THE UNIVERSITY OF CHICAGO

TRUNCATING THE HILBERT SPACE: TOPICS ON HARDWARE-EFFICIENT QUANTUM CONTROL, ERROR CORRECTION, AND TOMOGRAPHY

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE PRITZKER SCHOOL OF MOLECULAR ENGINEERING
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

BY
MING YUAN

CHICAGO, ILLINOIS
AUGUST 2025

ABSTRACT

Over the past century, people have observed the development of quantum mechanics from a conceptual mystery to a powerful framework for information processing. However, the quantum information remains fragile due to the ubiquitous noise. To address the challenge, device physicists continue to improve pulse-level control schemes for better performance in physical operations, while information scientists employ resource redundancy to protect the quantum states from direct damage caused by noise. Bridging these efforts, the hardware-efficient paradigm exploits the specific hardware structures to suppress the error or reshape the error pattern, which facilitates the error correction or characterization in the next step.

On the other hand, the Hilbert space is usually intractable due to its size. In this dissertation, I will provide several case studies to achieve hardware-efficiency through the truncation of Hilbert space. First, we truncate the Hilbert space of a resonator through a destructive interference, streamlining pulse design for high-fidelity operations. Then, we utilize driven-dissipative processes to autonomously stabilize subspaces in resonators or atomic systems, which provides an encoded qubit with a structured error channel. We also design operations that preserve such error structures, as they are essential for the next-level error correction. Finally, a collective truncation in a multipartite system or constraints on specific subsets of states in the Hilbert space can also provide efficiency in state characterization tasks. I will present several experimentally relevant examples to justify this claim.

To my parents

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

Pursuing the Ph.D. is a long journey, and it is indeed a tough one for me. I cannot finish it on my own without the help of my teachers, friends, and family. Here I would like to express my gratitude to all those who supported me along the way.

First and foremost, I would like to thank my advisor, Prof. Liang Jiang, who brought me into the field of quantum control and error correction. I am continually impressed by his smartness when discussing academic problems together, as he can always quickly grasp the key message I want to convey and offer wise suggestions. On the other hand, he is constantly kind and supportive in every aspect, especially when I experienced a hard time and felt pessimistic about my research and future. He always encourages me to talk and collaborate with others, and indeed, lots of beautiful results shown in this thesis came from collective efforts. I would say that Liang is a role model as both a scientist and a mentor, and it is a great honor for me to work with him over the past years. I always feel lucky about this.

Second, I want to acknowledge all the committee members for my thesis and candidacy exams. Prof. Aashish Clerk provided plenty of insightful suggestions for me on the Kerr universal control project (see Chapter 2) and helped a lot with my postdoc application. Prof. Hannes Bernien guided me on preparing mini-lectures and problem sets when I served as a teaching assistant with him. I was always inspired by his enthusiasm for teaching. Prof. Zoe Yan also provided sufficient help for my thesis defense. I also want to thank Prof. David Schuster, who offered experimental points of view on several projects I was involved in (Chapter 2 and 6). From the collaboration with his group, I gradually learned to think as an experimentalist instead of merely focusing on the abstract theoretical models.

Jiang Group is a large and collaborative place where I can easily find brilliant people with different expertise to talk with. Alireza Seif taught me a lot about numerical optimization skills, especially the neural network approach. Qian Xu is a reliable expert on several aspects

of quantum error correction, ranging from hardware-efficient encoding to the most recent qLDPC code. He helped me to form an understanding of cat code in the early time. Yat Wong is the cleverest person I have ever seen, who can quickly solve any math problem we encounter in research. Senrui Chen and I joined the group in the same year. Although our background and research focus are quite different (quantum information vs. hardware physics, see Chapter 1), we are still glad to share our progress and happiness with each other. Pei Zeng, my deskmate, is an enthusiast in every aspect of quantum. He is always willing to discuss any academic problem and share his opinions with me. Mikhail Mamaev and Anjun Chu brought their knowledge on atomic physics to the dark spin-cat project (Chapter 4) and its follow-up. They also helped with my job search. Zhaoyou Wang taught me several interesting topics on quantum optics and also supervised me on driving when we took road trips together. Xuanhui Mao was an undergrad in our group who also helped with the spin-cat. I am glad to share some of my understanding with her during this. Finally, I would also like to thank other collaborators in the group, including Changchun Zhong, Nam Mannucci, Fangxin Li, Han Zheng, Su-un Lee, Kento Tsubouchi; and friends including Philip Rozpedek, Changhun Oh, Bikun Li, Hyukgun Kwon, Guo Zheng, Debayan Bandyopadhyay, Yuexun Huang, and the younger generation of group members.

