\dagger - \frac{1}{2} \{ A_k^\dagger A_j, \rho \} \Big). \tag{1.36}$$

One can go a step further and define the superoperator $\mathcal{L} = -i\mathcal{H} + \mathcal{D}$, often called the Lindbladian superoperator, which has both Hamiltonian and dissipative parts, $\mathcal{H}$ and $\mathcal{D}$ respectively. The Lindblad equation then becomes $\dot{\rho} = \mathcal{L}(\rho)$. If $\mathcal{D} = 0$, the Lindblad equation reduces to the quantum Liouville equation for a closed system. As the name would suggest, however, a non-vanishing $\mathcal{D}$ generally introduces a non-unitary, irreversible character to the dynamics.

Notice that the sum in Eq. (1.34) runs over $j$ and $k$. It is an important fact that one can always get rid of the $j \neq k$ cross-terms in this sum by expressing $\mathcal{D}$ in terms of new operators. Concretely, let $W$ be a unitary matrix which diagonalizes $D$:

$$W^\dagger D W = \mathrm{diag}(\gamma_1, \gamma_2, \dots), \tag{1.37}$$

where $\{\gamma_j\}$ are the eigenvalues of $D$. Such a $W$ is guaranteed to exist. Then, one can re-write $\mathcal{D}$ as

$$\mathcal{D}(\rho) = \sum_{j\geq 1} \Big( L_j \rho L_j^\dagger - \frac{1}{2}\{L_j^\dagger L_j, \rho\} \Big), \tag{1.38}$$

in terms of the "jump operators"

$$L_j = \sqrt{\gamma_j} \sum_{k\geq 1} w_{kj} A_j. \tag{1.39}$$

These are so named because one can interpret each $L_j \rho L_j^\dagger$ term in $\mathcal{D}$ as describing the occurrence of a discrete jump/error $L_j$ on the system with some probability per unit time, determined by $\gamma_j$, within an otherwise-unitary dynamics [16, 17]. (The $-\frac{1}{2}\{L_j^\dagger L_j, \rho\}$ terms ensure proper normalization when no such jump occurs.)

This "diagonal" (i.e., having no cross-terms) form of $\mathcal{D}$ allows us to straightforwardly understand QEC in the language of Lindblad dynamics. To do so, consider the channel $\mathcal{K}$ describing Lindblad evolution for some short time $t$, which is given formally by $\mathcal{K} = e^{\mathcal{L}t}$ (assuming $\mathcal{L}$ is time-independent, otherwise the expression should include a time-ordered integral). In general, Kraus operators of $\mathcal{K}$ will depend on $t$ in complicated ways. However, we can expand them in powers of $t$ and solve for the leading-order parts quite easily. Consider some initial state $\rho$ evolving under Eq. (1.34) for a short time $t$. To first order in the evolution time $t$ the state becomes

$$\begin{aligned} \mathcal{K}(\rho) &= \rho + t \underbrace{\Big[ \Big( -iH - \frac{1}{2}\sum_{j\geq 1} L_j^\dagger L_j \Big)\rho + \rho \Big( -iH - \frac{1}{2}\sum_{j\geq 1} L_j^\dagger L_j \Big)^\dagger + \sum_{j\geq 1} L_j \rho L_j^\dagger \Big]}_{\mathcal{L}(\rho)} + O(t^2) \\ &= K_0 \rho K_0^\dagger + \sum_{j\geq 1} K_j \rho K_j^\dagger + O(t^2), \end{aligned} \tag{1.40}$$

where

$$K_0 = I - t\Big( iH + \frac{1}{2}\sum_{j\geq 1} L_j^\dagger L_j \Big) + O(t^2) \qquad \text{and} \qquad K_{j\geq 1} = \sqrt{t} L_j + O(t^{3/2}). \tag{1.41}$$

(If $\mathcal{L}$ were time-dependent, e.g., when expressed in a rotating frame, we would simply

get time-averaged operators in the expressions above.) Because Kraus operators always act on $\rho$ from both sides as $K_j \rho K_j^\dagger$, expanding $\mathcal{K} = e^{\mathcal{L}t} = \mathcal{I} + \mathcal{L}t + \ldots$ in powers of $t$ produces Kraus operators in integer powers of $\sqrt{t}$. We cannot generally hope to correct $\mathcal{K}$ to all orders in $t$, just as we could not do so for all powers of $p$ in Example 2. As in that example though, we can aim to correct the most damaging errors perturbatively, to order $O(t)$. This amounts to demanding that

$$PK_0^\dagger K_j P \propto P + O(t^{3/2}), \tag{1.42}$$

$$PK_j^\dagger K_k P \propto P + O(t^{3/2}), \tag{1.43}$$

and

$$PK_0^\dagger K_0 P \propto P + O(t^{3/2}), \tag{1.44}$$

for all $j, k \geq 1$. In terms of operators appearing in the Lindblad equation, these requirements are equivalent to

$$PL_j P \propto P, \tag{1.45}$$

$$PL_j^\dagger L_k P \propto P, \tag{1.46}$$

and

$$PHP \propto P, \tag{1.47}$$

respectively [18, 19]. This means that to suppress the dissipative part of a Lindblad dynamics to leading order in time, one should use a code for which $\{I, L_j\} \subset \mathscr{E}$. If one also wants to suppress the Hamiltonian part of the dynamics through QEC, one also needs $H \in \mathscr{E}$. (This may or may not be desirable, and can also be done e.g., through dynamical decoupling, as we will see in Part II.) The idea is to perform QEC recoveries frequently as compared to the relevant timescale set by $\{L_j\}$ and/or $H$ (often in a rotating frame), so as to keep uncorrected $O(t^2)$ terms from becoming important. One could also use a code which corrects to higher orders in $t$ through a broader $\mathscr{E}$, thus further reducing—though never completely nullifying—the logical error rate.

**Example 3.1: Independent Dephasing**

Consider for illustration the purely dissipative Lindblad equation

$$\frac{d\rho}{dt} = \frac{1}{2T_2^*} \sum_{j=1}^{3} \Big( Z_j \rho Z_j - \rho \Big) \tag{1.48}$$

on $n = 3$ qubits. One can easily show that the resulting dynamics coincides with the channel $\mathcal{K}_1^{\otimes 3}$ from Example 2, with $p = (1 - e^{-t/T_2^*})/2$. The 3-qubit repetition code can therefore largely suppress this decoherence when recoveries are repeated frequently compared to the dephasing time $T_2^*$. Over longer timescales, however, higher-order errors start to become significant, causing substantial decoherence at the logical level.

**Example 3.2: Independent Dephasing with a Hamiltonian**

Suppose there is also a Hamiltonian component to the above dynamics, namely:

$$\frac{d\rho}{dt} = -i[H, \rho] + \frac{1}{2T_2^*} \sum_{j=1}^{3} \Big( Z_j \rho Z_j - \rho \Big). \tag{1.49}$$

for $H = \frac{1}{2}(\omega_1 Z_1 + \omega_2 Z_2 + \omega_3 Z_3)$. An initial logical state of the repetition code $\rho_{\text{L}}$ that is corrected after a time $t$ will become

$$\rho_{\text{L}} \mapsto \mathcal{R}(e^{\mathcal{L}t} \rho_{\text{L}}) = \rho_{\text{L}} + O(t^2), \tag{1.50}$$

regardless of the exact values of $\omega_j$ and $1/T_2^*$, provided all are sufficiently small compared to $t^{-1}$ for this perturbative expansion to be meaningful. This insensitivity of the code to the precise dynamics is because the Kraus operators of $e^{\mathcal{L}t}$, truncated to order $O(t)$, are contained in the code's error "net" $\mathscr{E} = \text{span}\{I, Z_1, Z_2, Z_3\}$. Note however, that these short-time Kraus operators are now not equal to—nor proportional to—the error operators $I, Z_1, Z_2$, and $Z_3$; they only have the same span. This illustrates a further separation between the two.

**Example 3.3: Correlated Dephasing with a Hamiltonian**

Finally, consider the Lindblad equation

$$\frac{d\rho}{dt} = -i[H,\rho] + \frac{1}{2T_2^*}\sum_{j,k=1}^{3}\Big(Z_j\rho Z_k - \frac{1}{2}\{Z_jZ_k,\rho\}\Big), \tag{1.51}$$

where

$$H = \frac{\omega}{2}(Z_1 + Z_2 + Z_3). \tag{1.52}$$

We will see how such dynamics can arise in Part II. Upon diagonalizing $\mathcal{D}$, one finds that it has but a single non-vanishing jump operator, $L_1 \propto Z_1 + Z_2 + Z_3 \propto H$. Of course, the repetition code can still reverse this dynamics to leading order in time for logical states. Here, however, not only are $H$ and $L_1$ not Pauli operators, they (together with $I$) do not span $\mathscr{E}$ for this code. Rather, they span only a subset of $\mathscr{E}$ (namely, $\text{span}\{I, Z_1 + Z_2 + Z_3\}$), suggesting that the repetition code might be overkill here; i.e., that its error "net" is bigger than needed. This is indeed the case. Consider instead the following code on any two qubits: $|0_{\text{L}}\rangle = |01\rangle$ and $|1_{\text{L}}\rangle = |10\rangle$. Not only does this code satisfy the Knill-Laflamme condition, it is a "decoherence-free subspace" (DFS) of this dynamics, meaning that there is no need to actually correct errors (i.e., $\mathcal{R} = \mathcal{I}$), since the codespace is immune to them from the start. Note that this code/DFS does not correct for single-qubit phase flips, by design. Rather, it uses our knowledge of the dynamics to cast a more targeted "net" $\tilde{\mathscr{E}}$, and achieves (in principle) a vanishing logical error rate while reducing the overhead by half. This example illustrates clearly the distinction between the error operators around which a code is designed and the physical Kraus operators to be corrected. Moreover, it shows that choosing error operators carefully in light of the decoherence at hand, and designing new QEC codes from these error operators, can be quite beneficial. We will expand broadly on this approach in later chapters.

### 1.2.4 Choosing Error Operators

Based on this last example, the reader may wonder why most of the codes we have considered so far were constructed using bit or phase flips as error operator. Assuming nature has no special affinity for Pauli matrices, why are they so central to these and many other QEC codes? It is not because they are the Kraus operators describing decoherence in all leading quantum devices. Rather, it is because they have convenient mathematical properties, which in turn, lead to QEC codes with nice features; notably, these codes require remarkably little knowledge of the decoherence mechanisms against which they protect. This is a potentially critical feature for implementing QEC in large systems, where fully characterizing decoherence is all but impossible. Of course, this generality comes at a cost: the Pauli-centric approach to QEC can require prohibitive overheads in many devices, and can be largely incompatible with near-term applications like quantum sensing. This would seem to severely limit the utility of these codes in existing and emerging quantum devices. In this thesis we will largely consider codes built around non-Pauli errors, designed specifically for such devices. Before getting into these, however, it is useful to establish a baseline by briefly discussing Pauli-based QEC codes.

Strong, many-body interactions are rare in quantum systems. This means we generically expect quantum devices to couple to their environments in a predominantly local way; that is, through an interaction Hamiltonian that acts non-trivially on only one qubit (or perhaps few qubits) at a time, e.g.,

$$H_{\text{int}} \approx \sum_j \left(I \otimes \cdots \otimes I \otimes \underbrace{H_j}_{\text{qubit } j} \otimes I \cdots \otimes I\right) \quad \otimes \underbrace{H_{E_j}}_{\text{environment}} . \tag{1.53}$$

Recall that $\{I, X, Y, Z\}$ is a basis for all Hermitian matrices on single qubits, of which $H_j$ is one. Similarly, any 2-qubit Hermitian matrix can be written as a linear combination of tensor products of $I, X, Y$ and $Z$, and so on for multi-qubit Hamiltonians. This means that, quite generally, a QEC code built around an error "net" $\mathscr{E}$ comprising all single-qubit Paulis can suppress the effects of local coupling to an environment,

as in Eq. (1.53), to leading order in time. More broadly, if $\mathscr{E}$ includes all Paulis on $\leq w$ qubits (i.e., of weight $\leq w$), the resulting code can suppress the effects of $w$-local environmental couplings to leading order in time, and those of single-qubit couplings to higher orders. Such a code could also potentially suppress unwanted interactions between qubits, or control imperfections, due to a similar argument.

It is worth pausing for a moment to reflect on this remarkable result. It means that QEC codes designed to correct local Pauli errors could suppress decoherence in quantum devices without one ever knowing precisely how this decoherence arose, provided it was not pathologically nonlocal. Of course, the rates and types of logical errors would generally depend on the nature of the decoherence in question.

Let us briefly discuss an important family of QEC codes based on Pauli error operators, called stabilizer codes [7, 19, 20]. First, notice that any $n$-fold tensor product of $I$, $X$, $Y$ and $Z$ (e.g., $X \otimes I \otimes Z \otimes I$ for $n = 4$) has at most two eigenspaces, with eigenvalues $\pm 1$. Moreover, any two such operators either commute and anti-commute. In stabilizer codes, the codespace $\mathscr{C}_0$ is defined as the intersection of $\lambda = +1$ eigenspaces of several such operators $S_1, S_2, \ldots$, all of which commute with one another. That is,

$$\mathscr{C}_0 = \big\{ |\psi\rangle : S_j |\psi\rangle = |\psi\rangle \ \forall \, S_j \big\}. \tag{1.54}$$

If one can find $\{S_j\}$ such that either $E_k^\dagger E_\ell$ anti-commutes with some $S_j$, or $E_k^\dagger E_\ell$ is in the group generated by $\{S_j\}$ under multiplication, for all Pauli error operators $E_k, E_\ell \in \mathscr{E}$, the resulting $\mathscr{C}_0$ will satisfy the Knill-Laflamme condition for $\mathscr{E} = \text{span}\{E_k\}$. Repetition codes are stabilizer codes: the $n = 3$ bit-flip code, for instance, is specified by $S_1 = Z_1 Z_2$ and $S_2 = Z_2 Z_3$, and the corresponding phase-flip code by $S_1 = X_1 X_2$ and $S_2 = X_2 X_3$. The smallest code that can correct arbitrary single-qubit errors is also a stabilizer code, which uses $n = 5$ physical qubits to encode $k = 1$ logical qubit [21, 22], as specified by

$$\begin{aligned} S_1 &= X_1 Z_2 Z_3 X_4 \qquad & S_2 &= X_2 Z_3 Z_4 X_5 \\ S_3 &= X_1 X_3 Z_4 Z_5 \qquad & S_4 &= Z_1 X_2 X_4 Z_5. \end{aligned} \tag{1.55}$$

Note that it is common not to include ancillas when reporting the number of qubits $n$ in a code. We will follow this convention throughout this thesis.

Stabilizers codes have a number of desirable properties, besides casting a wide net $\mathscr{E}$. For one, they provide a very efficient description of the codespace itself. For instance, the operators in Eq. (1.55) are quite simple, whereas the logical states they imply are rather cumbersome. Stabilizer codes also provide a nice framework for encoded operations, and for analyzing fault tolerance: self-stabilizing QEC which can also handle control and measurement imperfections, and ancilla errors [19]. In this thesis we will focus on a limited family of encoded operations, and often assume a regime in which decoherence on idling qubits is the dominant error source. Accordingly, we will not elaborate on these two features here.

One final feature of stabilizer codes that will be directly relevant to us is the way in which one can detect/correct errors. The detection step of measuring which subspace $\mathscr{C}_0, \mathscr{C}_1, \mathscr{C}_2, \ldots$ the system is in corresponds to measuring whether it is in the $+1$ or $-1$ eigenspace of each $S_j$. There is a common trick for doing this: one performs a controlled-$S_j$ operation using an ancilla in state $|+\rangle$ as the control. Because $S_j$ is a tensor product of Paulis, this controlled-$S_j$ can be straightforwardly be decomposed into a sequence of two-qubit controlled operations. For a general system state $\alpha |\psi_+\rangle + \beta |\psi_-\rangle$, where $S_j |\psi_\pm\rangle = \pm |\psi_\pm\rangle$, the combined effect on ancilla $\otimes$ system is:

$$\begin{aligned} |+\rangle \left(\alpha |\psi_+\rangle + \beta |\psi_-\rangle\right) = \frac{1}{\sqrt{2}}\Big[ |0\rangle \left(\alpha |\psi_+\rangle + \beta |\psi_-\rangle\right) + |1\rangle \left(\alpha |\psi_+\rangle + \beta |\psi_-\rangle\right)\Big] \\ \updownarrow \text{ controlled-}S_j \\ \frac{1}{\sqrt{2}}\Big[ |0\rangle \left(\alpha |\psi_+\rangle + \beta |\psi_-\rangle\right) + |1\rangle \left(\alpha |\psi_+\rangle - \beta |\psi_-\rangle\right)\Big] = \alpha |+\rangle |\psi_+\rangle + \beta |-\rangle |\psi_-\rangle . \end{aligned} \tag{1.56}$$

One then applies a Hadamard gate to the ancilla (which maps $|0\rangle \leftrightarrow |+\rangle$ and $|-\rangle \leftrightarrow |1\rangle$), and measures it in the $\{|0\rangle, |1\rangle\}$ basis. With probability $|\alpha|^2$ one gets 0 and projects the system into the $+1$ eigenspace of $S_j$, otherwise the system is projected into the $-1$ eigenspace. This trick is sometimes called a "phase kickback," and it constitutes a simple instance of a more general procedure called quantum phase estimation [23]. The full error syndrome can be found by performing such measurements

for each $S_j$ (the order is immaterial since $[S_j, S_k] = 0$). One can then correct the inferred error by applying appropriate Pauli operators to the errant system qubit(s). Notice that this is precisely how the recovery is implemented in Fig. 1-4. The main appeal of this procedure, for our purposes, is that it gives the recovery an explicit form in terms of 1- and 2-qubit operations. In devising and analyzing QEC codes tailored to physical decoherence processes in subsequent chapters, we will see that certain features of stabilizer codes, such as simple recoveries using phase kickbacks, carry over straightforwardly in some settings and less so in others.

# Chapter 2

# Efficient QEC of Dephasing Induced by a Common Fluctuator

Consider a register of $n$ qubits for which the dominant decoherence mechanism is independent relaxation of each qubit to its ground state $|0\rangle$. This might occur, for instance, due to thermalization with an environment at low temperature compared to the qubits' energy gaps. To leading order in time, one could understand this process as $\sigma_- = |0\rangle\langle 1|$ acting independently on each qubit with some probability per unit time. What is the smallest number of qubits $n$ required to correct the resulting decoherence through QEC? One can easily check that no $n = 3$ repetition code will work; a $\sigma_-$ error is somehow more damaging than a bit or phase flip (for one, $\sigma_-$ has no inverse). Notice, though, that $\sigma_- \propto X + iY$. Therefore as per the previous chapter, one could realize a protected logical qubit with $n = 5$ physical qubits through the code defined by Eq. (1.55). However, this code is overkill to some extent: it can correct arbitrary (complex) linear combinations of $X$, $Y$, and $Z$ acting locally, whereas the physics here specifies a fixed combination $X+iY$. This additional structure suggests that one look for a QEC code tailored to the errors that actually occur, rather than hypothetical combinations of $X$, $Y$, and $Z$. Indeed, such a code was discovered in Ref. [24], which uses only $n = 4$ qubits. It has since been generalized to efficiently encode multiple logical qubits for larger $n$ [25].

More recently, it has been shown that the surface code, a promising type of large-

scale QEC code, can have a substantially higher fault-tolerance threshold when errors along $Z$ occur at higher rates than those along $X$ and $Y$, and one exploits this structure [26, 27]. This is a welcome discovery, since dephasing (broadly, $Z$ errors) is the strongest form of decoherence in many types of qubits. (Whether this can remain meaningfully true also during gates is less clear [28].) Similarly, codes based on field excitations in microwave resonators which deal efficiently with photon loss—often the dominant decoherence mechanism in such devices—have been very successful of late [29]. These three examples illustrate an important point: while QEC does not necessarily require a detailed understanding of the decoherence in a device, it can nevertheless be very advantageous to exploit whatever understanding one has. In large devices implementing the surface code, for instance, one may know fairly little about the decoherence mechanisms—perhaps simply that $Z$ errors are more common than others. In smaller, nearer-term devices, however, one often has a good understanding of the physics underlying the dominant decoherence mechanisms. In either regime, it is likely worthwhile to tailor the code for the decoherence at hand in proportion to one's understanding of it.

This is not a new idea; for instance, Nielsen and Chuang counsel ([7] §10.6.4):

> "A simple but important guiding principle is to choose your codes well. ...It may pay handsome dividends at the first level of concatenation[1] to use a code optimized to protect against the type of noise known to occur in the particular physical system being used for implementation."

While this idea may be uncontroversial, few concrete instances of such codes are known beyond the aforementioned ones. In this chapter we will construct and analyze a new family of such codes. They are designed to correct a particular decoherence mechanism that dominates in many nitrogen-vacancy (NV) center experiments, and can be important in several other platforms as well. They are fundamentally different than the 5-qubit or repetition codes from the previous chapter, in that they are

[1] One can create larger QEC codes by "concatenating" smaller ones to achieve better protection, by treating the logical qubits of a first code as physical qubits of a second one. We will touch on this approach only briefly in this thesis, focusing instead on the first level mentioned here.

not applicable in all devices, i.e., they are not one-size-fits-all. In fact, that is the point. We will find them to be far more efficient than conventional codes precisely because they correct the narrow range of errors describing the dominant decoherence mechanism in these devices. In the language of Chapter 1, these new codes cast a targeted net.

## 2.1 Decoherence Mechanism

The decoherence mechanism we will seek to correct is one in which a register of qubits dephases primarily due to energy-eigenstate-preserving (i.e., longitudinal) coupling of each qubit to a common fluctuator, which in turn exchanges energy with an external environment. That is, we consider a Hamiltonian

$$H = H_f^0 + \frac{1}{2}\sum_{j=1}^{n} \omega_j Z_j + H_f^{\text{int}} \otimes \sum_{j=1}^{n} g_j Z_j \tag{2.1}$$

for the fluctuator (labeled $f$) and the register qubits (labeled by $1 \leq j \leq n$), where $[H_f^0, H_f^{\text{int}}] = 0$. This type of coupling arises often as an effective description of a more general interaction. Consider, for instance, a qubit coupled to the fluctuator through some generic interaction Hamiltonian. In the interaction picture (i.e., the rotating frame generated by $H_f^0 + \frac{\omega_j}{2} Z_j$), the eigenstate-preserving term in Eq. (2.1) would remain time-independent, whereas the other terms would be oscillatory. If the qubit-fluctuator coupling is weak compared to $\omega_j$ and to the energy eigenvalue differences of $H_f^0$, these oscillating terms would average out quickly (on a timescale determined by the two latter energy scales), and have little impact. We could therefore neglect them to a good approximation—regardless of our choice of frames—leading to Eq. (2.1).

As described above, the fluctuator in this model jumps incoherently between its energy eigenstates $\{|\ell\rangle_f\}$, e.g., as

$$\frac{d\rho}{dt} = -i[H, \rho] + \sum_{k\ell} \gamma_{k\ell} \Big( A_{k\ell} \rho A_{k\ell}^\dagger - \frac{1}{2}\{A_{k\ell}^\dagger A_{k\ell}, \rho\} \Big), \tag{2.2}$$

where $A_{k\ell} = |k\rangle\langle\ell|_f \otimes I$ describes a jump from the fluctuator state $|\ell\rangle_f$ to $|k\rangle_f$ with rate $\gamma_{k\ell}$, and acts trivially on the register qubits. We will assume throughout that the fluctuator begins in some ensemble of energy eigenstates, rather than in a coherent superposition thereof, and therefore does not become entangled with the register. Notice that these Lindblad jump operators do not directly decohere the qubits; rather, they do so indirectly through the fluctuator. For this model to describe the dominant decoherence mechanism on the qubits, these must interact weakly with the broader environment, but strongly with the common fluctuator, which in turn couples strongly to said broader environment. Indeed, this is often true of nuclear spin qubits near an NV center in diamond at room temperature—the setting that motivated this work. In this platform, the nuclear spins are often used as a long-lived quantum memory, while the NV is used for sensing or to establish remote entanglement. The motivation here is to extend the lifetime of this quantum memory.

Carbon-13 nuclei in diamond have spin-1/2 and small magnetic dipole moments; they are therefore well-isolated in the absence of nearby defects in the diamond lattice. This gives them long coherence times even at room temperature, but also makes them slow to control and difficult to polarize at this temperature (i.e., to prepare in a pure, or nearly pure, state). The NV center is a type of defect in diamond with an effective spin-1 electronic degree of freedom, together with higher-energy states that will not be of direct concern here. Unlike the nuclear spin-$\frac{1}{2}$'s, it can be controlled rapidly and can easily be polarized, but it has much shorter coherence times ($T_2^{*(e)} \sim 4\,\mu$s [30]). An NV center can couple strongly to nearby $^{13}$C nuclear spins ($\lesssim 10\,$MHz [31]), allowing one to polarize and control these latter spins indirectly via the NV electronic spin (henceforth simply called the NV). Unfortunately, the NV often then becomes the limiting source of decoherence for these nuclear spins at room temperature. Specifically, coupling to the broader environment causes the NV to jump stochastically between its $|0\rangle$, $|-1\rangle$ and $|1\rangle$ energy eigenstates on a timescale $T_1^{(e)} \sim 5\,$ms [30]. These jumps perturb the nearby nuclear spins, setting the main limit on their coherence times. This reflects a central challenge in quantum engineering: the speed with which one can control quantum systems and the speed at which they

decohere are often closely related. This challenge will be a central theme in Part II.

The Hamiltonian for an NV coupled to $n$ nearby $^{13}$C nuclear spins can be expressed as

$$H = \underbrace{\Delta S_z^2 + \gamma_e B S_z}_{H_f^0} + \frac{\gamma_C B}{2} \sum_{j=1}^{n} Z_j + \frac{1}{2} S_z \otimes \sum_{j=1}^{n} (A_{zx}^{(j)} X_j + A_{zz}^{(j)} Z_j), \tag{2.3}$$

where $S_z = |1\rangle\langle 1| - |-1\rangle\langle -1|$ is the spin-1 equivalent of $Z$ on the NV, $\Delta$ is the NV zero-field splitting, $\gamma_e$ and $\gamma_C$ are the electronic and nuclear gyromagnetic ratios respectively, $B$ is the magnetic field component along an axis defined by the NV, and $A_{zx}^{(j)}$ and $A_{zz}^{(j)}$ are the hyperfine coupling strengths between the NV and the $j^{\text{th}}$ $^{13}$C nuclear spin. If one has nuclear spins with $A_{zx}^{(j)} \ll A_{zz}^{(j)}$, the previous equation immediately reduces to Eq. (2.1). Otherwise, one could use a strong magnetic field so that $\gamma_C B \gg A_{zx}$, to the same effect. (We will revisit this point later in the chapter.) Indeed, the reason why the NV couples to the nearby nuclei only via $S_z$ here is because one can neglect interaction terms which don't commute with $H_f^0$ (called non-secular terms) due to the large zero-field splitting $\Delta$, as per the argument above [32]. When $A_{zx}^{(j)}$ is negligible, the main form of decoherence in these carbon nuclei is dephasing due to the incoherent NV jumps, with a characteristic timescale $T_2^{*(n)} = 1.5 T_1^{(e)}$ [33]. Therefore, from the NV's point of view, the diamond is the environment. From the point of view of the nearby $^{13}$Cs, however, the environment is essentially just the NV—a highly structured environment.

The model in Eqs. (2.1) and (2.2), which we call *common-fluctuator dephasing* (CFD), can also be a significant source of decoherence in quantum dots and in superconducting qubits dispersively coupled to a common resonator[2] with non-zero effective temperature [33–46]. In these settings, as with spins in diamond, the register qubits are often read out and/or initialized via the fluctuator, imposing a lower limit on the desirable fluctuator-qubit coupling strengths, and making CFD a potentially important decoherence mode. Note that CFD does not generally produce a decoherence-free subspace (DFS).

[2]Such coupling induces an effective $\sigma_+\sigma_- + \sigma_-\sigma_+$ interaction between qubits, which could be suppressed through dynamical decoupling or through large detunings between qubits.

For the sake of generality, we now revert to the language of Eqs. (2.1) and (2.2). Moving to the interaction picture, the Hamiltonian (2.1) becomes

$$\tilde{H} = \sum_{\ell} \lambda_\ell \left|\ell\right\rangle\!\left\langle\ell\right|_f \otimes H_E, \tag{2.4}$$

where $H_f^{\text{int}} = \sum_\ell \lambda_\ell \left|\ell\right\rangle\!\left\langle\ell\right|_f$ and $H_E := \sum_{j=1}^n g_j Z_j$. When the fluctuator is in state $\left|\ell\right\rangle_f$, qubit $j$ has an effective Hamiltonian $\lambda_\ell\, g_j Z_j$ in the rotating frame. Jumps of the fluctuator therefore induce spatially-correlated random telegraph noise in the register, which causes dephasing [47, 48]. We will mostly be concerned here with the spatial correlations inherent in this noise. It is also interesting, however, to reflect on its temporal correlations for a moment. The noise source here—a single fluctuator—represents the opposite extreme of a large diffuse environment producing Gaussian noise through the central limit theorem, often with a $1/f$-like spectrum. (This is true at least when the couplings are strong compared to the rate of jumps in the fluctuator, which is the relevant regime with NV centers.) Indeed, the register qubits in this model have a highly structured environment. We will discuss temporal noise correlations in more detail in Chapter 4.

The standard QEC approach to correct dephasing is with the repetition code, built around error operators $E_j$'s which are Pauli $Z$ operators on at most $w$ qubits (and $I$ on the rest). There are $\sum_{m=0}^{w} \binom{n}{m}$ such matrices; a simple counting argument (the quantum Hamming bound applied to phase noise) therefore suggests that $n \geq 2w+1$ physical qubits are required to protect $k = 1$ logical qubit from arbitrary phase errors of weight $\leq w$ [7]. Indeed, the repetition code saturates this bound: the smallest instance uses $n = 3$ for $w = 1$, has logical states $\left|0_{\text{L}}\right\rangle = \left|{+}{+}{+}\right\rangle$ and $\left|1_{\text{L}}\right\rangle = \left|{-}{-}{-}\right\rangle$ where $\left|\pm\right\rangle := \frac{1}{\sqrt{2}}(\left|0\right\rangle \pm \left|1\right\rangle)$, and corrects for $\mathscr{E} = \text{span}\{I, Z_1, Z_2, Z_3\}$ as we have seen in the previous chapter. It can correct CFD as follows: In any run of the experiment, the register evolves over time $\delta t$ as

$$U(\theta) = e^{-i\theta H_E} \tag{2.5}$$

for some random variable $\theta \in [\delta t\, \lambda_{\min}, \delta t\, \lambda_{\max}]$ that depends on the fluctuator's trajectory. Specifically, if we define the stochastic process $\lambda(t)$ such that $\lambda(t) = \lambda_\ell$ when the fluctuator is in state $\ell$ at time $t$, then

$$\theta = \int_0^{\delta t} \lambda(t) dt. \tag{2.6}$$

For small $\theta = O(\delta t)$ (often reducible through dynamical decoupling[3] [33, 49–51], discussed further in Chapter 4), $U(\theta)$ can be approximated as

$$U(\theta) = I - i\theta H_E + O(\theta^2). \tag{2.7}$$

Since $\theta H_E \in \mathscr{E}$ regardless of $\theta$, the 3-qubit repetition code recovery performed after a time $\delta t$ corrects dephasing at order $O(\theta) = O(\delta t)$. More generally, $H_E^q$ contains Paulis of weight $\leq q$, so correcting to order $O(\theta^q) = O(\delta t^q)$ with the repetition code requires $n = 2q + 1$ qubits (for $k = 1$).

While the value of $\theta$ is unknown and varies from one run to the next, the coupling strengths $g_j$ are often fixed and well characterized. This suggests designing a code that corrects expressly for $\mathscr{E} = \text{span}\{I, H_E, H_E^2, \ldots, H_E^q\}$, and depends on the $\{g_j\}$ in a particular device. As per the previous chapter, such a code would cast a more targeted net by exploiting our understanding of the dominant decoherence mechanism. A similar counting argument as above suggests that such a code would require $q + 1$ subspaces to protect a logical qubit to order $O(\theta^q) = O(\delta t^q)$ in the time between recoveries, and therefore require

$$n = \lceil 1 + \log_2(q + 1) \rceil \tag{2.8}$$

qubits—an exponentially smaller overhead. We now proceed to construct a family of such codes for general $q$ and arbitrary coupling strengths $\{g_j\}$. We focus in particular on the $q = 1$ case, where one logical qubit is encoded in two physical qubits rather than three. We also construct recovery and logical operations for this code, which

[3] This could be on the register qubits and/or on the fluctuator itself.

can be implemented using a constant number of one- and two-qubit operations, much like in stabilizer codes.

## 2.2 Code Construction

We will construct these device-adapted QEC codes by enforcing the Knill-Laflamme condition for $k = 1$ logical qubit, in terms of as-yet-unknown codewords $|0_{\text{L}}\rangle$ and $|1_{\text{L}}\rangle$. For $k = 1$ and $\mathscr{E} = \text{span}\{H_E^j\}_{j=0}^q$, this amounts to demanding

$$\langle 0_{\text{L}}| H_E^m |0_{\text{L}}\rangle = \langle 1_{\text{L}}| H_E^m |1_{\text{L}}\rangle \tag{2.9}$$

$$\langle 0_{\text{L}}| H_E^m |1_{\text{L}}\rangle = 0 \tag{2.10}$$

for $0 \leq m \leq 2q$, where we consider values of $q$ that saturate the ceiling in Eq. (2.8) (that is, $q = 2^{n-1} - 1$). Finding a QEC code that corrects this $\mathscr{E}$ therefore requires finding logical states $|0_{\text{L}}\rangle$ and $|1_{\text{L}}\rangle$ that satisfy Eqs. (2.9) and (2.10). We begin with the ansatz

$$|0_{\text{L}}\rangle = \sum_{j=0}^{2^n-1} r_j e^{i\theta_j} |j\rangle \qquad\qquad |1_{\text{L}}\rangle = \sum_{j=0}^{2^n-1} r_{(2^n-1-j)} e^{i\phi_j} |j\rangle\,, \tag{2.11}$$

for $r_j, \theta_j, \phi_j \in \mathbb{R}$, where we use $|j\rangle$ to denote the $n$-bit binary representation of the integer $j$. That is, we fix the amplitudes of $|1_{\text{L}}\rangle$ to be those of $|0_{\text{L}}\rangle$ in reverse order. Notice that Eq. (2.11) always satisfies (2.9) for even $m \geq 0$, since $X^{\otimes n} H_E^m X^{\otimes n} = (-1)^m H_E^m$. For odd $m$:

$$\langle 0_{\text{L}}| H_E^m |0_{\text{L}}\rangle = -\langle 1_{\text{L}}| H_E^m |1_{\text{L}}\rangle = \vec{z} \cdot \vec{v}_m, \tag{2.12}$$

where $\vec{z}$, $\vec{v}_m \in \mathbb{R}^{q+1}$ are defined as $z_i = \langle i| Z_{\text{L}} |i\rangle$, with $Z_{\text{L}} := |0_{\text{L}}\rangle\langle 0_{\text{L}}| - |1_{\text{L}}\rangle\langle 1_{\text{L}}|$, and $(\vec{v}_m)_i = \langle i| H_E^m |i\rangle$ for $i \in [0, q]$ and odd $m \in [0, 2q]$. Therefore, Eq. (2.9) is satisfied for all relevant $m$ if $\vec{z} \perp \text{span}\{\vec{v}_m\}$. We can always find such a $\vec{z}$ ($\neq \vec{0}$) since the $\vec{v}_m$'s have dimension $q + 1$ but there are only $q$ of them, so they cannot form a complete basis. One approach is to construct a matrix $V$ with $\vec{v}_m$'s as columns; then, $I - VV^+$

projects onto span$\{\vec{v}_m\}^\perp$ (where $+$ and $\perp$ denote the pseudoinverse and orthogonal complement, respectively) and therefore has at least one real eigenvector $\vec{u}$ with unit eigenvalue[4]. Taking $\vec{z} = \vec{u}/||\vec{u}||_1$ satisfies Eq. (2.9) since $\vec{u} \cdot \vec{v}_m = 0$ automatically. Finally, building upon a technique developed in Ref. [52] for optimization, we pick $r_j$'s as

$$(r_j, r_{(2^n-1-j)}) = \begin{cases} (0, \sqrt{z_j}), & \text{if } z_j \geq 0 \\ (\sqrt{-z_j}, 0), & \text{if } z_j < 0. \end{cases} \tag{2.13}$$

This choice ensures that $\langle j|0_{\text{L}}\rangle$ or $\langle j|1_{\text{L}}\rangle$ vanishes for every $j$, thus satisfying Eq. (2.10). We now have normalized logical states that form a valid QEC code for all $q \geq 1$. Notice that the components of $|0_{\text{L}}\rangle$ and $|1_{\text{L}}\rangle$ generically have unequal amplitudes $r_j$ by necessity, in marked contrast with classical error-correcting codes and most known QEC codes. The phases $\theta_j$ and $\phi_j$ can be chosen arbitrarily—we demonstrate a convenient choice below. The performance of these codes on $n \leq 5$ qubits is shown in Fig. 2-1 using an illustrative model of a normally-distributed $\theta$ (which could arise, e.g., by approximating the random telegraph noise here as a Gaussian process, see Chapter 4). In addition, we give the pseudothresholds for $n = 2$ and 3 under the same model in Section 2.3. Note that while these codes correct only for phase noise (which is classical, in a sense), they are nonetheless *quantum* error-correcting codes, as they protect quantum states by a means which has no classical analog, to our knowledge.

The remarkable performance of device-adapted codes in Fig. 2-1 reflects two underlying effects. First, a generic $n$-qubit state decoheres faster with larger $n$, as a rule of thumb. Since our adapted codes correct errors to the same order in $\delta t$ (or equivalently, $\theta$) as repetition codes using smaller $n$, they are less vulnerable to decoherence from the start. Second, the $y$-axis in Fig. 2-1 is the effective error probability $p$ averaged over all coupling strengths $g_j \in [0, 1]$. Unlike repetition codes, our codes can yield a DFS when one exists, in which case $p = 0$. Such DFS's become more common for larger $n$, since it becomes more likely that two elements of a random $(g_1, \ldots, g_n) \in [0, 1]^n$ be nearly equal. This further drives down the average $p$ as $n$

[4] Alternatively, the modified Gram-Schmidt procedure provides a less intuitive but more numerically stable method.

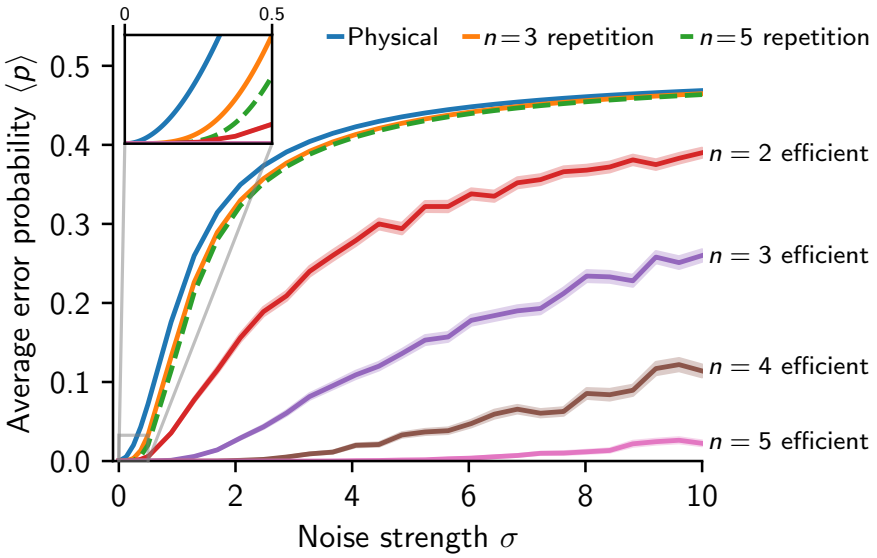



Figure 2-1: We assume that the effect of the quantum fluctuator is to impart a random phase, $\theta$, which follows a Gaussian distribution $\theta \sim \mathcal{N}(0, \sigma)$ with standard deviation $\sigma$. By normalizing the $g_j$'s to lie in $[0, 1]^n$, $\sigma$ describes the noise strength. CFD followed by a QEC recovery (if applicable) results in an effective phase- or bit-flip channel $\rho \mapsto (1-p)\rho + p\, A\rho A$, where $A = Z$ for the physical qubits, $X_{\text{L}} = |0_{\text{L}}\rangle\langle 1_{\text{L}}| + |1_{\text{L}}\rangle\langle 0_{\text{L}}|$ for the repetition codes, and $Z_{\text{L}}$ for hardware-efficient codes. The average infidelity, average trace distance and diamond distance to $I$ are all $\propto p$. As the performance of all strategies shown depends on $\{g_j\}$, we plot the average of $p$ over $\{g_j\} \in [0, 1]^n$. The error bands for the hardware-efficient codes denote the standard error of the mean from Monte Carlo integration. More details on the numerical implementation are given in Section 2.3 and Appendix A.

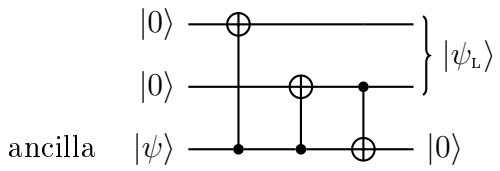



Figure 2-2: An encoding procedure for the toy model code in Eq. (2.14), where $|\psi_{\text{L}}\rangle = \alpha\,|0_{\text{L}}\rangle + \beta\,|1_{\text{L}}\rangle$ for arbitrary $\alpha$ and $\beta$.

grows.

To understand the smallest, i.e., $n = 2$ qubit version of this code, it is useful first to consider a toy model in which decoherence arises entirely differently. Consider a 2-qubit register where the first qubit is subject only to $X_1$ errors, and the second qubit experiences no errors. Of course, the easiest approach would be to simply not use the first, noisy qubit. For our purposes, however, imagine dealing with $\mathscr{E} = \text{span}\{I, X_1\}$ using a stabilizer QEC code with $S = Z_1 Z_2$, whose codewords we can choose as

$$|0_{\text{L}}\rangle = |00\rangle \qquad\qquad |1_{\text{L}}\rangle = |11\rangle\,. \tag{2.14}$$

An unknown state $|\psi\rangle$ on an ancilla could then be encoded using the circuit in Fig. 2-2 (and the process could be reversed by applying the operations backwards). A parity (i.e., $Z_1 Z_2 = S$) measurement on the two qubits would reveal whether an error occurred, that is, whether the state is in $\mathscr{C}_0 = \text{span}\{|0_{\text{L}}\rangle\,, |1_{\text{L}}\rangle\}$ or $\mathscr{C}_1 = \text{span}\{|0_{\text{E}}\rangle\,, |1_{\text{E}}\rangle\}$, where

$$|0_{\text{E}}\rangle := |10\rangle = X_1\,|0_{\text{L}}\rangle \qquad\qquad |1_{\text{E}}\rangle := |01\rangle = X_1\,|1_{\text{L}}\rangle\,. \tag{2.15}$$

One could then apply $X_1$ if needed, as shown in Fig. 2-3.

We now analyze explicitly the smallest case of our QEC code, which uses $n = 2$ qubits. As we will see, the encoding and recovery operations are very similar to those in the toy example above, even though the underlying decoherence is completely different. We will label the register qubits 1 and 2 such that $|g_1| \geq |g_2|$. Note that here—and in general—$H_E = g_1 Z_1 + g_2 Z_2$ is a combination of weight-1 Pauli operators, not a weight-2 Pauli. This $H_E$ gives $\vec{v}_1 = (g_1 + g_2, g_1 - g_2)^\top$. The matrix

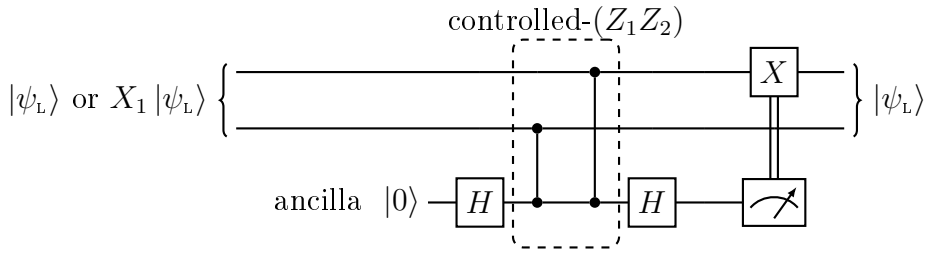



Figure 2-3: A recovery procedure for the toy model code in Eq. (2.14), where $|\psi_{\text{L}}\rangle = \alpha|0_{\text{L}}\rangle + \beta|1_{\text{L}}\rangle$ for arbitrary $\alpha$ and $\beta$, and $H$ denotes a Hadamard gate.

$I - VV^+$ has only a 1-dimensional eigenspace with unit eigenvalue, spanned by $\vec{u} = (-g_1 + g_2, g_1 + g_2)^\top$, where $\vec{u} \cdot \vec{v}_1 = 0$. If $g_1 > 0$ we find $r_1 = r_3 = 0$ and

$$r_0 = c\sqrt{g_1 - g_2} \qquad\qquad r_2 = c\sqrt{g_1 + g_2}, \tag{2.16}$$

where $c = 1/\sqrt{||\vec{u}||_1}$. This gives logical states

$$|0_{\text{L}}\rangle = |\chi_0\rangle|0\rangle \qquad\qquad |1_{\text{L}}\rangle = |\chi_1\rangle|1\rangle \tag{2.17}$$

with

$$\begin{aligned} |\chi_0\rangle &= c\Big(\sqrt{|g_1 - g_2|}\, e^{i\theta_0}\ |0\rangle + \sqrt{|g_1 + g_2|}\, e^{i\theta_2}\ |1\rangle\Big) \\ |\chi_1\rangle &= c\Big(\sqrt{|g_1 + g_2|}\, e^{i\phi_1}\ |0\rangle + \sqrt{|g_1 - g_2|}\, e^{i\phi_3}\ |1\rangle\Big). \end{aligned} \tag{2.18}$$

The $g_1 < 0$ case gives the same result up to a relabeling of $|0_{\text{L}}\rangle \leftrightarrow |1_{\text{L}}\rangle$. This code corrects for $\mathscr{E} = \text{span}\{I, H_E\}$; by design, however, it does not correct for $Z_1Z_2$, nor $Z_1$ or $Z_2$ individually, none of which are in $\mathscr{E}$. Rather, it corrects CFD with fewer qubits than the smallest repetition code precisely because we have chosen not to correct individual Pauli operators.

