THE UNIVERSITY OF CHICAGO

# TOWARDS PRACTICAL FAULT-TOLERANT QUANTUM COMPUTING

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE PRITZKER SCHOOL OF MOLECULAR ENGINEERING

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

BY

QIAN XU

CHICAGO, ILLINOIS

JUNE 2024

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# ACKNOWLEDGMENTS

First, I would like to thank my advisor Liang Jiang for his strong support during my PhD study. Under his influence, I learned how to be not only a good researcher but also an understanding and supportive person. I appreciate the environment where I can freely explore new ideas/projects following my interests while getting enough support.

Second, I want to thank all of my amazing collaborators, from whom I learnt various topics across different fields: Pei Zeng, Daohong Xu, Pablo Bonilla, Christopher Pattison, Nithin Raveendran, Dolev Bluvstein, Jonathan Wurtz, Bane Vasić, Mikhail Lukin, Hengyun Zhou, Guo Zheng, Yu-Xin Wang, Peter Zoller, Aashish Clerk, Nam Mannucci, Alireza Seif, Aleksander Kubica, Steven Flammia, Ming Yuan, Alireza Seif, Haoxiong Yan, Nam Mannucci, Bernard Ousmane Sane, Rodney Van Meter, Andrew Cleland, Peter Leviant, Serge Rosenblum, Harald Putterman, Joseph Iverson, Oskar Painter, Fernando Brandão, Kyungjoo Noh, Filip Rozpędek, and Saikat Guha.

Thirdly, I thank my office friends and colleagues for the time we spent together at leisure. I also thank my basketball friends and all the memorable times we had together at the basketball court.

Finally, I sincerely thank my parents, Qinbao Xu, and Guangfang Kan, for their unspeakable personal support. In addition, I truly appreciate the tremendous love and support from my partner, You Zu. I always feel recharged when being together with her and I will always value her accompany in my dark times. I could not imagine that I could make it through the long PhD journey without the love and support of my family and partner.

# ABSTRACT

Quantum computers hold the promise of solving classically intractable problems. However, quantum systems are intrinsically faulty due to the environment-induced noise and decoherence. As such, fault-tolerant quantum computers, which can run a quantum computation successfully even if the physical operations are faulty, are needed for solving utility-scale problems. Building such a fault-tolerant quantum computer requires a fault-tolerant scheme, which replaces target quantum circuits with new circuits that involve extra qubits and more gates but are robust against all physical faults with the help of quantum error correction. The state-of-the-art fault-tolerant scheme, which we refer to as the baseline scheme, encodes qubits into planar surface-code patches and performs encoded logical operations via code deformations and lattice surgeries. However, its large space-time overhead, which requires millions of physical qubits and takes days for large computations, poses a formidable challenge for scaling quantum computing to practical levels. In this thesis, we introduce new fault-tolerant schemes by leveraging new hardware and coding features not previously considered by the baseline scheme. Specifically, we present new schemes for bosonic systems that feature an infinite-dimensional Hilbert space, design noise-tailored topological codes, and investigate low-overhead schemes on hardware with long-range connectivity using the recently developed quantum low-density-parity-check codes. Our results demonstrate that by incorporating these hardware-specific schemes and employing improved quantum codes, we can drastically reduce the overhead barrier towards practical fault-tolerant quantum computing.

# CHAPTER 1

# INTRODUCTION

## 1.1 The field of study: fault-tolerant quantum computing

Unlike their classical counterparts, quantum computers operate on principles of quantum mechanics, leveraging unique properties such as superposition and entanglement [14, 15]. These distinct mechanisms enable quantum computers to tackle problems that are computationally intractable for classical systems. For instance, Shor's seminal work [16] demonstrated that factoring large integers, a task requiring exponential time for classical computers, can be accomplished in polynomial time on a quantum computer. This illustrates the remarkable potential of quantum computing to solve complex problems efficiently.

Successfully executing a large quantum circuit necessitates an extremely low error rate per operation, typically below $10^{-10}$ for practical quantum algorithms [17, 18, 19, 20]. However, unlike classical systems, quantum systems inherently suffer from noise induced by the environment and decoherence [15]. State-of-the-art quantum devices exhibit error rates ranging from approximately 0.1% to 1% [21, 22, 23, 24]. Closing such a significant gap in operational error rates is unlikely achievable solely through reductions in physical error rates. Instead, the field must focus on developing quantum computers capable of supporting *fault-tolerant quantum computing*, wherein computations can succeed despite noisy physical operations. Fortunately, the advancement of *fault-tolerant schemes* has demonstrated the feasibility of this task [25].

A fault-tolerant scheme replaces a quantum circuit designed on bare physical qubits with a larger circuit capable of "simulating" the original circuit [26]. This simulation aims to produce the same output statistics with a high success probability, even in the presence of physical noise. The cornerstone of all fault-tolerant schemes lies in *quantum error correction* (QEC) [27, 28, 15], which actively corrects physical errors by introducing extra redundant

qubits, thereby expanding the circuit's width. In this thesis, our focus will be on code-centric fault-tolerant schemes, which are built upon specific QEC codes (see Sec.1.2 and Sec.1.3 for an introduction to QEC codes).

At a high level, a fault-tolerant scheme based on a QEC code $\mathcal{C}$ operates as follows:

1. Each qubit of the original circuit $C$ is replaced with a logical qubit encoded in the code $\mathcal{C}$.

2. Each operation of $C$, such as state preparation, unitary gates, and measurements, performed on physical qubits, is replaced with an $\mathcal{C}$-encoded and error-corrected logical operation on the corresponding logical qubits.

It's important to note that such a code-centric scheme is independent of the specific circuit $C$, and the key task lies in designing a universal set of error-corrected logical operations for the code $\mathcal{C}$. This differs from recent circuit-centric fault-tolerant schemes [29, 30, 25, 31, 32], which do not rely on a specific QEC code initially. Instead, they aim to directly design a new circuit $C'$ capable of error correction by leveraging its intrinsic redundancy [30].

It's worth emphasizing that designing a universal set of error-corrected operations is distinct from conventionally defined QEC, which typically realizes error-corrected idling operations (for quantum memory) assuming perfect error-recovery operations. Designing fully error-corrected (or fault-tolerant) operations capable of tolerating all operational faults and extending beyond the memory level is significantly more complex. For instance, the simplest fault-tolerant operation is a logical gate implemented by transversal physical gates (where each qubit is involved in only one physical gate), followed by a full round of memory QEC. While such a logical operation is trivially fault-tolerant, as minimal errors are injected and propagated by the transversal gates, which can then be corrected by the QEC gadget (assuming it's also fault-tolerant), it's important to note that transversal gates do not generate a universal gate set, as per the Eastin–Knill theorem [33]. Hence, further efforts are required for developing code-centric fault-tolerant schemes. For more details about the philosophy,

recent developments, challenges, and opportunities of fault-tolerance, readers are referred to Ref. [25].

In Sec.1.4, we will introduce a paradigmatic fault-tolerant scheme based on the surface code[34, 35, 17]. This scheme has been a leading candidate for fault-tolerant quantum computing, owing to several advantages: a relatively high noise threshold, compatibility with geometrically local operations, well-developed logical operations, a clear understanding of resource requirements, and efficient circuit compiling for implementing logical algorithms. We will refer to this scheme as the *baseline fault-tolerant scheme*.

Early proof-of-principle demonstrations of the baseline scheme have been realized in state-of-the-art devices (see, e.g., Refs.[21, 36, 37]). However, as we will discuss in detail in Sec.1.4, significant challenges exist in scaling up quantum computers using the baseline scheme.

Therefore, the primary goal of this thesis is to address the challenges posed by the baseline scheme and propose new fault-tolerant schemes capable of surpassing its performance.

## 1.2 Quantum error correction

The paradigmatic problem of quantum error correction can be framed as transmitting quantum information reliably through some noisy channel $\mathcal{N}_\gamma$, which is a completely positive and trace-preserving (CPTP) map between quantum states (see Sec. 2.1.1 for a more elementary introduction of quantum channels). The general strategy is as follows [38]:

1. Encode the source Hilbert space $\mathcal{H}_S$ carrying the information redundantly into a larger physical Hilbert space $\mathcal{H}$ via some encoding channel: $\mathcal{E}_\mathcal{C} : \mathcal{L}(\mathcal{H}_S) \to \mathcal{L}(\mathcal{H})$, where $\mathcal{L}(\bullet)$ represents the space of lienar operators on $\bullet$. Note that this encoding channel $\mathcal{E}_\mathcal{C}$ uniquely determines a QEC code $\mathcal{C}$.

2. Let the encoded quantum state transmit through the noisy channel $\mathcal{N}_\gamma : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H})$, where $\gamma$ is some parameter characterizing the noise strength.

3. Apply a recovery channel $\mathcal{R} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H})$ that can, to the best, cancel the effect of $\mathcal{N}_\gamma$.

4. Apply a decoding channel $\mathcal{D} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H}_S)$, which is typically the inverse of the encoding channel, to dictate the encoded information back into the source Hilbert space.

The composite channel

$$\mathcal{N} := \mathcal{D} \circ \mathcal{R} \circ \mathcal{N}_\gamma \circ \mathcal{E}_\mathcal{C} \tag{1.1}$$

finally maps from $\mathcal{H}_S$ onto $\mathcal{H}_S$ and should be close to an identity channel $\mathcal{I}_{\mathcal{H}_S}$ if the noise rate $\gamma$ is not too large. Note that in order for the recovery to be successful, it is essential to use a Hilbert space $\mathcal{H}$ that has a larger dimension than the source space $\mathcal{H}_S$ (this is the so-called redundancy).

There are different ways of performing the encoding, depending on the code we choose and the physical system we encode the information into. For example, we have $\mathcal{H}_S = \mathbb{C}_2$ and $\mathcal{H} = \mathbb{C}_2^{\otimes n}$ for a $[[n, 1, d]]$-stabilizer qubit code that encodes a single logical qubit into $n$ physical qubits, whereas $\mathcal{H} = \mathcal{H}_{\text{CV}}$ (the infinite-dimensional Hilbert space of an oscillator) for a single-mode bosonic code that encodes a single logical qubit into an oscillator (as we will introduce in Chapter 2).

## 1.3 Stabilizer codes

In this section, we briefly review the basics of a representative family of QEC code, called stabilizer codes, that admit a particularly simple encoding and recovery channel (see Eq. (1.1)). The purpose of this section is mainly to introduce the notation and terminologies that we will use later in this thesis. For more details about the stabilizer codes, readers are referred to Ref. [39].

A $[[n, k, d]]$-stabilizer encodes a $2^k$-dimensional source (or logical) space (can be associated

with $k$ logical qubits) into the $2^n$-dimensional space of $n$ qubits. Specifically, we define a $2^k$-dimensional subspace of the $n$-qubit Hilbert space as the common $+1$ eigenspace of an Abelian stabilizer group $S \subset \mathcal{P}^n$ that does not contain $-I$, where $\mathcal{P}^n$ denotes the $n$-qubit Pauli group. The stabilizer group $S$ is generated by a set of $r = n - k$ stabilizer generators, i.e. $S = \langle S_1, S_2, \cdots, S_r \rangle$. We also refer interchangeably to these generators as checks. A CSS stabilizer code is a stabilizer code with stabilizer generators that can be partitioned into $Z$-types (only containing Pauli $Z$ operators) and $X$-types (only containing Pauli $X$ operators). The $Z$ and $X$ checks can be represented as binary check matrices $H_Z$ and $H_X$, respectively, where the columns are associated with the qubits and each row is associated with a stabilizer generator. The commutativity of the stabilizers requires that $H_Z H_X^T = 0 \bmod 2$.

Given a stabilizer code with a stabilizer group $S$, the nontrivial logical operators are given by

$$\mathcal{L} = N(S) \backslash S, \tag{1.2}$$

where $N(\bullet)$ denotes the normalizer of $\bullet$ with respect to the Pauli group. The distance $d$ of the code is defined as the minimum Hamming weight of the nontrivial logical operators, i.e. $d := \min_{L \in \mathcal{L}} |L|$.

The stabilizer codes are designed against stochastic Pauli error channels:

$$\mathcal{N}_p := \sum_{Q \in \mathcal{P}_n} p(Q) Q \bullet Q^\dagger \tag{1.3}$$

where $p$ is some probability distribution of different Pauli errors. For such a Pauli channel, the recovery channel for the stabilizer codes is particularly simple: (1) One first measures all the $r$ stabilizer generators and records the outcomes as an error syndrome $s \in \mathbb{F}_2^r$. This corresponds to projecting the system to a certain error subspace. (2) One applied a Pauli correction that is the output of some decoder $D : \mathbb{F}_2^r \to \mathcal{P}_n$, which decodes from the syndrome $s$. The total recovery channel is thus the sum of all the projections over all the error subspaces

followed by the corresponding feedback correction. If using an optimal decoder, the code can correct any Pauli error with a Hamming weight $\leq \lfloor (d-1)/2 \rfloor$.

## 1.4 The baseline fault-tolerant scheme with surface codes

In this section, we briefly sketch the baseline fault-tolerant scheme using surface codes and physical qubits. Please refer to Refs. [18, 40, 20] for more details.

As illustrated in Fig. 45 of Ref. [18], the baseline scheme performs fault-tolerant quantum computing on a $2D$ grid of physical qubits with only nearest-neighbor couplings. The logical qubits are encoded using tiles of surface code patches. See Sec. 3.1.2 for more details about surface codes. Single-qubit logical operations can be performed by deforming each patch. Two-qubit entangling operations between two logical qubits can be performed by first bringing the patches to adjacent positions (via continuous patch deformation) and then merging the corresponding patch boundaries.

Such a baseline scheme is appealing due to its various advantages, such as: (1) It only requires local nearest-neighbor couplings between qubits. (2) It has a relatively high threshold ($\sim 1\%$ [17, 41, 42]) for the entire computation. (3) The logical operations are well understood and can even be simply described as a game of deforming and merging different piles of qubits [10, 18]. (4) Fault tolerance is easily understood using the topological picture, i.e. fault tolerance is achieved as long as boundaries supporting nonlocal logical errors are well separated in both space and time. (5) Decoding is easy and efficient: the error syndromes are in the form of a stream of defects in space-time tubes and decoding is performed by simply pairing these space-time defects using the efficient matching algorithm [17, 43].

However, there are also severe challenges when applying this baseline scheme for large-scale fault-tolerant quantum computing, i.e. involving thousands of logical qubits and $> 10^{10}$ gates:

1. The threshold ($\sim 1\%$) is not high enough for near-term devices with bare physical

qubits. To quickly suppress the logical errors, the physical error rate should be not only below the threshold but also deeply below the threshold. Since state-of-the-art devices have error rates only slightly below 1% [21, 22, 23], further increase of the code threshold or reduction of physical error rates are in demand.

2. **The scheme comes with a large space-time overhead**. Given an algorithmic circuit with $k$ logical qubits and $T$ non-Clifford gates, the baseline scheme has a "logical space-time volume", defined as the number of logical qubits times the number of logical cycles (most of the bottom-layer logical operations, e.g. a Pauli product measurement, take one logical cycles), roughly $2kT$ [18, 40]. Let $d$ denote the distance of each surface code patch. Since each logical qubit requires $\Omega(d^2)$ physical qubits, and each logical cycle takes a time of $\Omega(d)$, the total space-time volume of the fault-tolerant circuit, i.e. the number of physical qubits times the runtime of the circuit, is given $\Omega(kTd^3)$. As $k$ and $T$ increase with the size of the computation, $d$ also needs to increase to sufficiently suppress the total failure rate of the circuit. This translates to enormous space-time overhead due to the $d^3$ factor. State-of-the-art estimates show that implementing useful quantum algorithms, e.g. the factoring algorithm, requires millions of qubits and takes days.

## 1.5 Overview of the thesis and reading guide

As discussed in Sec. 1.4, the baseline fault-tolerant scheme using surface codes faces major challenges and, in particular, its large space-time overhead makes large-scale quantum computing a formidable task.

To improve upon the baseline fault-tolerant scheme, an important observation is that there are several strong constraints (or standard assumptions) people have put on the hardware and the QEC code when designing the scheme and estimating its resource overhead. As shown in Table 1.1, it is assumed that the base physical units are simply physical qubits

with two energy levels, the noise model is the symmetric depolarizing noise model (with equally distributed Pauli error rates), the connectivity between qubits is constrained to nearest-neighbor couplings, and the code (surface code) has vanishing encoding rate (the ratio between the encoded logical qubits and the block length) asymptotically. The hardware and coding assumptions reflect the constraints for the leading physical devices (superconducting circuits with transmon qubits [44, 21]) and the leading QEC code (surface code), respectively. However, as various new types of hardware advance with distinct characteristics and noise profiles, we might be able to design new hardware-tailored fault-tolerant schemes that outperform the baseline. Also, new breakthroughs in coding theory and fault-tolerant designs may enable us to explore better schemes with drastically better QEC codes. The new hardware and coding features that we exploit in this thesis to design better fault-tolerant schemes are summarized in Table 1.1.

| | Base physical units | Noise model | Connectivity | Encoding rate |
|---|---|---|---|---|
| Baseline scheme | Bare qubits (two-level) | Depolarizing noise | Nearest-neighbor | Vanishing |
| New schemes | Oscillators (infinite-dimensional) | Biased noise | Long-range | Constant |

Table 1.1: Summary of key hardware and coding features that we utilize in the thesis to design new fault-tolerant schemes that improve upon the baseline scheme.

An overview of the thesis structure is shown in Fig. 1.1.

In Chapter 2, we will consider oscillators with an infinite-dimensional Hilbert space as the base physical units. Compared to bare two-level qubits, an oscillator intrinsically provides the Hilbert space redundancy required for QEC (see Sec. 1.2). We exploit such redundancy and design bosonic QEC schemes that are more hardware-efficient, i.e. consuming fewer physical units, than the baseline.

In Chapter 3, we will consider physically relevant noise models that are more structured than the depolarizing noise. For instance, the so-called bosonic qubits [38, 45, 4, 46] have a

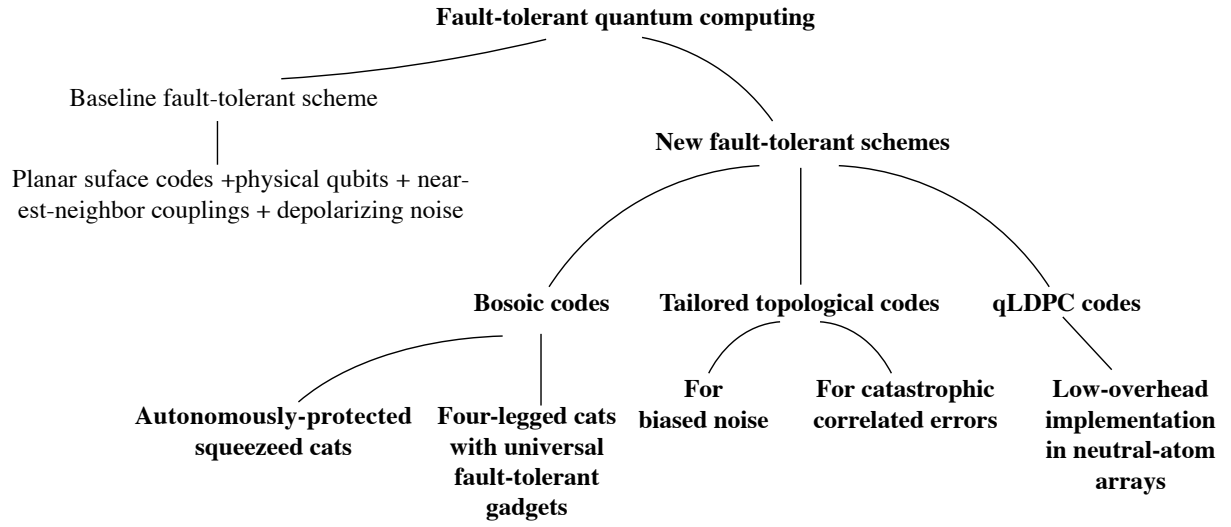


Figure 1.1: Structure of the thesis.

highly biased Pauli noise channel, i.e. one type of Pauli error is dominant over the rest. We exploit such noise structures and design noise-tailored topological codes with significantly boosted performance.

In Chapter 4, we consider non-local hardware connectivity, which can be realized in recently-advanced platforms such as reconfigurable neutral arrays [47, 22]. We show that when combining the non-local connectivity with the recently developed quantum low-density-parity-check (qLDPC) codes featuring a constant encoding rate [12, 48], we can design fault-tolerant schemes with significantly lower space overhead.

Finally, in Chapter 5, we provide concluding remarks and discuss caveats of our schemes and alternative future pathways toward fault-tolerant quantum computing.

# CHAPTER 2

# BOSONIC CODES

In this chapter, we will explore novel fault-tolerant schemes utilizing bosonic systems, or harmonic oscillators [38]. Compared to the baseline scheme using surface codes with physical qubits, these bosonic fault-tolerant schemes offer several advantages:

1. Hardware Efficiency: QEC in bosonic systems can be more hardware-efficient. Unlike physical two-level systems, an oscillator possesses an infinite-dimensional Hilbert space, allowing for redundant encoding of error-protected logical qudits into single physical units. Note that logical information can also be encoded collectively into many oscillators, which can further boost the QEC performance [49, 50], or even enable an error-corrected logical oscillator [51].

2. Simplicity: QEC in bosonic systems is simpler due to a simpler noise structure and fewer errors to correct compared to many physical two-level systems, such as the transmon. For instance, as we will discuss in Sec. 2.1, one typically only needs to correct the dominant photon loss errors in most bosonic systems. The structured noise allows for the design of simple and tailored QEC codes capable of efficiently correcting dominant errors.

3. Autonomous QEC: The local nature of bosonic QEC enables autonomous QEC by engineering a local reservoir. For single-mode bosonic codes, the QEC recovery channels only require local operations on a single oscillator. Such recovery channels can be implemented using active measurements and feedback, or equivalently, by engineering a local reservoir (only coupled to the target oscillator) that passively and continuously extracts the entropy [52]. Such an autonomous scheme can reduce the complexity of QEC and be more robust against operational errors. Note that similar autonomous

schemes for surface codes, or other stabilizer codes, are much more challenging since they would involve reservoirs that couple multiple qubits.

However, challenges arise when utilizing bosonic systems for full-fledged fault-tolerant quantum computing: (1) Achieving fault tolerance against *all* physically relevant errors, i.e. ancilla errors, is challenging. (2) Implementing fault-tolerant logical operations and extending encoded bosonic modes beyond the memory level poses significant challenges.

In Sec. 2.1, we will introduce the basics of bosonic QEC, covering typical noise channels and representative bosonic codes. We will discuss in detail the challenges faced by existing fault-tolerant schemes with bosonic systems.

In Sec. 2.2, we will present a new scheme for autonomous QEC and fault-tolerant quantum computing using squeezed cat qubits. Our scheme autonomously suppresses common types of bosonic errors using low-order non-linear interactions between bosonic modes, enabling low-overhead fault-tolerant quantum computing.

In Sec. 2.3, we will introduce a new set of universal and fault-tolerant logical operations for the four-leg cat code. We overcome the challenge of implementing fault-tolerant logical gates for this code by introducing discrete-variable ancillae and achieving fault tolerance for the hybrid system.

## 2.1 Basics of bosonic quantum error correction

In this section, we review some basic aspects of bosonic error correction, covering common noise channels of bosonic systems (Sec. 2.1.1) and representative bosonic codes (Sec. 2.1.2). Finally, we discuss the challenges of fault-tolerant quantum computing using existing bosonic codes.

### *2.1.1 Noise channels*

Typical bosonic noise channels include excitation loss, heating, dephasing, and Gaussian random shifts. For most bosonic systems such as the superconducting cavities [53] and optics [54], the excitation loss is the dominant error source. In the following, we provide a brief introduction to the loss channel. We refer the readers to Refs. [38, 55] for the rest of the noise channels.

#### Bosonic loss channel

Before we introduce the loss channel, let us briefly first introduce some notations and terminologies for generic quantum channels. The notations closely follow Ref. [56].

Let $\mathcal{H}_X$ be the Hilbert space of a system $X$, $\mathcal{L}(\mathcal{H}_X)$ the space of linear operators over $\mathcal{H}_X$ [57], and $\mathcal{D}(\mathcal{H}_X)$ the space of density matrices. A linear map $\mathcal{N}_{A\to B}\colon \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_B)$ is called a *quantum channel* from $A$ to $B$ if it is completely positive and trace-preserving (CPTP), i.e. if for any ancillary system $E$, the composite map

$$(\mathcal{N}\otimes\mathcal{I})_{AE\to BE} : \mathcal{L}(\mathcal{H}_A)\otimes\mathcal{L}(\mathcal{H}_E) \to \mathcal{L}(\mathcal{H}_B)\otimes\mathcal{L}(\mathcal{H}_E), \tag{2.1}$$

where $\mathcal{I}$ denotes an identity channel (on $E$), maps density matrices to density matrices.

Any quantum channel $\mathcal{N}_{A\to B}$ admits a *Kraus representation*:

$$\mathcal{N}_{A\to B} = \sum_i K_i \bullet K_i^\dagger, \tag{2.2}$$

where $K_i\colon \mathcal{H}_A \to \mathcal{H}_B$ are a set of Kraus operators representing $\mathcal{N}_{A\to B}$, and $\bullet$ denotes any operator in $\mathcal{L}(\mathcal{H}_A)$.

Let $H_{\mathrm{CV}} = \mathrm{span}\{|n\rangle\}_{n=0}^{\infty}$ denotes the infinite-dimensional Hilbert space of an oscillator, where $|n\rangle$ denotes a Fock state with $n$ excitations, and $\hat{a} := \sum_{n=1}^{\infty}\sqrt{n}\,|n-1\rangle\langle n|$ the an-

nihilation operator acting on $\mathcal{H}_{CV}$. The bosonic loss channel that maps $\mathcal{L}(\mathcal{H}_{\mathrm{CV}})$ back to $\mathcal{L}(\mathcal{H}_{\mathrm{CV}})$ can be defined using the Lindblad master equation [58]

$$\frac{d}{dt}\hat{\rho}(t) = \kappa\mathcal{D}\left[\hat{a}\right]\left(\hat{\rho}(t)\right), \tag{2.3}$$

where $\mathcal{D}[\hat{O}](\bullet) := \hat{O} \bullet \hat{O}^{\dagger} - \frac{1}{2}\{\hat{O}, \bullet\}$. By integrating Eq. (2.3), we obtain

$$\hat{\rho}(t) = \mathcal{N}_L\left[\gamma\right]\left(\hat{\rho}(0)\right), \tag{2.4}$$

where $\gamma = 1 - e^{-\kappa t}$ is the transmitivity, and $\mathcal{N}_L\left[\gamma\right]$ is the CPTP map representing the loss channel with a transmitivity $\gamma$.

The Kraus representation for $\mathcal{N}_L\left[\gamma\right]$ is given by [38]:

$$\mathcal{N}_L\left[\gamma\right] = \sum_{n=0}^{\infty} \hat{L}_k \bullet \hat{L}_k^{\dagger}, \quad \hat{L}_k = \sqrt{\frac{\gamma^k}{k!}}(1-\gamma)^{\frac{\hat{n}}{2}}\hat{a}^k. \tag{2.5}$$

### *2.1.2 Bosonic cat codes*

In this section, we briefly review a representative family of single-mode bosonic codes - the rotation-symmetric codes [59], and a sub-family of codes - the cat codes [60]. We refer the readers to Refs. [61] for another family of common single-mode bosonic codes — the GKP code [62].

#### Rotation-symmetrical codes

A $N$-fold rotation-symmetrical (RS) code is stabilized by a discrete bosonic rotation operator:

$$\hat{R}_N := e^{i(2\pi/N)\hat{n}}, \tag{2.6}$$

where $\hat{n}$ is the bosonic photon-number operator and $N$ is some positive integer. This indicates that the code possesses a $N$-fold rotation symmetry since all the codewords are invariant under a $2\pi/N$ phase rotation.

However, a single stabilizer in Eq. (2.6) does not uniquely specify a two-dimensional subspace of an oscillator. Thus, one needs to specify another degree of freedom to define a logical qubit. More specifically, we can uniquely define a $(N, |\Theta\rangle)$-RS code with codewords [59]

$$\begin{aligned} \left|+_{N,\Theta}\right\rangle &:= \frac{1}{\sqrt{\mathcal{N}_+}} \sum_{m=0}^{2N-1} e^{i(m\pi/N)\hat{n}} \left|\Theta\right\rangle, \\ \left|-_{N,\Theta}\right\rangle &:= \frac{1}{\sqrt{\mathcal{N}_-}} \sum_{m=0}^{2N-1} (-1)^m e^{i(m\pi/N)\hat{n}} \left|\Theta\right\rangle, \end{aligned} \tag{2.7}$$

where $\mathcal{N}_+, \mathcal{N}_-$ are normalization constants. Note that each of the codeword is simply a superposition of rotated copies of $|\Theta\rangle$. As such, $|\Theta\rangle$ can be viewed as a base state, which, together with the rotation symmetry $N$, uniquely define a rotion-symmetrical code. As an example, taking $|\Theta\rangle = |\alpha\rangle$, where $|\alpha\rangle$ denotes a coherent state with amplitude $\alpha$, we obtain the well-known cat codes with $2N$ legs [38, 59].

A $(N, |\Theta\rangle)$-RS code is always capable of detecting $N-1$ loss errors, since the codewords $\left|0_{N,\Theta}\right\rangle$ and $\left|1_{N,\Theta}\right\rangle$ in Eq. (2.7) only support on every $2kN$ and $(2k+1)N$ fock state (for $k = 0, 1, \cdots$), respectively. Such a photon number separation by $N$ between codewords is intrinsically due to the $N$-fold rotation symmetry and irrespective of the choice the base state $|\Theta\rangle$.

However, the capability of correcting loss errors will depend on the base state $|\Theta\rangle$ since the distortion of the codespace after $p$ ($p \leq 1$) losses will depend on $\Theta$:

$$\hat{P}_{N,\Theta}(\hat{a}^{\dagger p}\hat{a}^p)\hat{P}_{N,\Theta} = \epsilon_0 \hat{\bar{I}} + \sum_{i\in\{X,Y,Z\}} \epsilon_i(p, N, \Theta)\hat{\bar{P}}_i, \tag{2.8}$$

where $\hat{P}_{N,\Theta}$ denotes the projection operator onto the codespace, $\hat{\bar{I}}$ the logical indentity, and

$\hat{\bar{P}}_X, \hat{\bar{P}}_Y, \hat{\bar{P}}_Z$ the three types of logical Pauli operator of the code. $\epsilon_i(p, N, \Theta)$ will be some complex functions that depends on not only $N$, but also $|\Theta\rangle$. If $\epsilon_i(p, N, \Theta) \neq 0$ fir some $i$, the $(N, |\Theta\rangle)$ code can only approximately correct $p$ photon losses and the amplitude of $\epsilon_i(p, N, \Theta)$ characterizes the code's QEC capability as an approximate error-correcting code [63].

As an example, a carefully-designed binomial code [64], whose base state $|\Theta\rangle$ has binomial coefficients in the Fock basis, can perfectly correct photon losses up to a certain order. In contrast, a cat code, with a coherent state as the base state, can only approximately correct loss errors. Nevertheless, the error coefficients $\{\epsilon_i(p, N, |\alpha\rangle)\}$ of a cat code get exponentially suppressed by the mean photon number of the code $\bar{n} := \mathrm{tr}(\hat{P}_{N,\alpha}, \hat{n})$ , which is approximately $|\alpha|^2$ (see Sec. 2.3.1 for more details). As such, a cat code can also perfectly detect and correct $N-1$ photon losses in the large-cat limit ($\gg 1$).

### Cat codes

As shown in the previous section, a $2N$-leg cat can be defined as a $N$-fold RS code with a base state $|\alpha\rangle$. Such a cat can detect $N-1$ photon loss errors, and can also approximately correct them when $\alpha \gg 1$.

An advantage of the cat codes, compared to other RS codes, is that it is relatively simple to prepare its logical states, which are simply superposition of coherent states. In fact, the four-leg cat, capable of correcting a single-photon loss, was the first QEC code demonstrating break-even error suppression, i.e. the lifetime of a logical qubit is longer than its physical counterpart, experimentally [65].

Another advantage of the cat code is that it is relatively easy to engineer a dissipator that stabilizes the codespace. As shown in Ref. [45], a $2N$-leg cat code can be stabilized by the following Markovian dissipation:

$$\frac{d}{dt}\hat{\rho} = \mathcal{D}[a^{2N} - \alpha^{2N}]\hat{\rho}. \tag{2.9}$$

More concretely, Eq. (2.9) has a $2N$-dimensionsal coherent steady-state subspace, which is spanned by $\{e^{i(2\pi/Nm)\hat{n}}\,|\alpha\rangle\}_{m=0}^{2N-1}$ [45]. Obviously, the codespace of the cat is a subspace of this steady-state subspace, and thus being protected by the engineered dissipator $\mathcal{D}[a^{2N} - \alpha^{2N}]$.

As an example, a two-leg cat (with $N = 1$) can be stabilized by a dissipator $\mathcal{D}[a^2 - \alpha^2]$, which can be engineered by a two photon loss $\mathcal{D}[a^2]$, together with a coherent two photon drive [4]. It can be realized in, e.g. superconducting circuits [66], using only third-order nonliearities. Such dissipatively protected two-leg cats are referred to as dissipative cats, and are considered as a promising candidates for realizing low-overhead fault-tolerant quantum computing [4, 5]. Although this dissipative cat cannot correct any photon loss errors and end up with uncorrected phase-flip errors, the bit-flip errors, which are associated with flipping the coherent states from $|\alpha\rangle$ to $|-\alpha\rangle$, are exponentially suppressed by the dissipation. Then, one can obtain a logical qubit with [4]

$$\gamma_{\text{phase}} \propto \kappa_1 \bar{n}, \quad \gamma_{\text{bit}} \propto \bar{n} e^{-2\bar{n}}, \tag{2.10}$$

where $\gamma_{\text{phase}}$ and $\gamma_{\text{bit}}$ denotes the logical phase- and bit-flip error rate, respectively, and $\bar{n} = |\alpha^2|$ denotes, again, the mean photon number. Since the bit-flip rate is exponentially suppressed while the phase-flip rate is only linearly amplified, this code can be viewed as a bosonic analog of the repetition code. More importantly, such exponential suppression against one type of error is achieved without active QEC, but instead using engineered dissipation. As such, these cat qubits can be viewed as a new type of physical qubits with certain intrinsic error protection and a highly biased noise channel. The noise bias can be viewed as a resource for reducing the resource overhead [67, 68], when concatenating these cat qubits with an outer qubit code [69, 5].

The two-leg cat can be alternatively stabilized by a Hamiltonian with Kerr nonlineari-

ties [70, 46]:

$$H_{\text{Kerr}-\text{cat}} = -K(a^{\dagger 2} - \alpha^2)(a^2 - \alpha^2). \tag{2.11}$$

The codespace of the cat is now the degenerate ground-state manifold of the engineered Hamiltonian, which is gapped from all other excited states by a gap $\sim 4K|\alpha|^2$. Such a cat protected by a gapped Hamiltonian is referred to as the Kerr-cat qubit. The same codespace can be also protected by a combination of gapped Hamiltonian and engineered dissipation [71, 72]. We refer to all these cat qubits with passive stabilization as stabilized cat qubits.

These stabilized cat qubits all feature a highly suppressed bit-flip rate and, consequently, a high noise. Moreover, the noise bias of these stabilized cats can be preserved in the circuit level by using carefully-designed bias-preserving operations [46]. Therefore, it is feasible to design fault-tolerant quantum computing schemes with lower overhead using biased-noise cat qubits [5, 69, 73, 74].

### *2.1.3 Challenges of fault-tolerant quantum computing using bosonic cat codes*

Although the family of bosonic cat codes are promising candidates for low-overhead fault-tolerant quantum computing, significant challenges are present:

1. All though the cat code can, in principle, exponentially suppress the bit-flip errors while also correcting finite-order photon losses (with $N \geq 2$), simultaneous suppression of both errors is challenging to achieve in practice. For the stabilized two-leg cats (with $N = 1$), although exponential bit-flip error suppression has been demonstrated with the passive stabilization [66], loss-induced phase-flip errors are uncorrectable using this code. Consequently, very low photon-loss rates (or alternatively, very high stabilization strength) are required for achieving fault tolerance. For instance, it is estimated in Ref. [5] that the fault tolerance threshold for $\kappa_1/\kappa_2$, where $\kappa_1$ denotes the single-photon loss and $\kappa_2$ denotes the engineered two-photon dissipation rate, is in the

level of $10^{-4}$, which is far below what has been achieved ($\sim 10^{-1}$ [66]) so far. For cats with more legs (e.g. $N = 2$), although loss-induced phase-flip errors can be suppressed to a higher order by measuring the modular photon number [65], bit-flip errors become harder to correct. On the one hand, passive QEC involves high-order dissipator (see Eq. (2.9)) that is challenging to realize; On the other hand, existing active QEC schemes, such as the teleportation-based QEC scheme in Ref. [59], involve high-order nonlinear interactions between bosonic modes, as well as phase measurements that are challenging to engineer.

2. Due to the highly nonlinear nature of the cat codes, implementing fault-tolerant logical operations is challenging. Existing schemes either require strong nonlinear interactions between bosonic modes (see e.g. [4, 59]), or use discrete-variable ancillae that could inject errors to the bosonic modes [75].

## 2.2 Fault-tolerant quantum computation with autonomously-protected squeezed cat qubits

The standard QEC procedure relies on actively measuring the error syndromes and performing feedback controls [15]. However, such adaptive protocols demand fast, high-fidelity coherent operations and measurements, which poses significant experimental challenges. At this stage, the error rates in the encoded level are still higher than the physical error rates in current devices due to the errors during the QEC operations [24, 37, 76, 77]. To address these challenges, we may implement QEC non-adaptively via engineered dissipation – an approach called autonomous QEC (AutoQEC) [52, 78]. Such an approach avoids the measurement imperfection and overhead associated with the classical feedback loops. AutoQEC schemes that can greatly suppress dephasing noise in bosonic systems have been both theoretically investigated and experimentally demonstrated using the two-component

cat code [45, 4, 70, 71, 66, 79, 80]. However, AutoQEC against excitation loss, which is usually the dominant error source in a bosonic mode, remains challenging. It requires either large nonlinearities that are challenging to engineer (e.g., the multiphoton processes needed for $n$-fold rotation-symmetrical codes with $n \geq 4$ [45, 81, 59]) or couplings to an intrinsically nonlinear DV system [82, 83, 84] that is much noisier than the bosonic mode.

In this work, we propose an AutoQEC scheme against excitation loss with low-order nonlinearities and accessible experimental resources. Our scheme is, in principle, device-independent and readily implementable in superconducting circuits and trapped-ion systems. The scheme is based on the squeezed cat (SC) encoding [85], which involves the superposition of squeezed coherent state. We introduce an explicit AutoQEC scheme for the SC against loss errors by engineering a nontrivial dissipation, which simultaneously stabilizes the SC states and corrects the loss errors. We show that the engineered dissipation is close to the optimal recovery obtained using a semidefinite programming [1, 2, 3]. Notably, our proposed dissipation can be implemented with the same order of nonlinearity as that required by the two-component cat, which has been experimentally demonstrated in superconducting circuits [86, 66, 87] and shown to be feasible in trapped-ion systems [88].

Furthermore, we show that similar to the stabilized cat qubits, the stabilized SC qubits also possess a biased noise channel (with one type of error dominant over others), with an even larger bias (defined to be the ratio between the dominant error rate and the others) $\sim e^{\bar{n}^2}$ (compared to $\sim e^{\bar{n}}$ for the cat), where $\bar{n}$ denotes the mean excitation number of the codewords. Consequently, we can concatenate the stabilized SC qubits with a DV code tailored towards the biased noise to realize low-overhead fault tolerant QEC and quantum computation [89, 90, 91, 92, 93, 68]. We develop a set of operations for the SC that are compatible with the engineered dissipation and can preserve the noise bias needed for the concatenation. Compared to those for the cat [4], these operations suffer less from the loss errors because of the AutoQEC. Moreover, they can be implemented faster due to a larger

effective dissipation gap and a cancellation of the leading-order non-adiabatic errors. In combination, the access to higher-quality operations leads to much better logical performance in the concatenated level using the SC qubits. For instance, we can achieve one-to-two orders of magnitude improvement in the $\kappa_1/\kappa_2$ threshold, where $\kappa_1$ is the excitation loss rate and $\kappa_2$ is the engineered dissipation rate, for the surface-SC and repetition-SC scheme (compared to surface-cat and repetition-cat, respectively). Furthermore, the repetition-SC can achieve a logical error rate as low as $10^{-15}$, which already suffices for many useful quantum algorithms [17, 94], even using a small SC with $\bar{n} = 4$ under a practical noise ratio $\kappa_1/\kappa_2 = 10^{-3}$.

We note that aspects of the SC encoding were also recently studied in Ref. [85], with an emphasis on the enhanced protection against dephasing provided by squeezing (a point already noted in Refs. [95, 96, 97]). Unlike our work, Ref. [85] neither explored the enhanced noise bias provided by squeezing, nor exploited the ability to concatenate the SC code with outer DV codes using bias-preserving operations; as we have discussed, these are key advantages of the SC approach. Our work also goes beyond Ref. [85] in providing an explicit, fully autonomous approach to SC QEC that exploits low-order nonlinearities, and it is compatible with several experimental platforms. In contrast, Ref. [85] studied an approach requiring explicit syndrome measurements and a formal, numerically-optimized recovery operation. It was unclear how such an operation could be feasibly implemented in experiment. We also note that the SC has also been studied in the context of quantum transduction [98] (a very different setting than that considered here) and preparation of SC states has been recently demonstrated experimentally [99].

### *2.2.1 Squeezed cat encoding and subsystem decomposition*

**Squeezed cat encoding**

The codewords of the SC are defined by applying a squeezing along the displacement axis

(which is taken to be real) to the cat codewords:

$$|\mathrm{SC}^{\pm}_{r,\alpha'}\rangle := \hat{S}(r)|C^{\pm}_{\alpha'}\rangle \tag{2.12}$$

where $|C^{\pm}_{\alpha'}\rangle := \mathcal{N}_{\pm}(|\alpha'\rangle + |-\alpha'\rangle)$ with $\mathcal{N}_{\pm} = \frac{1}{\sqrt{2(1\pm e^{-2\alpha'^2})}}$ being normalization factors, and $\hat{S}(r) := \exp\left[\frac{1}{2}r(\hat{a}^2 - \hat{a}^{\dagger 2})\right]$ is the squeezing operator. The above codewords with even ($|\mathrm{SC}^{+}_{r,\alpha'}\rangle$) and odd ($|\mathrm{SC}^{-}_{r,\alpha'}\rangle$) excitation number parity are defined to be the X-basis eigenstates. Similar to other bosonic codes [100], the performance of the SC code is related to the mean excitation number $\bar{n}$ of the codewords:

$$\begin{aligned}\bar{n} &:= \frac{1}{2}(\langle \mathrm{SC}^{+}_{r,\alpha'}|\hat{a}^{\dagger}\hat{a}|\mathrm{SC}^{+}_{r,\alpha'}\rangle + \langle \mathrm{SC}^{-}_{r,\alpha'}|\hat{a}^{\dagger}\hat{a}|\mathrm{SC}^{-}_{r,\alpha'}\rangle) \\ &= \alpha'^2\left(\coth 2\alpha'^2 \cosh 2r - \sinh 2r\right) + \sinh^2 r\end{aligned} \tag{2.13}$$

For a SC code with fixed $\bar{n}$, according to Eq. (2.13), the amplitude $\alpha'$ of the underlying coherent states varies with the squeezing parameter $r$ as

$$\alpha' \approx \sqrt{\bar{n} - \sinh^2 r}e^{r}, \tag{2.14}$$

which holds for the regime of interest where $\alpha' > 1$. See a graphic illustration of the interdependency between $\bar{n}, \alpha'$ and $r$ in Fig. 2.1. Note that $\alpha'$ is closely related to how separated in phase space the two $Z$-basis SC states are, which determines their resilience against local error processes [100, 101]. At fixed $\bar{n}$, $\alpha'^2$ can be written as a concave quadratic function of $e^{2r}$, $\alpha'^2 \approx -\frac{1}{4}e^{4r} + (\bar{n} + \frac{1}{2})e^{2r} - \frac{1}{4}$, which has a maximum $\alpha'^2_{\max} = \bar{n}^2 + \bar{n}$ (see Fig. 2.1(a)). In Fig. 2.1(b), we show how the average photon number is varied as a function of squeezing while maintaining a constant dissipation gap, which is proportional to $\alpha'^2$, fixed. It is then apparent that we can achieve a maximum suppression of average photon number at an intermediate squeezing while enjoying the same protection from dissipation gap as the

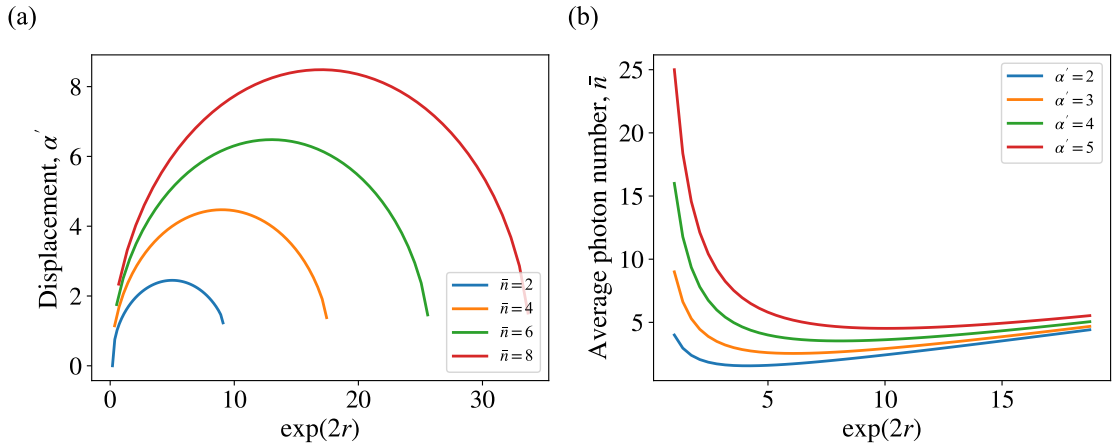


Figure 2.1: Relations between the displacement, $\alpha'$, and the average photon number, $\bar{n}$, as a function of the exponential of squeezing, $e^{2r}$, of a SC code, $|\mathrm{SC}_{\alpha',r}\rangle$. (a) The displacement varies with the squeezing as $\alpha' = \sqrt{\bar{n} - \sinh^2 r}e^r$ while fixing the average photon numbers. The maximum achievable displacements are $\sqrt{\bar{n}^2 + \bar{n}}$. (b) The average photon number varies with the squeezing as $\bar{n} = \alpha'^2 e^{-2r} + \sinh^2 r$ while fixing the displacements and, therefore, the dissipation gap.

cat code.

**Subsystem decomposition**

We can divide the Hilbert space of a bosonic mode $\mathcal{H}_{\mathrm{CV}} = \mathrm{span}\{|n\rangle, n = 0, 1, 2..\}$ into two orthogonal subspaces:

$$\mathcal{H}_{\mathrm{CV}} = \mathcal{H}_{\mathrm{even}} \oplus \mathcal{H}_{\mathrm{odd}},$$

where $\mathcal{H}_{\mathrm{even}}$ ($\mathcal{H}_{\mathrm{odd}}$) is the $+1$ ($-1$) eigenspace of the parity operator $\hat{\Pi} := e^{-i\hat{n}\pi}$. Since both $\mathcal{H}_{\mathrm{even}}$ and $\mathcal{H}_{\mathrm{odd}}$ are isomorphic to $\mathcal{H}_{\mathrm{CV}}$, we can decompose $\mathcal{H}_{\mathrm{CV}}$ into two subsystems:

$$\mathcal{H}_{\mathrm{CV}} = \mathbb{C}_L^2 \otimes \mathcal{H}_g,$$

where $\mathbb{C}_L^2$ is a Hilbert space of dimension 2 (which we refer to as a qubit), and $\mathcal{H}_g \simeq \mathcal{H}_{\mathrm{CV}}$ is a Hilbert space of infinite dimension (which we refer to as a gauge mode). Under this decomposition, $|+\rangle_L \otimes |\psi\rangle_g$ ($|-\rangle_L \otimes |\psi\rangle_g$) is an even-(odd-) parity state for any $|\psi\rangle_g \in \mathcal{H}_g$. The choice of the gauge mode basis is not unique. For instance, we can choose a basis for

$\mathcal{H}_g$ based on the modular decomposition [102] of the number operator: $|+\rangle_L \otimes |m\rangle_g := |0+2m\rangle$ ($|-\rangle_L \otimes |m\rangle_g := |1+2m\rangle$). For the squeezed cat (SC), it is convenient to work with another basis related to the squeezed coherent states. We first define a set of non-orthonormalized states $|\psi_{n,\pm}\rangle := \mathcal{N}_{n,\pm}\hat{S}(r)\left[\hat{D}\left(\alpha'\right) \pm (-1)^n \hat{D}\left(-\alpha'\right)\right]|n\rangle$, where $\mathcal{N}_{n,\pm}$ is the normalization factor. Note that states with different parity $\pm$ are orthogonal with each other. We then apply the Gram-Schmidt orthonormalization procedure to the states within each parity branch (starting from $|\phi_{0,\pm}\rangle$ and then increasing $n$) to obtain a set of orthonomalized states $\{|\Phi_{n,\pm}\rangle\}$. Finally, we define a subsystem basis as

$$|\pm\rangle_L \otimes |\tilde{n} = n\rangle_g := |\Phi_{n,\pm}\rangle. \tag{2.15}$$

The choice of this subsystem basis can describe the SC more efficiently than the Fock basis since $|\pm\rangle_L \otimes |\tilde{n} = 0\rangle_g$ coincides with the SC codewords. Furthermore, the physical states of a stabilized SC usually evolve within a subspace with low excitation in the gauge mode. As such, we can apply a truncation to the gauge mode and perform the analysis within a truncated $2d$-dimensional (with $d$ being a small integer) subspace of $\mathcal{H}_{\text{CV}}$: $\text{span}\{|\pm\rangle_L \otimes |\tilde{n} = n\rangle_g, n = 0, 1, .., d-1\}$.

Note that the states $|\Phi_{n,\pm}\rangle$ are equivalent to the shifted Fock basis introduced in Ref. [5] by a global squeezing transformation. For example, we can obtain the expression of the bosonic annihilation operater by applying a squeezing transformation to Eq. (C15) in Ref. [5], where it is decomposed as $\hat{a}_{\text{sf}} \approx \hat{Z}_L \otimes (\alpha + \hat{\tilde{a}})$, and obtain

$$\hat{a} \xrightarrow{\hat{S}(r)} \hat{S}^\dagger(r)\hat{a}_{\text{sf}}\hat{S}(r) \tag{2.16}$$

$$= \cosh r\hat{a}_{\text{sf}} - \sinh r\hat{a}_{\text{sf}} \tag{2.17}$$

$$= \hat{Z}_L \otimes (e^{-r}\alpha' + \cosh r\hat{\tilde{a}} - \sinh r\hat{\tilde{a}}^\dagger) + \mathcal{O}(e^{-2\alpha'^2}), \tag{2.18}$$

where $\hat{Z}_L$ is the Pauli Z operator acting on the logical qubit, and $\hat{\tilde{a}} = \sum_{n=0}^{\infty}\sqrt{n+1}|\hat{\tilde{n}} =$

$n\rangle_g\langle\hat{\tilde{n}} = n+1|$ is the annihilation operator acting on the gauge mode.

**Error-correction capability**

Here, we analyze the capability of the squeezed cat encoding for correcting excitation loss ($\hat{a}$), heating ($\hat{a}^\dagger$), and dephasing ($\hat{a}^\dagger\hat{a}$) errors. We first explain why the SC code can correct the loss errors by analyzing the Knill–Laflamme error correction conditions [103, 104] and evaluating the QEC matrices [105]. Consider a pure loss channel with a loss probability $\gamma$, the leading-order Kraus operators are $\{\hat{I}, \sqrt{\gamma}\hat{a}\}$. The detectability of a single excitation loss is quantified by the matrix:

$$\begin{aligned}\hat{P}_{\text{code}}\hat{a}\hat{P}_{\text{code}} &= e^{-r}\alpha'\frac{q+q^{-1}}{2}\hat{Z}_c + ie^{r}\alpha'\frac{q-q^{-1}}{2}\hat{Y}_c \\ &\approx \sqrt{\bar{n}-\sinh^2 r}\hat{Z}_c - ie^{r}\alpha' e^{-2\alpha'^2}\hat{Y}_c,\end{aligned} \tag{2.19}$$

where $\hat{P}_{\text{code}}$ is the projection onto the code space, $\hat{Z}_c := \hat{Z}_L \otimes |0\rangle_g\langle 0|$ ($\hat{Y}_c := \hat{Y}_L \otimes |0\rangle_g\langle 0|$) is the Pauli $Z$ ($Y$) operator in the code space, and $q := \sqrt{\frac{1-e^{-2\alpha'^2}}{1+e^{-2\alpha'^2}}}$. See Supplementary Note 2 of Ref. [106] for a detailed derivation. The approximation in the second line is made in the regime of interest where $e^{-2\alpha'^2} \ll 1$.

Eq. (2.19) indicates that a single excitation loss mostly leads to an undetectable logical phase-flip error with a probability that decrease with the squeezing parameter $r$, which can be better understood by considering the action of the decomposed $\hat{a}$ operator (Eq. (2.18)) on the codeword $\hat{a}(|+\rangle_L \otimes |0\rangle_g) = |-\rangle_L \otimes \sqrt{\bar{n}}(\sqrt{\eta}|0\rangle_g - \sqrt{1-\eta}|1\rangle_g)$, where

$$\eta := (\bar{n} - \sinh^2 r)/\bar{n}. \tag{2.20}$$

As shown in Fig. 1, after a single excitation loss, the branch of the population (with ratio $\eta$) that stays in the ground state of the gauge mode leads to undetectable logical phase-flip errors. In contrast, the other branch (with ratio $1-\eta$) that goes to the first excited gauge state is in principle detectable. The detectable branch is also approximately correctable since

$\hat{P}_{\text{code}}\hat{a}^\dagger\hat{a}\hat{P}_{\text{code}} \approx \bar{n}\hat{I}_c + \mathcal{O}(e^{-2\alpha'^2})\hat{X}_c$. Therefore, we expect that we can suppress the loss-induced phase flip errors by a factor $\eta$ that decreases with the squeezing $r$. Moreover, the $\hat{X}_c$ and $\hat{Y}_c$ terms in the QEC matrices for both loss, heating, and dephasing are exponentially suppressed by $\alpha'^2$. As shown in Eq. (2.14), $\alpha'^2$ can be greatly increased by adding squeezing (with $\alpha'^2_{\text{max}} = \bar{n}^2 + \bar{n}$). Consequently, we expect that the SC can also have significantly enhanced noise bias compared to the cat.

### *2.2.2 Autonomous quantum error correction*

While we have shown that the SC encoding can, in principle, detect and correct the loss errors, it remains a non-trivial task to find an explicit and practical recovery channel. In this section, we provide such a recovery channel, showing surprisingly that it requires only experimental resources that have been previously demonstrated.

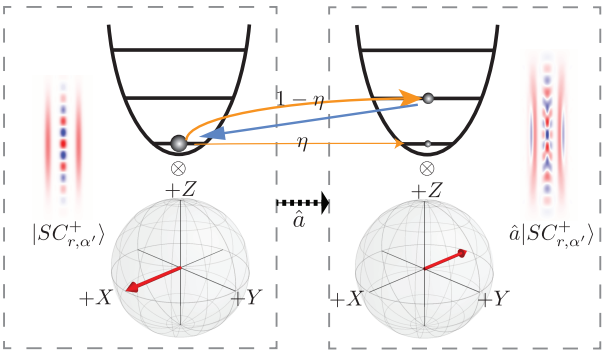


Figure 2.2: The illustration of a SC that suffers from a single excitation loss and then approximately corrects it. Each dashed box represents a state (visualized by the Wigner function) of the SC, which is decomposed as a product of a logical qubit and a gauge mode. A single excitation loss corrupts the codeword $|+\rangle_c$ (left) into the state $\hat{a}|+\rangle_c/\sqrt{\langle+|_c\hat{a}^\dagger\hat{a}|+\rangle_c}$ (right). During such a process, a phase flip happens on the logical qubit, and a fraction $1-\eta$ of the gauge mode population gets excited (indicated by the thick orange arrow). The excited population can be detected and then corrected, as indicated by the blue arrow.