In our group, I have benefited from having great opportunities to work with external collaborators. Kevin He and Srivatsan Chakram from the multimode team are reliable colleagues who successfully demonstrated our efficient tomography idea with their device (Chapter 6), while Andreas Kruckenhauser took the lead on wrapping up the dark spin-cat results. I want to mention all others who have collaborated with me at some point, including Andrew Lingenfelter, Akel Hashim, Long B. Nguyen, Rick van Bijnen, Prof. Peter Zoller, Jordan Huang, and Yufeng Wu.

Besides, I want to thank all the theoretical and experimental colleagues in the UChicago Quantum Community who would like to talk with me and treat me as a friend, especially

Yuxin Wang, Mingxing Yao, Yinxing Zang, and others who share the same office with me. Here is also a special thanks to another deskmate of mine, Kaiwen Gui, who always encouraged me to actively reach out and develop a professional network. It is indeed an important skill to survive in our society.

Outside academia, I want to acknowledge every roommate of mine, including Yu Gui, Zhouyu Shen, Chao Guo, Shaozhen Yang, and Yuxuan Zhu. Yu and Zhouyu are also classmates of mine at USTC who helped with my initial setup. Chao is like a senior brother who taught me to live for a living. Shaozhen is a passionate experimentalist on AMO physics, who taught me to drive the car and always managed to cheer me up when I was struggling. Yuxuan just came to Chicago during my last year of study, so I can tell him my experience of surviving here. I would also like to thank other friends who came to UChicago together with me from USTC, in particular, Sijia Chen and Erting Tang, who invited me to celebrate Chinese New Year together every past year, so that I would not feel alone.

I sincerely wish everyone I mentioned above a bright future.

Finally, I would like to thank my parents, Yanying Sun and Peizhao Yuan, whom I have not seen in the past four years since I came to Chicago. Despite knowing nothing about what I am working on, they still provide endless love and support to me. This thesis is dedicated to them.

CHAPTER 1

INTRODUCTION

1.1 The world of quantum

Quantum mechanics, a theory initially designed to explain the light-matter interaction at the microscopic scale, has now become the cornerstone of the whole modern physics. The mysterious features allowed by quantum, including the superposition of classically distinguishable states (a cat both alive and dead, noticed by E. Schrödinger [1]) and the entanglement over separated parties ("spooky action at a distance", by A. Einstein [2]), have drawn endless debates on the philosophical level, while the validity of the theory has been demonstrated via enormous sophisticated experiments over the past 100 years.

Those unconventional properties from the quantum world also allure people to develop better strategies that utilize them for practical purposes. More than 40 years ago, R. P. Feynman proposed the idea of simulating the physics of a many-body quantum system with a computer whose elements follow the quantum mechanical laws [3]. In 1994, P. Shor proposed a factoring algorithm with polynomial complexity when using a quantum computer [4], which is believed to be a hard problem in classical computing and therefore plays an important role in cryptography. The success of Shor's algorithm undoubtedly demonstrates the power of quantum computing and inspires generations of people to seek more applications where quantum can offer advantages. We have observed considerable progress in quantum communication, quantum simulation, quantum computing, and quantum sensing. The concept of quantum is gradually accepted by the public, while the advantages of quantum technology are expected to alter the lives of everyone.

However, the promising applications of quantum cannot be achieved without the ability to manipulate quantum states in actual physical systems. In the past 30 years, people have considered a variety of platforms that provide potential opportunities to process quantum

information, from nuclear magnetic resonance (NMR) [5–7] to the recent state-of-the-art platforms including trapped ions [8–12], neutral atoms [13–16], superconducting circuit [17–20], solid-state defects [21–24] and others [25–27]. Each platform has its own pros and cons due to the differences in the physical structures and mechanisms to control the quantum states. DiVincenzo has pointed out several criteria that a physical system should meet for quantum computing [28], which include the scalability of the system, the ability to preserve quantum coherence and control the quantum states (initialize the states, entangle different parties, and finally measure them). It is still an important open question to find the most suitable platforms for quantum computing or other quantum application tasks.