Observe that Eqs. (2.17) and (2.18) reduce to a DFS in the limit where one exists ($|g_1| = |g_2|$), but this is in practice rare. More generally, notice that the choice $\theta_0 = \phi_1 + \pi = -\theta_2 = -\phi_3 = \vartheta$ for arbitrary $\vartheta$ proves convenient: First, it gives $\langle\chi_0|\chi_1\rangle = 0$, which means an ancilla state $|\psi\rangle$ can be encoded onto the register

through a circuit like that in Fig. 2-2, but with $X_1$ replaced by

$$U_x := |\chi_0\rangle\langle\chi_1| + |\chi_1\rangle\langle\chi_0| \,, \tag{2.19}$$

a $\pi$ rotation about some axis determined by $g_1$, $g_2$ and $\vartheta$, as shown in Fig. 2-4. Moreover, this choice of phases results in a simple action of $H_E$ on logical states:

$$\begin{aligned} H_E \, |0_{\text{L}}\rangle &\propto |\chi_1\rangle \, |0\rangle =: |0_{\text{E}}\rangle \\ H_E \, |1_{\text{L}}\rangle &\propto |\chi_0\rangle \, |1\rangle =: |1_{\text{E}}\rangle \,, \end{aligned} \tag{2.20}$$

which closely resembles the effect of $X_1$ in the toy model, but in the $\{|\chi_0\rangle \, , |\chi_1\rangle\}$ basis on qubit 1. Both lines have the same proportionality constant, and we have defined the error states $|0_{\text{E}}\rangle$ and $|1_{\text{E}}\rangle$. We emphasize that since $H_E$ cannot generically be decomposed as a tensor product, it maps most separable states to entangled states; Eq. (2.20)—wherein the first qubit is "flipped" by $H_E$—is due to our choice of $|0_{\text{L}}\rangle$ and $|1_{\text{L}}\rangle$. Second, consider the orthogonal projectors $P = |0_{\text{L}}\rangle\langle 0_{\text{L}}| + |1_{\text{L}}\rangle\langle 1_{\text{L}}|$ and $P_{\text{E}} = |0_{\text{E}}\rangle\langle 0_{\text{E}}| + |1_{\text{E}}\rangle\langle 1_{\text{E}}|$ onto $\mathscr{C}_0 = \text{span}\{|0_{\text{L}}\rangle \, , |1_{\text{L}}\rangle\}$ and $\mathscr{C}_1 = \text{span}\{|0_{\text{E}}\rangle \, , |1_{\text{E}}\rangle\}$ respectively ($\mathscr{H} = \mathscr{C}_0 \oplus \mathscr{C}_1$). One can detect an error non-destructively by measuring parity in the $|\chi_i\rangle \, |j\rangle$ basis, which can be done by performing phase estimation/kickback on

$$S = P - P_{\text{E}} = U_z \otimes Z \tag{2.21}$$

with an ancilla, rather than on $Z \otimes Z$ as in the toy example [23]. Crucially, the choice of phases in $|0_{\text{L}}\rangle$ and $|1_{\text{L}}\rangle$ makes $S$ separable here, where $U_z := |\chi_0\rangle\langle\chi_0| - |\chi_1\rangle\langle\chi_1|$ is a $\pi$ rotation about a different axis. This means that the controlled-$S$ (c$S$) operation used to measure the error syndrome can be implemented through a pair of two-qubit operations (c$U_z$ and c$Z$), rather than a more challenging 3-qubit operation. If an error is detected, it can be corrected by applying $U_x$ to qubit 1. (Both $U_x$ and $U_z$ could be synthesized out of a constant number of Pauli rotations, or implemented directly, e.g., by driving qubit 1 off resonance [53].) The full recovery procedure, which corrects CFD to leading order, is shown in Fig. 2-5. It is identical to that in

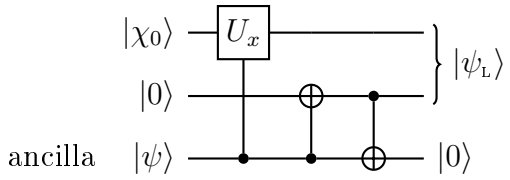



Figure 2-4: An encoding procedure for $n = 2$ qubits where $|\psi_{\rm L}\rangle = \alpha\,|0_{\rm L}\rangle + \beta\,|1_{\rm L}\rangle$ for arbitrary $\alpha$ and $\beta$. The unitaries $U_x$ is a $\pi$ rotation about an axis on the Bloch sphere determined by $g_1$, $g_2$ and $\vartheta$. Notice that if the NV electronic spin is used an ancilla, the only 2-qubit gates in this circuit are between it and the nuclear spins (as opposed to between two weakly-coupled nuclear spins).

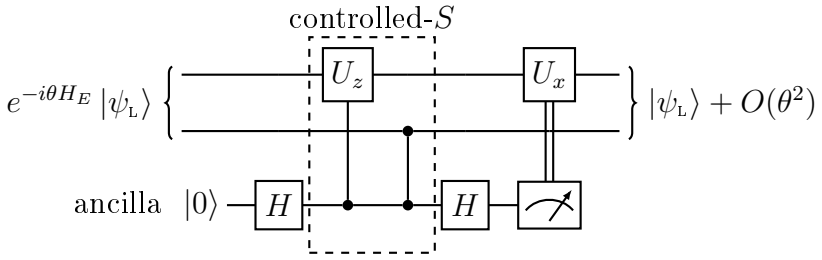



Figure 2-5: A recovery procedure for $n = 2$ qubits where $H$ denotes a Hadamard gate, and $\theta$ is a random variable. Notice, as in Fig. 2-4, that if the NV electronic spin is used an ancilla, the only 2-qubit gates in this circuit are between it and the nuclear spins (as opposed to between two weakly-coupled nuclear spins) [32]. Note that in NV implementations it may be advantageous to avoid measuring the NV through repetitive readout to detect errors, as this can be disruptive to nearby nuclear spins. Instead, one could perform a controlled-$U_x$ followed by an NV reset.

Fig. 2-3, but with $X_1 \to U_x$ and $Z_1 \to U_z$. Note that $S$ behaves like a stabilizer, in the sense of its action on $\mathscr{C}_0$ and $\mathscr{C}_1$. It does not, however, fit in the usual QEC stabilizer formalism since $\{H_E, S\} \neq 0$ generically, because $\{H_E, S\}|\psi\rangle = 0$ for $|\psi\rangle \in \mathscr{C}_0$ but not for $|\psi\rangle \in \mathscr{C}_1$ [20]. This is because $H_E$ maps $\mathscr{C}_0$ to $\mathscr{C}_1$ without distortion, but not vice-versa, as $H_E$ is not generically a tensor product of Paulis. (Neither is $S$.) In spite of these unusual features, the procedure for feeding back on $S$ in Fig. 2-5 is largely the same as that of the usual stabilizer formalism. Finally, there is a simple way to implement any logical unitary $U_{\rm L}$ in this code: apply the corresponding physical $U$ to qubit 2 followed by a recovery.

The logical states derived above are also valid for all $q > 1$ (i.e., $n > 2$ qubits), but the corresponding recovery and logical operations are generally more involved.

Generically, the analogues[5] of $S$ in (2.21) need not be separable for any choice of $\theta_j$ and $\phi_j$. One might still synthesize them with one- and two-qubit operations, perform phase kickback through optimal control, or implement a QEC recovery via more general channel-engineering techniques [54–57]. More efficient solutions could even be found by analyzing specific experimental scenarios. One approach could be for example to use devices with $\{g_j\}$ chosen so that the recovery and logical operations (or good approximations of them, at least) can be conveniently implemented. One could also correct to a slightly lower order $q$ [i.e., maintaining $n = O(\log q)$ but not saturating the ceiling in Eq. (2.8)]; this would yield a continuous family of possible $\vec{z}$'s [cf. Eq. (2.13)], among which one might find codes with convenient QEC operations.

## 2.3 Code Performance

We now turn our attention to evaluating the codes we have constructed. We will compare their effective error rates to those of bare physical qubits and repetition codes, and examine how they perform under various imperfections and uncertainties. Like all QEC codes under this noise model, ours will perform differently for different values of the fluctuator-qubit coupling strengths $\{g_j\}$. As part of our experimental efforts, we have simulated specific candidate $g_j$'s. Here, however, it will often be more useful to report average performance over all possible coupling strengths for an idealized fluctuator, as we have done in Fig. 2-1. Monte Carlo integration is by far the most effective technique we found to perform this averaging. It involves constructing codewords and recoveries automatically for many random $\{g_j\}$. As the number of qubits $n$ grew, this process was limited not by the usual exponential growth in complexity, but by numerical instabilities. Even if numerical errors arose only in narrow regions of $\{g_j\}$-space, these could bias our Monte Carlo results. We have therefore taken great care to fend off such numerical issues for $n \leq 5$. In large part this has meant extensive analytical manipulations, which allowed us to avoid numerics

[5] e.g., $S_1 = P + P_1 - P_2 - P_3$ and $S_2 = P - P_1 + P_2 - P_3$, which could be measured sequentially to identify an error for $n = 3$

until the very last step. We present the details of this process below. Wherever possible, we have also compared our Monte Carlo results with exact or asymptotic expressions found (often with some difficulty) in Mathematica. These comparisons can be found in Appendix A, together with further details concerning our Monte Carlo integration.

### 2.3.1 Probability Distribution of $\theta$ and Effective Channel Form

In the interaction picture, CFD causes the register of qubits to evolve by $U(\theta) = e^{-i\theta H_{\text{E}}}$ in any run of the experiment, where $\theta$ is a random variable that depends on the fluctuator's dynamics, its coupling to the register, and the elapsed time. The fluctuator can behave differently in each run of the experiment, producing a different value of $\theta$ in each realization, and thus causing decoherence in the register.

As far as the register is concerned, the fluctuator's behavior can be fully captured by specifying a (classical) probability distribution for $\theta$. Generically, this distribution will depend on the underlying physics of the fluctuator. For instance, a fluctuator consisting of an NV at room temperature under dynamical decoupling can produce a somewhat different distribution for $\theta$ than a fluctuator consisting of a microwave resonator at cryogenic temperature. This variation between different systems precludes a fully general performance analysis not just of the QEC codes introduced here, but also of existing codes under CFD. Instead we consider a simple probability distribution for $\theta$ for the sake of illustration, both here and in Fig. 2-1 of the previous section.

In Fig. 2-1 we choose $\theta \sim \mathcal{N}(0, \sigma)$ to be normally distributed with zero mean and standard deviation $\sigma$. (A non-zero mean of $\theta$ could always be absorbed into the rotating frame to give the aforementioned distribution.) This choice has two main motivations:

1. It depends only on a single parameter $\sigma$ that can straightforwardly be interpreted as the noise strength, since we normalize the coupling strengths to lie in $[0, 1]$.

2. It induces a generally monotonic loss of coherence in the register with the noise strength $\sigma$. In contrast, more complicated distributions, such as those arising from random telegraph noise in certain regimes, can induce an oscillating loss of coherence. Such oscillations would becloud some of the analyses below. (For instance, there is some ambiguity in defining pseudothresholds when the physical and logical infidelities, or the like, display small oscillations and intersect several times.)

For the sake of comparison, however, we plot in Fig. 2-6 quantities analogous to those in Fig. 2-1, but with a different underlying distribution of $\theta$. Notice that the curves are deformed slightly from those in Fig. 2-1, but are qualitatively similar.

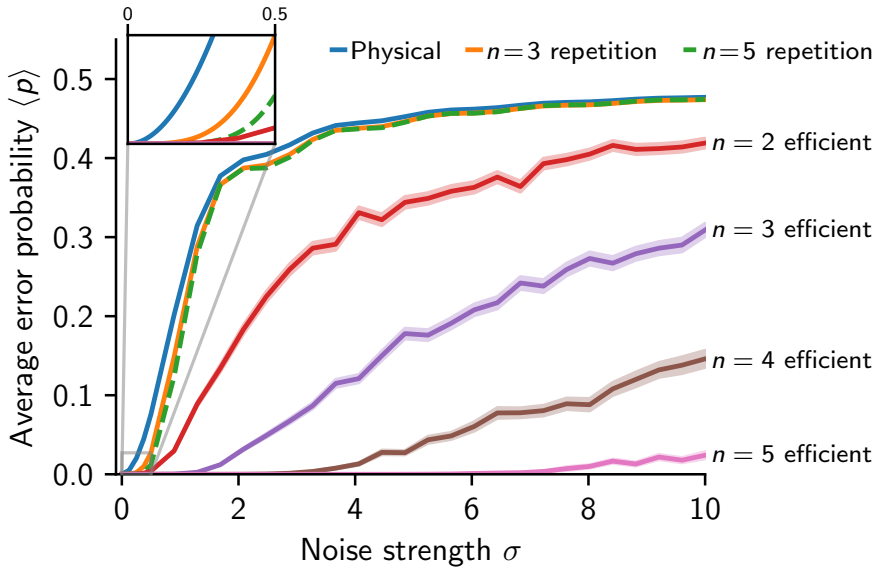



Figure 2-6: We assume that the effect of the quantum fluctuator is to impart a random phase, $\theta$. In contrast with Fig. 2-1, we consider here a uniformly distributed $\theta \sim \text{unif}(-\theta_{\max}, \theta_{\max})$ for some maximum rotation angle $\theta_{\max} \geq 0$. To simplify the comparison with Fig. 2-1, we plot $\langle p \rangle$ versus the standard deviation $\sigma = \theta_{\max}/\sqrt{3}$, rather than versus $\theta_{\max}$ directly. By normalizing the $g_j$'s to lie in $[0,1]^n$, $\sigma$ describes the noise strength. CFD followed by a QEC recovery (if applicable) results in an effective phase- or bit-flip channel $\rho \mapsto (1-p)\rho + p\, A\rho A$, where $A = Z$ for the physical qubits, $X_{\text{L}}$ for the repetition codes, and $Z_{\text{L}}$ for hardware-efficient codes. The average infidelity, average trace distance and diamond distance to $I$ are all $\propto p$. As the performance of all strategies shown depends on $\{g_j\}$, we plot the average of $p$ over $\{g_j\} \in [0,1]^n$. The error bands for the hardware-efficient codes denote the standard error of the mean from Monte Carlo integration.

Returning to $\theta \sim \mathcal{N}(0,\sigma)$, we now show that the effective logical channel describing CFD followed by a recovery has the form

$$\rho_{\text{L}} \mapsto (1-p)\rho_{\text{L}} + pZ_{\text{L}}\rho_{\text{L}}Z_{\text{L}}, \tag{2.22}$$

where $p$ depends on $\sigma$, $n$ and $\{g_j\}$. The average effect of $U(\theta)$ can be readily calculated for $\theta \sim \mathcal{N}(0,\sigma)$. First, we define the superoperators $\mathcal{H}_{\text{E}}$ and $\mathcal{U}_\theta$ by their action on a generic matrix $A$ as $\mathcal{H}_{\text{E}}(A) := [H_{\text{E}}, A]$ and $\mathcal{U}_\theta(A) := U(\theta)AU(\theta)^\dagger = e^{-i\theta\mathcal{H}_{\text{E}}}(A)$. To find $\langle\mathcal{U}\rangle := \int_{\mathbb{R}} \mathcal{U}_\theta\, p(\theta)\, d\theta$, the superoperator describing the register's average evolution, we begin by diagonalizing $\mathcal{H}_{\text{E}}$. Notice that if $H_{\text{E}}\,|j\rangle = E_j\,|j\rangle$, then $\{|j\rangle\langle k|\}_{j,k=0}^{2^n-1}$ is a complete eigenbasis for $\mathcal{H}_{\text{E}}$. In particular:

$$\mathcal{H}_{\text{E}}\big(\,|j\rangle\langle k|\,\big) = (E_j - E_k)\,|j\rangle\langle k|\,, \tag{2.23}$$

and so the spectral decomposition of $\mathcal{H}_{\text{E}}$ reads

$$\mathcal{H}_{\text{E}}(A) = \sum_{j,k=0}^{2^n-1} (E_j - E_k)\,\langle j|\,A\,|k\rangle\;\,|j\rangle\langle k|\,. \tag{2.24}$$

Therefore,

$$\langle\mathcal{U}\rangle(A) = \sum_{j,k=0}^{2^n-1} \langle e^{-i\theta(E_j-E_k)}\rangle\;\,\langle j|\,A\,|k\rangle\;\,|j\rangle\langle k|\,. \tag{2.25}$$

Finally, we use that the characteristic function of $\theta \sim \mathcal{N}(0,\theta)$ is $\langle e^{i\theta x}\rangle = e^{-\frac{\sigma^2 x^2}{2}}$ for any $x \in \mathbb{R}$ to arrive at

$$\langle\mathcal{U}\rangle(A) = \sum_{j,k=0}^{2^n-1} e^{-\sigma^2(E_j-E_k)^2/2}\;\,\langle j|\,A\,|k\rangle\;\,|j\rangle\langle k|\,, \tag{2.26}$$

or more compactly: $\langle\mathcal{U}\rangle = \exp(-\sigma^2\mathcal{H}_{\text{E}}^2/2)$. This last expression is convenient for numerical analyses, as it removes the need to average over many realizations of $\theta$ numerically. For $\theta \sim \text{unif}(-\theta_{\max}, \theta_{\max})$, a similar calculation gives $\langle\mathcal{U}\rangle = \text{sinc}(\mathcal{H}_{\text{E}}\,\theta_{\max})$.

We consider throughout the usual (i.e., transpose) recovery channel $\mathcal{R}(\rho) =$

$\sum_j U_j^\dagger P_j \rho P_j U_j$, where $U_j$ and $P_j$ are defined as in Section 1.2.2. A straightforward calculation allows one to express $\mathcal{R}$ directly in terms of $P$ and $H_{\rm E}$ as

$$\mathcal{R}(\rho) = \sum_{j,k=0}^{q} (M^+)_{jk} P H_{\rm E}^j \rho H_{\rm E}^k P, \tag{2.27}$$

where $M = (m_{jk})_{j,k=0}^q$ is the code matrix $P H_{\rm E}^j H_{\rm E}^k P =: m_{jk} P$ and $M^+$ is the Moore-Penrose pseudoinverse of $M$ (which reduces to $M^{-1}$ when $M$ is invertible). By construction, $m_{jk} = 0$ when $j + k$ is odd, i.e., when $j$ and $k$ have different parities. Furthermore, the $m_{jk}$'s are equal along anti-diagonal bands of $M$ with $j + k = \text{const}$. It follows that every eigenvector $\vec{u}$ of $M$ can be chosen such that $u_i = 0$ either for all even $i$ or for all odd $i$. This implies that $(M^+)_{jk} = 0$ when $j$ and $k$ have different parities, so the only non-vanishing terms in Eq. (2.27) will have $j$ and $k$ both even or both odd[6].

Notice that $(\mathcal{R}\mathcal{U}_\theta)(\rho_{\rm L})$ will comprise a weighted sum of terms with the form $\mathcal{R}(H_{\rm E}^\ell \rho_{\rm L} H_{\rm E}^m)$. Using that

$$P H_{\rm E}^m P = \frac{1}{2} \operatorname{tr}(Z_{\rm L}^m H_{\rm E}^m) Z_{\rm L}^m \tag{2.28}$$

from Eqs. (2.11)–(2.13), and that $\rho_{\rm L} = P \rho_{\rm L} P$ for encoded states, we find

$$\mathcal{R}(H_{\rm E}^\ell \rho_{\rm L} H_{\rm E}^m) = \frac{1}{4} \sum_{j,k=0}^{q} (M^+)_{jk} \operatorname{tr}(Z_{\rm L}^{j+\ell} H_{\rm E}^{j+\ell}) \operatorname{tr}(Z_{\rm L}^{k+m} H_{\rm E}^{k+m}) Z_{\rm L}^{j+\ell} \rho_{\rm L} Z_{\rm L}^{k+m}. \tag{2.29}$$

It follows immediately that $\mathcal{R}\langle\mathcal{U}\rangle = \langle\mathcal{R}\mathcal{U}_\theta\rangle$ must have the form

$$\langle\mathcal{R}\mathcal{U}_\theta\rangle(\rho_{\rm L}) = a_0 \rho_{\rm L} + a_1 Z_{\rm L} \rho_{\rm L} + a_2 \rho_{\rm L} Z_{\rm L} + a_3 Z_{\rm L} \rho_{\rm L} Z_{\rm L} \tag{2.30}$$

[6] One can also use this observation to separate $M$ into two submatrices with $(j,k) = (\text{even, even})$ or (odd, odd), for which the pseudoinverses can be computed separately and then combined to give $M^+$. This is not only faster, tends to be more numerically stable for larger $n$.

for appropriate $a_j$'s. In particular,

$$\langle \mathcal{R}\mathcal{U}_\theta \rangle(\rho_{\text{L}}) = \sum_{\ell,m=0}^{\infty} \frac{(-1)^\ell \, i^{\ell+m} \, \langle \theta^{\ell+m} \rangle}{\ell! \, m!} \mathcal{R}(H_{\text{E}}^\ell \, \rho_{\text{L}} \, H_{\text{E}}^m). \tag{2.31}$$

For $\theta \sim \mathcal{N}(0,\sigma)$, $\langle \theta^{\ell+m} \rangle = 0$ if $\ell + m$ is odd, i.e., if $\ell$ and $m$ have different parities. Therefore, the only non-vanishing terms in Eq. (2.31) have $\ell$ and $m$ both even or both odd. Comparing with Eq. (2.29), one immediately sees that the $a_1$ and $a_2$ cross-terms vanish. Finally, since $\langle \mathcal{R}\mathcal{U}_\theta \rangle$ is completely positive and trace-preserving (CPTP), we can write $a_0 = 1 - p$ and $a_3 = p$ for some $p \in [0,1]$, as claimed[7] in Eq. (2.22). For $n = 2$ the resulting function $p$ is

$$p = \frac{1}{4g_1^2} \, e^{-2(g_1+g_2)^2\sigma^2} \Big[ 2g_1^2 e^{2(g_1+g_2)^2\sigma^2} + (g_1 - g_2)g_2 - 2(g_1^2 - g_2^2) e^{2g_1(g_1+2g_2)\sigma^2} \\ - e^{8g_1 g_2 \sigma^2} g_2(g_1 + g_2) \Big], \tag{2.32}$$

where we have assumed $|g_1| \geq |g_2| \geq 0$ without loss of generality. The expressions for $p$ quickly become complicated for larger $n$.

More straightforwardly, the effect of CFD on the physical qubit $j$ can be shown to have the same form, except with $p = [1 - \langle \cos(2g_j\theta) \rangle]/2$, where $\langle \cos(2g_j\theta) \rangle = e^{-2g_j^2\sigma^2}$ for $\theta \sim \mathcal{N}(0,\sigma)$. Similarly, the logical channels for the $n = 3$ and $n = 5$ qubit repetition codes (i.e., phase-flip codes) can be shown to have the form of a bit-flip channel $\rho_{\text{L}} \mapsto (1-p)\rho_{\text{L}} + pX_{\text{L}}\rho_{\text{L}}X_{\text{L}}$, for different functions $p$. Alternatively, they could be expressed as phase-flip channels with the same $p$'s by using the labeling convention $|\pm_{\text{L}}\rangle = |\pm\rangle^{\otimes n}$. For $n = 3$,

$$p = \frac{1}{16} \Big( 8 - 4e^{-2g_1^2\sigma^2} - 4e^{-2g_2^2\sigma^2} - 4e^{-2g_3^2\sigma^2} + e^{-2(g_1+g_2+g_3)^2\sigma^2} + e^{8(g_1+g2)g_3\sigma^2 - 2(g_1+g_2+g_3)^2\sigma^2} \\ + e^{8g_2(g_1+g_3)\sigma^2 - 2(g_1+g_2+g_3)^2\sigma^2} + e^{8g_1(g_2+g_3)\sigma^2 - 2(g_1+g_2+g_3)^2\sigma^2} \Big). \tag{2.33}$$

[7] The same argument holds for symmetric distributions about $\theta = 0$ more generally, such as $\theta \sim \text{unif}(-\theta_{\max}, \theta_{\max})$.

The corresponding expression for $n = 5$ is more complicated. Notice that while the codewords for the repetition code do not depend on $\{g_j\}$, Eq. (2.33) does, via $\langle \mathcal{U} \rangle$.

Many common measures of performance, e.g., average infidelity, average trace distance or diamond distance from the identity channel, have the form $p \times$ const. for bit- or phase-flip channels. It is therefore highly convenient that the channels above have this form, as one can capture all of these performance measures at once (up to constant pre-factors) simply by considering $p$. This avoids any ambiguity in choosing one performance measure over others.

### 2.3.2 Pseudothresholds and Non-Commuting Interaction Terms

We now examine the pseudothresholds for our hardware-efficient codes with $\theta \sim \mathcal{N}(0, \sigma)$. That is, for a given $n$, we find the largest $\sigma$ for which $p$ of our codes is smaller than the $p$'s of all physical qubits. Naturally, this pseudothreshold value, which we denote $\sigma_{\text{th}}$, depends on $\{g_j\}$. As discussed above, the pseudothresholds derived from $p$ are the same as those from average infidelity, and average trace and diamond distance from the identity.

The left panel of Fig. 2-7 shows the pseudothresholds for the efficient code with $n = 2$. Notice that $\sigma_{\text{th}}$ diverges around the region where $g_1 = g_2$, as the codespace becomes a decoherence-free subspace (DFS) here. In order to put these values of $\sigma_{\text{th}}$ in perspective, we repeat this analysis for the 3-qubit repetition code. Its pseudothreshold under CFD will depend on $\{g_1, g_2, g_3\}$, making it difficult to visualize. To get around this difficulty, we fix $\max\{g_j\} = 1$ and leave the other coupling strengths free. One could interpret this as expressing $\sigma_{\text{th}}$ and $\{g_j\} \setminus \{g_{\max}\}$ in units of $g_{\max}$. To simplify comparison between two- and three-qubit codes, the right panel of Fig. 2-7 shows $\sigma_{\text{th}}$ in this way for the $n = 2$ efficient code. Fig. 2-8 then shows the pseudothresholds for the $n = 3$ repetition code (left) and efficient code (right) under CFD. We used different color maps to emphasize that these plots depict very different ranges of $\sigma_{\text{th}}$, since their features would be obscured if they were shown on a single shared color map with the same scale. The pseudothresholds in Figs. 2-7 and 2-8 assume perfect operations, exact knowledge of $\{g_j\}$, and no other decoherence besides the CFD.

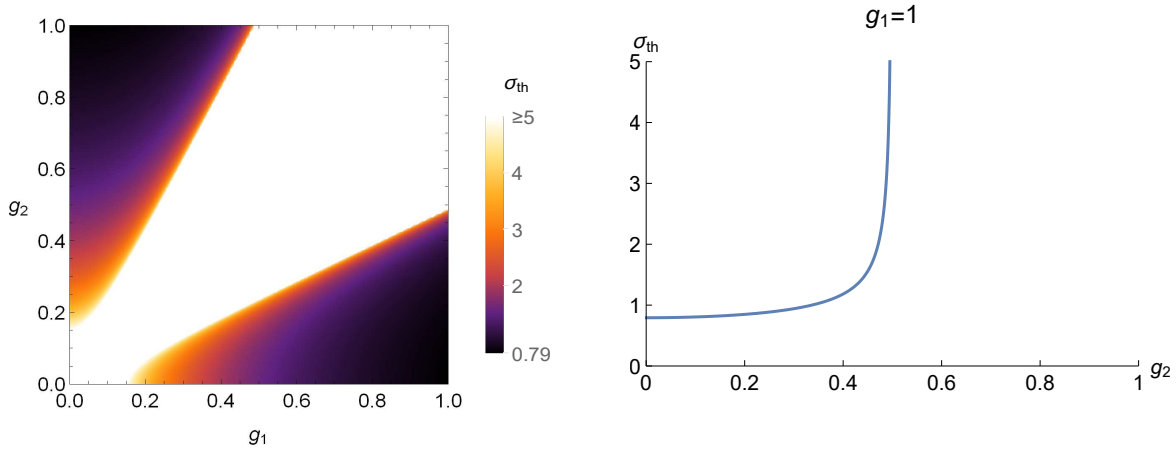



Figure 2-7: The pseudothresholds for the $n = 2$ efficient code. Without loss of generality, we consider $g_1, g_2 \geq 0$. Left: the pseudothreshold vs. the coupling strengths $(g_1, g_2)$. Right: a slice of the left panel with $g_1$ fixed to 1, for comparison with Fig. 2-8.

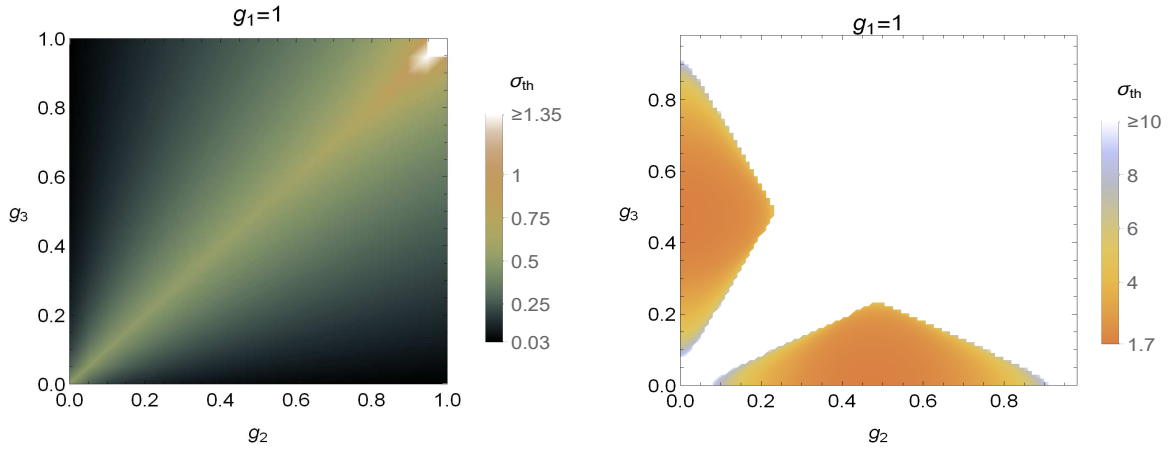



Figure 2-8: Left: the pseudothresholds for the $n = 3$ repetition code. Right: the pseudothresholds for the $n = 3$ efficient code. In both panels $g_{\max} = g_1$ is fixed to 1. Without loss of generality, we consider $g_1, g_2 \geq 0$.

As discussed Section 2.1, there are experimental settings in which the fluctuator's state affects not only the energy gaps of the register qubits, but also their quantization axes (e.g., due to nonzero $A_{zx}^{(j)}$ in Eq. (2.3)). This not only dephases the qubits, but more generally depolarizes them. Our hardware-efficient codes do not correct such decoherence at present; instead, the underlying interaction terms can be suppressed at the hardware level by increasing the qubits' free energy gaps, as we have discussed. Or, they could be corrected at the "software" level by concatenating these codes with a more conventional QEC code. Fig. 2-9 shows the impact of such terms on the

pseudothresholds of the $n = 2$ efficient code.

We model the effect of these depolarizing terms (here for 2 qubits) by the Hamiltonian

$$H_{\text{E}} = g_1^{\perp} X_1 + g_1^{\parallel} Z_1 + g_2^{\perp} X_2 + g_2^{\parallel} Z_2, \tag{2.34}$$

choosing the non-commuting ($\perp$) components to lie along $X_1$ and $X_2$, without loss of generality, as in Eq. (2.3). In the interaction picture, the total Hamiltonian (cf. Eq. (2.4)) takes the form

$$\tilde{H}_{\text{int}}(t) = \sum_{\ell} \lambda_\ell \, |\ell\rangle\langle\ell|_f \otimes \Big( g_1^{\perp} \big[ \cos(\omega_1 t) X_1 - \sin(\omega_1 t) Y_1 \big] + g_1^{\parallel} Z_1 + g_2^{\perp} \big[ \cos(\omega_2 t) X_2 - \sin(\omega_2 t) Y_2 \big] + g_2^{\parallel} Z_2 \Big). \tag{2.35}$$

In each run of the experiment, the register's dynamics is therefore generated by

$$\tilde{H}_{\text{eff}}(t) = \lambda(t) \Big( g_1^{\perp} \big[ \cos(\omega_1 t) X_1 - \sin(\omega_1 t) Y_1 \big] + g_1^{\parallel} Z_1 + g_2^{\perp} \big[ \cos(\omega_2 t) X_2 - \sin(\omega_2 t) Y_2 \big] + g_2^{\parallel} Z_2 \Big), \tag{2.36}$$

where the stochastic process $\lambda(t)$ is defined such that $\lambda(t) = \lambda_\ell$ when the fluctuator is in state $\ell$ at time $t$. When $g_j^{\perp} = 0$, this reduces to the case from the previous sections, with $g_j \leftrightarrow g_j^{\parallel}$ and $\theta = \int_0^{\delta t} \lambda(t) dt$, where $\delta t$ is the elapsed time between recoveries. In the more general setting here, however, the $U$ from Eq. (2.5) can instead be expressed as a Magnus series, which we truncate to leading order under the assumption that the couplings are weak compared to $\delta t^{-1}$:

$$U \approx \exp\Big[ -i \int_0^{\delta t} \tilde{H}_{\text{eff}}(t) \, dt \Big]. \tag{2.37}$$

For the sake of illustration, we take $\omega_1 = \omega_2 =: \omega$ and $g_1^{\perp}/g_1^{\parallel} = g_2^{\perp}/g_2^{\parallel}$. Moreover, as per the hardware-level approach described above, we will assume that $\omega$ is made large compared to the coupling strengths and $\delta t^{-1}$. We do not, however, make any assumptions about the relative strengths of $g_j^{\perp}$ and $g_j^{\parallel}$.

The integral in Eq. (2.37) contains two kinds of terms: (i) those considered in the

previous sections, of the form $\theta\,(g_1^{\parallel}Z_1 + g_2^{\parallel}Z_2)$, and (ii) rapidly oscillating terms of the form

$$g_j^{\perp}\int_0^{\delta t}\lambda(t)\big[\cos(\omega t)X_j - \sin(\omega t)Y_j\big]dt. \tag{2.38}$$

As with $\theta$, here we want to describe these integrals, $\xi_x = \int_0^{\delta t}\lambda(t)\cos(\omega t)dt$ and $\xi_y = \int_0^{\delta t}\lambda(t)\sin(\omega t)dt$, as random variables with a given distribution. Defining the Hamiltonian and propagator superoperators, $\mathcal{H}$ and $\mathcal{U}$, as done Section (2.3.1), we would find that the $X_j$ and $Y_j$ components (normalized by $\delta t$) give quadratures of the power spectral density of $\lambda$ at $\omega$, $S_\lambda(\omega)$. (We will review the formalism of stochastic processes in more detail in Chapter 4.) As they represent quadratures of the spectrum at a (shifted) frequency $\omega$, we can thus model $\xi_{x,y}$ as independent of $\theta$ and of each other, identically distributed, and as both following a normal distribution with zero mean and variance $S_\lambda(\omega)\delta t/2$ (so that the total power $S_\lambda(\omega)$ is split evenly between both quadratures). Concretely, we take

$$U = \exp\left\{-i\Big[\sum_{j=1}^{2} g_j^{\perp}\big(\xi_x X_j + \xi_y Y_j\big) + \theta\sum_{j=1}^{2} g_j^{\parallel}Z_j\Big]\right\}, \tag{2.39}$$

and average over $\theta \sim \mathcal{N}(0,\sigma)$ and $\xi_x, \xi_y \sim \mathcal{N}(0, \sqrt{\frac{S_\lambda(\omega)\delta t}{2}})$ (all independent). Notice that the non-commuting terms' importance depends not only on the relative strengths of $g_j^{\perp}$ and $g_j^{\parallel}$, but also on $S_\lambda(\omega)$, which is a property of the fluctuator. In fact, using the scaling property of normal distributions, Eq. (2.39) can be re-written as

$$U = \exp\left\{-i\sum_{j=1}^{2} g_j^{\parallel}\Big[\Xi_x X_j + \Xi_y Y_j + \theta Z_j\Big]\right\}, \tag{2.40}$$

where $\Xi_x, \Xi_y \sim \mathcal{N}\left(0, \sqrt{\frac{S_\lambda(\omega)\delta t}{2}}\frac{g_j^{\perp}}{g_j^{\parallel}}\right)$. Therefore, just as $\sigma$ sets the strength of the commuting noise, $\sqrt{S_\lambda(\omega)\delta t}\,g_j^{\perp}/g_j^{\parallel}$ (which is independent of $j$) sets the strength of the non-commuting (i.e., non-secular) noise[8]. Fig 2-9 shows the resulting pseudothresh-

[8]The factor of $1/\sqrt{2}$ drops out since there are two independent quadratures that contribute.

olds for different ratios

$$r = \frac{\sqrt{S_\lambda(\omega)\delta t}\ g_j^\perp / g_j^\parallel}{\sigma} \tag{2.41}$$

of these two quantities. The ratio $r$ describes the relative importance of the two noise types: $r \ll 1$ when commuting ($\parallel$) noise dominates, while $r \gg 1$ when non-commuting ($\perp$) noise dominates.

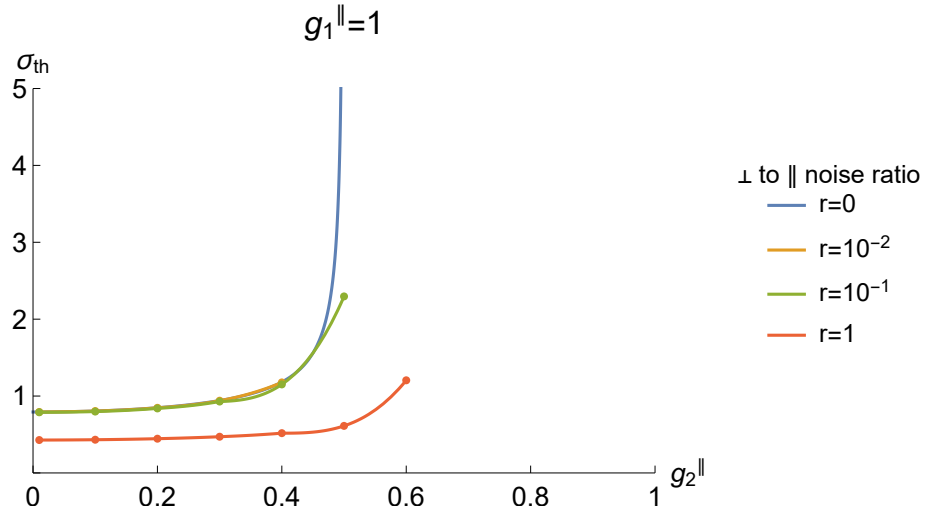



Figure 2-9: The analogous pseudothresholds to Fig. 2-7, but under non-commuting interaction terms of various strengths. The $r = 0$ curve is the same as in Fig. 2-7. The $r > 0$ curves were calculated using the average infidelity (as opposed to $\langle p \rangle$), as their logical channels no longer have a simple phase-flip form. These curves appear to diverge like the $r = 0$ curve, but at larger values of $g_2^\parallel$. Only when the non-commuting ($\perp$) terms become dominant (i.e., when $r > 1$, not shown) do the pseudothresholds start to behave erratically and become ill-defined. The pseudothresholds for $r > 0$ were computed at discrete intervals; the interpolating curves serve to guide the eye.

Pseudothresholds under CFD become difficult to visualize for $n \geq 4$. Moreover, since those of efficient codes diverge around DFS's, we cannot average $\sigma_{\text{th}}$ over $\{g_j\}$. Therefore, in order to analyze the sensitivity of efficient codes to calibration errors (i.e., uncertainty) in $\{g_j\}$ versus $n$, we will revert to plotting $\langle p \rangle$ versus $\sigma$, as in Fig. 2-1. $p$ is always bounded, so the average $\langle p \rangle$ over $\{g_j\}$ is always well-defined, in contrast with $\sigma_{\text{th}}$.

### 2.3.3 Sensitivity to Calibration Errors

We now analyze the effects of calibration error (i.e., uncertainty) in $\{g_j\}$. Since the encoding and recovery operations for our codes depend explicitly on $\{g_j\}$, error in measuring these coupling strengths will generally hurt overall performance. We model such errors through additive Gaussian noise on $g_j$: that is, if $g_j$ is the true coupling strength, we suppose that one instead estimates $g_j + \delta g_j$, where $\delta g_j \sim \mathcal{N}(0, \Delta g_j)$. For the sake of illustration, we assume furthermore that the errors $\delta g_j$ are independent across qubits, and that the measurement precision $\Delta g_j$ is the same for all $j$. We then estimate $\langle p \rangle$ through Monte Carlo averaging both over $\vec{g} \in [0,1]^n$ and $\delta g_j$.

Our goal is to analyze the sensitivity of efficient codes to measurement precision $\Delta g_j$ as a function of $n$. To this end, we begin by computing analogues of Fig. 2-1 (which assumes $\Delta g_j = 0$) with increasing calibration error. The results are shown in Fig. 2-10. Notice that at $\Delta g_j = 0.01$ our efficient codes all perform only slightly worse than in Fig. 2-1. As $\Delta g_j$ increases further, however, we begin to see saturation: that is, there is an $n_{\max}$ (decreasing with $\Delta g_j$) above which there is no further improvement. Finally, when one has almost no knowledge of the coupling strengths at $\Delta g_j = 0.5$, the efficient codes are no longer effective for any $n$. This behavior is expected: our codes achieve a high level of protection by exploiting knowledge of the noise. In the limit where we lose this knowledge, we necessarily also lose the protection. Note that the measurement precision reported in Ref. [31] for nuclear spins near an NV center corresponds to $\Delta g_j \sim 0.01$.

Fig. 2-11 presents similar information in a different way. It shows explicitly the behavior of $\langle p \rangle$ vs $n$ at different representative noise strengths and levels of calibration error. In each panel, we fix $n$ and examine the dependence of $\langle p \rangle$ on $\Delta g_j$. In the low- and intermediate-noise regimes, this dependence is well-described by $\langle p \rangle = A_n \, \Delta g_j^2 + B_n$, whereas at high noise $\langle p \rangle = A_n \, \Delta g_j + B_n$ gives a better fit. The coefficients $A_n$ (which correspond to $d\langle p \rangle / d(\Delta g_j^2)$ and $d\langle p \rangle / d(\Delta g_j)$ respectively, and were found using smaller steps of $\Delta g_j$ than shown in Figs. 2-10 and 2-11) are plotted versus $n$ in Fig. 2-12. Notice that the sensitivity to calibration uncertainty always increases

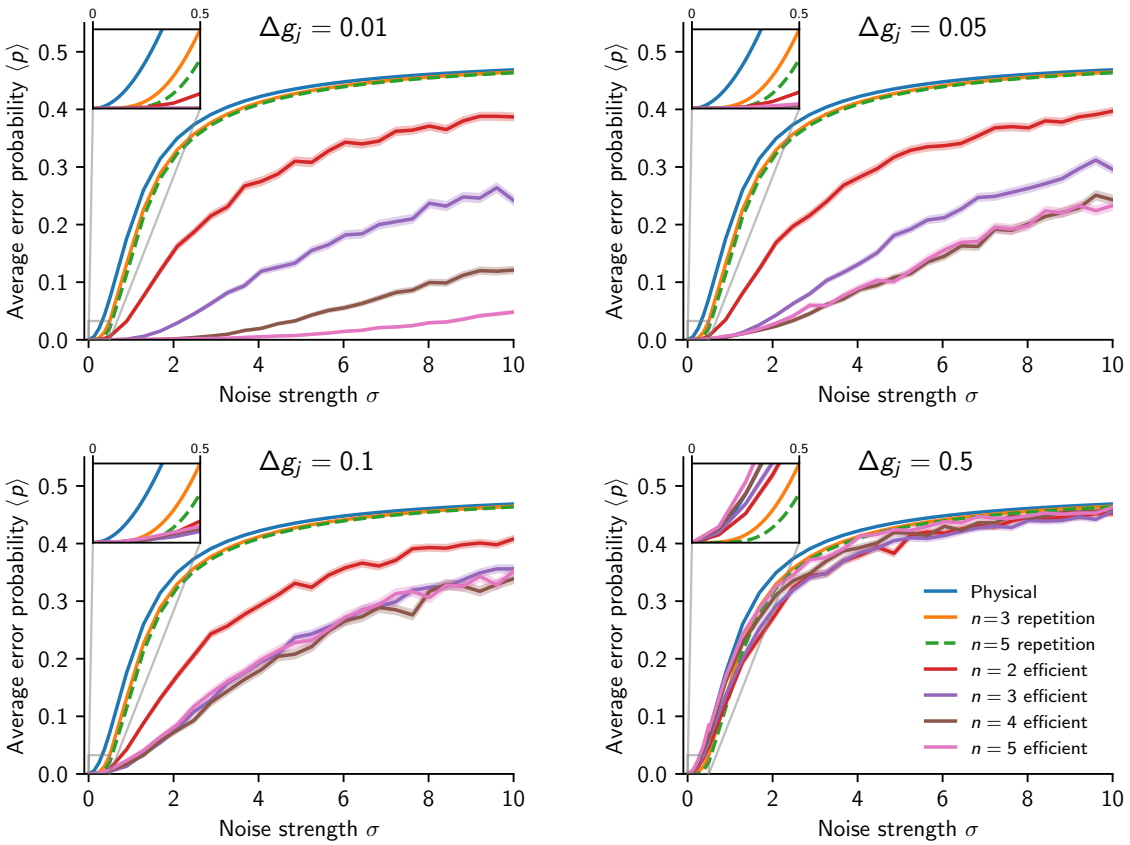



Figure 2-10: The quantities analogous to those in Fig. 2-1, but where with increasing measurement uncertainty $\Delta g_j$ in the coupling strengths between register qubits and the common resonator. Error bands denote standard error of the mean from Monte Carlo integration.

with $n$. This is expected, since adding more qubits with $\Delta g_j > 0$ introduces more uncertainty to the system. The asymptotic behavior of the sensitivity versus $n$ is less clear. In the high-noise regime, for instance, it is consistent with an exponential increase. In the low- and intermediate-noise regimes, however, the scaling appears sub-exponential (and perhaps even sub-linear in the latter regime).