As shown by the blue arrow in Fig. 2.2, we can, in principle, perform photon counting measurement on a probe field that is weakly coupled to the gauge mode, and apply a feedback parity flip $\hat{Z}_L$ on the logical qubit upon detecting an excitation in the probe field [107].

Such measurement and feedback process can be equivalently implemented by applying the dissipative dynamics as described by Lindblad master equation $\frac{d\hat{\rho}}{dt} = \kappa_2 \mathcal{D}[\hat{F}]$, with the jump operator $\hat{F}$ given by

$$\hat{F} = (\hat{Z}_L \otimes \hat{\tilde{I}})\hat{S}(r)(\hat{a}^2 - \alpha'^2)\hat{S}^\dagger(r), \tag{2.21}$$

and $\mathcal{D}[\hat{A}]\hat{\rho} := \hat{A}\hat{\rho}\hat{A}^\dagger - \frac{1}{2}\{\hat{A}^\dagger\hat{A}, \hat{\rho}\}$. When $\alpha' \gg 1$, $\hat{F} \propto \hat{Z}_L \otimes \hat{\tilde{a}}$ represents a logical phase flip conditioned on the gauge mode losing an excitation. In the Fock basis, such an operator can be approximately given by

$$\hat{F} \approx \frac{1}{\alpha'}\hat{S}(r)(c_1\hat{a} + c_2\hat{a}^\dagger)(\hat{a}^2 - \alpha'^2)\hat{S}^\dagger(r), \tag{2.22}$$

with $c_1 + c_2 = 1$.

In Sec. 2.2.5, we propose two reservoir-engineering approaches to implement such a non-trivial dissipator using currently accessible experimental resources. We sketch the main ideas here. The first approach utilizes three bosonic modes that are nonlinearly coupled. As shown in Fig. 2.6(a), a high-quality mode $b$ and a lossy mode $c$, together, serve as a nonreciprocal bath [108] that provides a directional interaction $e^{-i\theta\hat{Z}_L} \otimes \hat{\tilde{a}}$ from the gauge mode to the logical qubit in the storage mode $a$. Such a coupled system can be physically realized in, e.g., superconducting circuits [86, 87]. The second approach couples a bosonic mode nonlinearly to a qutrit $\{|g\rangle, |e\rangle, |f\rangle\}$. As shown in Fig. 2.7, the bosonic mode is coupled to the $gf$ transition via $\hat{S}(r)(\hat{a}^2 - \alpha'^2)\hat{S}^\dagger(r)|f\rangle\langle g| + h.c.$ and to the $ef$ transition via $\hat{Z}_L|e\rangle\langle f| + h.c.$. By enhacing the decay from $|e\rangle$ to $|g\rangle$, we can obtain the effective dissipator $\hat{F}$ by adiabatically eliminating both $|e\rangle$ and $|f\rangle$. Such a system can be physically realized in, e.g., trapped-ion system [88].

With the engineered dissipator in Eq. (2.21), the SC can be autonomously protected from excitation loss, heating and dephasing. We now derive the error channel of the dissipatively stabilized SC qubit in the memory level. The dynamics of the system are described by the

Lindblad master equation:

$$\frac{d\hat{\rho}}{dt} = \kappa_2 \mathcal{D}[\hat{F}]\hat{\rho} + \kappa_1(1+n_{\text{th}})\mathcal{D}[\hat{a}]\hat{\rho} + \kappa_1 n_{\text{th}}\mathcal{D}[\hat{a}^\dagger]\hat{\rho} + \kappa_\phi \mathcal{D}[\hat{a}^\dagger\hat{a}]\hat{\rho}. \tag{2.23}$$

The logical phase-flip and bit-flip error rates of the SC under the dynamics described by Eq. (2.23) are given by:

$$\gamma_Z = [\kappa_1(1+2n_{\text{th}}) + \kappa_\phi e^{-2r}](\bar{n} - \sinh^2 r), \tag{2.24}$$

$$\gamma_{X,Y} = \kappa_\phi \frac{(\bar{n}-\sinh^2 r)e^{2r}(\sinh^2 2r/4 + \cosh 4r)}{2\sinh\big(2(\bar{n}-\sinh^2 r)e^{2r}\big)}, \tag{2.25}$$

where $\gamma_{X,Y}$ denotes the sum of the logical $X$ and $Y$ error rates, which we refer to as the bit-flip rate for simplicity. In the following, we provide a detailed derivation of Eqs. (2.24) (2.25).

The phase-flip error rate can be easily derived by analyzing the errors under the subsystem decomposition. The loss and heating errors are in the form $\hat{a} \approx \hat{Z}_L \otimes (e^{-r}\alpha' + \cosh r\hat{\tilde{a}} - \sinh r\hat{\tilde{a}}^\dagger)$, $\hat{a}^\dagger \approx \hat{Z}_L \otimes (e^{-r}\alpha' + \cosh r\hat{\tilde{a}}^\dagger - \sinh r\hat{\tilde{a}})$. They both contribute to the phase-flip rate via the undetectable term $e^{-r}\alpha'\hat{Z}_L = \sqrt{\eta\bar{n}}\hat{Z}_L$ (the detectable part associated with the $\hat{Z}_L \otimes \hat{\tilde{a}}^\dagger$ term is approximately correctable by $\hat{F}$). The dephasing is in the form $\hat{a}^\dagger\hat{a} \approx \hat{I}_L \otimes [e^{-2r}\alpha'^2 + e^{-2r}\alpha'(\hat{\tilde{a}}+\hat{\tilde{a}}^\dagger) + \cosh^2 r\hat{\tilde{a}}^\dagger\hat{\tilde{a}} + \sinh^2 r\hat{\tilde{a}}\hat{\tilde{a}}^\dagger - \cosh r\sinh r(\hat{\tilde{a}}^2 + \hat{\tilde{a}}^{\dagger 2})]$. It contributes to the phase-flip rate dominantly by the $e^{-2r}\alpha'\hat{I}_L \otimes \hat{\tilde{a}}^\dagger$ term, which creates an excitation in the gauge mode that is subsequently destroyed by $\hat{F}$ with a residual phase flip. Therefore, the dephasing contributes to the phase-flip rate by $\kappa_\phi e^{-4r}\alpha'^2 = \kappa_\phi e^{-2r}\eta\bar{n}$.

Since the bit-flip error rate is exponentially small in $\alpha'$, the subsystem decomposition is insufficient to obtain an analytical expression of it. Thus, we derive the bit-flip error rate using the conserved quantities of the system [45, 109]. To facilitate the analysis, we first neglect the $\hat{Z}_L$ term in the dissipator in Eq. (2.21) since it does not contribute to the bit-flip

rate, and then analyze the system dynamics in the squeezed frame:

$$\frac{d\hat{\rho}_s}{dt} = \kappa_2 \mathcal{D}[\hat{a}^2 - \alpha'^2]\hat{\rho}_s + \kappa_\phi \mathcal{D}[\hat{a}_s^\dagger \hat{a}_s]\hat{\rho}_s, \tag{2.26}$$

where $\hat{A}_s := \hat{S}^\dagger(r)\hat{A}\hat{S}(r)$ for any operator $\hat{A}$. Note that we we consider the dephasing here, which is the dominant source for the bit-flip errors. The two conserved quantities associated with the dominant dissipator $\hat{a}^2 - \alpha'^2$ are

$$\begin{aligned} \hat{J}_{++} &= \sum_{n=0}^{\infty} |2n\rangle \langle 2n| , \\ \hat{J}_{+-} &= \sqrt{\frac{2\alpha'^2}{\sinh 2\alpha'^2}} \sum_{q=-\infty}^{\infty} \frac{(-1)^q}{2q+1} I_q\left(\alpha'^2\right) \hat{J}_{+-}^{(q)}, \end{aligned} \tag{2.27}$$

where $I_q(\cdot)$ is the modified Bessel function of the first kind, and $\hat{J}_{+-}^{(q)} = \frac{\left(\hat{a}^\dagger \hat{a} - 1\right)!!}{\left(\hat{a}^\dagger \hat{a} + 2q\right)!!} \hat{J}_{++} \hat{a}^{2q+1}$ for $q \geq 0$ and $\hat{J}_{+-}^{(q)} = \hat{J}_{++} \hat{a}^{\dagger 2q+1} \frac{\left(\hat{a}^\dagger \hat{a}\right)!!}{\left(\hat{a}^\dagger \hat{a} + 2|q| - 1\right)!!}$ for $q < 0$. The steady state coherence of the system initialized in $\hat{\rho}(0)$ can be computed through $c_{++}(\infty) = \mathrm{tr}\left\{J_{++}^\dagger \rho(0)\right\}$ and $c_{+-}(\infty) = \mathrm{tr}\left\{J_{+-}^\dagger \rho(0)\right\}$. Thus, we compute the bir-flip rate perturbatively by considering the dephasing in the squeezed frame,

$$\gamma_{X,Y} = -\kappa_\phi \mathrm{tr}\left\{J_{+-}^\dagger \mathcal{D}\left[\hat{S}^\dagger(r)\, \hat{a}^\dagger \hat{a} \hat{S}(r)\right] \left|C_\alpha^+\right\rangle \left\langle C_\alpha^-\right|\right\}, \tag{2.28}$$

which is then simplified to Eq. (2.25).

In the regime where $e^{-r} \ll 1$ and $\gamma_Z$ is mainly contributed by excitation loss, we can simplify Eqs. (2.24) and (2.25) as

$$\gamma_Z \approx \eta \bar{n} \kappa_1, \gamma_{X,Y} \approx \frac{9}{16} \kappa_\phi \alpha'^2 e^{-2\alpha'^2} e^{4r}, \tag{2.29}$$

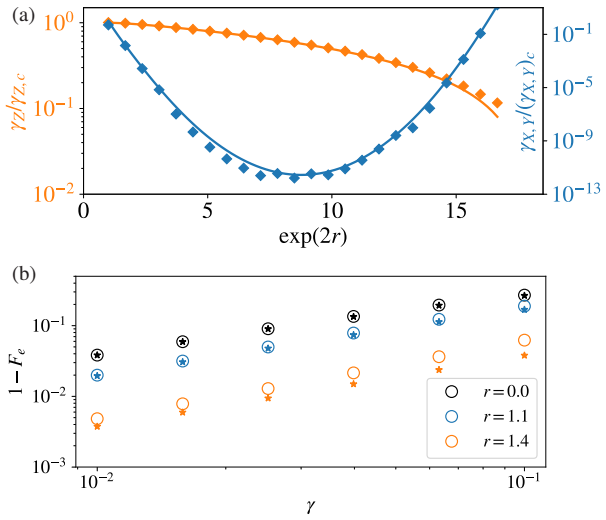


Figure 2.3: Error rates of the stabilized SC. (a) The phase $\gamma_Z$ (orange) and bit $\gamma_{X,Y}$ (cyan) error rate of the dissipatively stabilized SC as a function of squeezing $r$ under the parameters $\bar{n} = 4, \kappa_1 = 100\kappa_\phi = \kappa_2/100, n_{\text{th}} = 0.01$. The solid lines represent the analytical expressions Eqs. (2.24) and (2.25) while the diamonds represent the numerically extracted values. All the error rates are normalized by those of the dissipative cat $\gamma_{Z,c}, (\gamma_{X,Y})_c$, which are given by Eqs. (2.24) and (2.25) with $r = 0$. (b) The entanglement infidelity of a joint loss and recovery channel varying with the loss probability $\gamma$ for the SC encoding with $\bar{n} = 4$. The recovery channel is either the engineered dissipation (the circles) or the optimal recovery channel determined by an SDP program [1, 2, 3] (the stars).

where

$$\alpha' \approx \sqrt{4\eta(1-\eta)}\bar{n}. \tag{2.30}$$

As plotted in Fig. 2.3(a), fixing $\bar{n}$, $\gamma_Z$ decreases monotonically with the squeezing $r$ (unless $r$ approaches the maximum squeezing $r_{\max} \approx \sinh^{-1}(\sqrt{\bar{n}})$. See Methods of Ref. [106] for details.) as the undetectable portion $\eta$ of the loss-induced errors decreases (see Eq. (2.20)). The change of $\gamma_{X,Y}$ with $r$ (or equivalently, $\eta$) is roughly captured by the change in the displacement amplitude $\alpha'$ (see Eq. (2.30)), and $\gamma_{X,Y}$ takes the minima roughly when $\alpha'$ reaches the maxima $\alpha'_{\max} = \sqrt{\bar{n}^2 + \bar{n}}$. Note that the minimal bit-flip rate of the SC enjoys a more favorable scaling $\gamma_{X,Y} \propto e^{-2\bar{n}^2}$ with $\bar{n}$, compared to $\gamma_{X,Y} \propto e^{-2\bar{n}}$ for the cat, so that the SC can have a much larger noise bias under the same excitation number constraint.

In principle, one needs to consider the tradeoff between $\gamma_Z$ and $\alpha'$ and choose the optimal

$\eta$ depending on the tasks of interest. Smaller $\eta$ leads to better protection from excitation losses, which is preferred by, e.g., the idling operations. Larger $\alpha'$, on the other hand, leads to a larger noise bias and a widened effective dissipation gap, which can support faster operations, e.g., the bias-preserving CNOT gate introduce in the next section. Here, the effective dissipation gap is defined as the the excitation gap of the effective Hamiltonian $\hat{H}_{\text{eff}} = \frac{1}{2}\kappa_2 \hat{F}^\dagger \hat{F} = \frac{1}{2}\kappa_2 \hat{S}^{(r)}(\hat{a}^{\dagger 2} - \alpha'^2)(\hat{a}^2 - \alpha'^2)\hat{S}^\dagger(r)$, which characterizes the leakage rate and the non-adiabatic error rate under a Hamilotnian perturbation [109, 4, 5]. Since $\hat{H}_{\text{eff}}$ is the same as that for a cat with a displacement $\alpha'$ up to a unitary transformation, the effective dissipation gap for the dissipative SC is $2\kappa_2\alpha'^2$. In the following, we fix $\bar{n} = 4$ and $\eta = 1/4$ if not specified otherwise, which corresponds to a squeezing of $r = 1.32$ (11.5 dB). Such a parameter choice leads to $\gamma_Z \approx \kappa_1$, which removes the enhancement factor $\bar{n}$ present for the stabilized cats [4, 70] (for $\bar{n} = 4$). Meanwhile, $\alpha'^2 \approx \frac{3}{4}\bar{n}^2$ provides a sufficiently large noise bias and a large effective dissipation gap.

In Fig. 2.3(b), we benchmark the performance of our Auto-QEC scheme against loss errors by comparing it to the optimal recovery channel given by a semidefinite programming (SDP) method [1, 2, 3]. We consider the composite channel $\mathcal{N} = \mathcal{D} \cdot \mathcal{N}_\gamma \cdot \mathcal{E}$, where $\mathcal{E}$ denotes the encoding map from a qubit to the SC, $\mathcal{N}_\gamma$ denotes a Gaussian pure loss channel (corresponding to Eq. (2.23) with $\kappa_2 = \kappa_\phi = n_{\text{th}} = 0$) with loss probability $\gamma := \kappa_1 t$, and $\mathcal{D}$ denotes the recovery channel either using the autonomous QEC with the dissipator Eq. (2.21) or the optimal recovery channel. We use the entanglement fidelity $F_e := \langle\Phi^+|(\mathcal{N}\otimes\mathcal{I})(|\Phi^+\rangle\langle\Phi^+|)|\Phi^+\rangle$, where $|\Phi^+\rangle$ denotes a Bell state for the logical qubit and an ancilla qubit, as the error metric for the composite channel. We evaluate the entanglement infidelity (EI) $1-F_e$ as a function of the loss probability $\gamma$. Note that the EI is the objective function for the SDP. As shown in Fig. 2.3(b), the EI obtained using the Auto-QEC is close to the optimal EI, especially in the low-$\gamma$ regime, demonstrating that our proposed autonomous QEC scheme is close to optimal for correcting excitation loss errors. We note that it is crucial to have

the phase-flip $\hat{Z}_L$ correction in the dissipator $\hat{F}$ in order to correct the loss-induced phase-flip errors. Otherwise, a simple dissipator $\hat{S}(r)(\hat{a}^2 - \alpha'^2)\hat{S}^\dagger(r)$ directly generalized from the dissipative cat would still give an unsuppressed phase-flip rate $\gamma_Z = \kappa_1 \bar{n}$.

We note that the SC encoding also emerges as the optimal or close-to-optimal single-mode bosonic code through a bi-convex optimization (alternating SDP) procedure for a loss and dephasing channel with dephasing being dominant, as shown in Ref. [56].

### *2.2.3 Bias-preserving operations*

To apply the autonomously protected SC for computational tasks, we need to develop a set of gate operations that are compatible with the engineered dissipation. Furthermore, the operations should preserve the biased noise channel of the SC, which can be utilized for resource-efficient concatenated QEC and fault-tolerant quantum computing [67, 68, 93, 5, 69]. Following the literature for the cat and the pair-cat [4, 46, 110, 111], we develop a set of bias-preserving operations for the SC:

$$\mathcal{B} = \{\mathcal{P}_{|\pm\rangle_c}, \mathcal{M}_X, X, Z(\theta), ZZ(\theta), \text{ CNOT, Toffoli}\}, \tag{2.31}$$

which suffice for many concatenated QEC schemes (e.g. concatenation with the repetition codes or the surface codes). The detailed design of each operation is presented at the end of this section.

Overall, the bias-preserving operations for the SC can achieve much higher fidelity (lower dominant $Z$-type error rates) than those for the cat for the following two reasons: (1). The operations suffer less from the excitation loss errors, which are (partially) autonomously corrected. (2). The non-adiabatic errors are significantly suppressed by the $\hat{Z}_L$ correction in the dissipator $\hat{F}$ (see Eq. (2.21)) and the enlarged effective dissipation gap ($\propto \alpha'^2$), so that the gate operations could be implemented faster. In Fig. 2.4, we show the total $Z$-type

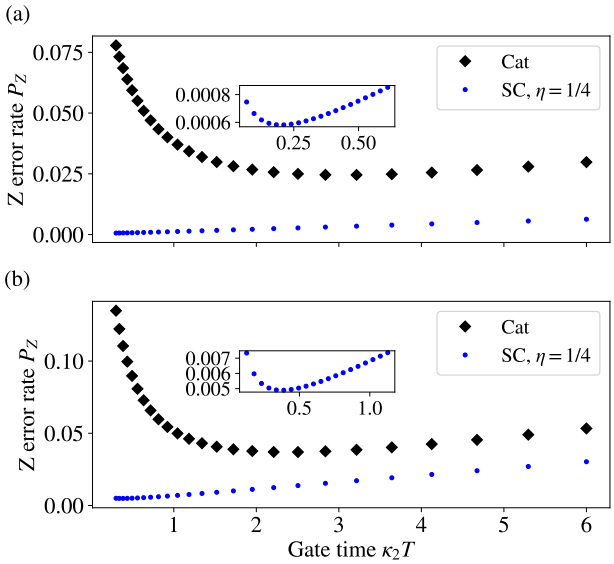


Figure 2.4: Error rates of the bias-preserving $Z$-rotation and CNOT for the stabilized SC. We plot the total $Z$ error probability of the $Z(\pi)$ gates (a) and the CNOT gates (b) versus the gate time. For the CNOT gate, $p_Z := p_{Z_c} + p_{Z_t} + p_{Z_c Z_t}$ is the sum of the control-mode, target-mode, and the correlated phase flip rates. $\kappa_1/\kappa_2$ is fixed at $10^{-3}$. The blue lines represent the gates on the cat qubits [4], and the red lines represent our proposed gates on the SC qubits with $\eta = 1/4$. $\bar{n}$ is chosen as 4 for both cat and SC. The insets are the zoomed-in error rates of the SC gates around the optimal gate times. As detailed in Methods, the $Z(\pi)$ gate requires a linear drive of strength $\frac{\pi}{8T}$ ($\frac{\pi}{4T}$) for the cat (SC). The CNOT gate requires a nonlinear coupling between the control and the target mode of strength $\frac{\pi}{8T}$ ($\frac{\pi}{4T}$) for the cat (SC).

error rates for the Z-axis rotation $Z(\theta)$ and the CNOT gate as a function of the gate time. Compared to the cat gates in Ref. [4] with the same $\bar{n}$, the SC $Z(\theta)$ (CNOT) gate can achieve a 42.0 (7.56) times reduction in the lowest error rates. While we have fixed $\eta = 1/4$ as mentioned in last section, it is not necessarily the optimal choice of the squeezing. In fact, with $\eta$ approaching 1/2, we obtain even lower errors at faster gate times.

We note that compared to the cat stabilized by $\hat{a}^2 - \alpha^2$ in the literature, a simple extension to a SC stabilized by $\hat{S}(r)(\hat{a}^2 - \alpha'^2)\hat{S}^\dagger(r)$ can also lead to improvement in the gate speed and fidelities due to the enlarged effective dissipation gaps. However, adding the extra phase flip in the dissipator brings a much more significant improvement due to the suppression of the loss-induced errors and the leading-order non-adiabatic errors. See Sec. 2.2.6 for more

details.

Now, we present the detailed design and error analysis for the bias-preserving gates for the SC in Eq. (2.31).

**$Z$-axis rotation $Z(\theta)$**

Similarly to the cat, the $Z$-axis rotation $Z(\theta)$ can be generated by a resonant linear drive

$$\hat{H}_Z = \frac{\theta}{4\alpha' T} e^r (\hat{a} + \hat{a}^\dagger), \tag{2.32}$$

in the presence of the engineered dissipation in Eq. (2.21) for a time $T$. In the subsystem basis, $\hat{H}_Z \approx \frac{\theta}{4\alpha' T} \hat{Z}_L \otimes (2\alpha' + \hat{\tilde{a}} + \hat{\tilde{a}}^\dagger)$. The total phase flip error probability of the $Z$ rotation is $p_Z = p_Z^{\mathrm{NA}}(T) + \kappa_1 \eta \bar{n} T$, where the second term represents the loss-induced phase flips and the first term represents the non-adiabatic errors due to the non-adiabatic excitation $\hat{Z}_L \otimes \hat{\tilde{a}}^\dagger$ in $\hat{H}_Z$. We note that compared to the parity-preserving dissipator $\mathcal{D}[\hat{I}_L \otimes \hat{\tilde{a}}]$, which is used in the literature for the cat (by applying a driven two-photon dissipation), the parity-flipping dissipator $\hat{F}$ in Eq. (2.21) can significantly reduce the non-adiabatic errors induced by $\hat{Z}_L \otimes \hat{\tilde{a}}^\dagger$. The reason is that the majority of the parity flips associated with the non-adiabatic transitions can be flipped back through the application of the dissipator. The remaining errors with a fraction $\xi$ leads to the residual non-adiabatic error $p_Z^{\mathrm{NA}}$ proportional to $\xi$ (see the previous Methods section). Under the adiabatic limit $\frac{\theta}{4\alpha' T} \ll 4\kappa_2 \alpha'^2$, the system's evolution under the dissipator $\hat{F}$ can be approximated by the dynamics of the density matrix $\hat{\rho}_{\mathrm{trunc}}$ with a truncated 2-level gauge basis:

$$\kappa_2 \mathcal{D}[\hat{F}]\hat{\rho} \approx 4\kappa_2 \alpha'^2 ((1-\xi)\mathcal{D}[\hat{Z}_L \otimes \hat{\tilde{a}}] + \xi \mathcal{D}[\hat{I}_L \otimes \hat{\tilde{a}}])\hat{\rho}_{\mathrm{trunc}}. \tag{2.33}$$

Performing first-order adiabatic elimination [112] on the gauge excited state results in an effective Z error rate $\frac{\xi \theta^2}{16\kappa_2 \alpha'^4 T}$. Notice that adiabatic elimination does not capture the higher-order errors and the result only holds under the adiabatic limit. A more accurate expression

can be derived through solving the ordinary differential equations of the two level system. As a result, the modified non-adiabatic error has the form:

$$p_Z^{\mathrm{NA}}(T) = \frac{\xi\theta^2}{16\kappa_2{\alpha'}^4T^2}(c_1T + c_2\frac{e^{-2\kappa_2{\alpha'}^2T}-1}{2\kappa_2{\alpha'}^2}). \tag{2.34}$$

Performing numerical fit, we obtain $c_1 = 1.5, c_2 = 1.8$.

**CNOT gate**

The CNOT gate is implemented by applying the engineered dissipation only on the control mode and a Hamiltonian term that drives a phase rotation on the target mode conditioned on the states of the control mode:

$$\begin{aligned} \tfrac{d}{dt}\hat{\rho} &= \kappa_2\mathcal{D}[\hat{F}_c]\hat{\rho} - i[\hat{H}_{\mathrm{CNOT}},\hat{\rho}], \\ \hat{H}_{\mathrm{CNOT}} &= \tfrac{\pi}{4\alpha' T}\left[e^r(\hat{a}_c+\hat{a}_c^\dagger) - 2\alpha'\right](\hat{a}_t^\dagger\hat{a}_t - {\alpha'}^2), \end{aligned} \tag{2.35}$$

where $\hat{F}_c$ denotes the engineered dissipator in Eq. (2.21) on the control mode. The noise terms are not shown for simplicity. We note that compared to the standard CNOT gate on the cat [4, 5], we turn off the dissipation on the target mode during the gate to circumvent the need for high-order coupling terms between the two modes. Although the target mode temporarily loses the protection against the excitation loss, we can still implement a high-quality gate if the gate time is short enough and the leakage on the target mode can be subsequently returned to the code space without introducing too many errors. Similar strategy and insights have been made in Ref. [72]. To deal with the non-adiabatic transitions on the target mode, which preserve the parity, we apply a parity-preserving dissipation $\kappa_2\mathcal{D}[\hat{S}(r)(\hat{a}_t^2-{\alpha'}^2)\hat{S}^\dagger(r)]$ on the target mode (while the control mode is, as always, protected by the parity-flipping dissipation) for a time $T_{\mathrm{cool}}$. In our simulations, we fix the cooling time $T_{\mathrm{cool}} = 8\times\frac{1}{4\kappa_2{\alpha'}^2}$ to ensure that the leakage is suppressed to below 0.5%. Using the

Pauli-twirling approximation, the Z-type errors of the CNOT gate are

$$\begin{aligned} p_{Z_c} &= \kappa_1 \eta \bar{n} \left(T + T_{\text{cool}}\right) + p_Z^{\text{NA}}(T), \\ p_{Z_t} &= \kappa_1 \bar{n} \left(\frac{T}{2} + T_{\text{cool}}\right), \\ p_{Z_c Z_t} &= \frac{1}{2} \kappa_1 \bar{n} T, \end{aligned} \tag{2.36}$$

where $p_{Z_c}, p_{Z_t}$ and $p_{Z_c Z_t}$ denote the $Z$ error on the control, target mode and the correlated $Z$ error, respectively. They sum to the total Z error probability $p_Z = \kappa_1 \bar{n}(1+\eta)(T+T_{\text{cool}}) + p_Z^{\text{NA}}(T)$. Note that, unlike the $Z$ rotation, the CNOT gate does not enjoy a full suppression of the loss-induced errors (by a factor $\eta$) due to the lack of the engineered dissipation on the target mode during the gate. The non-adiabatic error $p_Z^{\text{NA}}(T)$ on the control mode has a similar form as Eq. (2.34):

$$p_Z^{\text{NA}}(T) = \frac{\xi \pi^2}{16 \kappa_2 \alpha'^2 T^2} \left(1.5T + 0.6 \frac{e^{-2\kappa_2 \alpha'^2 T} - 1}{2\kappa_2 \alpha'^2}\right) \tag{2.37}$$

We also present the non-$Z$ error rate of the CNOT gate here. Note that the CNOT gate has a significantly larger non-$Z$ error rate than all other bias-preserving operations in $\mathcal{B}$. As discussed numerically in Ref. [5], the non-$Z$ error of a cat's CNOT gate scales approximately as $1.8 \frac{e^{-2\alpha^2}}{\alpha^2} \frac{1}{\kappa_2 T}$. For our CNOT gate on the SC, we find a similar expression

$$p_{X,Y} = 5.57 \frac{e^{-2\alpha'^2}}{\alpha'^2} \frac{1}{\kappa_2 T}, \tag{2.38}$$

in the regime where $\kappa_2 T > 1$. Note that for shorter gate time, we cannot find a simple expression for $p_{X,Y}$ and a numerical simulation of the gate has to be performed to determine $p_{X,Y}$.

**$ZZ$ rotation $ZZ(\theta)$**

The $ZZ$ rotation can be implemented by applying the following beam-splitter Hamiltonian

to two dissipatively stabilized modes: $\hat{H}_{ZZ} = \frac{\pi}{4\alpha'^2} e^{2r}(\hat{a}_1\hat{a}_2^\dagger + \hat{a}_1^\dagger\hat{a}_2)$.

**Toffoli gate**

The Toffoli gate can be implemented similarly to the CNOT gate:

$$\frac{d}{dt}\hat{\rho} = \kappa_2\mathcal{D}[\hat{F}_{c_1}] + \kappa_2\mathcal{D}[\hat{F}_{c_2}] - i[\hat{H}_{\text{Tof}}, \hat{\rho}],$$
$$\hat{H}_{\text{Tof}} = -\frac{\pi}{8\alpha'^2 T}((e^r\hat{a}_{c_1} - \alpha')(e^r\hat{a}_{c_2}^\dagger - \alpha') + \text{h.c.}) \times (\hat{a}_t^\dagger\hat{a}_t - \alpha'^2),$$

where $\hat{F}_{c_1}$ and $\hat{F}_{c_2}$ are the parity-flipping dissipator on the $c_1$ and $c_2$ mode, respectively.

**$X$ gate**

The X gate can be implemented by adiabatically tuning the phases of the stabilized code states $e^{-i\frac{\pi}{T}t\hat{a}^\dagger\hat{a}}|\text{SC}^{\pm}_{r,\alpha'}\rangle$ so that a $\pi$ phase rotation is implemented in time $T$. By adding a counterdiabatic drive $\hat{H}_X = \frac{\pi}{T}\hat{a}^\dagger\hat{a}$ the non-adiabatic effects are completely suppressed and the X gate can be implemented arbitrarily fast in principle.

**$X$-basis state preparation**

To prepare the X-basis eigenstate $|\text{SC}^{+}_{r,\alpha'}\rangle$, we initialize the system into the vacuum state $|0\rangle$, and turn on the parity-preserving dissipation $\kappa_2\mathcal{D}[\hat{S}(r)(\hat{a}^2 - \alpha'^2)\hat{S}^\dagger(r)]$. Since the parity is a conserved quantity, we obtain the even-parity state $|\text{SC}^{+}_{r,\alpha'}\rangle$ at a time $t \gg 1/4\kappa_2\alpha'^2$. We note that such a process is not protected from photon losses, since the dissipation does not correct the loss-induced stochastic phase flips. We will explore the implementation of more robust state preparation against losses in future work. One possible approach is to adiabatically inflate the SC from vacuum by tuning the dissipation (with an appropriate counter-adiabatic drive [70]), while maintaining the phase-flip correction.

**$X$-basis measurement**

To perform an X-basis measurement, we turn off the engineered dissipation and apply the standard QND bosonic parity measurement using a dispersive coupling between the bosonic mode and a transmon qubit [113, 114, 115]. A single photon loss during the dispersive coupling changes the parity and leads to stochastic measurement errors with a probability

that depends on when the loss jump happens. As such, the loss-induced measurement error probability associated with such a single measurement is $p_m = \frac{1}{2}\bar{n}\kappa_1 T$. To suppress the loss-induced error, we apply the QND parity measurement three times and apply the parity-flipping dissipation to correct the loss after each measurement. Finally, we perform a majority vote to obtain the measurement outcome. Such a protocol leads to a loss-induced measurement error probability $p'_m = \frac{3}{2}\eta\kappa_1\bar{n}T + \mathcal{O}((\kappa_1 T)^2)$, which maintains the $\eta$ suppression factor against photon loss.

### *2.2.4 Concatenated quantum error correction with squeezed cats*

With the bias-preserving operations, we can concatenate the SC with an outer discrete-variable code to suppress the logical error rates to the desired level. To compare the SC with the standard cat, we follow the literature and consider the concatenation with a repetition code [4] and a thin rotated surface code [5]. The surface-cat scheme can arbitrarily suppress the errors in a resource-efficient manner once the ratio between the loss rate $\kappa_1$ and the engineered dissipation rate $\kappa_2$ is below a certain threshold. The repetition-cat, on the other hand, cannot arbitrarily suppress the errors for a cat with constrained $\bar{n}$. Below a $\kappa_1/\kappa_2$ threshold, as the repetition code size increases, the logical $Z$ error rate is exponentially suppressed while the logical $X$ error is linearly amplified. Thus, a minimal total logical error rate is present.

The concatenated schemes with the cat face several challenges. First, the $\kappa_1/\kappa_2$ thresholds (e.g., $\sim 5 \times 10^{-4}$ for the surface-cat in Fig. 2.5(a). See also a comparable estimation in Ref. [5].) are very low because of the low-fidelity bias-preserving operations. Also, the minimal logical error probability of the repetition-cat (e.g., $\sim 10^{-2}$ for $\bar{n} = 4$, see Fig. 2.5(d)) is not low enough for fault-tolerant algorithms, except for cats with very large mean photon number, because of the limited noise bias.

In the following, we will show that these challenges can be overcome by using the dissi-

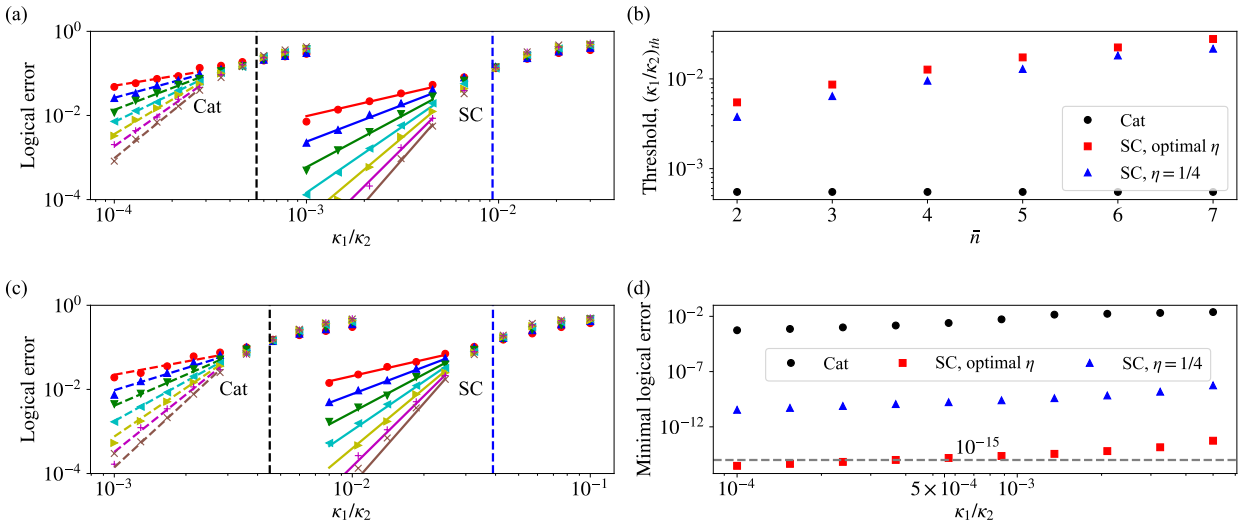


Figure 2.5: Logical errors of the SC and the cat concatenated with repetition codes or surface codes. (a) Surface code logical $Z$ error probabilities for a range of code distance $d_Z = 3, 5, 7, ..., 15$ (from red to brown) with fixed $d_X = 3$. The SC is fixed to $\bar{n} = 4, \eta = 1/4$. The dashed lines indicate the threshold values of $\kappa_1/\kappa_2$. (b) Surface code thresholds in $\kappa_1/\kappa_2$ varying with the average excitation number of the SC or the cat. (c) Repetition code logical $Z$ error probabilities for a range of code size $d_Z$. (d) Repetition code minimum total logical error probabilities, under the long gate time constraint $T \geq 1/\kappa_2$. Both the cat and the SC have an average excitation number $\bar{n} = 4$. The logical error probabilities for both the surface codes and the repetition codes are obtained from Monte Carlo simulations of $d_Z$ code cycles and one final round of perfect stabilizer measurement. We use the same minimum-weight-perfect-matching (MWPM) decoder as described in Ref. [5]

pative SC. The $\kappa_1/\kappa_2$ thresholds for both the surface code and the repetition code can be significantly improved by concatenating with the dissipative SC. Moreover, the repetition-SC can reach sufficiently low logical error probability $\sim 10^{-15}$ even with a small SC $\bar{n} = 4$ (see Fig. 2.5(d)). It is worth noticing that the thresholds for concatenated cat code shown in Fig. 2.5(a, c) are approximately independent of the size of the cat since the optimal CNOT gate error is independent of $\bar{n}$ for cat code.

We first consider the concatenation of the SC with a $d_X$ by $d_Z$ thin surface code. Similar to Ref. [5], we fix the X distance $d_X$ to 3, which suffices to suppress the logical $X$ error rate, and increase the Z distance $d_Z$ to suppress the logical $Z$ error rate. At fixed $\eta = 1/4$, we obtain the logical $Z$ error probability for $d_Z$ code cycles as a function of $\kappa_1/\kappa_2$ for different

$d_Z$, as shown in Fig. 2.5(a). The physical error rates of each physical operation involved in the surface-code QEC are presented in Supplementary Note 5 of Ref. [106]. We obtain a $\kappa_1/\kappa_2$ threshold at 0.93%, which is around 20 times higher than that of the surface-cat [5]. Note that by optimizing the choice of the squeezing, the maximum threshold we obtained for $\bar{n} = 4$ is around 1.2%. Moreover, in Fig. 2.5(b) we show that this threshold can be further increased to about 3% by increasing $\bar{n}$ to 7. Note that the $\kappa_1/\kappa_2$ threshold of the surface-cat remains almost the same when increasing $\bar{n}$. We attribute the increase of the $\kappa_1/\kappa_2$ threshold (for the concatenated SC schemes) to the reduced physical operation error rates (see the previous section).

Next, we consider the concatenation of the SC with a repetition code with size $d_Z$. As shown in Fig. 2.5(c), we obtain a 3.9% $\kappa_1/\kappa_2$ threshold for the logical $Z$ error rate (again, see Supplementary Note 5 in Ref. [106] for the physical error rates used for the simulation), which is roughly 9 times higher than that of the repetition-cat. Below the $\kappa_1/\kappa_2$ threshold, as previously mentioned, a minimal total logical error rate is present. To obtain the minimal total logical error rate (by optimizing over $d_Z$), we find approximate expressions for the logical $Z$ and $X$ error probabilities in the sub-threshold regime ($\kappa_1/\kappa_2 < 10^{-3}$):

$$\begin{aligned} p_L^Z &\approx 0.059 d_Z \left( \frac{p'_Z}{0.056} \right)^{0.48 d_Z}, \\ p_L^X &\approx 2 d_Z (d_Z - 1) p_{X,Y}, \end{aligned} \tag{2.39}$$

where $p'_Z := p_{Z_t} + p_{Z_c Z_t}$ denotes the sum of the target mode and the correlated phase-flip rate of the CNOT gate (phase flips on the control mode have negligible contribution to the logical error rate for the repetition code), $p_{X,Y}$ the total non-$Z$ error rates of the CNOT gate (the total rates of all the two-qubit Pauli errors that do not contain $Z$ terms). $p'_Z$ and $p_{X,Y}$ are in general functions of the CNOT gate time $\kappa_2 T$. To obtain simple expressions for them, we restrict the CNOT gate time to be $\kappa_2 T \geq 1$, which limits the nonadiabatic leakage

during the gate. In this regime, we have $p'_Z \approx \kappa_1 \bar{n} T$, $p_{X,Y} \approx 5.57 \times \frac{e^{-2\alpha'^2}}{\alpha'^2} \frac{1}{\kappa_2 T}$. Note that we do not see the contribution from the loss rate $\kappa_1$ to $p_{X,Y}$ since for fast gate, $p_{X,Y}$ is dominated by the nonadiabatic errors.

In Fig. 2.5(d), we plot the minimal total logical error probability $p_L = p_L^Z + p_L^X$ of the repetition-SC by optimizing $d_Z$ and $\kappa_2 T$ for $\bar{n} = 4$ and $\eta = 1/4$. As a comparison, we also include minimal logical error probabilities of the repetition-cat with $\bar{n} = 4$ using the physical error rates in Ref. [5]. When $\kappa_1/\kappa_2 \geq 10^{-3}$, the optimal gate error is no longer attained under the long gate time constraint, $\kappa_2 T \geq 1$ for the SC. Therefore, in that regime, the SC results can be understood as an upper bound of the minimum total logical error rates. For a practical noise ratio $\kappa_1/\kappa_2 = 10^{-3}$, the minimal logical error probability of the repetition-SC can reach $\sim 10^{-15}$, which suffices for many useful quantum computational tasks [17, 94]. In contrast, the logical error probability of the repetition-cat (with $\bar{n} = 4$) can only reach $\sim 10^{-2}$, which is far from being useful. Even with a larger cat size of $\bar{n} = 8$ as considered in, e.g., Ref. [5], the minimum logical error probability is still roughly $\sim 10^{-5}$ at such a noise ratio. To reach a similar level of logical error probability as the repetition-SC, we need either a much larger cat with $\bar{n} \gg 10$ (with the repetition code), or a more sophisticated outer code, e.g., the surface code. We attribute the drastic reduction in the minimal logical error rate of the repetition-SC to the significantly enhanced noise bias, or equivalently, the reduced physical bit-flip rates of the SC, which are exponentially suppressed by $\bar{n}^2$, instead of $\bar{n}$ for the cat.

### *2.2.5 Physical implementation*

In this section, we present the details of the two approaches implementing the dissipator in Eq. (2.21) for autonomously protecting ths SC qubits. Before describing our recipes, it is worth discussing the challenges involved here. The most straightforward method to realizing a generic Lindblad dissipator $\mathcal{D}[\hat{F}]$ is to couple the system to an auxiliary reservoir mode $c$

(with decay rate $\kappa_c$) via a coupling Hamiltonian $g(\hat{F}\hat{c}^\dagger + h.c.)$. In the limit where mode $c$ acts as a Markovian environment for the system, i.e. $\kappa_c \gg g$, we realize the target dissipator $\hat{F}$ with an effective dissipation rate $4g^2/\kappa_c$. For the dissipator in Eq. (2.21), this simple route requires a strong fourth-order nonlinear coupling, which has not been demonstrated yet due to the experimental challenges.

Here we present two approaches for realizing the desired nonlinear dissipator using accessible experimental resources: The first approach utilizes three nonlinearly-coupled bosonic modes, which can be physically realized in, e.g., superconducting circuits [86, 87]; The second approach couples a bosonic mode nonlinearly to a qutrit, which can be physically realized in, e.g., trapped-ion system [88].

**Implementation in superconducting circuits**

The first approach only requires third-order nonlinearities, which can be experimentally generated by four-wave fixing, to implement our desired dissipator, making use of a more structured engineered dissipation proposed in Ref. [108]. Under the subsystem decomposition of the storage mode $a$ encoding the SC, one can realize a general nonlinear dissipator of the form $\mathcal{D}[e^{-i\theta\hat{Z}_L} \otimes \hat{\tilde{A}}]$ (with an angle $\theta$), by coupling a gauge-mode operator $\hat{\tilde{A}}$ and an auxiliary mode $b$ to the input and output ports of a directional waveguide, respectively, and introducing a dispersive interaction between an auxiliary mode $b$ and the logical qubit: $\hat{H}_{\text{disp.}} = \frac{\lambda}{2}\hat{Z}_L\hat{b}^\dagger\hat{b}$. For the dissipator in Eq. (2.21), we choose $\hat{\tilde{A}} = \hat{\tilde{a}}$. The physical interactions (in the Fock basis) can be obtained from the mapping $\hat{\tilde{a}} \to \frac{1}{2\alpha'}\hat{S}(r)\left(\hat{a}^2 - \alpha'^2\right)\hat{S}^\dagger(r)$, and $\hat{Z}_L \to \frac{1}{2\alpha'}\hat{S}(r)\left(\hat{a} + \hat{a}^\dagger\right)\hat{S}^\dagger(r)$, which means that we need a nonlinear coupling between the storage mode $a$ and the waveguide port. While it is challenging to directly achieve this using e.g. a physical circulator, the directional dynamics can be synthetically engineered by adding another reservoir mode $c$. The whole setup is illustrated in Fig. 2.6(a), whose

dynamics is given by master equation

$$\frac{d}{dt}\hat{\rho} = -i[\hat{H}_{\text{disp.}} + \hat{H}_{\text{tun.}}, \hat{\rho}] + \kappa_c \mathcal{D}[\hat{c}], \tag{2.40}$$

where the tunnel coupling Hamiltonian $\hat{H}_{\text{tun.}}$ of the total system-reservoir is given by

$$\hat{H}_{\text{tun.}} = J_{ab}\hat{\tilde{a}}^\dagger \hat{b} + (J_{ac}\hat{\tilde{a}} - iJ_{bc}\hat{b})\hat{c}^\dagger + h.c., \tag{2.41}$$

$$J_{ab} = \sqrt{\Gamma_a \Gamma_b}/2,\ J_{ac} = \sqrt{\Gamma_a \kappa_c}/2,\ J_{bc} = \sqrt{\Gamma_b \kappa_c}/2. \tag{2.42}$$

In the regime where the joint $b, c$ modes act as a Markovian reservoir for mode $a$, i.e. $\kappa_c \gg \sqrt{\Gamma_a \Gamma_b}$ and $\Gamma_b \gg \Gamma_a$, we can adiabatically eliminate both $b$ and $c$ to obtain an effective dissipator (using the effective operator formalism [112]), as

$$\frac{d}{dt}\hat{\rho} = \Gamma_a \mathcal{D}\left[\frac{i\lambda \hat{Z}_L - \Gamma_b}{i\lambda \hat{Z}_L + \Gamma_b}\hat{\tilde{a}}\right]\hat{\rho}. \tag{2.43}$$

Setting $\lambda = \Gamma_b$, we obtain the desired dissipator $\hat{Z}_L \otimes \hat{\tilde{a}}$ to stabilize the SC (see Supplementary Note 6 of Ref. [106] for a detailed derivation).

When deriving Eq. (2.43), we require the physical setup Eq. (2.40) to operate in the regime where adiabatic elimination remains valid. It is thus natural to ask what are the imperfections given realistic physical parameters, i.e. when the decay rates $\kappa_c$, $\Gamma_b$ of auxiliary modes $b, c$ cannot be infinitely large. In that case, one can show the dominating error due to finite reservoir bandwidth is due to the finite decay rates $\kappa_c$ and $\Gamma_b$, and it is preferable to set $\kappa_c \sim \Gamma_b$ to optimize over hardware resources (see Supplementary Note 6 of Ref. [106] for details). In this regime, the extra error introduced by physical implementation is determined by the ratio $\Gamma_a/\Gamma_b$, which heuristically describes the branching ratio between the logical qubit population that does not undergo the parity flip (uncorrected error) and the population that does (corrected error) whenever a gauge mode excitation decays into the environment.

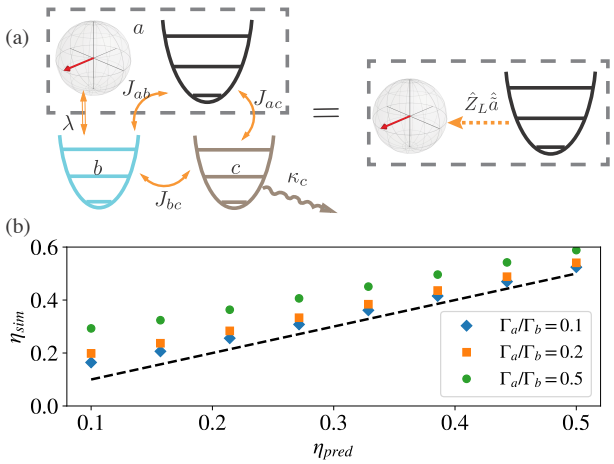


Figure 2.6: Physical realization of the stabilized SC in superconducting circuits. (a) Realization of the parity-flipping dissipator $\hat{Z}_L \otimes \hat{\tilde{a}}$ using three nonlinearly coupled bosonic modes. (b) Comparison between the numerically extracted $\eta$ ($\eta_{\text{sim}}$) and the theoretically predicted $\eta$ ($\eta_{\text{pred}}$ in Eq. (2.20)) for a range of finite $\Gamma_a/\Gamma_b$. The dashed line indicates the ideal case where $\eta_{\text{sim}} = \eta_{\text{pred}}$.

More specifically, as shown in Supplementary Note 6 of Ref. [106], we can approximately derive the discrepancy between the desired suppression factor for the loss-induced phase flip rate $\eta_{\text{pred}}$ (using Eq. (2.20)) and the numerically extracted (achievable) value $\eta_{\text{sim}}$, as $\eta_{\text{sim}} - \eta_{\text{pred}} = (1 - \eta_{\text{pred}})(\Gamma_a/2\Gamma_b)$. As shown in Fig. 2.6(b), by setting $\Gamma_a/\Gamma_b = 0.1$, we can realize the desired $\eta$ within 50% accuracy.

To make the required nonlinearity more clear, we can also explicitly write down the physical Hamiltonian Eq. (2.40) in the Fock basis:

$$\begin{aligned} \hat{H}_{\text{tun.}} &= \frac{J}{2\alpha'}\hat{a}_s^{\dagger 2}(\hat{b} + \hat{c}) - \frac{J}{2}\alpha'(\hat{b} + \hat{c}) - \frac{i}{2}\kappa_c \hat{b}\hat{c}^\dagger + h.c., \\ \hat{H}_{\text{disp.}} &= \frac{\kappa_c e^r}{4\alpha'}(\hat{a} + \hat{a}^\dagger)\hat{b}^\dagger\hat{b}, \end{aligned} \tag{2.44}$$

where $\hat{a}_s = \cosh r\hat{a} + \sinh r\hat{a}^\dagger$ is the squeezed annihilation operator, and $J := \sqrt{\Gamma_a\kappa_c}/2$. We have assumed that $\kappa_c = \Gamma_b = \lambda$. $\hat{H}_{\text{tun.}}$ involves several cubic nonlinear couplings between the $a, b$ modes and between the $a, c$ modes. In addition, $\hat{H}_{\text{tun.}}$ requires a resonant linear coupling between the $b, c$ modes and some linear drives with strength $\frac{J\alpha'}{2}$ on the $b, c$

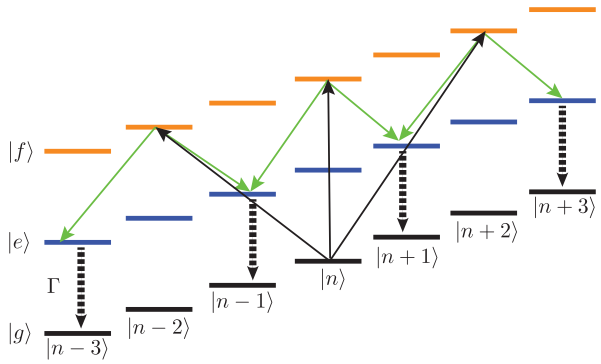


Figure 2.7: Laser configuration for the coupling Hamiltonian in Eq. (2.46) for implementing the SC in trapped-ion system. The motional mode of the ion is coupled to three internal states via the sideband transitions, represented by the black and green arrows. Starting from $|g\rangle \otimes |\psi\rangle$ ($|\psi\rangle$ is an arbitrary motional state), the system goes through a two-step coherent transition $|g\rangle \otimes |\psi\rangle \to |f\rangle \otimes \hat{F}_1|\psi\rangle \to |e\rangle \otimes \hat{F}_2\hat{F}_1|\psi\rangle$ (indicated by the black and the green solid arrows, respectively) and decays rapidly to $|g\rangle \otimes \hat{F}_2\hat{F}_1|\psi\rangle$ (indicated by the black dashed arrows). Here $\hat{F}_1 \propto \hat{S}(r)(\hat{a}^2 - \alpha'^2)\hat{S}^\dagger(r)$ and $\hat{F}_2 \propto \hat{S}(r)\hat{a}\hat{S}^\dagger(r)$. Adiabatically eliminating the $|e\rangle, |f\rangle$ states, we obtain the effective dissipator on the motional mode $\hat{F} = \hat{F}_2\hat{F}_1$.

modes that pump energy into the system. Note that all the nonlinear terms are cubic, which have been experimentally demonstrated in superconducting circuits [86, 87]. The maximum nonlinear coupling strength is $J_m := J \sinh 2r/2\alpha'$. Now we can write $\kappa_2$ as a function of $J_m$ and $\kappa_c$. Comparing the dissipator $\sqrt{\Gamma_a}\hat{Z}_L \otimes \hat{\tilde{a}}$ in Eq. (2.43) with the dissipator $\sqrt{\kappa_2}\hat{F} \approx \sqrt{\kappa_2}2\alpha'\hat{Z}_L \otimes \hat{\tilde{a}}$, we have $\kappa_2 = \Gamma_a/4\alpha'^2$, so that $\kappa_2 = \frac{4J_m^2}{\kappa_c \sinh^2 2r}$. If we take $\Gamma_a = \epsilon_0\kappa_c$ (and correspondingly, $\kappa_c = \frac{4\alpha' J_m}{\sqrt{\epsilon_0}\sinh 2r}$), where $\epsilon_0 < 1$ is related to the adiabatic elimination condition discussed above, we can obtain $J_m = \alpha' \sinh 2r\kappa_2/\sqrt{\epsilon_0}$.

**Implementation in trapped-ion systems**

Now we present the second approach for implementing the dissipator $\hat{F} = \frac{1}{\alpha'}\hat{S}(r)\hat{a}(\hat{a}^2 - \alpha'^2)\hat{S}^\dagger(r)$ using a coupled boson-qutrit system. Note that a simpler dissipator stabilizing a cat $\hat{a}^2 - \alpha^2$ was obtained using a coupled boson-qubit system in trapped-ion platform in Ref. [88]. However, the dissipator $\hat{F}$ cannot be directly engineered using their approach since there are many frequency-degenerate terms, e.g., $\hat{a}$ and $\hat{a}^\dagger\hat{a}^2$, that cannot be independently controlled by a single sideband drive. To resolve this, we generalize their approach by introducing a third internal level of the ion, and implementing the dissipator $\hat{F}$ in two steps

associated with different electronic transitions. Specifically, we use the motional mode of the ions in a 1D harmonic trap as the bosonic mode, which is coupled to three internal levels $|g\rangle, |e\rangle$ and $|f\rangle$ via several laser beams:

$$\frac{d}{dt}\hat{\rho} = -i[\hat{H}_{\text{eff}}, \hat{\rho}] + \mathcal{J}\hat{\rho}, \tag{2.45}$$

where $\hat{H}_{\text{eff}} = \nu\hat{a}^\dagger\hat{a} + \omega_e|e\rangle\langle e| + \omega_f|f\rangle\langle f| + \frac{1}{2}\Omega_0(|f\rangle\langle g|e^{-i\omega_f t} + h.c.) + \hat{H}_{\text{coup}} - i\frac{\Gamma}{2}|e\rangle\langle e|$, with

$$\begin{aligned}\hat{H}_{\text{coup}} = &\sum_{i=1}^{3}\Omega_i \cos\left[\eta_0(\hat{a}+\hat{a}^\dagger)\right](|f\rangle\langle g|e^{-i(\omega_f+\delta_i)t} + h.c.) \\ &+ \sum_{i=4}^{5}\Omega_i \sin\left[\eta_0(\hat{a}+\hat{a}^\dagger)\right](|e\rangle\langle f|e^{-i(\omega_e-\omega_f+\delta_i)t} + h.c.),\end{aligned} \tag{2.46}$$

and

$$J\hat{\rho} = \Gamma\int_{-1}^{1} duN(u)e^{-i\eta_0 u\left(\hat{a}+\hat{a}^\dagger\right)} \times |g\rangle\langle e|\hat{\rho}|e\rangle\langle g|e^{i\eta_0 u\left(\hat{a}+\hat{a}^\dagger\right)}. \tag{2.47}$$

Here $\nu$ is the trap frequency, $\eta_0$ the Lamb-Dick parameter, $\Gamma$ the engineered decay rate from $|e\rangle$ to $|g\rangle$, and $N(u)$ the normalized dipole pattern. $\hat{H}_{\text{coup}}$ describes the coupling between the motional mode and the internal states, illustrated in Fig. 2.7, and $\mathcal{J}\hat{\rho}$ describes the spontaneous emission of the ion from $|e\rangle$ to $|g\rangle$ and its associated momentum kicks. The drive with amplitude $\Omega_0$ in $\hat{H}_{\text{eff}}$ comes from a laser that is coupled to the ion along a constrained (transverse) direction, thereby only driving the internal transitions. By tuning the laser detunings $\delta_1 = -2\nu, \delta_2 = 2\nu, \delta_3 = 0, \delta_4 = -\nu$, and $\delta_5 = \nu$, and choosing appropriate driving strength $\{\Omega_i\}$ (see Supplementary Note 6 of Ref. [106] for details), we can obtain a coupling Hamiltonian (neglecting the fast-rotating terms):

$$\hat{H}_{\text{coup}} = \Omega'_{gf}\hat{S}(r)(\hat{a}^2 - \alpha'^2)\hat{S}^\dagger(r)|f\rangle\langle g| + \Omega'_{ef}\frac{1}{\alpha'}\hat{S}(r)\hat{a}\hat{S}^\dagger(r)|e\rangle\langle f| + h.c. \tag{2.48}$$

In the regime where $2\alpha'\Omega'_{gf} \ll \Gamma$, $\Omega'_{gf} \ll \Omega'_{ef}$, we can obtain a reduced dynamics on the

motional mode by adiabtically eliminating the $|e\rangle, |f\rangle$ states:

$$\frac{d}{dt}\hat{\rho}_m = \kappa_2 \mathcal{D}[\hat{F}]\hat{\rho}_m, \tag{2.49}$$

where $\hat{\rho}_m$ is the reduced density matrix on the motional mode. Through numerically simulations we find that we can obtain the dissipator $\hat{F}$ with the desired rate by setting $\Omega'_{ef} = 0.5\Gamma, \Omega'_{gf}/\Omega'_{ef} = 1/20$. A large $\kappa_2$, therefore, demands large $\Gamma$ and driving strength. Note that we have assumed that $\Gamma$ and $\{\Omega_i\}, i = 1, 2, 3, 4, 5$ are much smaller than $\nu$, so that the off-resonant terms can be safely neglected (secular approximation). In practice, however, one might be able to go beyond this weak-drive regime by carefully cancelling the effects from the off-resonant terms. We have also neglected the effects from the momentum kicks here, which only lead to a small increase in the phase-flip suppression factor $\eta \to \eta + \mathcal{O}(\eta_0^2)$. See Supplementary Note 6 of Ref. [106] for a more detailed analysis. We stress that our proposed approach requires the same order of nonlinearity as that required by a two-component cat, which has been considered to be feasible in trapped-ion system [88].