Things are not always perfect when manipulating the quantum states. The central quantum system cannot be isolated from the environment, which will inject noise into the system and decohere the quantum states. The imprecision of the control pulses may also cause deviation from the ideal outcomes. To reduce the effects of noise, people who work on the hardware physics level or the circuit level have their own methods. On the physical level, we can engineer the shapes or the sequences of the control pulses to perform desired operations with optimized fidelity, gate time, or the suppression of given types of noise. This field of study is categorized as *robust quantum control*. On the circuit level, people try to build the *fault-tolerant* (FT) architecture [29] based on *quantum error correction* (QEC) [30, 31]. The logical states are encoded in a redundant number of physical qubits. The noise from one qubit may not directly cause the corruption of logical information but trigger the change of syndromes, which can be detected from the syndrome measurements and then corrected by the following recovery operations. Further, when the physical error rate is below a certain threshold, the FT schemes confirm the ability to get arbitrarily small logical error by using a reasonable amount of overhead, which can be achieved through code concatenation [32, 33] or simply expanding the code patch [34].

It is widely believed that we are now in the Noisy Intermediate-Scale Quantum (NISQ)

era, as suggested by J. Preskill in 2018 [35]. Till now, people have built quantum processors with up to $O(10^3)$ qubits [16] while the infidelity from an entangling gate as well as the state preparation and measurement (SPAM) is at the level of 0.1% - 1% [11, 15, 20]. A variety of experiments have been performed to show the quantum advantages [18, 25]. However, despite sufficient progress having been made in QEC demonstration [15, 20], there is still a long way to go toward scalable, fault-tolerant quantum computation. People who work on the device physics level and the circuit level should work together to make this dream come true.

1.2 Hardware-efficiency in quantum control and error correction

To preserve the fragile quantum information, the device physicists and the information scientists may have their own specialties in this problem. As a student with a physics background, please allow me to explain my perspective in a bottom-up approach: I will start with the pulse-level quantum control that aims for high-fidelity physical operations despite the presence of noise. Then, I will discuss the opportunities to engineer the structures of qubit errors based on suitable choices of hardware encoding, which will be easier for upper-level QEC codes to correct. Finally, quantum control is needed again to preserve the error structure when executing the gates. All the protocols here are built upon exploiting the physical properties of specific platforms. As a result, people use the word hardware-efficiency to name such a paradigm.

1.2.1 Quantum control

The first question in quantum control is always the controllability of the system. Basically, people want to know what unitary operations can be achieved by manipulating the accessible degrees of freedom. Though in principle it is a math problem to compute all achievable generators that will form the allowed dynamical group [36, 37], practically people will use

the numerical optimal control to find suitable pulses for the desired gates [38–40].

However, the actual Hamiltonian we implement may deviate from the ideal one due to the control imperfection. We call it coherent noise since the output state will be different from the desired one by a unitary process. Composite or shaped pulses are designed to solve this problem, which work well in scenarios including the imprecise amplitude or frequency in qubit control [41, 42], and the reduction of coherent leakage into outer subspace [43]. Other noise, like the thermal relaxation, comes from the coupling with the environment [44]. We name it incoherent noise since it will decohere the quantum state. There are some pulse-level techniques like the dynamical-decoupling method [45, 46] to effectively isolate the central system and reduce the noise from dephasing. However, there may not be a ubiquitous solution for every problem. Numerical optimization on the pulse sequences is needed again to get a higher fidelity while taking the noise into account [47]. Meanwhile, modern optimization schemes based on reinforcement-learning have been developed for the quantum control tasks [48], which are now widely adopted in different platforms to improve the performance of the desired quantum process [49, 50].

On the other hand, a heuristic understanding of the hardware physics may help to simplify the search for high-fidelity control sequences. For example, empirically we can mitigate the incoherent error by accelerating the gates, if the noise rate is not affected at the same time. This applies to a resonator that either has self-nonlinearity or is coupled with another nonlinear device [51, 52]. The operation speed is boosted through cavity displacement while the effective loss rate is unchanged. Then we can only focus on the dynamics in the displaced frame while optimizing the sequences of the displacement parameter for desired operations. In this way, we successfully explore the large photon-number regime to speed up the gate without dealing with the high-dimensional Hilbert space in numerical optimal control.

In Chapter 2, I will justify the above picture with an engineered truncation in the displaced frame, which is first proposed in a Kerr resonator with both linear and parametric

drives on it [53]. From the control aspect, the truncation not only provides the universal controllability within the truncated subspace but also greatly reduces the computational overhead to find the desired drives. I will also show that, in principle, only linear drives are sufficient to realize the truncation from a dynamical point of view. This observation further reduces the hardware complexities, albeit at the expense of higher power demand.