The upper limit of $n = 5$ in these plots is due to numerical instabilities during Monte Carlo averaging. When the coupling strengths $\{g_j\}$ admit a DFS, the matrix $M$ pertaining to Eq. (2.27) becomes trivial, as does the recovery $\mathcal{R}$. Physically, this is a very fortunate situation, as it gives a good quantum memory without the need for active error correction. The same is true when $\{g_j\}$ give only an approximate DFS. However, while physically convenient, this latter case is problematic for the

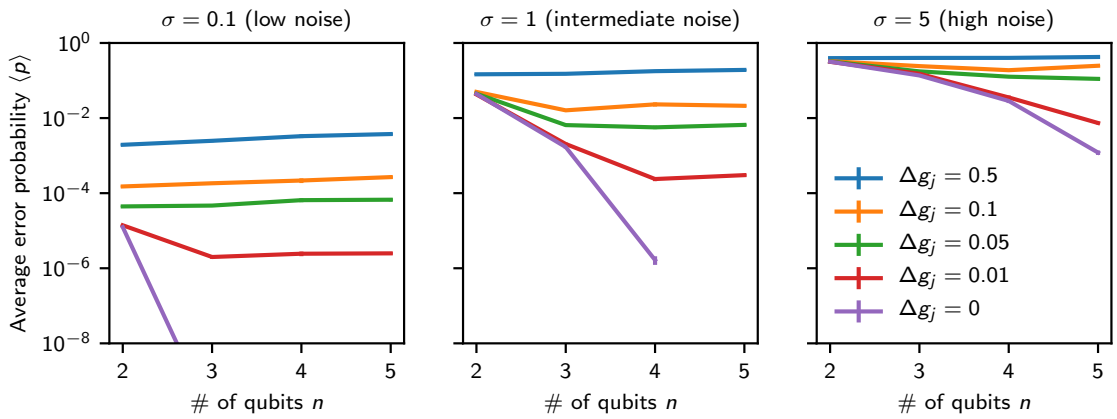



Figure 2-11: Average performance of efficient QEC codes by $\Delta g_j$ in three illustrative noise strength regimes. Error bars denote standard error of the mean from Monte Carlo integration, and lines are to guide the eye.

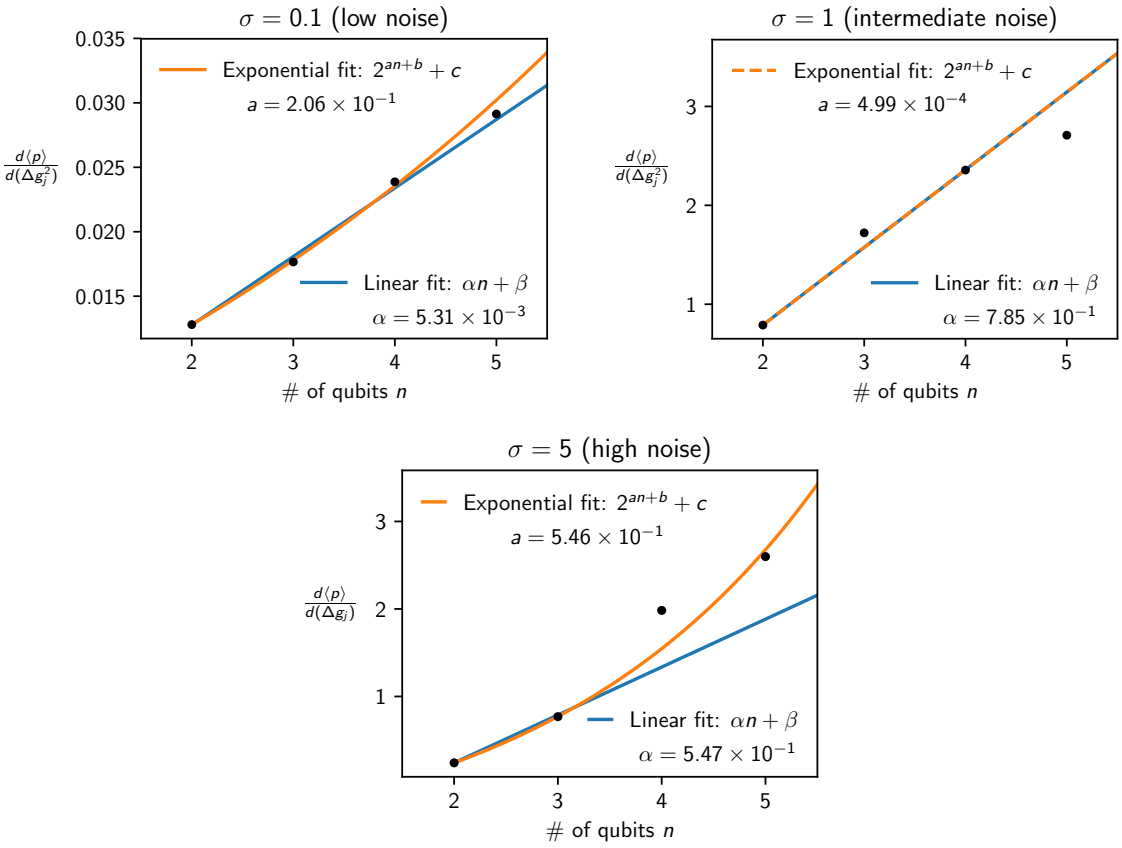



Figure 2-12: Sensitivity of efficient codes to calibration errors versus $n$ in three illustrative noise strength regime. The blue and yellow curves are the best linear and exponential fits, respectively.

Monte-Carlo averaging we perform here, as it gives an ill-conditioned code matrix $M$. This, in turn, can produce large numerical errors in $p$ when generated and inverted automatically as a subroutine of Monte Carlo integration. Such approximate DFS's become more frequent as $n$ increases (since it becomes more likely that two elements of a random $\vec{g} \in [0,1]^n$ be nearly equal), thus making it difficult to compute *average* $p$'s for $n \gtrsim 6$.

## 2.4 Discussion

The noise-adapted QEC codes we introduced involve a trade-off: they correct CFD very efficiently at the cost of leaving most other errors uncorrected. For instance, errors during gates, due to miscalibration of $g_j$'s, or from decoherence beyond CFD will generally affect the logical state. Accordingly, these codes are manifestly not fault-tolerant in their current form [58]. Crucially though, they offer such a large error budget under strong CFD—as evidenced by the gaps between QEC codes and physical qubits in Fig. 2-1—that this trade-off can easily be worthwhile, much like the targeted correction of photon loss in Ref. [29]. Indeed, as we showed in Section 2.3, the gap survives even in the presence of large miscalibration of the $g_j$'s. Fault-tolerance could still be achieved using implementation-specific methods as in Ref. [59]. In the long-term, concatenation could potentially reach fault-tolerance, using our noise-adapted codes at the lowest level of encoding to protect against the dominant error source, and more conventional codes at higher levels. This is precisely the strategy described in the quote from Nielsen and Chuang [7] at the start of this chapter. Even more importantly, our codes could have a near-term impact in applications such as quantum sensing and communication, where long-lived quantum memories are useful even when they are not fault-tolerant. We emphasize, however, that these codes are designed expressly for small- and medium-scale qubit registers, and that the exponential reduction in overhead should be understood to apply only in such devices. For one, there is typically a maximum $n$ above which CFD no longer dominates. Also, while the error budget always increases with $n$ in principle, so too do the

effects of gate errors, miscalibration of $g_j$'s and decoherence beyond CFD, as more qubits introduce more error channels. Conversely, this growing sensitivity suggests an unconventional quantum sensing scheme to measure $\{g_j\}$ for large $n$, by variationally adjusting one's estimates to maximize code performance. In the nearer term, however, these imperfections will likely set a maximum $n$ in any particular device beyond which one achieves no further gains, depending on their relative importance compared to CFD, as observed in Section 2.3.

The QEC codes presented in this chapter could be generalized in several ways. First, they can readily be made to correct dephasing due to multiple common fluctuators given enough qubits, at the cost of correcting to lower order in $\delta t$. Similarly, they can correct spatially-correlated phase noise beyond that arising from common fluctuators. As we will see in Part II for instance, classical white noise in the energy gaps of register qubits leads to Lindblad jump operators $L_j = \sqrt{\lambda_j}\,\vec{c}_j \cdot (Z_1, \ldots, Z_n)$, where $\{\sqrt{\lambda_j}\,\vec{c}_j\}$ describes the noise's normal modes[9] [60]. In the limit of spatially uncorrelated noise the $L_j$'s become Pauli $Z$ operators; however, correlated noise produces $L_j$'s with unequal amplitudes $\sqrt{\lambda_j}$. When the noise correlations are appreciable, it could be advantageous to use a QEC code that corrects the stronger noise modes (those with large $\lambda_j$'s) to higher order in $\delta t$ than the weaker ones (smaller $\lambda_j$'s) through an appropriate choice of $V$.

It may also be possible to extend the codes presented here for the setting where a fluctuator's state affects not only the energy gap of each qubit, but also the direction of its Hamiltonian (i.e., its quantization axis), as in Eq. (2.3) [61]. As discussed previously, eigenstate-preserving coupling arises frequently in practice because a large detuning between a weakly-coupled qubit and fluctuator suppresses non-commuting parts of their interaction Hamiltonian. However, when the coupling to the fluctuator is comparable to the internal Hamiltonian, such as for some nuclear spins near defects in diamond, there can remain significant non-commuting terms leading to $H_E \sim$

[9] One could approximate the CFD considered in this chapter by instead defining a single non-vanishing jump operator $L \propto H_E$ acting directly on the register. This would capture the spatial correlations inherent in CFD, though not the temporal correlations in general. (Specifically, it would approximate random telegraph noise as white noise.)

$\sum_j \vec{g}_j \cdot \vec{\sigma}_j$ in Eq. (2.4). We have analyzed this effect's impact on code performance in Section 2.3. Extending the codes introduced here to this more general setting would make them even more widely applicable to near-term experiments, but at the cost of larger overheads, since they would need to contend with a substantially larger space of possible errors. It may be more practical instead to suppress non-commuting interaction terms at the hardware level by increasing the energy gaps $\omega_j$ of the register qubits, or at the "software" level through concatenation.

Another interesting generalization would be to efficiently encode $k > 1$ logical qubits, which seems plausible based on the counting argument used throughout involving the dimension of $\mathscr{H}$ versus $\mathscr{E}$. (Note, however, that this counting argument should be taken only as a guide to where new QEC codes might exist, rather than as a guarantee that they will exist. For instance, this argument suggests the possibility of a 3-qubit amplitude damping code, cf. the 4-qubit version discussed at the beginning of this chapter. However, numerical evidence suggests that no such code exists [62].) Finally, it would be interesting to use the tools presented here to design codes for other common error sources, such as other types of decoherence or control/measurement errors.

Our results demonstrate that it is possible to find noise-adapted QEC codes with a well-defined advantage (here exponential) over known, general codes. It is commonly argued that QEC will be of little use in Noisy Intermediate-Scale Quantum (NISQ) devices due to its prohibitive overhead [2]. Device-adapted QEC codes are a promising way to reduce this overhead, although to date they have mostly relied on numerical and variational techniques discussed in more detail in Chapter 3. These codes can therefore lack transparency in terms of what advantage they offer, and when [63–67] (see also [19] Ch. 13 and [68]). In contrast, the codes introduced here exhibit a clear reduction in overhead under a well-characterized and common type of noise. New QEC codes of this type could provide a middle ground between small-scale uncorrected devices and large-scale fault-tolerant ones, where the dominant decoherence mechanisms are tamed through specialized codes with only modest overheads. This view of near-term QEC as quantum "firmware" rather than "software" suggests

a possible interplay between theory and experiment, whereby NISQ hardware and efficient QEC codes both guide each other's development.

# Chapter 3

# Robustness-Optimized QEC

As we saw in the last chapter, it can be advantageous in the near-term to trade the wide net of conventional QEC for a more tailored approach, which exploits knowledge of the dominant error mechanisms to achieve better error suppression [24, 69–72]. Chapter 2 and Part II focus on adapting QEC protocols to the intrinsic decoherence in idling quantum devices. Of course, errors can also arise in other ways, due to faulty controls or measurements, for instance, whose structure could also be exploited. In this chapter, we will therefore take preliminary steps in developing tailored QEC protocols for these latter types of errors.

In Chapter 2, we constructed tailored QEC codes analytically. A more common approach, called optimization-based QEC, aims to do so numerically [25, 63, 64, 66–68, 73–81] (see also Ref. [19], Chapter 13, for a review). It works by mapping the search for good QEC protocols (i.e., codes and recoveries) to an optimization problem, whose solution gives a protocol tailored for a particular type of noise. There are several ways to perform this mapping, some of which enable efficient optimization, as well as a degree of robustness to uncertainties in the error model [19, 65, 82]. While the resulting protocols often lack an intuitive structure, they hold promise for near-term devices, and perhaps as a first level of encoding in larger devices [74].

To date, optimization-based QEC has been largely synonymous with channel-adapted QEC; that is, the focus has been on adapting QEC protocols to the quantum channels describing intrinsic decoherence in idling devices. A notable feature

in some recent, pre-fault-tolerant experiments, however, is that errors due to imperfect QEC recoveries (i.e., measurement and feedback operations) comprise a significant share of the logical errors [29, 83]. In other words, there is ample room to improve QEC performance in near-term experiments by minimizing the impact of such recovery errors, in the spirit of Refs. [84, 85]. This suggests a new type of optimization-based QEC, orthogonal to channel-adapted QEC: rather than tailoring QEC protocols to the intrinsic decoherence between recoveries, one could instead find protocols which are optimally robust against imperfections in the recoveries themselves. This is a fundamentally different task; instead of finding an optimal way to suppress errors inherent to a device, it involves devising protocols that perform optimally under imperfect implementation. We demonstrate this latter approach, which we call *robustness-optimized QEC*, by maximizing the robustness of the $n = 3$ repetition code to syndrome measurement errors in the associated recovery.

## 3.1 Setting

We consider, for illustration, the task of preserving a logical qubit using three physical qubits subject to phase noise, which is the dominant kind of decoherence in many types of quantum devices [51, 86–90]. For simplicity, we will not let the QEC code itself vary in the optimization here; rather, we will use the phase-flip code, with codewords

$$|0_{\text{L}}\rangle = |{+}{+}{+}\rangle \qquad |1_{\text{L}}\rangle = |{-}{-}{-}\rangle\,, \tag{3.1}$$

where $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$ [91–94] (see also Ref. [19] Ch. 21 and references therein). That is, we will not tailor the code itself, but rather the associated control operations. For the sake of generality, then, we will assume decoherence with no particular spatial structure akin to that in the previous chapter, and focus simply on the impact of $\sigma_z$ errors on the qubits. These can be detected non-destructively by measuring $\{P, P_1, P_2, P_3\}$, where $P = |0_{\text{L}}\rangle\langle 0_{\text{L}}| + |1_{\text{L}}\rangle\langle 1_{\text{L}}|$ and $P_j = Z_j P Z_j$ are rank-2 orthogonal projectors. ($Z_j$ denotes the Pauli matrix $\sigma_z$ on qubit $j$.) A $Z_j$ error will transform the logical state $|\psi_{\text{L}}\rangle = \alpha\,|0_{\text{L}}\rangle + \beta\,|1_{\text{L}}\rangle$ into range$(P_j)$ in a way that can be reversed

by applying $Z_j$. The quantum channel describing this ideal recovery procedure is

$$\mathcal{R}_{\text{ideal}}(\rho) = \sum_{j=0}^{3} U_j^\dagger P_j \rho P_j U_j, \tag{3.2}$$

where $P_0 = P$, $U_0 = I$, $U_j = Z_j$ for $j \geq 1$ [7], as introduced in Chapter 1. Note that throughout this thesis we consider the conceptually-simple QEC strategy in which errors are physically corrected upon detection, as opposed to more sophisticated strategies using Pauli/Clifford frames [95, 96]. Moreover, we will assume that syndrome measurements are not repeated multiple times as part of the same recovery, e.g., due to the time this would take.

Suppose, however, that the measurement process is imperfect, and reports the wrong result uniformly with some probability $p_{\text{meas}}$, e.g., due to an error on an uncorrected ancilla. That is, a general state may be projected into range($P_j$) in the usual way, but the measurement device sometimes reports it to be in range($P_k$) for $k \neq j$. Feeding back on this faulty syndrome would cause a logical error. The channel describing this imperfect recovery is[1]:

$$\mathcal{R}_{\text{faulty}}(\rho) = (1 - p_{\text{meas}})\, \mathcal{R}_{\text{ideal}}(\rho) + \frac{p_{\text{meas}}}{3} \sum_{\substack{i,j=0 \\ i \neq j}}^{3} U_j^\dagger P_i \rho P_i U_j. \tag{3.3}$$

Note that $p_{\text{meas}}$ is the total measurement error probability, which may encompass the individual error probabilities from measurements on several ancilla qubits.

How can the phase-flip code be made more robust to such imperfections in the recovery? One can imagining two extreme strategies which work well in different regimes:

**Strategy A - Conventional QEC** If $p_{\text{meas}}$ is sufficiently small, a good strategy is to periodically perform $\mathcal{R}_{\text{faulty}}$, and simply accept the performance degradation due to non-zero $p_{\text{meas}}$.

[1] Note that a syndrome measurement error is not equivalent to a $Z_j$ error on a data qubit, since it has no effect in the absence of feedback.

**Strategy B - Quantum Zeno Effect** If $p_{\text{meas}}$ is sufficiently large, it may be better not to actively correct phase errors at all. Instead, one could suppress them—independent of $p_{\text{meas}}$—through the quantum Zeno effect by repeatedly measuring $\{P_j\}$ without feedback [97–100].

Which of these represents the better approach will depend both on $p_{\text{meas}}$ and on the total amount of time, $\Delta t$, for which one wants to preserve the logical state.

More generally, however, one could interpolate between Strategies A and B as follows: with probability $p_{\text{fb}}$ perform $\mathcal{R}_{\text{faulty}}$, and with probability $1-p_{\text{fb}}$ measure the parity $\{P_j\}$ but do not feed back. This corresponds to the channel

$$\mathcal{R}_{\text{opt}}(\rho) = p_{\text{fb}}\,\mathcal{R}_{\text{faulty}}(\rho) + (1-p_{\text{fb}})\sum_{j=0}^{3} P_j \rho P_j. \tag{3.4}$$

Strategies A and B then correspond to $p_{\text{fb}} = 1$ and 0 respectively. Instead of adopting either strategy entirely, we will treat $p_{\text{fb}}$ as a free parameter, and find the optimal value which maximizes robustness to recovery imperfections. For certain values of $p_{\text{meas}}$ and $\Delta t$, we find that intermediate values of $p_{\text{fb}}$ outperform both extreme strategies.

## 3.2 Decoherence Model and Objective Function

A common and simple model for the phase noise is a Lindblad equation with $Z_j$ jumps, as in Example 3.1 of Chapter 1. This would be equivalent to the qubits' energy gaps being subject to a zero-mean Gaussian white noise process (see Chapter 4), and would suppress single-qubit coherence as $\left|\langle 0|\,\rho_j\,|1\rangle\right| \propto \exp(-t/T_2^*)$ for some characteristic dephasing time $T_2^*$ [60, 101]. While this is a common idealization of realistic decoherence, it is unsuitable here. The quantum Zeno effect—which has been observed in several experiments, including some which preserve subspaces of dimension $\geq 1$, see e.g., [102–106]—does not occur in the pathological limit where the phase noise has infinite power at high frequencies. This is precisely the limit described by the aforementioned Lindblad model, and so repeated measurements of $\{P_j\}$, no matter how frequent, would not preserve a logical state in this model.

Adopting such a model would make it largely pointless to optimize $p_{\text{fb}}$.

A more realistic model for some experiments, which displays a Zeno effect and in turn a rich landscape in $p_{\text{fb}}$, is dephasing due to low-frequency noise in the qubits' energy gaps. Such noise suppresses single-qubit coherence as $\exp[-(t/T_2^*)^2]$, which is more typical in many experiments with slowly-evolving environments[2] [107, 108]. Concretely, we assume that in a suitable frame the qubits evolve as

$$H(t) = \frac{1}{2}\sum_{j=1}^{3} \omega_j(t) Z_j, \tag{3.5}$$

where the $\omega_j$'s are independent quasi-static noise processes that are approximately constant over $[0, \Delta t]$ but vary between runs of the experiment. More precisely, we take $\omega_j$ to be a zero-mean, stationary Gaussian stochastic process with a constant autocorrelation function

$$\langle \omega_j(t)\, \omega_j(0) \rangle = \frac{2}{(T_2^*)^2}, \tag{3.6}$$

where $\langle \cdot \rangle$ denotes a (classical) average over realizations of $\omega_j$. That is, the power spectrum of $\omega_j$ goes as $S_{\omega_j}(\nu) \propto \delta(\nu)$. While the dynamics in each run of the experiment is unitary, the average dynamics is not, which leads to dephasing. Note that dynamical decoupling could refocus this noise, although we will not consider it here in order to isolate the effects of QEC [49, 50, 109]. In practice, however, it could be beneficial to use dynamical decoupling in conjunction with the present QEC scheme.

We suppose that one can perform $\mathcal{R}_{\text{opt}}$ $N \geq 1$ times, equally spaced, during the interval $[0, \Delta t]$ (with the first $\mathcal{R}_{\text{opt}}$ occurring at time $\delta t = \Delta t/N$ and the last at $\Delta t$). To describe the effect of this procedure, we first define the superoperator $\mathcal{V}_t(\rho) := V_t\, \rho\, V_t^\dagger$, where

$$V_t := \exp\left[-i \int_0^t H(t')\, dt'\right]. \tag{3.7}$$

Then, if the system is prepared in the initial logical state $\rho_{\text{L}} = |\psi_{\text{L}}\rangle\langle\psi_{\text{L}}|$, its final state

[2] This is the same Ramsey decay as produced by $1/f$-type noise, where most of the noise power is at low frequencies. See Chapter 4 for an introduction.

after performing $N$ repetitions of $\mathcal{R}_{\text{opt}}$ in the interval $[0, \Delta t]$ is

$$\rho_f = \left\langle \left( \mathcal{R}_{\text{opt}} \, \mathcal{V}_{\Delta t/N} \right)^N \right\rangle (\rho_{\text{L}}). \tag{3.8}$$

We will use the quantum fidelity $F = \langle \psi_{\text{L}} | \, \rho_f \, | \psi_{\text{L}} \rangle$ as a measure of performance. More precisely, we use the fidelity averaged over all initial logical states, $\overline{F}$, as a figure of merit/objective function when optimizing the robustness. For $N = 1$ recovery (at a final time $\Delta t$), we have

$$\begin{aligned} \overline{F}_1 = {} & \frac{1}{6}\big[1 + p_{\text{fb}}(3 - 4p_{\text{meas}})\big] + \frac{1}{2} e^{-2(\Delta t/T_2^*)^2}(1 - p_{\text{fb}}) \\ & + \frac{1}{4} e^{-(\Delta t/T_2^*)^2}\big[1 + p_{\text{fb}}(1 - 2p_{\text{meas}})\big] + \frac{1}{12} e^{-3(\Delta t/T_2^*)^2}\big[1 + p_{\text{fb}}(2p_{\text{meas}} - 3)\big]. \end{aligned} \tag{3.9}$$

We were able to find analytic expressions for $\overline{F}_N$ with $1 \leq N \leq 10$, although for $N \geq 2$ the expressions quickly become lengthy. (The expression for $\overline{F}_{10}$, for instance, contains 4588 terms. It, along with the other $\overline{F}_N$'s, can be found in the Mathematica notebook included in the supplementary material of Ref. [110].). Average fidelities for $N \geq 11$ are not only difficult to compute, but they are of limited relevance to near-term experiments where control limitations and other sources of error impose a limit on $N$. Moreover, even in the longer term, the number of recoveries within an interval $[0, \Delta t]$ must be limited if there is to be time left over to perform logical operations on the encoded state (since recoveries will not be instantaneous in practice).

## 3.3 Results

We will treat $\Delta t$ and $p_{\text{meas}}$ as fixed in any given experiment, which leaves the parameters $N$ and $p_{\text{fb}}$ to be optimized. The dependence of $\overline{F}_N$ on these parameters, for a particular $\Delta t$ and $p_{\text{meas}}$, is illustrated in Fig. 3-1. For this $\Delta t$ and $p_{\text{meas}}$, the most robust strategy is a hybrid of Strategies A and B, which outperforms the two extremes. Perhaps counter-intuitively, this means that the average fidelity is *increased* here by introducing extra randomness into $\mathcal{R}_{\text{opt}}$ through the choice of $0 < p_{\text{fb}} < 1$.

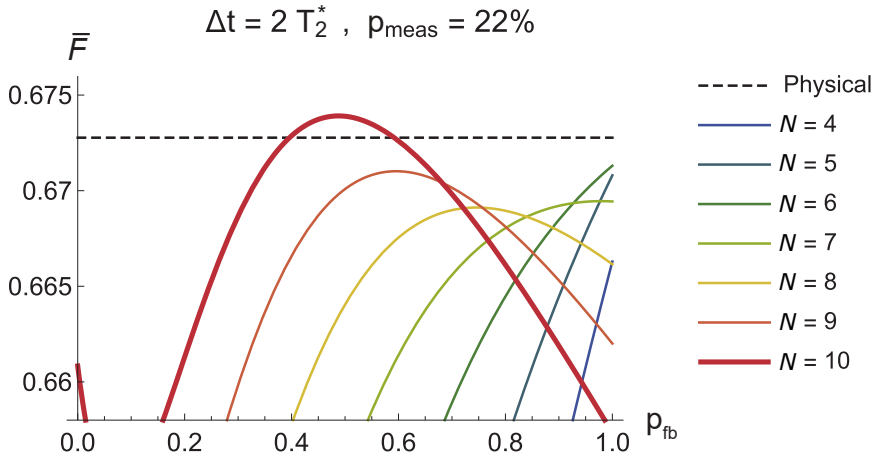



Figure 3-1: The average fidelity versus $p_{\text{fb}}$ and $N$ for $\Delta t = 2T_2^*$ and $p_{\text{meas}} = 0.22$. The solid lines denote $\overline{F}_N$ for $N \geq 4$; the curves for $N \leq 3$ are not visible as they are too low. The dashed line is the fidelity of single physical qubit under the same noise. The optimal strategy of those considered, that is, the $N \in [1, 10]$ and $p_{\text{fb}} \in [0, 1]$ combination producing the highest fidelity, uses $N = 10$ (bold red line) and $p_{\text{fb}} = 0.488$ to achieve a fidelity of $\overline{F}_{\text{max}} = 0.674$.

More generally, for each $(\Delta t, p_{\text{meas}})$, we optimize $\overline{F}_N$ over both $N$ and $p_{\text{fb}}$. The optimal $p_{\text{fb}}$, shown in Fig. 3-2, has three distinct "phases" in the parameter range considered. As anticipated above, when $p_{\text{meas}}$ is sufficiently small the optimal strategy is to perform conventional recoveries ($p_{\text{fb}} = 1$) and simply accept the occasional faults that these introduce. Conversely, when $p_{\text{meas}}$ is sufficiently large (and/or $\Delta t$ is sufficiently small), it is better to avoid feedback entirely and simply preserve the logical state using a Zeno effect from repeated parity measurements. We observe a sharp transition between these two optimal strategies in much of the parameter space. Mathematically, this is due to the maxima of $\overline{F}_N$ often occurring on the boundary of $\{p_{\text{fb}} \in [0, 1]\}$ rather than in the interior. Remarkably, however, there is a finite region where the transition is not sharp, which exhibits a third "phase" corresponding to optimal $p_{\text{fb}}$'s near 0.5 (though not always exactly equal to 0.5, see e.g., Fig. 3-1). The $\Delta t$ and $p_{\text{meas}}$ from Fig. 3-1 are from this region.

The maximum values of $\overline{F}_N$ and the optimal $N$'s resulting from this same optimization are shown in the left and center panels of Fig. 3-3. As one might expect, the fidelity decays gradually with increasing $\Delta t$ and $p_{\text{meas}}$. The choice of $N$ is more complex, as the same optimal $N$ can represent different strategies depending on the

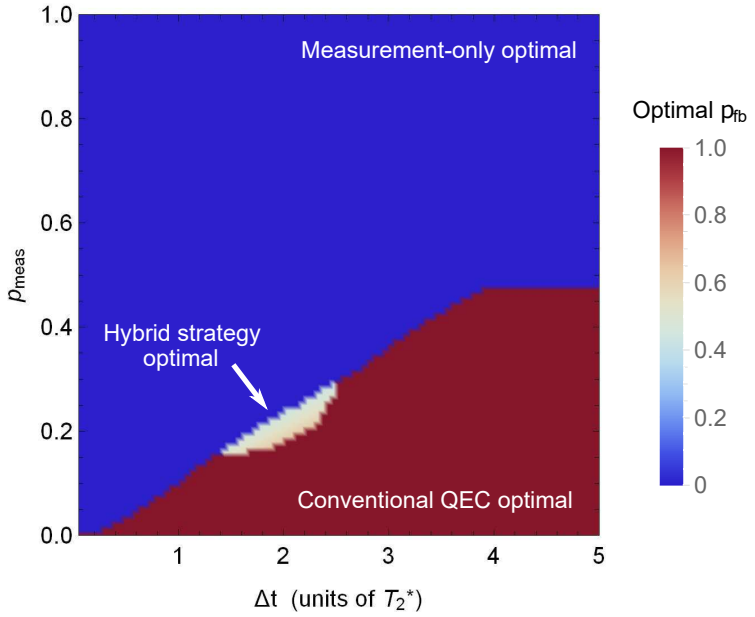



Figure 3-2: The optimal $p_{\text{fb}}$ for different values of $\Delta t$ and $p_{\text{meas}}$, after encoding. The best $p_{\text{fb}}^{(N)}$ for each $\overline{F}_N$ was found separately; this figure shows the one giving the highest value of $\overline{F}_N$. $p_{\text{fb}} = 1$ gives the conventional QEC strategy of measurement and feedback, whereas $p_{\text{meas}} = 0$ uses no feedback, relying instead on a quantum Zeno effect from repeated parity measurements.

corresponding $p_{\text{fb}}$. For instance, using a large $N$ is optimal both when $p_{\text{meas}}$ is small and when it is large (compared to $\Delta t$). In the former regime one has $p_{\text{fb}} = 1$, so a large $N$ reduces the buildup of uncorrectable errors of weight 2 and 3 due to phase noise. In the latter regime $p_{\text{fb}} = 0$, so a large $N$ means frequent measurements and therefore a stronger Zeno effect. Between these two regimes, moderate values of $N$ are optimal, as they provide some correction without too many recovery faults. Finally, for large $\Delta t$ and large $p_{\text{meas}}$ we find small $N$ to be optimal. This is likely an artifact of considering only $N \leq 10$: $\lim_{N\to\infty} \overline{F}_N = 1$ for all $\Delta t$ and $p_{\text{meas}}$, so if we allowed unbounded $N$ the Zeno strategy would always be optimal in principle. However, for large $\Delta t$, $N \leq 10$ measurements are insufficient to produce a strong Zeno effect, so the next-best strategy is to use faulty recoveries sparingly. Note finally that for large $\Delta t$ and/or $p_{\text{meas}}$, including some values where a hybrid strategy is shown to be optimal in Fig. 3-2, it may be better not to perform encoding at all (see Appendix B).

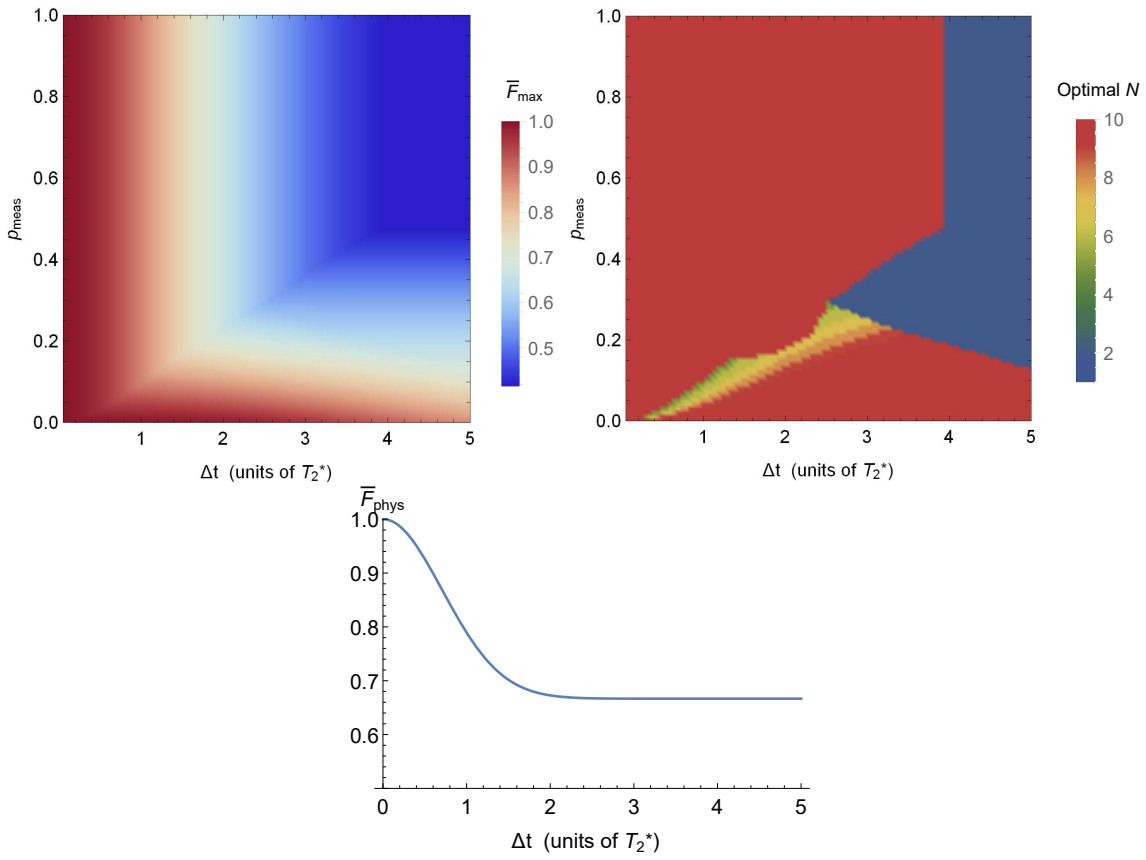



Figure 3-3: Top left: The maximum fidelity achievable by optimizing over $p_{\text{fb}} \in [0, 1]$ and $1 \leq N \leq 10$. Top right: The optimal $N$ which gives this maximum fidelity. Bottom: For comparison, the fidelity for a single physical qubit subject to the same noise.

## 3.4 Discussion

We have shown that one can optimize the robustness of small, pre-fault-tolerant QEC protocols to recovery errors, in analogy to how such protocols have previously been optimized for specific decoherence channels. Whereas the latter approach is often called channel-adapted QEC, we term ours robustness-optimized QEC. Errors from QEC recoveries have formed an appreciable fraction of the total logical errors in recent experiments [29, 83]. This suggests that there is much to be gained by optimizing for robustness against such errors in addition to optimizing for the decoherence inherent in particular devices. While fault-tolerant methods could handle such errors in the longer term, the present strategy is specifically intended for nearer-term, pre-fault-tolerant experiments [111, 112].

These results raise a number of further questions and possibilities, which we divide into technical points and points of strategy. First the technical points. As in previous works on optimization-based QEC, there is some ambiguity here in choosing a figure of merit. We have used average fidelity for convenience; however, the optimization could give slightly different results/strategies if we had chosen a different objective function, e.g., trace distance to the identity [113]. Moreover, there is often little reason to favor one particular performance measure over another *a priori* (see [19], Chapter 13). It would be useful to better understand how such effects affect schemes of the sort considered here. Similarly, the robust QEC strategies found here are robust against a particular type of error during recovery, which we chose as a generic illustration—they are not a panacea[3]. Different types of recovery errors will likely require different models and optimization mappings than the ones used here, which may need to be worked out case-by-case. Fortunately, there is less ambiguity with this choice, since the dominant error sources in current experiments are often well-understood (see, e.g., [29, 83]). There is likely more room for optimization in more detailed fault models, e.g., where the probability of measurement errors is outcome-dependent, or when such errors are predominantly due to decoherence of ancillas (rather than limited

[3]In particular, our fault model is different—and simpler—than the dominant recovery imperfections in [29, 83].

measurement fidelity, for instance) [91, 94]. Indeed, noise that is highly structured can often be dealt with more efficiently in general [24, 26, 27, 69, 70, 72, 109]. Finally, previous works on channel-adapted QEC have introduced sophisticated mappings which result in convex/bi-convex optimization problems that are efficiently solvable. Developing analogous tools for robustness-optimized QEC would enable the analysis of more complex codes and even more realistic noise models (such as $1/f$ noise) than those analyzed here (see [114] and references therein).

As for the points of strategy: First, rather than optimizing the probability of performing feedback, one could instead optimize over deterministic strategies of the form "feedback, no feedback, feedback, ...". This would most likely improve performance, but at the cost of transforming a continuous optimization problem into a potentially more expensive combinatorial one. Second, while we have only optimized the form of the recovery here, it may be advantageous to optimize both the code and the recovery, as is common in channel-adapted QEC [114]. Indeed, it would be interesting to combine the distinct approaches taken in this chapter and in the previous one. Moreover, one could think of changing the recovery's structure more generally, e.g., by using different $U_j$'s in Eqs. (3.2) and (3.3). (However, we have had limited success with this approach to date.) Finally, it may be possible to build upon the existing machinery of channel-adapted QEC by incorporating tools from robust or stochastic optimization, which can find near-optimal solutions to problems that are robust against imperfections in implementation [115] (see also [116] for an introduction). There appears to be ample room for new approaches to optimization-based QEC in light of recent experimental progress.

# Part II

# Application-Adapted Quantum Error Correction

*"In order to seek truth, it is necessary once in the course of our life to doubt, as far as possible, of all things."*

– René Descartes

# Chapter 4

# Introduction

In the second part of this thesis we will develop and analyze QEC codes that are adapted for quantum sensing applications. We begin here by introducing two topics that will be of central importance in Chapters 5 and 6: dephasing due to classical noise, and quantum sensing.

## 4.1 Dephasing From Classical Noise

A common explanation for why quantum systems decohere is that they become entangled with their environment [14]. Since we can't generally measure this environment, a gradual buildup of system-environment entanglement looks to us like gradual decoherence of the system. Strictly speaking, this picture should always be correct if one defines the environment broadly enough (assuming quantum mechanics gives an accurate description of nature at all relevant scales). Often, however, it is overkill. For many quantum devices, the environment (or at least part of it) is effectively classical, and its impact on the system can be effectively captured by environment-dependent scalar parameters in the system Hamiltonian. For instance, the energy gap of a spin-$\frac{1}{2}$ system depends on the external magnetic field, which can usually be modeled classically.

The catch is that, in practice, even a classical environment is typically noisy, meaning it exhibits random fluctuations which are manifest as random changes in

the system Hamiltonian. As we mentioned briefly in Chapter 1, this results in a system dynamics that is unitary in any run of an experiment, but which varies from one run to the next, leading to a non-unitary dynamics on average. Since we typically repeat experiments many times to build up measurement statistics, it is this average dynamics that we see. Therefore, while decoherence is in part a manifestly quantum phenomenon, it is also in part a consequence simply of living in a noisy (and largely classical) world. In fact, a broad family of decoherence processes can understood classically in this way, see, e.g., Refs. [117–120].

In this section we will introduce the basic tools for analyzing decoherence due to classical noise. Such physical details of decoherence may be of little relevance for conventional quantum error correction (QEC). As in Part I, however, we will find that these details open the door to powerful noise suppression schemes which could not have been discovered from coarser noise models.

In general, classical noise from the environment can impact quantum systems in a wide range of ways. We will mostly be concerned with the common case where classical fluctuations in each qubit's energy gap induce dephasing [101, 121]. Concretely, we consider for now a single qubit whose energy gap depends on the environment (say, the $z$ component of the local magnetic field):

$$H(t) = \frac{\omega_0 + \delta\omega(t)}{2} Z. \tag{4.1}$$

Here $\omega_0$ is the average energy gap, and $\delta\omega(t)$ encodes the fluctuations in the environment during a given run of the experiment. Without loss of generality, we will assume $\delta\omega(t)$ to have zero mean. Moreover, we will assume for simplicity that $\omega_0$ is constant, although the results below can easily be generalized for a time-dependent $\omega_0$.

Solving the Schrödinger equation for Eq. (4.1), we find that the qubit evolves as

$$\underbrace{\begin{pmatrix} a & b \\ b^* & c \end{pmatrix}}_{\rho(0)} \mapsto \underbrace{\begin{pmatrix} a & b\, e^{-i\omega_0 \Delta t} E(\Delta t) \\ b^*\, e^{i\omega_0 \Delta t} E(\Delta t)^* & c \end{pmatrix}}_{\rho_{\delta\omega}(\Delta t)} \tag{4.2}$$

in the computational basis, where

$$E(\Delta t) = \exp\left[-i\int_0^{\Delta t} \delta\omega(t)dt\right] \tag{4.3}$$

is the excess phase accumulated due to $\delta\omega$ over $[0, \Delta t]$. If we repeat this experiment many times to gather measurement statistics, resetting the qubit to the same initial state $\rho(0)$ each time, we will get a different function $\delta\omega$ with every run. This will produce a different excess phase $E(\Delta t)$ and therefore a different state $\rho_{\delta\omega}(\Delta t)$ each time. The measurement statistics, however, depend not just on a single $\rho_{\delta\omega}(\Delta t)$ from one repetition, but rather, on the average state

$$\langle\rho(\Delta t)\rangle = \begin{pmatrix} a & b\, e^{-i\omega_0\,\Delta t}\langle E(\Delta t)\rangle \\ b^*\, e^{i\omega_0\,\Delta t}\langle E(\Delta t)\rangle^* & c \end{pmatrix}, \tag{4.4}$$

where $\langle\,\cdot\,\rangle$ denotes an average over realizations (rather than a quantum-mechanical expectation value). To see why, note that the probability of getting a measurement outcome described by an operator $M_j$ is formally

$$\begin{aligned} p_j &= \int_{\text{all }\delta\omega} \Pr(\text{outcome } j \,|\, \delta\omega)\; p(\delta\omega)\, d(\delta\omega) \\ &= \int_{\text{all }\delta\omega} \text{tr}\left[M_j\, \rho_{\delta\omega}(\Delta t)\right]\; p(\delta\omega)\, d(\delta\omega) = \text{tr}\left[M_j\, \langle\rho(\Delta t)\rangle\right], \end{aligned} \tag{4.5}$$

where $p(\delta\omega)$ is the probability density of getting a particular $\delta\omega$ in an experiment, and where the meaning of this integral will be made more precise below. Since $\langle E(\Delta t)\rangle$ is the average of many complex numbers on the unit circle, we expect that it will tend towards zero, suppressing the qubit's phase coherence as $t$ grows.

### 4.1.1 Gaussian Noise

There is no exact closed-form expression for $\langle E(\Delta t)\rangle$ in general. One approach instead is to calculate it case-by-case for different noise processes. For instance, one can find an exact expression for $\langle E(\Delta t)\rangle$ under random telegraph noise [where $\delta\omega(t)$ jumps at random between a discrete set of values] of the kind encountered in Chapter 2

[47, 48, 122]. Fortunately, there is a simple expression for $\langle E(\Delta t)\rangle$ for a broad and common type of noise; namely, that where $\delta\omega$ is a Gaussian process [123, 124]. Such noise can arise, for instance, as the aggregate of many independent and identically distributed (IID) noise processes affecting a qubit, due to the central limit theorem.

We will focus on such Gaussian noise in Chapters 5 and 6, for which calculating $\langle E(\Delta t)\rangle$ will involve three steps:

1. Breaking the interval $[0, \Delta t]$ into $N$ small timesteps $\delta t$, and computing a discretized approximation to $\langle E(\Delta t)\rangle$.
2. Finding the exact $\langle E(\Delta t)\rangle$ by taking the continuum limit $\delta t \to 0$.
3. Expressing $\langle E(\Delta t)\rangle$ in terms of the noise power spectral density.

Finding the average excess phase involves averaging over all possible $\delta\omega$'s, weighted by the probability of each, as in Eq. (4.5). Formally:

$$\langle E(\Delta t)\rangle = \int_{\text{all } \delta\omega} \underbrace{\exp\left[-i\int_0^{\Delta t} \delta\omega(t)dt\right]}_{E(\Delta t) \text{ for a particular } \delta\omega} p(\delta\omega)\, d(\delta\omega). \tag{4.6}$$

This is very similar to a path integral, except that we need not impose any particular boundary conditions on $\delta\omega$. To make Eq. (4.6) well-defined, we will discretize time so as to integrate over vectors rather than functions, before ultimately taking the continuum limit. Concretely, we divide $[0, \Delta t]$ into $N$ intervals of length $\delta t = \Delta t/N$, and define points in time $t_k = k\,\delta t$. We also define a vector $\overrightarrow{\delta\omega}$ whose elements are the values of a realization $\delta\omega$ at these discrete points:

$$\overrightarrow{\delta\omega} = \Big[\delta\omega(t_1), \delta\omega(t_2), \ldots, \delta\omega(t_N)\Big]^\top. \tag{4.7}$$

In this discretized picture, each run of the experiment gives a vector $\overrightarrow{\delta\omega}$ with probability (density) $p(\overrightarrow{\delta\omega})$ rather than a function $\delta\omega(t)$ with probability (density) $p(\delta\omega)$. The key step is to assume that $\overrightarrow{\delta\omega}$ follows a multivariate Gaussian distribution, e.g., arising from the central limit theorem. This is what we mean by "Gaussian" noise.

Recall that a scalar $x$ follows a zero-mean Gaussian/normal distribution if the

probability density function is given by

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right), \tag{4.8}$$

where $\sigma^2$ is the variance. Similarly, a vector $\vec{x}$ is normally distributed (with zero mean) if it has a probability density

$$p(\vec{x}) = \frac{1}{\sqrt{\det(2\pi\Sigma)}} \exp\left(-\frac{1}{2}\,\vec{x}^{\top}\Sigma^{-1}\vec{x}\right), \tag{4.9}$$

where $\Sigma$ is the covariance matrix, defined as

$$\Sigma_{ij} = \langle x_i x_j \rangle, \tag{4.10}$$

which generalizes the variance $\sigma^2 = \langle x^2 \rangle$ in the 1-dimensional case.