### *2.2.6 Comparison with related schemes*

Although in this work we benchmark the performance of the concatenated codes as a function of $\kappa_1/\kappa_2$ for both the cat and the SC, it might be of different difficulty level to engineer the same dissipation rate $\kappa_2$ for the cat and the SC, depending on the hardware implementation. Therefore, we can compare the performance of the concatenated codes as a function of $\kappa_1/M$, where $M$ is the physical rate that is most challenging to engineer in practice. Here we focus on the implementation in superconducting circuits.

For example, a potential hardware challenge is to engineer strong nonlinear couplings. In this case, we can compare the concatenated codes as a function of $\kappa_1/J_m$, where $J_m$ denotes the largest nonlinear coupling strength required. For the cat, $J_m$ is simply given

by $g_2$, the strength of the two-photon exchange Hamiltonian $g_2\left(\hat{a}^2\hat{b}^\dagger + h.c.\right)$. Assuming an adiabaticity constraint $\epsilon$, the lossy mode can be adiabatically eliminated, and $J_m = \sqrt{\bar{n}}\kappa_2/2\epsilon$ [5, 86]. For the SC, as shown in the Methods, the maximum nonlinear coupling strength is given by $J_m = \alpha' \sinh 2r\kappa_2/\sqrt{\epsilon_0}$. Here, $\epsilon_0$ is the relevant adiabatic condition for our proposed stabilization scheme using three bosonic modes.

Using these relations, we can change the horizontal axis in Fig. 2.5(a) to $\kappa_1/J_m$ and obtain about a 3.5 times increase in the $\kappa_1/J_m$ threshold for the surface-SC compared to the surface-cat. Furthermore, results in Fig. 2.5(a) are obtained by optimizing the parameters, such as the squeezing $r$ and the gate times, with the target function set to be the threshold in $\kappa_1/\kappa_2$ . If the target function is set to the threshold in $\kappa_1/J_m$ instead, it is likely that the optimal code parameters are different, and the corresponding threshold could be further increased. Based on these considerations, we expect that the SC should maintain advantages over the cat even considering the experimental constraints (which will be hardware-specific) in the circuit level. We leave it to future work on optimizing the hardware design and quantifying the hardware-specific improvement.

To better understand the novelty and necessity of the partity-flipping dissipator $\hat{F}$ we introduced in Eq. (2.21), we compare it with a parity-preserving dissipator

$$\hat{F}' = \hat{S}(r)(\hat{a}^2 - \alpha'^2)\hat{S}^\dagger(r) \approx 2\alpha' \hat{I}_L \otimes \hat{\tilde{a}}, \tag{2.50}$$

which is a straightforward extension from $\hat{a}^2 - \alpha^2$ that stabilizes the cat. Such a dissipator was recently considered in Ref. [116] for stabilizing the SC. We show that the extra phase-flip correction in $\hat{F}$ is essential for reducing SC's error rate in both the memory level and gate operations, which then leads to better logical performance in the concatenated level.

In the memory level, the change of a parity flip on the dissipator does not affect the bit-flip error rate we derived in Eq. (2.25). So a SC stabilized by $\hat{F}'$ can also have a favorable scaling between the minimal bit-flip rate and $\bar{n}$: $\gamma_{X,Y} \propto e^{-2\bar{n}^2}$. Nevertheless, $\hat{F}'$ lacks the

parity flip $Z_L$ that corrects the detectable portion of the loss-induced errors, as shown clearly from Fig. 2.2 (the missing of the blue arrow). Therefore, a SC stabilized by $\hat{F}'$ is not capable of correcting the loss errors. As such, it suffers from the same phase-flip error rate as a cat, $\gamma_Z = \kappa_1\bar{n}$.

Regarding the gate operations, we take the Z rotation and the CNOT gate as examples. For the Z rotation, a SC stabilized by $\hat{F}'$ only enjoys a suppression in the non-adiabatic errors by the the increased adiabatic gap, $4\kappa_2\alpha'^2$, compared to conventional cat of the same $\bar{n}$. In contrast, a SC stabilized by $\hat{F}$ corrects the leading-order non-adiabatic error in $1/\alpha'^2$, since the the extra $\hat{Z}_L$ in $\hat{F}$ compensates the parity-flip associated with the non-adiabatic transition (to the leading order). The residual errors are proportion to the correction factor, $\xi \propto 1/\alpha'^2$ (see Eq. (2.34)). Therefore, while the minimal $Z(\theta)$ gate error for the SC with $\hat{F}'$ is roughly suppressed by a factor $1/\bar{n}$ compared to the cat, that for the SC with $\hat{F}$ is suppressed by an $1/\bar{n}^2$ factor (see Table 2.1).

| Normalized gate error | SC with $\hat{F}'$ | SC with $\hat{F}$ |
|---|---|---|
| $Z(\theta)$ | $1/(\bar{n}+1)$ | $\sim \bar{n}^{-2}$ |
| CNOT | $2/\sqrt{\bar{n}+1}$ | $\sim \bar{n}^{-3/2}$ |

Table 2.1: Optimal gate error rate of the SC gates compared to the cat. All errors are normalized by the optimal gate errors of the cat, which are given by $p_{Z(\theta)} = \frac{\theta}{2}\sqrt{\frac{1}{\bar{n}}\frac{\kappa_1}{\kappa_2}}$ and $p_{\text{CNOT}} = \frac{\pi}{2\sqrt{2}}\sqrt{\frac{\kappa_1}{\kappa_2}}$ [5]. The definitions of $\hat{F}$ and $\hat{F}'$ are given in Eq. (2.21) and Eq. (2.50) respectively. The optimal gate errors for SC are reached at $\eta \approx \frac{1}{2}$. The optima $Z(\theta)$ gate time for SC with $\hat{F}'$ and $\hat{F}$ are approximately $\frac{\pi}{4\sqrt{\kappa_1\kappa_2}}\bar{n}^{-5/2}$ and $\frac{\pi}{4\sqrt{3\kappa_1\kappa_2}}\bar{n}^{-7/2}$ respectively. The gate times for CNOT are approximately $\frac{\pi}{4\sqrt{2\kappa_1\kappa_2}}\bar{n}^{-3/2}$ and $\frac{\pi}{12\sqrt{\kappa_1\kappa_2}}\bar{n}^{-5/2}$ respectively. Since the cooling time is mostly assumed to be constant in our gate scheme, it is neglected for simplicity. Here, for simplicity, we only provide the scaling of the gate errors with $\bar{n}$ for the SC.

The errors of CNOT operation can be analyzed in a similar fashion. Due to the enlarged adiabatic gap, the minimal $Z$ error rate of our SC gate with $\hat{F}'$ is a factor of $\frac{2}{\sqrt{\bar{n}+1}}$ smaller than that of the cat gate [4]. For the mean excitation number we consider, $\bar{n} = 4$, this

factor is only slightly less than 1. However, with the parity-flipping dissipator $\hat{F}$, the gate error enjoys a $\eta$ suppression in the loss errors and an additional $\propto 1/\alpha'^2$ suppression in the non-adiabatic error. Combining these advantages, the CNOT gate error ratio with that of the cat roughly scales as $\bar{n}^{-3/2}$ (see Table 2.1).

Since the fault-tolerant threshold is mostly limited by errors of the CNOT and the idling operation, the thresholds of the concatenated SC schemes using $\hat{F}'$ is comparable to that of the concatenated cat scheme even at optimal squeezing for small mean excitation number. As such, having the extra phase-flip correction in the dissipator $\hat{F}$ is crucial for concatenated QEC and fault-tolerant quantum computing.

The stabilized cat qubits have been considered as a candidate for hardware-efficient, fault-tolerant, and scalable computation tasks in superconducting circuits [5, 69]. The dissipative SC, which we show has an overall advantage over the cat, could play an important role along this direction.

The dissipative SC could also find its application in trapped-ion systems. On the one hand, encoding into the motional states of the ions provides an alternative approach for storing and protecting the quantum information. How to process the information (e.g., implementing quantum gates) remains to be explored. On the other hand, if the information is stored in the internal states of the ions (the conventional approach), the bosonic codes like the SC could lead to more robust information processing. One could utilize multi-species ions [117, 118] with multiple levels [119] and dissipatively protect the motional modes while leaving a subset of the ions' internal states that carry the information intact. The protected motional modes can, for instance, be used for scalable, parallel, and high-quality entangling gates mediated by localized phonon modes [120].

## 2.3 Fault-tolerant operation of rotational-symmetrical codes with discrete-variable ancillae

While robust quantum memories based on single-mode bosonic codes have been experimentally demonstrated with improved memory lifetime [65, 121, 122], realizing error-corrected operations on these bosonic qubits remains a formidable task. One of the primary complexities stems from the weak non-linear interactions inherent in bosonic modes, necessitating the use of discrete-variable ancillae in systems such as circuit quantum electrodynamics (circuit QED) platform [123, 124]. However, a significant challenge arises in this hybrid system, as errors in the ancillae tend to propagate back to the bosonic mode, potentially compromising the encoded quantum information [125]. To address this issue, several methods have been developed to maintain precise control over the bosonic mode even in the presence of noisy ancillary systems [126, 7, 6]. Nevertheless, a comprehensive fault-tolerance framework for this hybrid system, along with guidelines for constructing fully fault-tolerant protocols using advanced quantum control concepts, remains conspicuously absent. Consequently, while universal error-detection operations on bosonic qubits have been constructed [127, 128] and demonstrated [129], achieving a complete set of error-corrected operations has remained a significant challenge.

In this chapter, we bridge this gap by introducing a fault-tolerance framework tailored to the hybrid system composed of bosonic data modes and discrete-variable ancillae. Inspired by concatenated qubit codes [26], we identify essential properties for gadgets encoded in bosonic codes (referred to as "level-1" gadgets) in Sec. 2.3.2. These properties play a crucial role in determining the fault tolerance of a level-1 circuit, where the overall failure probability must be suppressed to a certain order of the physical error rate. Furthermore, we demonstrate how the defined fault tolerance can be achieved through the integration of bosonic QEC with compatible quantum control techniques. Specifically, in Sec. 2.3.3, we establish a connection between a generalized version of path-independent control [6] (referred to as GPI) and fault

tolerance, highlighting the importance of GPI operations as fundamental building blocks for error-corrected gadgets.

As an application of these fault-tolerant tools, in Sec. 2.3.4, we construct universal error-corrected gadgets using GPI operations for the four-legged cat qubit [60, 130, 38]. These gadgets can tolerate a single photon loss and an arbitrary ancilla fault, while only relying on dispersive coupling between bosonic modes and ancillae [131, 75, 7] and beam-splitter (BS) coupling between bosonic modes [132, 133]. Importantly, these coupling mechanisms have been experimentally demonstrated with strong coupling strengths. Each level-1 logical qubit, encoded in a four-legged cat code, utilizes only a single bosonic mode and a three-level ancilla, featuring the hardware efficiency of bosonic QEC. We numerically demonstrate the second-order error suppression for the level-1 gadgets. Moreover, we show that using a teleportation gadget that pumps energy into the system and suppresses phase-rotation errors, a robust cat-encoding memory is feasible even in the presence of finite $\chi$ mismatches in the circuit quantum electrodynamics (cQED) platform with current experimental parameters [7]. Compared to former constructions involving only bosonic modes [59], our constructed operations can be more practical by leveraging the strong interaction between the bosonic modes and the qudit ancillae in e.g., the cQED platform. See Sec. 2.3.6 for more details.

Finally in Sec. 2.3.5, we present a practical and hardware-efficient architecture for fault-tolerant quantum computing by concatenating the four-legged cat qubits with an outer qubit code. While we primarily focus on the four-legged cat code throughout this work, we discuss in Sec. 2.3.6 that the fault-tolerant schemes developed herein can be readily adapted to other rotation-symmetric bosonic codes [59].

### *2.3.1 Hybrid system of bosonic data modes and descrete-variable ancillae*

We first introduce some notations. We denote $[k] := \{1, 2, \cdots, k\}$ as the set of integers from 1 to $k$. We denote $\left[\int_{t_h} dt_h\right]_{h\in[k]} := \int_{t_k} d_{t_k} \int_{t_{k-1}} dt_{k-1} \cdots \int_{t_1} dt_1$ as the multiple integral over

variables in $\{t_h\}_{h\in[k]}$, and similarly $\left[\sum_{a_h}\right]_{h\in[k]} := \sum_{a_k}\sum_{a_{k-1}}\cdots\sum_{a_1}$ as the sum over variables in $\{a_h\}_{h\in[k]}$. We denote $A \propto B$ if there exists some $c \in \mathbb{C}$ such that $A = cB$. We denote $\mathcal{T}$ as the time ordering operator.

## Bosonic code, four-legged cat

Single-mode bosonic error-correcting codes encode logical information into a subspace of the infinite-dimensional Hilbert space of an oscillator. Among them, the four-legged cat code [60, 130, 38] encodes a single logical qubit and has codewords

$$|\mu_L\rangle = c_\mu \left[|\alpha\rangle + |-\alpha\rangle + (-1)^\mu(|i\alpha\rangle + |-i\alpha\rangle)\right], \tag{2.51}$$

where $\mu = 0/1$, $|\gamma\rangle$ denotes a coherent state with an amplitude $\gamma \in \mathbb{C}$, and

$$c_\mu = 1/(2\sqrt{2\exp\bigl(-|\alpha|^2\bigr)(\cosh|\alpha|^2 + (-1)^\mu\cos|\alpha|^2)})$$

are normalization constants. Given any quantum code encoding a single logical qubit, we denote $P_c := |0_L\rangle\langle 0_L| + |1_L\rangle\langle 1_L|$ as the projection onto the codespace, and $\bar{X}_c, \bar{Y}_c, \bar{Z}_c$ the logical $X$-, $Y$-, $Z$-Pauli operators respectively.

The capability of an error-correcting code to correct a given set of errors $\mathbf{E}$ is given by the Knill-Laflamme (KL) condition [104]: $P_c E_i^\dagger E_j P_c \propto P_c$ for any $E_i, E_j \in \mathbf{E}$. More specifically, we can evaluate the $2\times 2$ QEC matrix $\epsilon^c_{jk}$ for any pair of errors $E_j, E_k$ [38]:

$$P_c E_j^\dagger E_k P_c = \epsilon^c_{jk}, \tag{2.52}$$

where $\epsilon^c_{jk}$ can be parametrized as $\epsilon^c_{jk} = c^c_{jk}P_c + x^c_{jk}\bar{X}_c + y^c_{jk}\bar{Y}_c + z^c_{jk}\bar{Z}_c$, where $c^c_{jk}, x^c_{jk}, y^c_{jk}, z^c_{jk} \in \mathbb{C}$. The KL condition is satisfied if $x^c_{jk} = y^c_{jk} = z^c_{jk} = 0$ for any $j$ and $k$.

Consider the four-legged code and an error set containing a single-photon loss $\mathbf{E} = \{I, a\}$,

where $a$ denotes the annihilation operator. First, we have $P_c a P_c = 0$, indicating that a single-photon loss is perfectly detectable. Second,

$$P_c a^\dagger a P_c = \bar{n} P_c + \frac{\delta n}{2} \bar{Z}_c, \tag{2.53}$$

where $\bar{n} := (\langle 0_L | a^\dagger a | 0_L \rangle + \langle 1_L | a^\dagger a | 1_L \rangle)/2$ denotes the mean photon number and $\delta n := \langle 0_L | a^\dagger a | 0_L \rangle - \langle 1_L | a^\dagger a | 1_L \rangle$ denotes the photon number difference between the two codewords. For an arbitrary $\alpha$, $\delta n \neq 0$, indicating the a single photon loss is not perfectly correctable. However, $\delta n = O(e^{-2\alpha^2})$ as $\alpha \gg 1$ and a single-photon loss is approximately correctable for large $\alpha$. Furthermore, $\delta n = 0$ is exactly satisfied at a discrete set of finite $\alpha$ [8], which we refer to as sweet spots. Similarly, one can show that for a continuous set of phase-rotation errors $\mathbf{R} = \{e^{i\theta a^\dagger a}\}_{\theta \in [-\theta_m, \theta_m]}$, the KL condition is approximately satisfied for large $\alpha$ if $\theta_m < \pi/4$ [59]. First, $P_c e^{-i\theta_1 a^\dagger a} e^{i\theta_2 a^\dagger a} P_c = c_{12} P_c + z_{12} \bar{Z}_c$ for any $\theta_1, \theta_2 \in \mathbf{R}$ since $e^{i(\theta_2 - \theta_1) a^\dagger a}$ preserves the photon number. Next,

$$\begin{aligned} z_{12} &= \left( \langle +_L | e^{i(\theta_2 - \theta_1) a^\dagger a} | -_L \rangle + \langle -_L | e^{i(\theta_2 - \theta_1) a^\dagger a} | +_L \rangle \right) /2 \\ &\approx \left( \langle i\alpha | \alpha e^{i(\theta_2 - \theta_1)} \rangle + \langle -i\alpha | \alpha e^{i(\theta_2 - \theta_1)} \rangle \right) /2 + h.c., \end{aligned} \tag{2.54}$$

where the the approximation utilizes that $|+_L\rangle \approx (|\alpha\rangle + |-\alpha\rangle)/\sqrt{2}$ and $|-_L\rangle \approx (|i\alpha\rangle + |-i\alpha\rangle)/\sqrt{2}$ for large $\alpha$. Obviously, $z_{12} \to 0$ as $\alpha \gg 1$ as long as $|\theta_2 - \theta_1| \neq \pi/2$, which holds if $\theta_m < \pi/4$.

To conclude, the four-legged cat code can approximately correct a single photon loss and a continuous set of phase rotations with amplitude smaller than $\pi/4$ (for large $\alpha$). In fact, cat codes serve as numerically optimized codes for certain regimes of a bosonic channel with both loss and dephasing errors [56].

## Open quantum system and Markovian quantum evolution

A noisy Markovian evolution of a quantum system is described by a Lindblad master equation:

$$\frac{d\rho}{dt} = \mathcal{L}(t)\rho = -i[H(t), \rho] + (\sum_j \mathcal{D}[\sqrt{\gamma_j} J_j])\rho, \tag{2.55}$$

where $H(t)$ is the system Hamiltonian and $\mathcal{D}[O] = O \bullet O^\dagger - \frac{1}{2}\{O^\dagger O, \bullet\}$ is a Lindblad superoperator associated with a jump operator $O$, and $\gamma_j$ is the jump rate for the error $J_j$.

Denote $H_{\text{eff}}(t) := H(t) - \frac{i}{2}\sum_j \gamma_j J_j^\dagger J_j$ as the effective Hamiltonian, and $\mathcal{S} := \sum_j \gamma_j J_j \bullet J_j^\dagger$ as the superoperator describing quantum jumps. The Lindbladian $\mathcal{L}(t)$ can be rewritten as $\mathcal{L}(t) = -i[H_{\text{eff}}(t), \bullet] + \mathcal{S}$. Then, the joint quantum channel, given by the time integral of Eq. (2.55), admits a Dyson expansion with respect to $\mathcal{S}$ [6]:

$$\rho(t) = \mathcal{G}(t, 0)\rho(0) = \sum_{q=0}^{\infty} \mathcal{G}_q(t, 0)\rho(0), \tag{2.56}$$

where $\mathcal{G}_0(t, 0) = \mathcal{W}(t, 0) := W(t, 0) \bullet W^\dagger(t, 0)$, with $W(t, 0) := \mathcal{T} \exp\left[-i \int_0^t H_{\text{eff}}\left(t'\right) dt'\right]$, describes evolution without any quantum jumps, and

$$\mathcal{G}_q(t) = \left[\int_{t_h=0}^{t} dt_h\right]_{h\in[q]} \mathcal{T}\big(\mathcal{W}(t, t_q)\mathcal{S} \cdots \mathcal{S}\mathcal{W}(t_2, t_1)\mathcal{S}\mathcal{W}(t_1, 0)\big), \tag{2.57}$$

where $\mathcal{G}_q$ ($q \geq 1$) describes the evoluation with $q$ quantum jumps. We refer to such an expansion as the jump expansion and and $\mathcal{G}^{[n]} := \sum_{q=0}^{n} \mathcal{G}_q$ as the $n$-th order truncation of $\mathcal{G}$ under the jump expansion.

For quantum error correction, we care about the expansion of the channel $\mathcal{G}$ in terms of the small noise parameter $p := \gamma_i t$ given an evolution time $t$ (here, we have assumed equal noise rates for simplicity):

$$\mathcal{G}(t, 0) = \sum_{q'} p^{q'} \mathcal{G}'_{q'}(t, 0). \tag{2.58}$$

Such an expansion can be obtained by Dyson expanding $\mathcal{G}$ with respect to the full Lindblad superoperators of the noise ($\sum_j \mathcal{D}[\sqrt{\gamma_j}J_j]$) in Eq. (2.55), instead of their quantum-jump components $\mathcal{S}$. We refer to such an expansion of $\mathcal{G}$ as its noise expansion, and $\mathcal{G}'^{[n]} := \sum_{q'=0}^{n} \mathcal{G}'_q$ as the $n$-th order truncation of $\mathcal{G}$ under the noise expansion.

Observe that $\mathcal{G}^{[n]} = \mathcal{G}'^{[n]} + O(p^{n+1})$, i.e. the $n$-th order truncation of a channel $\mathcal{G}$ in terms of its jump expansion or its noise expansion is equivalent up to $n$-order of $p$. Since $\mathcal{G}^{[n]}$ is easier to evaluate for the purpose of this work, we will mostly consider the jump expansion of channels. A nice property of a channel's jump expansion is that it is automatically in a Kraus form:

$$\mathcal{G}(t,0) = \sum_{q=0} \left[\int_{t_h=0}^{t} dt_h\right]_{h\in[q]} \left[\sum_{j_h=1}^{N}\right]_{h\in[q]} G_q(\{t_h\},\{j_h\}) \bullet G_q^\dagger(\{t_h\},\{j_h\}), \tag{2.59}$$

where

$$G_q(\{t_h\},\{j_h\}) := \mathcal{T}W(T,t_q)E_{j_q}\cdots E_{j_2}W(t_2,t_1)E_{j_1}W(t,0). \tag{2.60}$$

One can, therefore, view $G_q(\{t_h\},\{j_h\})$ as a Kraus operator of the channel with discrete indices $q,\{j_h\}$ and continuous indices $\{t_h\}$.

## Ancilla-assisted bosonic operations

As shown in Fig. 2.8(a), we consider gadgets for some bosonic code consisting of a sequence of ancilla-assisted operations (AAOs). For each AAO, a $d_A \geq 2$ ancilla $A$ is initialized in some initial state $|i\rangle_A$, interacts with the bosonic mode $C$ for some time $T$, and is finally measured in some basis $\mathcal{B}_A$. We consider continuous Markovian interactions between $A$ and $C$, which is described by the Lindblad master equation in Eq. (2.55) with a Hamiltonian $H_{AC}(t)$ that acts on the joint system, a set of bosonic jump operators $\{\sqrt{\kappa_i}F_i\}$, and a set of ancilla jump operators $\{\sqrt{\gamma_j}J_j\}$. We allow adaptively choosing the later AAOs using the earlier ancilla measurement outcomes. Note that a direct operation on the bosonic mode

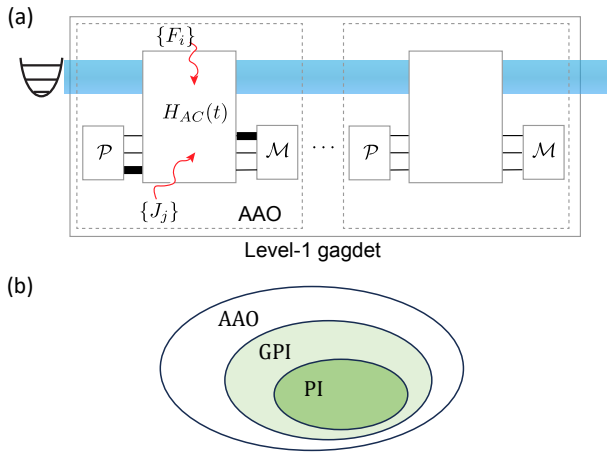


Figure 2.8: (a) Illustration of a level-1 bosonic gadget consisting of a sequence of ancilla-assisted operations. For each AAO, the ancilla is initialized to some state $|i\rangle$ and measured in some basis $\mathcal{B}_A$. The later AAOs can be performed adaptively using the earlier ancilla measurement outcomes. (b) Illustration of the AAO, GPI and PI operations. As a special case of AAO, the GPI operations with bosonic QEC can handle bosonic errors induced by relevant ancilla faults. The previous PI operations [6] can be regarded as a special GPI without bosonic QEC, which are designed to avoid any introduction of bosonic errors due to relevant ancilla faults.

can be viewed as a trivial AAO with the ancilla being initialized in some state $|i\rangle$, idling for the evolution time, and measured in $|i\rangle$ without any errors.

Such an AAOs-composed bosonic gadget forms a channel $\mathcal{N}$ on the bosonic mode, which can be decomposed as $\mathcal{N} = \mathcal{N}_n \circ \mathcal{N}_0$, where $\mathcal{N}_0$ is the target bosonic operation and $\mathcal{N}_n = \sum_k N_k \bullet N_k^\dagger$ is a data noise channel represented by a set of noise Kraus operators $\{N_k\}$. Fault tolerance essentially concerns the relation between the physical channels $\{\mathcal{G}\}$ and the resultant bosonic channel $\mathcal{N}$. More specifically, we need to study how the noise in $\mathcal{G}$, which we refer to as faults, propagates to the data errors $\{N_k\}$ in $\mathcal{N}_n$. We will need to quantify the size of the faults and the data errors and design AAOs such that small faults only propagate to small data errors. We refer to a physical channel $\mathcal{G}$ that contains no more than $t$ faults as its $n$-th order truncation $\mathcal{G}^{[n]}$. To quantify the size of the data bosonic errors, we need to specify a bosonic error-correcting code and an error basis. In this work, we will primarily focus on the cat codes [60, 130, 45] and a basis we termed loss-dephasing basis, which is

closely related to photon loss and bosonic dephasing errors.

### Noise channel

Typical bosonic errors include excitation loss ($a$), heating ($a^\dagger$), and bosonic dephasing ($a^\dagger a$). For such errors, a natural basis to use is $\{e^{-i\theta a^\dagger a}a^k, a^{\dagger k'}e^{i\theta' a^\dagger a}\}_{k,k'\in\mathbb{N};\theta,\theta'\in(-\pi,\pi]}$, which is a complete basis for all single-mode bosonic operators. Neglecting heating errors $a^\dagger$, which are typically small [7, 127], the relevant errors are then spanned by $\{E_k(\theta) := e^{-i\theta a^\dagger a}a^k\}_{k\in\mathcal{N},\theta\in(-\pi,\pi]}$, which we refer to as the loss-dephasing basis.

A four-legged cat code can only correct errors $E_k(\theta)$ with small $k$ and $|\theta|$ (see Sec. 2.3.1). This motivates us to quantify the size of $E_k(\theta)$ as $|E_k(\theta)|_w := (k, |\theta|) \in \mathbb{N} \times [0, \pi]$. We compare the size of two errors by introducing a partial order with respect to the proper cone $R_+^2 := [0, \infty) \times [0, \infty)$, i.e. $|E_{k'}(\theta')|_w \geq |E_k(\theta)|_w \leftrightarrow (k' - k, |\theta'| - |\theta|) \in R_+^2$. We say that a bosonic noise channel $\mathcal{N}_n$ contains at most $(k, \theta)$ errors if all its Kraus operators have size at most $(k, \theta)$, and a state $\left|\phi'\right\rangle$ is at most $(k, \theta)$ far from a target state $|\phi\rangle$ if there exists a $\mathcal{N}_n$ containing at most $(k, \theta)$ errors such that $\left|\phi'\right\rangle$ is in the support of $\mathcal{N}_n(|\phi\rangle\langle\phi|)$. With this quantification of error sizes, for $\alpha \gg 1$, the four-legged cat can correct errors $|E_k(\theta)|_w \leq (1, \pi/4)$ [38]. Fig. 2.9(a) illustrates the 2-dimensional error space indicated by the number of photon loss and dephasing angle.

## *2.3.2 Fault tolerance of the hybrid system*

In this section, we formalize a fault-tolerance framework for the hybrid system with bosonic modes and discrete-variable ancillae in the context of concatenated codes [26]. Since the single-mode cat code alone cannot arbitrarily suppress logical errors, one needs to concatenate it with an outer qubit code for fault-tolerant quantum computing. That is, we will have three levels of gadgets. The level-0 gadgets are the physical operations; The level-1 gadgets are encoded in the cat code and the level-2 gadgets are encoded in the outer qubit

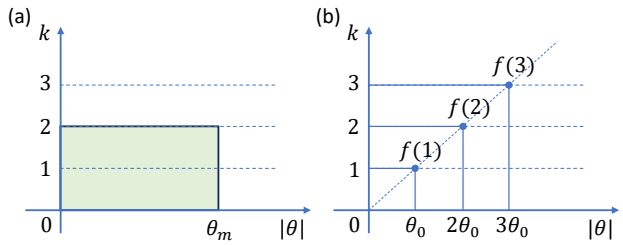


Figure 2.9: Illustration of bosonic error decomposition and the error propagation function $f(m)$. (a) The bosonic loss-dephasing error can be expanded by the basis $E_k(\theta)$. By defining a partial order of the size $E_k(\theta)$, the bosonic error $E_k(\theta)$ with at most $(k, \theta_m)$ error can be illustrated as the green region in the plot. Here $k = 2$. (b) Suppose $m$ faults occur during the gate implementation. To capture the propagation of faults to the final bosonic error, we introduce a function $f(m) = (m, m\theta_0 \bmod \pi)$ as a upper bound of the induced final loss and dephasing errors.

code. A quantum circuit is fault-tolerantly executed using level-2 gadgets, and each level-2 gadget is executed using a level-1 circuit with multiple level-1 gadgets. In order for each level-2 gadget (or equivalent, a level-1 circuit) to be executed with a failure rate $O(p^{t+1})$, which suppresses the physical error rate $p$ to certain order $t + 1$, the level-1 gadgets suffice to satisfy the following properties:

First, there exists a function $f : \mathbb{N} \to \mathbb{N} \times [0, \pi]$ that satisfies:

1. $f(m_1) \leq f(m_2) \leftrightarrow m_1 \leq m_2$ if $m_1, m_2 \leq t$.

2. $f(m_1 + m_2) = f(m_1) + f(m_2)$ if $m_1 + m_2 \leq t$.

Roughly speaking, $f(m)$ constraints the maximal size of data errors that $a$ faults during a protocol can propagate to. For instance, for a bosonic code that can correct phase rotations smaller than $\theta_{\max}$, we will choose $f(m) = (m, m\theta_0 \mod \pi)$ for some $\theta_0 \in [0, \theta_{\max}/t]$, which constraints that $m$ faults can propagate to at most $m$ photon losses and phase rotations of size at most $m\theta_0$. We illustrate such an error propagation constrained by $f$ in Fig. 2.9(b).

Given $f$ and $t$, we then define $t$-FT fault-tolerant gadgets, including gate, error correction, state preparation, and measurement, for the hybrid system by generalizing the definitions in Ref. [26] for qubits. We remark that, the following FT definitions are related to the choice

of the function $f$.

**Definition 1** ($t$-FT gate)**.** *A gate is $t$-FT if it satisfies: For an input codeword that has an error of size $(k, \theta)$, if at most $m$ faults occur during the gate with $(k, \theta) + f(m) \leq f(t)$, the output state is at most $(k, \theta) + f(m)$ far from the codespace; Furthermore, ideally decoding the output state gives the same codeword as first ideally decoding the input state and then applying the ideal gate.*

Note that this condition for the gate corresponds to the combination of Property 2 and Property 3 of Ref. [26].

**Definition 2** ($t$-FT QEC)**.** *A QEC gadget is $t$-FT if it satisfies:*

1. *For an input codeword with an error of size $(k, \theta)$, if at most $m$ faults occur during the protocol with $(k, \theta) + f(m) \leq f(t)$, ideally decoding the output state gives the same codeword as ideally decoding the input state.*

2. *For at most $m$ faults during the protocol with $f(m) \leq f(t)$, no matter the size of the error on the input state, the output state is at most $f(m)$-far from the code space.*

Note that conditions (i) and (ii) correspond to Properties 1 and 0 of Ref. [26], respectively.

State preparation and measurement are special cases of FT gates:

**Definition 3** ($t$-FT state preparation)**.** *A state-preparation gadget is 1-FT if it satisfies: If at most $m \leq t$ faults occur during the protocol, the output state is at most $f(m)$-far from the target state; Furthermore, ideally decoding the output state gives the ideal target state.*

**Definition 4** ($t$-FT measurement)**.** *A measurement gadget is $t$-FT if it satisfies: For an input codeword that has an error of size $(k, \theta)$, if at most $m$ faults occur during the gate with $(k, \theta) + f(m) \leq f(t)$, the measurement is equivalent to applying the ideal measurement to the input codeword.*

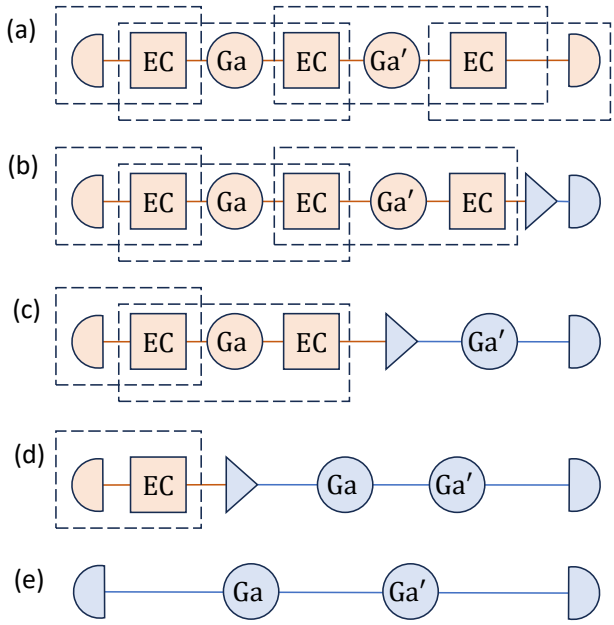


Figure 2.10: Reduction of a FT level-1 circuit to the ideal circuit.

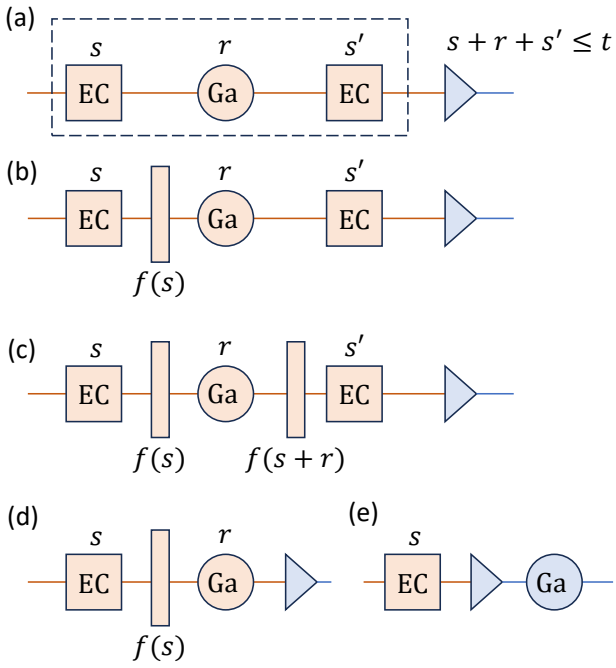


Figure 2.11: Reduction of the extended rectangular to an ideal gadget.

Based on the definition of the $t$-FT gadgets, we have the following proposition.

**Proposition 5.** *Using $t$-FT level-1 gadgets, any level-1 circuit has a failure rate $O(p^{t+1})$, where $p \in [0, 1)$ is the physical error rate, i.e., the probability that one fault happens in the gadget.*

*Proof.* We follow the extended-rectangle formalism in Ref. [26]. Without loss of generality, we consider an ideal quantum circuit in Fig. 2.10(e). Here we take the single-qubit level-1 circuit as an example. In practice, we realize this circuit using the noisy $t$-FT level-1 gadgets shown in Fig. 2.10(a). To analyze the fault-tolerance property of this circuit, we draw several dashed boxes to cover the whole circuit. The dashed boxes are called extended rectangles exRec. For a quantum gate, an extended rectangle exRec (a dashed box in Fig. 2.10(a)) is a composition of a front EC gadget, a gate and a rear EC gadget, i.e. $\text{exRec} = \text{EC} \circ \text{Ga} \circ \text{EC}$.

We say that any operation Op is $t$-good if it contains no more than $t$ faults. In what follows, we show that if all the dashed boxes in Fig. 2.10(a) are $t$-good, we can reduce the noisy circuit to the ideal circuit following the stream in Fig. 2.10. To this end, we introduce the ideal decoder ID (the blue triangles in Fig. 2.10 and 2.11), which performs an ideal recovery given a bosonic code. We also introduce a $(k, \theta)$-filter $[(k, \theta)]\text{F}$ (the orange thin rectangles in Fig. 2.11) which performs a projection onto the space spanned by all states that can be obtained by acting on a codeword with an error no larger than $(k, \theta)$.

First of all, we notice that if the last box in Fig. 2.10(a) is $t$-good, then based on the definition of $t$-FT QEC and measurement, we can equivalently convert the circuit in Fig. 2.10(a) to (b).

Then, we follow the procedures in Fig. 2.11 to reduce the extended gadgets of quantum gates to the ideal gadgets: Denote the faults that occur in the front EC gadget, the gate gadget and the rear EC gadget to be $s, r, s'$, respectively. Since the dashed box is $t$-good, we have $s + r + s' \leq t$. Fig. 2.11(a) and (b) are equivalent due to the second requirement of FT QEC in Def. 2; (b) and (c) are equivalent due to the first requirement of the FT gate

in Def. 1; (c) and (d) are equivalent due to the first requirement of FT QEC in Def. 2; (d) and (e) are equivalent due to the second requirement of the FT gate in Def. 1. Then we can transform the circuit from Fig. 2.10(b) to (d) using the reduction in Fig. 2.11.

Finally, we use the property of FT state preparation to reduce Fig. 2.10(d) to (e). The argument is similar to the ones for the extended gadgets of quantum gates in Fig. 2.11.

We assume that errors represented by quantum jumps occur independently in the level-1 gadgets. Consider a level-1 circuit composed of many $t$-FT level-1 gadgets that can be grouped into extended rectangles (see e.g., Fig. 2.10(a)). If there are at most $t$ quantum errors in each extended rectangle, we can convert it to an ideal gadget. In that case, only when more than $t$ errors occur in the same extended rectangles at the same time can one logical error happen, which owns a probability of $O(p^{t+1})$. □

In the following, we focus on constructing FT level-1 bosonic gadgets that satisfy the above definitions by integrating bosonic quantum error correction and quantum controls. More specifically, given a bosonic code $\mathcal{C}$ that can correct loss and phase-rotation errors, e.g. the cat code, we try to design error-corrected $\mathcal{C}$-encoded gadgets by carefully engineering the Hamiltonian of their composing AAOs so that physical faults propagate controllably to data errors. An analogous example in the context of qubit fault-tolerance is the use of transversal gates [134], which guarantees that a single circuit fault only propagates to a single data error per code block. However, this quantum control task is more sophisticated when involving bosonic modes as we need to consider complicated continuous evolution in a larger Hilbert space.

In order for a level-1 gadget to be FT, it has to tolerate both bosonic faults and ancilla faults. Tolerance against bosonic faults can be achieved relatively easily by using the error-transparency control [135], or more generally, the error-closure control [136]. Tolerance against ancilla errors is usually harder to achieve since some DV ancilla errors tend to propagate uncontrollably and a small ancilla fault could lead to a catastrophic error in

the bosonic mode. Fortunately, path-independent control [6, 137, 7] has been proposed for controlling the ancilla faults propagation to the bosonic mode. However, the previously defined PI condition [6] is more stringent than what is required by fault tolerance. Thus in the next section, we will generalize the PI control, relax its requirement, and rigorously connect its generalized version to fault-tolerance analyzed in this section.

### *2.3.3 Generalized path-independent operations*

We first review the PI control proposed in Ref. [6]. Again, we consider a Markovian interaction between a bosonic mode $C$ and a $d \geq 2$-level ancilla $A$ described by Eq. (2.55)todo: re, where only the ancilla noises are considered, i.e.

$$\frac{d\rho}{dt} = -i[H_{AC}(t), \rho] + \sum_j \mathcal{D}[\sqrt{\gamma_j} J_j]\rho \tag{2.61}$$

where $J_j$ are some jump operators acting only on the ancillary system. The ancilla is initialized in some initial state $|i\rangle_A$, and measured in a certain basis $\{|r\rangle_A\}$ after an interaction time $T$. Let $\mathcal{G}(T)$ denote the joint channel generated by the Lindblad master equation in Eq. (2.61) for a time $T$. With a slight abuse of notation, we may neglect the subscripts $A$ or $C$ labeling the ancilla or the bosonic mode for states or operators without ambiguity. We denote $\langle\langle r|\mathcal{G}|i\rangle\rangle := \langle r| \mathcal{G}(|i\rangle \langle i| \otimes \bullet) |r\rangle$ as the (unnormalized) channel on the bosonic mode conditioned on the initial and final ancilla state $|i\rangle$ and $|r\rangle$ [137]. A PI gate is defined as follows.

**Definition 6** (PI gate)**.** *An ancilla-assisted gate $\mathcal{G}(T)$ is PI in an ancilla basis $\mathcal{B}_A$ with an ancilla initial state $|i\rangle$ if for any $|r\rangle \in \mathcal{B}_A$,*

$$\langle\langle r|\mathcal{G}(T)|i\rangle\rangle \propto U_{ri} \bullet U_{ri}^\dagger, \tag{2.62}$$

*where $U_{ri}$ is some $r$-dependent unitary on the bosonic mode.*

The PI condition in Eq. (2.62) implies that each conditional channel does not contain any errors (it is a unitary channel without information loss) propagated from the ancilla, although the unconditional channel might. In other words, no information is lost to the environment by performing such a noisy operation if the ancilla measurement outcome is accessible. See Fig. 2.12 for an illustration of such a PI gate.

Note that the PI condition in Eq. (2.62) is for the joint channel, which could be hard to evaluate directly. As such, Ref. [137] provided an easy-to-evaluate algebraic condition for the Hamiltonian $H_{AC}(t)$ and the jump operators $\{J_j\}$ in order for $\mathcal{G}$ to satisfy Eq. (2.62), which we present in Appendix A of Ref. [138].

The PI definition in Def. 6 considers an infinite number of ancilla-faults since it examines the full $\mathcal{G}(T)$. In practice, when the ancilla noises are small, by correcting only a finite number of ancillary faults, we can suppress the real logical error rates to a higher order. As such, we define the following finite-order PI gate that only concerns a finite truncation of $\mathcal{G}(T)$:

**Definition 7** (Finite-order PI gate)**.** *An ancilla-assisted gate is $n$-PI in an ancilla basis $\mathcal{B}_A$ with an ancilla initial state $|i\rangle$ if for any $|r\rangle \in \mathcal{B}_A$ and any $k \leq n$,*

$$\langle\langle r|\mathcal{G}^{[k]}(T)|i\rangle\rangle \propto U_{ri} \bullet U_{ri}^{\dagger}, \tag{2.63}$$

*where $U_{ri}$ is some $r$-dependent unitary on the bosonic mode.*

In Appendix A of Ref. [138], we present an algebraic condition for the Hamiltonian and jump operators in order for $\mathcal{G}$ to satisfy Eq. (2.63).

The PI condition, even with a finite order, is demanding since it requires the conditional channels to be exactly unitary channels and thus allows no error propagation at all. However, observe that if the bosonic mode is protected by some bosonic codes, fault-tolerance can still be achieved even if we allow error propagations, as long as the propagated errors are small

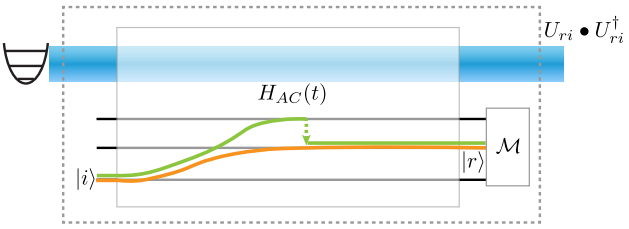


Figure 2.12: Illustration of a PI gate. Given an ancilla initial state $|i\rangle$ and a measurement basis $\mathcal{B}_A$, the bosonic mode undergoes a $r$-dependent unitary $U_{ri}$ for any ancilla measurement outcome $|r\rangle \in \mathcal{B}_A$, independent of the different system paths (see e.g. the green and orange curves, where an ancilla relaxation happens for the green curve).

and correctable. Based on this observation, we generalize the PI control and present a less stringent condition that, nevertheless, still complies with the idea of fault-tolerance:

**Definition 8** (GPI operation)**.** *Given a single-mode bosonic code with a codespace projection $P_c$, we say that an ancilla-assisted operation is $n$-th order generalized path-independent (GPI) in an ancilla basis $\mathcal{B}_A$ with an initial ancilla state $|i\rangle$ if for any $|r\rangle \in \mathcal{B}_A$ and $k \leq n$,*

$$\langle\langle r|\mathcal{G}^{[k]}(T)|i\rangle\rangle \propto (\sum_s K_{ri}^s \bullet K_{ri}^{s\dagger}), \tag{2.64}$$

*where $\{K_{ri}^s\}_s$ satisfies the KL condition, i.e. $P_c K_{ri}^{s\dagger} K_{ri}^{s'} P_c \propto P_c$ for any $K_{ri}^s, K_{ri}^{s'} \in \{K_{ri}^s\}_s$.*

Note that any conditional channel $\langle\langle r|\mathcal{G}^{[k]}(T)|i\rangle\rangle$ can be written in the form of Eq. (2.64), with a set of $(r,i)$-dependent Kraus operators $\{K_{ri}^s\}_s$. The condition that $\{K_{ri}^s\}_s$ satisfies the KL condition implies that the conditional channel $\langle\langle r|\mathcal{G}^{[k]}(T)|i\rangle\rangle$ contains only correctable errors.

The GPI condition generalizes from the PI condition in Def. 7 from the following two aspects. First, the GPI condition considers any operation (any CPTP map) to the bosonic mode as a target, while the PI condition only considers unitary gates. Second, the GPI condition allows finite propagation of ancilla faults to the bosonic mode for each conditional channel, as long as the propagated errors are correctable by the bosonic code. See Fig. 2.8(b)

for an illustration of the relation between ancilla-assisted operations, GPI operations and PI operations.

In Appendix A of Ref. [138], we present an algebraic condition for GPI operations again by only examining the Hamiltonian and jump operators. Note that we directly present the GPI condition in the finite-order form, which is of practical relevance.

## GPI examples

Here, we provide two examples of GPI operations for the four-legged cat code.

**GPI SNAP gate with a three-level $\chi$-mismatched ancilla**

As an example, we consider the photon-number selective phase (SNAP) gate [131] in circuit-QED systems. In the rotating frame, a three-level ancilla with levels $\{|g\rangle, |e\rangle, |f\rangle\}$ is dispersively coupled to a bosonic mode via the Hamiltonian

$$H_0 = -(\chi_f |f\rangle\langle f| + \chi_e |e\rangle\langle e|) \otimes a^\dagger a, \tag{2.65}$$

and the ancilla is frequency-selectively driven between $|g\rangle$ and $|f\rangle$ states:

$$H_c(t) = \Omega \sum_{n=0}^{N} e^{-i(n\chi_f t - \phi_n)} |f\rangle\langle g| + h.c., \tag{2.66}$$

where $\vec{\phi} := (\phi_0, \phi_1, \cdots, \phi_N)$ is some phase vector that we can choose. We consider ancilla jump operators $\{J_1 = \sqrt{\gamma}|e\rangle\langle f|, J_2 = \sqrt{\gamma}|g\rangle\langle e|, J_3 = \sqrt{\gamma}\sum_{s\in\{e,f\}} \Delta_s |s\rangle\langle s|\}$, where $J_1$ describes the ancilla decays from $|f\rangle$ to $|e\rangle$, $J_2$ the ancilla decay from $|e\rangle$ to $|g\rangle$, and $J_3$ an ancilla dephasing with arbitrary phases $\Delta_e, \Delta_f \in \mathbb{C}$. We will use this error model throughout the paper whenever using a three-level ancilla.

In the interaction picture associated with $H_0$, the system Hamiltonian reads

$$\tilde{H} = \Omega \left[ |f\rangle\langle g| \otimes S(\vec{\phi}) + h.c. \right], \tag{2.67}$$

where $S(\vec{\phi}) := \sum_{n=1}^{N} e^{i\phi_n} |n\rangle \langle n|$ applies a number dependent phase shift $\vec{\phi}$ to the bosonic mode. Note that we have performed the rotating wave approximation by assuming $\Omega \ll \chi_f$. Denote $\Delta\chi := \chi_f - \chi_e$ as the $\chi$ mismatch. The ancilla jump operators transform to $\tilde{J}_1(t) = \sqrt{\gamma}|e\rangle\langle f| \otimes e^{-i\Delta\chi t a^\dagger a}$, $\tilde{J}_2(t) = \sqrt{\gamma}|g\rangle\langle e| \otimes e^{-i\chi_e t a^\dagger a}$, and $\tilde{J}_3 = J_3$.

We initialize the ancilla in $|g\rangle$ and let the system evolve for a time $T = \pi/2\Omega$, and measure the ancilla in the $\{|g\rangle, |e\rangle, |f\rangle\}$ basis. In the absence of errors, the ancilla will end up in $|f\rangle$ while the central system undergoes the target number-dependent phase shifts $S(\vec{\phi})$, i.e. $\langle f|e^{-i\tilde{H}_c T}|g\rangle = S(\vec{\phi})$. With ancilla errors, we can explicitly write down the truncated conditional channels (in the interaction picture) up to the first order:

$$\begin{aligned}
&\langle\langle g|\tilde{\mathcal{G}}^{[1]}(T)|g\rangle\rangle \propto \mathcal{I}, \\
&\langle\langle f|\tilde{\mathcal{G}}^{[1]}(T)|g\rangle\rangle \propto S(\vec{\phi}) \bullet S^\dagger(\vec{\phi}), \\
&\langle\langle e|\tilde{\mathcal{G}}^{[1]}(T)|g\rangle\rangle \propto \int_{t=0}^{T} dt\ e^{-i\Delta\chi t a^\dagger a} S(\vec{\phi}) \bullet S^\dagger(\vec{\phi}) e^{i\Delta\chi t a^\dagger a},
\end{aligned} \tag{2.68}$$

If there is no $\chi$-mismatch, i.e. $\Delta\chi = 0$, this gate is strictly a 1-PI gate (see Eq. (2.63)); If there is a finite $\chi$-mismatch, the gate is no longer PI. Nevertheless, for a bosonic code that can correct phase rotations in the range $[-\theta_m/2, \theta_m/2]$ (e.g. $\theta_m = \pi/2$ for the four-legged cat), the gate is still a 1-GPI gate if $\Delta\chi T \leq \theta_m$ (see Eq. (2.64)).

In Appendix A of Ref. [138], we show that one can verify the GPI property of this SNAP gate more easily without calculating the conditional channels by checking a set of algebraic conditions for the Hamiltonian and jump operators. Also, in Appendix C of Ref. [138], we present another GPI SNAP scheme using a qutrit and an extra flag qubit, which can tolerate even larger $\chi$-mismatch $\Delta\chi T$.

**GPI parity measurement with a three-level $\chi$-mismatched ancilla**

As another example of GPI operations, we consider the parity measurement for correcting photon loss errors [75] using a three-level $\chi$-mismatched ancilla.

The system Hamiltonian (in the rotating frame) is $H_0 = -(\chi_e|e\rangle\langle e| + \chi_f|f\rangle\langle f|) \otimes a^\dagger a$

(without ancilla drives). We denote $|\pm\rangle$ as $(|g\rangle \pm |f\rangle)/\sqrt{2}$. The ancilla is initialized in $|+\rangle$ and measured in the basis $\{|+\rangle, |-\rangle, |e\rangle\}$.

In the absence of ancilla errors, the operation performs a projection onto the even (odd) subspace of the bosonic mode conditioned on the ancilla measured in $|+\rangle$ ($|-\rangle$):

$$\begin{aligned} \langle\langle +|\mathcal{G}^{[0]}|+\rangle\rangle &= P_+ \bullet P_+, \\ \langle\langle -|\mathcal{G}^{[0]}|+\rangle\rangle &= P_- \bullet P_-, \end{aligned} \tag{2.69}$$

where $P_\pm := (I \pm e^{-i\pi a^\dagger a})/2$ is the projection on the even/odd parity subspace of the bosonic mode.

In the presence of ancilla errors $\{J_1 = \sqrt{\gamma}\,|e\rangle\langle f|$, $J_2 = \sqrt{\gamma}\,|g\rangle\langle e|$, $J_3 = \sqrt{\gamma}\sum_{s\in\{e,f\}} \Delta_s\,|s\rangle\langle s|\}$, we move to the interaction picture associated with $H_0$. Now, the system Hamiltonian is 0 and the ancilla jump operators read $\tilde{J}_1(t) = \sqrt{\gamma}|e\rangle\langle f| \otimes e^{-i\Delta\chi t a^\dagger a}$, $\tilde{J}_2(t) = \sqrt{\gamma}|g\rangle\langle e| \otimes e^{-i\chi_e t a^\dagger a}$, and $\tilde{J}_3 = J_3 = \sqrt{\gamma}\sum_{s\in\{e,f\}} \Delta_s\,|s\rangle\langle s|\}$, same to those in the previous SNAP example. Here, without loss of generality, we set $\Delta_f = -1$.