1.2.2 Quantum error correction

Another approach to deal with the physical noise is through error correction, which fundamentally relies on the introduction of redundancy. Similar to the standard QEC where people encode one or few logical qubits into many physical qubits, the hardware-efficient QEC approach suggests that we can also encode the logical information into a multi-level physical system, and use the redundant levels in the Hilbert space to detect and correct the error.

From the device perspective, the first candidate to provide level redundancy is a resonator, which contains an infinite number of levels with equally spaced energy. Two important code designs are the cat code [54–61] and the Gottesman-Kitaev-Preskill (GKP) code [50, 62–65], which are both constructed from the extension of coherent states while exploring the rotational [66] or translational symmetries of the phase space, respectively. With this, the cat code can correct finite order of dephasing noise, while GKP outperforms the others in correcting the photon losses [67].

In this thesis, I will discuss different variants of the cat code. Let me take one more step on it. The codeword of the cat qubit is encoded by two separated coherent states $|\pm\alpha\rangle$. The code subspace can either be autonomously stabilized under an engineered driven-dissipative process [54], or confined by a large energy gap provided from a Kerr nonlinear Hamiltonian [57]. Since most physically relevant noise like photon loss or gain only acts locally on the phase space, it will be difficult for the noise to flip one code state into another.

Therefore, the bit-flip (X) error in the cat qubit is suppressed by the code construction directly, which leaves a Pauli bias in the error channel. The biased error structure could greatly simplify the upper-level QEC construction, since there is only one type of error (Z) left to correct. A repetition code will be sufficient for this task [59, 61, 68].

However, there is still one caveat in the above argument, which comes from the requirement that the Pauli error bias needs to be preserved even after gate operations. Solutions to this problem are first proposed in two theoretical works based on either dissipative cat [68] or the Kerr cat [69] framework. In these approaches, a set of bias-preserving (BP) operations was designed to prevent the generation of bit-flip errors from physical noise during gate execution. Further, the BP operation set including CX and Toffoli (CCX) gates is proved to be sufficient for universal computing with the upper-level repetition code [68], which concludes the full architecture based on the biased error pattern.

On the other hand, people can keep improving the ability of error correction from the hardware encoding. A photon loss can be corrected with the 4-legged cat [54]. Meanwhile, a multimode generalization named pair-cat has been proposed to continuously monitor the photon loss during operations, and there will still be Pauli error bias coming from the uncorrectable noise [70]. In Chapter 3, I will discuss how to generalize the BP gate construction from the cat code to the pair-cat, while keeping the fault-tolerance against a photon loss error.

There are also Fock-state-based designs to tackle the photon loss [71–73], with the dual-rail encoding being the simplest construction [74–76]. By encoding the logical qubit into the single-photon subspace across two different modes, the photon loss event can be detected through monitoring the population in the ground state. Although the qubit information is erased in the noisy process, the error-detecting feature helps us to identify which dual-rail qubit suffers the error [77, 78]. Such an error structure is called the erasure error. The location information provided by the erasure errors is highly beneficial for the subsequent

decoding steps in the upper-level QEC code, since a code with distance d can tolerate $(d-1)$ many erasure errors, twice as many as the random Pauli errors [79]. Meanwhile, the simplicity in the dual-rail structure also facilitates high-fidelity operations [80], which, together with the erasure error model, makes it a competitive design in the hardware-efficient QEC code family.

Within the past ten years, we have also observed the fast development of the neutral atom platform in the field of quantum information processing. The rich structure of atomic inner levels also motivates the development of novel hardware-efficient QEC schemes. The noise model here is also different from the bosonic counterpart, as the dominant error mechanism here will be the Rydberg decay during the execution of entangling (CZ) gates. These features first led to the proposal on the BP operations, where two code basis are designed to have a large separation in magnetic quantum number [81]. Moreover, a metastable encoding makes it possible to detect the Rydberg decay after the entangling operations, which again converts the leakage into the erasure type [82]. A further modified scheme tries to combine the two merits together, where both the location (erasure) and the Pauli type (Z error bias) will be available to us if the leakage is detected [83]. Such a biased erasure error is extremely structured and will provide around a 10% threshold in the XZZX surface code simulation, in contrast to the 1% result using the depolarizing noise model. The threshold can be further improved if we can find a native biased erasure CX design [83, 84], though its physical construction is still an open question.

In Chapter 4, I will borrow the concept of cat encoding from bosonic systems to atomic levels. Here we use two antipodal spin-coherent states (SCS) in the ground hyperfine manifold to encode the qubit, which is the same as that in [85]. In our protocol, the code subspace can be autonomously stabilized while utilizing the spontaneous emission from the excited levels. Therefore, our encoding is named the dark spin-cat. We can design single-qubit operations here by analogy with the bosonic counterpart, while the existing biased-erasure

CZ protocol based on the Rydberg interaction can also be seamlessly applied to the spin-cat qubit. We then propose a possible solution to execute the CX gate on the spin-cat with the biased-erasure error structure, though it could be practically demanding.