We can approximate the excess phase in a given run through a Riemann sum:

$$E(\Delta t) = \exp\left[-i\sum_{j=1}^{N} \delta\omega(t_j)\,\delta t + O(1/N)\right] = \exp(-i\delta t\ \vec{1}\cdot\overrightarrow{\delta\omega}) + O(1/N), \tag{4.11}$$

where $\vec{1} := (1, \ldots, 1)^{\top}$. The discretized version of Eq. (4.6) is therefore

$$\langle E(\Delta t)\rangle \approx \int_{\mathbb{R}^N} \underbrace{e^{-i\delta t\,\overrightarrow{\delta\omega}\cdot\vec{1}}}_{\approx E(\Delta t)} \underbrace{\frac{1}{\sqrt{\det(2\pi\Sigma)}} e^{-\overrightarrow{\delta\omega}^{\top}\Sigma^{-1}\overrightarrow{\delta\omega}/2}}_{p(\overrightarrow{\delta\omega})}\, d(\overrightarrow{\delta\omega}), \tag{4.12}$$

where we get exact equality in the continuum limit. This integral can be easily evaluated through the substitution $\overrightarrow{\delta\omega} = \vec{u} - i\delta t\,\Sigma\vec{1}$, which gives

$$\langle E(\Delta t)\rangle \approx e^{-\delta t^2\,\vec{1}^{\top}\Sigma\vec{1}/2} \frac{1}{\sqrt{\det(2\pi\Sigma)}} \int_{\mathbb{R}^N + i\delta t\,\Sigma\vec{1}} e^{-\vec{u}^{\top}\Sigma^{-1}\vec{u}/2}\, d\vec{u}, \tag{4.13}$$

where we use the shorthand

$$\mathbb{R}^N + i\delta t\,\Sigma\vec{1} := \{\vec{x} + i\delta t\,\Sigma\vec{1} \,|\, \vec{x} \in \mathbb{R}^N\}. \tag{4.14}$$

One can show (e.g., by diagonalizing $\Sigma^{-1}$) that the remaining integral in Eq. (4.13) is independent of $\delta t$; we are therefore free to integrate over $\mathbb{R}^N$ (i.e., take $\delta t = 0$ in the domain of integration) rather than $\mathbb{R}^N + i\delta t\,\Sigma\vec{1}$. Comparing with Eq. (4.9), normalization dictates that this integral equal $\sqrt{\det(2\pi\Sigma)}$, so

$$\langle E(\Delta t)\rangle \approx e^{-\delta t^2\ \vec{1}^\top \Sigma \vec{1}/2}, \tag{4.15}$$

where we get equality in the $N \to \infty$ limit. Notice that

$$\delta t^2\ \vec{1}^\top \Sigma \vec{1} = \sum_{i,j=1}^{N} \Sigma_{ij}\,\delta t^2 = \sum_{i=1}^{N}\sum_{j=1}^{N} \langle \delta\omega(t_i)\,\delta\omega(t_j)\rangle\,\delta t^2 \to \int_0^{\Delta t}\int_0^{\Delta t} \langle \delta\omega(t_1)\delta\omega(t_2)\rangle\,dt_1 dt_2 \tag{4.16}$$

in this limit. Taking $N \to \infty$ (or equivalently $\delta t \to 0$) in Eq. (4.15), we arrive at the exact expression for the average excess phase due to Gaussian noise:

$$\langle E(\Delta t)\rangle = \exp\left[-\frac{1}{2}\int_0^{\Delta t}\int_0^{\Delta t} \langle \delta\omega(t_1)\,\delta\omega(t_2)\rangle\,dt_1 dt_2\right]. \tag{4.17}$$

$\langle \delta\omega(t_1)\,\delta\omega(t_2)\rangle$ is called the autocorrelation function of $\delta\omega$, and it quantifies how strongly correlated the fluctuations are at different times $t_1$ and $t_2$. One could interpret this function—cautiously[1]—as quantifying how much "memory" the environment has. We can simplify things a bit by assuming that the noise is stationary, meaning that it depends only on time differences, not on absolute time [125]. (This is a common, though sometimes questionable assumption.) The previous expression can then be rewritten as

$$\langle E(\Delta t)\rangle = \exp\left[-\frac{1}{2}\int_0^{\Delta t}\int_0^{\Delta t} A_{\delta\omega}(t_1 - t_2)\,dt_1 dt_2\right] = \exp\left[-\int_0^{\Delta t}\big(\Delta t - t\big)\,A_{\delta\omega}(t)dt\right], \tag{4.18}$$

where we use $A_{\delta\omega}(t) = \langle \delta\omega(t)\delta\omega(0)\rangle$ to denote the autocorrelation, which is even in $t$. Notice that $A_{\delta\omega}$ is real-valued, so $\langle E(\Delta t)\rangle$ does not contribute to the qubit's phase—it

[1] One must keep in mind, however, that in this model the environment never acquires any information about the system's state. Therefore, even an environment with an autocorrelation that decays slowly with $|t_1 - t_2|$ has no knowledge of prior system states.

just suppresses the coherences. This is not always true for non-Gaussian noise, such as random telegraph noise, where even zero-mean fluctuations can contribute a net phase [126, 127].

It is convenient to convert this expression to the frequency domain, and express the average excess phase in terms of the noise power spectral density [125], often simply called $\delta\omega$'s power spectrum, which we define as the Fourier transform of $A_{\delta\omega}$:

$$S_{\delta\omega}(\nu) = \frac{1}{2\pi}\int_{-\infty}^{\infty} A_{\delta\omega}(t)e^{-i\nu t}dt. \tag{4.19}$$

Eq. (4.18) can expressed in terms of $S_{\delta\omega}$ (by writing $A_{\delta\omega}$ as the inverse Fourier transform of $S_{\delta\omega}$) as

$$\langle E(\Delta t)\rangle = \exp\left[-\int_0^{\infty} S_{\delta\omega}(\nu)F_{\Delta t}(\nu)\,d\nu\right], \tag{4.20}$$

where

$$F_{\Delta t}(\nu) = \frac{4}{\nu^2}\sin^2\left(\frac{\nu\Delta t}{2}\right) \tag{4.21}$$

is called the filter function for the evolution. (We have made use here of the fact that the Fourier transform of a real even function is also real and even.) This frequency-domain expression is convenient for two main reasons: First, $S_{\delta\omega}$ has a nice physical interpretation. Second, $F_{\Delta t}$ can be modified by applying control pulses to the qubit in the interval $[0, \Delta t]$, so as to suppress dephasing. We will treat the first point here, and the second in Section 4.1.3.

Imagine decomposing $\delta\omega(t)$ for a given run of the experiment into a Fourier series

$$\delta\omega(t) = \sum_{m=-\infty}^{\infty} a_m\, e^{i\nu_m t}, \tag{4.22}$$

where the coefficients are given by

$$a_m = \frac{1}{\Delta t}\int_0^{\Delta t} \delta\omega(t)\, e^{-i\nu_m t}dt \tag{4.23}$$

for frequencies $\nu_m = 2\pi m/\Delta t$. Since $\delta\omega$ will be different from one run to the next, we should think of the Fourier coefficients as random variables. We can quantify how

important the coefficients around some frequency $\nu$ are, on average, by the mean squared magnitude of the Fourier coefficients in a frequency band $[\nu, \nu + \Delta\nu]$:

$$\frac{1}{\Delta\nu} \sum_{m \text{ in band}} \langle |a_m|^2 \rangle, \tag{4.24}$$

In the limit of large $\Delta t$ (so the spacing between Fourier modes vanishes) and $\Delta\nu \to 0$, Eq. (4.24) is equal to $S_{\delta\omega}(\nu)$ by the Wiener–Khinchin theorem [128, 129]. In brief:

$$\begin{aligned} \langle a_m^* a_m \rangle &= \frac{1}{\Delta t^2} \int_0^{\Delta t} \int_0^{\Delta t} \langle \delta\omega(t_1) \delta\omega(t_2) \rangle \, e^{-i\nu_m (t_1 - t_2)} dt_1 dt_2 \\ &= \frac{1}{\Delta t^2} \int_{-\Delta t}^{\Delta t} \big(\Delta t - |t|\big) \, A_{\delta\omega}(t) \, e^{-i\nu_m t} dt \\ &= \int_{-\infty}^{\infty} S_{\delta\omega}(\nu') \, \mathrm{sinc}^2 \left[ \frac{(\nu' - \nu_m)\Delta t}{2} \right] d\nu', \end{aligned} \tag{4.25}$$

where we expressed $A_{\delta\omega}$ as the inverse Fourier transform of $S_{\delta\omega}$ to arrive at the last line. The Fourier series frequencies $\nu_m$ are all spaced apart by $2\pi/\Delta t$, so there are on average $\Delta t \, \Delta\nu / 2\pi$ of them in the band $[\nu, \nu + \Delta\nu]$. Therefore:

$$\lim_{\substack{\Delta\nu \to 0 \\ \Delta t \to \infty}} \frac{1}{\Delta\nu} \sum_{m \text{ in band}} \langle |a_m|^2 \rangle = \lim_{\Delta t \to \infty} \frac{\Delta t}{2\pi} \int_{-\infty}^{\infty} S_{\delta\omega}(\nu') \, \mathrm{sinc}^2 \left[ \frac{(\nu' - \nu)\Delta t}{2} \right] d\nu'. \tag{4.26}$$

Note finally that $\Delta t \, \mathrm{sinc}^2(x \Delta t / 2)$ approaches $2\pi\delta(x)$ for large $\Delta t$, meaning the expression above equals $S_{\delta\omega}(\nu)$ in the large $\Delta t$ limit, as claimed.

### 4.1.2 Common Power Spectra

Every reasonably well-behaved stochastic process has a power spectrum. (For instance, random telegraph noise has a Lorentzian spectrum.) In general, however, stochastic processes can also have non-trivial polyspectra, which are higher-order generalizations of $S_{\delta\omega}$ [126, 127]. Gaussian noise is unique in that all of its higher-order spectra vanish, meaning that it is completely determined by $S_{\delta\omega}$ (and also by its mean, which we have absorbed here into $\omega_0$). It is therefore common to refer to certain noise processes simply by the shape of $S_{\delta\omega}$, which tacitly assumes the noise

to be stationary and Gaussian in the sense of the derivation above.

In practice, the power spectrum of phase noise in many experiments is complicated, and difficult to accurately model *ab initio*. Instead, one must often measure it to achieve a detailed understanding of the noise. As a first approximation, however, it is useful to analyze certain archetypal power spectra, and their resulting coherence decay functions $\langle E(\Delta t)\rangle$.

### White noise

Consider first the case of a flat power spectrum: $S_{\delta\omega}(\nu) = \text{const}$. This is called white noise, in analogy to white light, because all frequencies are equally important. Plugging this into Eq. (4.20), we find that it leads to an exponential decay in qubit coherence

$$\langle E(t)\rangle = e^{-t/T_2^*}, \tag{4.27}$$

for some characteristic timescale $T_2^*$ that depends on the noise strength. We use $t$ here for simplicity to denote the duration of each repetition, now that there is no risk of confusion.

Putting this into Eq. (4.4), we see that the average dynamics is also the solution to the Lindblad equation with $Z$ jumps, as claimed in Chapter 3:

$$\frac{d\rho}{dt} = -i[H_0, \rho] + \frac{1}{2T_2^*}\Big(Z\rho Z - \rho\Big), \tag{4.28}$$

where $H_0 = \frac{\omega_0}{2}Z$ is the average Hamiltonian. (We will write simply $\rho$ rather than $\langle\rho\rangle$ when the intended meaning is clear.) This is not a coincidence. Moving back into the time domain [cf. Eq.(4.17)], white noise has an autocorrelation function $A_{\delta\omega}(t) \propto \delta(t)$, meaning that there is no memory at all, i.e., the noise is totally local in time. Likewise, the Lindblad equation is also local in time, as we saw in Chapter 1.

White noise is inherently unphysical. One way of seeing this is that the area under a flat power spectrum is infinite, meaning that the noise has infinite power at arbitrarily high frequencies. We could deal with this by imposing an artificial cutoff on $S_{\delta\omega}(\nu)$ at high frequencies, which would give the same decay shape of $\langle E(t)\rangle$ up

to a correction that scales inversely with the cutoff frequency. For our purposes, however, it suffices to note that while white noise provides a useful model, one should take care not to use it—and the Lindblad equation more broadly—in calculations where the answer depends critically on the noise's unphysical features. For instance, this is why we could not use white noise to examine Zeno-like effects in Chapter 3.

### $1/f^\alpha$ noise

A ubiquitous type of noise in many quantum and classical devices is $1/f$ noise, where the power spectrum goes as $1/|\nu|$, or more generally as $1/|\nu|^\alpha$ for some $\alpha \sim 1$. [We use $\nu$ rather than $f$ for noise frequencies. Also, it is customary to write simply $S_{\delta\omega}(\nu) \propto 1/\nu^\alpha$ and leave the absolute value implicit.] Such noise can arise when $\delta\omega$ is a linear combination of many weak random telegraph processes, for instance [130, 131]. It is interesting to note that this model is the opposite extreme of that which arose in Chapter 2, where dephasing was primarily due to a single, strongly-coupled fluctuator, rather than a diffuse bath of weakly-coupled ones.

Unfortunately, plugging $S_{\delta\omega}(\nu)$ directly into Eq. (4.20) generally gives a divergent integral. For $\alpha \geq 1$ the problem is the opposite of that with white noise: divergent noise power at low frequencies. A common way to address this is to argue that in an experiment of duration $t$, we can't possibly be sensitive to noise at frequencies much lower than $1/t$. Accordingly, we are justified in imposing a low-frequency cutoff at $c/t$, where $c$ is some unitless constant. Integrating $S_{\delta\omega}(\nu) \times F_t(\nu)$ over $\nu \geq c/t$ gives an average excess phase of

$$\langle E(t) \rangle = e^{-(t/T_2^*)^{\alpha+1}}, \tag{4.29}$$

where $T_2^*$ is a complicated function of the noise strength, $c$ and $\alpha$, that we can fit empirically.

Some remarks are in order. First, notice that for "pure" $1/f$ noise ($\alpha = 1$), the qubit loses phase coherence as $e^{-(t/T_2^*)^2}$. Unfortunately, this common decay profile is sometimes colloquially called "Gaussian" dephasing [as opposed to the exponential dephasing in Eq. (4.27)]. It is crucial to remember that "Gaussian" in this sense has

virtually nothing to do with the underlying assumption of Gaussian noise. Second, observe that pure $1/f$ noise gives the same coherence decay profile as the quasi-static noise [i.e., $S_{\delta\omega}(\nu) \propto \delta(\nu)$] considered in Chapter 3, which is why we used the latter as a rough approximation to more realistic noise. Finally, note that $1/f^{\alpha}$ noise reduces to white noise for $\alpha = 0$.

### 4.1.3 Dynamical Decoupling

Consider dephasing due to stationary Gaussian noise with power predominantly at low frequencies. To a good approximation, such noise would cause a qubit's energy gap to vary between runs of an experiment, but would not cause it to change significantly during any given run. The qubit would dephase approximately as $\langle E(t) \rangle \approx e^{-(t/T_2^*)^2}$ over $[0, t]$ if allowed to evolve freely. Suppose instead that one applied an $X$ gate[2] at time $t/2$, and another at $t$. (We will assume all such pulses to be instantaneous.) The net effect of these pulses would be to effectively reverse the direction of rotation over $[t/2, t]$, so that any phase acquired in the first half of the interval would be canceled out (or "echoed" out, or "refocused") during the second half, thus suppressing dephasing.

Consider now this same approach but for noise with a generic power spectrum. The low-frequency components would still be largely suppressed; in fact, they could be suppressed more effectively by adding more pulses during the interval. The high frequency noise components, on the other hand, would not generally be suppressed. Rather, certain high-frequency components could be amplified if they are resonant with the pulse spacings.

The strategy described here is a simple instance of a technique called dynamical decoupling (DD) [49, 50], which could be used as a first line of defense before quantum error correction. As mentioned above, it has a nice interpretation in terms of the frequency-domain expression for $\langle E(t) \rangle$ in Eq. (4.20). If one repeats the derivation in Section 4.1.1 but with $\pi$ pulses interleaved in $[0, t]$, one would get extra functions multiplying the integrand of Eq. (4.17), which toggle between $\pm 1$ at times that coincide with the pulses. The net result in the frequency domain would be a different filter

[2] Or equivalently $Y$, or any other $\pi$ rotation around an axis on the equator of the Bloch sphere.

function $F_t$ in Eq. (4.20), which depends on the number of pulses applied and their timing. This allows one to engineer a filter function through a careful choice of pulse sequence so that $F_t$ is small at frequencies where $S_{\delta\omega}$ is large [and vice versa, since the area under the curve $F_t(\nu)$ is independent of the pulse sequence], thus minimizing the overlap in Eq. (4.20) and lengthening the coherence time [108, 109]. One could think of this process as creating an effective, smoothed power spectrum, defined by the overlap of the physical power spectrum $S_{\delta\omega}$ and the engineered filter function $F_t$.

Notice that the channel described by Eq. (4.2) is exactly equivalent to a unitary evolution by $H_0$ for a time $t$ and an application of the phase-flip channel $\rho \mapsto (1-p)\rho + pZ\rho Z$ with $p = [1 - \langle E(t)\rangle]/2$. Yet, without considering the physical origins of the aggregate channel, as we have just done, it would be quite difficult to see the potential of DD. This would be a costly oversight: take for example an NV electronic spin, whose dephasing time under a simple DD sequence (denoted $T_2$) can be $\sim 10^3$ times longer than that under free evolution (denoted $T_2^*$) [132]. One could argue that the device-adapted QEC codes in Chapter 2 are analogous to the picture of DD introduced here; they key difference being that the latter exploits temporal noise correlations, while the former exploit spatial correlations. We will develop this analogy further in Chapters 5 and 6 as we compare the roles of QEC and DD in quantum sensing.

## 4.2 Quantum Sensing

Quantum sensing refers to the use of quantum phenomena to measure external physical quantities. It exploits what is otherwise a central weakness of many quantum systems: their strong susceptibility to external disturbances. Much like quantum computers, quantum sensors can be realized using a number of different physical platforms; in fact, there is substantial overlap in the platforms used for both. Specifically, quantum sensors can be realized using certain spin qubits (such as NV centers), trapped ions, and superconducting qubits, among others. They can measure several different quantities depending on the platform used, including electromag-

netic fields, temperature and rotation, with remarkable sensitivity. In contrast with quantum computers, however, quantum sensors are expected to be a relatively near-term quantum technology. A thorough review of quantum sensing can be found in Ref. [108].

The typical strategy for quantum sensing is to estimate some parameter affecting the sensor's dynamics, which depends on its local environment. This could be, for instance, the local magnetic field, which affects the sensor's energy eigenvalues. In this approach, the sensor is prepared in some known initial state, allowed to evolve, and the unknown quantity is ultimately inferred from measurements of the sensor's final state through many repetitions [133, 134].

Like quantum computers, quantum sensors can comprise subsystems of varied dimensionality, although most commonly both are composed of multiple qubits. We will assume this structure throughout. Not only are many types of qubits inherently sensitive to their environments, but the performance of an $n$-qubit sensor can improve faster with $n$ than classical physics would allow. Moreover, such sensors can offer other practical advantages: for instance, the small size of certain quantum sensors can allow for exceptional spatial resolution.

We will begin by discussing how a single qubit can be used for quantum sensing, and later extend our results to various multi-qubit strategies. There are two canonical single-qubit strategies: one in which the quantity of interest becomes manifest as a relative phase between energy eigenstates (often called Ramsey sensing), and another in which it is manifest in the transition rate between energy eigenstates (often called Rabi sensing) [108]. We will focus on the former strategy, which is both common and flexible, although the results derived in this chapter and subsequent ones could be generalized to the latter strategy as well.

Concretely, the sensing scheme we consider allows one to measure an external quantity which couples to a sensing qubit via its energy gap. That is, suppose the qubit's Hamiltonian is

$$H(t) = \frac{\gamma\omega(t)}{2} Z, \tag{4.30}$$

where $\omega(t)$ is a variable, extrinsic quantity to be measured (e.g., a local magnetic field component), and $\gamma$ is a constant, intrinsic parameter—sometimes called the transduction parameter—quantifying the sensor's susceptibility to changes in $\omega(t)$ (e.g., the qubit's magnetic dipole moment). The strategy is to

1. Prepare an initial state $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$, perhaps by first preparing $|0\rangle$ and performing a Hadamard gate.

2. Let the sensor evolve by $H(t)$ for a time $\Delta t$, which leads to a time-evolved state

$$\frac{1}{\sqrt{2}}\left(|0\rangle + \exp\left[i\gamma\int_0^{\Delta t}\omega(t)dt\right]|1\rangle\right). \tag{4.31}$$

   up to a global phase.

3. Apply a control pulse to rotate the state by $\pi/2$ about some axis on the equator of the Bloch sphere. The optimal choice of this axis depends on the accumulated phase, meaning that it may need to be chosen adaptively. We will assume here instead that one has an initial estimate of $\omega(t)$, which is to be refined through quantum sensing. One can then define a rotating frame based on this initial estimate, in which the Hamiltonian keeps the form of that in Eq. (4.30) but with $\int_0^{\Delta t}\omega(t)dt \ll 1$. In this frame the $\pi/2$ pulse should rotate about $X$. A measurement in the computational basis then returns 1 with probability

$$p_1 = \frac{1}{2} - \frac{1}{2}\sin\left[\gamma\int_0^{\Delta t}\omega(t)dt\right] \tag{4.32}$$

   and 0 otherwise.

One repeats this process many times to estimate $p_1$, which in turn yields an estimate for $\int_0^{\Delta t}\omega(t)dt$.

Notice that one needs[3] a large $\gamma$ so that $p_1$ depends strongly on $\omega(t)$. A large $\gamma$ comes at a cost, however. In practice, the form of $p_1$ will gradually deviate from that

[3]Cf. Eq. (2.3), where the gyromagnetic ratio of an NV electronic spin is $\sim 10^3$ times larger than that of a $^{13}$C nuclear spin. We therefore expect an NV to be a far better magnetometer than a $^{13}$C nuclear spin, which is indeed true.

in Eq. (4.32) due to decoherence in the qubit; we expect that for $\Delta t$ much longer than the decoherence timescale, the value of $p_1$ will carry little information about $\omega$, thus limiting sensitivity. *A priori* there could be many physical processes contributing to such decoherence, which might vary greatly based on the device in question. When $\gamma$ is large, however, background fluctuations in $\omega(t)$, of the type analyzed in Section 4.1, induce rapid dephasing, which is often the limiting source of decoherence in quantum sensors. That is, in practice $\omega(t)$ could be decomposed as

$$\omega(t) = \omega_0(t) + \delta\omega(t), \tag{4.33}$$

where $\omega_0(t) := \langle\omega(t)\rangle$ is the deterministic part and $\delta\omega(t)$ is the stochastic part causing dephasing. [Without loss of generality we can choose $\delta\omega(t)$ to have zero mean.] This latter part is present to some extent in all qubits, although it need not be the dominant source of decoherence in general. The rate of dephasing due to $\delta\omega$ grows with $\gamma$, however. This means that for large enough $\gamma$ such dephasing becomes the dominant kind of decoherence, regardless of other device physics. This is the case in a wide range of quantum devices used for sensing, precisely because of their intended application. In this chapter, as well as Chapters 5 and 6, we will therefore assume background noise in $\omega$ to be the limiting source of decoherence. Under this assumption, $p_1$ is well-approximated by

$$p_1 = \frac{1}{2} - \frac{1}{2}\langle E(\Delta t)\rangle \sin\left[\gamma \int_0^{\Delta t} \omega(t)dt\right] \tag{4.34}$$

when one accounts for decoherence, assuming that $\langle E(\Delta t)\rangle \geq 0$ (as is the case for stationary Gaussian noise). We will explore in Chapters 5 and 6 how application-adapted QEC codes can be used to suppress such decoherence.

### 4.2.1 Sensitivity

There are quantum sensing schemes both for time-dependent and static, i.e., DC signals $\omega_0(t)$. We will discuss the time-dependent case in Section 4.2.4, focusing here

instead on the DC case in which $\int_0^{\Delta t} \omega_0(t)dt = \omega_0 \Delta t$, where we have used $\omega_0$ to denote the constant average value of $\omega(t)$.

The sensitivity[4] $\eta$ offered by a sensor is typically defined as the smallest $\omega_0$ that gives a signal-to-noise ratio (SNR) of at least 1, normalized by the total measurement time. This means (somewhat unfortunately) that a lower sensitivity is better. We now calculate the sensitivity offered by the Ramsey measurement scheme above under the assumption that the dominant source of decoherence is background noise in $\omega(t)$ producing $\langle E(t) \rangle \in \mathbb{R}_{\geq 0}$ (such as Gaussian noise).

Suppose that one repeats the Ramsey sensing scheme with a single qubit $N_{\text{tot}}$ times so as to estimate $p_1$, and ultimately $\omega_0$. The number of measurements yielding 1, $N_1$, follows a binomial distribution with mean $\mu = N_{\text{tot}}\, p_1$ and standard deviation $\sigma = \sqrt{N_{\text{tot}}\, p_1(1-p_1)}$. It is natural to estimate $p_1$ with the maximum likelihood estimator $\hat{p}_1 = N_1/N_{\text{tot}}$. We will take as the "signal" in the SNR the absolute mean of $\hat{p}_1 - 1/2$, namely $\left|p_1 - 1/2\right|$, where the $1/2$ offset is because $\langle \hat{p}_1 \rangle|_{\omega_0=0} = 1/2$. Similarly, we will take the "noise" to be the standard deviation in $\hat{p}_1 - 1/2$, which is

$$\frac{\sigma}{N_{\text{tot}}} = \sqrt{\frac{p_1(1-p_1)}{N_{\text{tot}}}}. \tag{4.35}$$

Note that the choice of rotation axis in the Ramsey scheme discussed above affects both the signal and noise strengths. Our choice gives the optimal compromise.

The SNR is therefore

$$\begin{aligned} \text{SNR} &= \left| p_1 - \frac{1}{2} \right| \sqrt{\frac{N_{\text{tot}}}{p_1(1-p_1)}} \\ &= \omega_0 \gamma \Delta t \, \langle E(\Delta t) \rangle \sqrt{N_{\text{tot}}} + O\big[(\omega_0 \gamma \Delta t)^3\big], \end{aligned} \tag{4.36}$$

so the smallest $\omega_0$ giving unit SNR is approximately

$$\omega_{0,\text{min}} = \frac{1}{\gamma \Delta t \, \langle E(\Delta t) \rangle \sqrt{N_{\text{tot}}}} = \frac{1}{\gamma \Delta t \, \langle E(\Delta t) \rangle} \sqrt{\frac{\Delta t}{t_{\text{tot}}}}, \tag{4.37}$$

[4] In other contexts this is called the noise-equivalent power.

where we have expressed the number of cycles $N_{\text{tot}}$ in terms of the cycle duration $\Delta t$ and the total sensing period $t_{\text{tot}} = N_{\text{tot}}\,\Delta t$. Note that this expression assumes $\omega_{0,\min}\,\gamma\Delta t \ll 1$, or equivalently, $N_{\text{tot}} \gg \langle E(\Delta t)\rangle^{-2}$. Moreover, we are assuming instantaneous state preparation and measurement, and neglecting any measurement error; relaxing these idealizations leads to simple corrections in the expression above [108].

While $\omega_{0,\min}$ tells us something about the sensor's performance, it is, by itself, a flawed figure of merit because it depends explicitly on $t_{\text{tot}}$. This means one can achieve an arbitrarily low $\omega_{0,\min}$ in any (unbiased) sensor through a long enough $t_{\text{tot}}$. Instead, we get a better figure of merit by normalizing $\omega_{0,\min}$ by $t_{\text{tot}}$; here this means multiplying by $\sqrt{t_{\text{tot}}}$ rather than dividing by $t_{\text{tot}}$, as one might have initially expected. The resulting quantity is the sensitivity:

$$\eta_1(\Delta t) := \omega_{0,\min}\sqrt{t_{\text{tot}}} = \frac{1}{\gamma\sqrt{\Delta t}\,\langle E(\Delta t)\rangle}, \tag{4.38}$$

which has units of $[\omega_0] \times \sqrt{\text{time}}$, or more commonly $[\omega_0]/\sqrt{\text{frequency}}$, which is why, e.g., the sensitivity of magnetometers is often reported in otherwise-mysterious units of $\text{T}/\sqrt{\text{Hz}}$. The subscript "1" here refers to the use of a single qubit. Again, note that this expression for $\eta_1(\Delta t)$ only makes sense when $N_{\text{tot}} \gg \langle E(\Delta t)\rangle^{-2}$.

It remains to calculate how long each cycle should be. We will denote the optimal sensitivity as

$$\eta_1 := \min_{\Delta t} \eta_1(\Delta t), \tag{4.39}$$

here and as $\eta$ (with no $\Delta t$ dependence) more broadly. A longer $\Delta t$ means more time for $\omega_0$ to imprint on the sensor, but also more decoherence, and a smaller $N_{\text{tot}}$ for any given $t_{\text{tot}}$. Assuming $1/f^\alpha$ noise as in Section 4.1.2 for $\alpha \geq 0$, so that at $\langle E(t)\rangle = \exp[-(t/T_2^*)^{\alpha+1}]$, the optimal (i.e., minimal) sensitivity is

$$\eta_1 = \frac{(2e\beta)^{1/2\beta}}{\gamma\sqrt{T_2^*}}, \tag{4.40}$$

where $\beta := 1 + \alpha$, which uses a cycle length of

$$\Delta t_{\text{opt}} = T_2^* \, (2\beta)^{-1/\beta}. \tag{4.41}$$

This expression for $\eta_1$ is therefore valid for $N_{\text{tot}} \gg e^{\alpha+1}$. We will be particularly concerned with white noise ($\alpha = 0$) in the coming chapters, for which these quantities reduce to $\eta_1 = \gamma^{-1}\sqrt{\frac{2e}{T_2^*}}$ and $\Delta t_{\text{opt}} = T_2^*/2$, assuming $N_{\text{tot}} \gg 1$.

These results highlight an important feature of quantum sensing: it works best in a relatively high-noise regime. That is, the optimal duration of each cycle, $\Delta t_{\text{opt}}$, is comparable to the coherence time $T_2^*$. In particular, for $1/f$ noise and white noise, the best sensitivity is achieved by letting the qubit coherence decay by $1 - \langle E(\Delta t_{\text{opt}})\rangle = 22\%$ and $39\%$ respectively in each cycle. This holds true regardless of $T_2^*$; that is, if one increased $T_2^*$ (e.g., through QEC), the corresponding $\Delta t_{\text{opt}}$ would increase accordingly. The fact that quantum sensors naturally tolerate some noise is very appealing, especially when one considers the gap between current noisy devices and the ultra-low noise rates required for many quantum computing applications.

### 4.2.2 Quantum Cramér-Rao Bound

A quantum sensor's capacity to estimate $\omega_0$ can also be quantified in a more formal way. In general, consider estimating an unknown quantity $\omega_0$ through $N_{\text{tot}}$ repeated independent measurements of duration $\Delta t$, using a sensor that outputs a number $x_j$ in each run depending on $\omega_0$. Let $\hat{\omega}_0 = \hat{\omega}_0(x_1, \ldots, x_{N_{\text{tot}}})$ be an unbiased estimator of $\omega_0$, meaning that $\langle \hat{\omega}_0 \rangle = \omega_0$. Then, the variance of $\hat{\omega}_0$, which formally quantifies how well one can estimate $\omega_0$ from the measurement results, is bounded as

$$\text{var}(\hat{\omega}_0) \geq \frac{1}{N_{\text{tot}} \, F}, \tag{4.42}$$

where $F$ is called the Fisher information [135], and is defined as

$$F = -\Big\langle \frac{\partial^2}{\partial \omega_0^2} \ln \Pr(x_j | \, \omega_0) \Big\rangle, \tag{4.43}$$

where $\Pr(x_j|\omega_0)$ is the likelihood of the sensor outputting $x_j$ in the $j^{\text{th}}$ measurement given an underlying value of $\omega_0$. Eq. (4.42) is called the Cramér-Rao bound [136, 137]. If one takes as $\hat{\omega}_0$ the maximum likelihood estimator (that is, if $\hat{\omega}_0$ is taken as the value most likely to produce the observed data), which we will assume from now on, then the bound is asymptotically saturated for large $N_{\text{tot}}$. That is [138, 139]:

$$\text{var}(\hat{\omega}_0) = \frac{1}{N_{\text{tot}}\, F}\big[1 + O(1/N_{\text{tot}})\big]. \tag{4.44}$$

For a quantum sensor, the Fisher information $F$, and in turn $\text{var}(\hat{\omega}_0)$, depends on which measurement one performs on the time-evolved sensor state. By optimizing $F$ over all possible measurements one gets the quantum Fisher information $\mathcal{F}$ [140]. For a time-evolved sensor state $\rho_{\omega_0} = \rho_{\omega_0}(\Delta t)$ which depends on $\omega_0$, this quantity is given by [141]

$$\mathcal{F} = 2\int_0^\infty \text{tr}\left[\left(e^{-\lambda\,\rho_{\omega_0}}\,\frac{\partial \rho_{\omega_0}}{\partial \omega_0}\right)^2\right] d\lambda. \tag{4.45}$$

Substituting $F$ for $\mathcal{F}$ in the Cramér-Rao bound gives the quantum Cramér-Rao bound. For a sensing qubit under stationary Gaussian phase noise, as we considered in Section 4.2.1, the quantum Fisher information for an initial state $\rho_0 = |+\rangle\langle+|$ that has evolved for time $\Delta t$ is

$$\mathcal{F} = \big[\langle E(\Delta t)\rangle\, \gamma \Delta t\big]^2. \tag{4.46}$$

Note that here the implicit optimization over possible measurements in $\mathcal{F}$ is equivalent to our earlier choice of the optimal $\pi/2$ rotation axis in the Ramsey scheme.

When devising quantum sensing schemes, one may be inclined to maximize $\mathcal{F}$. However, this is usually a flawed objective function in practice, i.e., when there is decoherence. Fundamentally, the quantity to be optimized (here, minimized) is

$$\text{var}(\hat{\omega}_0) \sim \frac{1}{N_{\text{tot}}\,\mathcal{F}} = \frac{\Delta t}{\mathcal{F}}\, t_{\text{tot}}^{-1}, \tag{4.47}$$

assuming $N_{\text{tot}} \gg 1$. As in our earlier derivation of the sensitivity $\eta$, we see that $\text{var}(\hat{\omega}_0)$

can be made arbitrarily small by making the total sensing time $t_{\text{tot}}$ sufficiently large, regardless of sensor performance. For a fixed $t_{\text{tot}}$, minimizing $\text{var}(\hat{\omega}_0)$ amounts not to maximizing $\mathcal{F}$, but rather $\mathcal{F}/\Delta t$, which here is equal to

$$\frac{\mathcal{F}}{\Delta t} = \left[\langle E(\Delta t)\rangle \, \gamma\sqrt{\Delta t}\right]^2. \tag{4.48}$$

While we have arrived at this quantity in a fairly abstract way, it has a simple physical interpretation. Namely, notice here by comparing with Eq. (4.38) that $\mathcal{F}/\Delta t = \eta_1(\Delta t)^{-2}$, so $\max_{\Delta t}(\mathcal{F}/\Delta t) = \eta_1^{-2}$.

### 4.2.3 Scaling with $n$

The two derivations in Sections 4.2.1 and 4.2.2 both suggested the same figure of merit $\eta$. Accordingly, we will mostly treat this quantity as the natural figure of merit in Chapters 5 and 6. There is however, one limiting situation in which $\eta$ is not a useful metric: that where there is no decoherence. Notice in Eqs. (4.38) and (4.48) that if $\langle E(\Delta t)\rangle$ does not decrease with $\Delta t$, $\eta(\Delta t)$ can be made arbitrarily small (or $\mathcal{F}/\Delta t$ arbitrarily large) by using a long $\Delta t$. In principle, this eventually makes $N_{\text{tot}} = t_{\text{tot}}/\Delta t$ too small for the derivations in Sections 4.2.1 and 4.2.2—which assumed some decoherence—to make sense. Rather, in this idealized case where the quantum Fisher information grows as $\mathcal{F} = (\xi\Delta t)^2$, it is natural to use the coefficient $\xi$ as a figure of merit instead. For a single noiseless qubit, for instance, $\xi = \gamma$. In introducing multi-qubit sensing schemes below we will begin by briefly examining the scaling of $\xi$ with the number of qubits $n$, before ultimately returning to the more broadly relevant $\eta$ as a figure of merit.

The quantum Fisher information has the important property that $\mathcal{F}(\rho_{\omega_0}^{\otimes n}) = n\mathcal{F}(\rho_{\omega_0})$ [108]. This means that if one performs Ramsey sensing on $n$ identical noiseless qubits in parallel, one gets an $\sqrt{n}$-fold enhancement in $\xi$. This is the same scaling as one gets using $n$ classical sensors in parallel. Such scaling is often called (somewhat confusingly) the standard quantum limit. If instead, however, one prepares a GHZ

state on $n$ such qubits, the dynamics will be

$$\frac{1}{\sqrt{2}}\Big( |0\ldots0\rangle + |1\ldots1\rangle \Big) \mapsto \frac{1}{\sqrt{2}}\Big( |0\ldots0\rangle + e^{i\gamma n\omega_0\Delta t}\,|1\ldots1\rangle \Big), \tag{4.49}$$

which looks like that of a single qubit but with an effective $n$-fold enhancement in the transduction parameter $\gamma$, resulting in an $\xi = n\gamma$; a distinctly quantum-mechanical scaling that is sometimes called Heisenberg scaling. We now return to the more realistic case where there is decoherence, in which the sensitivity $\eta$ is the natural figure of merit.

All of the multi-qubit sensing schemes considered in this thesis will involve a relative phase accumulating between two orthogonal states at a rate proportional to $\omega_0$, and a loss of phase coherence between these states on a timescale proportional to $T_2^*$. Rather than calculate the sensitivity $\eta$ offered by each scheme from scratch, we can map them to effective single-qubit dynamics from which $\eta$ can be found immediately. Concretely, suppose the effective dynamics arising from an $n$-qubit sensing scheme looks like that of a single qubit with an average energy gap $A\gamma\omega_0$, and which dephases under $1/f^\alpha$ noise ($\alpha \geq 0$) on a timescale $T_2^*/B$ for constants $A$ and $B$ that depend on the scheme. The optimal sensitivity for general $A$ and $B$ is then

$$\eta = \frac{\sqrt{B}}{A}\,\eta_1, \tag{4.50}$$

which can represent a net enhancement or degradation compared to $\eta_1$ depending on the relative values of $A$ and $B$. As noted above, a GHZ sensing scheme gives $A = n$. Unfortunately, as we will see in Chapter 5, states like those in Eq. (4.49) dephase on a timescale of $nT_2^*$ (i.e., $B = n$) under spatially-uncorrelated phase noise, which leads only to a $1/\sqrt{n}$ enhancement in sensitivity compared to a single physical qubit. Likewise, one can see from Section 4.2.1 that using $n$ qubits in parallel produces the same $1/\sqrt{n}$ enhancement in sensitivity. In other words, even weak spatially-uncorrelated phase noise destroys the Heisenberg scaling offered by GHZ states, instead reducing it to the standard quantum limit. (We will see in subsequent chapters that these two schemes do not generally give the same $\eta$ when the noise has non-trivial spatial

correlations).

This result highlights one of the central challenges in quantum sensing: it is not enough to simply amplify the signal if the noise is also amplified by a corresponding amount [142]. The same goes for attenuating the noise. Rather, one must find methods to carefully filter noise from the signal, so as to produce a net improvement in sensitivity.

### 4.2.4 AC Signals

A common way to performing such filtering is with dynamical decoupling, as introduced in Section 4.1.3. That is, rather than allow the sensor to evolve freely over $[0, \Delta t]$, one instead applies a series of control pulses (or continuous controls) during this interval.

To see how such pulses interact with the deterministic part of $\omega(t)$, namely $\omega_0(t)$, consider the AC signal

$$\omega_0(t) = \Omega \sin(2\pi f t + \phi), \tag{4.51}$$

where $f$ and $\phi$ are known and the amplitude $\Omega$ is to be measured. Notice first that in the absence of control pulses, $\omega_0(t)$ does not lead to a steady accumulation of phase in the sensor like a DC signal; rather, $\int_0^{\Delta t} \omega(t) dt$ oscillates with $\Delta t$. Instead, one could rectify the signal by applying $\pi$ pulses about some axis in the $xy$-plane at times $t_\star$ where $\omega_0(t_\star) = 0$, which amounts to a form of DD. The effects would be twofold. First, these added pulses would lead to a gradual accumulation of phase in the sensor as $\int_0^{\Delta t} |\omega_0(t)| dt$. Second, they would define a filter function $F_{\Delta t}$ which suppresses some noise frequencies and amplifies others [108, 143]. (This strategy can easily be generalized for situations when the signal's phase is not known [144].)

In fact, the interaction of such pulse sequences with a general $\omega_0(t)$ can be understood using a similar formalism to that of filter functions for $\delta\omega$. Broadly, one can design sequences which amplify both signal and noise in certain frequency bands on the sensor, while suppressing both at other frequencies. A simple and powerful way to sense AC signals at known frequencies, then, is to use DD sequences which create

an effective bandpass filter around the signal frequency, and which (approximately) reject most other frequencies. One can use this approach more broadly to measure Fourier components of time-dependent signals.

While this approach can give remarkably good sensitivity, it has some notable limitations. First, it is mostly incompatible with DC and low-frequency signals, since the DD passband cannot be placed at arbitrarily low frequencies. Second, DD cannot suppress high-frequency noise components, so these continue to limit sensitivity [109]. Third, this technique of creating an effective narrow frequency filter is cumbersome when measuring wideband signals arising in some applications.

In Chapters 5 and 6 we will develop application-adapted QEC codes for sensing which filter noise from signal in order to enhance sensitivity. They work in a way that is analogous to DD, albeit with an essential difference: they act as filters in the spatial domain, whereas DD acts as a filter in the time/frequency domain. That is, DD exploits temporal noise correlations, whereas our QEC codes exploit spatial noise correlations. As we will see, QEC constitutes a complementary—or perhaps even an alternative—technique to DD for sensing; since the two exploit different properties of the signal and noise, they have both different strengths and different limitations.

# Chapter 5

# Spatial Noise Filtering through QEC for Quantum Sensing

Noise in candidate systems for quantum technologies often has structure. In small- and intermediate-scale systems, this structure can often be measured or derived from first principles, at least in part. A central theme in this thesis so far has been that while conventional QEC works regardless of such structure (mostly), it can be very advantageous to incorporate any known noise structure into QEC schemes from the start. In this chapter and in Chapter 6, we will apply this principle in the context of quantum sensing. Rather than adapt QEC codes to particular devices to maximize noise suppression (as in Chapter 2), we will instead adapt codes for the unique challenges posed by sensing applications, across many devices, so as to optimize sensitivity.

## 5.1 Error-Corrected Quantum Sensing

As we discussed in Chapter 4, one can improve the sensitivity of a quantum sensor through control schemes in which the signal gets amplified more than the noise, or, the noise gets attenuated more than the signal. We can think of such schemes as filters. Dynamical decoupling (DD) sequences provide a family of such filters, which separate signal from noise based on their respective temporal structures. A complementary

family of filters based instead on quantum error correction (QEC) has been proposed more recently, which do not suppress noise on the basis of frequency, and therefore do not share the same limitations as DD [145–148].

The canonical scheme for error-corrected quantum sensing (ECQS), illustrated here with a Lindblad description, is: (i) to prepare a superposition of logical energy eigenstates, (ii) to let the sensor evolve for a time $\delta t$ under the Liouvillian $\mathcal{L} = -i\mathcal{H}_0 + \mathcal{D}$, where the Hamiltonian superoperator $\mathcal{H}_0(\rho) = [H_0, \rho]$ is proportional to the parameter one wants to estimate, and (iii) to apply a recovery operation $\mathcal{R}$ which seeks to correct the effects of the noise from

$$\mathcal{D}(\rho) = \sum_i L_i \rho L_i^\dagger - \frac{1}{2}\{L_i^\dagger L_i, \rho\}, \tag{5.1}$$

where $\{L_i\}$ are the Lindblad jump operators describing decoherence. (The recovery is approximated as being instantaneous.) Steps (ii) and (iii) are repeated until (iv) the final state is read out after a total time $\Delta t$. In the limit where (ii)–(iii) are fast and repeated many times ($\delta t \to 0$ with $\Delta t$ finite), the sensor evolves stroboscopically as

$$\mathcal{L}_{\text{eff}} = -i\mathcal{R}\mathcal{H}_0 + \mathcal{R}\mathcal{D} + O\big(||\mathcal{L}||\,\delta t\big) \tag{5.2}$$

according to Chernoff's theorem (Ref. [149] p. 241, see also Ref. [150]). If $\mathcal{R}\mathcal{D} = 0$ but $\mathcal{R}\mathcal{H}_0 \neq 0$ on logical states then ECQS can approach a noiseless sensing limit by making $\delta t$ sufficiently short compared to the noise strength. As in Chapter 4, one then repeats this overall scheme $N_{\text{tot}} \gg 1$ times.

For ECQS to provide such a noiseless $\delta t \to 0$ limit ($\mathcal{R}\mathcal{D}|_{\text{code}} = 0$), the jump operators for the sensor must satisfy the Knill-Laflamme condition applied to Lindblad operators as in Eqs. (1.45) and (1.46), namely:

$$PL_iP \propto P \qquad\qquad PL_i^\dagger L_j P \propto P \tag{5.3}$$

for all $i$ and $j$, where $P = P^\dagger$ projects onto the codespace. For the signal to survive in this limit ($\mathcal{R}\mathcal{H}_0|_{\text{code}} \neq 0$), however, $H_0$ must not be fully correctable; that is, it

must not satisfy the Knill-Laflamme condition applied to the Hamiltonian part of $\mathcal{L}$ in Eq. (1.47), namely:

$$PH_0P \not\propto P. \tag{5.4}$$

Naturally then, $H_0$ must not be a linear combination of $L_i$, $L_i^\dagger$ and $L_i^\dagger L_j$ terms. Indeed, it was recently proven that there exists a code with projector $P$ satisfying these ECQS conditions, Eqs. (5.3) and (5.4), if and only if $H_0 \notin \mathscr{S}$ where

$$\mathscr{S} = \text{span}\{I, L_i, L_i^\dagger, L_i^\dagger L_j\} \tag{5.5}$$

is the so-called Lindblad span [151, 152].

An archetypal example of ECQS to measure a DC signal $\omega_0$ was proposed in Refs. [147, 148, 153]: it involves a three-qubit sensor with $H_0 = \sum_{i=1}^{3} \frac{\gamma\omega_0}{2} Z_i$ subject to independent bit-flip errors on each qubit ($L_i \propto X_i$). Initializing the sensor in $|+_{\text{L}}\rangle = \frac{1}{\sqrt{2}}(|0_{\text{L}}\rangle + |1_{\text{L}}\rangle)$, where $|0_{\text{L}}\rangle = |000\rangle$ and $|1_{\text{L}}\rangle = |111\rangle$, the errors can be detected and corrected by the bit-flip code recovery $\mathcal{R}$ [154], while the signal $\omega_0$ imprints through $H_0$ as a relative phase in the encoded state. The signal is unimpeded by the frequent applications of $\mathcal{R}$ as it couples to the qubits through a different operator than the noise, giving $H_0 \notin \mathscr{S}$, and hence $\mathcal{R}\mathcal{D} = 0$ but $\mathcal{R}\mathcal{H}_0 \neq 0$ on logical states. Therefore, the scheme enables near-noiseless sensing of $\omega_0$ for short $\delta t$. To our knowledge, all explicit ECQS schemes for multi-qudit sensors prior to Ref. [60] (on which this chapter is based) operate similarly, correcting only for noise which couples to the sensor via different operators than the signal [145–148, 153, 155–162].

It is useful to recall the analogy from Chapter 1, of QEC codes as defining a "net" which catches (and ultimately undoes) a family of disturbances on the system, both unitary and non-unitary. Eq. (5.3) tells us what this net should encompass if it is to correct all errors to order $O(\delta t)$. Quantum sensing applications, however, pose an additional constraint in the form of Eq. (5.4), which also tells us what this net must *not* encompass. These conditions are easily satisfied when the signal and the dominant decoherence act on the sensor through different operators, as in the previous example. Suppose instead, however, that $H_0$ were the same but the jump

operators were $L_i \propto Z_i$. In this case Eqs. (5.3) and (5.4) could not be satisfied, as no code could filter noise from signal since $H_0 \in \mathscr{S}$. The issue here is that the signal and the noise are too similar, and cannot be completely distinguished at order $O(\delta t)$ by QEC, even in principle.