We can calculate the noise expansion of the joint channel up to the first order (see Eq. (2.59)):

$$\tilde{\mathcal{G}}^{[1]}(T) = W(T,0) \bullet W^\dagger(T,0) + \gamma \int_{t=0}^{T} G_1(t,1) \bullet G_1^\dagger(t,1) + \gamma \int_{t=0}^{T} G_1(t,3) \bullet G_1^\dagger(t,3), \tag{2.70}$$

where $W(t_2,t_1) := \exp\left[-iH_{\text{eff}}(t_2-t_1)\right]$ with $H_{\text{eff}} := -\frac{i}{2}\sum_{j=1}^{3} \tilde{J}_j^\dagger \tilde{J}_j = -\frac{i}{2}\gamma[(1+|\Delta_e|^2)\,|e\rangle\langle e| + 2\,|f\rangle\langle f|]$, and $G_1(t,j) := W(T,t)\tilde{J}_j(t)W(t,0)$. Note that we have dropped the term associated with the first-order quantum jump with $\tilde{J}_2$, which is zero when the ancilla starts from $|+\rangle$. Going back to the lab frame, the truncated channel is $\mathcal{G}^{[1]}(T) = \tilde{\mathcal{G}}^{[1]}(T) \circ [U_0(T) \bullet U_0^\dagger(T)]$,

where $U_0(T) := e^{-iH_0T}$. Then we can calculate the truncated conditional channels:

$$\begin{aligned}
\langle\langle +|\mathcal{G}^{[1]}|+\rangle\rangle &= [(1-\frac{p}{2})P_+ + \frac{p}{2}P_-] \bullet [(1-\frac{p}{2})P_+ + \frac{p}{2}P_-] + pP_- \bullet P_- + O(p^2) \\
\langle\langle -|\mathcal{G}^{[1]}|+\rangle\rangle &= [(1-\frac{p}{2})P_- + \frac{p}{2}P_+] \bullet [(1-\frac{p}{2})P_- + \frac{p}{2}P_+] + pP_+ \bullet P_+ + O(p^2) \\
\langle\langle e|\mathcal{G}^{[1]}|+\rangle\rangle &= \frac{p}{2T}\int_{t=0}^{T} dt e^{-i(\Delta\chi t+\pi)a^\dagger a} \bullet e^{i(\Delta\chi t+\pi)a^\dagger a} + O(p^2),
\end{aligned} \tag{2.71}$$

where $p := \gamma T$. For a four-legged cat with $\alpha \gg 1$, Eq. (2.71) satisfies the GPI condition as long as $\Delta\chi T < \pi/2$. Note that the first two terms in Eq. (2.71) imply that one might obtain wrong parity measurement outcomes with a probability $O(p)$ if the ancilla is measured in $|\pm\rangle$. Such effective measurement errors can be suppressed to the second order by repeating the above parity measurement three times and performing majority voting, which will be discussed in the next section when we rigorously consider fault tolerance.

## Connection between GPI and fault-tolerance

Here, we establish the connection between GPI quantum control and the hybrid fault-tolerance defined in Sec. 2.3.2. Let the bosonic mode be encoded in some bosonic code with a code projection $P_c$.

**Proposition 9.** *Each AAO contained in a $t$-FT level-1 gadget with an ancilla initial state $|i\rangle$ and an ancilla measurement basis $\mathcal{B}_A$ has to be $t$-GPI with respect to $|i\rangle$, $\mathcal{B}_A$, and the code projection $P_c$.*

*Proof.* Any $t$-FT gadget requires that if any $m \leq t$ faults occur during the protocol, the output is guaranteed to be correct. However, if one AAO is not $t$-GPI, there exists an ancilla measurement outcome $r$, conditioned on which the bosonic channel (see Eq. (2.64)) contains non-correctable errors. As a result, the final output can no longer be guaranteed to be correct. □

Conversely, we can combine pieces of $t$-GPI operations to get a $t' \leq t$-FT gadgets, as shown in Fig. 2.8. In order to make $t' = t$, there are extra requirements for the AAOs, which are typically easier to satisfy than GPI. Instėad of making rigorous statements about these requirements for generic gadgets, we will make case studies when constructing concrete FT gadgets. Nevertheless, we comment on some high-level ideas used to design the extra ingredients that can be combined with GPI to achieve fault tolerance here: (i) Operations are error transparent/closure against bosonic errors (see Appendix B of Ref. [138]); (ii) The propagation from ancilla faults to bosonic errors is linear; (iii) There exists at least one ancilla state $|r\rangle$ such that the ideal conditional bosonic channel $\langle\langle r|\mathcal{G}^{[0]}|i\rangle\rangle$ gives the target operation.

As the first example, we construct 1-FT Z-axis rotation $Z(\theta)$ for the four-legged cat using the GPI SNAP gate presented in Sec. 2.3.3. To implement a $Z(\theta)$ gate, we choose a GPI SNAP gate with $\Delta\chi T < \pi/2$ and

$$S(\vec{\theta}) = P_0 + P_3 + e^{i\theta}(P_2 + P_1), \tag{2.72}$$

where $P_j := \sum_{i=0} |4i + j\rangle \langle 4i + j|$. We consider the same ancilla jump errors as those presented in Sec. 2.3.3. In addition, we consider a single jump operator $a$ representing a single photon loss for the bosonic mode. Then we implement the 1-FT $Z(\theta)$ gate based on Algorithm 1 below. The 3-level ancilla basis is denoted by $|g\rangle$,$|e\rangle$ and $|f\rangle$ according to Eq. (2.65).

Now, we verify that the above protocol satisfies the definition of a 1-FT gate in Def. 1. Here, we choose $f(m) = (m, m\Delta\chi T/2)$ with $\Delta\chi T/2 < \pi/4$. Suppose the input error is of size $(k, \theta_0)$ and there are $m$ faults during the protocol. There are two cases where $(k, \theta_0) + f(m) \leq f(1)$. First, $m = 0$ and $(k, \theta_0) \leq (1, \Delta\chi T/2)$. Obviously, the gate is error-transparent to the phase rotation $e^{-i\theta a^\dagger a}$, i.e. it simply commutes through the gate and remains at the output, since it commutes with the system Hamiltonian (see Eq. (2.65) and (2.66)). Moreover, as

**Algorithm 1:** 1-FT $Z(\theta)$ gate

```
o ← e. // o records the ancilla measurement outcome
while o ≠ f do
    Prepare the anilla in the |g⟩ state, apply the GPI SNAP gate with S(θ⃗) in
     Eq. (2.72) for a time T = π/2Ω, and measure the ancilla in the |g⟩, |e⟩, |f⟩ basis
     with an outcome o ∈ {g, e, f}.
    if o = e then
        Apply a phase rotation e^{iΔχTa†a/2} to the bosonic mode.
```

we shown in Appendix B of Ref. [138], the gate is also error-transparent to a single photon loss $a$ when using the form of $S(\vec{\phi})$ in Eq. (2.72). Therefore, the input $(k \leq 1, \theta \leq \Delta\chi T/2)$ error simply remains at the output and stays correctable. Second, $m = 1$ and $(k, \theta) = (0, 0)$. In this case, either an ancilla dephasing, or an ancilla decay, or a single photon loss occurs during the protocol. A single ancilla dephasing might cause the ancilla ending in $|g\rangle$ instead of $|f\rangle$ but does not add any error to the bosonic mode; A single ancilla decay from $|f\rangle$ to $|e\rangle$ only causes a correctable phase rotation with an angle $|\delta\theta| \leq \Delta\chi T/2 < \pi/4$ [1]; A single-photon loss simply remains at the output and stays correctable.

As the second example, we construct a 1-FT QEC protocol for correcting a single-photon loss. Note that we will present a full EC gadget correcting both photon loss and dephasing errors in the next section. The protocol utilizes the 1-GPI parity measurement presented in Sec. 2.3.3, with a $\chi$ mismatch $\Delta\chi T < \pi/2$.

Now, we verify that the protocol in Alg. 2 satisfies the definition of a 1-FT QEC in Def. 2. Similar to the previous $Z(\theta)$ gate example, we choose $f(m) = (m, m\Delta\chi T/2)$. Assume there is an input error of size $(k, 0)$ and $m$ faults occur during the protocol. Note that since we are only correcting single photon losses for now, we assume the input has no dephasing errors. To verify condition (i) in Def. 2, we consider either $k = 1$, $m = 0$ or $k = 0$, $m = 1$. In the earlier case, the single photon loss can be perfectly corrected and the output has no error; in

1. An ancilla decay from $|f\rangle$ to $|e\rangle$ induces phase rotations in the range of $[0, \Delta\chi T]$ (see Eq. (2.68)). The feedback rotation by $-\Delta\chi T/2$ then shifts the phase rotation window to $[-\Delta\chi T/2, \Delta\chi T/2]$

**Algorithm 2:** 1-FT photon-loss correction

$o_i \leftarrow e$ for $i \in \{1,2,3\}$. // $\{o_i\}_{i\in[3]}$ `record three consecutive parity measurement outcomes`

**for** $i \leftarrow 1$ **to** 3 **do**

  **while** $o_i = e$ **do**

    Prepare the anilla in the $|+\rangle$ state, apply the dispersive coupling for a time $T = \pi/\chi_f$, and measure the ancilla in the $= \{|+\rangle, |-\rangle, |e\rangle\}$ basis with an measurement outcome $o_i$.

    **if** $o_i = e$ **then**

      Apply a phase rotation $e^{i\Delta\chi T a^\dagger a/2}$ to the bosonic mode.

Apply a parity correction based on the majority voting over $\{o_i\}_{i\in[3]}$.

the latter case, we consider either an ancilla dephasing, an ancilla decay, or a single photon loss. An ancilla dephasing may flip a single parity measurement outcome but does not affect the final majority voting; A single ancilla decay only causes a correctable phase rotation with amplitude $\leq \Delta\chi T/2 < \pi/4$, which is a correctable error; A single photon loss during the protocol either gets corrected or goes undetected but remains as a correctable single photon loss at the output.

For condition (ii) in Definition 2, one simply observes that a single photon loss error at the input can be detected and then corrected (although a logical error may happen when combined with another photon loss during the protocol), while a single photon loss or an ancilla decay can cause at most a $f(1) = (1, \Delta\chi T/2)$ error that can go undetected and remains at the output.

### *2.3.4 Fault-tolerant operations of four-legged cat code*

In this section, we focus on the four-legged cat and construct universal 1-FT level-1 gadgets that can correct a single-photon loss and any single ancilla fault, using GPI operations.

The universal operation set we consider is

$$\mathcal{S} = \{\mathrm{EC}, Z(\theta), X(\phi), XX(\delta), \mathcal{P}_{|+_L\rangle}, \mathcal{M}_Z, \mathcal{M}_X\}, \tag{2.73}$$

including error correction, $Z$-axis rotation, $X$-axis rotation, $XX$ rotation ($\exp(-i\delta XX/2)$), state preparation in the $X$ basis, measurement in the $Z$ basis, and measurement in the $X$ basis.

One essential element for our construction is the GPI SNAP gate and GPI parity measurement described in Sec. 2.3.3, respectively. Recall that both of these two operations use a three-level ancilla, which is dispersive coupled to the bosonic mode via $-(\chi_e |e\rangle\langle e| + \chi_f |f\rangle\langle f|) \otimes a^\dagger a$, potentially with a $\chi$ mismatch $\Delta\chi := \chi_f - \chi_e$. Denote the gate time for the SNAP gates as $T$ and that for a parity measurement as $T_P$. Typically $T \gg T_P$ since the driving strength $\Omega$ for the SNAP gate (see Eq. (2.66)) is much smaller than $\chi_f$ in order for the rotating-wave-approximation to hold [7]. We choose $f(m) = (m, m\Delta\chi T/2)$ with $\Delta\chi T/2 < \pi/8$ [2] for proving the fault-tolerance of the gadgets. Unless specially noted, all the SNAP gates and parity measurement gadgets we use have a $\chi$ mismatch $\Delta\chi$.

Similar to the previous sections, we consider $\{a, |e\rangle\langle f|, |g\rangle\langle e|, \sum_{s\in\{e,f\}} \Delta_s |s\rangle\langle s|\}$ as the errors, representing a single photon loss, an ancilla decay from $|f\rangle$ to $|e\rangle$, an ancilla decay from $|e\rangle$ to $|g\rangle$, and an ancilla dephasing, respectively.

**Z-axis rotation**

A 1-FT $Z$-axis rotation with an arbitrary angle ($\theta$) using GPI SNAP gate is presented in Alg. 1 in the previous section. Note that a 1-FT logical gate using strictly PI SNAP gate (with no $\chi$ mismatch) has been experimentally demonstrated for a small binomial bosonic code [7]. Here, the main difference is that our protocol allows a finite $\chi$ mismatch.

2. Here, we choose $\Delta\chi T/2 < \pi/8$ instead of $\pi/4$ so that the relative phase shift between two data bosonic modes is smaller than $\pi/4$, which will be used to prove the fault tolerance of the two-qubit logical gate ($XX$ rotation)

**X-axis rotation**

In the large $\alpha$ limit, a $X$-axis rotation is given by

$$X(\phi) \approx e^{i\phi} \left|C_\alpha^+\right\rangle \left\langle C_\alpha^+\right| + \left|C_{i\alpha}^+\right\rangle \left\langle C_{i\alpha}^+\right|, \tag{2.74}$$

where $\left|C_\beta^\pm\right\rangle := c_\beta^\pm(|\beta\rangle \pm |-\beta\rangle)$, with $c_\beta^\pm$ being normalization constants. We implement $X(\phi)$ by adding a phase $\phi$ to the subspace spanned by the two coherent states $|\alpha\rangle$ and $|-\alpha\rangle$. As illustrated in Fig. 2.13(a), we first displace the cavity by $\alpha$ and apply a phase shift to the vacuum $S(\vec{\phi}) = e^{i\phi} |0\rangle \langle 0| + I - |0\rangle \langle 0|$ using the SNAP gate (see Sec. 2.3.3). Then we displace the cavity by $-2\alpha$ and apply another $S$. Finally, we displace the cavity by $\alpha$ back to the codespace. The joint operation is:

$$\begin{aligned} U_X &= D(\alpha)S(\vec{\phi})D(-2\alpha)S(\vec{\phi})D(\alpha) \\ &= [D(\alpha)S(\vec{\phi})D(\alpha)^\dagger][D(-\alpha)S(\vec{\phi})D(\alpha)^\dagger] \\ &\approx e^{i\theta}P_{\pm\alpha} + I - P_{\pm\alpha}, \end{aligned} \tag{2.75}$$

where $P_{\pm\alpha} := |\alpha\rangle \langle\alpha| + |-\alpha\rangle \langle -\alpha| = \left|C_\alpha^+\right\rangle \left\langle C_\alpha^+\right| + \left|C_\alpha^-\right\rangle \left\langle C_\alpha^-\right|$.

We now show that this gate is 1-FT if the $\chi$-mismatch during the SNAP gates is zero. Assume there is a $(k, \delta\theta)$ input error and $m$ faults occur during the gate. Again, for 1-FT gate (see Def. 1), we only need to consider either $(k = 0, \delta\theta = 0)$, $m = 1$, or $(k \leq 1, \delta\theta \leq \Delta\chi T/2)$, $m = 0$.

First, we consider a single fault occurring during $U_X$. A single-photon loss simply commutes through the entire gate since the two SNAP gates $S$ are error-transparent (see Appendix B of Ref. [138]) and $D(\alpha)$ commutes with $a$ up to a constant. A single-ancilla decay or dephasing during the $S$ gate does not cause any error to the bosonic mode when assuming perfect $\chi$ matching. Therefore, a single fault during the gate causes at most a $(1, 0) < f(1)$-error at the output, which is correctable.

Second, we consider a ($k \leq 1, \delta\theta \leq \Delta\chi T/2$) input error $e^{i\delta\theta a^\dagger a}a^k$. We first notice that $U_X e^{i\delta\theta a^\dagger a}a^k P_c \propto a^k U_X e^{i\delta\theta a^\dagger a}P_c$ since $U_X$ is error-transparent for $a^k$ (see Eq. (2.75)). Here, $P_c := |+_L\rangle\langle+_L| + |-_L\rangle\langle-_L| \approx |C_\alpha^+\rangle\langle C_\alpha^+| + |C_{i\alpha}^+\rangle\langle C_{i\alpha}^+|$ is the projector onto the code space of the four-legged cat. Then we only need to make sure that $U_X$ is also error-transparent to dephasing $e^{i\delta\theta a^\dagger a}$. Let $E := U_X e^{i\delta\theta a^\dagger a}U_X^\dagger$ be the effective error that $e^{i\delta\theta a^\dagger a}$ propagates to after the gate. $E$ satisfies

$$EP_c = e^{i\delta\theta a^\dagger a}P_c + (1-e^{-i\phi})(P_{\pm\alpha} - I)e^{i\delta\theta a^\dagger a}|C_\alpha^+\rangle\langle C_\alpha^+| + (e^{i\phi}-1)P_{\pm\alpha}e^{i\delta\theta a^\dagger a}|C_{i\alpha}^+\rangle\langle C_{i\alpha}^+|, \tag{2.76}$$

where we can see that $U_X$ is not error-transparent against the dephasing due to the last two terms of Eq. (2.76). Fortunately, we can make it approximately error-transparent by modifying the SNAP gate:

$$S(\vec{\phi}) \to e^{i\phi}(P_{[s]}) + I - P_{[s]}, \tag{2.77}$$

where $P_{[s]} := \sum_{i=0}^s |i\rangle\langle i|$ is the projection onto the $s$-neighborhood of vacuum. Then the gate unitary becomes $U_X \to e^{i\phi}P_{\pm\alpha,s} + I - P_{\pm\alpha,s}$, where $P_{\pm\alpha,s} := D(\alpha)P_{[s]}D^\dagger(\alpha) + D(-\alpha)P_{[s]}D^\dagger(-\alpha)$ is the projection onto a neighborhood of $|\alpha\rangle$ and $|-\alpha\rangle$. Now, the effective error for the dephasing error becomes

$$EP_c = e^{i\delta\theta a^\dagger a}P_c + (1-e^{-i\phi})(P_{\pm\alpha,s} - I)e^{i\delta\theta a^\dagger a}|C_\alpha^+\rangle\langle C_\alpha^+| + (e^{i\phi}-1)P_{\pm\alpha,s}e^{i\delta\theta a^\dagger a}|C_{i\alpha}^+\rangle\langle C_{i\alpha}^+|. \tag{2.78}$$

For $|\delta\theta| \leq \Delta\chi T/2 < \pi/8$, we can choose $s = O(|\alpha|^2)$ such that the last two terms vanish in the $\alpha \gg 1$ limit, i.e.,

$$\begin{aligned}\langle C^+_{\alpha e^{i\delta\theta}}|P_{\pm\alpha,s}|C^+_{\alpha e^{i\delta\theta}}\rangle &\to 1,\\ \langle C^+_{i\alpha e^{i\delta\theta}}|P_{\pm\alpha,s}|C^+_{i\alpha e^{i\delta\theta}}\rangle &\to 0.\end{aligned} \tag{2.79}$$

Then we have $EP_c \approx e^{i\delta\theta a^\dagger a}P_c$ and the gate is error-transparent to dephasing as well.

Note that 1-fault-tolerance can no longer be rigorously attained (even in the larger-$\alpha$

limit) if using SNAP gates $S$ with a finite $\chi$-mismatch. Taking the second $S$ gate as an example, and suppose it has a $\chi$-mismatch $\Delta\chi'$, a single ancilla decay could cause a $e^{i\delta\theta' a^\dagger a}$ phase rotation with $|\delta\theta'| \leq \Delta\chi' T/2$ after $S$, which propagates to $e^{-i\delta\theta'[a^\dagger a+\alpha(a+a^\dagger)+\alpha^2]}$ after the later displacement. The extra displacement error after the gate is uncorrectable. Thus a single ancilla fault during the $X$-rotation can cause a first-order logical error with a probability $cp$, where $c$ is a constant depending on $\Delta\chi' T$. Nevertheless, if $\Delta\chi' T$ is small enough, the coefficient $c$ can be made comparable or even smaller than $p$, and we can still achieve good error suppression in practical regimes, as is shown in later numerical results in Fig. 2.14(a).

**XX rotation**

For large $\alpha$, the $XX$ rotation reads

$$\begin{aligned} XX(\delta) \approx e^{i\delta}(\left|C_\alpha^+, C_\alpha^+\right\rangle \left\langle C_\alpha^+, C_\alpha^+\right| + \left|C_{i\alpha}^+, C_{i\alpha}^+\right\rangle \left\langle C_{i\alpha}^+, C_{i\alpha}^+\right|) \\ + (\left|C_\alpha^+, C_{i\alpha}^+\right\rangle \left\langle C_\alpha^+, C_{i\alpha}^+\right| + \left|C_{i\alpha}^+, C_\alpha^+\right\rangle \left\langle C_{i\alpha}^+, C_\alpha^+\right|). \end{aligned} \tag{2.80}$$

We implement $XX(\delta)$ by adding a phase $\delta$ to the subspace spanned by $|\pm\alpha, \pm\alpha\rangle$ and $|\pm i\alpha, \pm i\alpha\rangle$. As illustrated in Fig. 2.13(b), we interfere the two modes through a $50:50$ beamsplitter, apply the number dependent phase shift $S(\vec{\delta}) = e^{-i\delta}|0\rangle\langle 0| + I - |0\rangle\langle 0|$ to both ports, and then interfere through another $50:50$ beamsplitter:

$$U_{XX} = \mathrm{BS}(\frac{\pi}{2})^\dagger (S \otimes S)\mathrm{BS}(\frac{\pi}{2}), \tag{2.81}$$

where $\mathrm{BS}(\theta) := \exp\left[\frac{\theta}{2}(ab^\dagger - a^\dagger b)\right]$ with $a$ and $b$ denoting the annihilation operator of the two involved modes, respectively.

To understand the effect of $U_{XX}$, we consider a coherent-state input $|\alpha, \beta\rangle$. The first BS

interferes the two coherent states:

$$\mathrm{BS}\,|\alpha,\beta\rangle = \left|(\alpha+\beta)/\sqrt{2},(\alpha-\beta)/\sqrt{2}\right\rangle, \tag{2.82}$$

We take the approximation $S\,|\gamma\rangle \approx e^{-i\delta\mathbf{1}[\gamma=0]}\,|\gamma\rangle$, where $\mathbf{1}[x]$ is the indicator function. Then the two SNAP gates in Eq. (2.81) add a nontrivial phase to the R.H.S. of Eq. (2.82) if $\alpha=\beta$ or $\alpha=-\beta$:

$$(S\otimes S)\mathrm{BS}\,|\alpha,\beta\rangle = e^{-i\delta(\mathbf{1}[\alpha=\beta]+\mathbf{1}[\alpha=-\beta])}\left|\frac{\alpha+\beta}{\sqrt{2}},\frac{\alpha-\beta}{\sqrt{2}}\right\rangle. \tag{2.83}$$

Finally, the last BS restores the input coherent state potentially with an extra phase:

$$U_{XX}\,|\alpha,\beta\rangle = e^{-i\delta(\mathbf{1}[\alpha=\beta]+\mathbf{1}[\alpha=-\beta])}\,|\alpha,\beta\rangle\,. \tag{2.84}$$

We remark that, when $\alpha$ and $\beta$ are only chosen from a set of discrete values $\{\alpha_i\}_i$ which are well-separated in the phase space, Eq. (2.84) provides an exact expression of the action of $U_{XX}$. The rigorous form of $U_{XX}$ is given Appendix B of Ref. [138]. To conclude, a two-mode coherent state accumulates a nontrivial phase if and only if the two coherent states have matched amplitudes and aligned/anti-aligned phases. Let $P_{\pm(i)\alpha}$ be the projection onto a four-dimensional subspace spanned by $|\alpha\rangle, |-\alpha\rangle, |i\alpha\rangle, |-i\alpha\rangle$, we then have

$$P_{\pm(i)\alpha}\otimes P_{\pm(i)\alpha}U_{XX}P_{\pm(i)\alpha}\otimes P_{\pm(i)\alpha} = e^{i\delta}(P_{\pm\alpha}\otimes P_{\pm\alpha}+P_{\pm i\alpha}\otimes P_{\pm i\alpha})+(P_{\pm\alpha}\otimes P_{\pm i\alpha}+P_{\pm i\alpha}\otimes P_{\pm\alpha}). \tag{2.85}$$

Note that Eq. (2.85) implies $P_c^{(AB)}U_{XX}P_c^{(AB)} = XX(\theta)$ where $P_c^{AB} = P_c^{(A)}\otimes P_c^{(B)}$ is the projector onto the collective code space of 4-legged cat on bosonic modes $A$ and $B$.

Now, we prove this $XX(\theta)$ gate is 1-FT according to Def. 1. We first consider the case where there is an input error $e^{i(\delta\theta_a a^\dagger a+\delta\theta_b b^\dagger b)}b^{k_b}a^{k_a}$ with $k_a,k_b\le 1$ and $|\delta\theta_a|,|\delta\theta_b|\le \Delta\chi T/2 < \pi/8$, but no fault during the gate. $b^{k_b}a^{k_a}$ simply commutes through the gate when acting on the code space due to the error-transparency form of $U_{XX}$ in Eq. (2.85).

Similar to proof for the $X$-axis rotation in Eq. (2.78), one can show that $U_{XX}$ is also approximately error-transparent to the phase rotation $e^{i(\delta\theta_a a^\dagger a+\delta\theta_b b^\dagger b)}$ by replacing $S \to e^{-i\delta}(\sum_{i=0}^{s}|i\rangle\langle i|)+I-(\sum_{i=0}^{s}|i\rangle\langle i|)$. We put the proof in Appendix **??**. As a result, the input error commutes through the gate and remains correctable.

To complete the proof that the $U_{XX}$ is 1-FT, we also need to show that for a perfect input state and any single fault during $U_{XX}$, each of the two output modes has an error of size at most $f(1)=(1,\Delta\chi T/2)$. As shown in Appendix B of Ref. [138], a single-photon loss during the gate propagates to an error of the form $c_1 a + c_2 b$, where $c_1, c_2 \in \mathbb{C}$, due to the error transparency of the SNAP gates and the error closure of the BSs. By using a $\chi$-matched ancilla for each SNAP gate, any single ancilla fault does not propagate to any bosonic data errors.

We note that similar to the $X$-axis rotation, the $XX$ rotation is not strictly 1-FT if there is a finite $\chi$-mismatch when executing the SNAP gates, as an induced phase rotation would propagate to uncorrectable errors after the last BS. Nevertheless, as we show numerically in Fig. 2.14, high-fidelity $XX$ rotation can still be realized in practical regimes even with a finite but small $\chi$-mismatch.

**X-basis state preparation**

The $+1$ $X$ basis eigenstate is a two-legged cat state with an even photon parity $|+_L\rangle = |C_\alpha^+\rangle = c_\alpha^+(|\alpha\rangle+|-\alpha\rangle)$. Observe that $|+_L\rangle \propto P_+|\alpha\rangle$, i.e. the even-parity projection of a coherent state $|\alpha\rangle$. Thus, we can prepare the even cat state by first preparing a coherent state $|\alpha\rangle$, and then performing a non-destructive parity measurement to project it to an even cat state (up to a parity frame update). For 1-FT state preparation, unlike the 1-FT photon-loss correction protocol in Alg. 2, we do not need to repeat the parity measurement three times as it allows a noisy output with up to $f(1)=(1,\Delta\chi T/2)$ error for up to a single fault during the parity measurement (see Def. 3). Concretely, we implement the following protocol:

**Algorithm 3:** 1-FT $X$-basis state preparation

Prepare the bosonic mode in the coherent state $|\alpha\rangle$.
$o \leftarrow e$ `// records the parity measurement outcome`
**while** $o = e$ **do**
    Prepare the ancilla in the $|+\rangle$ state, apply the dispersive coupling for a time $T = \pi/\chi_f$, and measure the ancilla in the $= \{|+\rangle, |-\rangle, |e\rangle\}$ basis with an measurement outcome $o$.
    **if** $o = e$ **then**
        Apply a phase rotation $e^{i\Delta\chi T a^\dagger a/2}$ to the bosonic mode.
Apply a parity correction if $o = -$.

Note that the $X$-basis state preparation here allows a finite $\chi$-mismatch.

**Z-basis measurement**

The $Z$-basis measurement admits the form of measuring photon number modulo 4. In order to obtain the correct logical measurement outcome in the presence of a single-photon loss, as required by Def. 4, we insert a non-destructive parity measurement before each logical $Z$ measurement. The full FT protocol is presented in Alg. 4.

Note that each modulo-4 photon number measurement $o_{i,b}$ is conditioned on the parity measurement outcome $o_{i,a}$, i.e. we distinguish the photon number between 0 mod 4 and 2 mod 4 for even parity ($o_{i,b} = +$) and between 3 mod 4 and 2 mod 4 for odd parity ($o_{i,a} = -$).

To verify that the 1-FT measurement condition in Def. 4 holds, one simply observe that a single photon loss before the measurement can be captured by the parity measurements, and any single fault during the measurement protocol can only cause at most one measurement error on one of $\{o_{i,b}\}_{i=1,2,3}$, and thus does not affect the majority voting. Note that the $Z$-basis measurement here can also allow a finite $\chi$-mismatch between the ancilla and the bosonic mode, as dephasing errors commute with the measurements.

**X-basis measurement**

The $X$-basis measurement amounts to distinguishing the phase of the coherent states modulo

**Algorithm 4:** 1-FT $Z$-basis measurement

**for** $i \leftarrow 1$ **to** 3 **do**
  $o_{i,a} \leftarrow e$;
  **while** $o_{i,a} = e$ **do**
    Prepare the anilla in the $|+\rangle$ sate, apply the dispersive coupling for a time $T = \pi/\chi_f$, and measure the ancilla in the $\{|+\rangle, |-\rangle, |e\rangle\}$ basis with an measurement outcome $o_{i,a}$.
  $o_{i,b} \leftarrow e$;
  **while** $o_{i,b} = e$ **do**
    **if** $o_{i,a} = +$ **then**
      Prepare the anilla in the $|+\rangle$ sate, apply the dispersive coupling for a time $T = \pi/2\chi_f$, and measure the ancilla in the $\{|+\rangle, |-\rangle, |e\rangle\}$ basis with an measurement outcome $o_{i,b}$.
    **else**
      Prepare the anilla in the $|+\rangle$ sate, apply the dispersive coupling for a time $T = \pi/2\chi_f$, apply an ancilla phase rotation $e^{-i\frac{\pi}{2}|f\rangle\langle f|}$, and measure the ancilla in the $\{|+\rangle, |-\rangle, |e\rangle\}$ basis with an measurement outcome $o_{i,b}$.
Obtain the logical measurement outcome as the majority voting of $\{o_{i,b}\}_{i=1,2,3}$.

$\pi$. We achieve this by interfering the mode $a_i$ with another mode $b_i$ in a coherent state $|\alpha\rangle$ through a $50:50$ beam splitter and measuring if the two output modes $a_o, b_0$ have less than $s$ photons. We obtain a logical $-$ if both modes have more than $s$ photons and a logical $+$ otherwise, i.e. we implement the following POVMs:

$$\begin{aligned} M_- &= [I_{a_o} - \sum_{i=0}^{s}(|i\rangle_{a_o}\langle i|)] \otimes [I_{b_o} - \sum_{i=0}^{s}(|i\rangle_{b_o}\langle i|)] \\ &\approx (\hat{I}_{a_i} - \sum_{i=0}^{s}(|\alpha, i\rangle_{a_i}\langle\alpha, i|))(I_{a_i} - \sum_{i=0}^{s}(|-\alpha, i\rangle_{a_i}\langle -\alpha, i|)), \\ M_+ &= I - M_- \\ &\approx \sum_{i=0}^{s}(|\alpha, i\rangle_{a_i}\langle\alpha, i|) + \sum_{i=0}^{s}(|-\alpha, i\rangle_{a_i}\langle -\alpha, i|)), \end{aligned} \tag{2.86}$$

where each subscript labels the mode that a state or an operator belongs to.

Measuring if one mode has less than $s$ photons can be realized by dispersively coupling

it to a qubit, driving the qubit from $|g\rangle$ to $|e\rangle$ conditioned on the mode having less than $s$ photons, and measuring the qubit in the $|g\rangle, |e\rangle$ basis. In the interaction picture associated with the dispersive coupling, the Hamiltonian reads

$$\tilde{H}_{AC} = \Omega \left( |e\rangle \langle g| \otimes P_{[s]} + h.c. \right). \tag{2.87}$$

Recall that $P_{[s]} := \sum_{i=0}^{s} |i\rangle \langle i|$. In the absence of errors, the 0-th order conditional operations are

$$\begin{aligned} \langle\langle e|\mathcal{G}^{[0]}(T)|g\rangle\rangle &= P_{[s]} \bullet P_{[s]} + O(p), \\ \langle\langle g|\mathcal{G}^{[0]}(T)|g\rangle\rangle &= (I - P_{[s]}) \bullet (I - P_{[s]}) + O(p). \end{aligned} \tag{2.88}$$

A single fault will affect the measurement outcome or cause a bosonic error diagonal in the Fock basis. The former can be tolerated by repeating the above measurement three times and performing majority voting, while the latter simply commutes with the measurements and does not affect the (later) measurement outcome.

To check this $X$-basis measurement scheme is 1-FT according to Def. 4, we also need to verify that the measurement outcome is correct for any input error $(k, \theta) \leq (1, \Delta\chi T/2)$. First, a single-photon loss does not affect the measurement outcome since $a$ does not change the phase of any coherent states. Second, a small phase rotation rotates $|\alpha\rangle$ to $\left|\alpha e^{i\theta}\right\rangle$. Similar to the argument for the $X$-axis rotation, the $X$-basis measurement outcome is correct as long as the POVM $M_+$ captures $\left|\pm\alpha e^{i\theta}\right\rangle$ but not $\left|\pm i\alpha e^{i\theta}\right\rangle$.

**Error correction**

To correct both loss and dephasing errors, i.e. data errors with $(k > 0, |\theta| > 0)$, we employ a Knill-type QEC [139] using a teleportation circuit shown in Fig. 2.13(c). A fresh ancilla bosonic mode $b$ is initialized in $|+\rangle$ state and gets entangled with the data mode $a$ via a $XX$ rotation along with singe-mode rotations. The data mode $a$ is then measured in the $Z$ basis, where the measurement outcome is used to apply a feedback $Z$ operation on the $b$ mode. All the gadgets here are 1-FT using previous constructions. Moreover, they are error-

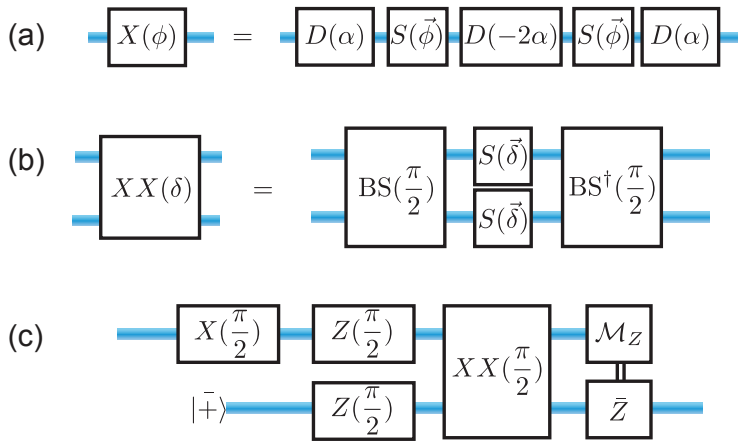


Figure 2.13: Illustration of the $X$-axis rotation (a), $XX$ rotation (b), and teleportation-based EC (c) in the level-1 gadgets $\mathcal{S}$ for the four-legged cat.

transparent to any input error on the $a$ mode smaller than $f(1) = (1, \Delta\chi T/2)$. Therefore, the input data error simply commutes through all the gates and does not propagate to the $b$ mode. Furthermore, the 1-FT $Z$-basis measurement is correct for an error smaller than $f(1)$. Therefore, such an input error can be corrected by the EC gadget.

To verify the 1-FT EC conditions, we need to further show that a single fault during the teleportation circuit only leads to a correctable residual error of size at most $f(1)$ at the output of the $b$ mode. Since we are using 1-FT gates, the output for the $a$ or $b$ mode (before the $Z$ measurement) has an error at most $f(1)$.

As shown in Fig. 2.14(a), we numerically evaluate the average infidelity of the teleportation gadget in Fig. 2.13(c). In the absence of $\chi$ mismatch (see the blue curve), we show that it has an error rate that scales as $(\kappa/\Omega)^2$, manifesting the fault tolerance of its composing gadgets, which cover the entire $\mathcal{S}$ other than the $X$-basis measurement. There is an error floor in the low $\kappa/\Omega$ regime, which is exponentially suppressed by $|\alpha|^2$, due to the finite-size imperfection of the $X$ rotation and the $XX$ rotation. In the presence of a finite $\chi$ mismatch, a rigorous second-order error suppression can no longer be attained due to the induced random phase rotations during the $X$- and $XX$-rotation gates. However, sufficient error suppression can still be achieved with a finite but small $\chi$-mismatch in practically

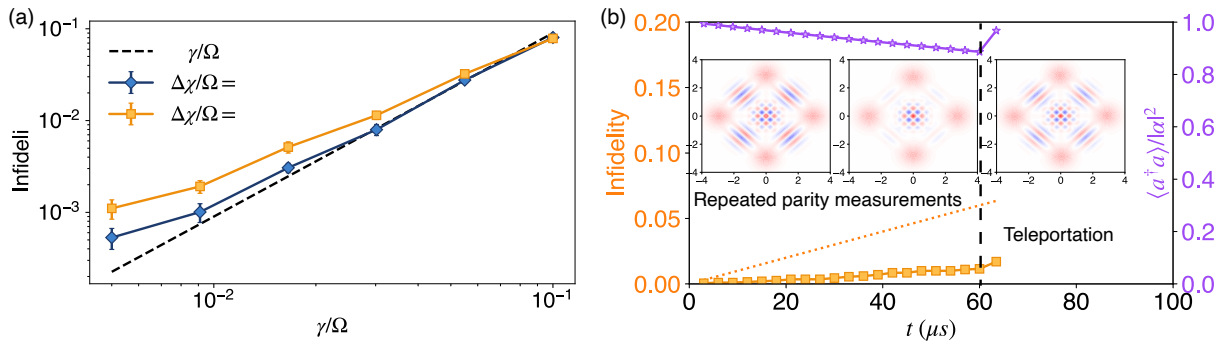


Figure 2.14: (a) Average infidelities of an error-correction gadget using teleportation in Fig. 2.13(c) as a function of $\gamma/\Omega$ with perfect $\chi$ matching (blue line) or finite $\chi$ mismatches (orange line). Here, we use experimental parameters from Ref. [7] for the coherent interaction strengths $\chi_f = 2\pi \times 1\text{MHz}$, $\Omega = 0.3\chi_f$, $g_{\text{BS}} = 2\chi_f$. We consider single-photon loss, ancilla decay from $f$ to $e$, ancilla decay from $e$ to $g$, and ancilla dephasing $\mathcal{D}[|e\rangle\langle e| + 2|f\rangle\langle f|])$ with rates $\kappa$, $\gamma_{f\to e}$, $\gamma_{e\to g}$, and $\gamma_\phi$, respectively. We assume the ancilla error rates are much larger than the cavity loss rate and set $\gamma_{f\to e} = \gamma_{e\to g} = \gamma$, $\gamma_\phi = \gamma/4$, and $\kappa = \gamma/10$ [7]. We choose $\alpha = 2.9$, which is a sweet spot for the four-legged cat that minimizes the back action of photon loss [8]. (b) The accumulation of average infidelity and decay of mean photon number $\langle a^\dagger a\rangle$ for 40 rounds of repeated parity measurements (infidelities are shown for every two rounds) followed by teleportation. We use the same coherent parameters $\chi_f, \Omega$ and $g_{\text{BS}}$ as in (a), a finite $\chi$ mismatch $\Delta\chi = \Omega/10$, and the experimental error rates from Ref. [7]: $\kappa = 2\text{KHz}$, $\gamma_{f\to e} = \gamma_{e\to g} = \gamma = 20\text{KHz}$ and $\gamma_\phi = 5\text{KHz}$ (with the same ratios between these error rates are in (a)). The teleportation pumps energy into the system and suppresses the random phase rotations caused by $\Delta\chi$. The three Wigner plots depict the density matrix at the input, before and after the teleportation respectively.

relevant regimes (see the orange and gree curves).

In practice, where photon loss is typically the predominant error source, repeated parity measurements that correct photon losses (see Alg. 2) suffice for extending the lifetime of the cats. Such a robust memory that reaches the break-even point has been experimentally demonstrated [65]. However, only parity measurements are not enough to protect the cats during long computational tasks as the mean photon number would keep decaying (the parity measurement and gates in $\mathcal{S}$ are energy-preserving operations that commute with $a^\dagger a$) due to deterministic energy loss to the environment. We propose to solve this problem by inserting the teleportation gadget periodically in between certain rounds of parity measurements, which pumps energy into the system and restores the amplitude of the cats. Furthermore,

the teleportation can suppress the accumulation of random phase rotations if, for example, there is some finite $\chi$-mismatch or small cavity dephasing errors $\kappa_\phi \mathcal{D}[a^\dagger a]$. We demonstrate such effects numerically in Fig. 2.14(b).

Note that the teleportation-based QEC scheme for rotation-symmetrical codes was first proposed in Ref. [59]. Compared to their teleportation circuit, our circuit in Fig. 2.13 uses different gates and measurements and, in particular, avoids (i) nonlinear interaction between bosonic modes and (ii) phase measurements that are challenging to implement in practice.

### *2.3.5 Concatenated QEC with four-legged cats*

With the set of 1-FT level-1 gadgets in $\mathcal{S}$, we can concatenate the level-1 logical qubits (four-legged cats) with a level-2 qubit code for arbitrary error suppression. We show such a concatenated architecture in Fig. 2.15. The basic elements for each level-1 qubit are simply a storage bosonic mode and a three-level ancilla that are dispersively coupled. The ancilla is used for the fault-tolerant operation of the bosonic qubit in each storage mode, including state preparation, photon-loss correction, gates, and measurements (see Sec. 2.3.4). In addition, a small number of extra bosonic modes shared by neighboring storage modes, which we refer to reservoir modes, are used to pump energy into the storage modes periodically via teleportation (see Fig. 2.13(c)).

The level-2 QEC requires certain couplings between level-1 qubits. Importantly, we can achieve this by introducing only BS coupling between nearest-neighbor storage bosonic modes (see Fig. 2.15) for 2D stabilizer codes. The level-2 syndrome-extraction circuits are made of non-destructively measurement of high-weight Pauli operators, featuring a sequence of two-qubit entangling gates such as the CNOT gate. In Fig. 2.16(a), we show how one can get a level-1 CNOT gate using 1-FT single-mode and two-mode rotations in $\mathcal{S}$. Although the complied circuit is long with a depth 6, we note that one can usually reduce the depth per CNOT gate when considering the full stabilizer measurement circuits. As an example, we

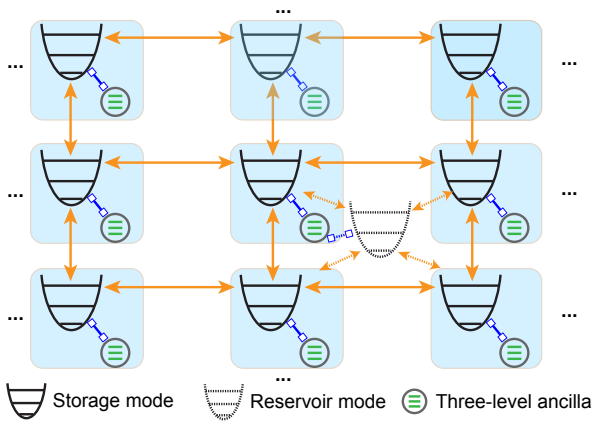


Figure 2.15: Hardware layout for concatenated 2D codes with four-legged cats. Each level-1 logical qubit (blue box) consists of a storage bosonic mode and a three-level ancilla, which are dispersively coupled. BS coupling between neighboring storage bosonic modes is required for the level-2 QEC. In addition, reservoir modes (with only one shown here as an example) shared between neighboring storage modes are used to pump energy into the system via teleportation (see Fig. 2.13(c)).

can measure a weight-$n$ $X$ Pauli operator using a circuit of depth $2n + 4$ (see Fig. 2.16(b)). We leave the evaluation and optimization of the error rates of level-1 gates, e.g. the CNOT gate, as well as the threshold and resource overhead of concatenated codes to future work. Nevertheless, we remark that each CNOT gate (se Fig. 2.16(a)) uses similar gadgets as those for teleportation (see Fig. 2.13(c)), and each CNOT gate in a syndrome extraction depth (see Fig. 2.16(b)) has a similar depth as the teleportation on average, we expect the CNOT gates have a similar error rate as the teleportation shown in Fig. 2.14(b). Using this rough estimates, a gate error rate below $10^{-2}$, which corresponds to the gate error threshold for the surface codes, is achievable using the current circuit-QED hardware.

To sum up, our construction of $\mathcal{S}$ in this work enables a practical, hardware-efficient architecture for fault-tolerant quantum computing, which features only one bosonic mode and one qutrit per level-1 logical qubit and only requires low-order physical interactions (dispersively coupling and beam-splitter interaction) that have been experimentally demonstrated. Furthermore, realizing high-fidelity level-1 gadgets with error rates below the threshold requirement of the level-2 codes is promising for near-term experimental demonstrations.

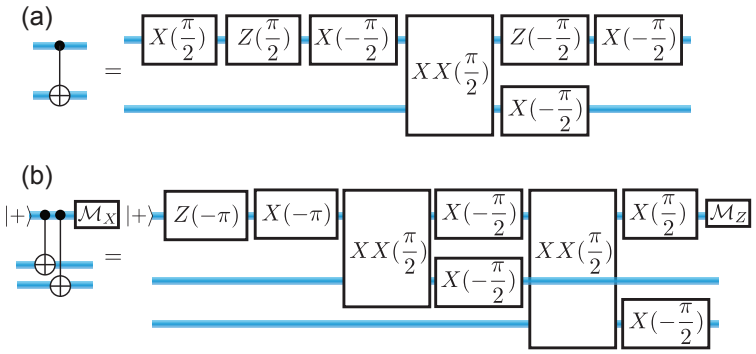


Figure 2.16: Compilation of level-1 CNOT (a) and a stabilizer $X^{\otimes 2}$ measurement circuit (b) using our constructed 1-FT level-1 gadgets in $\mathcal{S}$.

### *2.3.6 Discussion and comparison to related schemes*

The fault-tolerant gadgets $\mathcal{S}$ that we develop in Sec. 2.3.4 for the four-legged cat can be applied to other rotation-symmetric codes [59], whose codewords are invariant under a $N$-fold rotation $R = \exp\left[i(2\pi/N)a^\dagger a\right]$. Taking $N = 2$ for example, an arbitrary code with a two-fold rotation symmetry have codewords

$$\begin{aligned} |+_\Theta\rangle &\approx \frac{1}{\sqrt{\mathcal{N}_+}}(I + e^{i\pi a^\dagger a})|\Theta\rangle, \\ |-_\Theta\rangle &\approx \frac{1}{\sqrt{\mathcal{N}_-}}(e^{i\pi a^\dagger a/2} + e^{i3\pi a^\dagger a/2})|\Theta\rangle, \end{aligned} \tag{2.89}$$

where $\mathcal{N}_\pm$ are normalization constants, and the approximation holds when the base state $|\Theta\rangle$ is localized in phase space, i.e. $\langle\Theta| e^{i\pi a^\dagger a/2} |\Theta\rangle \approx 0$. The fault-tolerant gadgets in $\mathcal{S}$ can be applied to such an arbitrary rotation-symmetric code with a localized base state $|\Theta\rangle$, except that for the $X$-basis state preparation in Alg. 3, we need to replace the initial state with $|\Theta\rangle$ in the first step. In particular, the $X$ rotation and $XX$ rotation still work since they are based on the phase-space locality of the base state.

In Tab. 2.2, we compare our construction of fault-tolerant gadgets for rotation-symmetrical codes that can correct photon losses with those in the literature. In particular, compared to the gadgets in Ref. [59] using only bosonic modes, our gadgets using qutrit ancillae avoid the

| Gadgets | Prior schemes | Our scheme |
|---|---|---|
| Error correction | PI parity measurement [75]; Two one-bit teleportation with two ancillary bosonic modes [59] Engineered dissipation [106, 84] | GPI parity measurement + one-bit teleportation with a shared ancillary mode |
| $Z$-type gates | PI SNAP gate [7, 6]; Self-Kerr $(a^\dagger a)^2$ Hamiltonian [59] | GPI SNAP gate |
| $X$-type gates | Teleported Hadamard gate with an ancillary bosonic mode [59] | $X$-axis rotation using cavity displacements and SNAP gates |
| Entangling gate | CZ gate using cross-Kerr $a^\dagger a \otimes b^\dagger b$ [59] | $XX$ rotation using beam-splitter + SNAP gates |
| $X$-basis measurement | Phase measurement [59] | Beam splitter + SNAP gates |

Table 2.2: Comparison of different constructions of fault-tolerant gadgets for rotation-symmetrical codes that can correct photon losses. We denote $Z$-type gates as those that preserve the photon number (alternatively, those that add photon-number dependent phases), and $X$-type gates as those that do not preserve the photon number.

demand of nonlinear interaction between bosonic modes and the phase measurement, which are both challenging to engineer in practice.

# CHAPTER 3

# TAILORED TOPOLOGICAL CODES

Topological stabilizer codes, particularly those that are geometrically local in two dimensions, have emerged as the leading candidate for fault-tolerant quantum computing. These codes offer several advantages: (1) Local Interactions: They require only local interactions between qubits, with syndrome extraction circuits being geometrically local and of constant depth. (2) High Threshold: They exhibit a relatively high threshold, such as approximately 1% for planar surface codes [17, 41, 42]. (3) Efficient decoders: They are equipped with efficient decoders, such as the matching decoder for surface codes [43]. (4) Well-Studied Logical Operations: Their logical operations are well understood and compatible with geometric constraints, such as lattice surgery operations for surface codes [10].

In fact, the baseline fault-tolerant scheme considered in this thesis (see Sec. 1.4) is based on the planar surface code, one of the most popular topological codes.

However, implementing standard fault-tolerant schemes based on these topological codes on existing hardware still faces key challenges:

1. Although the circuit-level threshold of leading topological codes ($\sim 1\%$) is relatively high among known QEC codes, it may not be sufficient for demonstrating deep error suppression in the far-below-threshold regime using near-term devices. State-of-the-art devices have so far only demonstrated physical error rates slightly below 1% [21, 22]. There are significant challenges in further reducing the physical error rates while scaling up the system sizes.

2. Space Overhead: Existing topological codes come with a large space overhead, which may hinder scalability. For example, a distance of approximately 30 is required for sufficiently suppressing errors using surface codes [17, 20, 18, 19], resulting in a space overhead of approximately 1000.

3. Correlated Non-Local Errors: It is challenging for topological codes to handle correlated non-local errors, which are realistic noise sources in many physical platforms [140, 141, 142, 143]..

To address these challenges, we can leverage certain hardware characteristics or noise profiles that are often overlooked by standard fault-tolerant schemes. For instance, systems like stabilized cat/squeezed-cat qubits (see Chapter 2) often exhibit biased noise channels instead of the symmetrical depolarizing noise model considered by standard fault-tolerant schemes when deriving the $\sim 1\%$ threshold.

In Sec. 3.2, we will present a family of topological codes tailored to biased noise, aiming to boost thresholds and reduce space overheads. This section will aim to address the first and second challenges.

In Sec. 3.3, we will introduce a new scheme using concatenated surface codes and distributed QEC to combat catastrophic correlated errors, addressing the third challenge.

Before presenting the new results in Sec.3.2 and Sec.3.3, we will review some basics of topological codes in Sec. 3.1.

## 3.1 Basics of topological codes

In this section, we provide a brief review of the topological stabilizer codes. We first review the definitions and characteristics of topologcal codes and list a few representative code families in Sec. 3.1.1. We then dive into a particular type of $2D$ topological code — the surface code, and review its construction and connection to homological algebra [144].

### *3.1.1 Definitions, characteristics, and representative codes*

Topological stabilizer codes are broadly referred to as stabilizer codes such that [145] (i) physical qubits are arranged in a finite-dimensional lattice of variable size, (ii) the stabilizer

generators $\{S_i\}$ have geometrically local support on the lattice and (iii) logical operators have support comparable to the lattice size. These topological stabilizer codes are connected to the topological phases of matter by viewing the codespace as the ground state manifold of a local gapped many-qubit Hamiltonian $H = -\sum_i S_i$. The resulting condensed matter system is said to be topologically ordered [146]: there exists a ground state degeneracy with a non-local origin.

The major characteristic of topological codes is that they protect information against local errors by encoding it into global degrees of freedom associated with the underlying lattice (or a manifold). At the same time, topological codes require only geometrically local syndrome extraction for error correction and thus seem well-suited for many experimental systems, where locality is an important constraint. In addition, topological codes tend to have relatively high thresholds [147].

However, the geometrical locality of the topological codes also puts some constraints on their performance and computational power. For example, the parameters of a $D$-dimensional topological code are constrained by the bound [148, 149]:

$$kd^{\frac{2}{D-1}} \leq O(n), \tag{3.1}$$

where $k$, $n$, and $d$ denote the number of logical qubits, the number of physical qubits, and the code distance, respectively.

Furthermore, for any $D$-dimensional topological code, the group of logical gates implemented via local quantum circuits of constant depth must be contained in the $D$-th level of the Clifford hierarchy $\mathcal{C}_D$, where $\mathcal{C}_1 = \mathcal{P}_n$ (the Pauli group) and the $D$-level is defined recursively as the set of all unitaries which map the Pauli group onto unitaries in the $(D-1)$-level [150, 33, 151].

Typical representatives of topological codes in $2D$ and $3D$ include the toric code [34, 35] and the color code [152] (the latter is equivalent to the former up to local unitary

transformations [153]).

### *3.1.2 Surface codes*

#### Construction

As shown in Fig. 3.1, a surface code is defined on a $2D$ lattice, or a tiling of a $2D$ surface, where the qubits are associated with the edges and the $Z$ ($X$) checks are associated with the plaquettes (vertices). Each plaquette $Z$ check checks four qubits on its boundary edges and each vertex $X$ check checks four qubits on its adjacent edges. A planar surface code has four open boundaries and logical operators are associated with error chains that connect the opposite boundaries. A planar surface code has two pairs of opposite open boundaries, called the "smooth" and "rough" boundaries, respectively. We refer to a $Z$ ($X$) string operator as a Pauli $Z$ ($X$) operator that is supported on a chain of qubits and only triggers the $X$ ($Z$) checks on its two endpoints. The planar surface code encodes a single logical qubits, whose logical $Z$ operator is associated with an $Z$ string that connects the two rough boundaries while $X$ operator is associated with a $X$ string that connects the two smooth boundaries. See Ref. [17] for more details about the surface code construction, and Ref. [35] for its connection to topological physics.

The terminology "surface codes" is also often referred to a generalized class of codes defined on a general tilled surface, or manifold $\mathcal{M}$ (e.g. the surface of a torus), where the qubits and checks are defined analogously as edges and plaquettes (vertices). This generalization can be best understood using the phonological description of the codes, as presented in Sec. 3.1.2.

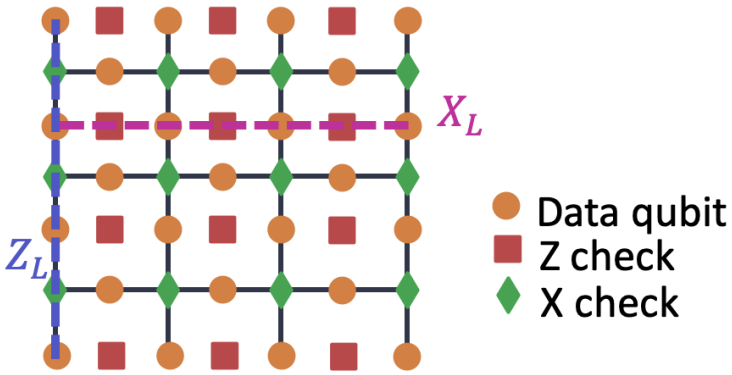


Figure 3.1: Illustration of a planar surface code. Given a tiling of a 2D surface, the qubits are associated with the edges and the $Z$ and $X$ checks are associated with the plaquettes and the vertices, respectively. A planar surface code has four open boundaries and logical operators are associated with error chains that connect the opposite boundaries.

## Connection to homological algebra

Here, we briefly discuss how the general class of surface codes can be described using the homological algebra. We refer the readers for more details in Ref. [154].

Any CSS stabilizer code with check matrices $H_X$ and $H_Z$ can be described as a chain complex:

$$C_2 \xrightarrow{\partial_2} C_1 \xrightarrow{\partial_1} C_0, \tag{3.2}$$

where each $C_i$ is a vector space over $\mathbb{F}_2$, and each $\partial_i$ is a linear map $\partial_i : C_i \to C_{i-1}$ that is also called a boundary operator. The vectors in $C_2, C_1$, and $C_0$ are associated with the $Z$ stabilizers, the qubits, and the $X$ stabilizers of the code, respectively. With a proper choice of basis, the boundary maps are given by the parity-check matrices:

$$\partial_2 = H_Z^T, \quad \partial_1 = H_X. \tag{3.3}$$

The boundary maps satisfy $\partial_1 \partial_2 = 0$ since $H_X H_Z^\intercal = 0$, which is guaranteed by the commutation relation between the $X$ and $Z$ stabilizers. The kernel of $\partial_1$, $\ker(\partial_1)$, is associated with the Pauli-Z operators that commute with all the $X$ stabilizers. The image of $\partial_2$,

$\mathrm{im}(\partial_2) \cong \mathrm{rs}(H_Z)$ (the row space of $H_Z$), is associated with all the Pauli-$Z$ operators that belong to the stabilizer group. Therefore, the first homology group, $H_1 := \ker(\partial_1)/\mathrm{im}(\partial_2)$, is associated with the logical $Z$ operators of the code. Similarly, the logical $X$ operators are associated with the first cohomology group $H_1' := \ker(\partial_1')/\mathrm{im}(\partial_2')$, where the co-boundary maps are given by $\partial_1' = H_Z$ and $\partial_2' = H_X^T$.

A nice feature of representing a quantum code using a chain complex is that the logical operators are now associated with the homology/cohomology groups of the complex, which could be better understood using the homology theory.

To describe the generalized surface codes, we can further embed the chain complex in Eq. (3.2) into a Manifold $\mathcal{M}$. Specifically, given a tilling of $\mathcal{M}$, we associate the basis of $C_2$, $C_1$, and $C_0$ to the plaquettes, edges, and vertices of the tilled $\mathcal{M}$, respectively. Then the boundary maps $\partial_2$ and $\partial_1$ are naturally given by the boundaries of plaquettes and edges (the endpoints), respectively. Now, the logical operators are associated with the homology/cohomology groups of the manifold $\mathcal{M}$, which are a topological property independent of the specific tilling. For example, if $\mathcal{M}$ is a torus, the two generators of the homology group are associated with the two inequivalent non-contractable loops on the torus, representing the two logical $Z$ operators of the two encoded logical qubits.

## 3.2 Tailored topological codes for biased noise

Actively correcting unstructured errors, such as depolarizing noise, is challenging because topological codes designed for such errors have a relatively low threshold and require large resource overhead [17, 18, 155, 156]. However, physically relevant errors typically have certain structures, which can be exploited to design QEC schemes that are less demanding. As an example, many physical systems, such as bosonic systems encoded in cat codes [60, 66, 79, 45, 46, 106], have a noise channel biased towards dephasing (see Chapter 2). One can then take advantage of the noise bias and design tailored topological codes that have a boosted

performance against the biased noise [157, 89, 90, 91, 92, 69]. In particular, Ref. [92] shows that the so-called XZZX surface codes with XZZX-type stabilizers exhibit exceptionally high thresholds as well as reduced resource overhead when the noise is biased towards dephasing.

As the error rates of physical systems readily approach or even fall below the fault tolerance threshold [44, 158, 159, 160], it is the resource overhead that ultimately limits the practical application of QEC schemes. The analysis of the XZZX surface codes in Ref. [92] showed very promising thresholds for that class of codes, but the question of resource overhead was only briefly addressed.

In this work, we focus on designing QEC codes and schemes that can reduce the resource overhead for fault-tolerant QEC under experimentally relevant biased noise. More specifically, given a (finite) noise bias, we aim to find codes that use as few qubits as possible to suppress the logical error rate to a target level. As we will show in Sec. 3.2.1, instead of numerically extracting the logical error rates, we can characterize the performance of different codes against biased Pauli noise by estimating their effective code distance $d'$, which takes the bias into consideration and serves as an analogy to the code distance $d$ in the biased-noise setting. The notion of effective distance was introduced in Ref. [161], and here we use an alternative (though related) definition. For the physical error rate $p \ll 1$, the effective code distance $d'$ approximately determines how the logical error rate $p_L$ scales with $p$, i.e., $p_L \sim p^{d'/2}$, and it thus serves as a good proxy for the logical error rate. Now our task simply becomes finding codes that use the minimal number of qubits $n$ to reach a target effective distance $d'$ among certain code families. Given a target effective distance, we can then characterize the efficiency of a code by the code size required to achieve that effective distance.

To construct highly qubit efficient codes, we start from the observation that the well-known five-qubit code [39], with stabilizers comprising the cyclic permutation of $XZZXI$, is the smallest code among all possible codes with its effective distance (3 and 5, respectively)

for both depolarizing and infinitely biased Pauli Z noise. This indicates that codes with XZZX-type stabilizers could potentially be resource efficient over a wide range of bias [157]. In Sec. 3.2.2, We generalize the five-qubit code by introducing a family of XZZX cyclic codes, which inherit the cyclic structure and all have weight-four XZZX-type stabilizers. These cyclic codes can reach the optimal effective distance $n = d'$ against infinitely biased noise since they exhibit a repetition-code structure under pure Pauli $Z$ noise.

To facilitate the analysis of their performance under finite-bias noise, we map them to a family of topological codes in Sec. 3.2.2. Concretely, by wrapping the cyclic codes around a torus, we find that these codes belong to a family of XZZX generalized toric codes (GTCs), first introduced in Ref. [162] (albeit called checkerboard and non-bipartite rotated toric codes). The GTCs are constructed by first drawing a square qubit lattice with faces representing the XZZX stabilizers, and then identifying qubits that differ by a periodicity vector within the span of two basis periodicity vectors $\vec{L}_1, \vec{L}_2$ (see Fig. 3.2c). A GTC is therefore specified by its periodic boundary condition induced by $\vec{L}_1$ and $\vec{L}_2$. As shown in Sec. 3.2.3, the GTCs share similarly high thresholds with the XZZX surface codes, which we attribute to the local equivalence of their check operators on a torus. Using our tailored efficient decoders, the code-capacity thresholds of the GTCs roughly track the Hashing bound (what is achievable with random coding [103, 163]), and their phenomenological thresholds increase from 3.5% to 10% when the bias parameter (which we will introduce later) increases moderately from 1 to 4.

More importantly, because of the nontrivial boundary conditions (meaning nontrivial choices of $\vec{L}_1, \vec{L}_2$), the GTCs can be more resource efficient than the XZZX surface codes with either the open or closed rectangular boundaries considered in Ref. [92]. We derive the effective distance of the GTCs using topological (or geometrical) tools in Sec. 3.2.2, and from this, we can optimally choose $\vec{L}_1, \vec{L}_2$ if given the value of a bias parameter $\omega$ (defined later). The optimal codes satisfy $n = d'^2/2\omega$, which indicates that the tailored GTCs require re-

sources that scale quadratically in the target effective distance, similar to the standard surface codes (for depolarizing noise), but with a reduction by a factor of $2\omega$. We note that nontrivial choices of the periodicity vectors can be viewed as applying boundary twists to toric codes with periodicity vectors aligned with the square qubit lattice [162, 164, 93]. Ref. [93] showed that certain boundary twists can improve the performance of the conventional XZZX toric codes under infinite bias. As we will show later, the GTCs family we consider covers their constructions and include more codes with a larger effective distance-block size ratio. Furthermore, we provide tools to systematically optimize the code performance by adaptively choosing the periodicity vectors (or equivalently, applying boundary twists) given any finite bias.

In Sec. 3.2.3, combining the analysis for the cyclic codes and the GTCs, we obtain the optimal performance for the XZZX codes, given a bias parameter $\omega$. For $n \leq 2\omega$, the optimal codes are those with cyclic (repetition-code) structures, and the optimal resource-distance dependence $n = d'$ is achieved. For $n > 2\omega$, the optimal codes are the GTCs with an optimized layout and have a quadratic resource scaling $n = {d'}^2/2\omega$ with an extra reduction by the factor of $2\omega$.

Lastly in Sec. 3.2.3, we show that we can preserve the large effective distance of the tailored XZZX codes and maintain the scaling of the logical error rate $p_L \sim p^{d'/2}$ in the fault-tolerant regime by using only *one* flag qubit, which is recently introduced for low-overhead fault-tolerant QEC [165, 166, 167, 168].

### *3.2.1 Effective code distance for asymmetric Pauli noise*

Here, we consider error correction under an i.i.d. asymmetric Pauli channel $\mathcal{E}(\rho) = (1-p)\rho + \sum_{\sigma\in\{X,Y,Z\}} p_\sigma \sigma\rho\sigma$, where $\{p_\sigma\}$ denotes an asymmetric probability distribution of three Pauli errors and $p = \sum_\sigma p_\sigma$ is the total error probability. The largest Pauli error probability is denoted as $p_m$, i.e. $p_m = \max_\sigma p_\sigma$. To estimate the performance of different

error-correcting codes under the asymmetric channel, we define the effective distance $d'$ of a stabilizer code as the minimum modified weight of logical operators, with the noise-modified weight of a Pauli $\sigma$ given by $\mathrm{wt}'(\sigma) \equiv \log p_\sigma / \mathcal{N}$, where $\mathcal{N}$ is a normalization factor; see Ref. [161] for an alternative but related definition of the effective code distance. To normalize the effective weight of the most probable Pauli error to 1, we choose $\mathcal{N} = \log p_m$. The effective weight of a $n$-qubit Pauli string $P = \bigotimes_{i=1}^{N} \sigma_i$ with non-identity support on $N$ qubits characterizes its error probability since, by definition, $\Pr(P) = p_m^{\sum_{i=1}^{N} \mathrm{wt}'(\sigma_i)} \times (1-p)^{n-N} \approx p_m^{\mathrm{wt}'(P)}$ (to leading order in $p$). The effective code distance, therefore, roughly characterizes how the logical error rate $p_L$ scales with the physical error rate $p$: In general $p_L$ is suppressed to certain order $r$ of $p$, i.e. $p_L \sim p^r$, and $r$ is approximately given by $d'/2$. Under depolarizing noise, the effective weight of a Pauli operator reduces to the Hamming weight and the effective code distance reduces to the code distance $d$. Under infinitely biased noise (pure $\sigma$ noise), the effective weight simply counts $\sigma$ as 1 and other Pauli operators as $\infty$, and the effective distance $d_\sigma$ of a code is the minimum Hamming weight of the logical operators consisting of only $\sigma$ and identity. Without loss of generality, in the rest of the paper, we will consider noise biased towards Pauli $Z$ errors, i.e. $p_Z \geq p_X, p_Y$, unless specially noted.

The introduction of the effective code distance greatly clarifies our task. We simply aim to find codes that can reach a large effective code distance using a small number of physical qubits. Given a family of codes, we optimize them by finding the codes that reach a given effective distance $d'$ with a minimum code size $n$, or equivalently the codes that can reach the largest effective distance $d'$ given a code size $n$.

### *3.2.2 The XZZX codes*

In this section, we present the details of the XZZX code family that we consider in this work. We start from a set of cyclic codes generalized from the five-qubit code and show that they are qubit-efficient against infinitely biased Pauli noise due to their repetition-code

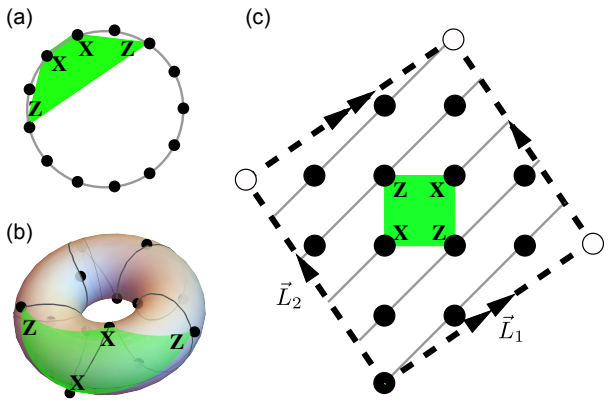


Figure 3.2: Map from the $\mathcal{S}(13,2,1)$ code (a) to a GTC with $\vec{L}_1 = (3,2), \vec{L}_2 = (-2,3)$ (b)(c). The qubits are represented by black dots and labeled along the solid grey string. The stabilizer generators are represented by green plaquettes (only one generator is plotted and others are obtained by shifting along the grey string). (b) and (c) are obtained by wrapping the grey string around the torus. (c) is the 2D layout of (b), with opposite sides of the parallelogram enclosed by $\vec{L}_1$, $\vec{L}_2$ identified.

structure. We then map the cyclic codes to a larger family of topological codes, GTCs, to study their performance under finite-bias noise. Using topological tools, we show how to efficiently derive the effective code distance for the GTCs under an arbitrary noise bias.

## The XZZX cyclic codes

The five-qubit code, with stabilizers generated by cyclic permutations of $ZXXZI$, is the smallest code with distance $d = 3$. Moreover, it is also the smallest code with $d_Z = 5$, where $d_Z$ denotes the effective distance under pure Pauli $Z$ noise since it exhibits a repetition-code structure for pure Pauli $Z$ noise. Therefore, it is a resource-efficient code for both depolarizing and infinitely biased Pauli $Z$ noise. To find larger codes that are efficient under a wide range of noise bias, we can generalize the five-qubit code and consider a family of XZZX cyclic codes with stabilizer groups in the form

$$\mathcal{S}(n,a,b) = \langle Z_i X_{i\oplus a} X_{i\oplus a\oplus b} Z_{i\oplus 2a\oplus b} | \forall i \in \mathbb{Z}/n\mathbb{Z} \rangle, \tag{3.4}$$

where $n$ is the total number of qubits, $a$ and $b$ are positive integers, and $\oplus$ denotes addition modulo $n$. Each weight-four stabilizer generator is of XZZX type, with $a-1$ identities inserted between $Z$ and $X$ and $b-1$ identities inserted between two $X$s. We refer to the XZZX cyclic code defined in Eq. (3.4) as $\mathcal{S}(n,a,b)$. We note that $\mathcal{S}(7,1,3)$ has been considered in Ref. [157] and shown to have a good performance against $Z$-biased noise.

For $Z$-biased noise, we can introduce a parameter $\eta = p_Z/(p_X+p_Y)$, which ranges from $\frac{1}{2}$ to infinity, to characterize the noise bias. We aim to find codes that are efficient over a wide range of biases $\eta$. We attempt this by finding codes that can reach large effective code distance in the two extreme cases — under depolarizing noise ($\eta = \frac{1}{2}$) and pure Pauli $Z$ noise ($\eta = \infty$) — using only a small number of qubits. We may directly generalize the five-qubit code by keeping $a = b = 1$ and increasing $n$. However, in this way $d_Z$ increases while $d$ is fixed, e.g. $\mathcal{S}(13,1,1)$ has $d_Z = 13$ and $d = 3$. It turns out that to simultaneously increase $d_Z$ and $d$ we need to also modify the stabilizer structure, i.e. to change $a, b$. For example, the $\mathcal{S}(13,2,1)$ code has $d_Z = 13$ and $d = 5$. In general, it is easy to identify codes that have the maximal effective distance $d_Z = n$ (using $n$ qubits) under pure Pauli $Z$ noise since any code defined in Eq. (3.4) with $b$ and $n$ being coprime has a repetition-code structure by neglecting the $Z$ components in the stabilizers. However, it is nontrivial to identify codes that are also efficient against depolarizing or finite-bias noise. To accomplish this, we can wrap the cyclic codes on a torus and map them to a family of generalized toric codes (GTCs) introduced in Ref. [162]. When discussing GTCs in the following, we will use notations consistent with earlier literatures [162, 164].

## From the XZZX cyclic codes to the XZZX generalized toric codes

An XZZX generalized toric code GTC($\vec{L}_1, \vec{L}_2$) is a stabilizer code with qubits on a square lattice $\mathbb{Z}^2$ and stabilizers generators $\{S_{i,j} \equiv X_{i,j}Z_{i+1,j}Z_{i,j+1}X_{i+1,j+1} | i,j \in \mathbb{Z}\}$, with boundary conditions specified by the two basis periodicity vectors $\vec{L}_1, \vec{L}_2 \in \mathbb{Z}^2$: two points $\vec{u}, \vec{v} \in \mathbb{Z}^2$

are identified iff:

$$\vec{u} - \vec{v} \in \text{span}(\vec{L}_1, \vec{L}_2) := \{m_1\vec{L}_1 + m_2\vec{L}_2 \mid m_1, m_2 \in \mathbb{Z}\}. \tag{3.5}$$

GTC$(\vec{L}_1, \vec{L}_2)$ encodes $k$ logical qubits in $n$ physical qubits where $n = |\vec{L}_1 \times \vec{L}_2|$. If both periodicity vectors have even 1-norm, then $k = 2$. Otherwise, $k = 1$ [164, 169].

A GTC can be viewed as stabilizer codes defined on a graph $G(\vec{L}_1, \vec{L}_2)$ embedded on a torus [1], with qubits on vertices and stabilizers on plaquettes. The infinite square lattice $\mathbb{Z}^2$ acts as the universal cover of $G(\vec{L}_1, \vec{L}_2)$, with the covering map given by the boundary condition Eq. (3.5). A code is uniquely specified by the submodule of $\mathbb{Z}^2$ given by span$(\vec{L}_1, \vec{L}_2)$, the span of the two basis vectors $\vec{L}_1, \vec{L}_2$. Different choices of basis periodicity vectors give the same GTC so long as they are related by a unimodular transformation. A single Pauli $Z_{i,j}$ ($X_{i,j}$) anti commutes with two stabilizer generators that lie along the diagonal: $\{S_{i-1,j-1}, S_{i,j}\}$ ($\{S_{i-1,j}, S_{i,j-1}\}$). To facilitate the analysis, we define the diagonal axes, which we call the "XZ" axes, to be the axes corresponding to the $X$ and $Z$ error chains with respective basis vectors $\hat{x} := (-1, 1), \hat{z} := (1, 1)$. We note that the GTCs encoding two logical qubits, which can be obtained from the CSS toric codes [35] by applying local Hadamard transformations and twisting the boundary conditions, are considered for biased noise in Ref. [93]. In this work, however, we will focus on the GTCs encoding 1 logical qubit since they can reduce the required code size by roughly a factor of 2 for reaching a target effective code distance compared to their 2-logical-qubit counterparts (which will become clear later). The rectangular-lattice toric codes considered in Ref. [92] with $\vec{L}_1 = (d-1, 0), \vec{L}_2 = (0, d)$ belong to the GTCs encoding 1 logical qubit. However, Ref. [92] only considers this special instance and has a limited discussion on its performance when

1. An embedding of a graph $G(V, E)$ with vertices $V$ and edges $E$ on a manifold $\mathcal{M}$ is a map $\Gamma : V \cup E \to \mathcal{M}$. With the embedding, we can define the plaquettes (or faces) of the embedded graph as $F = \mathcal{M} \backslash \Gamma(E)$. See Ref. [164] for details. We refer to the graphs associated with the GTCs embedded graphs (on the torus) with well-defined vertices, edges and plaquettes, unless specially noted.

the bias is finite. In this work, we will systematically investigate the performance of the 1-logical-qubit GTCs by studying their effective distance and adaptively find the optimal codes given any noise bias.

The XZZX cyclic codes can be mapped to a subset of GTCs by wrapping the qubits around the torus along a certain direction. As an example, we show how the $\mathcal{S}(13,2,1)$ code can be mapped to the GTC with $\vec{L}_1 = (3,2), \vec{L}_2 = (-2,3)$ in Fig. 3.2. We also explicitly provide more general mappings from the XZZX cyclic codes to the GTCs in the following.

Given a XZZX cyclic code $\mathcal{S}(n,a,b)$, we try to find the corresponding GTC($\vec{L}_1, \vec{L}_2$) with $\vec{L}_1, \vec{L}_2$ determined by the parameters $n, a, b$. For a GTC that can be mapped to a cyclic code, it has to be cyclic along a certain direction $\hat{l}_1$, along which qubits are labeled. Suppose that this is the case and the two periodicity vectors of the GTC are given by $\vec{L}_1 = n\hat{l}_1, \vec{L}_2 = \hat{l}_2$, where $\hat{l}_1, \hat{l}_2$ are non-parallel vectors with coprime coordinates that satisfy $|\hat{l}_1 \times \hat{l}_2| = 1$. Without loss of generality, we consider the case when a stabilizer $Z_{-a}X_0X_bZ_{a+b}$ (all the indices are modulo $n$) of the $\mathcal{S}(n,a,b)$ is mapped to a ZXXZ stabilizer of the GTC supported on four qubits with coordinates $(0,1),(0,0),(1,1)$ and $(1,0)$, respectively. Since the qubits are labeled along the $\hat{l}_1$ direction, the $Z_{-a}X_0X_bZ_{a+b}$ stabilizer should also be supported on qubits with coordinates $-a\hat{l}_1, (0,0), b\hat{l}_1$, and $(a+b)\hat{l}_1$, respectively. Hence, the points $-a\hat{l}_1$ and $(0,1)$ should be identified (by the boundary conditions specified by $\vec{L}_1$ and $\vec{L}_2$), and the same goes for the points $b\hat{l}_1$ and $(1,1)$. Considering the above conditions, the map from $n, a, b$ to $\hat{l}_1, \hat{l}_2$ is then given by the following constrained equations:

$$-a\hat{l}_1 \sim (0,1), \quad b\hat{l}_1 \sim (1,1), \quad |\hat{l}_1 \times \hat{l}_2| = 1 \tag{3.6}$$

where $\vec{A} \sim \vec{B}$ means that $\vec{A}$ and $\vec{B}$ are identified on the torus defined by $\vec{L}_1, \vec{L}_2$, i.e. $\vec{A} - \vec{B} = m_1\vec{L}_1 + m_2\vec{L}_2$ for $m_1, m_2 \in \mathbb{Z}$. We conjecture that the solution to the above nonlinear Diophantine equations always exists. Instead of proving the conjecture in general, we construct the explicit solution in some special cases below.

1. If $\gcd(n,b) \mid (a+1)$ ($\gcd(n,b)$ devides $(a+1)$), we can choose $\hat{l}_1 = (m, m+1), \hat{l}_2 = b\hat{l_1} - (1,1)$, where $m$ is the solution of $bm - nk = a+1$ with variables $m, k$.

2. If $\gcd(n,a) \mid (b+1)$, we can choose $\hat{l}_1 = (1,m), \hat{l}_2 = -a\hat{l_1} - (0,1)$, where $m$ is the solution of $am - nk = b+1$ with variables $m, k$.

An alternative formulation of the mapping is to instead focus on the labeling of qubits. Given a cyclic code $\mathcal{S}(n,a,b)$, we can label qubits associated with a plaquette stabilizer in the following way. The qubits at $(x,y), (x,y+1), (x+1,y)$ and $(x+1,y+1)$ are labeled $i, i-a, i+a+b$ and $i+a$ ( mod $n$), respectively. We note that the label increases by $-a$ ( mod $n$) by moving one step vertically on the lattice, and the label increases by $a+b$ by moving one step horizontally. The solutions $\alpha, \beta \in \mathbb{Z}$ admitted by the modular equation

$$\Big[(a+b)\alpha - a\beta\Big] \mod n = 0 \tag{3.7}$$

form all possible periodicity vectors $\vec{L} = (\alpha, \beta) \in \text{span}(\vec{L}_1, \vec{L}_2)$ of the GTC corresponding to $\mathcal{S}(n,a,b)$.

As pointed out earlier, the map from the XZZX cyclic codes to the XZZX GTCs is injective. We verify this by showing that only a subset of GTCs is cyclic. A GTC is cyclic if and only if there exists a direction along which the GTC is cyclic, which gives the following condition for a GTC being cyclic:

**Proposition 10** (Cyclic condition for the GTCs)**.** *A* GTC$(\vec{L}_1, \vec{L}_2)$ *is cyclic if and only if there exists* $\vec{L} = (L_x, L_y) \in$ *span*$(\vec{L}_1, \vec{L}_2)$ *such that* $\gcd(L_x, L_y) = 1$.

*Proof.* We first prove the sufficiency. We say that a vector $(x,y) \in \mathbb{Z}^2$ is a coprime vector if $x$ and $y$ are coprime. If a coprime periodicity vector $\vec{L}$ exists, we first find a direction $\hat{l}_0$ (which is a coprime vector) that satisfies $|\hat{l}_0 \times \vec{L}| = 1$ and then let $L_1' = n\hat{l}_0, \vec{L}_2' = \vec{L}$. Then the code GTC$(\vec{L}_1', \vec{L}_2')$ defines the same GTC, which is cyclic along the $\hat{l}_0$ direction. We prove the

necessity by showing a contradiction. Suppose that a GTC($\vec{L}_1, \vec{L}_2$) is cyclic, but does not have coprime periodicity vectors. The condition for the GTC being cyclic implies that there exists $\vec{L}'_1, \vec{L}'_2 \in \text{span}(\vec{L}_1, \vec{L}_2)$ that satisfies $\vec{L}'_1 = n\hat{l}_0, |\vec{L}'_2 \times \hat{l}_0| = 1$ ($\hat{l}_0$ is a coprime vector). Then $\vec{L}'_2$ has to be coprime since $|\vec{L}'_2 \times \hat{l}_0| = 1$, which contradicts the assumption. □

Obviously, there are codes that do not satisfy the condition 10, thus not cyclic. For example, the GTC($(0, p), (p, 0)$) does not have any coprime lattice vectors for $p \geq 2$ and as a result, these square-lattice toric codes are not cyclic.

We note that the GTCs are a larger family of codes that also include non-cyclic codes. By mapping the cyclic codes to the GTCs, we benefit from the following two aspects:

1. We can use topological tools to efficiently obtain the effective code distance given a finite noise bias, which enables us to adaptively design the optimal codes;

2. We can design efficient decoders that can lead to similarly high thresholds as those for the XZZX surface codes [92].

## Deriving the effective code distance for the GTCs

Calculating the effective code distance is likely to be computationally intractable in general since even computing the distance of a classical linear code is NP-hard. However, logical operators of topological codes embedded in a manifold are easily identified with geometrical objects on the manifold, and therefore the effective code distance of the GTCs can be efficiently derived using topological tools. Recall that a GTC($\vec{L}_1, \vec{L}_2$) is defined on an embedded graph $G(\vec{L}_1, \vec{L}_2)$ on a torus. We first consider the case when the plaquettes of $G$ are two-colorable, i.e. one can consistently two color the plaquettes such that two plaquettes sharing the same edge have different colors. In this case, we can transform $G$ to a graph $G'$ in which qubits are associated with edges while stabilizers are associated with plaquettes and vertices (e.g. from Fig. 3.3(b) to Fig. 3.3(c)). Now $G'$ is the same as the Kitaev's construction [35]

and mathematically $G$ is the medial graph of $G'$. As a result, the GTCs associated with $G'$ can be described by the standard 2 chain complex for CSS toric codes [170], and the logical operators are associated with homologically nontrivial cycles on the torus. See Appendix C of Ref. [68] for more details. In fact, the two-colorable GTCs are equivalent to the CSS toric codes by local Hadamard transformation. The distance of these two-colorable codes can then be obtained by estimating the shortest length of the nontrivial cycles, and the effective distance under an asymmetric noise, as we will show later, equals the shortest length of the nontrivial cycles under a noise-modified distance metrics.

However, the GTCs of interest that encode 1 logical qubit, are defined by embedded graphs $G$ that are not two-colorable, when at least one of $\vec{L}_1, \vec{L}_2$ is odd in 1-norm. In this scenario, the graph $G$ cannot be consistently two-colored and as a consequence, it can not be directly transformed to a graph $G'$ corresponding to a 2 chain complex. Fortunately, we can still use the algebraic tools by constructing the "doubled" graph $G_d(\vec{L}_{1,d}, \vec{L}_{2,d})$ [164]: Without loss of generality, we assume $\vec{L}_1$ is even while $\vec{L}_2$ is odd in 1-norm [2]. We then obtain the doubled graph $G_d$ by combining two copies of $G$ together along $\vec{L}_1$. Topologically, this corresponds to taking two tori, cutting them open along the $\vec{L}_1$ cycle and gluing them together. As such, $G_d$ is embedded on the doubled torus, which is specified by two doubled periodicity vectors $\vec{L}_{1,d}, \vec{L}_{2,d}$ (similar as that the original torus is specified by the periodicity vectors $\vec{L}_1, \vec{L}_2$ via Eq. (3.5)) that are given by the following map: $\vec{L}_{1,d} = \vec{L}_1, \vec{L}_{2,d} = 2\vec{L}_2$. As an example, we show how we construct the doubled graph for a GTC with $\vec{L}_1 = (0, 2), \vec{L}_2 = (3, 0)$ in Fig. 3.3 (from (a) to (b)).

The doubled graph $G_d$ is now two colorable. We can then transform $G_d$ to a graph $G'_d$ with the original vertices in $G_d$ on the edges of $G'_d$. The black plaquettes in $G_d$ are transformed into vertices in $G'_d$. See the transformation from Fig. 3.3(b) to (c). We now associate the edges in $G'_d$ with Pauli operators up to phases and vertices/plaquettes with

2. If both $L_1$ and $L_2$ are odd in 1-norm, we can choose $L'_1 = L_1 + L_2, L'_2 = L_2$ that defines the same code. Now only $L'_2$ is odd in 1-norm

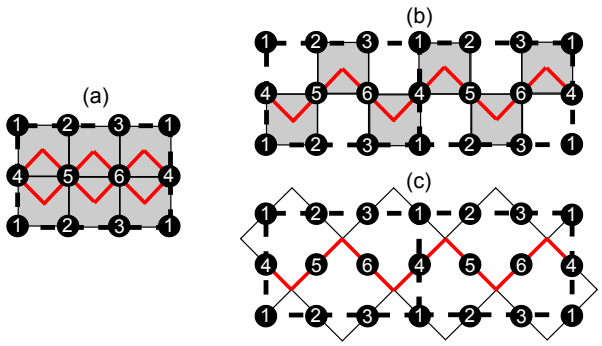


Figure 3.3: The construction of doubled graph for a non-two-colorable GTC with $\vec{L}_1 = (0,2), \vec{L}_2 = (3,0)$. (a) The original non-two-colorable graph $G$ that defines the GTC. There is a XZZX stabilizer on each of the shaded plaquettes. The red cycle depicts a logical operator $Y_4Y_5Y_6$ that cannot wrap back on itself after a single loop due to odd periodicity in the horizontal direction. (b) The doubled graph $G_d$ obtained by taking two copies of $G$ and gluing them horizontally. Now $G_d$ becomes two colorable and the stabilizers are only put on the shaded plaquettes. Consequently, a single loop is sufficient for the logical operator to wrap back on itself. (c) The equivalent graph $G'_d$ of $G_d$ where qubits are placed on the edges and stabilizers are placed on the vertices. Now the logical operator corresponds to a well-defined cycle that is homologically nontrivial on the doubled torus.

stabilizer generators. The idea of the doubling was introduced and analyzed in pure graph-based formalism in Ref. [164]. Here we formalize the codes associated with the doubled graph from the perspective of algebraic topology. As detailed in Appendix **??**, the representatives of logical operators $Z_L, X_L, Y_L$ are associated with the nontrivial elements in the first homology group, or equivalently, three homologically nontrivial loops, that are defined on the *doubled* torus. We note that because of the doubling, the non-two-colorable GTCs can effectively reduce the code size required for reaching a certain effective distance by a factor of 2 compared to their two-colorable counterparts.

With the above topological construction, we can then use geometrical method to calculate the effective code distance $d'$. This can be readily calculated when $X$ and $Z$ errors are independent. Therefore, we first consider independent Pauli X and Z noise, in which the probability distribution is given by $p_X = p_Z^\omega$, $p_Y = p_X p_Z = p_Z^{\omega+1}$ and $p = p_Z + p_X + p_Y$, where we assume $\omega \geq 1$. We note that here we use a bias parameter $\omega$ that is different from the parameter $\eta \equiv p_Z/(p_X + p_Y)$ commonly used in the literature [89, 90, 91, 92]. For

independent Pauli $X$ and $Z$ noise, we can convert $\eta$ to $\omega$ by $\omega = \frac{\log \eta}{\log 1/p_Z} + \frac{\log(1+1/p_Z)}{\log 1/p_Z}$, where $\omega$ depends on both $\eta$ and the error probability $p_Z$. Under such a noise model, the modified weights of Paulis are $\mathrm{wt}'(Z) = 1$, $\mathrm{wt}'(X) = \omega$ and $\mathrm{wt}'(Y) = \omega + 1$. Then, the effective code distance is given by the length of shortest homologically nontrivial cycle on the doubled torus with distance metrics being the rescaled 1-norm (in the XZ axes):

$$d' = \min_{m_1, m_2 \in \mathbb{Z}} \left\| m_1 \vec{L}_{1,d} + m_2 \vec{L}_{2,d} \right\|'_{xz,1}, \tag{3.8}$$

where the rescaled 1-norm of a vector $\alpha\hat{x} + \beta\hat{z}$ is given by $\|\alpha\hat{x} + \beta\hat{z}\|'_{xz,1} \equiv \omega|\alpha| + |\beta|$. We present an efficient algorithm with complexity $O(d'^2)$ to compute the effective distance $d'$ in Alg. 5.

**Algorithm 5:** Algorithm for calculating $d'$

**Input :** $\omega, \vec{L}_{1,d}, \vec{L}_{2,d}, \epsilon$
**Output:** $d'$

$\vec{L}'_{i,d} \leftarrow \vec{L}_{i,d} \begin{bmatrix} \omega & 0 \\ 0 & 1 \end{bmatrix}$ for $i = 1, 2$, $L \leftarrow (\vec{L}'^T_{1,d}, \vec{L}'^T_{2,d})^T$ `// ` $\vec{L}_{i,d}$ `are row vectors,` $L$ `is a` $2 \times 2$ `matrix`
$r \leftarrow 0, s \leftarrow 0$
**while** $s = 0$ **do**
    $r \leftarrow r + \frac{\epsilon}{2}$
    **for** *each* $k \in [0, \frac{2r}{\epsilon}]$ **do**
        $x, y \leftarrow -r + k\frac{\epsilon}{2}, k\frac{\epsilon}{2}$
        **if** $\|L \left[ L^{-1}(x, y)^T \right] - (x, y)^T\|_1 \leq \epsilon$ **then**
            `// [·] rounds the entries of a vector to integers.`
            $s \leftarrow 1$
        $x, y \leftarrow r - k\frac{\epsilon}{2}, k\frac{\epsilon}{2}$
        **if** $\|L \left[ L^{-1}(x, y)^T \right] - (x, y)^T\|_1 \leq \epsilon$ **then**
            $s \leftarrow 1$
$d' \leftarrow r$

It is worth noting that the choice of the shortest nontrivial cycle, which corresponds to

the logical operator with minimum effective weight, depends on the noise bias $\omega$. Moreover, the effective distance $d'$ for a given GTC also varies with $\omega$ and typically increases with $\omega$. As an example, we show how we can geometrically find the minimum-effective-weight logical operators and obtain the effective code distance for the [[13,1,5]] GTC under different bias in Fig. 3.4.

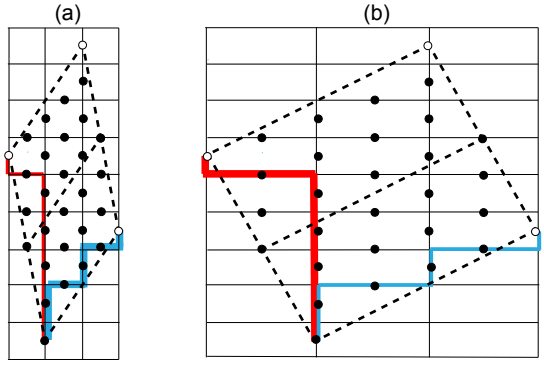


Figure 3.4: Geometric representation of the logical operators of the $[[13, 1, 5]]$ GTC with $\vec{L}_1 = (-1, 5) = 2\hat{x} + 3\hat{z}$, $\vec{L}_2 = (-3, 2) = -\frac{1}{2}\hat{x} + \frac{5}{2}\hat{z}$ for (a) $\omega = 1$ and (b) $\omega = 3$. The horizontal and vertical axes are aligned with the $X$ and $Z$ axes, respectively. The doubled graph is obtained by doubling along $\vec{L}_2$: $\vec{L}_{1,d} = \vec{L}_1, \vec{L}_{2,d} = 2\vec{L}_2 = -\hat{x} + 5\hat{z}$. The blue (red) cycle represents the logical operator associated with $\vec{L}_{1,d}$ ($\vec{L}_{2,d}$). The shortest nontrivial cycles $\vec{L}_m$ in the modified 1-norm determined by Eq. (3.8) are thickened. For (a) $\omega = 1$, $\vec{L}_m = \vec{L}_{1,d}$ and $d' = \|\vec{L}_{1,d}\|_{xz,1} = 5$; For (b) $\omega = 3$, $\vec{L}_m = \vec{L}_{2,d}$ and $d' = \|\vec{L}_{2,d}\|_{xz,1} = 8$.

To verify that the effective code distance estimated via Eq. (3.8) is indeed a good proxy for the code performance under the independent XZ noise model, we perform the Monte Carlo (MC) simulation using a MWPM decoder (see Appendix E of Ref. [171]) and fit the logical error rate by $p_L \propto p^r$. We compare the numerically fitted exponent $r$ with $\lfloor (d'+1)/2 \rfloor$, where $\lfloor \cdot \rfloor$ represents the floor operation, in Fig. 3.5.

We can see that for most of the codes $r$ agrees well with $\lfloor (d'+1)/2 \rfloor$ within the numerical uncertainty. The systematical deviation for some codes under $\omega = 3$ occurs because when $\omega > 1$, different Pauli operators have different effective weights and the effective weight of the most probable uncorrectable error associated with a logical operator with effective weight $w'$ is not necessarily (and in fact, only lowered bounded by) $\lfloor (w'+1)/2 \rfloor$. As a result, $r$ is,

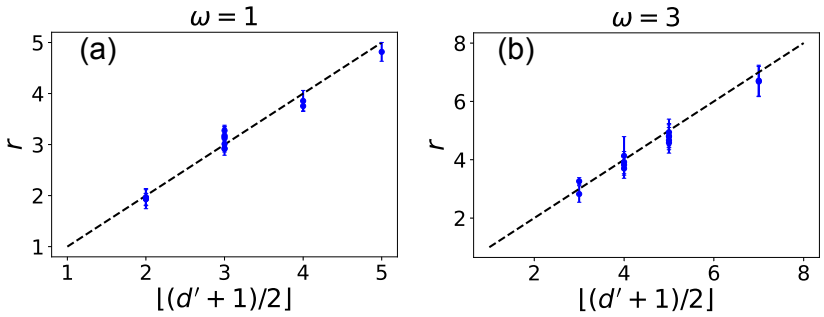


Figure 3.5: Verification of the effective code distance Eq. (3.8) as a good performance metrics for the GTCs. Given a bias parameter $\omega$, we numerically obtain the logical error rate $p_L$ for a set of randomly chosen GTCs using the MWPM decoders over a range of physical error rate $p$ (below the threshold), and fit the logical error rate by $p_L \propto p^r$. There is a good agreement between $r$ and $\lfloor (d'+1)/2 \rfloor$ for both (a) $\omega = 1$ and (b) $\omega = 3$. The range of the physical error rate $p$ used for the fitting is: (a) $p \in [0.02, 0.08]$ and (b) $p \in [0.06, 0.1]$.

in general, only lower bounded by $\lfloor (d'+1)/2 \rfloor$. In the following, we provide an improved approximation for $r$.

To more accurately characterizes how the logical error rate scales with the (total) physical error rate, we define the following effective half code distance: For an asymmetric Pauli channel with probability distribution $\{p_\sigma\}, \sigma \in \{X, Y, Z\}$ and total error probability $p$, the effective half distance $r'$ of a code is the minimum modified weight of any uncorrectable errors, with the noise-modified weight of a Pauli $\sigma$ given by: $\mathrm{wt}'(\sigma) \equiv \log p_\sigma / (\max_\sigma \log p_\sigma)$. With the above defined $r'$, the logical error rate (to the leading order) scales as $p_L \propto p^{r'}$. Note that for the depolarizing noise $r'$ is simply given by $r' = \lfloor (d'+1)/2 \rfloor$, which can be efficiently calculated by if $d'$ is known. However, for an asymmetrical Pauli channel, $r'$ not necessarily equals $\lfloor (d'+1)/2 \rfloor$ and can not be efficiently calculated in general. Instead of approximating $r'$ as $\lfloor (d'+1)/2 \rfloor$ in Fig. 3.5, which fails in some cases, we can adopt a better approximation of $r'$: (1) Find the logical operator $L_m$ with the minimum effective weight. (2) Approximate $r'$ as $r' \approx \min_{E \subset L_m} \mathrm{wt}'(E)$, where the minimization is over all the subsets of the logical operator. The above calculation can be done efficiently provided that $L_m$ can be efficiently located (or equivalently, $d'$ can be calculated efficiently).

### *3.2.3 Qubit-efficient fault-tolerant quantum error correction with the XZZX codes*

In this section, we investigate how the XZZX codes can be used for qubit-efficient fault-tolerant QEC. We numerically extract both the capacity thresholds and the phenomenological thresholds of the codes and show that the thresholds can be significantly enhanced by increasing the noise bias. In the sub-threshold regime, we present an adaptive design of the codes that minimizes the required qubit number for reaching a target effective distance given any noise bias. Furthermore, we propose a QEC scheme using flag qubits that can fault-tolerantly operate the XZZX codes under a practical circuit-level noise.

#### Code thresholds

The topological construction of the GTCs indicates that the GTCs could potentially have exceptionally high thresholds, which might be further boosted by having large bias. First we note that the GTCs are locally equivalent to the XZZX surface codes and differ mainly by boundary conditions. As a result, the code-capacity threshold, an asymptotic quantity for asymptotically-large code blocks, is the same for two code families if optimal decoders are applied. Next, we show that the GTCs have similarly high thresholds using our tailored efficient MWPM decoders.

We adopt the independent $X$ and $Z$ noise model and numerically extract the thresholds $p_c$ of GTCs using MWPW decoders for different bias parameter $\omega$. For each $\omega$, we extract the logical error rate $p_L$ as a function of physical error rate $p$ for a set of close-to-optimal codes (defined in Sec. 3.2.3) with $d' \leq 21$ from the Monte Carlo simulations. Then we estimate the threshold $p_c$ by fitting to the critical-exponent expression $p_L = A + Bx + Cx^2$, where $x := (p - p_c)d'^{1/\nu}$ [9]. In Fig. 3.6(a), we plot the thresholds of the GTCs when assuming perfect syndrome extractions. The thresholds match or even surpass the hashing bound (indicated by the dashed line), similarly as the XZZX surface codes. In Fig. 3.6(b), we plot

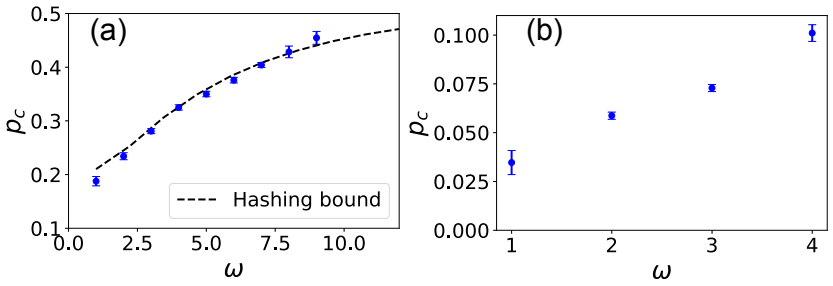


Figure 3.6: The thresholds $p_c$ of GTCs using MWPM decoders as functions of the bias parameter $\omega$, obtained by doing the critical exponent fit [9] on numerical data from MC simulations. (a) Code threshold assuming perfect syndrome measurements. The dashed line indicates the hashing bound. (b) Code threshold under a phenomenological error model, in which each syndrome measurement fails with a probability that equals the total error probability $p$ of the data qubits.

the thresholds under a phenomenological noise model, in which each syndrome measurement fails with a probability that equals to the total error probability $p$ of the data qubits. The phenomonological thresholds increase from 3.5% to 10% as $\omega$ increases from 1 to 4.

## Adaptive code design for qubit-efficient QEC

Under infinitely $Z$-biased noise, the optimal GTCs (encoding one logical qubit) should correspond to the cyclic codes with a repetition structure, whose effective distance reaches the optimal linear scaling with $n$, i.e. $d_Z = n$. Here we explicitly identify these cyclic GTCs. A GTC is cyclic if there is a cycle of length $n$ along certain direction $\hat{l}_0 \in \mathbb{Z}^2$ with $a, b$ being coprime, on the torus that goes through all the qubits without repetition, e.g. the grey string in Fig. 3.2. The qubits are then labeled along this cycle to be a XZZX cyclic code. Given such a direction $\hat{l}_0$, we say that the GTC is cyclic along $\hat{l}_0$. We then identify the GTCs that correspond to the XZZX cyclic codes with repetition structures by identifying the direction along which they are cyclic:

**Proposition 11.** *A GTC has a repetition-Z (X) structure iff it is cyclic along (1,1) ((-1,1)) direction; A GTC has a repetition-Y structure iff it is cyclic along (0,1) and (1,0) direction.*

*Proof.* The proof for repetition-$Z$ ($X$) structure is straightforward. For infinite $Z$ ($X$) noise, the GTCs are effectively single or disjoint sets of repetition codes obtained by removing the Pauli $Zs$ ($Xs$) in the stabilizers. A GTC is a single repetition code iff there are no logical operators consisting of only Pauli $X$s ($Z$s) that have weight smaller than $n$. Since the Pauli $Z$ ($X$) chains lie along $(1,1)$ ($(-1,1)$) direction, the above condition is equivalent to that there are no sub-cycles along the $(1,1)$ ($(-1,1)$) direction. Next we prove that it is necessary for a GTC to be cyclic along $(1,0)$ and $(0,1)$ direction in order to have a repetition-$Y$ structure. Suppose the code is not cyclic along either $(1,0)$ or $(0,1)$ direction, i.e. there are sub-cycles along that direction, then a Pauli-$Y$ string associated with a sub-cycle is a logical operators (with weight smaller than $n$), contradicting the assumption of the repetition-$Y$ structure. To prove the sufficiency, we first define a classical "Y-code" with parity-check matrix $H$, where each row of $H$ is associated with a stabilizer generator and $H_{i,j} = 1$ iff the action of $S_i$ on the $j$th qubit is non-identity. A GTC under pure $Y$ noise is then decoded as a Y code. Given the condition that a GTC is cyclic along both $(1,0)$ and $(0,1)$ direction, without loss of generality we can choose $\vec{L}_1, \vec{L}_2$ as $\vec{L}_1 = n(1,0), \vec{L}_2 = (-m,1)$, such that $\gcd(m,n) = 1$ (to ensure the code is cyclic along (0,1) direction). We then label the qubits along the $(1,0)$ direction and correspondingly, the $i$-th row of $H$ is: $H_{i,j} = 1$ for $j = i, i+1, i+m, i+1+m($ mod $n)$ and 0 otherwise. This is equivalent to (up to regrouping the stabilizer generators) the repetition code with $H_{i,j} = 1$ for $j = i, i+1$ since $m$ and $n$ are coprime. □

In the finite-bias regime, we can use our topological construction and geometrical methods to adaptively identify the optimal GTCs given any bias parameter $\omega$. We restrict ourselves to non-two-colorable GTCs since they are more resource efficient. The task is now to find the GTC that uses the smallest number of physical qubits $n$ to reach a given effective code distance $d'$. Given the geometrical interpretation of $d'$ and $n$:

$$d' = \min_{m_1,m_2\in\mathbb{Z}} \left\| m_1\vec{L}_{1,d} + m_2\vec{L}_{2,d} \right\|'_{xz,1},$$

and $n = \frac{1}{2}\left|\vec{L}_{1,d} \times \vec{L}_{2,d}\right|$, this task is equivalent to finding the densest packing of diamonds whose aspect ratio is given by the bias parameter $\omega$.

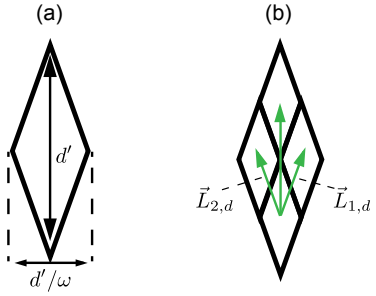


Figure 3.7: The densest packing of diamonds that correspond to the optimal choice of the GTCs given a bias parameter $\omega$. The diamonds to pack (a) have aspect ration $\omega$ and the densest packing pattern is the regular tiling (b).

As shown in Fig. 3.7(a), the diamond to pack has diagonals with length $d'$ and $d'/\omega$, respectively. Obviously, the densest packing pattern is the regular tiling of a surface using the diamonds (see Fig. 3.7(b)). Therefore, the optimal choice of $\vec{L}_{1,d}, \vec{L}_{2,d}$ is:

$$\vec{L}_{1,d}^{\mathrm{OPT}} = \tfrac{1}{2}d'\left(\tfrac{1}{\omega}\hat{x} + \hat{z}\right), \quad \vec{L}_{2,d}^{\mathrm{OPT}} = \tfrac{1}{2}d'\left(-\tfrac{1}{\omega}\hat{x} + \hat{z}\right). \tag{3.9}$$

Eq. (3.9) gives $n = d'^2/2\omega$, or $d' = \sqrt{2\omega n}$. However, the densest packing pattern is not always achievable since the codes have an additional constraint that $d' \leq n$, as there is always a logical operator consisting of all Pauli Zs of effective weight $n$. Hence, the optimal codes satisfy:

$$n = \begin{cases} d' & d' \leq 2\omega \\ d'^2/2\omega & d' > 2\omega \end{cases}, \tag{3.10}$$

or equivalently, $d' = \min(n, \sqrt{2\omega n})$. Eq. (3.10) should be viewed as an (tight) lower bound on the required size of our XZZX codes family for reaching a target effective distance (see the black dashed curve in Fig. 3.8). Furthermore, Eq. (3.10) provides the guiding principles for choosing the optimal codes: Given a noise bias $\omega$ and a target effective distance $d'$, the cyclic

GTCs with a repetition-code structure are optimal for $d' \leq 2\omega$; The GTCs with an optimized layout (optimal choice of $\vec{L}_1, \vec{L}_2$), which corresponds to the densest packing pattern of the associated diamonds, are optimal for $d' > 2\omega$.

However, due to the constraints that the periodicity vectors $\vec{L}_1, \vec{L}_2$ should be in $\mathbb{Z}^2$ and they produce a non-two-colorable graph, the optimal codes that saturate the upper bound are sparse in both $d'$ and $n$ (see the black dots in Fig. 3.8). To obtain more codes with good performance, we define the following close-to-optimal (CTO) codes by finding the codes that correspond to close-to-optimal packing patterns:

$$\vec{L}_{i,d}^{\mathrm{CTO}} = \vec{L}_{i,d}^{\mathrm{OPT}} + \delta l_i, \quad \|\delta l_i\|_{xz,1} \leq \Delta, \tag{3.11}$$

for $i = 1, 2$. Here $\Delta$ is a parameter that characterizes how far the packing pattern given by $\vec{L}_{i,d}^{\mathrm{CTO}}$ is deviated from the densest packing pattern (given by $\vec{L}_{i,d}^{\mathrm{OPT}}$). By increasing $\Delta$ we further relax the packing pattern and obtain more CTO codes. We note that due to the integer constraints on the periodicity vectors, we need to restrict $\omega$ to be integers in order to obtain optimal codes that exactly satisfy Eq. 3.9. However, we can still find a set of close-to-optimal codes that approximately satisfy Eq. 3.9 for any $\omega$. This could be done by either rounding $\omega$ to integers or finding periodicity vectors in $\mathbb{Z}^2$ that are close to the exact solutions of Eq. 3.9 within certain Manhattan distance (similar to Eq. 3.11).

In Fig. 3.8, we plot the required size $n$ of different codes as a function of the target effective distance $d'$ for $\omega = 1$ and 3. We compare the tailored GTCs with the unrotated planar XZZX surface code [92], whose aspect ratio is optimized according to $\omega$. For $d' \leq 2\omega$, the tailored GTCs enjoys the optimal scaling between $n$ and $d'$, i.e. $n = d'$, because some GTCs can have a repetition-code structure while the planar codes cannot; For $d' > 2\omega$, $n$ scales quadratically with $d'$ for both code families, but the planar codes require roughly four times more qubits than the tailored GTCs.

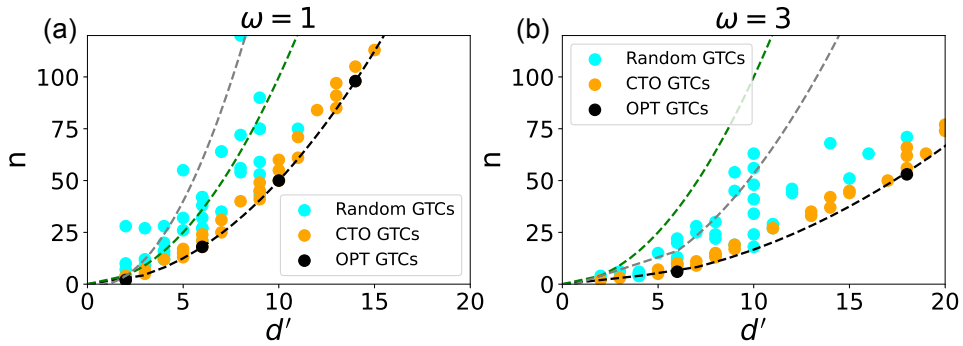


Figure 3.8: The scatter plots of the required size $n$ of different codes for achieving a target effective distance $d'$ under (a) $\omega = 1$, (b) $\omega = 3$. The black dashed line indicates the bound (Eq. (3.10)) $n = d'^2/2\omega$ for the GTCs. The green dashed line indicates the standard scaling $n = d'^2$ for $d'$ by $d'$ rotated planar surface codes. The grey dashed line indicates the scaling $n = \max[2d'^2/\omega - d'(1+1/\omega), 3d' - 2]$ for the unrotated planar surface codes with optimized aspect ratio. All the codes are of the XZZX type.

## Fault-tolerant quantum error correction

In this section we present a fault-tolerant QEC scheme for the GTCs using flag qubits. The main idea of the flag fault tolerance [165, 166, 167, 168] is that a small number of extra qubits (flags) are used to catch the bad ancilla errors that propagate to higher-weight data errors during the stabilizer measurements (e.g. a Pauli $X$ error on the syndrome qubit s that occurs after gate b, which is depicted by the red star, during the stabilizer measurement shown in Fig. 3.9). Therefore, the code distance is preserved under a circuit-level noise, i.e. a distance-$d$ code can tolerate up to $t = \lfloor (d-1)/2 \rfloor$ faults that occur at arbitrary locations during the protocol.

When considering the biased noise, we have shown that it is the effective code distance, which is typically larger than distance for tailored codes, that characterizes the code performance. As such, we need to design FT schemes that preserve the effective code distance. Following Refs. [26, 172] we give the following definition of $t'$-FTEC under biased noise: For $t' = (d'-1)/2$, an error correcting protocol using a stabilizer code with effective code distance $d'$ is $t'$ fault-tolerant if the following two conditions are satisfied:

1. For an input code word with error of effective weight $s'_1$, if any faults with effective

weight $s_2'$ occur during the protocol with $s_1' + s_2' \leq t'$, ideally decoding the output state gives the same code word as ideally decoding the input state.

2. For faults with effective weight $s'$ during the protocol with $s' \leq t'$, no matter how many errors are present in the input state, the output state differs from a code word by an error of at most effective weight $s'$.

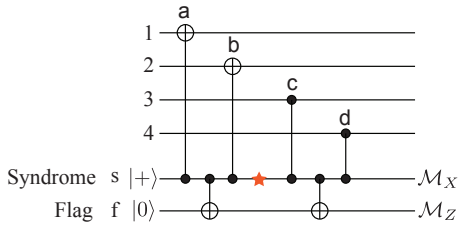


Figure 3.9: The flag circuit to extract a weight-4 stabilizer $X_1X_2Z_3Z_4$. We prepare a syndrome qubit s in $|+\rangle$ state, apply a sequence of $CX$ (a,b) and $CZ$ (c,d) gates between the s and the data qubits, and measure s in the $X$ basis to obtain the syndrome associated with $X_1X_2Z_3Z_4$. In addition, we apply two $CX$ gates between s and an extra flag qubit f to catch the bad (bit-flip) errors on the syndrome qubit. Without the flag qubit the bad errors after gates $b$ and $c$ can propagate to the data qubits with larger (effective) weight. With an extra flag qubit, the bad errors (e.g. a Pauli $X$ error depicted by the red star) can be detected and corrected adaptively.

For the GTCs, a subset of them can be made fault tolerant simply by appropriately scheduling the standard bare-ancilla syndrome extraction circuits (e.g. the circuit in Fig. 3.9 without the flag qubit), similar to the CSS surface code [173, 174]. However, other GTCs, e.g., the five-qubit code, require extra flag qubits (or other type of ancillas [134, 175, 139]) to be fault tolerantly operated since the hook errors of the ancilla circuits could reduce the effective code distances. See Appendix F of Ref. [171] for more details. Here we show that by using the flagged circuit in Fig. 3.9 with one extra flag qubit to measure the XZZX stabilizers, we can achieve the above-defined fault tolerance for *any* GTC. More precisely, we claim that by measuring the stabilizers using the flagged circuits (and unflagged circuits) in an appropriate sequence and applying proper decoding, a GTC with effective distance $d'$ can realize $t'$-FTEC, where $t' = (d' - 1)/2$.