This issue illustrates a deficiency in many ECQS schemes, of both practical and fundamental importance. As we discussed in Chapter 4, quantum sensors generally require large transduction parameters $\gamma$. This, however, makes them sensitive to background fluctuations in the quantity they seek to measure, which causes dephasing. When $\gamma$ is sufficiently large, this dephasing becomes the dominant type of decoherence, independent of other device physics. Indeed, such dephasing, in which the noise couples through the same operators as the signal, is the dominant type of decoherence in many quantum sensors by several orders of magnitude [51, 86–90]. While DD can suppress certain frequency components of this noise when sensing AC signals, most early ECQS schemes could not address it at all, to our knowledge. This apparent Achilles' heel would seem to severely limit the practical potential of QEC in many quantum sensors.

In this chapter we will show how application-adapted QEC could be used to filter out noise which couples to each qubit in a quantum sensor identically to the signal, and therefore improve sensitivity. Like DD, our QEC codes will exploit structure—specifically, correlations—that can be present in this background noise. In both cases, stronger noise correlations typically enable better sensitivity. While DD exploits temporal correlations, however, our codes will exploit spatial correlations. In a sense, our scheme is dual to dynamical decoupling for quantum sensing; its strengths and limitations are therefore "orthogonal" to those of DD. The key insight is that Eqs. (5.3) and (5.4) can be viewed not only as a condition on how signal and noise couple to the sensor, as with earlier ECQS codes, but also as a condition on the spatial profiles of each in the sensor. In the sections that follow, we will show how QEC could filter phase noise in sensors based on its spatial correlations, and illustrate this concept with representative examples. In Chapter 6 we will then develop a broad family application-adapted QEC codes to exploit the opportunity identified here. This final

step will allow us to systematically quantify the sensitivity enhancement offered by QEC in this setting.

## 5.2 Noise Model

We consider the task of measuring a DC signal $\omega_0$, inaccessible through DD, under decoherence that is predominantly from background noise $\delta\omega$. For simplicity, we assume a sensor comprising $n$ identical qubits, although this requirement could easily be relaxed at the cost of more cumbersome calculations. Concretely, following Section 4.2, we take the sensor's Hamiltonian to be

$$H(t) = \frac{\gamma}{2} \sum_{i=1}^{n} \omega_i(t)\, Z_i \tag{5.6}$$

in a suitable reference frame, where $\omega_i(t) = \omega_0 + \delta\omega_i(t)$. As in the previous chapter, $\omega_0 = \langle \omega_i(t) \rangle$ is the small, deterministic part of this unknown quantity (assumed to be the same on all qubits), and $\delta\omega_i$ describes zero-mean fluctuations on qubit $i$ (which can be different on each qubit). We will assume each qubit to have the same transduction parameter $\gamma$ to simplify the calculation, although we will comment on the more general case below. Moreover, we are tacitly assuming that $\gamma$ is sufficiently large that these fluctuations are the dominant source of decoherence.

Furthermore, we will assume that $\delta\omega_i$ is a stationary Gaussian white noise process with strength $\Gamma$ across all qubits. There are three reasons for this choice: First, it will allow us to easily move between the languages of classical dephasing and QEC. Second, it is a pessimistic assumption, since it is the canonical noise type that cannot be suppressed through DD. Third, our QEC codes will filter out noise based on its spatial correlations; assuming a complete lack of temporal correlations will therefore allow us to isolate the effects of QEC. Concretely, we take the noise autocorrelation function to be

$$\big\langle \delta\omega_i(t)\, \delta\omega_i(0) \big\rangle = 2\Gamma\delta(t) \tag{5.7}$$

on all qubits $i \in \{1, \ldots, n\}$. As we saw in Chapter 4, each qubit therefore loses phase

coherence as $\langle E(t)\rangle = e^{-t/T_2^*}$, with the same characteristic timescale

$$T_2^* = \frac{1}{\gamma^2\Gamma}. \tag{5.8}$$

This means that when operated alone, each qubit would give an optimal sensitivity of

$$\eta_1 = \gamma^{-1}\sqrt{\frac{2e}{T_2^*}} = \sqrt{2e\Gamma}, \tag{5.9}$$

which depends on the noise strength $\Gamma$ but not on the transduction parameter $\gamma$, in principle. Similarly, the $n$ qubits operated in parallel would give a sensitivity of

$$\eta_{\text{par}} = \frac{1}{\sqrt{n}}\,\eta_1. \tag{5.10}$$

Although we assume temporally-uncorrelated noise, we will allow general spatial correlations between $\delta\omega_i(t)$ and $\delta\omega_j(t)$. To make progress, we must make additional assumptions about the noise that are natural extensions of our assumptions so far: First, that $\delta\omega_i(t)$ and $\delta\omega_j(t)$ are jointly Gaussian and jointly stationary [163]. More concretely, we are assuming that not only is $\delta\omega_i(t)$ a Gaussian process, but so is $\delta\omega_i(t) \pm \delta\omega_j(t)$. Moreover, we assume not just that the autocorrelation functions $\langle\delta\omega_i(t_1)\,\delta\omega_i(t_2)\rangle$ depend only on $|t_1 - t_2|$, but that the same be true of the cross-correlation functions $\langle\delta\omega_i(t_1)\,\delta\omega_j(t_2)\rangle$. Furthermore, since we have assumed that the noise $\delta\omega_i$ on each qubit has no temporal correlations, we also naturally assume that $\langle\delta\omega_i(t)\,\delta\omega_j(0)\rangle \propto \delta(t)$ for $i \neq j$.

What range of cross-correlation strengths makes sense? The extreme cases are when $\delta\omega_i(t) = \pm\delta\omega_j(t)$; that is, when the fluctuations are exactly the same or exactly opposite on two different qubits. This leads to $\langle\delta\omega_i(t)\,\delta\omega_j(0)\rangle = \pm 2\Gamma\delta(t)$. Intermediate cases are possible too; e.g., if the fluctuations are independent on different qubits we get $\langle\delta\omega_i(t)\,\delta\omega_j(0)\rangle = 0$. Naturally then, the possible cross-correlation strengths between $\delta\omega_i$ and $\delta\omega_j$ are in the range $[-\Gamma, \Gamma]$. We can therefore write the auto/cross-

correlation functions compactly as

$$\big\langle \delta\omega_i(t)\, \delta\omega_j(0) \big\rangle = c_{ij}\, 2\Gamma\delta(t) \tag{5.11}$$

in terms of the overall noise strength $\Gamma$ and the coefficients $c_{ij} \in [-1, 1]$ describing the noise correlations between qubits $i$ and $j$. The extreme values of $c_{ij} = \pm 1$ describe identical $(+1)$ or opposite $(-1)$ fluctuations on either qubit, whereas $c_{ij} = 0$ means no correlation. Naturally $c_{ii} = 1$. Note that while this noise is separate from $\omega_0$, it may still arise from within the experimental device, e.g., from the surrounding nuclear bath in the case of spin qubits.

To see when and how QEC can enhance sensitivity, it will be useful to convert the average dynamics resulting from this noise model to a more explicit equation of motion. As we saw in Chapter 4, each qubit, when taken individually, evolves according to the Lindblad equation $\dot{\rho} = \mathcal{L}(\rho)$, where

$$\mathcal{L}(\rho) = \underbrace{-i\,[H_0, \rho]}_{\mathcal{H}_0(\rho)} + \underbrace{\frac{1}{2T_2^*}\big(Z\rho Z - \rho\big)}_{\mathcal{D}(\rho)} \tag{5.12}$$

and $H_0 = \frac{\gamma\omega_0}{2} Z$ is the average Hamiltonian. When there are no spatial noise correlations (i.e., when $c_{ij} = \delta_{ij}$), the overall equation of motion will simply be a Lindblad equation with this $\mathcal{L}$ acting separately on each qubit. For a more general noise correlation matrix $C = (c_{ij})$, though, we expect to get additional terms whose effects are visible in the dephasing rates of multi-qubit states. It turns out, in general, that the average $n$-qubit dynamics can be written as a Lindblad equation $\dot{\rho} = \mathcal{L}(\rho)$ where

$$\mathcal{L}(\rho) = \underbrace{-i\,[H_0, \rho]}_{\mathcal{H}_0(\rho)} + \underbrace{\frac{1}{2T_2^*} \sum_{i,j=1}^{n} c_{ij} \Big( Z_i \rho Z_j - \frac{1}{2}\{Z_i Z_j, \rho\} \Big)}_{\mathcal{D}(\rho)} \tag{5.13}$$

and $H_0 = \frac{\gamma\omega_0}{2} \sum_{i=1}^{n} Z_i$.

*Proof.* To see why this is, note first that a general $n$-qubit state $\rho$ can always be

decomposed as

$$\rho = \sum_{\vec{x}\vec{y}} \varrho_{\vec{x}\vec{y}} \, |\vec{x}\rangle\langle\vec{y}| \,, \tag{5.14}$$

where $|\vec{x}\rangle$ and $|\vec{y}\rangle$ are computational basis states and $\varrho_{\vec{x}\vec{y}} \in \mathbb{C}$ are coefficients. In the following proof, it will be convenient to label computational states not by bit-strings, but rather by their $Z_j$ eigenvalues. For instance, we will temporarily write the state $|01101\rangle$ as $|1,-1,-1,1,-1\rangle$. Accordingly, we will take $\vec{x}, \vec{y} \in \mathbb{R}^n$ to have elements of $\pm 1$.

First, note that

$$\mathcal{H}_0\big(\, |\vec{x}\rangle\langle\vec{y}| \,\big) = \frac{\gamma\omega_0}{2} \sum_{j=1}^{n} (x_j - y_j) \, |\vec{x}\rangle\langle\vec{y}| \,. \tag{5.15}$$

Similarly,

$$\begin{aligned} \mathcal{D}\big(\, |\vec{x}\rangle\langle\vec{y}| \,\big) &= \frac{1}{2T_2^*} \sum_{i,j=1}^{n} c_{ij} \left[ x_i y_j - \frac{1}{2}\big(x_i x_j + y_i y_j\big) \right] |\vec{x}\rangle\langle\vec{y}| \\ &= \frac{1}{2T_2^*} \left[ \vec{x}^{\top} C \vec{y} - \frac{1}{2}\big(\vec{x}^{\top} C \vec{x} + \vec{y}^{\top} C \vec{y}\big) \right] |\vec{x}\rangle\langle\vec{y}| \\ &= -\frac{1}{4T_2^*} (\vec{x} - \vec{y})^{\top} C (\vec{x} - \vec{y}) \; |\vec{x}\rangle\langle\vec{y}| \,, \end{aligned} \tag{5.16}$$

since $C^{\top} = C$. We can therefore write compactly

$$\mathcal{L}\big(\, |\vec{x}\rangle\langle\vec{y}| \,\big) = \underbrace{\left[ -\frac{i\gamma\omega_0}{2} \, \vec{1}^{\top} (\vec{x} - \vec{y}) - \frac{1}{4T_2^*} (\vec{x} - \vec{y})^{\top} C (\vec{x} - \vec{y}) \right]}_{:= \, \beta_{\vec{x}\vec{y}}} |\vec{x}\rangle\langle\vec{y}| \,, \tag{5.17}$$

where $\vec{1} := (1, \ldots, 1)^{\top}$ and $\beta_{\vec{x}\vec{y}} \in \mathbb{C}$. We now carry on from the other direction, so as to meet in the middle. Defining $\mathcal{U}_t = e^{-i \int_0^t \mathcal{H}(\tau) d\tau}$ in terms of the Hamiltonian superoperator $\mathcal{H}(t) = [H(t), \cdot\,]$, each component of $\rho$ evolves in a given run of the

experiment as

$$\begin{aligned}\mathcal{U}_t\big(\,|\vec{x}\rangle\langle\vec{y}|\,\big) &= e^{-i\int_0^t H(\tau)d\tau}\,|\vec{x}\rangle\langle\vec{y}|\,e^{i\int_0^t H(\tau)d\tau} \\ &= \exp\left[-\frac{i\gamma}{2}\sum_{j=1}^{n}\int_0^t \omega_j(\tau)x_j\,d\tau\right]|\vec{x}\rangle\langle\vec{y}|\left[\frac{i\gamma}{2}\sum_{k=1}^{n}\int_0^t \omega_k(\tau)y_k\,d\tau\right] \qquad (5.18)\\ &= \underbrace{\exp\left[-\frac{i\gamma\omega_0 t}{2}\sum_{j=1}^{n}(x_j-y_j)-\frac{i\gamma}{2}\int_0^t\sum_{j=1}^{n}\delta\omega_j(\tau)(x_j-y_j)d\tau\right]}_{:=e^{i\theta}}|\vec{x}\rangle\langle\vec{y}|\end{aligned}$$

We therefore get the average time-evolved version of each component by averaging $e^{i\theta}$ over realizations:

$$\langle e^{i\theta}\rangle = \exp\left[-\frac{i\gamma\omega_0 t}{2}\vec{1}\cdot(\vec{x}-\vec{y})\right]\langle\exp\left[-i\int_0^t f_{\vec{x}\vec{y}}(\tau)d\tau\right]\rangle, \qquad (5.19)$$

where

$$f_{\vec{x}\vec{y}}(\tau) = \frac{\gamma}{2}\sum_{j=1}^{n}\delta\omega_j(\tau)(x_j-y_j) \qquad (5.20)$$

is a stationary Gaussian stochastic process because of our earlier assumptions. We can therefore calculate $\langle e^{i\theta}\rangle$ using Eq. (4.17), substituting $f_{\vec{x}\vec{y}}$ for $\delta\omega$. By inspection $f_{\vec{x}\vec{y}}$ has zero mean, and its autocorrelation function is

$$\begin{aligned}\langle f_{\vec{x}\vec{y}}(t)f_{\vec{x}\vec{y}}(0)\rangle &= \frac{\gamma^2}{4}\sum_{j,k=1}^{n}(x_j-y_j)(x_k-y_k)\,\langle\delta\omega_j(t)\,\delta\omega_k(0)\rangle \\ &= \frac{\gamma^2\Gamma}{2}\sum_{j,k=1}^{n}(x_j-y_j)c_{jk}(x_k-y_k)\,\delta(t) \qquad (5.21)\\ &= \frac{1}{2T_2^*}(\vec{x}-\vec{y})^\top C(\vec{x}-\vec{y})\,\delta(t).\end{aligned}$$

Therefore, components of $\rho$ evolve on average as

$$\begin{aligned}\langle\mathcal{U}_t\rangle\big(\,|\vec{x}\rangle\langle\vec{y}|\,\big) &= \exp\left[-\frac{i\gamma\omega_0 t}{2}\,\vec{1}^\top(\vec{x}-\vec{y})-\frac{t}{4T_2^*}(\vec{x}-\vec{y})^\top C(\vec{x}-\vec{y})\right]|\vec{x}\rangle\langle\vec{y}| \qquad (5.22)\\ &= e^{t\,\beta_{\vec{x}\vec{y}}}\,|\vec{x}\rangle\langle\vec{y}|\,.\end{aligned}$$

It follows immediately from Eqs. (5.17) and (5.22) that the average time-evolved state $\langle \mathcal{U}_t(\rho) \rangle = \langle \mathcal{U}_t \rangle(\rho)$ satisfies

$$\begin{aligned} \frac{d}{dt}\langle \mathcal{U}_t \rangle(\rho) &= \sum_{\vec{x}\vec{y}} \varrho_{\vec{x}\vec{y}} \frac{d}{dt}\langle \mathcal{U}_t \rangle\big( |\vec{x}\rangle\langle\vec{y}| \big) = \sum_{\vec{x}\vec{y}} \varrho_{\vec{x}\vec{y}}\, \beta_{\vec{x},\vec{y}}\, \langle \mathcal{U}_t \rangle\big( |\vec{x}\rangle\langle\vec{y}| \big) \\ &= \sum_{\vec{x}\vec{y}} \varrho_{\vec{x}\vec{y}}\, \mathcal{L}\Big[ \langle \mathcal{U}_t \rangle\big( |\vec{x}\rangle\langle\vec{y}| \big) \Big] = \mathcal{L}\Big[ \langle \mathcal{U}_t \rangle\big(\rho\big) \Big], \end{aligned}$$

for the Liouvillian $\mathcal{L}$ in Eq. (5.13), as claimed. We now return to labeling computational basis states using bit-strings, rather than $\pm 1$'s. $\square$

## 5.3 Exploiting Spatial Noise Correlations

Having expressed the dynamics as a Lindblad equation allows us to use the framework introduced in Section 1.2.3. Specifically, we can remove the cross-terms of $\mathcal{D}$ in Eq. (5.13), thus expressing it in the form of Eq. (5.1)—that is, in terms of jumps $L_i$—by diagonalizing the correlation matrix $C \geq 0$ to yield operators $L_i = \sqrt{\lambda_i}\vec{v}_i \cdot \vec{Z}$. Here, $C\vec{v}_i = \lambda_i \vec{v}_i$ and $\vec{Z} = (Z_1, \ldots, Z_n)$. These $L_i$'s can be interpreted as the sensor's quantum jumps, while the $Z_i$'s in (5.13) cannot [19]. (This means that the no-go example in Section 5.1 with $L_i \propto Z_i$ implicitly made a very strong assumption: that the phase noise was completely devoid of spatial correlations.) One could also think of the $L_i$'s as describing "normal modes" of the noise. Crucially, the ECQS conditions, (5.3) and (5.4), deal with the quantum jump operators $L_i$'s, not the bare $Z_i$'s. This distinction is critical because it opens the possibility of engineering a sensor such that $C$ has a vanishing eigenvalue $\lambda_k = 0$, thus suppressing $L_k$. Generically, the Lindblad span $\mathscr{S}$ will then fail to contain $H_0$ due to this "missing" jump operator, opening the door for ECQS.

The requirement that $H_0 \notin \mathscr{S}$ for conditions (5.3) and (5.4) can be restated for signal and noise which couple identically to the sensor [in the sense of Eq. (5.6)] as

$$\vec{\gamma} \notin \mathrm{col}(C), \tag{5.23}$$

where $\vec{\gamma} := (\gamma, \ldots, \gamma)^\top \in \mathbb{R}^n$ is defined so that $H_0 = \frac{\omega_0}{2}\, \vec{\gamma} \cdot \vec{Z}$, and $\mathrm{col}(C)$ is the column space of $C$. Eq. (5.23) enforces two things: (i) that $\det(C) = 0$ and thus $L_k = 0$ for some $k$, and (ii) that $H_0$ is not composed only of non-vanishing $L_i$'s. It ensures that the signal and noise can be fully distinguished by the recovery operation on the basis of their respective spatial profiles (a requirement we will later relax).

*Proof.* We show here that Eq. (5.23) is equivalent to $H_0 \notin \mathscr{S}$ for the background noise described in Eq. (5.13), where $\mathscr{S}$ is the Lindblad span. Let $\{\vec{v}_i\} \subset \mathbb{R}^n$ be an orthonormal eigenbasis of $C$ such that $C\vec{v}_i = \lambda_i \vec{v}_i$, and define $\tilde{L}_i := \vec{v}_i \cdot \vec{Z} = \tilde{L}_i^\dagger$ and $L_i := \sqrt{\lambda_i}\tilde{L}_i = L_i^\dagger$. Notice that $\langle \tilde{L}_i, I\rangle = \langle \tilde{L}_i, \tilde{L}_j\tilde{L}_\ell\rangle = 0$ under $\langle A, B\rangle = \mathrm{tr}(A^\dagger B)$, so $\mathscr{S}$ can be decomposed into orthogonal subspaces as $\mathscr{S} = \mathscr{S}_1 \oplus \mathscr{S}_2$ where $\mathscr{S}_1 := \mathrm{span}\{L_i\}_{i\geq 1}$ and $\mathscr{S}_2 := \mathrm{span}\{I, L_iL_j\}_{i,j\geq 1}$. Having diagonalized $C$, we can express $H_0$ in terms of $\tilde{L}_i$'s (rather than $Z_i$'s) as $H_0 = \sum_{i=1}^n \alpha_i \tilde{L}_i$ for unique coefficients $\alpha_i = 2^{-n}\,\mathrm{tr}(\tilde{L}_i H_0) = \frac{\omega_0}{2}\,\vec{v}_i \cdot \vec{\gamma}$, implying that $H_0 \perp \mathscr{S}_2$. Therefore, $H_0 \notin \mathscr{S}$ if and only if (iff) $H_0 \notin \mathscr{S}_1$. This happens iff there is a $k$ such that $\lambda_k = L_k = 0$ and $\alpha_k \neq 0$, or equivalently, iff $\vec{v}_k \cdot \vec{\gamma} \neq 0$ for some $\vec{v_k} \in \ker(C)$. Finally, since $C = C^\top$ we have $\mathrm{col}(C) \oplus \ker(C) = \mathbb{R}^n$, and so $H_0 \notin \mathscr{S}$ iff $\vec{\gamma} \notin \mathrm{col}(C)$. □

While we have assumed for simplicity that each qubit has the same transduction parameter $\gamma$ [i.e., that $\vec{\gamma} = (\gamma, \ldots, \gamma)^\top$], the Lindblad span condition $H_0 \notin \mathscr{S}$ is equivalent to Eq. (5.23) more generally. This can be seen by noting that the proof above does not assume $\vec{\gamma}$ to be uniform. Indeed, while the results here and in Chapter 6 do not require a uniform $\vec{\gamma}$, we will generally focus on this case. As we mentioned above, one reason for this choice is simplicity. A more important reason, however, is that in analyzing the impact of distinct $\gamma_j$'s on each qubit, one risks taking the current noise model too seriously. For instance, if a sensor's dynamics were described perfectly by Eqs. (5.6) and (5.11), one would expect that doubling $\gamma_j$ on qubit $j$ would increase its dephasing rate by exactly a factor of 4, as per Eq. (5.8). More broadly, it would mean that whether or not $H_0$ is in the Lindblad span $\mathscr{S}$ is independent of the transduction parameters (assuming they are non-zero). In other words, it would mean that $H_0 \notin \mathscr{S}$ is a potential property of the noise $\{\delta\omega_j\}_j$, not of the qubits' response to this noise.

*Proof.* Suppose that a sensor's dynamics is exactly described by Eqs. (5.6) and (5.11), and that each qubit has its own (potentially distinct) transduction parameter $\gamma_i \neq 0$, leading to an overall average Hamiltonian $H_0' = \frac{\omega_0}{2}\vec{\gamma}'\cdot\vec{Z}$ for $\vec{\gamma}' := (\gamma_1, \ldots, \gamma_n)^\top$. Defining a matrix $W = \gamma^{-1}\mathrm{diag}(\vec{\gamma}')$, we can write $\vec{\gamma}' = W\vec{\gamma}$ for $\vec{\gamma} = (\gamma, \ldots, \gamma)^\top$. This change in transduction parameters would also change the non-unitary part of the sensor's dynamics in a very specific way, in principle. Concretely, the Lindblad coefficients in Eq. (5.13) would not longer be the elements of $C$, but rather, would be those of $C' = WCW$. (Unfortunately the elements $c'_{ij}$ of $C'$ are not bounded in general by $\pm 1$, so they lack a nice interpretation as correlation coefficients.) Observe, however, that $\ker(C') = \{W^{-1}\vec{x} \,|\, \vec{x} \in \ker(C)\}$, so $\vec{\gamma}' \notin \mathrm{col}(C') \iff \vec{\gamma} \notin \mathrm{col}(C)$. □

In practice, of course, there could be other sources of dephasing beyond $\{\delta\omega_j\}$, in which case $C$ would not transform with the $\gamma_j$'s exactly as in the last proof. Such scaling is not of direct interest here; the critical piece for QEC is simply that a dynamics described by Eq. (5.13) with a non-diagonal $C$ can easily arise. (Note, for instance, that the more general case where the qubits have different $T_2^*$'s can readily be treated in the present framework by using an adjusted—and less intuitive—definition of $C$.) In order to avoid pushing our noise model too far, we will therefore avoid questions of how $\mathcal{D}$ in Eq. (5.13) transforms with $\vec{\gamma}$ by focusing on the case where the qubit transduction parameters are all equal. Accordingly, we leave the simple task of adapting some of the calculations below for more general $\vec{\gamma}$'s to the interested reader, depending on how $\mathcal{D}$ changes with $\vec{\gamma}$ in their physical system of interest.

### 5.3.1 Negative Noise Correlations

Consider for example $n = 3$ sensing qubits with uniform negative noise correlations of $c_{ij} = -\varphi/2$ on each pair $(i \neq j)$, forming the correlation matrix

$$C_{\mathrm{neg}} = \begin{pmatrix} 1 & -\varphi/2 & -\varphi/2 \\ -\varphi/2 & 1 & -\varphi/2 \\ -\varphi/2 & -\varphi/2 & 1 \end{pmatrix}, \tag{5.24}$$

where $\varphi \in [0,1]$ encodes the noise correlation strength. (Specifically, $\varphi = 0$ produces vanishing correlations whereas $\varphi = 1$ gives the strongest correlations possible.) The eigenpairs of $C_{\text{neg}}$ are given by

$$\begin{aligned} \lambda_1 &= 1 + \varphi/2 & \vec{v}_1 &= \frac{1}{\sqrt{2}}(1,0,-1)^\top \\ \lambda_2 &= 1 + \varphi/2 & \vec{v}_2 &= \frac{1}{\sqrt{6}}(1,-2,1)^\top \\ \lambda_3 &= 1 - \varphi & \vec{v}_3 &= \frac{1}{\sqrt{3}}(1,1,1)^\top. \end{aligned} \tag{5.25}$$

Notice that the eigenvectors of $C_{\text{neg}}$ do not depend on $\varphi$, and that $\lambda_3$ is subdominant for any $\varphi > 0$, as shown in the left panel of Fig. 5-1. In the limiting case of $\varphi = 1$, we see that $C_{\text{neg}}\vec{\gamma} = \vec{0}$ for $\vec{\gamma} = (\gamma,\gamma,\gamma)$, and so $\vec{\gamma} \notin \text{col}(C_{\text{neg}})$.

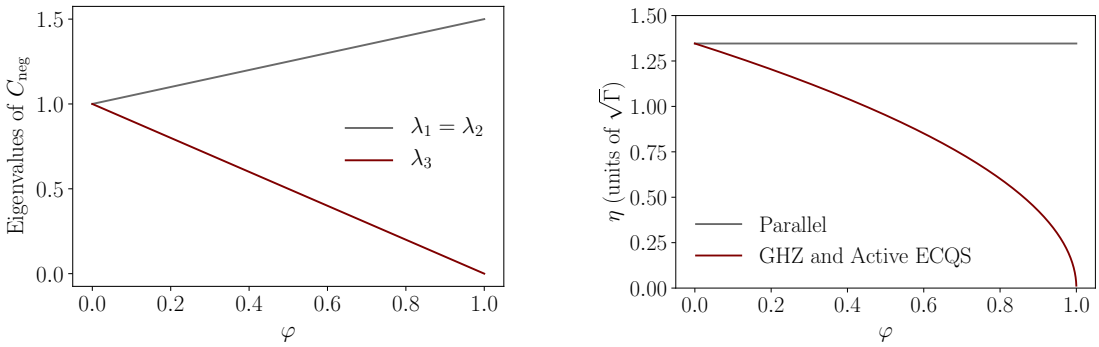



Figure 5-1: Left: The eigenvalues of $C_{\text{neg}}$ as a function of $\varphi$. Right: The achievable sensitivity with different schemes for a 3-qubit sensor under noise correlations $c_{ij} = -\varphi/2$, for all $i \neq j$. The active recovery and GHZ schemes use codewords (5.27) and (5.35) respectively, whereas the parallel scheme operates the qubits individually, without entanglement. The active and GHZ schemes give identical performance (in the regime of $\delta t \to 0$ for the former), and both outperform parallel sensing by a factor that grows with the correlation strength $\varphi$.

The jump operators in the ECQS conditions are not $Z_1$, $Z_2$ and $Z_3$, but rather

$$L_1 = \frac{\sqrt{2+\varphi}}{2}(Z_1 - Z_3), \qquad L_2 = \sqrt{\frac{2+\varphi}{12}}(Z_1 - 2Z_2 + Z_3), \qquad L_3 = \sqrt{\frac{1-\varphi}{3}}(Z_1 + Z_2 + Z_3), \tag{5.26}$$

found by diagonalizing $C_{\text{neg}}$. Observe that the global noise mode, $L_3$, becomes sub-

dominant for larger values of $\varphi$, until it vanishes completely when $\varphi = 1$, at which point $H_0 \notin \mathscr{S}$ as expected. Notice also that $P = |0_{\text{L}}\rangle\langle 0_{\text{L}}| + |1_{\text{L}}\rangle\langle 1_{\text{L}}|$ with logical states

$$|0_{\text{L}}\rangle = \frac{1}{\sqrt{3}}\big(\,|100\rangle + |010\rangle + |001\rangle\,\big) \qquad\qquad |1_{\text{L}}\rangle = \frac{1}{\sqrt{3}}\big(\,|011\rangle + |101\rangle + |110\rangle\,\big) \tag{5.27}$$

satisfies the conditions (5.3) and (5.4) when $\varphi \to 1$, despite the signal and noise both coupling to each sensing qubit identically. Contrast this with the $n = 3$ repetition code, which corrects for $Z_1, Z_2, Z_3$ and all linear combinations thereof, including $H_0$. Eq. (5.27) instead defines a weakened version of the repetition code, which corrects for linear combinations of $L_1$ and $L_2$, but not for $\vec{v}_3 \cdot \vec{Z} \propto Z_1 + Z_2 + Z_3$. In the language of Chapter 1, it forms a targeted "net," carefully chosen to catch the dominant noise but allow the signal to pass through. As per the derivation of the Knill-Laflamme condition in Chapter 1, the code in Eq. (5.27) partitions the sensor's Hilbert space into four orthogonal subspaces, $\mathscr{C}_0$, $\mathscr{C}_1$, $\mathscr{C}_2$ and $\mathscr{C}_3$, spanned by $|0_{\text{L}}\rangle$ and $|1_{\text{L}}\rangle$, and

$$\begin{aligned}
|0_1\rangle &= \frac{1}{\sqrt{2}}\Big(\,|100\rangle - |001\rangle\,\Big) & |1_1\rangle &= \frac{1}{\sqrt{2}}\Big(\,|110\rangle - |011\rangle\,\Big) \\
|0_2\rangle &= \frac{1}{\sqrt{6}}\Big(\,|001\rangle - 2\,|010\rangle + |100\rangle\,\Big) & |1_2\rangle &= \frac{1}{\sqrt{6}}\Big(\,-|011\rangle + 2\,|101\rangle - |110\rangle\,\Big) \\
|0_3\rangle &= |000\rangle & |1_3\rangle &= |111\rangle
\end{aligned} \tag{5.28}$$

respectively. As described in Chapter 1, the recovery associated with this code can be cast in the form of a projective measurement with unitary feedback. It can also be written as

$$\mathcal{R}(\rho) = P\rho P + \sum_{j=1}^{3} K_j \rho K_j^\dagger, \tag{5.29}$$

where $K_j = |0_{\text{L}}\rangle\langle 0_j| + |1_{\text{L}}\rangle\langle 1_j|$. One can show (see, e.g., [152]) that $\mathcal{RL}$ acting on logical states has a Hamiltonian term $\mathcal{RH}_0(\rho_{\text{L}}) = [H_{\text{eff}}, \rho_{\text{L}}]$ with

$$H_{\text{eff}} = PH_0P = \frac{\gamma\omega_0}{2} Z_{\text{L}}, \tag{5.30}$$

where $Z_{\text{L}} = |0_{\text{L}}\rangle\langle 0_{\text{L}}| - |1_{\text{L}}\rangle\langle 1_{\text{L}}|$, and a dissipative term $\mathcal{RD}$ with Lindblad operators

$$\{L_{\text{eff}}^{(i,j)}\}_{i,j} = \{PL_jP,\ \ K_iL_jP\}_{i,j}. \tag{5.31}$$

A direct calculation shows that the only nontrivial Lindblad jump operator (i.e., not proportional to $P$) is

$$L_{\text{eff}} = \sqrt{\frac{1-\varphi}{6T_2^*}}\, Z_{\text{L}}, \tag{5.32}$$

dropping the index labels to simplify the notation. The effective dynamics at the logical level, as $\delta t \to 0$, is therefore generated by $\mathcal{L}_{\text{eff}} = \mathcal{RL}|_{\text{code}}$ where

$$\mathcal{L}_{\text{eff}}(\rho_{\text{L}}) = -i[H_{\text{eff}}, \rho_{\text{L}}] + L_{\text{eff}}\,\rho_{\text{L}}\,L_{\text{eff}} - \frac{1}{2}\left\{L_{\text{eff}}^2, \rho_{\text{L}}\right\}. \tag{5.33}$$

Notice that $L_{\text{eff}} \to 0$ but $H_{\text{eff}}$ remains constant as $\varphi \to 1$, so the logical dynamics becomes less noisy for stronger correlations. In fact, in the limiting $\varphi \to 1$ case this code can fully correct the noise ($\mathcal{RD}|_{\text{code}} = 0$) while allowing the signal to still imprint on the logical states ($\mathcal{RH}_0|_{\text{code}} \neq 0$), resulting in a signal-dependent effective unitary dynamics.

More broadly, notice that for any positive $\varphi$ the code in Eq. (5.27) corrects all but the subdominant noise mode, and could therefore be expected to enhance performance even if $\varphi$ is smaller than 1. As we saw in Chapter 4, the performance of an imperfect quantum sensor is naturally quantified by the sensitivity it achieves; that is, the smallest signal it can detect per unit time. Sensitivity therefore also provides a way to benchmark ECQS schemes: a more effective scheme allows one to resolve a smaller signal per unit time, thus giving lower (i.e., better) sensitivity. Since the limiting logical dynamics in Eq. (5.33) can be mapped to the single-qubit case analyzed in Section 4.2.3, with $A = 1$ and $B = (1-\varphi)/3$, we immediately find that this code offers a sensitivity of

$$\eta_{\text{QEC}} = \sqrt{\frac{1-\varphi}{3}}\,\eta_1, \tag{5.34}$$

for general $\varphi \in [0, 1]$ as shown in the right panel of Fig. 5-1, using an initial state of

$|\psi_0\rangle = (|0_{\text{L}}\rangle + |1_{\text{L}}\rangle)/\sqrt{2}$. Notice that the sensitivity vanishes (i.e., approaches that of an ideal sensor) as $\varphi \to 1$, and outperforms the parallel sensing strategy for all $\varphi > 0$.

Another possible code for this $C_{\text{neg}}$ uses

$$|0'_{\text{L}}\rangle = |000\rangle \qquad |1'_{\text{L}}\rangle = |111\rangle\,. \tag{5.35}$$

When $\varphi \to 1$ it also satisfies the ECQS conditions, although its recovery procedure is trivial because $\text{span}\{|0'_{\text{L}}\rangle\,, |1'_{\text{L}}\rangle\}$ is a decoherence-free subspace (DFS) within which $H_0$ acts non-trivially [164]. (When $\varphi < 1$ it becomes an approximate DFS.) ECQS with this code is therefore a Greenberger-Horne-Zeilinger (GHZ) sensing scheme [165]. An analogous calculation to the one above shows that under $C_{\text{neg}}$, $(|0'_{\text{L}}\rangle + |1'_{\text{L}}\rangle)/\sqrt{2}$ evolves like a single qubit with $A = 3$ and $B = 3(1-\varphi)$. The optimal sensitivity offered by this GHZ/DFS scheme is therefore identical to that in Eq. (5.34) as shown in the right panel of Fig. 5-1, although this coincidence is particular to the present example, and will not occur more broadly (see, e.g., Section 5.3.2).

In general, a DFS is a code for which $|0_{\text{L}}\rangle$ and $|1_{\text{L}}\rangle$ are degenerate eigenvectors of all $L_i$. This means that $EP = \mu_E P$ for coefficients $\mu_E$, for all $E \in \mathscr{S}$, immediately satisfying condition (5.3). For states within the DFS to be sensitive to $\omega_0$ we also need $|0_{\text{L}}\rangle$ and $|1_{\text{L}}\rangle$ to be non-degenerate energy eigenstates, so that $H_0 P \not\propto P$, satisfying condition (5.4). Such a DFS satisfies the ECQS conditions—accordingly, our discussion of general ECQS encompasses DFS-enhanced sensing as a special case. In terms of the Knill-Laflamme condition, a DFS is a QEC code in which the code matrix has a rank of one. DFS-enhanced sensing is only possible for a small family of correlation matrices $C$ which we discuss below. A code designed for some general $C$, in contrast, will usually necessitate an active recovery $\mathcal{R}$. This can be implemented as described in Chapter 1, with details given in Chapter 6. For a logical state $\rho_{\text{L}} = P\rho_{\text{L}}P$ of a code satisfying Eqs. (5.3) and (5.4), such a recovery gives $\mathcal{R}\mathcal{D}(\rho_{\text{L}}) = 0$ and $\mathcal{R}\mathcal{H}_0(\rho_{\text{L}}) = [H_{\text{eff}}, \rho_{\text{L}}] \neq 0$ as desired, where $H_{\text{eff}} = PH_0P$ [152]. In other words, the sensor approaches noiseless evolution by $H_{\text{eff}}$ in the limit of frequent error detection/correction (i.e., $\delta t \to 0$).

### 5.3.2 Positive Noise Correlations

The codes in Section 5.3.1 [namely Eqs. (5.27) and (5.35)] exploit negative noise correlations to enhance sensitivity. However, quantum error correction can also enhance sensitivity under positive noise correlations (or indeed, a mix of positive and negative). Consider for illustration a 3-qubit sensor with positive noise correlations between each pair of qubits, described by

$$C_{\text{pos}} = \begin{pmatrix} 1 & 3\varphi/4 & 3\varphi/4 \\ 3\varphi/4 & 1 & \varphi/8 \\ 3\varphi/4 & \varphi/8 & 1 \end{pmatrix}. \tag{5.36}$$

Here, $\varphi \in [0,1]$ parameterizes the correlation strength in the same way as in Section 5.3.1. Positive noise correlations which are uniform across all qubit pairs yield subdominant noise modes orthogonal to $H_0$, and are therefore not straightforwardly amenable to ECQS. (Gradiometry schemes are a notable exception.) Accordingly, we choose for illustration here $C_{\text{pos}}$ describing non-uniform noise correlations, where the noise between qubits 2 and 3 is more weakly correlated than that between other pairs.

The eigenpairs of $C_{\text{pos}}$ are given by

$$\begin{aligned} \lambda_1 &= 1 - \varphi/8 & \vec{v}_1 &= \frac{1}{\sqrt{2}}(0,1,-1)^\top \\ \lambda_2 &= 1 + 9\varphi/8 & \vec{v}_2 &= \frac{1}{\sqrt{34}}(4,3,3)^\top \\ \lambda_3 &= 1 - \varphi & \vec{v}_3 &= \frac{1}{\sqrt{17}}(-3,2,2)^\top. \end{aligned} \tag{5.37}$$

As with $C_{\text{neg}}$, the eigenvectors above are independent of $\varphi$. Furthermore, the eigenvalue $\lambda_3$ is subdominant for all $\varphi > 0$, and vanishes in the limit of strong noise correlations ($\varphi \to 1$), as shown in the left panel of Fig. 5-2. In this limit $\vec{\gamma} \notin \text{col}(C_{\text{pos}})$. Note that $C_{\text{pos}}$ does not have a DFS when $\varphi \to 1$, unlike $C_{\text{neg}}$, so we are limited to ECQS schemes with active recoveries for this noise profile.

One can show that

$$|0_{\text{L}}\rangle = \sqrt{x}\Big(\,|001\rangle + |010\rangle\,\Big) + \sqrt{-\frac{7x}{3}+\frac{5}{6}}\,|100\rangle + \sqrt{\frac{x}{3}+\frac{1}{6}}\,|111\rangle \tag{5.38}$$
$$|1_{\text{L}}\rangle = \sqrt{\frac{x}{3}+\frac{1}{6}}\,|000\rangle + \sqrt{-\frac{7x}{3}+\frac{5}{6}}\,|011\rangle + \sqrt{x}\Big(\,|101\rangle + |110\rangle\,\Big)$$

is a valid ECQS code for any $x \in [0, 5/14] \setminus \{1/4\}$, in that it satisfies the Knill-Laflamme condition for $L_j = \sqrt{\lambda_j}\vec{v}_j \cdot \vec{Z}$, $j = 1, 2$, but not for $H_0$. (The free parameter $x$ could be chosen, for instance, so that the resulting code yields the simplest possible recovery $\mathcal{R}$.) As in Section 5.3.1, this code defines four 2-dimensional subspaces spanned here by the codewords and by

$$|0_1\rangle = \frac{1}{\sqrt{2}}\Big(\,|010\rangle - |001\rangle\,\Big) \qquad\qquad |1_1\rangle = \frac{1}{\sqrt{2}}\Big(\,|110\rangle - |101\rangle\,\Big),$$

$$|0_2\rangle = \frac{2}{\sqrt{5+14x}}\Big[\sqrt{x}\Big(\,|001\rangle + |010\rangle\,\Big) + \sqrt{\frac{5}{24}-\frac{7x}{12}}\,|100\rangle - \frac{5}{2}\sqrt{\frac{2x+1}{6}}\,|111\rangle\,\Big]$$
$$|1_2\rangle = \frac{2}{\sqrt{5+14x}}\Big[\frac{5}{2}\sqrt{\frac{2x+1}{6}}\,|000\rangle - \sqrt{\frac{5}{24}-\frac{7x}{12}}\,|011\rangle - \sqrt{x}\Big(\,|101\rangle + |110\rangle\,\Big)\Big] \tag{5.39}$$

$$|0_3\rangle = \frac{1}{\sqrt{42x+15}}\Bigg[\sqrt{\frac{15}{2}-6x(1+7x)}\big(\,|001\rangle + |010\rangle\,\big) - 7\sqrt{x(1+2x)}\,|100\rangle + \sqrt{x(5-14x)}\,|111\rangle\Bigg]$$
$$|1_3\rangle = \frac{1}{\sqrt{42x+15}}\Bigg[-\sqrt{x(5-14x)}\,|000\rangle + 7\sqrt{x(1+2x)}\,|011\rangle - \sqrt{\frac{15}{2}-6x(1+7x)}\big(\,|101\rangle + |110\rangle\,\big)\Bigg].$$

One can use these states to construct the Kraus operators of $\mathcal{R}$, as given in Eq. (5.29),

from which one finds

$$H_{\text{eff}} = \left(\frac{1-4x}{3}\right)\frac{\gamma\omega_0}{2}Z_{\text{L}}, \tag{5.40}$$

and a single non-trivial jump operator

$$L_{\text{eff}} = \sqrt{\left(\frac{17(1-4x)^2(1-\varphi)}{9}\right)\frac{1}{2T_2^*}}\; Z_L \tag{5.41}$$

at the logical level when $\delta t \to 0$. Comparing with Section 4.2.3, one can identify for this family of codes

$$A = \frac{1-4x}{3} \qquad\qquad B = \frac{17(1-4x)^2(1-\varphi)}{9}, \tag{5.42}$$

giving a sensitivity bound of

$$\eta_{\text{QEC}} = \sqrt{17(1-\varphi)}\,\eta_1 \tag{5.43}$$

for an initial state of $(|0_{\text{L}}\rangle + |1_{\text{L}}\rangle)/\sqrt{2}$, independent of $x$, as shown in the right panel of Fig. 5-2. For comparison, the GHZ scheme gives

$$\eta_{\text{GHZ}} = \frac{\sqrt{12+13\varphi}}{6}\,\eta_1 \tag{5.44}$$

for $C_{\text{pos}}$, which is worse than the parallel scheme for all $\varphi > 0$. Like in the previous example, the sensitivity offered by ECQS approaches the ideal case ($\eta_{\text{QEC}} \to 0$) as $\varphi \to 1$. Unlike the previous example, however, the codes presented here only outperform simpler schemes for very strong noise correlations $\varphi = 50/51 \approx 0.98$. It is unclear at present whether this high threshold is due to the choice of code, the spatial noise profile $C_{\text{pos}}$ itself, or both.

### 5.3.3 Robustness Analysis

Sections 5.3.1 and 5.3.2 examined the effects of the noise correlation strength ($\varphi$) on code performance. Before moving on from these two examples, let us consider

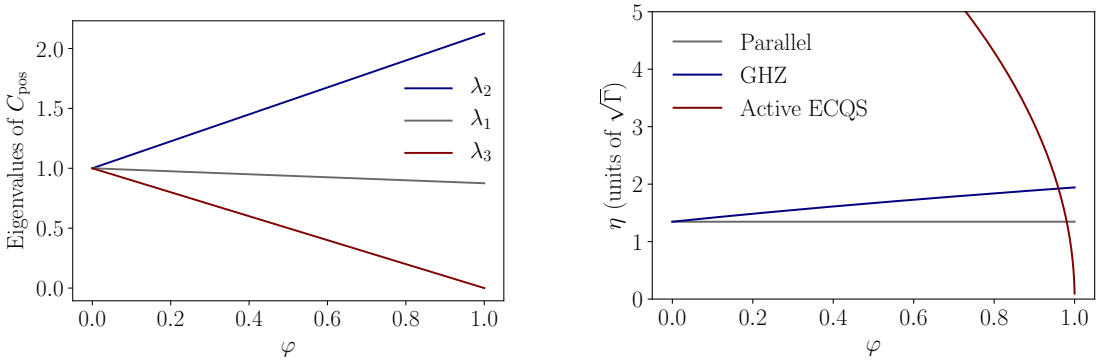



Figure 5-2: Left: The eigenvalues of $C_{\text{pos}}$ as a function of $\varphi$. Right: The optimal sensitivity achievable through ECQS with codewords (5.38), compared with that from parallel and GHZ schemes.

the question of robustness against uncertainty/errors in the measured $C$. That is, if the true noise correlation profile is given by $C$, but experimental imperfections lead one to measure $C' \neq C$ and design a code accordingly, how much will the sensor's performance suffer?

To address this question in the context of the two previous examples, we consider the family of correlation profiles

$$C_\phi = \begin{pmatrix} 1 & \phi & \phi \\ \phi & 1 & 2\phi^2 - 1 \\ \phi & 2\phi^2 - 1 & 1 \end{pmatrix}. \tag{5.45}$$

Here, $\phi \in [-1, 1]$ does not describe the strength of the correlations. Rather, it parameterizes a family of strongly-correlated noise profiles. In particular, $\phi = -1/2$ and $\phi = 3/4$ give $C_{\text{neg}}$ and $C_{\text{pos}}$ respectively with $\varphi = 1$.