We prove the claim in Appendix F of Ref. [171] and only sketch it here. For the flag-QEC scheme to work, two conditions have to be satisfied. (1) Errors with effective weight up to $t'$ on the ancilla qubits that propagate to higher (effective) weight errors have to be detected by the flag qubits. (2) The bad errors sharing the same flag pattern have to be distinguishable and correctable by the Knill-Laflamme conditions. The first condition is guaranteed by using the flag circuit Fig. 3.9 while the satisfaction of the second condition is in general code-specific and is proved for the GTCs in Appendix F of Ref. [171]. We leave the investigation of the GTCs performance under the circuit-level noise to future work.

### *3.2.4 Discussion*

#### Correlated Pauli $X$ and $Z$ noise

In this section, we discuss the performance of the GTCs under another physically relevant noise model, which we call the correlated Pauli $X$ and $Z$ noise: $p_X = p_Y = p_Z^\omega$ and $p_X + p_Y + p_Z = p$, where $\omega \geq 1$. This model does not assume the independence between $X$ and $Z$ errors and is widely considered in studying QEC under biased noise [89, 90, 91, 92]. Under such a model, the effective weights of the Pauli operators are $\mathrm{wt}'(Z) = 1$ and $\mathrm{wt}'(X) = \mathrm{wt}'(Y) = \omega$. It turns out that we can easily extend our analysis of the GTCs under the independent $X$ and $Z$ noise to the case where the $X$ and $Z$ error becomes correlated, and obtain similar results. Specifically, we have the following results:

(1) The effective code distance $d'_{\mathrm{cor}}$ of the GTCs under the correlated $X$ and $Z$ noise is close to that under the independent $X$ and $Z$ noise $d'$ for the same $\omega$, especially in the high-bias regime. More precisely, we have

$$\tfrac{\omega}{\omega+1} d' \leq d'_{\mathrm{cor}} \leq d'. \tag{3.12}$$

Both the upper and lower bounds are tight and they converge to be the same as $\omega$ increases.

Therefore, we can use Eq. (3.8) that we developed for calculating the $d'$ to approximately estimate the $d'_{\text{cor}}$ of the GTCs. In Fig. 3.10, we compare the bounds in Eq. (3.12) (orange bars) with the numerical estimates of the $d'_{\text{cor}}$ (green dots) for the close-to-optimal codes defined in Eq. (3.11). The numerical estimates of the $d'_{\text{cor}}$ are given by $2r-1$, where $r$ is the exponent in the expression $p_L \propto p^r$ that is extracted from fitting the MC simulations using tensor network decoders (see Appendix E of Ref. [171] for details of the decoders). The numerical estimates fall well within the theoretical bounds. For $\omega = 1$, the bounds Eq. (3.12) are relatively loose due to the factor $\frac{\omega}{\omega+1} = \frac{1}{2}$. As $\omega$ increases, the bounds become tighter and $d'$ serves as a good approximation for $d'_{\text{cor}}$ for all the codes.

(2) Based on (1), the optimal achievable performance of the GTCs under the biased model is close to Eq. (3.10). The optimal/close-to-optimal codes under the correlated $X$ and $Z$ noise are among the close-to-optimal codes under the independent $X$ and $Z$ noise. To identify the former, we need to look for codes in the latter family with the minimal number of $Y$s in the shortest logical operators. Consequently, these codes have effective distance $d'_{\text{cor}} \approx d'$, which, according to the upper bound in Eq. (3.12), is the largest achievable effective distance (given a bias and code size). Such optimal/close-to-optimal codes under the correlated $X$ and $Z$ noise can be identified in Fig. 3.10, as those with the numerically extracted effective code distance (green dots) close to Eq. (3.10) (dashed line).

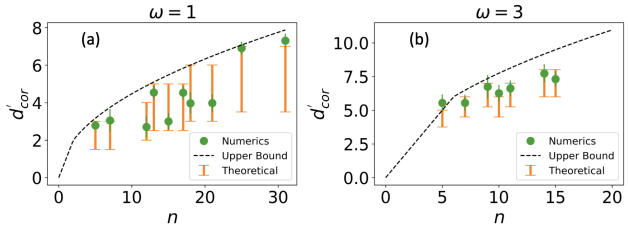


Figure 3.10: Performance of the close-to-optimal GTCs under the correlated Pauli $X$ and $Z$ noise for (a) $\omega = 1$ and (b) $\omega = 3$. The dashed line indicates the bound Eq. (3.10). The orange bars indicate the effective code distance bounds given in Eq. (3.12). The green dots are given by $2r-1$, where $r$ is the exponent in the expression $p_L \propto p^r$ that is extracted from fitting the MC simulations using tensor network decoders. The green bars indicate the standard numerical error. The range of the physical error rate $p$ used for the fitting is: (a) $p \in [0.01, 0.03]$ and (b) $p \in [0.04, 0.07]$.

## Practical applications

In practice, the noise channel in many physical systems is asymmetric, e.g. noise biased towards dephasing. Here we consider the stabilized cat qubits in bosonic systems, whose bit-flip rate is exponentially suppressed by the cat size $|\alpha|^2$ while phase-flip rate is only linearly amplified by $|\alpha|^2$, thereby supporting exponentially large noise bias [66, 79]. More importantly, the stabilized cats support a set of gates that preserve the bias of the noise [46, 4, 176], which are important for fault-tolerant QEC in the circuit level. According to Refs. [177, 69, 176], it is expected that the error rates in these systems can reach $p_Z \sim 10^{-2}, p_X \sim 10^{-6}$, which corresponds to $\omega \approx 3$. Under this bias, the optimized GTCs can act effectively as repetition codes with $d' = n$ for $n$ up to 6. For larger $n$, the GTCs remains resource efficient. As an example, the GTC with $\vec{L}_1 = (7, 5), \vec{L}_2 = (-2, 1)$ has an effective distance 9 using only 17 qubits. In contrast, to achieve the same distance with the standard surface code under the depolarizing noise one would need 81 qubits. Furthermore, if we are not restricted by local connectivity, by adding only two extra qubits we can fault-tolerantly operate this GTC using only 19 physical qubits.

Different from the surface codes, the GTCs require non-local connectivity if the qubits are placed in a planar layout. Such non-local connectivity has been demonstrated on current superconducting devices [178, 179, 180]. Furthermore, if we arrange the cyclic GTCs on a ring (see Fig. 3.2(a)), only quasi-local circular connectivity is required. Similar circular connectivity has been shown to be implementable in a resource-efficient manner in Ref. [181].

## Conclusion and outlook

In this work, we study the performance of a family of XZZX codes, including the XZZX cyclic codes and the XZZX GTCs, under biased noise. Facilitated by the topological tools, we show that the optimized XZZX codes can achieve a favorable effective distance - block size scaling. Furthermore, we show that the XZZX codes have remarkably high thresholds

that reach what is achievable by random codes, and they can fault tolerantly operated by adding one flag qubit. In combination, our results show that tailored XZZX codes give a qubit-efficient scheme for fault-tolerant QEC against biased noise.

So far, our discussion has mainly focused on the memory level. In future work, we will extend our analysis to fault-tolerant universal quantum computing with low overhead. We will investigate the implementation of fault-tolerant gates encoded in our tailored codes. Different sets of universal fault-tolerant gates on the five-qubit code have been developed and implemented experimentally. For instance, Gottesman's construction [182] is built on a three-qubit transversal Clifford gate and single-qubit Clifford gates and measurements. Ref. [183] constructs fault-tolerant CZ and T gate using the idea of pieceable fault tolerance, which has been recently demonstrated experimentally [76]. We believe it is possible to generalize these constructions to larger XZZX codes. Furthermore, a circuit level estimation of the error rates and a full analysis of the resource cost for fault-tolerant quantum computing will also be carried out.

In Ref. [161], it is shown that random Clifford deformations of the CSS surface codes can lead to better codes. It is worth investigating whether a similar transformation can boost the performance of our XZZX codes.

It is also possible to generalize the construction of the XZZX GTCs to the recently advanced quantum low-density-parity-check (LDPC) codes [184, 185, 186, 187, 12, 93], which have asymptotically finite coding rate and good block-length-to-distance scaling. The current construction of the quantum LDPC codes focuses mainly on homological CSS codes. The results in our paper indicate that non-CSS and non-homological construction might lead to more efficient codes against both symmetric and asymmetric noise.

## 3.3 Tailored topological codes for catastrophic correlated noise

Topological codes can exponentially suppress local and uncorrelated errors (or errors with finite-range correlations) by encoding quantum information non-locally in a many-body system (see e.g., Refs. [15, Chapter 10], [188, 189, 190]). Long-range correlated errors, however, e.g. due to coupling to a bosonic bath [140, 141, 142], can negatively impact the performance of QEC [191, 192].

Recently, it has been shown that a cosmic ray event (CRE) can cause catastrophic correlated errors in superconducting qubits [193, 194, 143, 77]. Upon impact of high-energy rays, phonons are created and spread in the substrate. These phonons then create quasiparticles in the superconducting material, which subsequently induces qubit decay [143]. Even though these events are rare, their effect is devastating as they cause fast correlated relaxation ($T_1$ error) in all the qubits in a chip that essentially erases the encoded quantum information [143], which is especially detrimental to long computational tasks that could take several hours [195]. Moreover, the adverse effect of CREs is not limited to superconducting qubits. Semiconductor spin qubits [196] and qubits based on Majorana fermions [197, 198] also suffer from the charge noise and quasiparticle poisoning that are resulted from CRE, respectively. One system-specific approach to reducing the impact of CREs is through changing the design of the device, for example, by introducing phonon and quasiparticle traps [199, 200, 201] and enhancing phonon relaxation in the device [198].

In this work, we take a different approach and use a distributed error correcting scheme with concatenated topological codes to detect and correct correlated errors by CREs. Distributed hardware architectures, connecting smaller nodes into a tightly-coupled system using an interconnect network, have been proposed to achieve scalability for a single computation [202, 203, 204, 205, 206, 207, 208, 209, 210]. Here, we repurpose these architectures to improve fault tolerance. Our approach is system independent and works as long as a quantum network can be built to share entanglement between separate chips. In a network of chips, a

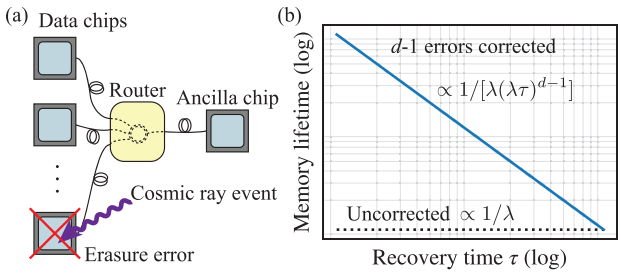


Figure 3.11: (a) Information is encoded in an error correcting code that is distributed across multiple data chips. The CRE-induced erasure errors on the data chips are corrected using ancilla-assisted syndrome measurements. (b) A code that corrects $d-1$ erasure errors, suppresses the rate of the CRE-induced catastrophic events.

CRE erases information from one chip, but as we show this event and the specific impacted chip can be detected (see Fig. 3.11a). Since the location of the error is now known, we can use erasure QEC to correct the errors and recover the information [211, 139, 212, 213, 214]. We present a low-overhead erasure QEC scheme that is fault tolerant against the CREs and discuss its implementation using superconducting chips connected with microwave links (see e.g., Refs. [215, 216, 217, 218]), and provide logical-error estimates in state-of-the-art experimental systems. Our analysis indicates that under reasonable assumptions, we can suppress the damage from these catastrophic events to higher order and reduce the CRE-induced logical error rate from 1 every 10 seconds in Ref. [143] to less than 1 per month.

We note that our scheme is not limited to correcting CREs, and it can be used more generally to deal with catastrophic correlated errors.

### *3.3.1 Setup - a two-level encoding*

We consider two levels of encoding on $n$ chips. The first level uses an error correcting code (e.g., a surface code [17]) to protect the information in each chip. In the second level, we concatenate this code with a $[[n, 1, d]]$ QEC code capable of correcting $d-1$ erasure errors [211], which is distributed across $n$ separate chips. The operations in the first level should be protected by the first-level code. Therefore, operations in that level are followed by

syndrome checks at every step. Upon a CRE impact on a specific chip, the qubits in a large region around the impact area experience a considerable reduction in their lifetime [143]. Consequently, most qubits in the impacted chip decay during an error correction cycle. The simultaneous decay of a large number qubits in a chip causes a significant increase in the number of error syndromes of the first-level encoding. The observation of a sharp jump in the number of error syndromes in a chip reveals the location of the erasure error in the second level [77, 219], which subsequently triggers error correction in the second level. We expect that by correcting $d-1$ errors we would be able to suppress the rate of catastrophic events to $\propto \lambda(\lambda\tau)^{d-1}$, where $\lambda$ is the CRE rate in a chip and $\tau$ is the time that it takes for the second-level error correction cycle (see Fig. 3.11b).

For example, we can use the [[4,1,2]] code [211] to correct single erasure errors. As shown in Fig. 3.12b, a single CRE event will trigger the QEC circuit to correct the erasure error and successfully restore the original encoding. However, if there is a second CRE erasure event during the erasure correction, the QEC circuit will fail to restore the encoded information, leading to a CRE-induced logical error rate proportional to $\lambda^2\tau$. Note that the QEC for $[[4,1,2]]$ is relatively simple because we only care about correcting single CRE errors and do not worry about CRE errors during the QEC operations. In order to use larger-distance codes, e.g., the $[[7,1,3]]$ code [220], to suppress the CRE errors to higher orders it is crucial to design the QEC circuit *fault tolerantly* so that possible CRE events during the QEC should not damage the encoded information.

### *3.3.2 Fault tolerant error correction for erasure errors*

We assume that by using sufficiently large surface codes in the first level, Pauli error rates due to the failure of the surface QEC are much lower than the rate of the CREs. As such, we only consider the errors induced by the CREs. For simplicity, we assume that a CRE-induced erasure error could propagate through a two-qubit gate and completely erase both

involved qubits [3]. Upon detecting erasure errors on a chip, we replace the erased chip with a reserve chip. The data qubits on the new chip are randomly initialized. Hence, their erasure errors are converted to detected Pauli errors randomly drawn from $\{I, X, Y, Z\}$ after the chip [3] replacement.[3] We propose a novel fault-tolerant QEC (FTQEC) scheme, called the erasure-flag scheme, that satisfies the fault-tolerant criteria [191, 172, 167]:

**Definition 12.** *For $t = d - 1$, an erasure QEC protocol is $t$ fault-tolerant using a distance-$d$ stabilizer code if the following two conditions are satisfied:*

1. *For an input codeword with an error of weight $s_1$, if $s_2$ erasures occur during the protocol with $s_1 + s_2 \leq t$, ideally decoding the output state gives the same codeword as ideally decoding the input state.*

2. *For $s$ erasures during the protocol with $s \leq t$, no matter how many errors are present in the input state, the output state differs from a codeword by an error of at most weight $s$.*

The scheme *adaptively* performs non-destructive stabilizer measurements using one ancilla qubit on an ancilla chip (Fig. 3.11a). A single erasure error that occurs on the ancilla could propagate into multiple data erasures on different data chips. However, we can detect the time and location of such bad errors. So similar to the flag FTQEC for Pauli errors [165, 166, 167], the access to this information enables us to design protocols that use minimal resources to tolerate the bad errors. Here, this information comes directly from the first-level QEC without requiring additional resources, e.g. flag qubits, in the second level.

The erasure-flag FTQEC protocol using a distance-$d$ code is implemented as follows. (i). Upon detecting erasure errors on the data qubits, replace the erased data qubits (chips), initialize the erasure-flag error set $\mathcal{E}$ that contains the detected data errors, set $s = 0$ that

3. In practice, the propagation of erasure errors at the surface-code level through the gates depends on both the microscopic detail of the erasures and the implementation of the gates. Here we consider the worst-case scenario, where a single erasure can completely erase all the qubits involved in a multi-qubit gate.

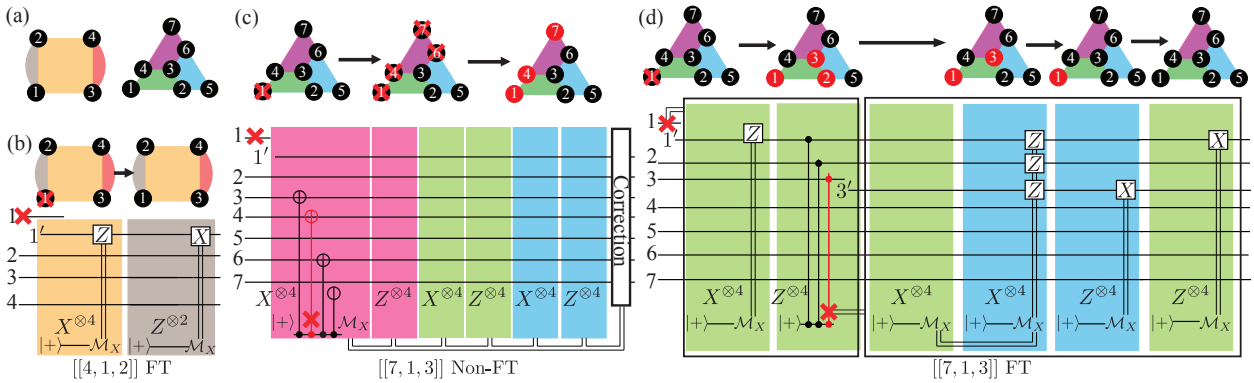


Figure 3.12: (a). The illustration of the [[4,1,2]] (left) and the [[7,1,3]] (right) code. The colored plaquettes represent stabilizer generators that have supports on the surrounding vertices (data qubits). (b) The FT circuit for the [[4,1,2]] code correcting one data erasure (red cross). Colored boxes represent an ancilla-assisted measurement of a stabilizer associated with a plaquette of the same color. The ancilla is initialized in $|+\rangle$, a sequence of $CX/CZ$ gates between the ancilla and the data qubits are applied (not shown), and the ancilla is measured in the Pauli $X$ basis. (c) The non-FT circuit for the [[7,1,3]] code that is *non-adaptive*. An initial erasure error on a data qubit triggers the circuit, which measures all the six stabilizers in a fixed sequence (pink-green-blue) and applies the correction at the end. We explicitly show the $CX$ gates in the first box ($X_3X_4X_6X_7$ stabilizer measurement) to illustrate an erasure error that propagates to multiple data errors and causes a logical failure. The top shows the evolution of the errors for the example trajectory. The red circles indicate qubits with potential Pauli errors (converted from the erasure errors). (d). The FT circuit for the [[7,1,3]] corrects the errors *adaptively*. Suppose another erasure happens during the $CZ$ gate (shown in red) between the ancilla and the third qubit while measuring $Z_1Z_2Z_3Z_4$. Upon detection of this error we stop the stabilizer measurement, discard and replace the ancilla and the thrid qubits and update the erasure-flag error set $\mathcal{E}$ to $\mathcal{E} = \{I, X_1\} \times \{I, P_3\} \times \{I, Z_1Z_2\}$, where $P_3$ indicates an arbitrary Pauli error on the third qubit. The correlated $Z_1Z_2$ error results from discarding the ancilla that is entangled with the first and the second qubits. We then measure the stabilizers $X_1X_2X_3X_4$, $X_2X_3X_5X_6$, $Z_2Z_3Z_5Z_6$, $Z_1Z_2Z_3Z_4$ to correct possible errors within $\mathcal{E}$.

counts the number of bad erasure errors that happen during the protocol and apply the following erasure-QEC. (ii). Measure a set of stabilizers of minimal size that can be used to correct the current $\mathcal{E}$. (A). If there are $s_{\text{new}}$ bad erasures detected in the middle of a stabilizer measurement with $s + s_{\text{new}} \leq d - 1$, stop the measurement immediately, update $s$ by adding $s_{new}$, replace the erased qubits (chips), update $\mathcal{E}$, and restart (ii). (B). Otherwise, apply a correction in $\mathcal{E}$ based on the measured syndromes.

The fault tolerance of the protocol is guaranteed by the following two key ingredients. (a). Bad erasures can be immediately detected so that we can keep track of the erasure-flag error set resulting from the bad errors. (b). The erasure-flag error set is correctable (different errors either have different syndromes or differ by a stabilizer) if there are fewer than $d$ faults. We show that the erasure-flag scheme can be applied to the four-qubit and seven-qubit codes using proper QEC circuits (Fig. 3.12). The FT circuit for the [[4,1,2]] code (Fig. 3.12b) corrects a single data erasure at the input. A non-FT circuit for the [[7,1,3]] code (Fig. 3.12c) is triggered by a data erasure error at the input and *non-adaptively* measures a full set of stabilizers in a fixed sequence. However, an extra erasure that occurs on the ancilla chip during a stabilizer measurement could propagate into multiple data errors and cause a logical failure. Therefore, the non-FT circuit fails to correct some consecutive double erasures. In contrast, the *adaptive* FT circuit (Fig. 3.12d), which keeps track of the possible error set and measures only a minimal set of stabilizers, can tolerate up to two consecutive erasures on arbitrary qubits. We note that the erasure-flag scheme is not guaranteed to work for any stabilizer code and it is mostly suitable for small codes or codes with certain structures, e.g. the topological surface codes with arbitrary distance (see Supplementary Material of Ref. [221]).

### *3.3.3 Analysis of the logical error rates*

Following Ref. [143], we model CREs on each chip by an independent and uncorrelated Poisson process $N(t)$, such that $P[N(t) = k] = (\lambda t)^k/k! \exp(-\lambda t)$, where $\lambda$ is the rate of the events [4]. The value of $\lambda$ depends on the geometry and other specifications of the chip [198], but for simplicity we use the reported value of $1/\lambda = 10$ s in Ref. [143]. Since the events in each chip are assumed to be independent, the introduction of additional chips increases the overall CRE rate linearly. Using the FT implementation of a QEC code that corrects

4. In principle, there might be spatial correlations between these events, but they can be treated independently as long as the correlation length is finite and the chips are separated beyond that length.

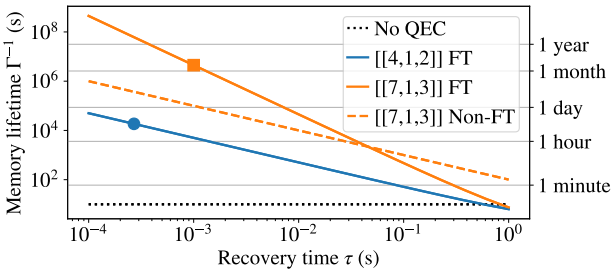


Figure 3.13: The solid (dashed) lines show the lower (upper) bound of the lifetime with (without) the fault-tolerant implementation. The dotted line shows the expected lifetime $\lambda^{-1}$ without error correction. The circle ($[[4,1,2]]$) and square ($[[7,1,3]]$) markers show estimates of the improved lifetime with error correction for maximal recovery time using experimentally feasible parameters.

$d-1$ erasure errors in a cycle, a catastrophic event might occur if there are more than $d-2$ additional events during the recovery time, $\tau$, following the first event that triggers error correction. Such a catastrophic event leads to a logical failure at the second level of encoding. The rate of these catastrophic events is obtained by taking the product of the rate of the CREs that trigger error correction and the probability that more than $d-2$ CREs happen in time $\tau$ following the first CRE. For a code over $n$ chips, the former is $n\lambda$. However, since we need an ancilla chip for our QEC scheme, the latter factor should be calculated using the rate $(n+1)\lambda$ (the reserve chips do not contribute to the logical error rates since they do not carry any logical information). Therefore, we find the rate of the catastrophic events, $\Gamma = n\lambda\{1-\exp[-(n+1)\lambda\tau]\sum_{k=0}^{d-2}[(n+1)\lambda\tau]^k/k!\}$. For $n\lambda\tau \ll 1$, we have $\Gamma \approx n\lambda[(n+1)\lambda\tau]^{d-1}/(d-1)!$, which shows the desired error suppression in this regime. While we considered the worst-case scenario, not all weight-$d$ (or higher) errors are catastrophic, and some are still correctable. Therefore, by considering the longest error correction and recovery time for $\tau$ (Fig. 3.12b and d), this analysis gives a lower bound on the memory lifetime, $\Gamma^{-1}$, limited by the $d-1$ coincident CREs within $\tau$ (solid lines in Fig. 3.13).

In contrast, for the non-FT implementation of the $[[7,1,3]]$ code, we obtain an upper bound on the memory lifetime (dashed line in Fig. 3.13). In this case, some double events

cause a logical failure. For an upper bound, we only consider the case where the first erasure error occurs on the edge chips in Fig. 3.12a. Following this event, depending on the affected chip, there are one or two stabilizer measurements during which an ancilla erasure can lead to logical failure. Therefore, we consider CREs on these 6 edge chips with the rate $6\lambda$ as triggering events and find the probability of an additional event on an ancilla during one of the stabilizer measurements. Since different stabilizer measurements (colored boxes in Fig. 3.12c) have the same number of inter-chip gates, we assume that they each take $\tau/6$. Therefore, we find the upper bound of $1/\{6\lambda[1-\exp(-\lambda\tau/6)]\}$ for $\Gamma^{-1}$ ( Fig. 3.13). In Supplementary Material of Ref. [221], we numerically verify that the logical error rates are suppressed to the desired order for both the FT and non-FT schemes. Furthermore, we show improved theoretical estimates for $\Gamma^{-1}$ under a specific model of the hardware operations.

### *3.3.4 Experimental implementation*

Our proposed scheme can be implemented experimentally in superconducting devices by coupling multiple data chips to an ancilla chip through a router [222, 223] (Fig. 3.11a). The ancilla chip is used to collect the syndrome information by coupling a syndrome patch to the data patches (all encoded in a surface code) associated with different stabilizers. The detail of coupling of the ancilla chip with one of the data chips is illustrated in Fig. 3.14. To implement an entangling gate, e.g. $CX$, between the syndrome patch S and a data patch D, we introduce an ancilla patch A on the ancilla chip and apply the measurement-based gate [10] (see the inset of Fig. 3.14). The measurement of joint Pauli operators $ZZ$ ($XX$) between the surface patches A and S (D) is implemented by lattice surgery [10], i.e. merging and then splitting the $Z$ ($X$) boundaries of the two involved patches. In our distributed architecture, we need to nonlocally merge the boundaries of the A and D patches that sit on different chips by adding new plaquettes (dashed boxes in Fig. 3.14) that connect the boundaries. Each of the new plaquette has two ancilla qubits (black dots in Fig. 3.14), each

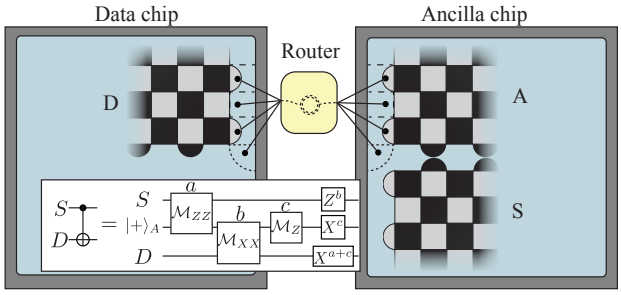


Figure 3.14: Detail of the coupling of the ancilla chip in Fig. 3.11(a) to a data chip. We show three surface patches S (syndrome), A (ancilla), and D (data). Each patch is encoded in a rotated surface code, with data qubits on the vertices and X-type (Z-type) syndrome qubits on the black (white) plaquettes. To extract error syndromes, a $CX$ gate between S and D patches is implemented by introducing the A patch and applying the measurement-based circuit shown in the inset.

sitting on one chip and is locally coupled to two data qubits on the boundary of the surface patches. To measure a new plaquette stabilizer, we apply a nonlocal $CX$ gate between the two ancilla qubits to create a Bell state $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, then apply two $CX/CZ$ gates between the ancilla qubits and their coupled data qubits, then apply another nonlocal CX gate between the ancillas and finally measure one of the ancillas. The non-local physical $CX$ gate between the ancilla qubits can be implemented by teleportation-based gates that use pre-shared and purified bell pairs between two chips [224, 225, 226, 180].

We can estimate the length of the outer QEC cycle and the corresponding upper bound of the logical error rate based on realistic experimental parameters in the superconducting architecture. The most time-consuming physical operations are the two-qubit gates ($\sim$ 100 ns [227, 228]), measurements ($\sim$ 200 ns [229]) and inter-chip state transfers ($\sim$ 100 ns [218, 217]). We assume that each surface patch is a $10 \times 10$ rotated surface code and each surface-level operation is followed by 10 repeated rounds of surface-QEC. For maximum parallelism for all the operations, we estimate that the maximum recovery time $\tau$ for the [[4,1,2]] ([[7,1,3]]) code correcting 1 (2) erasure errors is approximately 270 $\mu$s (1000 $\mu$s). See Sec. 3.3.5 for details. Based on these estimated recovery times, we obtain a lower bound of the memory lifetimes of approximately 5 hours using the four-qubit, and 51 days for

seven-qubit codes (markers in Fig. 3.13).

### *3.3.5 Comparison between the erasure-flag scheme and the Knill-QEC scheme*

We can adapt the Knill-QEC scheme [139] to our setting to correct catastrophic erasure errors on different chips. For a logical qubit encoded in an $n$-qubit code, we use three blocks of qubits indexed by different colors (black, blue, and orange). Each block is encoded using the $n$-qubit code, with each qubit in a distinct chip. In other words, there are $n$ chips in total, each containing three qubits of different colors (see Fig. 3.15(c) and (d)). To correct the erasure errors, we first prepare an encoded Bell pair between the blue and orange blocks and then teleport the information from the black block to the orange block by performing a Bell measurement on the black and blue blocks. The erasure-QEC succeeds if the logical information is not erased. That is, if we can express the logical $X$ and $Z$ operators using the remaining unaffected qubits. Faithful measurements of these logical $X$ and $Z$ operators recover the information in the black and blue blocks, respectively. Although most of the required operations are transversal, i.e. no inter-chip gates are required, the preparation of the logical $|00\rangle_L$ state for the blue and orange blocks is non-transversal (see Fig. 3.15(a)(b)). Consequently, an erasure error that happens in the middle of the logical-zero preparation could spread among different chips and erase multiple qubits in the blue/orange blocks. Fortunately, this problem can be resolved by verifying the state and redoing the preparation if necessary. Specifically, we can fault-tolerantly operate an $[[n, k, d]]$ code against erasure errors by implementing the following protocol:

(i) If there are data erasures detected at the beginning, replace the erased data qubits (chips) and apply the following erasure-QEC. (ii) Use the non-transversal circuit to prepare the logical state $|0_L\rangle$ for both the blue and the orange blocks. If there are bad erasures detected in the middle, replace the erased qubits (chips) and restart the state preparation.

Repeat until the logical $|00\rangle_L$ is prepared without bad erasures. Then apply transversal gates to prepare the bell state $|\Phi_L^+\rangle$ (between the blue and orange blocks). Next, perform the transversal Bell measurement between the black block and the blue block by measuring the intact qubits (which have not been erased) and teleport the information from the black block to the orange block.

As examples, we explicitly draw the fault-tolerant circuits for the $[[4,1,2]]$ ($[[7,1,3]]$) code correcting 1 (2) erasure errors in Fig. 3.15(c) (Fig. 3.15(d)). We note that the circuit for the $[[7,1,3]]$ is adaptive and we need to redo the preparation of $|00\rangle_L$ if there is erasure detected in the middle.

In the following, we provide a comprehensive comparison between the Knill and erasure-flag schemes in terms of their resource overhead, implementation speed, and requirements for the hardware layout. In general, the erasure-flag scheme is more resource-efficient, yet more complex and slower due to longer sequential circuits. The Knill-QEC, on the other hand, is simpler, faster, yet more resource-demanding and potentially requiring complex connectivity between different chips.

We first compare the resource overhead. For the $[[4,1,2]]$ code, the Knill-scheme requires 12 surface patches, whereas the erasure-flag scheme only requires 5 patches. In general, for an $[[n,k,d]]$ code, the Knill-scheme requires $3n$ patches, while the erasure-flag requires $n+1$ patches. If we take into account the ancilla patches used for lattice surgery, in the general case, the Knill-scheme requires $4n$ patches, while the erasure-flag requires $n+2$ patches.)

To compare the time requirement of these schemes, we roughly estimate the maximal cycle correcting 1 (2) erasure errors for the four-(seven-) qubit codes using the two schemes. As shown in Sec. 3.3.4, we take into account the following most time-consuming operations in the physical level for the superconducting architecture: two-qubit gates ($\sim$ 100 ns), measurements ($\sim$ 200 ns) and inter-chip state transfers ($\sim$ 100 ns). An inter-chip $CX/CZ$ gate using pre-shared and purified bell pair [226], which mainly consists of one step of lo-

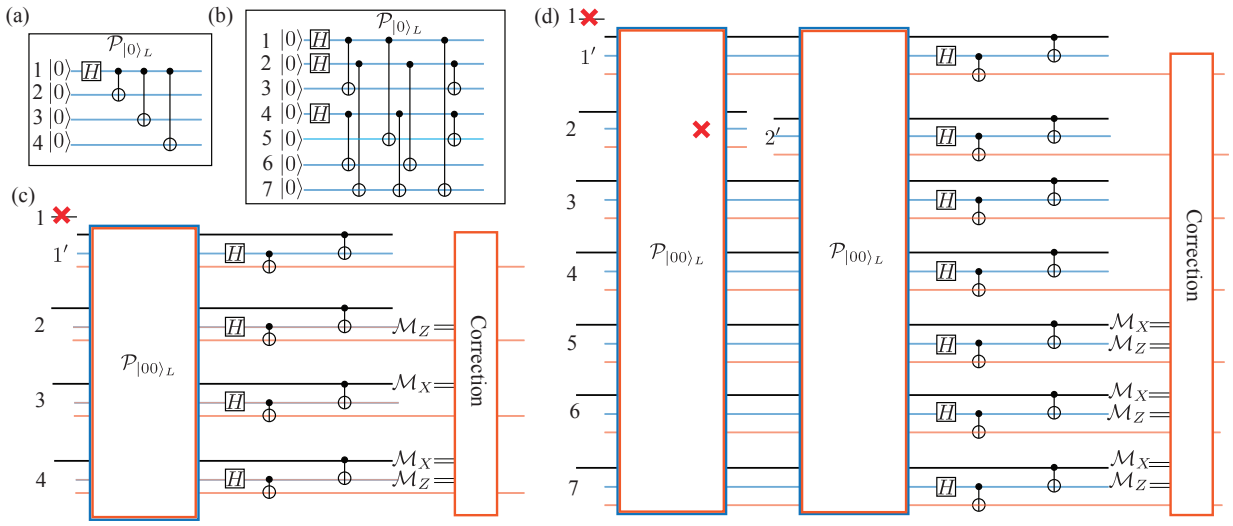


Figure 3.15: The Knill QEC circuits. The non-transversal circuits for the preparation of logical $|0\rangle$ for the (a) $[[4,1,2]]$ code, (b) $[[7,1,3]]$ code. (c). The FT circuit for the $[[4,1,2]]$ code correcting 1 data erasure at the input. (d). The FT circuit for the $[[7,1,3]]$ code correcting 1 data erasure at the input and 1 erasure during the preparation of $|00\rangle_L$. For (c) and (d), there are three blocks of qubits (black, blue and orange), each consisting of $n$ qubits encoded in an $n$-qubit code that are distributed over $n$ chips indexed by different numbers. An encoded bell pair between the blue and the orange blocks is created by preparing the logical state $|00\rangle_L$ non-transversally using the circuits (a)(b), and then applying a transversal (logical) Hadmard and a $CX$ gate. The information then is teleported from the black block to the orange block by performing a bell measurement between the black and the blue blocks. The bell measurement is implemented by first applying a transversal CX gate between the black and the blue blocks and measuring a logical $X$ ($Z$) operator on the black (blue) block, whose supported qubits have not been erased. The correction for the erasure errors is done by applying a Pauli frame update on the output orange block conditioned on the measurement outcomes.

cal entangling gates and one step of measurement, takes about $t_{\mathrm{NCX}} \approx 300$ ns. We then estimate the time for the operations at the surface-code level. We denote the time for a full surface-code QEC cycle as $t_{\mathrm{SQ}}$. A full cycle consists of 10 rounds of repeated stabilizer measurements (for a $10 \times 10$ rotated surface code), each consisting of four steps of parallel $CX$ ($CZ$) gates between the syndrome qubits and the data qubits and a measurement on the syndrome qubits. Therefore, $t_{\mathrm{SQ}} \approx 10 \times (4 \times 100 \text{ ns} + 200 \text{ ns}) \approx 6$ $\mu$s. The state preparation, measurement and single-qubit operations in the surface-code level take roughly $t_{\mathrm{SQ}}$, while the two-qubit gates based on lattice surgery are more time-consuming. A local

measurement-based $CX/CZ$ gate, which is introduced in Ref. [10] (see also Fig. 4 in the main text), mainly consists of one step of state preparation, two steps of joint Pauli measurement between two surface patches and one step of single-patch measurement. Each joint measurement between two surface patches is implemented by a merging and a splitting of the two patches, each followed by a full surface QEC cycle [10]. So in combination a local surface $CX/CZ$ gate takes about $t_{\mathrm{SCX}} \approx 6t_{\mathrm{SQ}} \approx 36$ $\mu$s. An inter-chip $CX/CZ$ gate at the surface-code level is implemented similarly, except that the merging between two surface patches on two different chips (the ancilla patch and the data patch in Fig. 4 in the main text) takes extra time due to the implementation of inter-chip physical $CX$ gates. As such, each inter-chip surface $CX/CZ$ gate takes a time $t_{\mathrm{NSCX}} \approx t_{\mathrm{SCX}} + 10t_{\mathrm{NCX}} \approx 39$ $\mu$s. Given the estimation of each surface-level operation, we can proceed to estimate the maximal cycle duration for the four- and seven-qubit codes, using the erasure-flag and the Knill circuits shown in Fig. 3.12 and Fig. 3.15, respectively. The estimates are summarized in Tab. 3.1. We note that these rough estimates strongly depend on the schemes and assumptions for different experimental operations, which are by no means optimal now. The Knill scheme is in general faster than the erasure-flag scheme since the latter requires more sequential operations using only one ancilla surface patch. However, we note that if we use more ancilla surface patches and measure different stabilizers in parallel, the erasure-flag scheme can be greatly accelerated.

In terms of the experimental layout, the erasure-flag scheme is preferred since it only requires the connectivity between the ancilla chip and the data chips, whereas the Knill-scheme might require complex connectivity between the data chips while preparing the logical state $|0\rangle_L$ (see Fig. 3.15(b)).

| | $[[4,1,2]]$ | $[[7,1,3]]$ |
|---|---|---|
| Erasure-flag | $5t_{\mathrm{SQ}} + 6t_{\mathrm{NSCX}} \approx 270\mu$s | $13t_{\mathrm{SQ}} + 24t_{\mathrm{NSCX}} \approx 1000\mu$s |
| Knill | $4t_{\mathrm{SQ}} + 2t_{\mathrm{SCX}} + 3t_{\mathrm{NSCX}} \approx 210\mu$s | $6t_{\mathrm{SQ}} + 2t_{\mathrm{SCX}} + 6t_{\mathrm{NSCX}} \approx 340\mu$s |

Table 3.1: Estimation of the maximal recovery time of a QEC cycle correcting 1 (2) erasure errors for the 4 (7)-qubit codes based on experimentally-relevant parameters. Each erasure-flag circuit shown in Fig. 2 in the main text consists of one round of erasure detection (surface QEC) at the beginning, which takes $t_{\mathrm{SQ}}$, and the measurements of multiple stabilizers. A measurement of a weight-$w$ stabilizer consists of the initialization of the ancilla qubit, $w$ sequential inter-chip CX/CZ gates between the ancilla qubit and the data qubits, and a measurement on the ancilla. So each weight-$w$ stabilizer measurement takes $T_w \approx 2t_{\mathrm{SQ}} + wt_{\mathrm{NSCX}}$. As shown in Fig. 2(b) in the main text, the worst-case erasure-flag circuit for the $[[4,1,2]]$ code correcting 1 erasure consists of one round of initial erasure detection, one weight-4 stabilizer measurement and one weight-2 stabilizer measurement. As such, it takes roughly $t_{\mathrm{SQ}} + T_4 + T_2 = 5t_{\mathrm{SQ}} + 6t_{\mathrm{NSCX}} \approx 270\mu$s. As shown in Fig. 2(d) in the main text, the worst-case FT erasure-flag circuit for the $[[7,1,3]]$ code correcting 2 erasures consists of one round of initial erasure detection and six weight-4 stabilizer measurements. As such, it takes roughly $t_{\mathrm{SQ}} + 6T_4 = 13t_{\mathrm{SQ}} + 24t_{\mathrm{NSCX}} \approx 1000\mu$s. On the other hand, as shown in Fig. 3.15(c)(d), each Knill circuit consists of one round of initial erasure detection ($t_{\mathrm{SQ}}$), possibly multiple rounds of logical state preparation $\mathcal{P}_{|00\rangle_L}$ ($T_{00}$), one round of Hadamard gates and two rounds of local CX gates ($t_{\mathrm{SQ}} + 2t_{\mathrm{SCX}}$). The preparation of logical state $|00\rangle_L$ is the most time-consuming part due to the sequential implementation of multiple inter-chip $CX$ gates. For the $[[4,1,2]]$ code, $T_{00} = 2t_{\mathrm{SQ}} + 3t_{\mathrm{NSCX}}$, and the maximal QEC cycle shown in Fig. 3.15(c) takes roughly $2t_{\mathrm{SQ}} + 2t_{\mathrm{SCX}} + T_{00} = 4t_{\mathrm{SQ}} + 2t_{\mathrm{SCX}} + 3t_{\mathrm{NSCX}} \approx 210\mu$s. For the [[7,1,3]] code, $T_{00} = 2t_{\mathrm{SQ}} + 3t_{\mathrm{NSCX}}$, and the maximal QEC cycle shown in Fig. 3.15(d) takes roughly $2t_{\mathrm{SQ}} + 2t_{\mathrm{SCX}} + 2T_{00} = 6t_{\mathrm{SQ}} + 2t_{\mathrm{SCX}} + 6t_{\mathrm{NSCX}} \approx 342\mu$s.

### *3.3.6 Discussion*

In principle, our scheme can be extended to universal fault-tolerant computing. Multiple logical qubits in the second-level code can be encoded transversely across all the data chips such that different base qubits comprising of each logical qubit sit on different chips and each chip contains exactly one base qubit from each logical qubit. Such an encoding is fault-tolerant since each CRE only erases one base qubit in each logical qubit, which can be corrected by performing QEC on different logical qubits independently. The implementation of universal fault-tolerant gates can be adapted from existing protocols [230]. Furthermore,

the resource overhead required for overcoming the CREs could be less than that required for the standard depolarizing noise. For example, we can prepare the magic states non-fault-tolerantly and verify them by performing erasure detection, without applying costly magic-state distillation [231, 230]. Additionally, we can use Knill-type QEC [139] that utilizes teleportation to correct erasure errors. We explore and compare that approach to ours in Sec. 3.3.5.

Lastly, we note that for now the second layer of QEC only suppresses the logical error rate due to the erasure errors. In the regime where the Pauli errors due to the failure of the surface codes are more detrimental, it is advantageous to minimize the total logical error rate by tailoring the outer codes to correct both the erasure and the Pauli errors. For instance, Ref. [214] shows that the surface code can handle a mixture of erasure and Pauli errors using a tailored decoder. We leave this topic to future work.

# CHAPTER 4

# QUANTUM LOW-DENSITY-PARITY-CHECK CODES

As discussed extensively in the previous chapters, standard fault-tolerant schemes for achieving fault-tolerant quantum computing, such as those based on the paradigmatic surface code, are generally very costly in terms of resource overhead, requiring millions of qubits to solve problems of interest [17, 18, 19, 20].

Recently, a new approach based on high-rate quantum low-density parity-check (qLDPC) codes has been proposed as a promising route to reduce the resources required. Unlike planar surface codes [35, 17, 18] that encode a single logical qubit per block, qLDPC codes can encode multiple logical qubits per block and achieve a much higher, asymptotically constant encoding rate [232, 233] as well as better distance scaling [234, 235, 236]. As shown in Gottesman's seminal work [48], it is possible to achieve fault-tolerant quantum computing with a constant space overhead, i.e. the ratio between the physical qubits and the logical qubits remain a constant when increasing the size of the logical circuit, using constant-rate qLDPC codes.

However, the research along this line of the qLDPC codes is still at its infant stage. Several fundamental open questions (the fault-tolerance of QEC, the capability of performing logical operations, etc.) need to be answered in order to better understand these codes, and several major challenges (finding practical codes, implementing non-local connectivity, etc.) need to be addressed to apply them in practice. See Sec. 4.2 for more details. In this chapter, we present a new work based on Ref. [237] that answers some of these open questions and contributes to bringing these abstract qLDPC codes into practical regimes.

In Sec. 4.1, we will introduce the tradeoff between geometrical locality and the error-correction capabilities when designing codes and performing QEC. As we will see, the strong constraints on the geometrically local topological codes motivate us to consider geometrically non-local qLDPC codes.

In Sec. 4.2, we will briefly review the definitions, development, and challenges of qLDPC codes. As we will show, although there are remarkable breakthroughs in the coding theory recently and qLDPC codes with superior parameters exist theoretically in the asymptotically large-block regime, applying them for fault-tolerant quantum computing in practice remains a formidable task.

In Sec. 4.3, we will present a new scheme for implementing some of the interesting qLDPC codes in practice and applying them for fault-tolerant quantum computing using the so-called reconfigurable atom arrays, which is a recently developed hardware that features certain long-range qubit connectivity and parallel classical controls [47, 22]. This section will be based on Ref. [237].

## 4.1 Tradeoff between geometrical locality and error-correction capabilities

In this section, we briefly review existing constraints on geometrically-local quantum codes, or topological codes that are described in detail in Sec. 3.1.

For a $[[n, k, d]]$ topological stabilizer code defined on a $D$-dimensional lattice, the code parameters are subject to the following contraints:

1. The Bravyi-Poulin-Terhal (BPT) bound [148]:

$$kd^{\frac{2}{D-1}} \leq O(n). \tag{4.1}$$

2. The Bravyi-Terhal (BT) bound [238]:

$$d \leq O(n^{\frac{D-1}{D}}). \tag{4.2}$$

3. The Haah's bound [239]:

$$k \leq O(n^{\frac{D-2}{D}}). \tag{4.3}$$

These bounds constrain the maximally achievable $k$, which characterizes the space overhead, and $d$, which roughly characterizes the achievable logical error rates. In particular, for $2D$ codes, we have

$$k \leq O(1), \quad d \leq O(\sqrt{n}), \tag{4.4}$$

which is saturated by the planar surface code. Consequently, we should not expect to do better than the surface code, if restricted in $2D$.

As $D$ increases, the upper bounds for $k$ and $d$ also increase. This indicates that it is possible to find better and better codes by increasing the embedding dimension. Indeed, $3D$ codes that can saturate these bounds and perform strictly better than the planar surface codes have been found recently [240, 241].

However, based on these bounds, it is impossible to find codes with a constant encoding rate ($k = \Theta(n)$) and growing distances $d > \Omega(1)$ in any finite dimension. This suggests that one would have to discard any locality constraint in order to find nontrivial codes with a constant encoding rate, which enables fault-tolerant quantum computing with a constant space overhead [48].

## 4.2 Definitions, development, and challenges of qLDPC codes

### Definition

The qLDPC codes are broadly defined as stabilizer codes with a choice of stabilzier generators such that: (1) each stabilizer generator checks $O(1)$ qubits and (2) each qubit is involved in $O(1)$ stabilizer generators. For CSS qLDPC codes, these two conditions translate to sparse check matrices, i.e. each of $H_X$ and $H_Z$ has bounded row and colume weights. Noe that this

| Code | $k$ | $d$ |
|---|---|---|
| Surface code [34] (1999) | 1 | $\Theta(\sqrt{n})$ |
| Freedman-Meyer-Luo [242] (2002) | 2 | $\Omega(\sqrt{n\sqrt{\log n}})$ |
| Hypergraph product codes [232] (2014) | $\Theta(n)$ | $\Theta(\sqrt{n})$ |
| Hyperbolic surface codes [243] (2016) | $\Theta(n)$ | $\Theta(\log n)$ |
| High-dimensional expander codes [244] (2020) | $\Theta(\sqrt{n})$ | $\Omega(\sqrt{n}\ \mathrm{polylog}(n))$ |
| Fibre bundle codes![184] (2020) | $\Theta(n^{\frac{3}{5}}/\mathrm{polylog}n)$ | $\Omega(n^{\frac{3}{5}}/\mathrm{polylog}n)$ |
| Lifted product codes [186] (2020) | $\Theta(n^{\alpha}\log n)$ | $\Omega(n^{1-\alpha}\log n)$ |
| PK codes [236] (2022)<br>Quantum Tanner codes [245, 246] (2022)<br>DHLV codes [247] (2022) | $\Theta(n)$ | $\Theta(n)$ |

Table 4.1: Development of qLDPC codes [12, 13].

broad definition does not contain any locality constraint, thus including quantum codes that cannot be embedded into any finite dimension and not constrained by the locality bounds in Sec. 4.1.

## Development

We provide a list of representative qLDPC codes in Table 4.1 (based on Ref. [12] and Ref. [13]) along the development of the so-called good qLDPC codes, i.e. qLDPC codes with a constant encoding rate ($k = \Theta(n)$) and linear distance scaling ($d = \Theta(n)$).

Codes with a constant encoding rate was first achieved by the Hypergraph product codes [232], which we will use and introduce in detail in Sec. 4.3. However, achieving linear distance had been challenging. A so-called distance barrier $d \leq O(\sqrt{n}\ \mathrm{polylog}(n))$ had not been broken for years until 2020. Finally, asymptotically good qLDPC codes were found in 2022.

### Challenges

Although qLDPC codes have superior code parameters asymptotically, major challenges are present in designing practical fault-tolerant schemes based on these codes.

First, in order to realize their appealing features, qLDPC codes require long-range connectivity between qubits, making their physical realization challenging [148, 248, 249]. While several proposals have been made for physical implementation of qLDPC codes in superconducting qubit architectures, the required long-range and multi-layer connectivity is considerably beyond both current and medium-term hardware capabilities [250, 251, 252].

In bringing qLDPC codes into practical use for full-fledged quantum computation, further challenges arise. A rigorous analysis of the circuit-level fault tolerance of qLDPC codes is lacking, despite some promising numerical evidence [251, 250]. Also, it is currently unclear if finite-size qLDPC codes can outperform surface codes in near- or medium-term devices with $\lesssim$10000 qubits and realistic physical error rates above $10^{-3}$. Since Gottesman's seminal results demonstrating that qLDPC codes can enable fault-tolerant quantum computing with constant space overhead [48], several practical gate constructions have recently been proposed [253, 254, 255, 256]. However, no studies of the circuit-level performance of these protocols have been carried out to date. In particular, it is not clear if the performance and low overhead of the qLDPC codes can be maintained during computation in a full circuit-level fault-tolerant setting.

## 4.3 Fault-tolerant quantum computation with qLDPC codes in reconfigurable atom arrays

In Ref. [237], we propose and analyze a realistic hardware-efficient neutral atom implementation of fault-tolerant quantum computation with high-rate qLDPC codes. We provide concrete experimental and theoretical blueprints, demonstrating their advantage over sur-

face codes starting from as few as several hundred physical qubits. Our proposal is based on reconfigurable atom arrays, a newly-developed hardware architecture for quantum computation with long-range, reconfigurable connectivity [47, 22]. We show how the product structure of many qLDPC codes [232, 234, 12] naturally matches the parallelism afforded by physical realizations of reconfigurable atom arrays, enabling their hardware-efficient implementation in a logarithmic number of steps. Through a combination of single-shot circuit-level threshold proofs and circuit-level simulations, we find competitive performance for hypergraph product (HGP) [232] codes and quasi-cyclic lifted product (LP) [234] codes, achieving error thresholds of around 0.6% under a circuit-level depolarizing noise model that neglects idling errors. Accounting for idling errors, which only have a minor contribution for the finite-size codes of our interest, we achieve an order of magnitude saving over the surface code with less than 3000 physical qubits (including ancillas) at a physical error rate of $10^{-3}$ (see Fig. 4.1 bottom panel and Table 4.2). We further extend this analysis to logical gate operation, numerically demonstrating that the high thresholds and good subthreshold scalingof high-rate qLDPC codes can be maintained during computation, paving the way to low-overhead fault-tolerant quantum computing.

### *4.3.1 Overview of qLDPC-based quantum computer*

An overview of our qLDPC-based approach to fault-tolerant quantum computation is shown in Fig. 4.1. It consists of a high-rate qLDPC memory block that reliably and efficiently stores the quantum information, a processor with computational logical qubits such as surface or color codes that perform logical gates, and mediating ancillae that interconnect the memory and processor. This allows us to take advantage of the dense storage capabilities of the qLDPC block while allowing flexible execution of quantum circuits. Adopting the conventional $[\![n, k, d]\!]$ notation for a code with $n$ physical qubits, $k$ logical qubits, and distance $d$, the qLDPC block using the HGP codes described below can provide a dense $[\![\Theta(k), k, d_{\text{mem}}]\!]$

encoding, where the memory distance $d_{\text{mem}} = O(\sqrt{k})$. This results in a constant encoding rate $k/n$ and a logical failure rate (LFR), defined as the probability that any of the logical qubits fails per code cycle, exponentially decaying with the code distance. We note that using LP codes with a higher encoding rate and better distance scaling (see Methods and Fig. 4.7) further reduces the space overhead at small sizes. Our processor consists of $m$ computational qubits of code parameters $[\![\Theta(d_{\text{comp}}^2), 1, d_{\text{comp}}]\!]$, where the code distance $d_{\text{comp}} = \Theta(\text{polylog}(kT))$, with $T$ being the depth of the logical circuit to execute, is chosen to produce a sufficiently low error rate. The mediating ancillae, one for each computational qubit (see Fig. 4.8), have code parameters $[\![O(d_{\text{comp}} d_{\text{mem}}), 1, \min(d_{\text{comp}}, d_{\text{mem}})]\!]$. By performing ancilla-assisted lattice surgery, the stored logical information can be teleported between any given pair of memory and computation qubits. Within this architecture, logical gates can be applied in parallel to a subset $m$ of the stored memory qubits in each logical circuit step. Selecting $m = O(k/d_{\text{comp}} d_{\text{mem}})$ ensures that the cumulative ancilla block overhead is $O(k)$, which does not exceed the overhead of the memory block. This choice leads to a constant encoding rate for quantum computation. For HGP codes with $d_{\text{mem}} = \Theta(\sqrt{k})$ and $d_{\text{comp}} = \Theta(\text{polylog}(kT))$, this implies $m = O(\sqrt{k}/\text{polylog}(kT))$, but if smaller code distance provides sufficient error suppression, the parallelism can be increased while maintaining constant encoding rate.

### *4.3.2 Implementation in Neutral Atom Arrays*

#### Overall scheme

We now describe the hardware-efficient implementation of this qLDPC-based architecture in the atom array platform. Here, qubits are encoded in long-lived spin degrees of freedom of an atom, with seconds-range coherence times [257, 258, 259, 47], high-fidelity single-qubit (>99.9%), two-qubit (>99.5%) control and mid-circuit measurements (>99.8%) [260, 261, 262, 257, 258, 47, 22]. By shuttling atoms around in optical tweezers, one can reconfigure

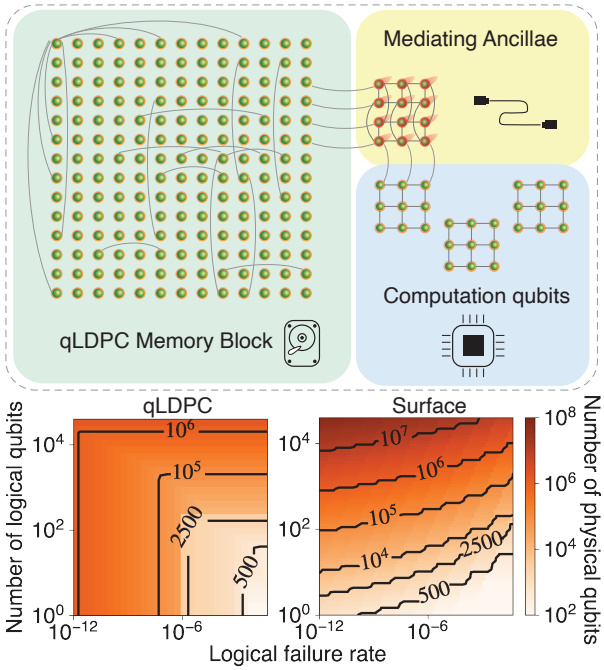


Figure 4.1: Architecture of a qLDPC-based fault-tolerant quantum computer using reconfigurable atom arrays. This consists of a qLDPC memory block, a processor with computational logical qubits, and mediating ancillae between the memory and the processor. The lower panel shows a contour plot of the number of physical qubits (including data and ancilla qubits) required by our architecture, at a $10^{-3}$ physical error rate, given a target number of logical qubits and a target logical failure rate, compared to the surface code. The qLDPC space overhead is given by the minimum of that for the LP codes shown in Fig. 4.7(b) with less than 1428 data qubits and that for HGP codes using an extrapolation of the numerical results in Fig. 4.7(a).

the processor connectivity during quantum evolution with minimal decoherence [47, 263] and realize parallel two-qubit gate operations with qubits across the whole system. The coherent shuttling approach features a high degree of parallelism, inherent to multiplexing with optical tools. A particularly powerful tool is the so-called acousto-optic deflectors (AODs), which can simultaneously control the position of rectangular grids composed of thousands of atoms, simply with two control waveforms for the X and Y coordinates [47]. These AOD tools are a key enabling technology for qubit transport in atom arrays, and support an inherent “product structure" that, as we will now show, is well-suited to the implementation of HGP and LP codes.

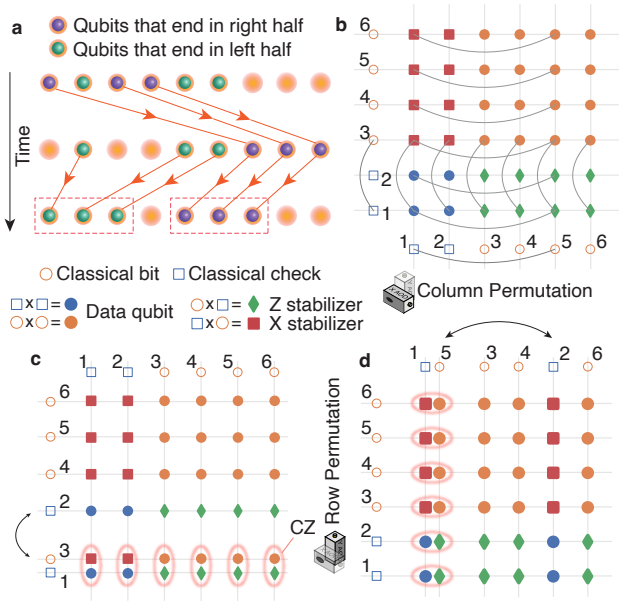


Figure 4.2: Efficient implementation of quantum LDPC codes with atom arrays. (a) Illustration of the algorithm to perform an arbitrary 1D log-depth rearrangement. We first move all atoms that need to end in the right half of the system to the right side, then compact each half into adjacent sites, so that there is sufficient workspace for subsequent steps. Moving atoms will be displaced perpendicularly during movement, to avoid collisions with static atoms in the same row, see Supplementary Movie. The same procedure can then be repeated on each half of the system recursively for depth $\log(L)$, where $L$ is the length of the atom array to be rearranged, resulting in the desired ordering. The algorithm uses 50% more static traps than the number of atoms as workspace. (b) Illustration of the HGP code, obtained as a product of two classical codes. Lines indicate that the parity check at the syndrome node involves the corresponding data node. (c,d) The required connectivity can be implemented via parallel row permutations, followed by parallel column permutations. Although we illustrate this for one pair of row/column interacting, the same permutations and CZs can be applied on multiple rows or columns in parallel.

We first describe an algorithm to implement good classical LDPC codes via efficient 1D atom rearrangements without atom collisions. Recall that a classical (LDPC) code is associated with a parity check matrix $H$, whose rows(columns) are associated with classical checks(bits). The $i$-th check is connected to the $j$-th bit if $H_{i,j} = 1$. We first lay out the classical LDPC code on a line and group syndrome extraction into layers of parallel entangling operations, resulting in an ordering of checks and bits such that the connected checks and bits in a given layer are neighboring. To achieve this ordering, we employ a

divide-and-conquer rearrangement strategy that only requires a number of steps logarithmic in the total number of checks and bits, as described in detail in Fig. 4.2(a), Sec. 4.3.2, and Supplementary Movie of Ref. [237]. This generalizes prior proposals for scrambling circuits [264] to arbitrary rearrangements. After sorting the atoms, a global laser pulse is applied to entangle neighboring checks and bits in parallel, before proceeding to the next layer of atom rearrangement, thus allowing the implementation of syndrome extraction in classical LDPC codes.

We next demonstrate how the product structure of one of the prototypical qLDPC codes (Fig. 4.2(b)), HGP codes, naturally matches the product structure of crossed AOD optical hardware, enabling its hardware-efficient implementation. We start from a pair of classical LDPC codes illustrated in the horizontal and vertical directions, with checks and bits denoted as blue squares and orange circles, respectively. To construct the resulting HGP code, we place a data qubit at each intersection of a horizontal check and a vertical check, or a horizontal bit and a vertical bit within a 2D grid [232] (Fig. 4.2(b)). $X$ stabilizer syndrome qubits and $Z$ stabilizer syndrome qubits are then placed at the intersection of a horizontal check and a vertical bit, or a vertical check and a horizontal bit, respectively. Importantly, the qubit connectivities are directly inherited from the underlying classical code in the horizontal and vertical directions, and thus the same entangling gates are applied across every row or column, matching well with the product structure of crossed AODs. Thus, by performing parallel row reordering in the vertical direction based on the vertical LDPC code, interleaved with entangling gates between data and stabilizer qubits (Fig. 4.2(c)), and then repeating the same in the horizontal direction (Fig. 4.2(d)), we can implement the syndrome extraction for HGP codes. The concrete syndrome extraction circuits are presented in Algs. 6-8 in the following Sec. 4.3.2.

## Details of the atom rearrangement

Here, we provide the details of the atom rearrangement, which is required for implementing the qLDPC codes (see Sec. 4.3.2).

The reconfigurable atom array platform features efficient, parallel control and rearrangement of large numbers of qubits, enabling the implementation of long-range connected quantum processors. As discussed in Sec. 4.3.2, optical tools such as crossed acousto-optic deflectors (AODs) can generate a rectangular grid of optical tweezers that can be reconfigured on the fly [47, 263, 265], allowing the control of large code blocks consisting of thousands of physical qubits with only a handful of classical controls.

However, the use of AODs for dynamic rearrangement comes with two key constraints: First, as the X and Y direction optical spots are controlled by separate AODs, the same operation needs to be applied across multiple rows and/or columns. Second, different rows of atoms cannot cross each other due to beating between RF tones and atom collisions, although they can be temporarily transferred and stored in static traps, such as those based on spatial light modulators (SLMs). Thus, the implementation of qLDPC codes with improved code parameters (number of encoded qubits, code distance), which often relies on pseudorandom expander graphs with complex connectivity graphs, requires the development of efficient atom rearrangement algorithms.

We provided a sketch of our atom rearrangement algorithm in Sec. 4.3.2. Now, we provide a more detailed description of it in Algs. 6-8. We also illustrate the algorithm with a movie in Supplemental Movie 1 of Ref. [237].

The first component, arbitrary 1D atom rearrangements with a number of steps that scales logarithmically, is described in detail in Alg. 6, illustrated in Fig. 4.2 and explicitly worked out for a small example in Fig. 4.3. Since successive layers each reduce the system size by half, the total number of layers required to achieve the desired rearrangement is $\lceil \log_2 L \rceil$. Thus, arbitrary rearrangements in very large systems can be achieved in a small

number of layers. We note that although this method shares some similarities to existing techniques such as bitonic sorting, the different constraints (comparators vs. parallel moves) lead to differences in the algorithm itself. Our algorithm can also be readily extended to use parallel qubit swaps [40, 266] at increasing distances as the basic primitive, with the same $O(\log L)$ scaling with system size.

The second component is the observation that the product structure of crossed AODs matches well with the product structure present in many qLDPC codes. In addition to the discussion in Sec. 4.3.2, we provide the details of the syndrome extraction circuit for HGP codes based on this observation in Alg. 7, which we refer to as the "product coloration circuit", as it makes use of coloration circuits for each of the component classical codes [250]. Note that the use of our product coloration circuit, as opposed to the coloration or cardinal circuits in Ref. [250], is necessary to fully exploit the parallel rearrangement capabilities across rows and columns. Here, the native entangling gate set of current atom array systems are diagonal [267, 260], so we use CZ gates and appropriate Hadamard rotations to perform syndrome extraction. Note that under global laser excitation and phase advances, any pair of qubits that are within a certain radius (known as the blockade radius) of each other will execute a CZ gate, while any individual qubits will undergo an identity gate. In order to compare our results against the literature, we use CNOT gates as the entangling gates in our simulations. This can be physically justified if the CZ gates are much noisier than the Hadamard gates.

The product coloration circuit separately extracts the $X$ and $Z$ syndromes, each requiring both a horizontal and vertical step. Thus, if the coloration of each of the classical codes involves $\Delta_C$ colors (for the codes constructed from $(3, 4)$-biregular graphs that we considered, $\Delta_C = 4$ [251]), the product coloration circuit will have $4\Delta_C$ entangling layers.

The product coloration circuit can also be applied to the LP codes we use in this work. As shown in Fig. 4.6(b), a LP code has the same product vertical (horizontal) connectivity

**Algorithm 6:** Arbitrary 1D Atom Rearrangement in a Logarithmic Number of Steps

**Input :** Final ordering $O = [o_i]$ $(i = 1...N)$ of all $N$ atoms, where $o_i$ is the final position of the atom that was initially at position $i$.

**Input :** Initial positions $A = [a_i]$ $(i = 1...N)$ of all $N$ atoms, with the positions ordered as $a_1 < a_2 < ... < a_N$.

**Input :** Positions $P = [p_j]$ $(j = 1...M)$ of all $M$ possible qubit locations, where $M \geq 3N/2$.

**Output:** Positions $C = [c_{s,i}]$ for the $i$th atom in the $s$th rearrangement step, for all $N$ atoms and all rearrangement steps.

**Function** `Rearrange`$(O, A, P)$**:**
  **if** $N = 1$ **then**
    $c_{1,1} \leftarrow p_1$
    **return** $[c_{s,i}]$
  $s \leftarrow 1$ `// Layer counter`
  **if** $a_N > p_N$ **then**
    `// Compactify atoms to the left to make space for subsequent moves`
    **for** $i \leftarrow 1$ **to** $N$ **do**
      $c_{s,i} \leftarrow p_i$
    $s \leftarrow s + 1$
  `// Determine whether each atom ends in the left or right half`
  $L, R \leftarrow []$
  `// Workspace separator for recursion`
  $X = \lfloor 3N/2 \rfloor$
  **for** $i \leftarrow 1$ **to** $N$ **do**
    **if** $O_i \leq \lfloor N/2 \rfloor$ **then**
      L.append$(i)$
      $c_{s,i} \leftarrow p_{N+len(L)}$ `// Move to right`
      $c_{s+1,i} \leftarrow X + len(L)$ `// Compactify`
    **else**
      R.append$(i)$
      $c_{s,i} \leftarrow a_i$ `// Stay in place`
      $c_{s+1,i} \leftarrow len(R)$ `// Compactify`
  $s \leftarrow s + 2$
  `// Recursive call on each half`
  $C_l \leftarrow$ `Rearrange`$(O[L], C[s-1, L], P[1..X])$
  $C_r \leftarrow$ `Rearrange`$(O[R], C[s-1, R], P[X+1..M])$
  $C[s.., L] \leftarrow C_l$
  $C[s.., R] \leftarrow C_r$
  **return** $C$
`// Main function`
`Rearrange`$(O, A, P)$

as a HGP code when flattening the inner nodes vertically (horizontally). Thus, the same product coloration circuit can be applied to the LP codes with an extra step of flattening the inner codes in between establishing the horizontal/vertical connections. As we use 3 by 5 base matrices with all weight-one entries, the product coloration circuit for the LP codes has an entangling gate depth of $4 \times 5 = 20$.

To further reduce the depth of the syndrome extraction circuit, we also propose a modification of the above circuit in Alg. 8 and Fig. 4.4, which we refer to as the pipelined product coloration circuit. Here, the main challenge is to choose a gate ordering such that the desired $X$ and $Z$ syndromes are correctly extracted. By performing pipelining and extracting the $X$ syndrome of the second round simultaneously with the $Z$ syndrome of the first round, we can ensure that the gate ordering is always valid, while reducing the number of entangling layers required to perform $d$ rounds of syndrome measurement to $(2d+2)\Delta_C$. This could be particularly relevant in further suppressing the effect of idling errors as well as improving the performance of logical gates, which in our scheme require $d$ rounds of repetition. However, our numerical simulations all make use of the product coloration circuit and we leave the exploration of other syndrome extraction circuits to future work.

### Estimation of the rearrangement time and the idling errors

We now estimate the scaling and quantitative experimental timescales of our rearrangement algorithm, demonstrating that the proposed hardware implementation is indeed achievable with existing experimental parameters.

For the atom rearrangement schemes described in Sec. 4.3.2, we can estimate the amount of time required to implement one round of stabilizer measurements for both the HGP codes and the LP codes using the technology that has been demonstrated in Ref. [47]. We assume a transfer time $\tau_t$ between a static SLM trap and dynamic AOD trap, a peak atom moving acceleration rate of $a_p$ with a cubic spline trajectory, and a uniform grid spacing $d$. For

**Algorithm 7:** Product Coloration Circuit for HGP Syndrome Extraction

**Input :** Edge colorations $\mathcal{C}_h, \mathcal{C}_v$ of Tanner graphs associated with horizontal and vertical classical codes $C_h$, $C_v$ that form the hypergraph product code.

**Output:** Measurement outcomes of all $X$ and $Z$ stabilizer generators.

`// X stabilizers`

Apply a Hadamard on all data qubits.

Prepare an ancilla in $|+\rangle$ for each $X$ stabilizer and move all $X$ ancilla qubits (red) into the LDPC grid region shown in Fig. 4.6. Do not include any $Z$ ancilla qubits (green).

**for** *direction* $f \in$ *{horizontal, vertical}* **do**

  **for** *color* $c \in \mathcal{C}_f$ **do**

    Apply algorithm 1 in direction $f$, across the whole grid, to bring each pair of qubits connected by an edge of color $c$ in direction $f$ together.

    Apply a CZ gate on each pair of neighboring qubits.

Apply a Hadamard on all data qubits.

Move $X$ ancilla qubits out of the grid region and measure them in the $X$ basis.

`// Z stabilizers`

Prepare an ancilla in $|+\rangle$ for each $Z$ stabilizer and move all $Z$ ancilla qubits (green) into the LDPC grid region shown in Fig. 4.6. Do not include any $X$ ancilla qubits (red).

**for** *direction* $f \in$ *{horizontal, vertical}* **do**

  **for** *color* $c \in \mathcal{C}_f$ **do**

    Apply algorithm 1 in direction $f$, across the whole grid, to bring each pair of qubits connected by an edge of color $c$ in direction $f$ together.

    Apply a CZ gate on each pair of neighboring qubits.

Move $Z$ ancilla qubits out of the grid region and measure them in the $X$ basis.

simplicity, we assume that the number of atoms on a line to be rearranged is a power of 2, $L = 2^k$. The algorithm in Alg. 6 will therefore have a depth of $k = \log_2 L$ layers. In order to provide enough workspace for shuttling, the total number of traps is $3L/2$.

The compactification step at scale $s$ requires a move of distance at most $sd/2$. Moving all target atoms to the right requires a move of distance at most $sd$. We can combine the two steps in a way that we pick up atoms for the next move and drop off atoms from the previous move at the same time; thus, each step requires on average one trap transfer between static and dynamic traps. As described in Ref. [47], using a cubic spline movement trajectory, a move of distance $l$ requires time $\sqrt{6l/a_p}$. The scaling with distance can be understood by

**Algorithm 8:** Pipelined Product Coloration Circuit for Multi-round HGP Syndrome Extraction

**Input :** Edge colorations $\mathcal{C}_h, \mathcal{C}_v$ of Tanner graphs associated with horizontal and vertical classical codes $C_h$, $C_v$ that form the hypergraph product code.
**Input :** Number of syndrome repetition rounds $d$.
**Output:** Measurement outcomes of all $X$ and $Z$ stabilizer generators.

```
// X stabilizers of the first round
```
Apply a Hadamard on data qubits in the bottom left block (orange) of Fig. 4.6.
Prepare an ancilla in $|+\rangle$ for each $X$ stabilizer and move all $X$ ancilla qubits (red) into the LDPC grid region shown in Fig. 4.6. Do not include any $Z$ ancilla qubits (green).
**for** *color* $c \in \mathcal{C}_h$ **do**
    Apply algorithm 1 in the horizontal direction, across all rows, to bring each pair of qubits connected by an edge of color $c$ in the horizontal direction together.
    Apply a CZ gate on each pair of neighboring qubits.

```
// Parallel syndrome extraction for d − 1 rounds
```
**for** $i \leftarrow 1$ **to** $d-1$ **do**
    **for** *direction* $f \in$ *{vertical, horizontal}* **do**
        Apply a Hadamard on all data qubits.
        **if** $f ==$ *vertical* **then**
            Measure any old $Z$ ancillas in the $X$ basis and prepare a fresh ancilla in $|+\rangle$ for each $Z$ stabilizer.
        **else**
            Measure any old $X$ ancillas in the $X$ basis and prepare a fresh ancilla in $|+\rangle$ for each $X$ stabilizer.
        Move all $X$ and $Z$ ancilla qubits into their appropriate positions in Fig. 4.6.
        **for** *color* $c \in \mathcal{C}_f$ **do**
            Apply algorithm 1 in direction $f$, across all columns, to bring each pair of qubits connected by an edge of color $c$ in direction $f$ together.
            Apply a CZ gate on each pair of neighboring qubits.

```
// Z stabilizers of the final round
```
Move $X$ ancilla qubits out of the grid region and measure them in the $X$ basis.
**for** *color* $c \in \mathcal{C}_v$ **do**
    Apply algorithm 1 in vertical direction, across the whole grid, to bring each pair of qubits connected by an edge of color $c$ in vertical direction together.
    Apply a CZ gate on each pair of neighboring qubits.
Move $Z$ ancilla qubits out of the grid region and measure them in the $X$ basis.

the fact that the derivative of acceleration (jerk) causes motional heating, and to maintain a fixed amount of heating in a move, the integrated heating (with dimensions of acceleration)

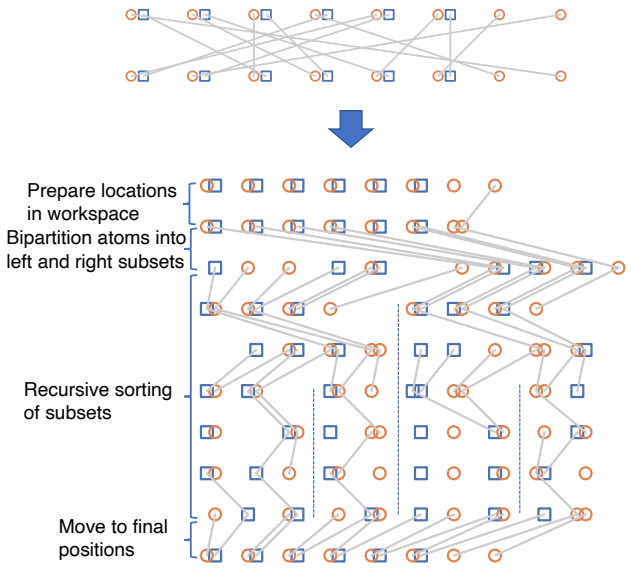


Figure 4.3: Efficient non-intersecting rearrangement in log-depth. By using a divide and conquer algorithm, we can perform an arbitrary 1D rearrangement in depth logarithmic in the number of qubits. Repeating this across the array yields an efficient implementation of the desired rearrangements, without requiring intersecting atom trajectories that may lead to additional loss and decoherence. Here, we illustrate the full set of movements required in a small example. Similar to the earlier figures, blue squares indicate classical checks and orange circles indicate classical bits. When a blue square and orange circle are moved to be neighboring at the end of the rearrangement, they execute an entangling gate. The top panel indicates the desired change of configuration, where the ordering of neighboring atoms in the top row needs to be modified to that in the bottom row via parallel rearrangement, as illustrated by the crossing gray lines. The bottom figure illustrates how we decompose the arbitrary rearrangement into a non-crossing rearrangement, where the gray lines no longer intersect.

should remain constant, thus giving a time scaling with distance that is the same as a constant acceleration profile.

The total time required for one layer of full rearrangement before each layer of entangling gates in the HGP syndrome extraction circuit (see Alg. 7) is thus

$$t_{\text{rearrange}} = 2k\tau_t + \sum_{i=0}^{k} \left( \sqrt{\frac{6 \cdot 2^i d}{a_p}} + \sqrt{\frac{3 \cdot 2^i d}{a_p}} \right) < 2\tau_t \log L + (3 + 2\sqrt{2}) \sqrt{\frac{6Ld}{a_p}}.$$

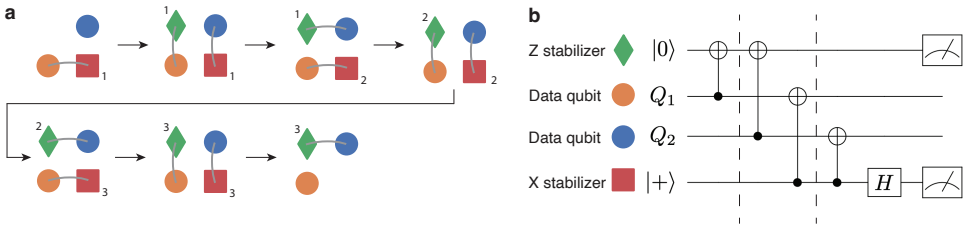


Figure 4.4: Illustration of ordering of operations in pipelined syndrome extraction. (a) Successive steps of entangling gates for the pipelined product coloration circuit described in Alg. 8, with $d = 3$ rounds of syndrome extraction. Numbers at the corners of the $X$ and $Z$ ancilla qubits denote the round of syndrome extraction they correspond to. (b) Illustration of a local circuit that data qubits and ancilla qubits see, with dashed lines indicating different circuit moments. As the $X$ stabilizer interacts with both qubits before the $Z$ stabilizer, the syndrome extraction order is valid. Similar analysis can be performed for the commutation relations with the next round of ancilla qubits.

Recent experiments have demonstrated parameters on the order of $\tau_t = 50\,\mu$s, $a_p = 0.02\,\mu\text{m}/\mu\text{s}^2$, $d = 5\,\mu$m. For a moderately sized code consisting of 10000 qubits (including both data and ancilla qubits), we have $L \approx 100$. The total trap transfer time is 0.7 ms and the atom movement time is 2.3 ms, for each gate layer. Assuming a $(3, 4)$-biregular graph for the underlying classical expander code, we need 8 rounds of rearrangement to measure one full round of stabilizers for Alg. 8, resulting in a total time overhead of 3 ms per rearrangement layer and 24 ms for a full round of syndrome extraction, a small fraction of the coherence time $T_c > 10$ s that has been demonstrated in neutral atom arrays [257, 258, 259]. This timescale is somewhat longer than the typical readout timescales, and thus the code cycle time will be dominated by the rearrangement time. The ancilla measurements can be pipelined to happen simultaneously with the atom rearrangements of the following round, and therefore will not increase the run time.

For LP codes, we need to first "flatten" the code layout before implementing the parallel rearrangement scheme. Since we use a fixed 3 by 5 protograph, the flattened rectangular array has dimensions roughly of $2n/8$ by 8. This can be achieved in log depth using a similar

divide-and-conquer algorithm that flattens the code by half each time. For example, as shown in Fig. 4.6(b), we flatten the codes vertically before establishing vertical connections between atoms via row permutations. The vertical connectivity is then the same as an HGP code, and we can use the efficient 1D rearrangement scheme described earlier. Therefore, we estimate the rearrangement time for LP codes by setting $L$ to $n/8$ in Eq. (4.5).

The rearrangement time in Eq. (4.5) determines the idling errors between sequences of entangling gates in a syndrome extraction circuit. In general, $t_{\text{rearrange}}(n)$ in Eq. (4.5) is a function of the code size $n$, as $L$ is a function of $n$. Setting the gate error rate $p_g$ as the characteristic physical error rate of our noise model, we rescale the idling error rate $p_i(n)$ together with $p_g$:

$$p_i(n) = t_{\text{rearrange}}(n)/T_c \times p_g/0.005, \tag{4.5}$$

where $T_c$ is the atom coherence time and 0.005 is the current CZ gate fidelity demonstrated in Ref. [260]. Note that the rescaling of $p_i(n)$ with $p_g$ in Eq. (4.5) can be justified by the fact that idling errors can improve together with the gate errors as hardware improves. For example, further improvements in coherence time can be achieved by detuning trapping light further and better magnetic field shielding. Even without further improvements, coherence times as long as 50s have been demonstrated in neutral atom systems [259], sufficient for our analysis at $p_g = 0.1\%$. For all the numerical estimations in this work, we use the upper bound for $t_{\text{rearrange}}(n)$ in Eq. (4.5) and use the experimental parameters listed below Eq. (4.5), with $T_c = 10$s.

In Supplement of Ref. [237], we numerically verify that the effect of the idling errors can be approximated by rescaling the gate errors $p_g \to p_g + 3p_i(n)$ using the product coloration circuit. Replacing the rescaled $p_g$ in the subthreshold scaling for HGP codes in Eq. (4.11), we can examine the effect of the idling errors on achievable logical failure rates. As shown in Fig. 4.5, the LFRs are first exponentially suppressed by the code size $n$ when $n$ is small and $3p_i(n) \ll p_g$, and gradually saturate and then increase as $3p_i(n) > p_g$ and finally approaches

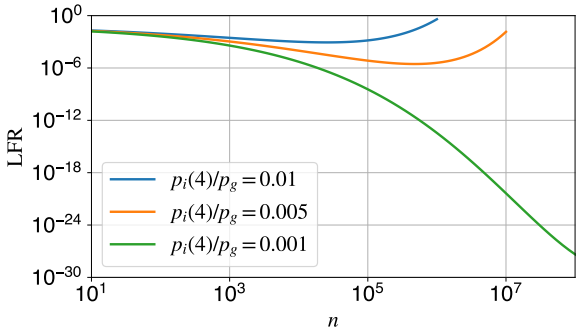


Figure 4.5: Achievable logical failure rates of the HGP codes with different idling error strengths.
We characterize the idling error strengths as the relative ratio between the idling error rate $p_i(n)$ at $n = 4$ and the gate error rate $p_g$. This idling error strength can potentially be reduced by, e.g., increasing the coherence time and accelerating the atom shuttling.

the gate error threshold. Using the relevant experimental parameters, the idling errors are negligible for $n$ up to $\sim 10^7$ and the LFRs can go below $10^{-24}$ (see the green curve), which already suffices for implementing practical quantum algorithms. Note that for even larger sizes, concatenation with another code can be employed to extend the effective coherence time [268, 269] and further suppress $p_i(n)$.

Note that our syndrome extraction circuit is not $2D$-local, since coupling two qubits does not incur an error proportional to their Euclidean distance. The effective syndrome extraction circuit depth is $\Delta[1 + 3p_i(n)/p_g]$, where the constant $\Delta$ is the entangling gate depth. In the regime where $3p_i(n) \ll p_g$, the circuit is quasi-nonlocal, with a constant circuit depth $\Delta$, whereas in the regime where $3p_i(n) \gg p_g$, the circuit is effectively 4D-local for the HGP codes according to the bound on the circuit depth in finite dimensions in Ref. [251]. Thus, although with the inclusion of idling errors, the logical failure rate cannot be arbitrarily suppressed asymptotically according to the no-go theorems in Ref. [251] and Ref. [270], constant overhead and fault tolerance can still be achieved at physically relevant sizes by utilizing the quasi-nonlocal connectivity in atom arrays.

To conclude, we shown in above that the total rearrangement time for implementing HGP codes scales as $O(\sqrt{L}) = O(\sqrt[4]{n})$, where $L$ is the length of the 2D atom array, similar to

the scaling for a constant acceleration trajectory. We have estimated that for a moderately-sized HGP code with 10000 qubits, each rearrangement layer between gates requires 3 ms, a small fraction of the coherence time $> 10$ s that has been demonstrated in neutral atom arrays [257, 258, 259]. These timescales can be significantly improved through innovations in optical technologies and compilation. Ancilla measurements can be pipelined, thereby not causing any increase in the cycle time. For very large codes, where idling errors are no longer negligible, concatenation with another code can also be employed to extend the effective coherence time [268]. Note that HGP codes based on expanding classical LDPC codes (also called quantum expander codes) have the single-shot property, and therefore only a single round of syndrome extraction is required to be fault-tolerant [271, 272, 145]. Although we have focused on the implementation of HGP codes, other families of qLDPC codes, such as lifted product (LP) codes, can also be implemented by adapting similar ideas.

### *4.3.3 qLDPC memory*

#### Code construction

We first describe how we construct the qLDPC codes that we will consider as quantum memories. We primarily focus on two families of qLDPC codes, which we describe in detail in this section. We leave the extension of these results to other families, such as asymptotically good codes [234, 235, 236, 245, 273, 274, 275], to future work.

The first family of codes are hypergraph product (HGP) codes [232], formed from the product of two classical LDPC codes. We have provided a geometric sketch of the code properties in the previous sections, and instead focus here on an alternative, algebraic description of the codes and provide more details of their code properties. Algebraically, if we denote the parity check matrix (where rows describe bits that should sum to an even number in the absence of errors) of the two underlying classical codes as $H_1 \in \mathbb{F}_2^{r_1 \times n_1}$, $H_2 \in \mathbb{F}_2^{r_2 \times n_2}$, then

the $X$ and $Z$ stabilizer check matrices for the HGP code can be written as

$$H_x = \begin{pmatrix} H_1^T \otimes I_{r_2} & I_{n_1} \otimes H_2 \end{pmatrix}, \tag{4.6}$$

$$H_z = \begin{pmatrix} I_{r_1} \otimes H_2^T & H_1 \otimes I_{n_2} \end{pmatrix}. \tag{4.7}$$

For classical $[n_i, k_i, d_i]$ linear codes defined by $r_i = n_i - k_i$ linearly-independent checks $(i = 1, 2)$, the resulting quantum code has parameters $[\![n_1 n_2 + r_1 r_2, k_1 k_2, \min\{d_1, d_2\}]\!]$. The surface code is a special case of hypergraph product codes, with the classical codes being repetition codes. However, by instead choosing classical codes with good vertex expansion , where $k_i = \Theta(n_i)$, $d_i = \Theta(n_i)$, the resulting quantum code (known as quantum expander codes) encodes a linear number of logical qubits $k = \Theta(n)$ and has distance $d = \Theta(\sqrt{n})$ [271]. Such classical expander codes can be obtained asymptotically, for example, from random biregular Tanner graphs, and will have sufficient vertex expansion with high probability [276]. Logical operators are inherited from the underlying classical code, and one can choose a basis such that each logical operator has support in only a single row or column [256, 277].

In this work, we follow the procedure of Ref. [278] and construct HGP codes by taking the hypergraph product of classical LDPC codes defined by $(3, 4)$-regular Tanner graphs, i.e. bipartite graphs with degree-3 bit nodes and degree-4 check nodes. By increasing the size of the graph, we obtain a family of HGP codes with a constant encoding rate $k/n \geq 0.04$. For each code size, we pick the classical code with the largest distance, Tanner graph girth at least 6 (length of the shortest cycle in the Tanner graph, obtained through rejection sampling without performing edge swaps), and the largest spectral gap (the gap between the largest two singular values of the classical check matrices) from randomly generated instances. It is known that the hypergraph product of vertex-expanding classical codes gives HGP codes that satisfy the confinement property, and support single-shot QEC [272, 271, 279].

The second family of codes we consider are quasi-cyclic lifted product (LP) codes [233,

235, 269], which can be viewed as a hypergraph product code followed by a symmetry reduction to reduce the number of required qubits [12]. Algebraically, a quasi-cyclic LP code is obtained from two base protographs (analogs of the classical codes in the HGP construction) associated with two base matrices $\mathbf{B}_1$ and $\mathbf{B}_2$ over the quotient polynomial ring $\mathbb{R}[x]/(x^l-1)$ [234]. Suppose the two base matrices are of size $m_{B_1} \times n_{B_1}$ and $m_{B_2} \times n_{B_2}$, respectively. We obtain two matrices (over the same polynomial ring) $\mathbf{B}_x$ and $\mathbf{B}_z$ by taking the hypergraph product:

$$\begin{aligned} \mathbf{B}_x &= \begin{pmatrix} \mathbf{B}_1^T \otimes \mathbf{I}_{m_{B_2}} & \mathbf{I}_{n_{B_1}} \otimes \mathbf{B}_2 \end{pmatrix}, \\ \mathbf{B}_z &= \begin{pmatrix} \mathbf{I}_{m_{B_1}} \otimes \mathbf{B}_2^T & \mathbf{B}_1 \otimes \mathbf{I}_{n_{B_2}} \end{pmatrix}. \end{aligned} \tag{4.8}$$

The $X$ ($Z$) check matrix $H_x$ ($H_z$) is then obtained by replacing each entry of $\mathbf{B}_x$ ($\mathbf{B}_z$) with its matrix representation as $l$ by $l$ circulant matrices, a process known as a lift. Specifically, we replace $x$ with a $l$ by $l$ square matrix $P$, where $P_{ij} = \delta_{i,j+1}$, and replace each polynomial in $x$ with the same polynomial in $P$. The code size is $N = l(n_{B_1}n_{B_2} + m_{B_1}m_{B_2})$ and the number of $X$ and $Z$ checks are $M_x = ln_{B_1}m_{B_2}$ and $M_z = lm_{B_1}n_{B_2}$, respectively. The encoding rate is lower bounded by $(N - M_x - M_z)/N = (n_{B_1}n_{B_2} + m_{B_1}m_{B_2} - m_{B_1}n_{B_2} - n_{B_1}m_{B_2})/(n_{B_1}n_{B_2} + m_{B_1}m_{B_2})$.

We can also describe the above construction using graphs. As an example, Fig. 4.6(b) shows a LP code using a 3 by 5 protograph associated with a base matrix $\mathbf{B} \in \{\mathbb{R}[x]/\left(x^2-1\right)\}^{3\times 5}$. The checks and bits of the protograph are illustrated by the big dashed nodes. The $i$-th dashed check node is connected to the $j$-th dashed bit node if $\mathbf{B}_{ij}$ is non-zero. A lift of the protograph is done by replacing each dashed node with its two inner solid nodes, and setting up the connectivity between the inner nodes according to the matrix representation of each ring element $\mathbf{B}_{ij}$. Eq. (4.8) corresponds to taking the hypergraph product between the protograph and itself, obtaining a grid of dashed nodes. Similar to the hypergraph product code, the connectivity between the dashed nodes (the entries of $\mathbf{B}_x$ and $\mathbf{B}_z$) is inherited

from $\mathbf{B}$. Then the qubits and the quantum checks are given by the inner nodes after the lift, and their connectivity is given by the matrix representation of $\mathbf{B}_x$ and $\mathbf{B}_z$. An important feature of the LP codes is that they still have some remaining product structure even after the lift. As shown in Fig. 4.6(b), when flattening the inner nodes vertically (horizontally), the vertical (horizontal) connectivity between the qubits and the checks for each column (row) is the same as the left (top) lifted classical code.

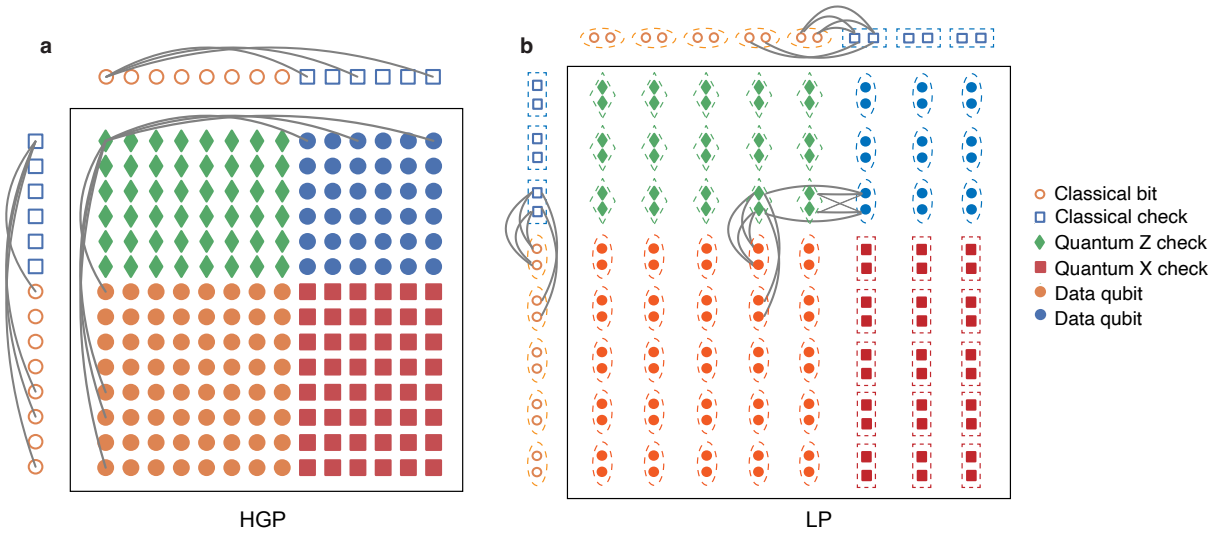


Figure 4.6: Product structure of HGP codes and LP codes. (a) The HGP code is constructed from two classical LDPC codes. The classical codes are illustrated on the left and top, where circles indicate classical bits and squares indicate classical checks. A data qubit is placed at each intersection of two classical bits (filled orange circles) and of two classical checks (filled blue circles). $Z$ stabilizer generators are placed at the intersection of horizontal bits and vertical checks, while $X$ stabilizer generators are placed at the intersection of horizontal checks and vertical bits. Each stabilizer is connected to data qubits along the same row or column, with the same connectivity as the classical codes, as illustrated for the top left Z stabilizer. We have omitted other connections for ease of visualization. (b) The LP code is constructed by taking a lift over the hypergraph product of two classical protographs. The protographs and their hypergraph product are indicated by the dashed nodes and the lift is illustrated by the multiple inner nodes within each dashed node. The inner connectivity between two dashed nodes is given by the circulant-matrix representation of the ring elements in Eq. (4.8). When flattening the inner nodes vertically (horizontally), the vertical (horizontal) connectivity between the qubits and the checks for each column (row) is the same as the left (top) lifted classical code.

For the LP codes constructed in this work, we choose a base matrix of dimension 3 by

5, where all entries are monomials, and obtain a family of codes with sizes up to 1428 by increasing the lift size $l$ from 16 to 42. The classical parity checks are optimized by choosing the base matrix entries over the quotient polynomial ring to obtain the best classical distance for the particular lift size $l$. The choice of the base matrix entries is also such that the girth is at least 8, and the distance of the lifted qLDPC codes matches the designed classical distances with a high probability. Allowing multiple polynomial terms for each base matrix entry and more protographs of different sizes gives more flexibility in LP code design and improves their distances. Such general code constructions and their impact on the proposed scheme will be explored in future work. Here, we explicitly provide the classical base matrices used to construct the four LP codes used in this work. Denoting $\mathbf{B}_d^l$ as a base matrix with a lift size $l$ and a classical code distance $d$ after the lift, the base matrices are

$$\mathbf{B}_{12}^{16} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & x^2 & x^4 & x^7 & x^{11} \\ 1 & x^3 & x^{10} & x^{14} & x^{15} \end{bmatrix}, \quad \mathbf{B}_{16}^{21} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & x^4 & x^5 & x^7 & x^{17} \\ 1 & x^{14} & x^{18} & x^{12} & x^{11} \end{bmatrix}, \tag{4.9}$$

$$\mathbf{B}_{20}^{30} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & x^2 & x^{14} & x^{24} & x^{25} \\ 1 & x^{16} & x^{11} & x^{14} & x^{13} \end{bmatrix}, \quad \mathbf{B}_{24}^{42} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & x^6 & x^7 & x^9 & x^{30} \\ 1 & x^{40} & x^{15} & x^{31} & x^{35} \end{bmatrix}. \tag{4.10}$$

These codes have an encoding rate lower bounded by 2/17 by counting the number of qubits minus the number of checks. For all the resource estimates involving these LP codes, we use $k \approx 0.38n^{0.85}$ that fits well on the above four codes. The quantum code distances are upper bounded by the classical code distances of the above (lifted) base matrices. Through an extensive search for minimum-weight logical operators using a GAP package [280], we believe these upper bounds are tight.

## Memory performance

We now analyze the fault-tolerant implementation of HGP and LP codes as a robust quantum memory. In the Supplement [281], we prove the existence of a circuit-level single-shot threshold for qLDPC codes with the linear confinement property [272], under a single-ancilla syndrome extraction circuit and a depolarizing noise model where error rates do not scale with instance size. The linear confinement property, which requires that for sufficiently small Pauli errors, the weight of the syndrome increases linearly with the (reduced) weight of the errors, holds for various qLDPC codes decodable by the small-set-flip-type decoders, including HGP codes with sufficiently expanding classical codes [281].

We next supplement this theoretical understanding with numerical simulations of HGP and LP codes at practically-relevant instance sizes [282], where we find competitive thresholds and LFRs for both codes. The details of the code constructions are shown in Fig. 4.2 and Methods. We use the product coloration circuit (Alg. 7) for syndrome extraction, a variation of the coloration circuit [250] that is more compatible with the product structure of the codes and hardware. The circuit uses a single ancilla for each stabilizer generator and has entangling-gate depth 16 and 20 for the HGP and LP codes we consider.

We construct a space-time circuit-level decoder based on the belief propagation and ordered statistics decoding (BP+OSD) algorithm [233, 283, 284, 285]. Specifically, for a QEC circuit with multiple cycles, we construct a bipartite decoding graph [284, 30] over a certain number of cycles, where the check nodes and variable nodes are associated with parities of stabilizer measurement outcomes and circuit faults, respectively. We apply BP decoding on this decoding graph to infer the circuit fault locations in all noisy code cycles, and apply the BP+OSD decoder in the final round to project back into the code space. For all memory simulations, we use space-time decoding graphs over three cycles, irrespective of the code size. Note that we observe improved logical error rates by increasing the number of cycles for decoding, indicating a tradeoff between the accuracy and the speed of our

decoder. We also remark that the BP decoder has similar performance as the BP+OSD decoder for decoding the noisy cycles while being much faster. Crucially, compared to prior phenomenological decoders that used a simpler decoding graph involving only independent data and measurement errors as bits and decoded over only one code cycle [278, 251], our space-time circuit-level decoder takes the full circuit details into account and can perform joint decoding on multiple QEC cycles, improving the threshold [281]. Moreover, the space-time decoding is also crucial for our simulations of logical operations in the next section, where repetitions of syndrome measurements are required for fault tolerance. See Methods of Ref. [237] for details of the decoder.

In the Supplement of Ref. [237], we find that the HGP and LP codes have a threshold of 0.63% and 0.62% respectively, under a depolarizing error model without idling errors. In the sub-threshold regime, assuming the absence of a decoding error floor, the LFRs of the two codes are well approximated by [281]

$$\begin{aligned} \mathrm{LFR(HGP)} &= 0.07(p_g/0.006)^{0.47n^{0.27}}, \\ \mathrm{LFR(LP)} &= 2.3(p_g/0.0066)^{0.11n^{0.60}}, \end{aligned} \tag{4.11}$$

indicating that finite-size LP codes have better subthreshold scaling than HGP codes. When also considering idling errors $p_i(n)$ associated with the atom rearrangement time overhead, which grow as $O(n^{1/4})$ for HGP codes and $O(n^{1/2})$ for LP codes (see Methods for details of the idling error model and the expression for $p_i(n)$ in Eq. (4.5)), there is no asymptotic code threshold [251]. However, we numerically observe that the effect of adding the idling errors can be approximated by rescaling the gate error $p_g \to p_g + 3p_i(n)$ using the product coloration circuit (see Supplement of Ref. [237]). The idling errors have a negligible contribution when $3p_i(n) \ll p_g$, which is the case for current experimental parameters and practically relevant code sizes (see Methods). Therefore, although there is no asymptotic threshold, constant overhead and fault tolerance can still be achieved at physically-relevant sizes by utilizing the

quasi-nonlocal connectivity in atom arrays. In Fig. 4.7(a) and (b), we show the simulated LFR versus the bare two-qubit gate error rate $p_g$ (which we refer to as the physical error rate for simplicity) for the HGP and LP codes, with the idling errors included and rescaled together with $p_g$. We find that good LFRs and subthreshold scaling are maintained for both codes in the presence of idling errors. The smallest instances of these codes that we simulate, involving hundreds of physical qubits, are readily within reach with experimentally demonstrated system sizes [286] and control capabilities [47, 22], and larger instances can be realized with improvements in trap power.

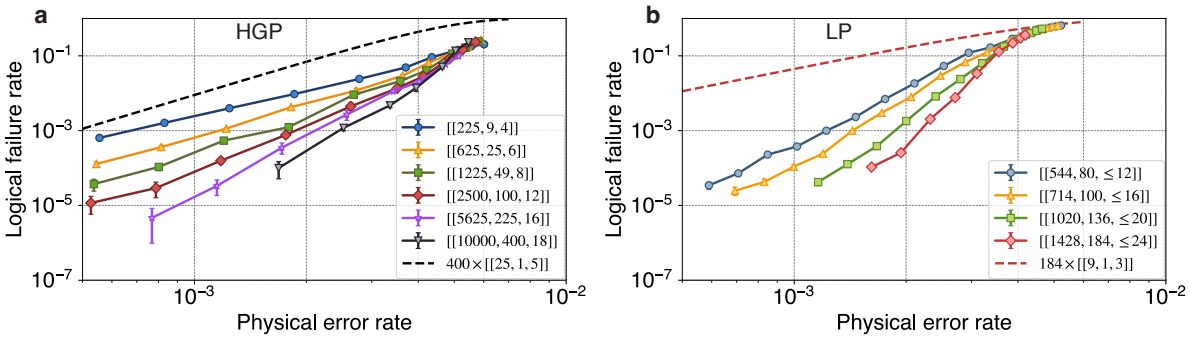


Figure 4.7: qLDPC memory performance. Logical failure rates as a function of physical error rate for the qLDPC memory using HGP codes (a) and LP codes (b), for a depolarizing error model that includes idling errors that increase with the code sizes. For the HGP codes, we take the hypergraph product of classical codes associated with random $(3, 4)$-regular bipartite graphs that have good expanding properties, which have an encoding rate lower bounded by $1/25$. For the LP codes, we choose 3 by 5 base matrices over a quotient polynomial ring and obtain a family of codes with sizes up to 1428 by increasing the lift size and optimizing the matrix entries. These LP codes have an encoding rate lower bounded by $2/17$, maintaining a higher encoding rate as well as better distances than HGP codes of the same sizes. See Methods for details of the code construction. We use a product coloration circuit for syndrome extraction (Alg. 7) and a space-time circuit-level decoder based on BP+OSD that decodes over every 3 code cycles, regardless of the code sizes. The LFRs are calculated using $\text{LFR} = 1 - (1 - p_L)^{1/m_c}$, where $p_L$ is the total logical failure probability over $m_c$ code cycles. We choose $m_c = 42$ for physical error rates below $4 \times 10^{-3}$ and $m_c = 12$ for physical error rates above $4 \times 10^{-3}$. $p_L$ is obtained from the Monte Carlo simulations, with a standard deviation $\sqrt{p_L(1 - p_L)/M}$, where $M$ denotes the number of samples. We also compare the largest HGP/LP code to surface codes of similar sizes and encoding rates (see the dashed lines), as discussed in the main text.

We now use Eq. (4.11) to estimate the total number of data and ancilla qubits $N$ needed

| Logical qubits | 25 | 80 | 180 | 400 |
|---|---|---|---|---|
| Logical failure rate | $10^{-3}$ | $10^{-4}$ | $2 \times 10^{-5}$ | $6 \times 10^{-6}$ |
| HGP code physical qubits (improvement over surface code) | 1235 (1×) | 4606 (2.8×) | 10760 (4.0×) | 19600 (6.9×) |
| LP code physical qubits (improvement over surface code) | 851 (1.4×) | 1367 (9.4×) | 2670 (16.2×) | |

Table 4.2: Total number of data and ancilla qubits required to reach target numbers of logical qubits and logical failure rates using HGP codes and LP codes, compared to using surface codes. We use ($\beta\times$) to indicate a $\beta$ times qubit saving compared to the surface codes by using the corresponding qLDPC codes. The physical error rate is set to be $10^{-3}$. The estimates for the HGP and LP codes are based on the numerical data in Fig. 4.7.

to reach a target number of logical qubits $k$ and a target LFR, demonstrating significant advantages of HGP and LP codes over the surface code. We rescale the gate error $p_g$ to approximate the presence of idling errors, and compare to the surface code subthreshold scaling formula $\text{LFR(surface)} = 1 - (1 - P_0)^k$, where $P_0 = 0.03(p_g/0.011)^{\lceil \lfloor \sqrt{n/k} \rfloor /2 \rceil}$, from Ref. [17, 41]. In Table 4.2, we present such estimates for finite-size HGP and LP codes that we directly simulated (see Fig. 4.7) at a realistic physical error rate of $10^{-3}$. Both HGP and LP codes outperform surface codes with as few as 25 logical qubits. At a moderate scale of less than 200 logical qubits, LP codes with less than 3000 physical qubits already achieve a qubit saving of over an order of magnitude. In the lower panel of Fig. 4.1, we estimate the space overhead of the HGP codes at a larger scale by extrapolation. We find that HGP codes can also achieve a qubit saving of over an order of magnitude at a scale of 1000 logical qubits and $10^5$ physical qubits.

### *4.3.4 Logical operations*

We now present a scheme inspired by Ref. [253] for performing fault-tolerant logical operations, and perform the first numerical simulation of logical gate performance on qLDPC codes using the space-time decoder developed above. We find that the threshold and logical performance remain almost unchanged when performing gates, indicating that the high

threshold and low overhead can be maintained for fault-tolerant quantum computation.

Our scheme is illustrated in Fig. 4.8(a). We teleport the logical information between the qLDPC memory and ancillary topological codes using a measurement-based circuit (Fig. 4.8 (b)), where the prescribed logical measurements are implemented using lattice surgery [10, 253, 11, 287]. Universal logical operations can then be performed in the topological codes using standard techniques. Our approach is applicable to any CSS qLDPC code, as proved in the Supplement of Ref. [237]. The choice of qLDPC code, whether it's an HGP or an LP code, does not impact the feasibility of our ancilla encoding for logical operations. Since each topological code patch and teleportation ancilla patch is much smaller than the qLDPC patch, as long as the number of such patches used is $O(\sqrt{k}/\mathrm{poly}\log(kT))$, the space overhead from ancilla patches will be sub-leading. Note that the scheme in Ref. [253] can also be used for teleportation between a qLDPC block and topological codes. In comparison, our scheme reduces the ancilla patch size by half and thus has smaller space overhead.

As an example, we consider the teleportation from a surface code patch to a HGP patch. As illustrated in Fig. 4.8(a), this is mediated by an additional ancilla logical qubit, formed by a hypergraph product of two classical codes associated with the logical operators of the two code patches, which enables joint logical measurements by merging code blocks. We perform $\min\{d_{\mathrm{comp}}, d_{\mathrm{mem}}\}$ rounds of syndrome extraction to ensure tolerance against measurement errors. We describe the scheme in more detail in Sec. 4.3.4, and prove the fault-tolerance of the scheme under data errors in the Supplement of Ref. [237].

To validate this method, we perform circuit-level simulations of the above teleportation, enabled by the space-time decoder described in the previous section. In our simulations, we use the teleportation circuit depicted in Fig. 4.8(b) with an initial state of $|\psi\rangle = |0\rangle$, focusing on errors during the merge and split operations in the $XX$ measurement to evaluate the performance of lattice surgery building blocks. As shown in Fig. 4.8(c), we observe similar LFRs and threshold crossings for the teleportation as for the memory. Note that each logical

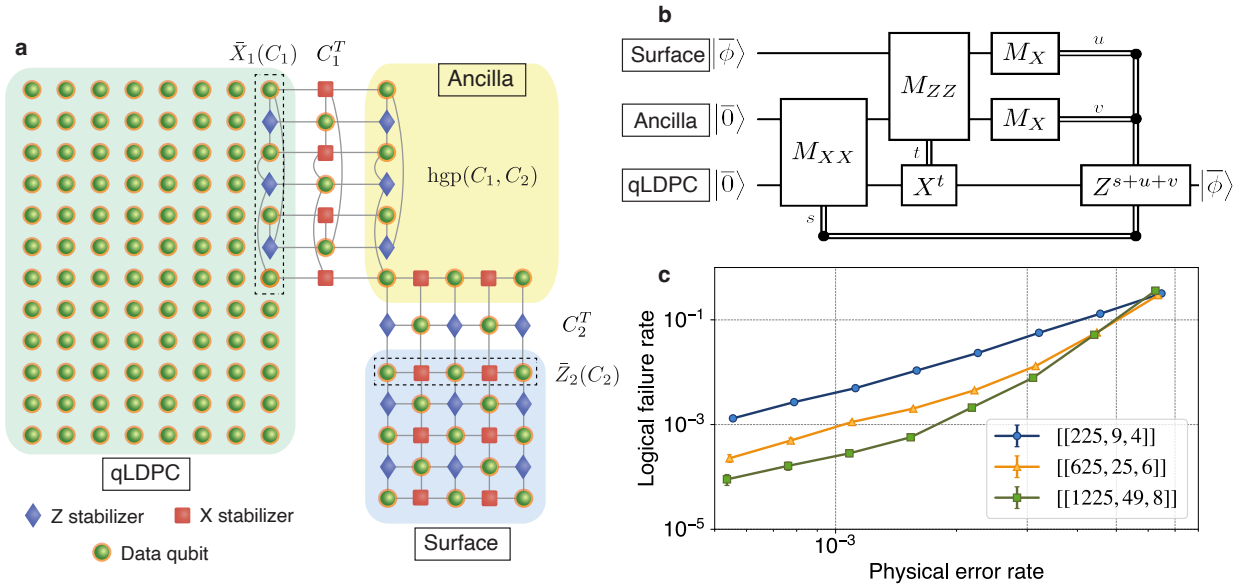


Figure 4.8: Fault-tolerant teleportation from surface to qLDPC code. We identify the logical $Z$ operator of the surface code, $\overline{Z}_2$, and the logical $X$ operator of one of the qLDPC code's logical qubits, $\overline{X}_1$. We associate these two logical operators with classical codes $C_1$ and $C_2$, by mapping the qubits supporting the logical operators to bits and the corresponding incident stabilizer generators to classical checks. We then construct an ancilla patch as the hypergraph product of $C_1$ and $C_2$, where the columns resemble $C_1$ and the rows resemble $C_2$. Direct lattice surgery between this ancilla patch and each of the surface and HGP codes is conducted by matching similar boundaries associated with the chosen logical operators. In between similar boundaries, an extra array of ancillary qubits and checks associated with the transpose of the classical code is inserted to mediate the surgery. All checks in the merged code commute with each other, as required, and the product of the stabilizers associated with the checks of the transposed code gives the required joint logical measurement, as in the case of standard surface code lattice surgery [10]. We elaborate on the lattice surgery procedure in Methods. (b) Measurement-based teleportation circuit [11]. Logical state $|\overline{\phi}\rangle$ is teleported from the surface code to one of the qLDPC's logical qubits. The joint Clifford measurements are conducted through lattice surgery as illustrated on panel (a). (c) Simulated logical failure rates (per code cycle) of the teleportation. Noise is added during the merge and split steps of the $XX$ lattice surgery. We decode with the same space-time circuit-level BP+OSD decoder used in the memory simulations. The corresponding surface codes paired with the three HGP codes have distances 3, 5, and 7. We record a logical failure if there is an error in any of the logical qubits of the qLDPC code after the teleportation scheme is complete. Denoting the total logical failure probability as $p_L$, we calculate the logical failure rate (per code cycle) as $\mathrm{LFR} = 1 - (1 - p_L)^{2d}$, where there are $2d$ cycles during the noisy $XX$ lattice surgery, and $d$ denotes the minimal code distance. The plotted physical error rates are rescaled to account for idling errors, as explained in Methods. Error bars obtained identically to Fig. 4.7.

qubit will fail at a rate much lower than the LFR of the entire block shown in Fig. 4.8(c), and the system is far below break-even at $10^{-3}$ physical error rate, even for the smallest code of size 225. This demonstrates that the high threshold and low resource overhead of the qLDPC code can be maintained at the computation level.