The eigenpairs of $C_\phi$ are given by

$$\begin{aligned} \lambda_1 &= 2\phi^2 + 1 & \vec{v}_1 &= \frac{1}{\sqrt{2\phi^2+1}}(1, \phi, \phi)^\top \\ \lambda_2 &= 2(1 - \phi^2) & \vec{v}_2 &= \frac{1}{\sqrt{2}}(0, 1, -1)^\top \\ \lambda_3 &= 0 & \vec{v}_3 &= \frac{1}{\sqrt{4\phi^2+2}}(2\phi, -1, -1)^\top. \end{aligned} \tag{5.46}$$

Unlike in previous sections, the subdominant eigenvalue $\lambda_3$ is identically zero here, as shown in the left panel of Fig. 5-3. However, the eigenvectors—i.e., the "shapes" of the noise modes—depend on $\phi$.

To analyze the robustness of the codes from previous sections, we apply them to noise with a spatial profile $C_\phi$, where $\phi$ is not necessarily the value for which they were designed. The codewords from Eq. (5.27) (designed for use with $C_{\text{neg}}$, i.e., $\phi = -1/2$) give

$$H_{\text{eff}} = \frac{\gamma\omega_0}{2} Z_{\text{L}} \qquad\qquad L_{\text{eff}} = \sqrt{\left(\frac{2\phi+1}{3}\right)^2 \frac{1}{2T_2^*}}\, Z_{\text{L}} \tag{5.47}$$

in the presence of noise with correlations $C_\phi$. Those from Eq. (5.38) (designed for use with $C_{\text{pos}}$, i.e., $\phi = 3/4$) give

$$H_{\text{eff}} = \left(\frac{1-4x}{3}\right) \frac{\gamma\omega_0}{2} Z_{\text{L}} \qquad\qquad L_{\text{eff}} = \sqrt{\left(\frac{(4x-1)(4\phi-3)}{3}\right)^2 \frac{1}{2T_2^*}}\, Z_{\text{L}}, \tag{5.48}$$

and thus a sensitivity independent of $x$. Comparing with Section 4.2.3, the factors $A$ and $B$, as well as the achievable sensitivities for these codes, are listed in the table below as functions of $\phi$. The sensitivities are also plotted in the right panel of Fig. 5-3.

| Scheme | $A$ | $B$ | $\eta$ |
|---|---|---|---|
| Parallel | - | - | $\sqrt{\frac{1}{3}}\,\eta_1$ |
| GHZ | 3 | $(2\phi+1)^2$ | $\left\lvert\frac{2\phi+1}{3}\right\rvert\,\eta_1$ |
| Active ECQS, Eq. (5.27) | 1 | $\left(\frac{2\phi+1}{3}\right)^2$ | $\left\lvert\frac{2\phi+1}{3}\right\rvert\,\eta_1$ |
| Active ECQS, Eq. (5.38) | $\frac{1-4x}{3}$ | $\left(\frac{(4x-1)(4\phi-3)}{3}\right)^2$ | $\lvert 4\phi-3\rvert\,\eta_1$ |

Some remarks are in order: (i) The active recovery using (5.27) performs identically to the GHZ scheme here, much like in Section 5.3.1. (ii) The performance of the codes in (5.38) is independent of $x$, as in Section 5.3.2. (iii) All codes perform best at the value of $\phi$ for which they were designed, and their performance deteriorates gradually away from this intended value. For reasons that are not clear at present, the code in (5.27) is more robust in $\phi$ than that in (5.38).

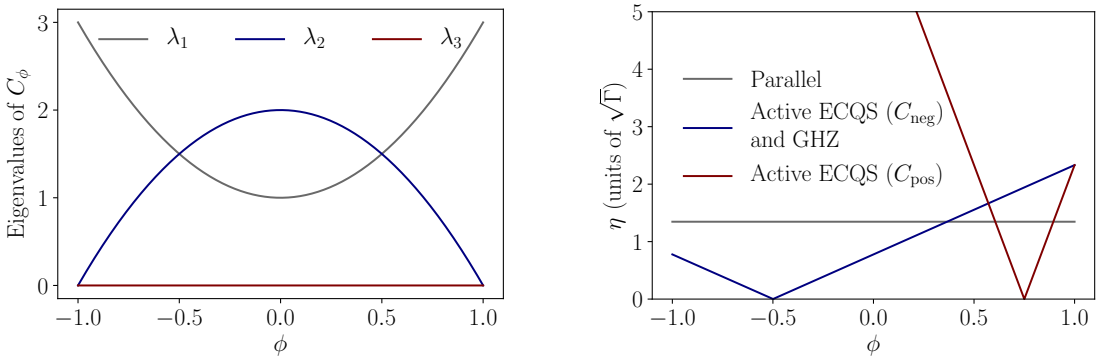



Figure 5-3: Left: The eigenvalues of $C_\phi$ as a function of $\phi$. Right: The optimal sensitivity offered by codes (5.27), (5.35), and (5.38) for arbitrary values of $\phi$, compared with a parallel sensing scheme.

## 5.4 Range of Applicability

Conditions (5.3) and (5.4) seem to impose a stringent requirement on the $C$'s amenable to ECQS under signal and noise which are "parallel," in that they couple to the sensor through the same operators (along the $z$ direction, i.e., the qubits' energy gaps, here). This need not be the case, however, since error correction can enhance quantum sensing even if it does not give a strictly noiseless limit. Notice in the right panels of Figs. 5-1 and 5-2, for instance, that ECQS can enhance sensitivity for $\varphi < 1$, even though conditions (5.3) and (5.4) are only satisfied exactly when $\varphi = 1$. More generally, if instead of satisfying (5.3) exactly, $PL_i^\dagger L_j P = m_{ij} P + O(\epsilon)$ (defining as shorthand $L_0 := I$), then $\mathcal{RD}(\rho_{\text{L}}) = O(\epsilon)$ for a logical $\rho_{\text{L}}$ instead of vanishing exactly [152]. If the time between successive recoveries is nonzero ($\delta t > 0$) as in most experiments, then decoherence will appear in the logical dynamics at order[1] $O(\delta t/T_2^*)$ in $\mathcal{L}_{\text{eff}}$. Provided $\epsilon \ll \delta t/T_2^*$, then, small violations of condition (5.3) will not appreciably change the degree to which quantum error correction suppresses noise in a sensor. This is true for generic ECQS schemes. When correcting noise which couples like the signal, in particular, allowing $\epsilon \neq 0$ enables codes which—by design—do not correct for errors $L_k$ with $||L_k|| \approx 0$, corresponding to an eigenvalue $\lambda_k$ of $C$ which is small but not exactly zero. Therefore in the present setting, relaxing condition (5.3)

[1] We write $O(t/T_2^*)$ rather than $O(||\mathcal{L}||\,\delta t)$ to simplify notation, assuming $\omega_0$ to be a small compared to $1/T_2^*$. If $\omega_0$ is not small, $O(\delta t/T_2^*)$ should be taken to mean $O(||\mathcal{L}||\,\delta t)$.

reduces the need for fine-tuned noise correlations.

For quantum error correction to filter noise from signal when $\delta t/T_2^*$ is finite, $\mathcal{R}$ must suppress the former more than the latter. Choosing $|0_{\text{L}}\rangle$ and $|1_{\text{L}}\rangle$ to be eigenstates of $H_{\text{eff}}$ with energies $E_0 > E_1$, the effective Hamiltonian takes the form $H_{\text{eff}} = \alpha P + \frac{A\gamma\omega_0}{2} Z_{\text{L}}$, where $P$ acts as $I$ over the codespace (i.e., trivially). Just as $\epsilon$ and $\delta t/T_2^*$ describe the extent to which noise is suppressed through frequent error correction, $A$ (cf. Section 4.2.3) describes the signal gain; that is, the fraction of the physical signal that survives at the logical level. Together with previous arguments about (5.3), we arrive at sufficient conditions in terms of this signal gain for error correction to enhance quantum sensing: $PL_i^\dagger L_j P = m_{ij} P + O(\epsilon)$, and $A \gg \epsilon$. We will make this more precise in Chapter 6. There is an analogy with dynamical decoupling to be drawn here: both quantum error correction and DD can significantly enhance sensing by partially filtering noise from the signal—they need not remove the noise entirely to be useful.

This analogy goes further: Just as DD sequences must be tailored to sense in a particular frequency band of interest, error-correcting codes must be tailored to $C$ and $\vec{\gamma}$ for a sensor to measure only in a particular spatial "mode," in order to correct for noise which couples locally in the same way as the signal. That is, a particular $\delta\omega$ and arrangement of sensing qubits (likely "baked in" to a device) will require a unique $P$. This is because the scheme depends on a code *not* correcting for $\vec{v}_k \cdot \vec{Z}$ with $\lambda_k \sim 0$, thus allowing the component of $H_0$ along $\vec{v}_k \cdot \vec{Z}$ to affect the logical states. Therefore, we expect that in experiment, codes will need to be tailored for individual devices, much like control sequences must be. We present here an initial method of doing so.

### 5.4.1 Numerical Code Search

When $\vec{\gamma} \notin \text{col}(C)$, Refs. [151, 152] provide recipes for codes which exactly satisfy the ECQS conditions, although these may require the sensor to contain up to $n$ noiseless ancilla qubits which do not couple to $\omega$, in addition to the $n$ sensing qubits. Here we take a different approach which naturally tolerates small violations of the ECQS

conditions, such as those discussed above. That is, it allows one to find codes which enhance sensitivity, irrespective of whether they provide a noiseless limit in theory. In contrast with several previous works [151, 152], our approach does not require the overhead of additional ancillas as part of the code. Specifically, for a given $C$, we map the task of finding a $P$ for (5.3) and (5.4) to an optimization problem, whose solutions are codes satisfying $PL_i^\dagger L_j P = m_{ij}P + O(\epsilon)$ for some code matrix $M = M^\dagger$, and giving a minimum signal gain of $A_{\min}$ that is freely adjustable:

$$\text{Minimize } F_{\text{tot}} = \sum_{E \in \mathscr{S}} F_E \text{ subject to } F_G > A_{\min}^2 \text{ and } \langle x|y\rangle = \delta_{xy}, \tag{5.49}$$

where $G := \frac{1}{2}\vec{\gamma}\cdot\vec{Z} = H_0/\omega_0$ and

$$F_E\big(\,|x\rangle\,,|y\rangle\,\big) = \big|\,\langle x|\,E\,|x\rangle - \langle y|\,E\,|y\rangle\,\big|^2 + 4\big|\,\langle x|\,E\,|y\rangle\,\big|^2. \tag{5.50}$$

Notice that $F_{\text{tot}}$ is non-negative with zeros where $P = |x\rangle\langle x| + |y\rangle\langle y|$ satisfies condition (5.3). In fact, solutions to $F_{\text{tot}} = 0$ with $A_{\min} = 0$ exactly satisfy the ECQS conditions and vice versa. Relaxing these conditions slightly, one can find codes approximately satisfying (5.3) and (5.4) by using $F_{\text{tot}} \leq \epsilon^2$ as a convergence criterion and $A_{\min} \gg \epsilon$. Note that the resulting codes can be quite general; for instance, they need not be stabilizer codes.

The objective function $F_{\text{tot}}$ may have several distinct zeros satisfying the constraints with $A_{\min} = 0$; for instance, the logical states in Eq. (5.27) and (5.35). [In other words, there can be more than one $P$ exactly satisfying Eqs. (5.3) and (5.4).] On the other hand, it will have no such zeros when $\vec{\gamma} \in \text{col}(C)$. More generally, for given $\epsilon$ and $A_{\min} \geq 0$, there may exist multiple regions where $F_{\text{tot}} \leq \epsilon^2$ subject to the constraints, or there may exist none for $C$ not amenable to ECQS with $n$ sensing qubits.

A similar approach to finding codes was recently used in Ref. [80], although to our knowledge it has not previously been used for ECQS. The factor of 4 in Eq. (5.50) was included specifically for the purpose of finding codes for sensing: While this factor is

irrelevant for enforcing that $E \in \mathscr{S}$ be corrected, a simple calculation shows that $F_G$ for a code gives exactly its signal gain squared. Therefore, requiring that $F_G > A_{\min}^2$ and $F_{\text{tot}} \leq \epsilon^2$ for some $A_{\min} \gg \epsilon$ is a transparent way of demanding that quantum error correction suppress noise much more strongly than the signal.

For a two-qubit sensor, $C$ has a zero eigenvalue only when $c_{12} = \pm 1$. The $c_{12} = 1$ case has $\vec{\gamma} \in \text{col}(C)$ and is therefore not amenable to ECQS. The $c_{12} = -1$ case, on the other hand, has $\vec{\gamma} \notin \text{col}(C)$. Therefore, $n = 2$ sensing qubits under strongly anti-correlated noise ($c_{12} \approx -1$) can benefit from ECQS—in fact, they can be used for DFS-enhanced sensing. ($c_{12} \approx 1$, however, could be useful for gradiometry, i.e., to measure a mean difference between the energy gap of each qubit.) For a three-qubit sensor a much broader family of $C$'s can satisfy the ECQS conditions. Using the mapping described above, the $C$'s for which Eq. (5.49) yielded codes approximately satisfying conditions (5.3) and (5.4) with no ancillas are shown in Fig. 5-4. Notice that DFS-enhanced sensing with $n = 3$ qubits is only possible for a small family of $C$'s. Therefore, while such schemes may be powerful [166–169], it could be exceedingly difficult to engineer the spatial noise correlations they require in many devices. Codes with active recoveries, in contrast, are much more broadly applicable. Chapter 6 will be devoted to finding closed-form QEC codes like those in Eqs. (5.27), (5.35) and (5.38), as a function of $C$ more generally. This will remove the need to rely on numerics to find ECQS protocols as we have done here, and will allow us to analyze the potential of QEC for phase noise filtering in quantum sensors with relative ease.

In an experiment, finding an appropriate code for ECQS first requires knowledge of the noise correlations encoded in $C$. For qubits $i$ and $j$, the coefficient $c_{ij}$ can be inferred by preparing the Bell state $\frac{1}{\sqrt{2}}(|0_i\rangle |0_j\rangle + |1_i\rangle |1_j\rangle)$ and measuring its pure dephasing rate $\tau_{ij}^{-1}$, which is related to $c_{ij}$ through $\tau_{ij}^{-1} = \frac{2}{T_2^*}(1+c_{ij})$, after subtracting the dephasing due to any relaxation that might also be present in practice. Note that the spatial noise correlations in our simple model can be described by a single matrix $C$ because we have assumed a lack of temporal noise correlations. More generally, these spatial correlations will be encoded in cross-correlation functions/cross-spectra, for which one would need to adapt the idealized framework used here.

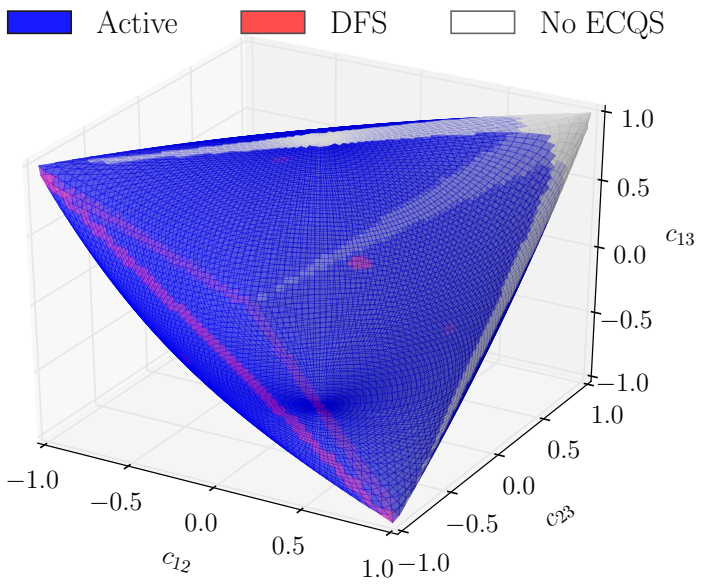



Figure 5-4: Values of $c_{12}$, $c_{23}$ and $c_{13}$ for which there exists a three-qubit code satisfying the ECQS conditions to a tolerance of $\epsilon = 10^{-5}$ and $A_{\min} = 10^{-1}$, and not requiring noiseless ancillas. The points $(c_{12}, c_{23}, c_{13})$ approximately satisfying conditions (5.3) and (5.4) form a tetrahedron-like surface. The portions of the surface in blue denote $C$'s for which ECQS is possible with an active (i.e., non-trivial) recovery. The red regions (enlarged for visibility) denote $C$'s for which DFS-enhanced sensing is possible, and the white regions denote $C$'s for which the optimization in Eq. (5.49) failed to converge to within the specified tolerance, either because the achievable signal gain is too small, or because of poor local minima in $F_{\text{tot}}$. The continuous red band comprises $C$'s for which noise on a pair of qubits is perfectly anti-correlated ($c_{ij} = -1$). Notice that ECQS is generically possible for both positive and negative noise correlations; it fails here only when $c_{ij} \approx 1$ for some pair of qubits ($i \neq j$), since this gives a missing/subdominant jump operator orthogonal to $H_0$.

## 5.5 Discussion

We have shown how error-corrected quantum sensing can filter noise from a signal when both couple to a sensor locally through the same operators. This stands in contrast with earlier explicit ECQS schemes, which were limited to correcting noise separate from the quantity to be measured, in that it couples differently to the sensor. In many quantum sensors such noise is sub-dominant, while the type of noise considered here is the limiting source of decoherence, and can only be partially filtered through DD. Our scheme relies on the observation that Eqs. (5.3) and (5.4) can be viewed as a condition on the spatial correlations of the signal and noise. This view

raises a close parallel between ECQS and DD: for signal and noise which couple identically to a sensor [in the sense of Eq. (5.6)], ECQS and DD can enhance sensitivity by acting as filters in the spatial and frequency domains, respectively. However, since these two schemes separate noise from signal on totally separate grounds, they are complementary, in that the limitations of one are not shared by the other. Finally, we proposed a numerical method of tailoring ECQS codes to the noise observed in specific devices. It yields not just codes that correct noise perfectly in the $\delta t \to 0$ limit, but also codes which can more generally improve sensitivity. We applied this method to sensors comprising 2 and 3 qubits—with no extra ancillas in the code—and showed how our error correction scheme could provide an advantage even in relatively small devices. We will introduce more sophisticated methods to this end in the following chapter.

Both ECQS and DD filter noise which couples locally like the signal by exploiting correlations in it: spatial correlations in the case of ECQS, and temporal correlations for DD. Accordingly, the effectiveness of both schemes depends on the degree to which noise in a sensor can made to have suitable correlations. (Note that different spatial noise correlations may lend themselves best to different sensing tasks. For instance, uniform positive correlations yield a dominant noise mode proportional to $H_0$. This makes them ill-suited for measuring small field values through ECQS, but well-suited for gradiometry.) Engineering appropriate spatial noise correlations is likely to be highly implementation-dependent, as it is with temporal correlations [132]. This is because the main sources of noise can be entirely different in different types of quantum sensors. While an analysis of achievable noise correlations in various types of sensors is beyond the scope of the present work, we note here simply that strong spatial correlations have been reported already in several experiments, e.g., Refs. [93, 166, 167, 170–172]. They also underlie the device-adapted QEC codes in Chapter 2.

The scheme presented here exploits spatial correlations to extract signal from a noisy background, which generically causes dephasing. (Of course, the signal and noise in question need not couple to the sensor via $Z$; in general, the qubits could

be decohered along any axis.) A similar approach may be possible with generalized amplitude-damping ($T_1$-type) errors, which are second only to phase errors as the dominant decoherence mode in many quantum sensors. That is, a sensor with qubits made to thermalize collectively, rather than individually, could be amenable to quantum error correction. The approach presented here could be combined with such a scheme—or with previous ECQS schemes—raising the intriguing prospect of a quantum sensor that is error-corrected against noise in all three spatial directions. While this is the norm in QEC proposals for quantum computing, it might have seemed *a priori* impossible for quantum sensing were it not for the results presented here.

# Chapter 6

# QEC Codes for Dephasing in Quantum Sensors

As we saw in the last chapter, quantum error correction (QEC) has the potential to enhance quantum sensing. This application, however, poses unique challenges for QEC. The main issue is that sensitivity is often limited by noise which acts on a sensor in a very similar way as the signal. In particular, when the signal imprints on the sensor as a relative phase, the dominant type of decoherence is often dephasing. This makes it difficult for QEC to suppress the latter without also suppressing the former by a corresponding amount. For most other applications, the goal of QEC is to effectively decouple a quantum system from its environment (while still allowing for certain control operations, of course). For sensing applications, however, the goal is fundamentally different: one wants to effectively decouple a sensor only from the parts of the environment that cause decoherence, while leaving it sensitive to those one wishes to measure. In other words, QEC should not act as an effective barrier between the system and its environment for quantum sensing, but rather, as a filter.

Fortunately, as we saw in Chapter 5, spatial correlations in background noise provide a means by which QEC could distinguish signal from noise, and preferentially suppress the latter. Specifically, the Lindblad jump operators arising from phase noise are only $Z_j$ phase flips in the limit where the noise has no spatial correlations. In the presence of spatial noise correlations, these jumps are instead linear combinations

of $Z_j$'s—analogous to normal modes—with unequal amplitudes. Remarkably, this is true even when the noise couples locally to each qubit, and is identically distributed on each. Such details are largely irrelevant for conventional QEC; for instance, repetition codes correct such noise to the same order in time regardless of its spatial correlations. For quantum sensing, however, this correlation-induced structure is critical.

There is a natural analogy between QEC and dynamical decoupling (DD) in this setting. DD can enhance sensitivity by allowing only signal and noise within a narrow frequency band to affect the sensor (roughly). Provided there is not too much noise power in the pass-band, this strategy can provide a substantial enhancement in sensitivity. Similarly, QEC could allow only signal and noise with a particular spatial structure to imprint at the logical level. That is, it could correct all but a subdominant normal mode of the noise, thus reducing dephasing without equally suppressing the signal. Like in DD, the net effect could be a substantial enhancement in sensitivity.

In Chapter 5 we identified this potential strategy, and analyzed some simple examples. We did not, however, take on the more difficult problem of finding suitable closed-form QEC codes (which do not require a large number of ancillas) beyond those specific examples. We will do so in this chapter. In particular, we will find application-adapted QEC codes as a function of the spatial noise correlations in quantum sensors. This will allow us to find a simple expression for the sensitivity offered by our scheme. To do so, we will first consider the ideal case of maximally strong spatial noise correlations (wherein the weakest normal mode of the noise vanishes entirely) to develop candidate codes, then generalize our results for realistic noise.

## 6.1 Closed-Form Codes

We will take two different approaches to developing QEC codes for sensing applications in this chapter. The first, covered in this section, involves transforming known QEC codes. The second, taken in Section 6.2, will instead involve developing efficient techniques to search for QEC codes numerically. Ultimately both techniques will be

successful, and will converge to produce the same codes.

We will use the same noise model as in Chapter 5, encompassed in Eq. (5.13), as a starting point for developing QEC codes. Recall that the associated Lindblad jump operators are

$$L_j = \sqrt{\lambda_j}\vec{v}_j \cdot \vec{Z} \tag{6.1}$$

where $C = (c_{jk})$ is the noise correlation matrix and $C\vec{v}_j = \lambda_j \vec{v}_j$, where $\{\vec{v}_j\}$ are chosen to be real and orthonormal, and $\vec{Z} = (Z_1, \ldots, Z_n)$. In this same notation

$$H_0 = \frac{\omega_0}{2}\vec{\gamma} \cdot \vec{Z} \tag{6.2}$$

for some appropriate $\vec{\gamma} \in \mathbb{R}^n$ encoding each qubit's transduction parameter. We will begin by assuming that $H_0$ is outside the Lindblad span $\mathscr{S}$, and construct codes initially for this limiting case. That is, we assume $\vec{\gamma} \notin \mathrm{col}(C)$, meaning that $C$ has at least one vanishing eigenvalue $\lambda_u = 0$ associated with an eigenvector $\vec{v}_u$ which is not orthogonal to $\vec{\gamma}$. (More precisely, if there are several vanishing eigenvalues, we assume $\vec{\gamma}$ has a non-trivial component in the span of the corresponding eigenvectors.) We use the index $u$ to mean "uncorrected," because we will seek to design QEC codes which cannot correct for $\vec{v}_u \cdot \vec{Z}$ by design, and in fact need not do so because the corresponding Lindblad jump operator $L_u = \sqrt{\lambda_u}\vec{v}_u \cdot \vec{Z} = 0$ vanishes.

### 6.1.1 Transforming the Dicke Code

Mathematically, the task at hand is to construct a QEC code with a single logical qubit such that the ECQS conditions in Eqs. (5.3) and (5.4) are satisfied for all $i, j \neq u$. (Equivalently, we can enforce these equations for all $i$ and $j$ since $L_u = 0$.) The resulting logical states should be functions of the correlation matrix $C$ and the transduction parameters $\vec{\gamma}$.

The first breakthrough to this end was due to Zhou [173], who generalized the code in Eq. (5.27). The codewords used there are sometimes known as Dicke states, and they can be viewed as describing a single excitation in the case of $|0_{\mathrm{L}}\rangle$, or a

single "hole," i.e., lack of excitation, in the case of $|1_{\text{L}}\rangle$ [174]. We will therefore informally refer to it as the Dicke code. Zhou's insight was to generalize this code by constructing codewords with one excitation/hole shared between $n$ qubits more generally, with weights/coefficients that depend on $C$ and $\vec{\gamma}$.

As per the previous chapter, the task of designing codes for this sensing application becomes non-trivial when dealing with $n \geq 3$ sensing qubits. For $n \geq 3$ then, the idea is to use as codewords

$$\begin{aligned} |0_{\text{L}}\rangle &= s_1 |100\cdots0\rangle + s_2 |010\cdots0\rangle + \cdots + s_n |000\cdots1\rangle \\ |1_{\text{L}}\rangle &= s_1 |011\cdots1\rangle + s_2 |101\cdots1\rangle + \cdots + s_n |111\cdots0\rangle = X^{\otimes n} |0_{\text{L}}\rangle , \end{aligned} \tag{6.3}$$

where $\vec{s} = (s_1, \ldots, s_n)^\top$ are coefficients to be determined. Notice first that $|0_{\text{L}}\rangle$ and $|1_{\text{L}}\rangle$ are orthogonal, and that for any vectors $\vec{x}, \vec{y} \in \mathbb{R}^n$

$$\langle 0_{\text{L}}| (\vec{x} \cdot \vec{Z})(\vec{y} \cdot \vec{Z}) |1_{\text{L}}\rangle = 0 \tag{6.4}$$

by inspection. Moreover, a direct calculation immediately reveals that

$$\langle 0_{\text{L}}| (\vec{x} \cdot \vec{Z})(\vec{y} \cdot \vec{Z}) |0_{\text{L}}\rangle = \langle 1_{\text{L}}| (\vec{x} \cdot \vec{Z})(\vec{y} \cdot \vec{Z}) |1_{\text{L}}\rangle . \tag{6.5}$$

Therefore, the ansatz in Eq. (6.3) automatically gives $P(\vec{x} \cdot \vec{Z})(\vec{y} \cdot \vec{Z})P \propto P$ for $P = |0_{\text{L}}\rangle\langle 0_{\text{L}}| + |1_{\text{L}}\rangle\langle 1_{\text{L}}|$, for any choice of coefficients $\vec{s}$. Similarly, this ansatz also gives

$$\langle 0_{\text{L}}| (\vec{x} \cdot \vec{Z}) |1_{\text{L}}\rangle = 0 \tag{6.6}$$

for any real $\vec{x}$. To correct the effects of $\{L_j\}_{j \neq u}$ to leading order in time, it remains only to enforce that

$$\langle 0_{\text{L}}| (\vec{x} \cdot \vec{Z}) |0_{\text{L}}\rangle = \langle 1_{\text{L}}| (\vec{x} \cdot \vec{Z}) |1_{\text{L}}\rangle \tag{6.7}$$

through an appropriate choice of coefficients $\vec{s}$. Indeed, a simple calculation shows

that

$$\langle 0_{\text{L}}| \left(\vec{x} \cdot \vec{Z}\right) |0_{\text{L}}\rangle = \sum_{j=1}^{n} x_j \big(1 - 2|s_j|^2\big) = - \langle 1_{\text{L}}| \left(\vec{x} \cdot \vec{Z}\right) |1_{\text{L}}\rangle \tag{6.8}$$

for this ansatz. In order for Eq. (6.7) to hold, the left- and right-hand sides must therefore both vanish; that is, we need the vector $\vec{r}$ with elements $r_j := (1 - 2|s_j|^2)$ to be orthogonal to $\vec{x}$. To correct $\{L_j\}_{j \neq u}$, we therefore need $\vec{r} \perp \vec{v}_j$ for all $j \neq u$; that is, $\vec{r} \in \ker(C)$.

We must also ensure, however, that $PH_0P \not\propto P$, which requires that $\vec{r} \cdot \vec{\gamma} \neq 0$. As per the results in Chapter 5, such an $\vec{r}$ is guaranteed to exist when $H_0$ is outside the Lindblad span, as we have assumed for now. The most straightforward choice is to take $\vec{r}$ parallel to the projection of $\vec{\gamma}$ onto $\ker(C)$:

$$\vec{r} \propto \text{proj}_{\ker(C)} \vec{\gamma} = (I - CC^+)\vec{\gamma}, \tag{6.9}$$

where the superscript + denotes the Moore-Penrose pseudoinverse[1], as in Chapter 2. We can find the right proportionality constant by enforcing normalization. Then, choosing the amplitudes of the coefficients in Eq. (6.3) as

$$|s_j|^2 = \frac{1}{2}\Big\{1 - \frac{n-2}{\vec{1}^{\top}(I - CC^+)\vec{\gamma}} \left[(I - CC^+)\vec{\gamma}\right]_j\Big\}, \tag{6.10}$$

for any complex phases, where $\vec{1} := (1, \ldots, 1)^{\top}$, would seem to give a QEC code satisfying Eqs. (5.3) and (5.4).

Notice, for instance, that this recipe immediately gives Eq. (5.27) upon plugging in $C_{\text{neg}}$ from Section 5.3.1 with maximally strong spatial correlations ($\varphi = 1$). Doing the same for $C_{\text{pos}}$ in 5.3.2 reveals a problem, however: it gives the nonsensical solution

$$|s_1|^2 = 2 \qquad |s_2|^2 = -1/2 \qquad |s_3|^2 = -1/2, \tag{6.11}$$

meaning that the ansatz in Eq. (6.3) does not work for this correlation matrix. Fortunately, there is an easy fix in this particular case: one can simply flip the quantization

[1]Note that $CC^+$ is the orthogonal projector onto $\text{col}(C)$, so $I - CC^+$ is that onto $\ker(C)$.

axes of some qubits (i.e., swap our definitions of $|0_j\rangle$ and $|1_j\rangle$ for certain qubits $j$) and try again.

Mathematically, this procedure is described by a diagonal matrix $Q$, with $q_{jj} = -1$ if the $j^{\text{th}}$ qubit is to be "flipped" (in terms of our labeling that is—nothing happens to the physical qubit), and $+1$ otherwise. This relabeling transforms $\vec{\gamma}$ and $C$ as

$$\vec{\gamma} \mapsto \vec{\gamma}\,' = Q\vec{\gamma} \qquad C \mapsto C' = QCQ. \tag{6.12}$$

The coefficients in Eq. (6.10) accordingly transform as

$$|s_j|^2 \mapsto |s'_j|^2 = \frac{1}{2}\Big\{1 - \frac{n-2}{\vec{1}^{\top} Q(I - CC^{+})\vec{\gamma}} \left[Q(I - CC^{+})\vec{\gamma}\right]_j\Big\}. \tag{6.13}$$

We can then attempt to choose $Q$ (i.e., choose whether the ground state of each qubit is labeled $|0\rangle$ or $|1\rangle$) so that each $|s'_j|^2$ is non-negative, and therefore describes a realizable QEC code. For $C_{\text{pos}}$ from Section 5.3.2 with $\varphi = 1$, the $2^3 = 8$ choices of $Q$ yield three distinct solutions (up to permutations of qubit and codeword labels). The first is the nonsensical solution in Eq. (6.11). The second, from flipping qubit 1, is

$$|s'_1|^2 = 2/7 \qquad |s'_2|^2 = 5/14 \qquad |s'_3|^2 = 5/14. \tag{6.14}$$

The third, from flipping qubit 2 or qubit 3, is

$$|s'_1|^2 = 0 \qquad |s'_2|^2 = 1/6 \qquad |s'_3|^2 = 5/6. \tag{6.15}$$

Plugging these into the ansatz in Eq. (6.3), one finds that they correspond to the code in Eq. (5.38) for $x = 5/14$ and $x = 0$ respectively. Therefore, the more general QEC code introduced in this section (up to qubit relabeling) reduces to some of the codes found "by hand" in the previous chapter. Of course, it is clearly not the only possible code [cf. Eqs. (5.35) and (5.38) for general $x$, which are not captured here]. This is fortunate, since the code developed in this section—even with qubit relabeling—is not defined for arbitrary correlation matrices $C$, even when $H_0$ is outside the Lindblad

span.

To see the limitations of the code introduced here, recall first the earlier definition $r_j := 1 - 2|s_j|^2$ for the coefficients in Eq. (6.3). Notice that $0 \leq |s_j|^2 \leq 1$ if and only if $||\vec{r}||_\infty \leq 1$, where the "infinity" norm is defined as $||\vec{r}||_\infty = \max_j |r_j|$. Allowing for arbitrary flipping/relabeling of the qubits, and substituting into Eq. (6.13), we can re-write this condition as

$$(n-2)\,\frac{||Q(I - CC^+)\vec{\gamma}||_\infty}{|\vec{1}^\top Q(1 - CC^+)\vec{\gamma}|} \leq 1. \tag{6.16}$$

Eq. (6.16) can be interpreted as follows: qubit flipping/relabeling can yield a valid QEC code of the type introduced in this section if and only if there exists a diagonal $Q = (\pm\delta_{jk})$ which satisfies the inequality. It can be simplified by noting that $||Q\vec{x}||_\infty = ||\vec{x}||_\infty$, and that the inequality is most easily satisfied when the denominator is largest, which occurs when we pick a $Q$ such that $\vec{1}^\top Q = (\pm 1, \ldots, \pm 1)$ has elements with the same signs[2] as $(I - CC^+)\vec{\gamma}$ (so that they all add "constructively"). That is:

$$\max_Q \left|\vec{1}^\top Q(I - CC^+)\vec{\gamma}\right| = ||(I - CC^+)\vec{\gamma}||_1, \tag{6.17}$$

where $||\vec{x}||_1 := \sum_j |x_j|$. Combining these observations, we conclude that the QEC code introduced above—even with spin flipping/relabeling—is well-defined if and only if

$$n - 2 \leq \frac{||(I - CC^+)\vec{\gamma}||_1}{||(I - CC^+)\vec{\gamma}||_\infty}. \tag{6.18}$$

Roughly speaking, then, one expects this code to become ill-defined more often for larger $n$, even when $H_0$ is outside the Lindblad span. Indeed, consider for instance a

[2]Equivalently, we could pick $Q$ to have the opposite signs; the result is the same.

sensor with $n = 5$ qubits for which $\vec{\gamma} = (\gamma, \gamma, \gamma, \gamma, \gamma)^\top$ and

$$C = \begin{pmatrix} 1 & \alpha & \alpha & \alpha & \alpha \\ \alpha & 1 & \alpha & \alpha & \alpha \\ \alpha & \alpha & 1 & \alpha & \alpha \\ \alpha & \alpha & \alpha & 1 & \beta \\ \alpha & \alpha & \alpha & \beta & 1 \end{pmatrix}, \tag{6.19}$$

for $\alpha = (1 + \sqrt{61})/60$ and $\beta = -9/10$. Here $H_0$ is outside the Lindblad span, which guarantees the existence of a QEC code that can perfectly filter signal from noise in principle. However, it violates the inequality in (6.18). This means—and indeed, one can check—that the code introduced in this section breaks down here, and cannot be fixed by relabeling any number of qubits.

Of course, the previous example is highly contrived, and the QEC code from this section could still often be useful in practice. The point, however, is that we are not justified in using this code to gauge the potential of error-corrected quantum sensing (ECQS) for arbitrary $n$, $\vec{\gamma}$ and $C$, because it is not always defined. To this end, we will instead develop a new code using a similar technique, which is always well-defined. It is with this new code that we will ultimately arrive at a general expression for the sensitivity offered by QEC.

### 6.1.2 Transforming the Repetition Code

Rather than using the Dicke code in Eq. (5.27) as a starting point for a new QEC code, we will instead use the repetition code. Consider now the ansatz (with $n \geq 3$)

$$|0_{\text{L}}\rangle = \bigotimes_{j=1}^{n} \Big[ \cos(\theta_j) |0_j\rangle + i \sin(\theta_j) |1_j\rangle \Big] \qquad |1_{\text{L}}\rangle = X^{\otimes n} |0_{\text{L}}\rangle . \tag{6.20}$$

These codewords have the same form as those of the repetition code, except that each qubit has been rotated[3] by some to-be-determined angled $\theta_j$.

[3] One could equivalently rotate about another axis, i.e., choose a different phase between $|0_j\rangle$ and $|1_j\rangle$, provided the relation between $|0_{\text{L}}\rangle$ and $|1_{\text{L}}\rangle$ is adjusted accordingly.

Let us verify that this code corrects all noise to order $O(\delta t)$, where $\delta t$ is the time between QEC recoveries, when $H_0$ is outside the Lindblad span for an appropriate choice of $\vec{\theta} = (\theta_1, \ldots, \theta_n)^\top$. We begin by showing that $|0_{\text{L}}\rangle$ and $|1_{\text{L}}\rangle$ are orthonormal. Normalization is clear. Orthogonality is apparent by noting that the components of $|0_{\text{L}}\rangle$ and $|1_{\text{L}}\rangle$ on qubit $j$ are

$$|0_{\text{L},j}\rangle = \cos(\theta_j)\,|0_j\rangle + i\sin(\theta_j)\,|1_j\rangle\,, \qquad |1_{\text{L},j}\rangle = i\sin(\theta_j)\,|0_j\rangle + \cos(\theta_j)\,|1_j\rangle \tag{6.21}$$

respectively, so that $|0_{\text{L}}\rangle = \otimes_{j=1}^n |0_{\text{L},j}\rangle$ and $|1_{\text{L}}\rangle = \otimes_{j=1}^n |1_{\text{L},j}\rangle$. Clearly $|0_{\text{L},j}\rangle$ and $|1_{\text{L},j}\rangle$ are orthogonal for all $j$, so $|0_{\text{L}}\rangle$ and $|1_{\text{L}}\rangle$ are also orthogonal.

Next, we examine terms of the form $P(\vec{x}\cdot\vec{Z})P$ for an arbitrary $\vec{x} \in \mathbb{R}^n$. For $n \geq 2$, the orthogonality of $|0_{\text{L},j}\rangle$ and $|1_{\text{L},j}\rangle$ implies that $\langle 0_{\text{L}}|\, Z_i\, |1_{\text{L}}\rangle = 0$. On the other hand,

$$\langle 0_{\text{L}}|\, Z_i\, |0_{\text{L}}\rangle = \cos(2\theta_i). \tag{6.22}$$

Likewise,

$$\langle 1_{\text{L}}|\, Z_i\, |1_{\text{L}}\rangle = \langle 0_{\text{L}}|\, X^{\otimes n} Z_i X^{\otimes n}\, |0_{\text{L}}\rangle = -\langle 0_{\text{L}}|\, Z_i\, |0_{\text{L}}\rangle = -\cos(2\theta_i). \tag{6.23}$$

Therefore

$$P(\vec{x}\cdot\vec{Z})P = \vec{x}\cdot\cos(2\vec{\theta})\, Z_{\text{L}}, \tag{6.24}$$

where the cosine is taken element-wise and $Z_{\text{L}} = |0_{\text{L}}\rangle\langle 0_{\text{L}}| - |1_{\text{L}}\rangle\langle 1_{\text{L}}|$. We can therefore enforce that $PL_jP \propto P$ for all non-vanishing $L_j$ by enforcing that $\cos(2\vec{\theta}) \perp \vec{v}_j$ for all eigenvectors $\vec{v}_j$ with nonzero eigenvalues; that is, by taking $\cos(2\vec{\theta}) \in \ker(C)$. To ensure that the signal survives at the logical level we also need $PH_0P \not\propto P$, meaning that $\vec{\gamma}\cdot\cos(2\vec{\theta}) \neq 0$. Mirroring our approach in Section 6.1.1, we can satisfy both conditions when $H_0$ is outside the Lindblad span by picking

$$\cos(2\vec{\theta}) \propto \text{proj}_{\ker(C)}\vec{\gamma} = (I - CC^+)\vec{\gamma}. \tag{6.25}$$

[Compare with Eq. (6.9).] More precisely, we can do so by picking angles $\vec{\theta}$ in our

repetition-code-like ansatz as

$$\vec{\theta} = \frac{1}{2}\arccos\left[\zeta \operatorname{proj}_{\ker(C)}\vec{\gamma}\right] = \frac{1}{2}\arccos\left[\zeta(I - CC^{+})\vec{\gamma}\right], \tag{6.26}$$

where $\zeta$ is an adjustable parameter in the range $[-\zeta_{\max}, \zeta_{\max}] \setminus \{0\}$ and

$$\zeta_{\max} = ||\operatorname{proj}_{\ker(C)}\vec{\gamma}||_{\infty}^{-1} = ||(I - CC^{+})\vec{\gamma}||_{\infty}^{-1} > 0. \tag{6.27}$$

This range of allowed $\zeta$'s was chosen so that the arccosine of each element in Eq. (6.26) is always well-defined, therefore always giving $PL_jP = 0$ and $PH_0P \not\propto P$ when $H_0 \notin \mathscr{S}$.

We now consider terms of the form $P(\vec{x}\cdot\vec{Z})(\vec{y}\cdot\vec{Z})P$, for arbitrary $\vec{x}, \vec{y} \in \mathbb{R}^n$. For $n \geq 3$, the orthogonality of $|0_{\text{L},i}\rangle$ and $|1_{\text{L},i}\rangle$ implies that $\langle 0_{\text{L}}| Z_jZ_k |1_{\text{L}}\rangle = 0$. We also have

$$\langle 1_{\text{L}}| Z_jZ_k |1_{\text{L}}\rangle = \langle 0_{\text{L}}| (X^{\otimes n}Z_jX^{\otimes n})(X^{\otimes n}Z_kX^{\otimes n}) |0_{\text{L}}\rangle = (-1)^2 \langle 0_{\text{L}}| Z_jZ_k |0_{\text{L}}\rangle . \tag{6.28}$$

Therefore, $P(\vec{x}\cdot\vec{Z})(\vec{y}\cdot\vec{Z})P \propto P$ automatically from our ansatz. We have therefore constructed a QEC code (or more properly, a family of codes) tailored for sensing applications, which satisfies the ECQS conditions in Eqs. (5.3) and (5.4) for all $n \geq 3$, $C$ and $\vec{\gamma}$ [provided $\vec{\gamma} \notin \operatorname{col}(C)$, i.e., $H_0$ is outside the Lindblad span].

## 6.2 Codes Through Semidefinite Programming

Before continuing our analysis of the QEC code found in Section 6.1.2 by deforming the repetition code, let us briefly recap a numerical technique developed by Zhou in Ref. [52] producing QEC codes for sensing. As we will see, the ansatz in Eq. (6.20) is optimal in terms of the sensing performance it offers when $H_0$ is outside the Lindblad span, and our choice of angles $\vec{\theta}$ is typically optimal. (At worst, it approximates the optimal choice, for which there is no known closed-form expression).

In Chapters 3 and 5 we made no efforts to formulate our optimization problems in

any special form with convenient mathematical properties. This is because there was no need to do so; it was sufficient for our purposes to construct generic objective/cost functions and apply general-purpose numerical optimization methods to them. In this section we will describe a different approach, using convex optimization, and more specifically, semidefinite programming.

An optimization problem is said to be convex if it can be written in the form

$$\text{Minimize } f(\vec{x}) \text{ subject to } g_i(\vec{x}) \leq 0 \text{ and } A_i\vec{x} = \vec{b}_i, \tag{6.29}$$

where $A_i$ and $\vec{b}_i$ are matrices and vectors respectively, and $f, g_j : \mathbb{R}^d \mapsto \mathbb{R}$ are convex functions, i.e.,

$$f\left[\phi\,\vec{x} + (1-\phi)\vec{y}\right] \leq \phi f(\vec{x}) + (1-\phi)f(\vec{y}) \tag{6.30}$$

for all $\vec{x}, \vec{y} \in \mathbb{R}^d$ and $\phi \in [0,1]$, and likewise for all $g_j$. (These functions could be defined only on a convex subset of $\mathbb{R}^d$, though we will not encounter that situation here.) Note that minimization problems can be converted to maximization problems and vice versa by replacing $f(\vec{x})$ with $-f(\vec{x})$.

Ref. [175] gives a thorough introduction to convex optimization, in which the authors write:

> "There is in general no analytical formula for the solution of convex optimization problems, but... there are very effective methods for solving them... If we can formulate a problem as a convex optimization problem, then we can solve it efficiently, just as we can solve a least-squares problem efficiently. With only a bit of exaggeration, we can say that, if you formulate a practical problem as a convex optimization problem, then you have solved the original problem."

The last point should be taken with a grain of salt when dealing with problems where the dimensionality $d$ can grow exponentially with the size of the physical system in question, as we are doing here. Nevertheless, it will be illuminating to encode the search for QEC codes for sensing into convex optimization problems because:

1. As stated above, there are very effective numerical methods for solving such problems.

2. Every minimum of a convex optimization problem is a global minimum. (In general there could be several distinct global minima, but they are all guaranteed to yield the same value of $f$.) There is therefore no possibility of getting stuck in a bad local minimum.

3. Many convex problems—and in particular, the ones we will encounter here—have an associated "dual problem" which may be simpler, and whose optimal value is equal to that of the original (or "primal") problem. This can provide partial information about the solutions to convex problems without having to solve them explicitly.

As in Section 6.1, we will consider here the ideal case where $H_0$ is outside the Lindblad span $\mathscr{S}$. This means there is guaranteed to exist a QEC code which can perfectly suppress the noise in the $\delta t \to 0$ limit, without equally suppressing the signal [151, 152]. Such codes offer vanishing (i.e., perfect) sensitivity $\eta$ in principle, because they can approach a quantum Fisher information scaling of $\mathcal{F} = (\xi \Delta t)^2$ with the sensing cycle length $\Delta t$. As we discussed in Section 4.2.3, a better figure of merit in this ideal scenario is therefore the coefficient $\xi$ that they offer. The best possible value of $\xi$ is given by [152]

$$\xi^2 = 4 \min_{\substack{S \in \mathscr{S} \\ S = S^\dagger}} ||H_0/\omega_0 - S||^2, \tag{6.31}$$

where $||A|| = \max_{|\psi\rangle} \langle\psi| A |\psi\rangle$ denotes the operator norm.

Zhou showed in Ref. [52] that solutions to the following optimization problem could be directly transformed into QEC codes for sensing which saturate Eq. (6.31):

$$\begin{gathered} \text{Maximize } \vec{x} \cdot \text{diag}(H_0/\omega_0) \text{ subject to} \\ ||\vec{x}||_1 \leq 2 \text{ and } \vec{x} \cdot \text{diag}(S) = 0 \text{ for all } S = S^\dagger \in \mathscr{S}, \end{gathered} \tag{6.32}$$

where $\mathrm{diag}(A)$ denotes the vector of diagonal elements of a matrix $A$. This is a convex optimization problem—in fact, it is a special kind called a semidefinite program[4]—in which $\vec{x}$ has dimension $d = 2^n$. Suppose $\vec{x}_\star$ is a solution to this problem. One can show that the states

$$|0_{\mathrm{L}}\rangle = \sum_{j=0}^{2^n-1} \sqrt{x_{\star j}^{(+)}}\,|j\rangle \qquad\qquad |1_{\mathrm{L}}\rangle = \sum_{j=0}^{2^n-1} \sqrt{x_{\star j}^{(-)}}\,|j\rangle\,, \tag{6.33}$$

satisfy Eqs. (5.3) and (5.4), where $\vec{x}_\star^{(\pm)}$ contains only the absolute values of positive/negative elements of $\vec{x}_\star$, with all other elements set to zero, such that $\vec{x}_\star = \vec{x}_\star^{(+)} - \vec{x}_\star^{(-)}$. (This is a very similar construction to that used in Section 2.2.) Moreover, by analyzing the dual of this problem, one can show that this code saturates Eq. (6.31).

In principle one could use Eq. (6.32) directly to devise QEC codes for particular $C$'s and $\vec{\gamma}$'s. (In fact, it can be also be used more broadly [52].) However, because the problem domain is exponentially large in the number of qubits $n$, this is not a scalable approach. Instead, one can start with the ansatz in Eq. (6.20) and use a different semidefinite program to find the angles $\vec{\theta}$. Not only will this new optimization problem be defined over $\mathbb{R}^n$ (as opposed to $\mathbb{R}^{2^n}$), but it will also saturate Eq. (6.31).

The semidefinite program in question is

$$\text{Maximize } \vec{x}\cdot\vec{\gamma} \text{ subject to } ||\vec{x}||_\infty \leq 1 \text{ and } C\vec{x} = \vec{0}, \tag{6.34}$$

where now $\vec{x} \in \mathbb{R}^n$. The solution $\vec{x}_\star$ to this problem can be used to form a QEC code using the ansatz in Eq. (6.20) with the angles $\vec{\theta} = \frac{1}{2}\arccos(\vec{x}_\star)$. Zhou showed in Ref. [52] that the resulting code saturates the bound[5] in Eq. (6.31). This is noteworthy for two reasons: First, it is a rare instance of a numerically-derived QEC code (or quantum protocol more broadly) in which the dimensionality of the problem is linear in the system size. This remarkable reduction in complexity compared to the semidef-

[4] In this setting a "program" refers to a an optimization problem, rather than a computer program.

[5] This does not guarantee that the solutions to Eqs. (6.32) and (6.34) will always produce the same QEC codes; only that these codes will offer the same performance.

inite program in Eq. (6.32) is due to the careful choice of ansatz in Eq. (6.20). By imposing the right structure before turning to numerical methods, we have removed illusory complexity and made this optimization-based approach scalable.

Second, one can see by inspection that the solution to Eq. (6.34) when $C$ has only one vanishing eigenvalue (the most likely case) is simply Eq. (6.26) with $\zeta = \zeta_{\max}$. In other words, our closed-form code from Section 6.1.2 is typically optimal when $H_0$ is outside the Lindblad span, and it approximates the optimal code in the rarer instances where $C$ has several vanishing eigenvalues. Notice that we have arrived at this conclusion without the need to actually perform numerics because the relevant problems could be formulated as well-behaved convex optimizations.

## 6.3 Sensitivity Afforded Under General Noise Correlations

In order to develop and analyze new QEC codes, we have so far made the simplifying assumption of maximally strong spatial noise correlations, leading to $H_0$ outside the Lindblad span. We now relax this assumption and turn our attention to the more general case of arbitrary $H_0$ and $C$ using the tools developed above. In light of the results in Sections 6.1 and 6.2, we will use the rotated-repetition-code ansatz in Eq. (6.20) as a starting point. We cannot claim that it offers optimal sensitivity in general; only that it is typically optimal in the limit of maximally strong noise correlations. On this basis, we expect it to perform at least reasonably well outside this limit.

Recall from Chapter 5 that in the generic case where $H_0$ is in the Lindblad span, there exists no QEC code that can completely suppress decoherence—even in principle—without equally suppressing the signal. As we saw in Sections 5.3.1 and 5.3.2, however, it is still possible for QEC to improve the signal to noise ratio and therefore the sensitivity—which is ultimately what matters—even if it cannot approach ideal performance in principle.

In this section we will find the sensitivity offered by QEC codes of the form in Eq. (6.20) for general spatial noise correlations. We will do so by finding the effective logical dynamics offered by these codes in the limit of frequent, instantaneous recoveries ($\delta t \to 0$). Ultimately, the effective dynamics will reduce to that of a dephasing qubit.

Recall from Section 5.2 that the optimal sensitivity for a single physical qubit in our noise model is

$$\eta_1 = \gamma^{-1}\sqrt{\frac{2e}{T_2^*}} = \sqrt{2e\Gamma} \tag{6.35}$$

in terms of the underlying noise strength $\Gamma$, and that for $n$ such qubits operating in parallel it is $\eta_{\text{par}} = \eta_1/\sqrt{n}$. As a warm-up, let us first find the sensitivity $\eta_{\text{GHZ}}$ offered by letting a GHZ state $(|0\ldots0\rangle + |1\ldots1\rangle)/\sqrt{2}$ evolve under $H_0 = \frac{\omega_0}{2}\vec{\gamma}\cdot\vec{Z}$ with arbitrary phase noise correlations $C$, for $\vec{\gamma} = (\gamma, \ldots, \gamma)^\top$. The energy difference between $|0\ldots0\rangle$ and $|1\ldots1\rangle$ is $\omega_0\,\vec{\gamma}\cdot\vec{1} = \omega_0 n\gamma$, so in the language of Eq. (4.50) we have $A = n$. Similarly, if each qubit dephases exponentially at a rate $1/T_2^* = \gamma^2\Gamma$, a superposition of $|0\ldots0\rangle$ and $|1\ldots1\rangle$ does so at a rate $B/T_2^*$ for $B = \vec{1}^\top C\vec{1}$. This may be faster or slower depending on $C$; for uncorrelated noise ($C = I$) it represents an $n$-fold increase in the dephasing rate. Comparing with Eq. (4.50), an initial GHZ state therefore gives a sensitivity of

$$\eta_{\text{GHZ}} = \frac{\sqrt{\vec{1}^\top C\vec{1}}}{n}\,\eta_1 \tag{6.36}$$

We can re-write this expression in a way that will facilitate comparison with the sensitivity offered by our QEC code. Namely, if we diagonalize $C$ as

$$C = V D_C V^\top, \tag{6.37}$$

where $V = (\vec{v}_1, \ldots, \vec{v}_n)$ is a matrix whose columns are orthonormal real eigenvectors of $C$ and $D_C = \text{diag}(\lambda_1, \ldots, \lambda_n)$, we have

$$\eta_{\text{GHZ}} = \frac{||D_C^{1/2}V^\top\vec{1}||_2}{n}\,\eta_1, \tag{6.38}$$

where $||\vec{x}||_2 = \sqrt{\sum_j |x_j|^2}$ is the usual 2-norm. Note that for uncorrelated noise

$$||D_C^{1/2} V^\top \vec{1}||_2 = \sqrt{n}, \tag{6.39}$$

thus negating any gains from entanglement. These expressions can easily be generalized for non-uniform $\vec{\gamma}$'s, although in doing so one must take care also to adjust the definition of $C$ in Eq. (5.13) appropriately, as discussed in Section 5.3.

We now perform an analogous—but more involved—analysis for the QEC code we developed in Section 6.1.2. In contrast with our initial treatment, however, we will not assume here that $C$ has a vanishing eigenvalue. This means that we cannot use Eq. (6.26) for the rotation angles $\vec{\theta}$ in our code (since $\ker(C) = \{\vec{0}\}$ generically). Instead, inspired by the examples in Sections 5.3.1 and 5.3.2, we will design our code so that a subdominant noise mode (rather than a vanishing one) is left uncorrected. The net result will be a non-vanishing signal at the logical level, as well as logical decoherence due to this purposely uncorrected noise mode. The hope will be to achieve a net enhancement in sensitivity at the logical level through such noise filtering.

Concretely, we will take

$$\vec{\theta} = \frac{1}{2} \arccos(\zeta \vec{v}_u), \tag{6.40}$$

where $u \in \{1, \ldots, n\}$ is the to-be-determined index of the Lindblad jump operator $L_u = \sqrt{\lambda_u} \vec{v}_u \cdot \vec{Z}$ that we will leave uncorrected, $\zeta \in [-\zeta_{\max}, \zeta_{\max}] \setminus \{0\}$ is again an adjustable parameter with $\zeta_{\max} = ||\vec{v}_u||_\infty^{-1}$, and the arccosine is again taken element-wise.

To find the sensitivity offered by the QEC code described above, we compute the sensor's effective Liouvillian, $\mathcal{L}_{\text{eff}} = \mathcal{R}\mathcal{L}\mathcal{P}$, under frequent QEC recoveries $\mathcal{R}$, where $\mathcal{L}$ is the sensor's Liouvillian and $\mathcal{P}(\rho) = P\rho P$ [60]. The usual QEC recovery in Section 1.2.2 (i.e., the transpose channel) results in population leakage out of the codespace due to the uncorrected error $L_u$, even when $\delta t \to 0$, which complicates the analysis. To prevent such leakage at leading order in $\delta t$, we modify the usual recovery so that the state is returned to the codespace after an $L_u$ jump, though perhaps with a logical error. This modification results in a Markovian, trace-preserving effective

dynamics over the two-dimensional codespace, given by $\mathcal{L}_{\text{eff}}$, from which it will be simple to find the afforded sensitivity. We now construct this $\mathcal{L}_{\text{eff}}$ explicitly.

We begin by constructing the recovery operation $\mathcal{R}$. The standard QEC recovery procedure (adapted from Section 1.2.2 for a Lindblad dynamics) is the following: The projector $P$ onto the codespace, together with the correctable jumps for our code $\{L_i\}_{i\neq u}$, define a set of rank-two projectors $\{P_i\}_{i\neq u}$ and corresponding unitaries $\{U_i\}_{i\neq u}$ [152]. To leading order in $\delta t$, the correctable jumps kick the sensor into the subspaces defined by the $P_i$'s. The standard/transpose recovery consists of measuring in $\{P, P_1, P_2, \dots\}$ (not containing a $P_u$), and applying $U_i^\dagger$ if the state is found in $\mathscr{C}_i := \text{col}(P_i)$, for $i \neq u$ [7, 19]. In the present setting, however, this procedure is problematic. The issue is that the uncorrected error $L_u$ can cause the state to jump into the "remainder" subspace $\mathscr{C}_{\text{R}} = \text{col}(P_{\text{R}})$, with projector $P_{\text{R}} = I - P - \sum_{i\neq u} P_i$. To avoid population leakage from the codespace into $\mathscr{C}_{\text{R}}$ at leading order in $\delta t$, we modify the usual procedure by returning the state to the codespace in the event of an error $L_u$, even though this error cannot—by design—be fully corrected. This gives an $\mathcal{L}_{\text{eff}}$ with non-trivial dynamics only in the 2-dimensional codespace, which lets us analyze the sensitivity using Eq. (4.50), i.e., as though it were a two-level system.

Concretely, let us first define Knill-Laflamme coefficients $m_{ij} \in \mathbb{R}$ by $PL_i^\dagger L_j P = m_{ij} P$, for all $i, j \in 1, \dots, n$. We also define the $n \times n$ code matrix $M = (m_{ij})_{i,j=1}^n$, and the $(n-1)\times(n-1)$ submatrix $\tilde{M} = (m_{ij})_{i,j\neq u}$, which is equal to $M$ with the $u^{\text{th}}$ row and column removed. (Recall that $u$ is the index of the noise mode left uncorrected.) Then, let $W$ be an orthogonal matrix diagonalizing $\tilde{M}$, such that $W^\top \tilde{M} W = \text{diag}(d_1, d_2, \dots)$. This $W$ lets us define new error operators $\{F_i\}_{i\neq u}$ such that

$$F_k = \sum_{i\neq u} w_{ik} L_i \qquad \text{and} \qquad L_j = \sum_{k\neq u} w_{jk} F_k, \tag{6.41}$$

which satisfy $PF_iF_jP = \delta_{ij} d_i P$. For $j \neq u$, we then use $F_j$ to define a unitary $U_j$ via polar decomposition, such that $F_j P = \sqrt{d_j} U_j P$, and finally $P_j = U_j P U_j^\dagger$. (If $d_j = 0$ take $U_j = 0$.) So far we have followed the standard construction. However, we now

define an additional unitary $U_u$ via the polar decomposition of $P_R L_u P$:

$$P_R L_u P = U_u \sqrt{(P_R L_u P)^\dagger (P_R L_u P)} = \sqrt{m_{uu} - \lambda_u \zeta^2 - \alpha}\, U_u P, \tag{6.42}$$

(taking $U_u = 0$ if $P_R L_u P = 0$) for some constant $\alpha$. Concretely, $\alpha$ is defined through

$$\begin{aligned} P L_u^\dagger P_{\text{R}} L_u P &= P L_u^2 P - (P L_u P)^2 - \sum_{i \in \mathcal{I}} P L_u U_i P U_i^\dagger L_u P \\ &= (m_{uu} - \lambda_u \zeta^2) P - \sum_{i \in \mathcal{I}} \frac{1}{|d_i|} P L_u L_i P L_i L_u P \\ &=: (m_{uu} - \lambda_u \zeta^2 - \alpha) P. \end{aligned} \tag{6.43}$$

where $\mathcal{I} = \{i \,|\, i \neq u,\, d_i \neq 0 \text{ and } 1 \leq i \leq n\}$. Our modified recovery channel then consists of measuring in $\{P, P_1, \ldots, P_n\}$, where $P_j := U_j P U_j^\dagger$ for all $j \in \{1, \ldots, j\}$ (now including $j = u$). [Eq. (6.42) immediately implies that these projectors satisfy $P_i P_j = \delta_{ij} P_i$ for all $0 \leq i, j \leq n$, where $P_0 := P$.] The correction step is then: If the state is found in the codespace, do nothing; if it is in $\mathscr{C}_i = \text{col}(P_i)$ for $1 \leq i \leq n$, apply $U_i^\dagger$. In other words:

$$\mathcal{R}(\rho) = \sum_{i=0}^{n} U_i^\dagger P_i \rho P_i U_i, \tag{6.44}$$

where $U_0 := I$.

Having defined $\mathcal{R}$, we now compute $\mathcal{L}_{\text{eff}}$. It is convenient to define superoperators $\mathcal{H}_0$ and $\mathcal{D}_j$ such that the sensor's Liouvillian takes the form

$$\mathcal{L}(\rho) = \underbrace{-i\,[H_0, \rho]}_{\mathcal{H}_0(\rho)} + \frac{1}{2T_2^*} \sum_{j=1}^{n} \underbrace{\Big( L_j \rho L_j - \frac{1}{2}\{L_j^2, \rho\} \Big)}_{\mathcal{D}_j(\rho)}, \tag{6.45}$$

where we assume $H_0 = \frac{\omega_0}{2} \vec{\gamma} \cdot \vec{Z}$ for $\vec{\gamma} = (\gamma, \ldots, \gamma)^\top$. Again, these results can easily be generalized for non-uniform $\vec{\gamma}$'s, although as discussed in Section 5.3, one must take care to define $C$ appropriately in this case.

We begin by computing $\mathcal{R}\mathcal{H}_0\mathcal{P}$. The fact that $P_j$'s are mutually orthogonal (for

$0 \leq j \leq n$) immediately implies that

$$\mathcal{R}\mathcal{H}_0\mathcal{P}(\rho) = [\mathcal{P}(H_0),\, \mathcal{P}(\rho)] = \frac{\omega_0\zeta(\vec{v}_u \cdot \vec{\gamma})}{2}[Z_{\text{L}}, \mathcal{P}(\rho)],$$

using (6.24). For a logical $\rho = \mathcal{P}(\rho)$ we may therefore write $\mathcal{R}\mathcal{H}_0\mathcal{P}(\rho) = [H_{\text{eff}}, \rho]$, where

$$H_{\text{eff}} = \frac{\omega_0\zeta(\vec{v}_u \cdot \vec{\gamma})}{2} Z_{\text{L}}. \tag{6.46}$$

We now turn to $\mathcal{R}\mathcal{D}_j\mathcal{P}$ for $j \neq u$, which we expect to vanish since the code corrects the associated error $L_j$ by design. Assuming a logical $\rho$ to simplify the notation, we have

$$\mathcal{R}\mathcal{D}_j\mathcal{P}(\rho) = \underbrace{PL_j\rho L_j P}_{\text{(i)}} - \underbrace{\frac{1}{2}P\{L_j^2, \rho\}P}_{\text{(ii)}} + \underbrace{\sum_{i=1}^{n} U_i^\dagger P_i L_j \rho L_j P_i U_i}_{\text{(iii)}} \tag{6.47}$$

where the other terms vanish because the $P_j$'s are all orthogonal. Term (i) vanishes since $PL_jP = 0$ for $j \neq u$, while term (ii) equals $m_{jj}\,\rho$. To evaluate term (iii), we first note that $U_u^\dagger P_u L_j P = 0$. The equality holds trivially when (6.42) vanishes, and when it does not we have

$$U_u^\dagger P_u L_j P = (m_{uu} - \lambda_u\zeta^2 - \alpha)^{-1/2}\; PL_uP_{\text{R}} \sum_{i\in\mathcal{I}} \frac{w_{ji}}{\sqrt{d_i}} P_i U_i = 0 \tag{6.48}$$

since $P_{\text{R}}P_i = 0$. This leaves

$$\begin{aligned}
\text{term (iii)} &= \sum_{i\in\mathcal{I}} U_i^\dagger P_i L_j P \rho P L_j P_i U_i \\
&= \sum_{i\in\mathcal{I}} \frac{1}{|d_i|} P F_i L_j P \rho P L_j F_i P \\
&= \sum_{i\in\mathcal{I}} \sum_{k,\ell\neq u} \frac{1}{|d_i|} w_{jk} w_{j\ell} \delta_{ik}\delta_{i\ell} d_k d_\ell \rho \\
&= m_{jj}\,\rho.
\end{aligned} \tag{6.49}$$

We therefore have the expected result $\mathcal{R}\mathcal{D}_j\mathcal{P} = 0$.

Finally, we turn our attention to $\mathcal{R}\mathcal{D}_u\mathcal{P}$, which we do not expect to vanish (unless

we are in the limit where $H_0 \notin \mathscr{S}$), since we have designed our code such that $L_u$ is uncorrectable. As before, we have

$$\mathcal{RD}_u\mathcal{P}(\rho) = \underbrace{PL_u\rho L_u P}_{\text{(I)}} - \underbrace{\frac{1}{2}P\{L_u^2,\rho\}P}_{\text{(II)}} + \underbrace{\sum_{i=1}^{n} U_i^\dagger P_i L_u \rho L_u P_i U_i}_{\text{(III)}}. \tag{6.50}$$

From (6.24), the first term simplifies to

$$\text{term (I)} = \lambda_u\zeta^2\, Z_{\text{L}}\rho Z_{\text{L}}, \tag{6.51}$$

while term (II) becomes $m_{uu}\,\rho$. Treating the $i \neq u$ and $i = u$ parts of term (III) separately, we have

$$\sum_{i\in\mathcal{I}} U_i^\dagger P_i L_u\rho L_u P_i U_i = \sum_{i\in\mathcal{I}} \frac{1}{|d_i|} PL_iL_uP\rho PL_uL_iP = \alpha\rho, \tag{6.52}$$

so

$$\mathcal{RD}_u\mathcal{P}(\rho) = \lambda_u\zeta^2 Z_{\text{L}}\rho Z_{\text{L}} + (\alpha - m_{uu})\rho + U_u^\dagger P_u L_u\rho L_u P_u U_u. \tag{6.53}$$

One immediately sees from (6.53) that were it not for the $P_u$ measurement and feedback that we have added to $\mathcal{R}$ (the last term in the above equation), $\mathcal{L}_{\text{eff}}$ would not be of Lindblad form over the codespace, due to leakage into $\mathscr{C}_{\text{R}} = \text{col}(P_{\text{R}})$. This is manifest through the mismatch between the $\lambda_u\zeta^2$ and $(\alpha - m_{uu})$ coefficients. By design, however, we have

$$\begin{aligned} U_u^\dagger P_u L_u P &= PU_u^\dagger L_u P \\ &= (m_{uu} - \lambda_u\zeta^2 - \alpha)^{-1/2}\; PL_uP_{\text{R}}L_uP \\ &= \sqrt{m_{uu} - \lambda_u\zeta^2 - \alpha}\; P \end{aligned} \tag{6.54}$$

from (6.42), which cancels the mismatched $\alpha - m_{uu}$ term in (6.53), giving the valid, trace-preserving Lindblad dissipator

$$\mathcal{RD}_u\mathcal{P}(\rho) = \lambda_u\zeta^2\big(Z_{\text{L}}\rho Z_{\text{L}} - \rho\big). \tag{6.55}$$

at the logical level.

In summary, then, the effective logical dynamics in the limit of frequent recoveries is generated by

$$\mathcal{L}_{\text{eff}}(\rho) = -i[H_{\text{eff}}, \rho] + L_{\text{eff}}\,\rho L_{\text{eff}}^\dagger - \frac{1}{2}\{L_{\text{eff}}^\dagger L_{\text{eff}}, \rho\}, \tag{6.56}$$

where $H_{\text{eff}} = \omega_0\zeta(\vec{v}_u \cdot \vec{\gamma})Z_{\text{L}}/2$ and $L_{\text{eff}} = \zeta\sqrt{\frac{\lambda_u}{2T_2^*}}Z_{\text{L}}$, with $Z_{\text{L}} = |0_{\text{L}}\rangle\langle 0_{\text{L}}| - |1_{\text{L}}\rangle\langle 1_{\text{L}}|$. In other words, at the logical level, the sensor accumulates a phase at a rate set by $\zeta$ and the overlap of $L_u$ with $H_0$, while also losing phase coherence at a rate set by $\zeta$ and $\lambda_u$. More specifically, the sensor's effective logical dynamics becomes that of a dephasing qubit with $A = \zeta|\vec{v}_u \cdot \vec{1}|$ and $B = \zeta^2\lambda_u$ from Eq. (4.50), giving

$$\eta_{\text{QEC}}^{(u)} = \frac{\sqrt{\lambda_u}}{|\vec{v}_u \cdot \vec{1}|}\,\eta_1 \tag{6.57}$$

when we leave mode $u \in \{1, \ldots, n\}$ uncorrected. Notice that the free parameter $\zeta$ cancels out here. Finally, the optimal choice of index $u$ is the one that minimizes the above quantity, giving:

$$\eta_{\text{QEC}} = \frac{1}{||D_C^{-1/2}V^\top\vec{1}||_\infty}\,\eta_1, \tag{6.58}$$

valid for arbitrary noise correlation profiles $C$ on $n \geq 3$ qubits. (Note that $D_C^{-1/2}$ is undefined when $C$ has a vanishing eigenvalue. In this case, Eq. (6.58) should be regularized by replacing $D_C \to D_C + \epsilon I$, evaluating the norm, then taking $\epsilon \to 0$.) Eq. (6.58) identifies the $C$'s for which this QEC scheme provides enhanced sensitivity over parallel and GHZ sensing [cf. Eq. (6.38)]. Notice that while $H_0$ is outside the Lindblad span only in a measure-zero set of $C$'s (on the boundary of the set of possible correlation matrices), QEC can enhance sensitivity over a much larger set, regardless of whether it can approach noiseless sensing in principle.

This result generalizes the behavior we observed in special cases considered in Sections 5.3.1 and 5.3.2. In fact, plugging in $C_{\text{neg}}$ from Section 5.3.1 gives precisely the same sensitivity as in Eq. (5.34) (using $u = 3$, independent of the correlation strength $\varphi$ and the adjustable parameter $\zeta$), which arose from *ad hoc* QEC codes

in the previous chapter. Whereas the code from Section 6.1.1 reduced to that in Eq. (5.27), the one considered here reduces to Eq. (5.35) for $\zeta = \pm\zeta_{\max}$, and produces new codes for general $\zeta$. [No value of $\zeta$ produces Eq. (5.27).]

On the other hand, plugging in $C_{\text{pos}}$ from Section 5.3.2 gives the same sensitivity as in Eq. (5.43) for correlation strengths $\varphi \geq 392/409 \approx 0.96$ (using $u = 3$, regardless of $\zeta$). For $\varphi < 392/409$, however, it is better to leave mode $u = 2$ uncorrected, whereas in Section 5.3.2 we used $u = 3$ for all correlation strengths $\varphi$. In this latter regime, taking $u = 2$ gives

$$\eta_{\text{QEC}} = \frac{\sqrt{17(8+9\varphi)}}{20}\,\eta_1, \tag{6.59}$$

which is substantially better than the sensitivity in Eq. (5.43), though slightly worse than the performance offered by GHZ and parallel sensing schemes. The codes resulting from $C_{\text{pos}}$ here are generically different from those in Eq. (5.38).

## 6.4 Illustration: Distance-Dependent Noise Correlations

Eq. (6.58) admits a broad range of $\eta_{\text{QEC}}$ vs. $n$ scalings due to the critical dependence of $\eta_{\text{QEC}}$ on $C = C(n)$, which could grow with $n$ in myriad different ways. Consider, for instance, a sensor comprising $m$ clusters of $\frac{n}{m}$ qubits, where each cluster's Hamiltonian is outside the Lindblad span, but where the noise has no inter-cluster correlations. In this case, one could use the code from Section 6.1.2 to make a noiseless sensor from each cluster in principle, and perform GHZ sensing at the logical level to get $\eta \propto 1/m \propto 1/n$ Heisenberg scaling. On the other hand, given an $n$-qubit sensor for which the total Hamiltonian is already outside the Lindblad span, adding an additional qubit which shares no noise correlations with the others has no impact on $\eta_{\text{QEC}}$. We now analyze an intermediate example between these extreme scalings in this section.

Consider a sensor comprising $n \geq 3$ identical qubits arranged in a ring, and equally

spaced from one another. Suppose that the noise correlation coefficient between two qubits depends only on the distance between them, so that neighboring qubits have a coefficient of $\alpha_1$, next-to-nearest neighbors have $\alpha_2$, and so on up to $\alpha_{\max}$ for the most distant qubits, as shown in the left panel of Fig. 6-1. (N.b., the special case of $\alpha_i = 0$ describes a lack of correlation.) We emphasize that in practice, the distance between two qubits is a poor predictor of how strongly correlated the noise in their gaps is. As discussed Part I, other factors, like proximity and relative orientations to nearby fluctuators, for instance, are often more important. Accordingly, this example is not meant to provide a particularly realistic model, but rather, an illustrative one which can be solved exactly, and which has normal noise modes (i.e., Lindblad jump operators) with a simple physical interpretation.

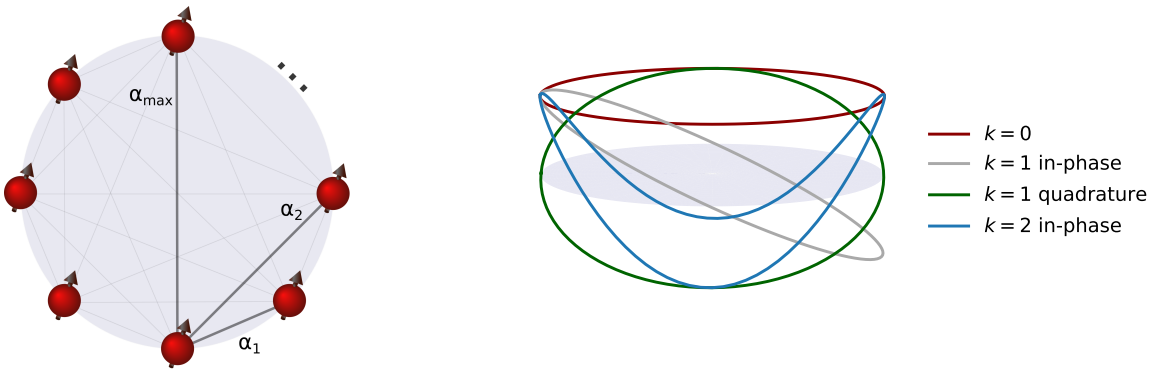



Figure 6-1: Left: A ring of $n \geq 3$ evenly-spaced probing qubits. The lines connecting each pair denote an arbitrary correlation strength $\alpha_i$ in the energy gap noise felt by either qubit. Right: The first few Fourier modes in the sensor resulting from the spatial noise correlations. (Qubits not shown.) The height of each point along a curve denotes the weight of $Z_j$ for qubit $j$ in the corresponding Lindblad jump operator, for modes $\vec{v}_0$, $\vec{v}_{1,0}$, $\vec{v}_{1,\pi/2}$, $\vec{v}_{2,0}$.

The qubits in this sensor are assumed to undergo Markovian dephasing described

by Eq. (5.13), with $\vec{\gamma} = (\gamma, \ldots, \gamma)^\top$ for simplicity and

$$C = \begin{pmatrix} 1 & \alpha_1 & \alpha_2 & & \alpha_2 & \alpha_1 \\ \alpha_1 & 1 & \alpha_1 & \cdots & \alpha_3 & \alpha_2 \\ \alpha_2 & \alpha_1 & 1 & & \alpha_4 & \alpha_3 \\ & \vdots & & \ddots & & \vdots \\ \alpha_2 & \alpha_3 & \alpha_4 & & 1 & \alpha_1 \\ \alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_1 & 1 \end{pmatrix}, \tag{6.60}$$

where we assign qubit numbers/labels sequentially based on their location. Notice that $C$ is a circulant matrix, so it is diagonalized by a discrete Fourier transform matrix [176]. This means that the eigenvectors of $C$ are simply the spatial Fourier modes on the ring. These are often chosen to be complex vectors of the form

$$\vec{w}_k = \frac{1}{\sqrt{n}}(1, \Omega_k, \Omega_k^2, \ldots, \Omega_k^{n-1})^\top, \tag{6.61}$$

where $\Omega_k := \exp(2\pi i k/n)$ for $k = 0, \ldots, n-1$. The corresponding eigenvalues are

$$\lambda_k = 1 + \alpha_1\Omega_k + \alpha_2\Omega_k^2 + \cdots + \alpha_2\Omega_k^{n-2} + \alpha_1\Omega_k^{n-1} \in \mathbb{R}. \tag{6.62}$$

Since these eigenvalues come in degenerate pairs ($\lambda_k = \lambda_{n-k}$ for $k \geq 1$), we can equivalently form a real eigenbasis for $C$ from $\vec{w}_k \pm \vec{w}_{n-k}$, in keeping with the convention from previous sections of using $\vec{v}_k \in \mathbb{R}^n$. The first such eigenvector is the $k = 0$, or constant Fourier mode

$$\vec{v}_0 = \frac{1}{\sqrt{n}}(1, \ldots, 1)^\top = \vec{1}/\sqrt{n}, \tag{6.63}$$

Higher wavenumbers $k \geq 1$ each describe a pair of eigenvectors with a $\pi/2$ phase offset on the ring:

$$\begin{aligned} \vec{v}_{k,0} &= \Big(1, \cos(\vartheta_k), \cos(2\vartheta_k), \ldots, \cos\big[(n-1)\vartheta_k\big]\Big)^\top \\ \vec{v}_{k,\pi/2} &= \Big(0, \sin(\vartheta_k), \sin(2\vartheta_k), \ldots, \sin\big[(n-1)\vartheta_k\big]\Big)^\top, \end{aligned} \tag{6.64}$$

where $\vartheta_k = 2\pi k/n$. The first few of these Fourier modes are illustrated in the right panel of Fig. 6-1. The key observation here is that all but the $k = 0$ noise mode are orthogonal to $\vec{\gamma}$. This forces us to take $\vec{v}_u = \vec{v}_0$ as the mode to leave uncorrected. Any other choice of $u$ would give $\eta_{\text{QEC}}^{(u)} = \infty$, and is therefore inadmissible. The eigenvalue of $C$ associated with this mode is given by

$$\lambda_u = \vec{v}_u^\top C \vec{v}_u = \frac{1}{n}\, \vec{1}^\top C \vec{1}. \tag{6.65}$$

From (6.58), this gives an ideal sensitivity of

$$\eta_{\text{QEC}} = \frac{\sqrt{\vec{1}^\top C\ \vec{1}}}{n}\, \eta_1 = \sqrt{\frac{1 + 2(\alpha_1 + \alpha_2 + \cdots + \alpha_{\max})}{n}}\, \eta_1 \tag{6.66}$$

Notice that the scaling of $\eta_{\text{QEC}}$ with $n$ here has two components: (i) a generic $1/\sqrt{n}$ improvement as one adds qubits, and (ii) a term that depends on how $\sum_i \alpha_i$ changes with $n$.

Comparing with $\eta_{\text{par}} = \eta_1/\sqrt{n}$, one sees that QEC provides an advantage over a parallel sensing scheme (i.e., one with initial state $|\psi_0\rangle = |+\rangle^{\otimes n}$) when $\alpha_1 + \cdots + \alpha_{\max} < 0$. Therefore, in this particular setting, there must be negative correlations in the noise (i.e., certain $\alpha_j < 0$) for QEC to provide an advantage over parallel sensing. This is not required in general however, as we have seen in Section 5.3.2.

Finally, we compare the sensitivity offered by QEC in this model with that offered by a GHZ scheme which uses an initial state $|\psi_0\rangle = \frac{1}{\sqrt{2}}(|0\ldots0\rangle + |1\ldots1\rangle)$. Eq. (6.38) immediately gives $\eta_{\text{GHZ}} = \eta_{\text{QEC}}$—therefore, in this particular example, our QEC scheme offers the same sensitivity as GHZ sensing (in the $\delta t \to 0$ limit) but does not surpass it. There is a simple explanation for this apparent coincidence: taking $\vec{v}_u = \vec{v}_0$ and $\zeta = \zeta_{\max} = \sqrt{n}$ gives logical states $|0_{\text{L}}\rangle = |0\ldots0\rangle$ and $|1_{\text{L}}\rangle = |1\ldots1\rangle$, resulting in the same initial state as the GHZ scheme. Moreover, a simple calculation shows that both of these logical states are in the kernel of every Lindblad jump with $k \geq 1$. In other words, $\text{span}\{|0\ldots0\rangle, |1\ldots1\rangle\}$ is a decoherence-free subspace of $\{L_k\}_{k\geq 1}$, and so the recovery $\mathcal{R}$ reduces to the identity channel (i.e., doing nothing) [60]. This means

that for this model, GHZ sensing is a special case of our error-correction scheme, for a particular choice of the adjustable parameter $\zeta$. Since $\eta_{\mathrm{QEC}}$ is independent of $\zeta$, it follows that $\eta_{\mathrm{QEC}} = \eta_{\mathrm{GHZ}}$ here. Of course, this is particular to the present model, and is not the case in general, even with purely positive noise correlations.

## 6.5 Discussion

In this chapter we proposed and analyzed a family of QEC codes which can filter background noise from a signal in quantum sensors. We found closed-form expressions for these codes in terms of the noise correlations in a sensor, and ultimately arrived at an expression for the sensitivity they can offer in principle. As expected, our codes are more likely to enhance sensitivity when there are strong noise correlations. It is not clear at present, however, why certain noise correlations seem to enable much better sensitivity than others.

While this chapter presents the natural continuation of the results in Chapter 5, it also raises new questions. For one, it is not clear whether, or when, the scheme we have proposed is optimal in the typical case where $H_0$ is in the Lindblad span. The fact that $\eta_{\mathrm{QEC}}$ for this scheme matches those from the examples in Chapter 5, which we analyzed using *ad hoc* methods, suggests that we may have reached a fundamental bound in these cases. On the other hand, we used a modified version of the standard QEC recovery here, so as to get a simple effective dynamics from which we could easily find the sensitivity. Our modified recovery, however, purposely causes logical errors rather than letting population build up in an unused subspace. It would be surprising if this mathematically-convenient strategy were optimal. Also, we did not analyze in detail what kind of circuits could implement the encodings or recoveries for the codes considered here. Naturally, this is would be an important next step if one is to implement these application-adapted codes in experiment. It would also be useful, to this end, to elucidate more concretely what noise correlations can be made to arise in different types of quantum sensors, and to what extent these are amenable to ECQS.

# Outlook

*"We live on an island surrounded by a sea of ignorance. As our island of knowledge grows, so does the shore of our ignorance."*

– John A. Wheeler

Quantum error correction (QEC) is often envisioned as a sort of quantum "software," which works somewhat independently of the underlying hardware. In both Parts I and II of this thesis, we illustrated how QEC might instead be used more like quantum "firmware." We showed that this firmware approach to QEC offers clear advantages over the conventional software approach in the near term. Yet, we have only just scratched the surface. Accordingly, we will discuss here some of the prospects and questions arising from the results in this thesis, beginning with those from Part I.

Beyond current experimental efforts to implement the codes from Chapter 2, the most natural extension of Part I would be to apply the techniques developed therein more broadly; that is, to devise device-adapted QEC codes for other error mechanisms, occurring in other types of quantum devices. These errors could include intrinsic decoherence as in Chapter 2, as well as measurement and control errors as in Chapter 3. Given the broad variety of current quantum devices, and the fact that device-adapted codes are—by design—not one-size-fits-all, there is likely a substantial amount of work to be done in this direction. More ambitiously, one could imagine a tighter feedback loop than currently exists between the design of quantum devices and that of QEC schemes to implement on them. A central theme of this thesis has been that noise can often be corrected more efficiently when it has structure, provided this structure is known and exploited. This suggests a new paradigm in which

one seeks not only to design less noisy quantum devices, but also to give whatever noise remains a particular structure that can be efficiently corrected through device-adapted QEC. One could imagine an iterative process whereby QEC codes and NISQ (noisy intermediate-scale quantum) devices guide each other's development.

This approach could perhaps be pushed even further by extending the techniques from Part I towards device-adapted fault tolerance. One strategy, mentioned in Chapter 2, would be to use modular architectures where device-adapted codes are used in each module as a first level of encoding, and the modules are then operated together using a conventional QEC code at the second level of encoding. Alternatively, it may be possible to incorporate a degree of fault tolerance directly into device-adapted codes. The underlying intuition is based on the recurring theme in this thesis that structured noise can be easier to manage than unstructured noise. In light of the results in Part I, it would not be surprising if the requirements for fault tolerance (e.g., thresholds) became much less stringent by incorporating device physics into the code from the start. Of course, this approach is unlikely to scale well to large devices, but it may provide a useful shortcut for initial demonstrations of fault tolerance.

We now turn our attention to the results from Part II, which raise a number of new questions about error-corrected quantum sensing. For instance, our results (and indeed, most other papers on the topic to date) are based on the effective sensor dynamics at leading order in the time between QEC recoveries. In experiments, however, higher-order effects could also be significant. These could bring both new challenges and new opportunities. For instance, higher-order terms in the effective dynamics can produce not only logical decoherence, but also additional Hamiltonian terms which could systematically bias the sensor's output [see Eq. (5.2)]. On the other hand, it could be that signal and noise which are difficult to distinguish at leading order in time through QEC could be distinguished more easily in these higher-order terms. More ambitiously, the techniques we used to develop application-adapted codes for sensing could perhaps also be used for other near-term application such as quantum annealing or analog quantum simulation, which also rely on manifestly analog dynamics.

The specialized codes in Part II could also provide fundamental insights into QEC more broadly. For instance, in Chapters 5 and 6 we encountered a strong tension between correcting local phase noise and allowing local unitaries to enact a logical unitary which encoded the signal (to leading order in time, at least). This difficulty seems reminiscent of the Eastin-Knill theorem, which says that a QEC code that can detect (arbitrary) single-qubit errors cannot have a universal, transversal encoded gate set [177]. (Of course, we had no interest in universal encoded operations in Part II, nor did we want to correct arbitrary local errors, but there still appears to be a similarity in spirit.) Indeed, recent work revisiting the Eastin-Knill theorem in the language of symmetries and approximate QEC hints further at a connection—see, e.g., Refs. [178, 179]. Notice, for instance, that the angles in Eq. (6.20) define single-qubit operators $\{O_j\}$ such that the separable unitary $\exp(-i\phi \sum_{j=1}^{n} O_j)$ enacts the logical unitary $\exp(-in\phi\, Z_{\text{L}})$. Moreover, the enhancement in sensitivity offered by these codes relies critically on the fact that they only partially correct the noise in question. It would be interesting, and perhaps mutually beneficial, to elucidate the connection between our application-adapted codes and the Eastin-Knill theorem.

Finally, there are some points common to both Parts I and II. In particular, we showed in this thesis that tailoring QEC codes for specific devices and applications can provide substantial benefits. In doing so, however, we took an extreme approach; that is, we devised codes that are highly adapted to particular error operators, and imposed almost no constraints *a priori* on their structure. For instance, we did not demand that the recoveries associated with our codes be easily implementable; instead, we occasionally got such structure for free, e.g., in Chapter 2. Having now illustrated the potential of QEC as quantum firmware, however, it is likely worthwhile going forward to impose a minimum structure on new QEC codes from the start. In other words, it is a good idea to demand up front that new codes have a certain locality, even if this means they cannot be as closely tailored to the error mechanisms at hand. The added ease of implementation from such constraints would likely justify the costs. What structure should be imposed depends on how these codes are to be implemented. For instance, codes intended to passively suppress errors through an energy penalty

could be constrained so that they are realizable using reasonably local couplings. Alternatively, codes to be implemented through active error detection and correction, as we have mostly considered here, could be constrained so that these procedures match the device connectivity. One could perhaps even demand that it be sufficient simply to track the error syndromes (and to modify future operations accordingly), rather than actively correcting errors as soon as they are detected. Again, the costs of imposing such constraints may be justified in practice. Perhaps the most useful QEC codes will ultimately lie somewhere between the highly customized ones developed here and conventional codes, which are highly structured.

Ideally, the techniques developed in this thesis will not remain relevant in the long term. That is, one hopes that quantum devices will eventually become so sophisticated that QEC—and indeed, quantum algorithms—can be implemented in a relatively hardware-agnostic way; i.e., as quantum software rather than firmware. It is becoming apparent, however, that getting to this point may require more than separate, steady advances in both quantum hardware and software. Rather, realizing the ultimate potential of quantum technologies may hinge upon our ability to first achieve intermediate goals, perhaps by means of special-purpose tools which are different from those we intend to use in the more distant future. The aim of this thesis has been to develop such tools. Some may object to this strategy, on the grounds that any steps off the most direct path to mature quantum technologies constitute wasted effort. I believe, on the contrary, that an indirect path, which winds through increasingly significant milestones, may in fact be the only viable one.

# Appendices

# Appendix A

# Appendix to Chapter 2

## A.1 Monte Carlo Averaging

We use Monte Carlo integration to compute $\langle p \rangle = \int_0^1 dg_1 \cdots \int_0^1 dg_n \, p$ in Figs. 2-1, 2-6, 2-10, 2-11 and 2-12. Specifically, we estimate $\langle p \rangle$ through

$$\langle p \rangle = \langle p(\sigma) \rangle \approx \frac{1}{N} \sum_{i=1}^{N} p(\vec{g}^{(i)}, \sigma) =: \mu, \tag{A.1}$$

where $\vec{g} = (g_1, \ldots, g_n)$ is sampled uniformly $N$ times from $[0,1]^n$. The sample variance of $p$ over $\vec{g} \in [0,1]^n$ is

$$\mathrm{Var}(p) = \frac{1}{N-1} \sum_{i=1}^{N} \left[ p(\vec{g}^{(i)}, \sigma) - \mu \right]^2, \tag{A.2}$$

so the standard error in approximating $\langle p \rangle$ by $\mu$ is $\sqrt{\mathrm{Var}(p)/N}$, which is shown as error bars/bands in these plots. To validate this numerical averaging, we compare in Fig. A-1 the Monte Carlo estimates of $\langle p \rangle$'s with analytical expressions [i.e., from integrating Eqs. (2.32), (2.33) etc.], when the latter can be calculated in Mathematica. The values of $\langle p \rangle$ for physical qubits and repetition codes shown in Figs. 2-1, 2-6, 2-10, 2-11 and 2-12 were found analytically to reduce unnecessary statistical noise in these plots. (That is, they use the yellow lines in Fig. A-1 rather than the blue dots.) Those for efficient codes are Monte Carlo averages.

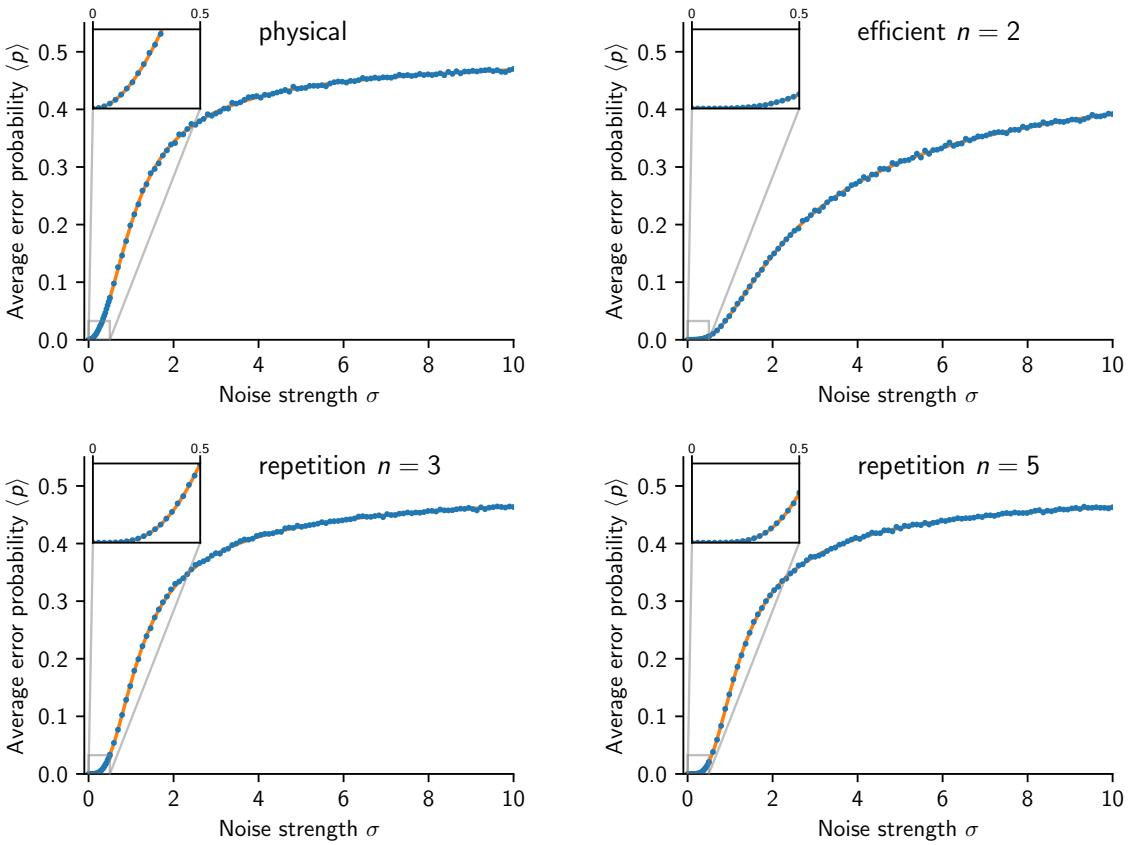



Figure A-1: Validation of Monte Carlo averaging. The yellow line is the exact $\langle p \rangle$ found analytically, and the blue dots are the Monte Carlo estimates of the same quantity. Error bars are too small to be seen due to the large number of samples used.