## Details of the teleportation scheme

Here, we describe in detail our teleportation scheme between the qLDPC code and the surface code. We select a logical $\bar{X}_1$ operator for the qLDPC code of minimum weight. This ensures it contains no sublogicals, i.e. inequivalent logical operators contained in its support, so that the scheme is fault-tolerant under data errors (see Supplement of Ref. [237] for proof). We then associated the qubit support of $\bar{X}_1$ of the qLDPC code ($\bar{Z}_2$ of the surface code) to the bits of a classical code $C_1$ ($C_2$), and associated the $Z$ ($X$) stabilizers of the qLDPC (surface) code with support on $\bar{X}_1$ ($\bar{Z}_2$) to the checks of $C_1$ ($C_2$). Denoting $H^1$ ($H^2$) as the check matrix for $C_1$ ($C_2$), $H^1_{ij}(H^2_{ij}) = 1$ if the $i$-th $Z$ ($X$) stabilizer checks the $j$-th qubit of $\bar{X}_1$ ($\bar{Z}_2$). We construct an ancilla code patch as a HGP code taking the hypergraph product of $C_1$ and $C_2$, which encodes a single logical qubit with a logical $X$ and $Z$ representative associated with the bits of $C_1$ and $C_2$, respectively.

Lattice surgery between the ancilla patch and the qLDPC (surface) code is realized by merging and splitting along $C_1$ ($C_2$), assisted by an extra array of ancillary qubits (see Fig. 4.8(a)). For a classical linear code $C$ with check matrix $H$, denote by $C^T$ its transposed code defined by the check matrix $H^T$. Taking the qLDPC-ancilla surgery as an example, we insert an extra array of $X$ stabilizers (initialized in $|+\rangle$) and qubits (initialized in $|0\rangle$) in the middle, associated with the checks and the bits of $C_1^T$, respectively. During the code merging, the Z stabilizers of the qLDPC and the ancilla patch associated with the $i$-th check of $C_1$ are each modified to include the middle qubit associated with the $i$-th bit of $C_1^T$; The middle $X$ stabilizer associated with the $j$-th check of $C_1^T$ checks the two qubits of the qLDPC

and surface codes associated with the $j$-th bit of $C_1$ as well as some middle qubits given by the incident relation of $C_1^T$. It is easy to verify that all the new stabilizers commute as the added/modified qubits and checks across the merged boundary form an HGP code with $C_1$ and a length-2 repetition code locally. The product of the middle $X$ stabilizers give the joint logical operator to measure. See Supplement of Ref. [237] for more algebraic details of the above lattice-surgery scheme, and a proof of its fault tolerance under data errors.

We perform the numerical simulations using Stim [282], and use the same space-time decoder described in the previous section for decoding. For ease of simulation, we add circuit depolarizing noise with no idling errors on the joint logical $XX$ measurement part of the teleportation circuit, in both the merging and splitting steps, and rescale the gate error rate $p_g \to p_g + 3p_i$ to obtain Fig. 4.8(c). As the merged code no longer supports single-shot error correction, we perform $d$ rounds of QEC after each of the merge and split steps of the $XX$ surgery, where $d = \min(d_1, d_2)$ is the distance of the teleportation scheme and $d_1$ and $d_2$ are the distances of the qLDPC code and surface code, respectively. We expect the logical failure rate of the $ZZ$ logical measurement to be nearly the same as that of the qLDPC memory if we use a computation surface code with a distance larger than that of the qLDPC code. For the logical $M_X$ measurement, we simulate a noiseless destructive measurement.

We found a gate error threshold (without idling errors) of $\sim 0.7\%$ by looking at the crossing of $d = 3, 5, 7$ schemes, which were constructed with $(d_1, d_2) = (4, 3), (6, 5), (8, 7)$, respectively. We attribute the slight increase in the threshold compared to the memory to the increase in the number of code cycles used by the space-time decoder (the memory simulations are decoded using only three cycles).

Note that the lattice surgery approach in Ref. [253] can also be used for teleportation between a qLDPC code and a surface code. Their approach essentially uses an ancilla patch formed by the hypergraph product of a length-$d$ repetition code and a "union" of $C_1$ and $C_2$ associated with the logical operators of the qLDPC code and the surface code respectively,

and directly performs a joint logical measurement on the qLDPC and surface code. The ancilla patch has size $2d^2$ (if using minimum-weight logical operators), which is twice as larger as our ancilla. Compared to their approach, our scheme has a lower space overhead but a larger temporal overhead overall.

### *4.3.5 Discussion and outlook*

By demonstrating large space overhead savings in practical regimes, good performance of logical gate operations, and providing a blueprint for their implementation with existing hardware capabilities, our work brings the use of high-rate qLDPC codes for fault-tolerant quantum computation into the practical regime.

Although our scheme shows a significant reduction in space overhead, it still carries substantial time overhead. This is because fault tolerance during gate operation requires $\Theta(d)$ QEC cycles, and the low encoding rate of the ancilla and computational code patches limit the logical parallelism when maintaining low space overhead. We expect certain compilations of quantum algorithms that have limited parallelism to be natural candidates for our architecture [20]. However, it would be interesting to carry out end-to-end algorithmic compilation with qLDPC codes to evaluate the full space-time cost, and understand which algorithms and compilations are most suited to our architecture. Another exciting avenue of research is to improve upon the QEC constructions used here, including alternative qLDPC code constructions with better properties, single-shot logical gate constructions [256, 255], as well as the use of other types of computational logical qubits that may support transversal non-Clifford gates [288] or have lower overhead. Moreover, just as topological codes have interpretations as topological phases of matter, it will be interesting to explore the connection between qLDPC codes and highly entangled states of matter [289], and the techniques described here may be useful for the exploration of novel exotic states of matter.

# CHAPTER 5

# CONCLUSION AND OUTLOOK

In this thesis, we have presented several new fault-tolerant schemes based on bosonic codes (Chapter 2), tailored topological codes (Chapter 3), and qLDPC codes (Chapter 4). These schemes exploit new hardware and coding features and could potentially outpeform the baseline scheme using planar surface codes, bare physical qubits, and nearest-neighbor connectivity. We emphasize again that fault tolerance is a field of study that contains, but is not limited to QEC. Instead of only focusing on designing new QEC codes with good parameters, more efforts in this thesis have been dedicated to designing fault-tolerant logical operations for full-fledged quantum computation. To achieve full fault tolerance in the operational level, we have exploited various fault-tolerant ideas such as autonomous QEC [52], path-independence quantum control [6], flag qubits [165], and single-shot QEC [145], and integrated techniques from various fields including classical and quantum coding theory, quantum control, open quantum systems, and reservior engineering, topological physics and AMO physics.

Although the new fault-tolerant schemes presented in this work provide promising alternative pathways toward fault tolerance, there are a few important caveats:

- Although the bosonic systems potentially support more hardware-efficient fault-tolerant schemes, they are generically more difficult to control compared to two-level systems. For instance, the stabilized cat/squeezed-cat qubits presented in Chapter 2 rely on a highly driven-dissipative system that involves nonlinear interactions and sophisticated drives. Such control complexities not only increase engineering difficulties but can often result in slower and noisier operations. So in practice, we might need to consider trade-offs between the control complexity and the error-suppression capability when choosing fault-tolerant schemes. A new effort that reflects upon such a philosophy is the recent development of the dual-rail bosonic qubits [290], which offers limited error-detection

capability but only requires extremely simple controls. Achieving much better error protection without sacrificing too much control complexity will be essential for bosonic systems to outperform bare qubits in practice.

- Although QEC codes, such as the topological coeds, can have boosted performance when tailored towards structured noise, creating and utilizing such noise structures does not come for free. For instance, although the code threshold $p_{\text{th}}$ could increase with the noise bias (see Chapter 3), the dominant physical error rate $p$ might also increase when trying to engineer a large bias and fully preserving them in the operational level. Since it is the ratio between $p$ and $p_{\text{th}}$ (not the $p_{\text{th}}$ itself) that matters ultimately, it is essential to have more structured noise without increasing the dominant error rates when considering noise-tailored fault-tolerant schemes in practice.

- Although the qLDPC codes (see Chapter 4) can significantly reduce the space overhead compared to surface codes, performing logical operations for these high-rate codes becomes generically more challenging. In addition, building long-range interactions for these non-local codes takes a longer time. So far, it seems that compared to using surface codes, using qLDPC codes is only a way of trading time for space. In the long run, particularly if the space resources are not strictly limited, it is not yet clear which scheme is superior. As such, more efforts toward designing qLDPC-based fault-tolerant schemes with not only a lower space overhead but also a lower space-time overhead overall, are in demand.

At the conclusion of this thesis, we discuss several future directions toward fault tolerance that extend beyond our current results.

Firstly, although we have focused on code-centric fault-tolerant schemes, the recent idea of circuit-centric schemes [29, 30, 25, 31, 32] provides alternative ways of achieving fault tolerance. Such schemes could correct faults in a circuit by directly identifying and utilizing

the inherent circuit redundancy, instead of starting from some QEC codes and building corresponding QEC and logical-operation circuits. This might open up new pathways towards fault-tolerant quantum computing.

Second, the experiences and lessons we learned from the bosonic codes can potentially be applied to systems beyond the bosonic systems. The advantage of bosonic modes for achieving fault tolerance can be fundamentally attributed to their multi-level structure that offers the necessary redundancy for error correction. Similar multi-level structures are also present in many other systems such as neutral atoms [291], ions [119], molecules [292], multi-level superconducting qubits, etc. The philosophy and techniques developed for bosonic codes might shine a light on designing hardware-efficient fault-tolerant schemes for such systems if sufficient quantum control can be realized on these multi-level objects.

Finally, the challenges of designing a practical fault-tolerant scheme that works versatilely for generic quantum computing and any physical platform call for new co-designed schemes that are tailored for specific hardware and computational tasks. For instance, while certain quantum codes favor Clifford gates, i.e. the surface codes, certain quantum codes support easy-to-implement non-Clifford gates [288, 293, 294], which make them suitable for some specific tasks such as the IQP sampling [295, 22]. As another example, while codes with local logical operations, e.g. lattice surgeries, are favored by the superconducting circuits, codes supporting reach transversal logical operations are preferred by devices with long-range connectivities [22]. We expect that, in the near future, co-design between hardware, codes, and algorithms will open up new possibilities for demonstrating robust quantum computing tasks beyond the NISQ era.

# REFERENCES


[1] M. Reimpell and R. F. Werner, Physical review letters **94**, 080501 (2005).

[2] A. S. Fletcher, P. W. Shor, and M. Z. Win, Physical Review A **75**, 012338 (2007).

[3] K. Noh, V. V. Albert, and L. Jiang, IEEE Transactions on Information Theory **65**, 2563 (2018).

[4] J. Guillaud and M. Mirrahimi, Physical Review X **9**, 041053 (2019).

[5] C. Chamberland, K. Noh, P. Arrangoiz-Arriola, E. T. Campbell, C. T. Hann, J. Iverson, H. Putterman, T. C. Bohdanowicz, S. T. Flammia, A. Keller, G. Refael, J. Preskill, L. Jiang, A. H. Safavi-Naeini, O. Painter, and F. G. Brandão, PRX Quantum **3**, 010329 (2022).

[6] W.-L. Ma, M. Zhang, Y. Wong, K. Noh, S. Rosenblum, P. Reinhold, R. J. Schoelkopf, and L. Jiang, Phys. Rev. Lett. **125**, 110503 (2020).

[7] P. Reinhold, S. Rosenblum, W.-L. Ma, L. Frunzio, L. Jiang, and R. J. Schoelkopf, Nature Physics **16**, 822 (2020).

[8] L. Li, C.-L. Zou, V. V. Albert, S. Muralidharan, S. M. Girvin, and L. Jiang, Phys. Rev. Lett. **119**, 030502 (2017).

[9] C. Wang, J. Harrington, and J. Preskill, Annals of Physics **303**, 31 (2003).

[10] C. Horsman, A. G. Fowler, S. Devitt, and R. Van Meter, New Journal of Physics **14**, 123011 (2012).

[11] N. P. Breuckmann, C. Vuillot, E. Campbell, A. Krishna, and B. M. Terhal, Quantum Science and Technology **2**, 10.1088/2058-9565/aa7d3b (2017).

[12] N. P. Breuckmann and J. N. Eberhardt, PRX Quantum **2**, 040101 (2021).

[13] V. V. Albert and P. Faist, eds., *The Error Correction Zoo* (2024).

[14] J. Preskill, Caltech Lecture Notes **7**, 1 (1999).

[15] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, 2010).

[16] P. W. Shor, in *Proceedings 35th annual symposium on foundations of computer science* (Ieee, 1994) pp. 124–134.

[17] A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, Physical Review A **86**, 032324 (2012).

[18] D. Litinski, Quantum **3**, 128 (2019).

[19] M. E. Beverland, P. Murali, M. Troyer, K. M. Svore, T. Hoeffler, V. Kliuchnikov, G. H. Low, M. Soeken, A. Sundaram, and A. Vaschillo, arXiv preprint arXiv:2211.07629 10.48550/arxiv.2211.07629 (2022).

[20] C. Gidney and M. Ekerå, Quantum **5**, 1 (2019).

[21] Nature **614**, 676 (2023).

[22] D. Bluvstein, S. J. Evered, A. A. Geim, S. H. Li, H. Zhou, T. Manovitz, S. Ebadi, M. Cain, M. Kalinowski, D. Hangleiter, J. P. B. Ataides, N. Maskara, I. Cong, X. Gao, P. S. Rodriguez, T. Karolyshyn, G. Semeghini, M. J. Gullans, M. Greiner, V. Vuletić, and M. D. Lukin, Nature 2023 , 1 (2023).

[23] L. Postler, S. Heu$\beta$en, I. Pogorelov, M. Rispler, T. Feldker, M. Meth, C. D. Marciniak, R. Stricker, M. Ringbauer, R. Blatt, *et al.*, Nature **605**, 675 (2022).

[24] L. Egan, D. M. Debroy, C. Noel, A. Risinger, D. Zhu, D. Biswas, M. Newman, M. Li, K. R. Brown, M. Cetina, *et al.*, Nature **598**, 281 (2021).

[25] D. Gottesman, arXiv preprint arXiv:2210.15844 (2022).

[26] P. Aliferis, D. Gottesman, and J. Preskill, arXiv preprint quant-ph/0504218 (2005).

[27] A. Peres, Physical review A **32**, 3266 (1985).

[28] P. W. Shor, Physical review A **52**, R2493 (1995).

[29] D. Bacon, S. T. Flammia, A. W. Harrow, and J. Shi, IEEE Transactions on Information Theory **63**, 2464 (2017).

[30] N. Delfosse and A. Paetznick, arXiv preprint arXiv:2304.05943 (2023).

[31] M. Davydova, N. Tantivasadakarn, S. Balasubramanian, and D. Aasen, arXiv preprint arXiv:2307.10353 (2023).

[32] X. Fu and D. Gottesman, arXiv preprint arXiv:2403.04163 (2024).

[33] B. Eastin and E. Knill, Physical review letters **102**, 110502 (2009).

[34] S. B. Bravyi and A. Y. Kitaev, arXiv preprint quant-ph/9811052 (1998).

[35] A. Y. Kitaev, Annals of Physics **303**, 2 (2003).

[36] S. Krinner, N. Lacroix, A. Remm, A. D. Paolo, E. Genois, C. Leroux, C. Hellings, S. Lazar, F. Swiadek, J. Herrmann, G. J. Norris, C. K. Andersen, M. Müller, A. Blais, C. Eichler, and A. Wallraff, Realizing repeated quantum error correction in a distance-three surface code (2021), arXiv:2112.03708 [quant-ph] .

[37] Y. Zhao, Y. Ye, H.-L. Huang, Y. Zhang, D. Wu, H. Guan, Q. Zhu, Z. Wei, T. He, S. Cao, *et al.*, Physical Review Letters **129**, 030501 (2022).

[38] 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).

[39] D. Gottesman, *Stabilizer codes and quantum error correction* (California Institute of Technology, 1997).

[40] D. Litinski and N. Nickerson, arXiv preprint arXiv:2211.15465 (2022).

[41] D. S. Wang, A. G. Fowler, and L. C. Hollenberg, Physical Review A **83**, 020302 (2011).

[42] A. M. Stephens, Physical Review A **89**, 022321 (2014).

[43] O. Higgott and C. Gidney, arXiv preprint arXiv:2303.15933 (2023).

[44] 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).

[45] M. Mirrahimi, Z. Leghtas, V. V. Albert, S. Touzard, R. J. Schoelkopf, L. Jiang, and M. H. Devoret, New Journal of Physics **16**, 045014 (2014).

[46] S. Puri, L. St-Jean, J. A. Gross, A. Grimm, N. E. Frattini, P. S. Iyer, A. Krishna, S. Touzard, L. Jiang, A. Blais, *et al.*, Science advances **6**, eaay5901 (2020).

[47] D. Bluvstein, H. Levine, G. Semeghini, T. T. Wang, S. Ebadi, M. Kalinowski, A. Keesling, N. Maskara, H. Pichler, M. Greiner, V. Vuletić, and M. D. Lukin, Nature 2022 604:7906 **604**, 451 (2022).

[48] D. Gottesman, arXiv preprint arXiv:1310.2984 (2013).

[49] B. Royer, S. Singh, and S. Girvin, PRX Quantum **3**, 010335 (2022).

[50] S. P. Jain, J. T. Iosue, A. Barg, and V. V. Albert, arXiv preprint arXiv:2302.11593 (2023).

[51] K. Noh, S. M. Girvin, and L. Jiang, Phys. Rev. Lett. **125**, 080503 (2020).

[52] J. Lebreuilly, K. Noh, C.-H. Wang, S. M. Girvin, and L. Jiang, arXiv preprint arXiv:2103.05007 10.48550/ARXIV.2103.05007 (2021).

[53] W.-L. Ma, S. Puri, R. J. Schoelkopf, M. H. Devoret, S. M. Girvin, and L. Jiang, Science Bulletin **66**, 1789 (2021).

[54] P. Kok, W. J. Munro, K. Nemoto, T. C. Ralph, J. P. Dowling, and G. J. Milburn, Reviews of modern physics **79**, 135 (2007).

[55] Kyungjoo Noh, *Quantum Computation and Communication in Bosonic Systems*, Ph.D. thesis, Yale University (2020).

[56] P. Leviant, Q. Xu, L. Jiang, and S. Rosenblum, Quantum **6**, 821 (2022).

[57] M. M. Wilde, *Quantum Information Theory*, 2nd ed (Cambridge University Press, 2013).

[58] H.-P. Breuer and F. Petruccione, *The theory of open quantum systems* (OUP Oxford, 2002).

[59] A. L. Grimsmo, J. Combes, and B. Q. Baragiola, Physical Review X **10**, 011058 (2020).

[60] P. T. Cochrane, G. J. Milburn, and W. J. Munro, Physical Review A **59**, 2631 (1999).

[61] A. J. Brady, A. Eickbusch, S. Singh, J. Wu, and Q. Zhuang, Progress in Quantum Electronics , 100496 (2024).

[62] D. Gottesman, A. Kitaev, and J. Preskill, Physical Review A **64**, 012310 (2001).

[63] G. Zheng, W. He, G. Lee, K. Noh, and L. Jiang, Bulletin of the American Physical Society (2024).

[64] M. H. Michael, M. Silveri, R. Brierley, V. V. Albert, J. Salmilehto, L. Jiang, and S. M. Girvin, Physical Review X **6**, 031006 (2016).

[65] 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).

[66] R. Lescanne, M. Villiers, T. Peronnin, A. Sarlette, M. Delbecq, B. Huard, T. Kontos, M. Mirrahimi, and Z. Leghtas, Nature Physics **16**, 509 (2020).

[67] J. P. Bonilla Ataides, D. K. Tuckett, S. D. Bartlett, S. T. Flammia, and B. J. Brown, Nature communications **12**, 1 (2021).

[68] Q. Xu, N. Mannucci, A. Seif, A. Kubica, S. T. Flammia, and L. Jiang, arXiv preprint arXiv:2203.16486 (2022).

[69] A. S. Darmawan, B. J. Brown, A. L. Grimsmo, D. K. Tuckett, and S. Puri, PRX Quantum **2**, 030345 (2021).

[70] S. Puri, S. Boutin, and A. Blais, npj Quantum Information **3**, 1 (2017).

[71] H. Putterman, J. Iverson, Q. Xu, L. Jiang, O. Painter, F. G. Brandão, and K. Noh, Physical Review Letters **128**, 110502 (2022).

[72] R. Gautier, A. Sarlette, and M. Mirrahimi, P R X Quantum **3**, 020339 (2022).

[73] E. Gouzien, D. Ruiz, F.-M. Le Régent, J. Guillaud, and N. Sangouard, Phys. Rev. Lett. **131**, 040602 (2023).

[74] D. Ruiz, J. Guillaud, A. Leverrier, M. Mirrahimi, and C. Vuillot, arXiv preprint arXiv:2401.09541 (2024).

[75] S. Rosenblum, P. Reinhold, M. Mirrahimi, L. Jiang, L. Frunzio, and R. J. Schoelkopf, Science **361**, 266 (2018).

[76] C. Ryan-Anderson, N. Brown, M. Allman, B. Arkin, G. Asa-Attuah, C. Baldwin, J. Berg, J. Bohnet, S. Braxton, N. Burdick, *et al.*, arXiv preprint arXiv:2208.01863 (2022).

[77] R. Acharya, I. Aleiner, R. Allen, T. I. Andersen, M. Ansmann, F. Arute, K. Arya, A. Asfaw, J. Atalaya, R. Babbush, *et al.*, arXiv preprint arXiv:2207.06431 (2022).

[78] J.-M. Lihm, K. Noh, and U. R. Fischer, Physical Review A **98**, 012317 (2018).

[79] A. Grimm, N. E. Frattini, S. Puri, S. O. Mundhada, S. Touzard, M. Mirrahimi, S. M. Girvin, S. Shankar, and M. H. Devoret, Nature **584**, 205 (2020).

[80] C. Berdou, A. Murani, U. Reglade, W. Smith, M. Villiers, J. Palomo, M. Rosticher, A. Denis, P. Morfin, M. Delbecq, *et al.*, arXiv preprint arXiv:2204.09128 (2022).

[81] S. Kwon, S. Watabe, and J.-S. Tsai, npj Quantum Information **8**, 1 (2022).

[82] B. Royer, S. Singh, and S. Girvin, Physical Review Letters **125**, 260509 (2020).

[83] P. Campagne-Ibarcq, A. Eickbusch, S. Touzard, E. Zalys-Geller, N. E. Frattini, V. V. Sivak, P. Reinhold, S. Puri, S. Shankar, R. J. Schoelkopf, *et al.*, Nature **584**, 368 (2020).

[84] J. M. Gertler, B. Baker, J. Li, S. Shirol, J. Koch, and C. Wang, Nature **590**, 243 (2021).

[85] D. S. Schlegel, F. Minganti, and V. Savona, Physical Review A **106**, 10.1103/physreva.106.022431 (2022).

[86] Z. Leghtas, S. Touzard, I. M. Pop, A. Kou, B. Vlastakis, A. Petrenko, K. M. Sliwa, A. Narla, S. Shankar, M. J. Hatridge, *et al.*, Science **347**, 853 (2015).

[87] S. Touzard, A. Kou, N. Frattini, V. Sivak, S. Puri, A. Grimm, L. Frunzio, S. Shankar, and M. Devoret, Physical review letters **122**, 080502 (2019).

[88] J. Poyatos, J. I. Cirac, and P. Zoller, Physical review letters **77**, 4728 (1996).

[89] D. K. Tuckett, S. D. Bartlett, and S. T. Flammia, Physical review letters **120**, 050505 (2018).

[90] D. K. Tuckett, A. S. Darmawan, C. T. Chubb, S. Bravyi, S. D. Bartlett, and S. T. Flammia, Physical Review X **9**, 041031 (2019).

[91] D. K. Tuckett, S. D. Bartlett, S. T. Flammia, and B. J. Brown, Physical review letters **124**, 130501 (2020).

[92] J. P. Bonilla Ataides, D. K. Tuckett, S. D. Bartlett, S. T. Flammia, and B. J. Brown, Nature communications **12**, 1 (2021).

[93] J. Roffe, L. Z. Cohen, A. O. Quintivalle, D. Chandra, and E. T. Campbell, arXiv preprint arXiv:2202.01702 (2022).

[94] J. O'Gorman and E. T. Campbell, Physical Review A **95**, 032338 (2017).

[95] R. Teh, P. Drummond, and M. Reid, Physical Review Research **2**, 043387 (2020).

[96] H.-Y. Lo, D. Kienzler, L. de Clercq, M. Marinelli, V. Negnevitsky, B. C. Keitch, and J. P. Home, Nature **521**, 336 (2015).

[97] H. Le Jeannic, A. Cavaillès, K. Huang, R. Filip, and J. Laurat, Physical Review Letters **120**, 073603 (2018).

[98] H.-K. Lau and A. A. Clerk, npj Quantum Information **5**, 1 (2019).

[99] X. Pan, J. Schwinger, N.-N. Huang, P. Song, W. Chua, F. Hanamura, A. Joshi, F. Valadares, R. Filip, and Y. Y. Gao, Physical Review X **13**, 021004 (2023).

[100] 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, Physical Review A **97**, 032346 (2018).

[101] M. Mirrahimi, Z. Leghtas, V. V. Albert, S. Touzard, R. J. Schoelkopf, L. Jiang, and M. H. Devoret, New Journal of Physics **16**, 045014 (2014).

[102] G. Pantaleoni, B. Q. Baragiola, and N. C. Menicucci, Physical Review Letters **125**, 040501 (2020).

[103] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Physical Review A **54**, 3824 (1996).

[104] E. Knill and R. Laflamme, Physical Review A **55**, 900 (1997).

[105] 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, Physical Review A **97**, 032346 (2018).

[106] Q. Xu, G. Zheng, Y.-X. Wang, P. Zoller, A. A. Clerk, and L. Jiang, npj Quantum Information **9**, 78 (2023).

[107] J. A. Gross, C. M. Caves, G. J. Milburn, and J. Combes, Quantum Science and Technology **3**, 024005 (2018).

[108] Y.-X. Wang, C. Wang, and A. A. Clerk, PRX Quantum **4**, 010306 (2023).

[109] V. V. Albert, arXiv preprint arXiv:1802.00010 (2018).

[110] Q. Xu, J. K. Iverson, F. G. Brandão, and L. Jiang, Physical Review Research **4**, 013082 (2022).

[111] M. Yuan, Q. Xu, and L. Jiang, arXiv preprint arXiv:2208.06913 (2022).

[112] F. Reiter and A. S. Sørensen, Physical Review A **85**, 032111 (2012).

[113] L. Sun, A. Petrenko, Z. Leghtas, B. Vlastakis, G. Kirchmair, K. Sliwa, A. Narla, M. Hatridge, S. Shankar, J. Blumoff, *et al.*, Nature **511**, 444 (2014).

[114] C. T. Hann, S. S. Elder, C. S. Wang, K. Chou, R. J. Schoelkopf, and L. Jiang, Physical Review A **98**, 022305 (2018).

[115] S. S. Elder, C. S. Wang, P. Reinhold, C. T. Hann, K. S. Chou, B. J. Lester, S. Rosenblum, L. Frunzio, L. Jiang, and R. J. Schoelkopf, Physical Review X **10**, 011001 (2020).

[116] T. Hillmann and F. Quijandría, arXiv preprint arXiv:2210.13359 (2022).

[117] I. V. Inlek, C. Crocker, M. Lichtman, K. Sosnova, and C. Monroe, Physical review letters **118**, 250502 (2017).

[118] C. Bruzewicz, R. McConnell, J. Stuart, J. Sage, and J. Chiaverini, npj Quantum Information **5**, 1 (2019).

[119] M. Ringbauer, M. Meth, L. Postler, R. Stricker, R. Blatt, P. Schindler, and T. Monz, Nature Physics **18**, 1053 (2022).

[120] T. Olsacher, L. Postler, P. Schindler, T. Monz, P. Zoller, and L. M. Sieberer, PRX Quantum **1**, 020316 (2020).

[121] V. Sivak, A. Eickbusch, B. Royer, S. Singh, I. Tsioutsios, S. Ganjam, A. Miano, B. Brock, A. Ding, L. Frunzio, *et al.*, Nature **616**, 50 (2023).

[122] Z. Ni, S. Li, X. Deng, Y. Cai, L. Zhang, W. Wang, Z.-B. Yang, H. Yu, F. Yan, S. Liu, *et al.*, Nature **616**, 56 (2023).

[123] A. Blais, R.-S. Huang, A. Wallraff, S. M. Girvin, and R. J. Schoelkopf, Phys. Rev. A **69**, 062320 (2004).

[124] S. Krastanov, V. V. Albert, C. Shen, C.-L. Zou, R. W. Heeres, B. Vlastakis, R. J. Schoelkopf, and L. Jiang, Phys. Rev. A **92**, 040303 (2015).

[125] R. W. Heeres, P. Reinhold, N. Ofek, L. Frunzio, L. Jiang, M. H. Devoret, and R. J. Schoelkopf, Nature communications **8**, 94 (2017).

[126] Y. Ma, Y. Xu, X. Mu, W. Cai, L. Hu, W. Wang, X. Pan, H. Wang, Y. Song, C.-L. Zou, *et al.*, Nature Physics **16**, 827 (2020).

[127] J. D. Teoh, P. Winkel, H. K. Babla, B. J. Chapman, J. Claes, S. J. de Graaf, J. W. Garmon, W. D. Kalfus, Y. Lu, A. Maiti, *et al.*, arXiv preprint arXiv:2212.12077 (2022).

[128] T. Tsunoda, J. D. Teoh, W. D. Kalfus, S. J. de Graaf, B. J. Chapman, J. C. Curtis, N. Thakur, S. M. Girvin, and R. J. Schoelkopf, PRX Quantum **4**, 020354 (2023).

[129] K. S. Chou, T. Shemma, H. McCarrick, T.-C. Chien, J. D. Teoh, P. Winkel, A. Anderson, J. Chen, J. Curtis, S. J. de Graaf, *et al.*, arXiv preprint arXiv:2307.03169 (2023).

[130] Z. Leghtas, G. Kirchmair, B. Vlastakis, R. J. Schoelkopf, M. H. Devoret, and M. Mirrahimi, Phys. Rev. Lett. **111**, 120501 (2013).

[131] R. W. Heeres, B. Vlastakis, E. Holland, S. Krastanov, V. V. Albert, L. Frunzio, L. Jiang, and R. J. Schoelkopf, Phys. Rev. Lett. **115**, 137002 (2015).

[132] B. J. Chapman, S. J. de Graaf, S. H. Xue, Y. Zhang, J. Teoh, J. C. Curtis, T. Tsunoda, A. Eickbusch, A. P. Read, A. Koottandavida, S. O. Mundhada, L. Frunzio, M. Devoret, S. Girvin, and R. Schoelkopf, PRX Quantum **4**, 020355 (2023).

[133] Y. Lu, A. Maiti, J. W. Garmon, S. Ganjam, Y. Zhang, J. Claes, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf, Nature Communications **14**, 5767 (2023).

[134] P. W. Shor, in *Proceedings of 37th conference on foundations of computer science* (IEEE, 1996) pp. 56–65.

[135] Y. Ma, Y. Xu, X. Mu, W. Cai, L. Hu, W. Wang, X. Pan, H. Wang, Y. P. Song, C.-L. Zou, and L. Sun, Nature Physics **16**, 827 (2020).

[136] T. Tsunoda, J. D. Teoh, W. D. Kalfus, S. J. de Graaf, B. J. Chapman, J. C. Curtis, N. Thakur, S. M. Girvin, and R. J. Schoelkopf, PRX Quantum **4**, 020354 (2023).

[137] W.-L. Ma, S.-S. Li, and L. Jiang, Phys. Rev. Res. **4**, 023102 (2022).

[138] Q. Xu, P. Zeng, D. Xu, and L. Jiang, arXiv preprint arXiv:2310.20578 (2023).

[139] E. Knill, Phys. Rev. A **71**, 042322 (2005).

[140] R. Alicki, M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Rev. A **65**, 062101 (2002).

[141] D. V. Averin and R. Fazio, Journal of Experimental and Theoretical Physics Letters **78**, 664 (2003).

[142] R. Klesse and S. Frank, Phys. Rev. Lett. **95**, 230503 (2005).

[143] M. McEwen, L. Faoro, K. Arya, A. Dunsworth, T. Huang, S. Kim, B. Burkett, A. Fowler, F. Arute, J. C. Bardin, *et al.*, Nature Physics **18**, 107 (2021).

[144] H. Cartan and S. Eilenberg, *Homological algebra*, Vol. 28 (Princeton university press, 1999).

[145] H. Bombin, Physical Review X **5**, 031043 (2015).

[146] X.-G. Wen, Physical Review B **40**, 7387 (1989).

[147] E. Dennis, A. Kitaev, A. Landahl, and J. Preskill, Journal of Mathematical Physics **43**, 4452 (2002).

[148] S. Bravyi, D. Poulin, and B. Terhal, Physical review letters **104**, 050503 (2010).

[149] S. Bravyi, M. B. Hastings, and S. Michalakis, Journal of mathematical physics **51** (2010).

[150] B. Zeng, A. Cross, and I. L. Chuang, IEEE Transactions on Information Theory **57**, 6272 (2011).

[151] T. Jochym-O'Connor, A. Kubica, and T. J. Yoder, Physical Review X **8**, 021047 (2018).

[152] H. Bombin and M. A. Martin-Delgado, Physical review letters **97**, 180501 (2006).

[153] A. M. Kubica, *The ABCs of the color code: A study of topological quantum codes as toy models for fault-tolerant quantum computation and quantum phases of matter* (California Institute of Technology, 2018).

[154] B. M. Terhal, Reviews of Modern Physics **87**, 307 (2015).

[155] R. Chao, M. E. Beverland, N. Delfosse, and J. Haah, Quantum **4**, 352 (2020).

[156] M. E. Beverland, A. Kubica, and K. M. Svore, PRX Quantum **2**, 020341 (2021).

[157] A. Robertson, C. Granade, S. D. Bartlett, and S. T. Flammia, Physical Review Applied **8**, 064004 (2017).

[158] L. Egan, D. M. Debroy, C. Noel, A. Risinger, D. Zhu, D. Biswas, M. Newman, M. Li, K. R. Brown, M. Cetina, *et al.*, arXiv preprint arXiv:2009.11482 (2020).

[159] Y. Wu, W.-S. Bao, S. Cao, F. Chen, M.-C. Chen, X. Chen, T.-H. Chung, H. Deng, Y. Du, D. Fan, *et al.*, Physical review letters **127**, 180501 (2021).

[160] Y. Zhao, Y. Ye, H.-L. Huang, Y. Zhang, D. Wu, H. Guan, Q. Zhu, Z. Wei, T. He, S. Cao, *et al.*, arXiv preprint arXiv:2112.13505 (2021).

[161] A. Dua, A. Kubica, L. Jiang, S. T. Flammia, and M. J. Gullans, arXiv preprint arXiv:2201.07802 (2022).

[162] A. A. Kovalev and L. P. Pryadko, in *2012 IEEE International Symposium on Information Theory Proceedings* (Cambridge, MA, USA, 2012) pp. 348–352.

[163] M. M. Wilde, *Quantum information theory* (Cambridge University Press, 2019).

[164] R. Sarkar and T. J. Yoder, arXiv preprint arXiv:2101.09349 (2021).

[165] R. Chao and B. W. Reichardt, Phys. Rev. Lett. **121**, 050502 (2018).

[166] R. Chao and B. W. Reichardt, npj Quantum Information **4**, 42 (2018).

[167] C. Chamberland and M. E. Beverland, Quantum **2**, 53 (2018).

[168] B. W. Reichardt, Quantum Science and Technology **6**, 015007 (2020).

[169] X.-G. Wen, Physical review letters **90**, 016803 (2003).

[170] S. Bravyi and M. B. Hastings, in *Proceedings of the Forty-Sixth Annual ACM Symposium on Theory of Computing*, STOC '14 (Association for Computing Machinery, New York, NY, USA, 2014) p. 273–282.

[171] Q. Xu, N. Mannucci, A. Seif, A. Kubica, S. T. Flammia, and L. Jiang, Phys. Rev. Res. **5**, 013035 (2023).

[172] D. Gottesman, in *Quantum information science and its contributions to mathematics, Proceedings of Symposia in Applied Mathematics*, Vol. 68 (2010) pp. 13–58.

[173] Y. Tomita and K. M. Svore, Physical Review A **90**, 062320 (2014).

[174] T. J. Yoder and I. H. Kim, Quantum **1**, 2 (2017).

[175] A. M. Steane, Physical Review Letters **78**, 2252 (1997).

[176] Q. Xu, J. K. Iverson, F. G. Brandão, and L. Jiang, Physical Review Research **4**, 013082 (2022).

[177] C. Chamberland, K. Noh, P. Arrangoiz-Arriola, E. T. Campbell, C. T. Hann, J. Iverson, H. Putterman, T. C. Bohdanowicz, S. T. Flammia, A. Keller, *et al.*, PRX Quantum **3**, 010329 (2022).

[178] X. Zhang, E. Kim, D. K. Mark, S. Choi, and O. Painter, arXiv preprint arXiv:2206.12803 (2022).

[179] A. Gold, J. Paquette, A. Stockklauser, M. J. Reagor, M. S. Alam, A. Bestwick, N. Didier, A. Nersisyan, F. Oruc, A. Razavi, *et al.*, npj Quantum Information **7**, 1 (2021).

[180] H. Yan, Y. Zhong, H.-S. Chang, A. Bienfait, M.-H. Chou, C. R. Conner, É. Dumur, J. Grebel, R. G. Povey, and A. N. Cleland, Physical Review Letters **128**, 080504 (2022).

[181] A. Antipov, E. Kiktenko, and A. Fedorov, arXiv preprint arXiv:2207.13356 (2022).

[182] D. Gottesman, Physical Review A **57**, 127 (1998).

[183] T. J. Yoder, R. Takagi, and I. L. Chuang, Physical Review X **6**, 031039 (2016).

[184] M. B. Hastings, J. Haah, and R. O'Donnell, in *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2021 (Association for Computing Machinery, New York, NY, USA, 2021) p. 1276–1288.

[185] N. P. Breuckmann and J. N. Eberhardt, IEEE Transactions on Information Theory **67**, 6653 (2021).

[186] P. Panteleev and G. Kalachev, IEEE Transactions on Information Theory **68**, 213 (2021).

[187] P. Panteleev and G. Kalachev, in *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2022 (Association for Computing Machinery, New York, NY, USA, 2022) p. 375–388.

[188] D. A. Lidar and T. A. Brun, *Quantum error correction* (Cambridge university press, 2013).

[189] J. P. Clemens, S. Siddiqui, and J. Gea-Banacloche, Phys. Rev. A **69**, 062313 (2004).

[190] N. H. Nickerson and B. J. Brown, Quantum **3**, 131 (2019).

[191] P. Aliferis, D. Gottesman, and J. Preskill, Quantum Info. Comput. **6**, 97–165 (2006).

[192] D. Aharonov, A. Kitaev, and J. Preskill, Phys. Rev. Lett. **96**, 050504 (2006).

[193] A. P. Vepsäläinen, A. H. Karamlou, J. L. Orrell, A. S. Dogra, B. Loer, F. Vasconcelos, D. K. Kim, A. J. Melville, B. M. Niedzielski, J. L. Yoder, *et al.*, Nature **584**, 551 (2020).

[194] L. Cardani, F. Valenti, N. Casali, G. Catelani, T. Charpentier, M. Clemenza, I. Colantoni, A. Cruciani, G. D'Imperio, L. Gironi, *et al.*, Nature communications **12**, 2273 (2021).

[195] C. Gidney and M. Ekerå, Quantum **5**, 433 (2021).

[196] F. A. Zwanenburg, A. S. Dzurak, A. Morello, M. Y. Simmons, L. C. L. Hollenberg, G. Klimeck, S. Rogge, S. N. Coppersmith, and M. A. Eriksson, Rev. Mod. Phys. **85**, 961 (2013).

[197] D. Rainis and D. Loss, Phys. Rev. B **85**, 174533 (2012).

[198] J. M. Martinis, npj Quantum Information **7**, 90 (2021).

[199] I. Nsanzineza and B. L. T. Plourde, Phys. Rev. Lett. **113**, 117002 (2014).

[200] U. Patel, I. V. Pechenezhskiy, B. L. T. Plourde, M. G. Vavilov, and R. McDermott, Phys. Rev. B **96**, 220501 (2017).

[201] F. Henriques, F. Valenti, T. Charpentier, M. Lagoin, C. Gouriou, M. Martínez, L. Cardani, M. Vignati, L. Grünhaupt, D. Gusenkova, *et al.*, Applied physics letters **115**, 212601 (2019).

[202] L. M. Duan, B. B. Blinov, D. L. Moehring, and C. Monroe, Quantum Inf. Comput. **4**, 165 (2004).

[203] L. Jiang, J. M. Taylor, A. S. Sørensen, and M. D. Lukin, Phys. Rev. A **76**, 062323 (2007).

[204] J. Kim and C. Kim, Quantum Information and Computation **9** (2009).

[205] Y. L. Lim, S. D. Barrett, A. Beige, P. Kok, and L. C. Kwek, Physical Review A **73**, 012304 (2006).

[206] N. H. Nickerson, Y. Li, and S. C. Benjamin, Nature communications **4**, 1756 (2013).

[207] D. K. L. Oi, S. J. Devitt, and L. C. L. Hollenberg, Phys. Rev. A **74**, 052313 (2006).

[208] R. Van Meter, T. D. Ladd, A. G. Fowler, and Y. Yamamoto, International Journal of Quantum Information **8**, 295 (2010).

[209] R. Van Meter and S. Devitt, IEEE Computer **49**, 31 (2016).

[210] C. Monroe, R. Raussendorf, A. Ruthven, K. R. Brown, P. Maunz, L. M. Duan, and J. Kim, Physical Review A **89**, 022317 (2014).

[211] M. Grassl, T. Beth, and T. Pellizzari, Phys. Rev. A **56**, 33 (1997).

[212] E. Knill, R. Laflamme, and G. Milburn, arXiv preprint quant-ph/0006120 (2000).

[213] M. Silva, arXiv preprint quant-ph/0405112 (2004).

[214] Y. Wu, S. Kolkowitz, S. Puri, and J. D. Thompson, arXiv preprint arXiv:2201.03540 (2022).

[215] P. Kurpiers, P. Magnard, T. Walter, B. Royer, M. Pechal, J. Heinsoo, Y. Salathé, A. Akin, S. Storz, J.-C. Besse, S. Gasparinetti, A. Blais, and A. Wallraff, Nature **558**, 264 (2018).

[216] C. J. Axline, L. D. Burkhart, W. Pfaff, M. Zhang, K. Chou, P. Campagne-Ibarcq, P. Reinhold, L. Frunzio, S. M. Girvin, L. Jiang, M. H. Devoret, and R. J. Schoelkopf, Nature Physics **14**, 705 (2018).

[217] Y. Zhong, H.-S. Chang, A. Bienfait, É. Dumur, M.-H. Chou, C. R. Conner, J. Grebel, R. G. Povey, H. Yan, D. I. Schuster, *et al.*, Nature **590**, 571 (2021).

[218] A. Gold, J. P. Paquette, A. Stockklauser, M. J. Reagor, M. S. Alam, A. Bestwick, N. Didier, A. Nersisyan, F. Oruc, A. Razavi, B. Scharmann, *et al.*, npj Quantum Information **7**, 142 (2021).

[219] Z. Chen, K. J. Satzinger, J. Atalaya, A. N. Korotkov, A. Dunsworth, D. Sank, C. Quintana, M. McEwen, R. Barends, P. V. Klimov, *et al.*, Nature **595**, 383 (2021).

[220] A. Steane, Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences **452**, 2551 (1996).

[221] Q. Xu, A. Seif, H. Yan, N. Mannucci, B. O. Sane, R. Van Meter, A. N. Cleland, and L. Jiang, Phys. Rev. Lett. **129**, 240502 (2022).

[222] H.-S. Chang, K. J. Satzinger, Y. P. Zhong, A. Bienfait, M.-H. Chou, C. R. Conner, É. Dumur, J. Grebel, G. A. Peairs, R. G. Povey, and A. N. Cleland, Applied Physics Letters **117**, 244001 (2020).

[223] C. Zhou, P. Lu, M. Praquin, T.-C. Chien, R. Kaufman, X. Cao, M. Xia, R. Mong, W. Pfaff, D. Pekker, and M. Hatridge, arXiv preprint 2109.06848 (2021).

[224] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, Physical Review Letters **76**, 722 (1996).

[225] Y. Wan, D. Kienzler, S. D. Erickson, K. H. Mayer, T. R. Tan, J. J. Wu, H. M. Vasconcelos, S. Glancy, E. Knill, D. J. Wineland, A. C. Wilson, and D. Leibfried, Science **364**, 875 (2019).

[226] K. S. Chou, J. Z. Blumoff, C. S. Wang, P. C. Reinhold, C. J. Axline, Y. Y. Gao, L. Frunzio, M. Devoret, L. Jiang, and R. Schoelkopf, Nature **561**, 368 (2018).

[227] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. Brandao, D. A. Buell, B. Burkett, Y. Chen, Z. Chen, B. Chiaro, R. Collins, W. Courtney, A. Dunsworth, E. Farhi, B. Foxen, A. Fowler, C. Gidney, M. Giustina, R. Graff, K. Guerin, S. Habegger, M. P. Harrigan, M. J. Hartmann, A. Ho, M. Hoffmann, T. Huang, T. S. Humble, S. V. Isakov, E. Jeffrey, Z. Jiang, D. Kafri, K. Kechedzhi, J. Kelly, P. V. Klimov, S. Knysh, A. Korotkov, F. Kostritsa, D. Landhuis, M. Lindmark, E. Lucero, D. Lyakh, S. Mandrà, J. R. McClean, M. McEwen, A. Megrant, X. Mi, K. Michielsen, M. Mohseni, J. Mutus, O. Naaman, M. Neeley, C. Neill, M. Y. Niu, E. Ostby, A. Petukhov, J. C. Platt, C. Quintana, E. G. Rieffel, P. Roushan, N. C. Rubin, D. Sank, K. J. Satzinger, V. Smelyanskiy, K. J. Sung, M. D. Trevithick, A. Vainsencher, B. Villalonga, T. White, Z. J. Yao, P. Yeh, A. Zalcman, H. Neven, and J. M. Martinis, Nature **574**, 505 (2019).

[228] Y. Wu, W.-S. Bao, S. Cao, F. Chen, M.-C. Chen, X. Chen, T.-H. Chung, H. Deng, Y. Du, D. Fan, *et al.*, Physical Review Letters **127**, 180501 (2021).

[229] J. Heinsoo, C. K. Andersen, A. Remm, S. Krinner, T. Walter, Y. Salathé, S. Gasparinetti, J.-C. Besse, A. Potočnik, A. Wallraff, and C. Eichler, Physical Review Applied **10**, 034040 (2018).

[230] S. Bravyi and A. Kitaev, Physical Review A **71**, 022316 (2005).

[231] E. Knill, arXiv preprint quant-ph/0402171 (2004).

[232] J. P. Tillich and G. Zemor, IEEE Transactions on Information Theory **60**, 1193 (2014).

[233] P. Panteleev and G. Kalachev, Quantum **5**, 585 (2019).

[234] P. Panteleev and G. Kalachev, IEEE Transactions on Information Theory **68**, 213 (2022).

[235] N. P. Breuckmann and J. N. Eberhardt, IEEE Transactions on Information Theory **67**, 6653 (2020).

[236] P. Panteleev and G. Kalachev, Proceedings of the Annual ACM Symposium on Theory of Computing , 375 (2022).

[237] Q. Xu, J. Ataides, C. A. Pattison, N. Raveendran, D. Bluvstein, J. Wurtz, B. Vasic, M. D. Lukin, L. Jiang, and H. Zhou, arXiv preprint arXiv:2308.08648 (2023).

[238] S. Bravyi and B. Terhal, New Journal of Physics **11**, 043029 (2009).

[239] J. Haah, SciPost Physics **10**, 011 (2021).

[240] D. J. Williamson and N. Baspin, arXiv preprint arXiv:2309.16503 (2023).

[241] T.-C. Lin, A. Wills, and M.-H. Hsieh, arXiv preprint arXiv:2309.16104 (2023).

[242] M. H. Freedman, D. A. Meyer, and F. Luo, in *Mathematics of quantum computation* (Chapman and Hall/CRC, 2002) pp. 303–338.

[243] N. P. Breuckmann and B. M. Terhal, IEEE transactions on Information Theory **62**, 3731 (2016).

[244] S. Evra, T. Kaufman, and G. Zémor, SIAM Journal on Computing , FOCS20 (2022).

[245] A. Leverrier and G. Zémor, arXiv preprint arXiv:2202.13641 10.48550/arxiv.2202.13641 (2022).

[246] S. Gu, C. A. Pattison, and E. Tang, in *Proceedings of the 55th Annual ACM Symposium on Theory of Computing* (2023) pp. 919–932.

[247] I. Dinur, M.-H. Hsieh, T.-C. Lin, and T. Vidick, in *Proceedings of the 55th annual ACM symposium on theory of computing* (2023) pp. 905–918.

[248] N. Baspin and A. Krishna, Quantum **6**, 711 (2021).

[249] N. Baspin and A. Krishna, Physical Review Letters **129**, 050505 (2022).

[250] M. A. Tremblay, N. Delfosse, and M. E. Beverland, Physical Review Letters **129**, 050504 (2022).

[251] N. Delfosse, M. E. Beverland, and M. A. Tremblay, arXiv preprint arXiv:2109.14599 (2021).

[252] A. Strikis and L. Berent, arXiv preprint arXiv:2209.14329 (2022).

[253] L. Z. Cohen, I. H. Kim, S. D. Bartlett, and B. J. Brown, Science Advances **8**, 10.1126/sciadv.abn1717 (2022).

[254] A. Krishna and D. Poulin, Physical Review X **11**, 011023 (2021).

[255] N. P. Breuckmann and S. Burton, arXiv preprint arXiv:2202.06647 (2022).

[256] A. O. Quintavalle, P. Webster, and M. Vasmer, arXiv preprint arXiv:2204.10812 (2022).

[257] A. Jenkins, J. W. Lis, A. Senoo, W. F. McGrew, and A. M. Kaufman, Physical Review X **12**, 021027 (2022).

[258] S. Ma, A. P. Burgers, G. Liu, J. Wilson, B. Zhang, and J. D. Thompson, Physical Review X **12**, 021028 (2022).

[259] K. Barnes, P. Battaglino, B. J. Bloom, K. Cassella, R. Coxe, N. Crisosto, J. P. King, S. S. Kondov, K. Kotru, S. C. Larsen, J. Lauigan, B. J. Lester, M. McDonald, E. Megidish, S. Narayanaswami, C. Nishiguchi, R. Notermans, L. S. Peng, A. Ryou, T. Y. Wu, and M. Yarwood, Nature Communications 2022 13:1 **13**, 1 (2022).

[260] S. J. Evered, D. Bluvstein, M. Kalinowski, S. Ebadi, T. Manovitz, H. Zhou, S. H. Li, A. A. Geim, T. T. Wang, N. Maskara, H. Levine, G. Semeghini, M. Greiner, V. Vuletic, and M. D. Lukin, arXiv preprint arXiv:2304.05420 (2023).

[261] S. Ma, G. Liu, P. Peng, B. Zhang, S. Jandura, J. Claes, A. P. Burgers, G. Pupillo, S. Puri, and J. D. Thompson, arXiv preprint arXiv:2305.05493 (2023).

[262] P. Scholl, A. L. Shaw, R. B.-S. Tsai, R. Finkelstein, J. Choi, and M. Endres, arXiv preprint arXiv:2305.03406 (2023).

[263] J. Beugnon, C. Tuchendler, H. Marion, A. Gaëtan, Y. Miroshnychenko, Y. R. Sortais, A. M. Lance, M. P. Jones, G. Messin, A. Browaeys, and P. Grangier, Nature Physics 2007 3:10 **3**, 696 (2007).

[264] T. Hashizume, G. S. Bentsen, S. Weber, and A. J. Daley, Physical Review Letters **126**, 200603 (2021).

[265] T. Dordevic, P. Samutpraphoot, P. L. Ocola, H. Bernien, B. Grinkemeyer, I. Dimitrova, V. Vuletić, and M. D. Lukin, Science **373**, 1511 (2021).

[266] M. Malinowski, D. T. C. Allcock, and C. J. Ballance, arXiv preprint arXiv:2305.12773 (2023).

[267] H. Levine, A. Keesling, G. Semeghini, A. Omran, T. T. Wang, S. Ebadi, H. Bernien, M. Greiner, V. Vuletić, H. Pichler, and M. D. Lukin, Physical Review Letters **123**, 170503 (2019).

[268] C. A. Pattison, A. Krishna, and J. Preskill, arXiv preprint arXiv:2303.04798 10.48550/arxiv.2303.04798 (2023).

[269] N. Raveendran, N. Rengaswamy, F. Rozpędek, A. Raina, L. Jiang, and B. Vasić, Quantum **6**, 767 (2022).

[270] N. Baspin, O. Fawzi, and A. Shayeghi, arXiv preprint arXiv:2302.04317 10.48550/arxiv.2302.04317 (2023).

[271] A. Leverrier, J. P. Tillich, and G. Zemor, Proceedings - Annual IEEE Symposium on Foundations of Computer Science, FOCS **2015-Decem**, 810 (2015).

[272] A. O. Quintavalle, M. Vasmer, J. Roffe, and E. T. Campbell, PRX Quantum **2**, 10.1103/prxquantum.2.020340 (2020).

[273] S. Gu, C. A. Pattison, and E. Tang, Proceedings of the Annual ACM Symposium on Theory of Computing , 919 (2022).

[274] I. Dinur, M.-H. H. Hsieh, T.-C. C. Lin, and T. Vidick, Proceedings of the Annual ACM Symposium on Theory of Computing , 905 (2022).

[275] T.-C. Lin and M.-H. Hsieh, arXiv preprint arXiv:2203.03581 10.48550/arxiv.2203.03581 (2022).

[276] T. Richardson and R. Urbanke, *Modern coding theory* (Cambridge University Press, 2008).

[277] A. O. Quintavalle and E. T. Campbell, IEEE Transactions on Information Theory **68**, 6569 (2022).

[278] A. Grospellier, L. Grouès, A. Krishna, and A. Leverrier, Quantum **5**, 432 (2021).

[279] O. Fawzi, A. Grospellier, and A. Leverrier, in *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing* (2018) pp. 521–534.

[280] L. P. Pryadko, V. A. Shabashov, and V. K. Kozin, Journal of Open Source Software **7**, 4120 (2022).

[281] *Supplementary Material*.

[282] C. Gidney, Quantum **5**, 497 (2021).

[283] J. Roffe, D. R. White, S. Burton, and E. Campbell, Physical Review Research **2**, 043423 (2020).

[284] O. Higgott and N. P. Breuckmann, PRX Quantum **4**, 020332 (2023).

[285] K. Y. Kuo and C. Y. Lai, npj Quantum Information 2022 8:1 **8**, 1 (2022).

[286] S. Ebadi, A. Keesling, M. Cain, T. T. Wang, H. Levine, D. Bluvstein, G. Semeghini, A. Omran, J. G. Liu, R. Samajdar, X. Z. Luo, B. Nash, X. Gao, B. Barak, E. Farhi, S. Sachdev, N. Gemelke, L. Zhou, S. Choi, H. Pichler, S. T. Wang, M. Greiner, V. Vuletić, and M. D. Lukin, Science **376**, 1209 (2022).

[287] H. Poulsen Nautrup, N. Friis, and H. J. Briegel, Nature Communications 2017 8:1 **8**, 1 (2017).

[288] H. Bombín, New Journal of Physics **17**, 10.48550/arxiv.1311.0879 (2013).

[289] A. Anshu, N. P. Breuckmann, and C. Nirkhe, arXiv preprint arXiv:2206.13228 (2022).

[290] J. D. Teoh, P. Winkel, H. K. Babla, B. J. Chapman, J. Claes, S. J. de Graaf, J. W. Garmon, W. D. Kalfus, Y. Lu, A. Maiti, *et al.*, Proceedings of the National Academy of Sciences **120**, e2221736120 (2023).

[291] I. Cong, H. Levine, A. Keesling, D. Bluvstein, S.-T. Wang, and M. D. Lukin, Physical Review X **12**, 021049 (2022).

[292] V. V. Albert, J. P. Covey, and J. Preskill, Physical Review X **10**, 031050 (2020).

[293] E. T. Campbell, Blog post (2016).

[294] A. Kubica, B. Yoshida, and F. Pastawski, New Journal of Physics **17**, 083026 (2015).

[295] L. Paletta, A. Leverrier, A. Sarlette, M. Mirrahimi, and C. Vuillot, arXiv preprint arXiv:2307.10729 (2023).