As further validation, we also plot the average error probability $\langle p \rangle$ in the ultra-low-noise regime in Fig. A-2, compared to the corresponding power law in $\sigma$.

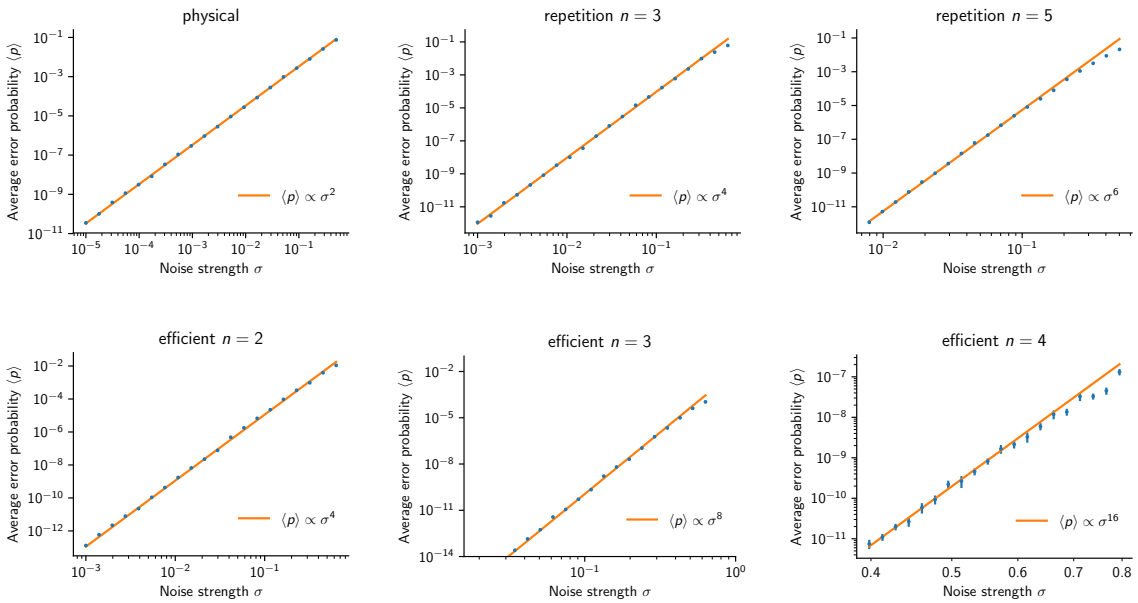
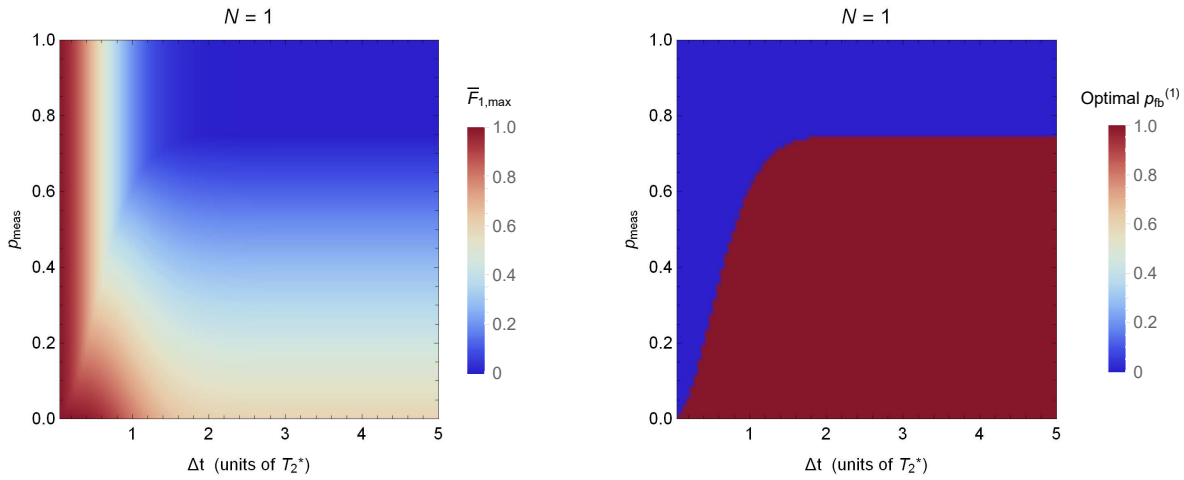


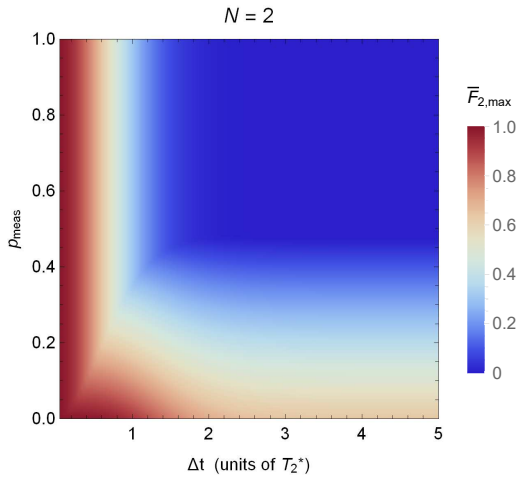

Figure A-2: Further validation of Fig. 2-1, in the ultra-low-noise regime. Each code is shown in a different panel for clarity (along with an unencoded qubit for reference). The blue dots are data points; error bars, when visible, denote standard error of the mean in Monte Carlo averaging. The yellow lines are the corresponding power laws whose slopes serve to guide the eye (their offsets, which were determined through fitting, are not of primary importance here). For clarity, the panels use different ranges of $\sigma$ (unlike the inset of Fig. 2-1), chosen to be as wide as possible, such that (i) they represent a perturbative regime ($\sigma \lesssim 1$) where $\langle p \rangle$ is described by a power law, and (ii) they do not produce unmanageably small values of $\langle p \rangle$ prone to significant numerical errors. The hardware-efficient code with $n = 5$ is not shown because it suppressed noise so strongly that it was difficult to find values of $\sigma$ satisfying both (i) and (ii).

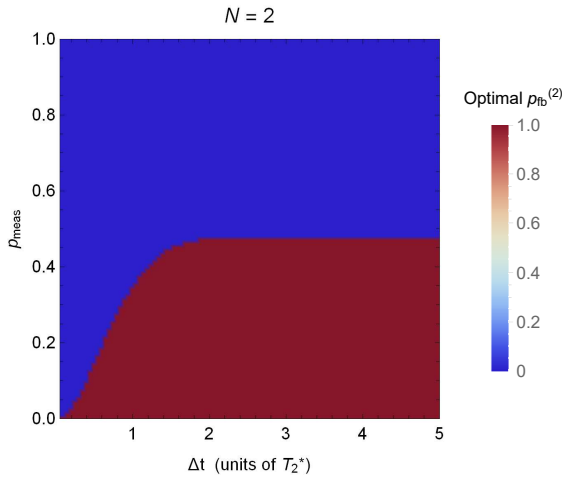
# Appendix B

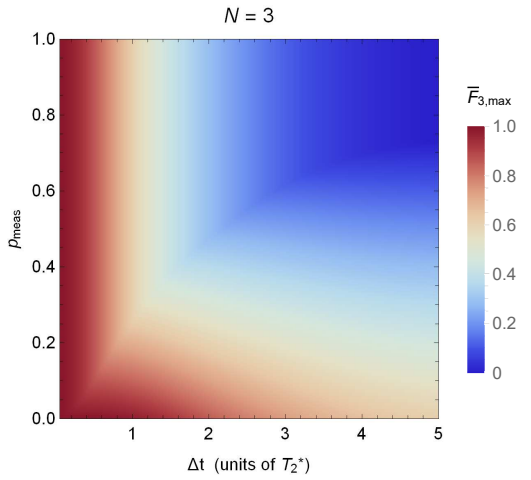
# Appendix to Chapter 3

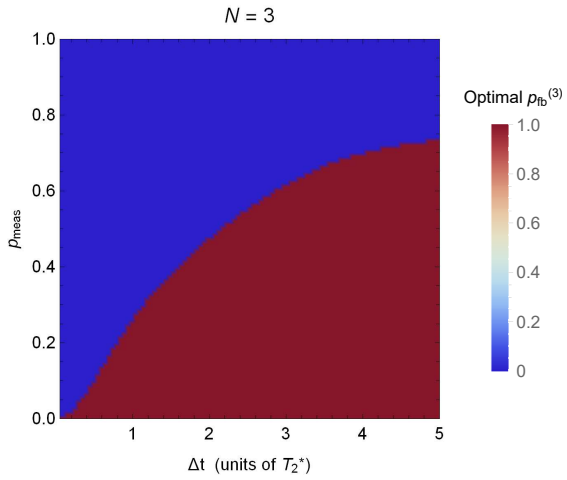
Figs. 3-2 and 3-3 show the results of an optimization performed first over $p_{\mathrm{fb}} \in [0, 1]$ for each $N$, and then over $1 \leq N \leq 10$. In Fig. B-1 we show the results from the first step of this optimization separately for each $N$.

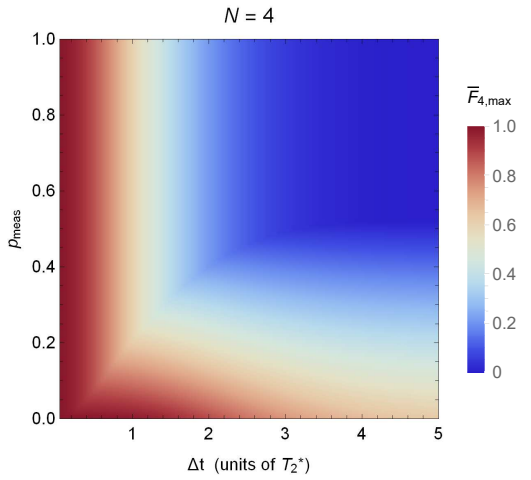


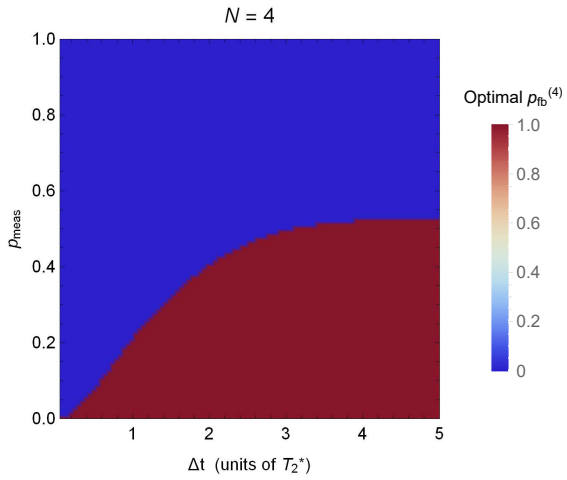

N = 2
p_meas
Δt (units of T2*)
F̄2,max

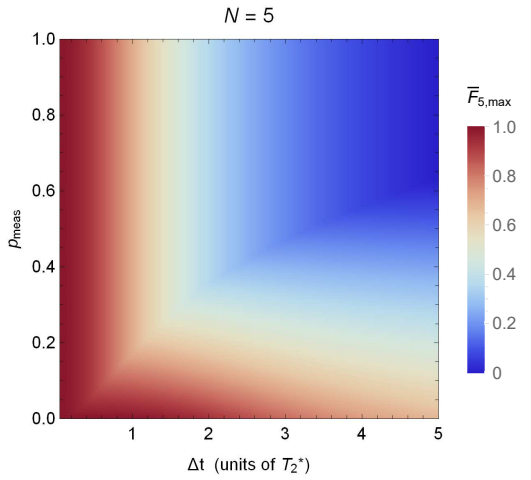

N = 2
p_meas
Δt (units of T2*)
Optimal pfb(2)

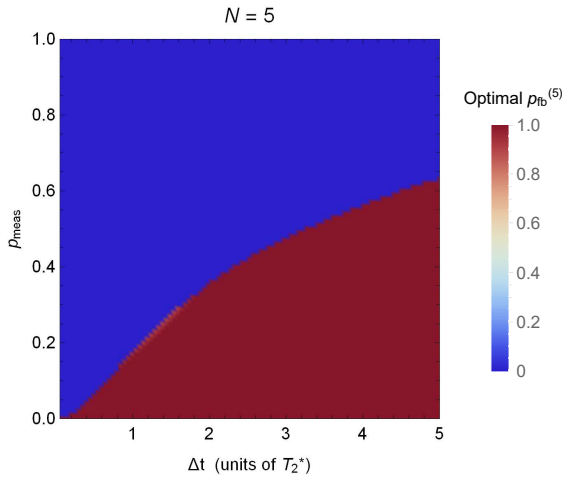

N = 3
p_meas
Δt (units of T2*)
F̄3,max

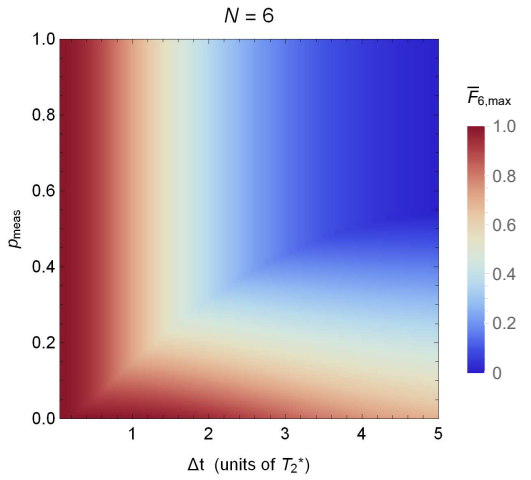

N = 3
p_meas
Δt (units of T2*)
Optimal pfb(3)

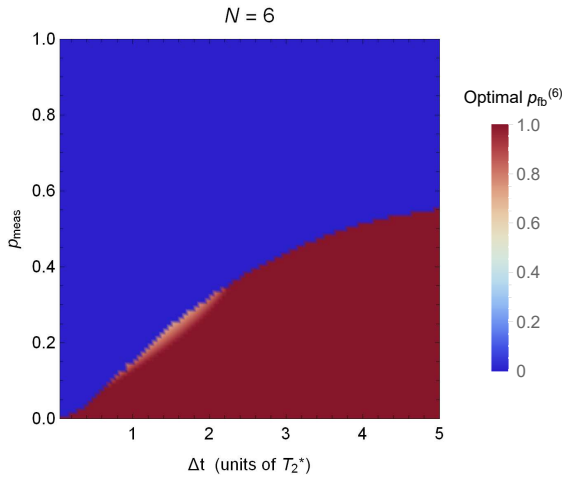

N = 4
p_meas
Δt (units of T2*)
F̄4,max

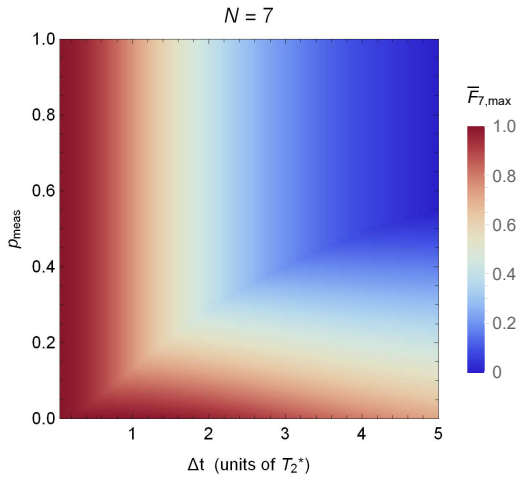

N = 4
p_meas
Δt (units of T2*)
Optimal pfb(4)

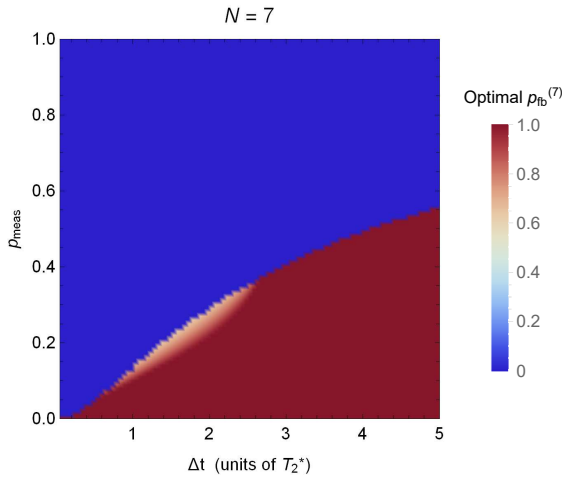

N = 5
p_meas
Δt (units of T2*)
F̄5,max


N = 5
p_meas
Δt (units of T2*)
Optimal p_fb(5)


N = 6
p_meas
Δt (units of T2*)
F̄6,max


N = 6
p_meas
Δt (units of T2*)
Optimal p_fb(6)


N = 7
p_meas
Δt (units of T2*)
F̄7,max


N = 7
p_meas
Δt (units of T2*)
Optimal p_fb(7)

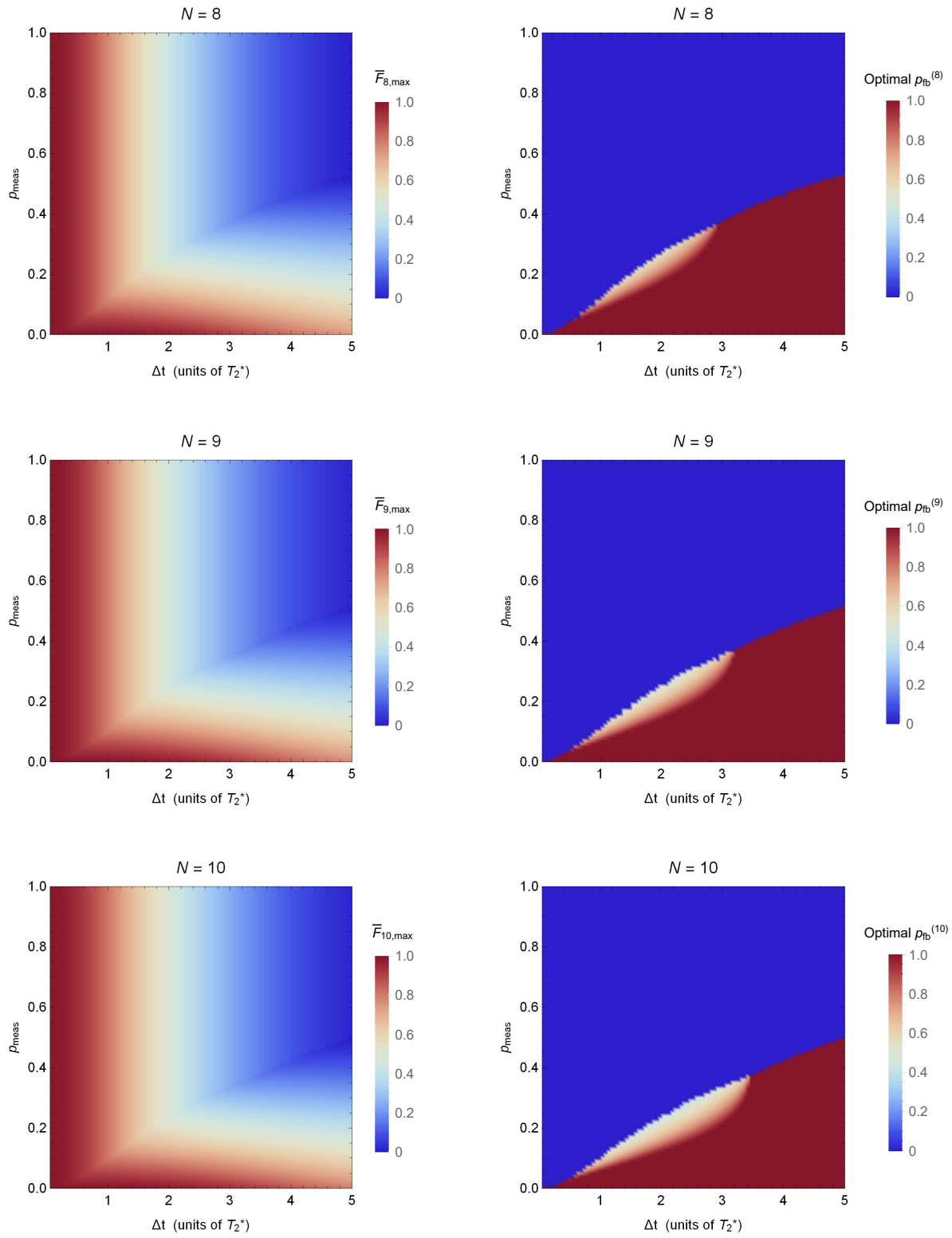



Figure B-1: The optimal $\overline{F}_N$ for each $1 \leq N \leq 10$ separately (left panels), and the corresponding $p_{\text{fb}}^{(N)}$ at which this fidelity is achieved (right panels). Note that the color bars in the left panels have a different scale than that in Fig. 3-3.

# Bibliography


[1] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. Brandao, D. A. Buell, *et al.*, Nature **574**, 505 (2019).

[2] J. Preskill, Quantum **2**, 79 (2018).

[3] A. Kay, arXiv:1809.03842 (2018).

[4] S. Nakajima, Prog. Theor. Phys. **20**, 948 (1958).

[5] R. Zwanzig, J. Chem. Phys. **33**, 1338 (1960).

[6] K. Noh, S. Girvin, and L. Jiang, arXiv:1903.12615 (2019).

[7] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, New York, NY, USA, 2000).

[8] H. K. Ng and P. Mandayam, Phys. Rev. A **81**, 062342 (2010).

[9] S. P. Jordan, E. Farhi, and P. W. Shor, Phys. Rev. A **74**, 052322 (2006).

[10] A. D. Bookatz, E. Farhi, and L. Zhou, Phys. Rev. A **92**, 022317 (2015).

[11] I. Marvian, arXiv:1602.03251 (2016).

[12] M. Marvian and D. A. Lidar, Phys. Rev. A **95**, 032302 (2017).

[13] G. Lindblad, Commun. Math. Phys. **48**, 119 (1976).

[14] H.-P. Breuer and F. Petruccione, *The theory of open quantum systems* (Oxford University Press on Demand, 2002).

[15] D. A. Lidar, arXiv:1902.00967 (2019).

[16] J. Dalibard, Y. Castin, and K. Mølmer, Phys. Rev. Lett. **68**, 580 (1992).

[17] H. J. Carmichael, Phys. Rev. Lett. **70**, 2273 (1993).

[18] C. Bény, Phys. Rev. Lett. **107**, 080501 (2011).

[19] D. Lidar and T. Brun, *Quantum Error Correction* (Cambridge University Press, 2013).

[20] D. Gottesman, arXiv:9705052 (1997).

[21] R. Laflamme, C. Miquel, J. P. Paz, and W. H. Zurek, Phys. Rev. Lett. **77**, 198 (1996).

[22] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Phys. Rev. A **54**, 3824 (1996).

[23] R. Cleve, A. Ekert, C. Macchiavello, and M. Mosca, Proc. R. Soc. A **454**, 339 (1998).

[24] D. W. Leung, M. A. Nielsen, I. L. Chuang, and Y. Yamamoto, Phys. Rev. A **56**, 2567 (1997).

[25] A. S. Fletcher, P. W. Shor, and M. Z. Win, IEEE Trans. Inf. Theory **54**, 5705 (2008).

[26] D. K. Tuckett, S. D. Bartlett, and S. T. Flammia, Phys. Rev. Lett. **120**, 050505 (2018).

[27] D. K. Tuckett, A. S. Darmawan, C. T. Chubb, S. Bravyi, S. D. Bartlett, and S. T. Flammia, Phys. Rev. X **9**, 041031 (2019).

[28] S. Puri, L. St-Jean, J. A. Gross, A. Grimm, N. Frattini, P. S. Iyer, A. Krishna, S. Touzard, L. Jiang, A. Blais, *et al.*, arXiv:1905.00450 (2019).

[29] N. Ofek, A. Petrenko, R. Heeres, P. Reinhold, Z. Leghtas, B. Vlastakis, Y. Liu, L. Frunzio, S. Girvin, L. Jiang, *et al.*, Nature **536**, 441 (2016).

[30] M. Hirose and P. Cappellaro, Nature **532**, 77 (2016).

[31] A. Dréau, J.-R. Maze, M. Lesik, J.-F. Roch, and V. Jacques, Phys. Rev. B **85**, 134107 (2012).

[32] L. Jiang, J. S. Hodges, J. R. Maze, P. Maurer, J. M. Taylor, D. G. Cory, P. R. Hemmer, R. L. Walsworth, A. Yacoby, A. S. Zibrov, and M. D. Lukin, Science **326**, 267 (2009).

[33] M. Chen, W. K. C. Sun, K. Saha, J.-C. Jaskula, and P. Cappellaro, New J. Phys. **20**, 063011 (2018).

[34] P. C. Maurer, G. Kucsko, C. Latta, L. Jiang, N. Y. Yao, S. D. Bennett, F. Pastawski, D. Hunger, N. Chisholm, M. Markham, D. J. Twitchen, J. I. Cirac, and M. D. Lukin, Science **336**, 1283 (2012).

[35] J. H. Shim, I. Niemeyer, J. Zhang, and D. Suter, Phys. Rev. A **87**, 012301 (2013).

[36] S. Zaiser, T. Rendler, I. Jakobi, T. Wolf, S.-Y. Lee, S. Wagner, V. Bergholm, T. Schulte-Herbrüggen, P. Neumann, and J. Wrachtrup, Nat. Commun. **7**, 12279 (2016).

[37] P. Bertet, I. Chiorescu, G. Burkard, K. Semba, C. J. P. M. Harmans, D. P. DiVincenzo, and J. E. Mooij, Phys. Rev. Lett. **95**, 257002 (2005).

[38] P. Bertet, I. Chiorescu, C. Harmans, and J. Mooij, arXiv:0507290 (2005).

[39] J. Gambetta, A. Blais, D. I. Schuster, A. Wallraff, L. Frunzio, J. Majer, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Phys. Rev. A **74**, 042318 (2006).

[40] J. Majer, J. Chow, J. Gambetta, J. Koch, B. Johnson, J. Schreier, L. Frunzio, D. Schuster, A. Houck, A. Wallraff, *et al.*, Nature **449**, 443 (2007).

[41] A. A. Clerk and D. W. Utami, Phys. Rev. A **75**, 042302 (2007).

[42] A. P. Sears, A. Petrenko, G. Catelani, L. Sun, H. Paik, G. Kirchmair, L. Frunzio, L. I. Glazman, S. M. Girvin, and R. J. Schoelkopf, Phys. Rev. B **86**, 180504 (2012).

[43] F. Yan, S. Gustavsson, A. Kamal, J. Birenbaum, A. P. Sears, D. Hover, T. J. Gudmundsen, D. Rosenberg, G. Samach, S. Weber, *et al.*, Nat. Commun. **7**, 12964 (2016).

[44] J.-H. Yeh, J. LeFebvre, S. Premaratne, F. C. Wellstood, and B. S. Palmer, J. Appl. Phys. **121**, 224501 (2017).

[45] F. Yan, D. Campbell, P. Krantz, M. Kjaergaard, D. Kim, J. L. Yoder, D. Hover, A. Sears, A. J. Kerman, T. P. Orlando, S. Gustavsson, and W. D. Oliver, Phys. Rev. Lett. **120**, 260504 (2018).

[46] Z. Wang, S. Shankar, Z. Minev, P. Campagne-Ibarcq, A. Narla, and M. Devoret, Phys. Rev. Applied **11**, 014031 (2019).

[47] S. Machlup, J. Appl. Phys. **25**, 341 (1954).

[48] C. Neuenhahn, B. Kubala, B. Abel, and F. Marquardt, Phys. Status Solidi B **246**, 1018 (2009).

[49] L. Viola and S. Lloyd, Phys. Rev. A **58**, 2733 (1998).

[50] M. Ban, J. Mod. Opt **45**, 2315 (1998).

[51] M. J. Biercuk, H. Uys, A. P. VanDevender, N. Shiga, W. M. Itano, and J. J. Bollinger, Nature **458**, 996 (2009).

[52] D. Layden, S. Zhou, P. Cappellaro, and L. Jiang, Phys. Rev. Lett. **122**, 040502 (2019).

[53] D. C. McKay, C. J. Wood, S. Sheldon, J. M. Chow, and J. M. Gambetta, Phys. Rev. A **96**, 022330 (2017).

[54] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, and S. J. Glaser, J. Magn. Reson. **172**, 296 (2005).

[55] P. de Fouquieres, S. Schirmer, S. Glaser, and I. Kuprov, J. Magn. Reson. **212**, 412 (2011).

[56] S. Lloyd and L. Viola, Phys. Rev. A **65**, 010101 (2001).

[57] C. Shen, K. Noh, V. V. Albert, S. Krastanov, M. H. Devoret, R. J. Schoelkopf, S. M. Girvin, and L. Jiang, Phys. Rev. B **95**, 134501 (2017).

[58] D. Nigg, M. Müller, E. A. Martinez, P. Schindler, M. Hennrich, T. Monz, M. A. Martin-Delgado, and R. Blatt, Science **345**, 302 (2014).

[59] S. Rosenblum, P. Reinhold, M. Mirrahimi, L. Jiang, L. Frunzio, and R. J. Schoelkopf, Science **361**, 266 (2018).

[60] D. Layden and P. Cappellaro, npj Quantum Inf. **4**, 30 (2018).

[61] C. D. Aiello and P. Cappellaro, Phys. Rev. A **91**, 042340 (2015).

[62] Y. Ouyang and C.-Y. Lai, arXiv:2001.03976 (2020).

[63] M. Reimpell and R. F. Werner, Phys. Rev. Lett. **94**, 080501 (2005).

[64] A. S. Fletcher, P. W. Shor, and M. Z. Win, Phys. Rev. A **75**, 012338 (2007).

[65] R. L. Kosut, A. Shabani, and D. A. Lidar, Phys. Rev. Lett. **100**, 020502 (2008).

[66] S. Taghavi, R. L. Kosut, and D. A. Lidar, IEEE Trans. Inf. Theory **56**, 1461 (2010).

[67] P. D. Johnson, J. Romero, J. Olson, Y. Cao, and A. Aspuru-Guzik, arXiv:1711.02249 (2017).

[68] K. Noh, V. V. Albert, and L. Jiang, IEEE Trans. Inf. Theory **65**, 2563 (2019).

[69] L. Ioffe and M. Mézard, Phys. Rev. A **75**, 032345 (2007).

[70] C. Cafaro and P. van Loock, Phys. Rev. A **89**, 022316 (2014).

[71] A. Robertson, C. Granade, S. D. Bartlett, and S. T. Flammia, Phys. Rev. Applied **8**, 064004 (2017).

[72] D. Layden, M. Chen, and P. Cappellaro, Phys. Rev. Lett. **124**, 020504 (2020).

[73] N. Yamamoto, S. Hara, and K. Tsumura, Phys. Rev. A **71**, 022322 (2005).

[74] A. S. Fletcher, P. W. Shor, and M. Z. Win, Phys. Rev. A **77**, 012320 (2008).

[75] R. L. Kosut and D. A. Lidar, Quantum Inf. Process **8**, 443 (2009).

[76] C. Bény and O. Oreshkov, Phys. Rev. Lett. **104**, 120501 (2010).

[77] J. Tyson, J. Math. Phys. **51**, 092204 (2010).

[78] C. Bény and O. Oreshkov, Phys. Rev. A **84**, 022333 (2011).

[79] M. H. Michael, M. Silveri, R. Brierley, V. V. Albert, J. Salmilehto, L. Jiang, and S. M. Girvin, Phys. Rev. X **6**, 031006 (2016).

[80] V. V. Albert, K. Noh, K. Duivenvoorden, D. J. Young, R. T. Brierley, P. Reinhold, C. Vuillot, L. Li, C. Shen, S. M. Girvin, B. M. Terhal, and L. Jiang, Phys. Rev. A **97**, 032346 (2018).

[81] L. Li, D. J. Young, V. V. Albert, K. Noh, C.-L. Zou, and L. Jiang, arXiv:1901.05358 (2019).

[82] G. Balló and P. Gurin, Phys. Rev. A **80**, 012326 (2009).

[83] L. Hu, Y. Ma, W. Cai, X. Mu, Y. Xu, W. Wang, Y. Wu, H. Wang, Y. Song, C.-L. Zou, *et al.*, Nat. Phys. **15**, 503 (2019).

[84] B. Rahn, A. C. Doherty, and H. Mabuchi, Phys. Rev. A **66**, 032304 (2002).

[85] C. Chamberland, J. Wallman, S. Beale, and R. Laflamme, Phys. Rev. A **95**, 042332 (2017).

[86] W. M. Witzel, M. S. Carroll, A. Morello, L. Cywiński, and S. Das Sarma, Phys. Rev. Lett. **105**, 187602 (2010).

[87] H. Bluhm, S. Foletti, I. Neder, M. Rudner, D. Mahalu, V. Umansky, and A. Yacoby, Nat. Phys. **7**, 109 (2011).

[88] M. W. Doherty, N. B. Manson, P. Delaney, F. Jelezko, J. Wrachtrup, and L. C. Hollenberg, Phys. Rep. **528**, 1 (2013).

[89] J. T. Muhonen, J. P. Dehollain, A. Laucht, F. E. Hudson, R. Kalra, T. Sekiguchi, K. M. Itoh, D. N. Jamieson, J. C. McCallum, A. S. Dzurak, *et al.*, Nat. Nanotechnol. **9**, 986 (2014).

[90] J.-L. Orgiazzi, C. Deng, D. Layden, R. Marchildon, F. Kitapli, F. Shen, M. Bal, F. R. Ong, and A. Lupascu, Phys. Rev. B **93**, 104518 (2016).

[91] J. Kelly, R. Barends, A. G. Fowler, A. Megrant, E. Jeffrey, T. C. White, D. Sank, J. Y. Mutus, B. Campbell, Y. Chen, *et al.*, Nature **519**, 66 (2015).

[92] D. Riste, S. Poletto, M.-Z. Huang, A. Bruno, V. Vesterinen, O.-P. Saira, and L. DiCarlo, Nat. Commun. **6**, 6983 (2015).

[93] P. Schindler, J. T. Barreiro, T. Monz, V. Nebendahl, D. Nigg, M. Chwalla, M. Hennrich, and R. Blatt, Science **332**, 1059 (2011).

[94] J. Cramer, N. Kalb, M. A. Rol, B. Hensen, M. S. Blok, M. Markham, D. J. Twitchen, R. Hanson, and T. H. Taminiau, Nat. Commun. **7**, 11526 (2016).

[95] D. P. DiVincenzo and P. Aliferis, Phys. Rev. Lett. **98**, 020501 (2007).

[96] C. Chamberland, P. Iyer, and D. Poulin, Quantum **2**, 43 (2018).

[97] A. Degasperis, L. Fonda, and G. C. Ghirardi, Il Nuovo Cimento A (1965-1970) **21**, 471 (1974).

[98] B. Misra and E. C. G. Sudarshan, J. Math. Phys. **18**, 756 (1977).

[99] L. Vaidman, L. Goldenberg, and S. Wiesner, Phys. Rev. A **54**, R1745 (1996).

[100] N. Erez, Y. Aharonov, B. Reznik, and L. Vaidman, Phys. Rev. A **69**, 062315 (2004).

[101] R. Kubo, J. Phys. Soc. Jpn. **17**, 1100 (1962).

[102] W. M. Itano, D. J. Heinzen, J. J. Bollinger, and D. J. Wineland, Phys. Rev. A **41**, 2295 (1990).

[103] M. C. Fischer, B. Gutiérrez-Medina, and M. G. Raizen, Phys. Rev. Lett. **87**, 040402 (2001).

[104] J. Bernu, S. Deléglise, C. Sayrin, S. Kuhr, I. Dotsenko, M. Brune, J. M. Raimond, and S. Haroche, Phys. Rev. Lett. **101**, 180402 (2008).

[105] F. Schäfer, I. Herrera, S. Cherukattil, C. Lovecchio, F. S. Cataliotti, F. Caruso, and A. Smerzi, Nat. Commun. **5**, 3194 (2014).

[106] A. Signoles, A. Facon, D. Grosso, I. Dotsenko, S. Haroche, J.-M. Raimond, M. Brune, and S. Gleyzes, Nat. Phys. **10**, 715 (2014).

[107] D. Suter and G. A. Álvarez, Rev. Mod. Phys. **88**, 041001 (2016).

[108] C. L. Degen, F. Reinhard, and P. Cappellaro, Rev. Mod. Phys. **89**, 035002 (2017).

[109] M. J. Biercuk, A. C. Doherty, and H. Uys, J. Phys. B **44**, 154002 (2011).

[110] D. Layden, L. R. Huang, and P. Cappellaro, Quantum Sci. Technol. (2020).

[111] A. G. Fowler, A. C. Whiteside, and L. C. L. Hollenberg, Phys. Rev. Lett. **108**, 180501 (2012).

[112] E. T. Campbell, Quantum Sci. Technol. **4**, 025006 (2019).

[113] P. Iyer and D. Poulin, Quantum Sci. Technol. **3**, 030504 (2018).

[114] A. Fletcher, *Channel-adapted quantum error correction*, Ph.D. thesis, Massachusetts Institute of Technology (2007).

[115] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski, *Robust optimization*, Vol. 28 (Princeton University Press, 2009).

[116] K. M. Teo, *Nonconvex robust optimization*, Ph.D. thesis, Massachusetts Institute of Technology (2007).

[117] O.-P. Saira, V. Bergholm, T. Ojanen, and M. Möttönen, Phys. Rev. A **75**, 012308 (2007).

[118] J. Helm and W. T. Strunz, Phys. Rev. A **80**, 042108 (2009).

[119] D. Crow and R. Joynt, Phys. Rev. A **89**, 042123 (2014).

[120] B. Gu and I. Franco, J. Chem. Phys. **151**, 014109 (2019).

[121] R. Kubo, M. Toda, and N. Hashitsume, *Statistical Physics II: Nonequilibrium Statistical Mechanics*, Springer Series in Solid-State Sciences (Springer Berlin Heidelberg, 2012).

[122] B. S. Abel, *Macroscopic superposition states and decoherence by quantum telegraph noise*, Ph.D. thesis, LMU Munich (2008).

[123] A. Sears, *Extending Coherence in Superconducting Qubits: From Microseconds to Milliseconds*, Ph.D. thesis, Yale University (2013).

[124] J. Preskill, "Notes on noise," .

[125] E. Parzen, *Stochastic Processes*, Classics in Applied Mathematics (Society for Industrial and Applied Mathematics (Philadelphia, PA), 1999).

[126] L. M. Norris, G. A. Paz-Silva, and L. Viola, Phys. Rev. Lett. **116**, 150503 (2016).

[127] Y. Sung, F. Beaudoin, L. M. Norris, F. Yan, D. K. Kim, J. Y. Qiu, U. von Lüpke, J. L. Yoder, T. P. Orlando, S. Gustavsson, *et al.*, Nat. Commun. **10**, 1 (2019).

[128] N. Wiener, Acta Math. **55**, 117 (1930).

[129] A. Khintchine, Math. Ann **109**, 604 (1934).

[130] J. Schriefl, Y. Makhlin, A. Shnirman, and G. Schön, New J. Phys. **8**, 1 (2006).

[131] P. Dutta and P. M. Horn, Rev. Mod. Phys. **53**, 497 (1981).

[132] A. Ajoy, Y. Liu, and P. Cappellaro, arXiv:1611.04691 (2016).

[133] V. Giovannetti, S. Lloyd, and L. Maccone, Phys. Rev. Lett. **96**, 010401 (2006).

[134] V. Giovannetti, S. Lloyd, and L. Maccone, Nat. Photonics **5**, 222 (2011).

[135] R. A. Fisher, Math. Proc. Camb. Philos. Soc. **22**, 700–725 (1925).

[136] H. Cramér, *Mathematical Methods of Statistics*, Princeton Landmarks in Mathematics and Physics (Princeton University Press, 1999).

[137] C. R. Rao, in *Breakthroughs in statistics* (Springer, 1992) pp. 235–247.

[138] S. L. Braunstein, J. Phys. A **25**, 3813 (1992).

[139] S. L. Braunstein, Phys. Rev. Lett. **69**, 3598 (1992).

[140] S. L. Braunstein and C. M. Caves, Phys. Rev. Lett. **72**, 3439 (1994).

[141] M. G. A. Paris, Int. J. Quantum Inf. **07**, 125 (2009).

[142] S. F. Huelga, C. Macchiavello, T. Pellizzari, A. K. Ekert, M. B. Plenio, and J. I. Cirac, Phys. Rev. Lett. **79**, 3865 (1997).

[143] J. Taylor, P. Cappellaro, L. Childress, L. Jiang, D. Budker, P. Hemmer, A. Yacoby, R. Walsworth, and M. Lukin, Nat. Phys. **4**, 810 (2008).

[144] M. Bal, C. Deng, J.-L. Orgiazzi, F. Ong, and A. Lupascu, Nat. Commun. **3**, 1 (2012).

[145] G. Arrad, Y. Vinkler, D. Aharonov, and A. Retzker, Phys. Rev. Lett. **112**, 150801 (2014).

[146] E. M. Kessler, I. Lovchinsky, A. O. Sushkov, and M. D. Lukin, Phys. Rev. Lett. **112**, 150802 (2014).

[147] R. Ozeri, arXiv:1310.3432 (2013).

[148] W. Dür, M. Skotiniotis, F. Fröwis, and B. Kraus, Phys. Rev. Lett. **112**, 080801 (2014).

[149] P. R. Chernoff, J. Funct. Anal. **2**, 238 (1968).

[150] D. Layden, E. Martín-Martínez, and A. Kempf, Phys. Rev. A **93**, 040301 (2016).

[151] R. Demkowicz-Dobrzański, J. Czajkowski, and P. Sekatski, Phys. Rev. X **7**, 041009 (2017).

[152] S. Zhou, M. Zhang, J. Preskill, and L. Jiang, Nat. Commun. **9**, 78 (2018).

[153] F. Reiter, A. S. Sørensen, P. Zoller, and C. Muschik, Nat. Commun. **8**, 1822 (2017).

[154] O. Oreshkov and T. A. Brun, Phys. Rev. A **76**, 022318 (2007).

[155] D. A. Herrera-Martí, T. Gefen, D. Aharonov, N. Katz, and A. Retzker, Phys. Rev. Lett. **115**, 200501 (2015).

[156] M. B. Plenio and S. F. Huelga, Phys. Rev. A **93**, 032123 (2016).

[157] T. Gefen, D. A. Herrera-Martí, and A. Retzker, Phys. Rev. A **93**, 032133 (2016).

[158] M. Bergmann and P. van Loock, Phys. Rev. A **94**, 012311 (2016).

[159] P. Sekatski, M. Skotiniotis, J. Kołodyński, and W. Dür, Quantum **1**, 27 (2017).

[160] T. Unden, P. Balasubramanian, D. Louzon, Y. Vinkler, M. B. Plenio, M. Markham, D. Twitchen, A. Stacey, I. Lovchinsky, A. O. Sushkov, M. D. Lukin, A. Retzker, B. Naydenov, L. P. McGuinness, and F. Jelezko, Phys. Rev. Lett. **116**, 230502 (2016).

[161] L. Cohen, Y. Pilnyak, D. Istrati, A. Retzker, and H. S. Eisenberg, Phys. Rev. A **94**, 012324 (2016).

[162] Y. Matsuzaki and S. Benjamin, Phys. Rev. A **95**, 032303 (2017).

[163] A. Leon-Garcia, *Probability, Statistics, and Random Processes for Electrical Engineering* (Pearson/Prentice Hall, 2008).

[164] D. A. Lidar, I. L. Chuang, and K. B. Whaley, Phys. Rev. Lett. **81**, 2594 (1998).

[165] D. M. Greenberger, M. A. Horne, and A. Zeilinger, "Going beyond bell's theorem," in *Bell's Theorem, Quantum Theory and Conceptions of the Universe* (Springer Netherlands, Dordrecht, 1989) pp. 69–72.

[166] C. F. Roos, M. Chwalla, K. Kim, M. Riebe, and R. Blatt, Nature **443**, 316 (2006).

[167] M. Chwalla, K. Kim, T. Monz, P. Schindler, M. Riebe, C. Roos, and R. Blatt, Appl. Phys. B **89**, 483 (2007).

[168] U. Dorner, New J. Phys. **14**, 043011 (2012).

[169] J. Jeske, J. H. Cole, and S. F. Huelga, New J. Phys. **16**, 073039 (2014).

[170] E. M. Fortunato, L. Viola, J. Hodges, G. Teklemariam, and D. G. Cory, New J. Phys. **4**, 5 (2002).

[171] T. Monz, P. Schindler, J. T. Barreiro, M. Chwalla, D. Nigg, W. A. Coish, M. Harlander, W. Hänsel, M. Hennrich, and R. Blatt, Phys. Rev. Lett. **106**, 130506 (2011).

[172] Y. Romach, C. Müller, T. Unden, L. J. Rogers, T. Isoda, K. M. Itoh, M. Markham, A. Stacey, J. Meijer, S. Pezzagna, B. Naydenov, L. P. McGuinness, N. Bar-Gill, and F. Jelezko, Phys. Rev. Lett. **114**, 017601 (2015).

[173] S. Zhou, personal communication (Oct. 3, 2017).

[174] R. H. Dicke, Phys. Rev. **93**, 99 (1954).

[175] S. Boyd and L. Vandenberghe, *Convex optimization* (Cambridge university press, 2004).

[176] P. Davis, *Circulant matrices*, Pure and applied mathematics (Wiley, 1979).

[177] B. Eastin and E. Knill, Phys. Rev. Lett. **102**, 110502 (2009).

[178] P. Faist, S. Nezami, V. V. Albert, G. Salton, F. Pastawski, P. Hayden, and J. Preskill, arXiv:1902.07714 (2019).

[179] A. Kubica and R. Demkowicz-Dobrzanski, arXiv:2004.11893 (2020).