From the physical aspect, there are still certain analogies between the bosonic cat and the dark spin-cat for us to explore. In Chapter 5, I will focus on the single-atom model with a generalized definition of the dark spin-cat. I will discuss the similarities of the stabilization mechanisms between the two systems, while using a semiclassical approach to analyze the bit-flip rate in the generalized spin-cat when it is subjected to the noise. I will explain if the common tricks for bit-flip suppression in bosonic cat can be transferred onto our spin-cat model, with a discussion on their limitations.

1.3 Efficient quantum state characterization

Both quantum control and error correction aim to reduce the effect of noise during quantum information processing. However, the resulting quantum state in the experiment may still be unknown to us. Sometimes we need to extract certain information from a given state, like the fidelity estimation task, where we want to see how close it is to the desired one. Sometimes we might be more ambitious to learn the full information of the state, where we call this task state tomography.

However, those tasks could be resource-demanding as the system size grows. An M-qubit quantum system has a $D = 2^M$ dimensional Hilbert space, and therefore it requires $(2^{M+1} - 2)$ number of real parameters to determine a pure quantum state in it. If the state has experienced some noise and cannot be treated as a pure one, we should use a density matrix to describe it, and the number of unknown real parameters will become $(2^{2M} - 1)$ in the most generic case. Such an exponentially growing number is the major challenge for people to perform full state tomography on an unknown state in a large system, since we need to get the expectation values with the same amount of linear-independent observables for

informational completeness. There have been demonstrations to fully reconstruct a quantum state with more than 10 qubits in an experiment [86], but we do not expect the allowed qubit number will increase too much in the future due to the exponential overhead. There have been proposals that use adaptive methods [87] or even machine learning techniques [88, 89] to optimize the choice of measurement basis each time, but none of these have shown the ability to resolve the exponentially growing overhead.

To achieve efficiency in the full state tomography, one should further exploit the structures of the unknown state. For example, if we know in prior that the density matrix has a low rank $r \ll D$, then the expectation values over $O[rD \text{ poly}(\log D)]$ different Pauli operators will be sufficient to fully reconstruct the state [90, 91]. Compared with the previous $O(D^2)$ requirement, such a compressed sensing technique indeed provides a scaling improvement on the resource overhead. Another direction is to explore the underlying physics of the state. To stress the local correlations in the physical system, one may use the form of a Matrix Product State (MPS) to represent a pure state [92–94], or a Matrix Product Operator (MPO) to represent a density matrix [95, 96]. In these cases, information from the local reduced density matrix could be sufficient for full state reconstruction, which makes the $O[\text{poly}(M)]$ overhead feasible. However, existing protocols either require a sophisticated post-processing method without a rigorous convergent proof [93, 96], or need extra assumptions on the states themselves [95].

In Chapter 6 of this thesis, I will show that a collective truncation over the multi-party Hilbert space may provide another opportunity to achieve efficient state tomography. For example, given a limited number of total photons in an M-mode resonator, or an M-qubit stabilizer state with a bounded-weight random Pauli error, the dimension of the relevant Hilbert subspace only scales polynomially with the system size M rather than exponentially. Further, if each matrix element can be estimated with a cost independent of the system size, then the overall scheme will be provably efficient. In addition to a theoretical justification,

we also experimentally demonstrate this idea using a multimode resonator [97]. The efficient tomography is performed in a maximally one-photon subspace across up to 4 different modes.

The fidelity estimation task is somewhat different, since there is only one parameter to learn. If arbitrary unitary operations are allowed before measurement, then we can always map the target state to a computational basis state and measure it thereafter to extract the fidelity. As a result, we need to confine the allowed gates and measurements before discussing the efficiency in fidelity estimation. For example, early studies showed that the fidelity with a given stabilizer state can be efficiently characterized using Pauli measurement only, where the measurement overhead is independent of the system size [98]. For a generic multi-qubit target state, it is still an open problem to see if we can use a poly(M) overhead to characterize the fidelity with it, when single-qubit measurements with adaptively chosen basis are allowed [99].

With the development of both hardware performance and the QEC protocols, now people gradually switch to manipulate the quantum state from physical to logical level [15, 100–102]. As the infidelity ϵ of the prepared (logical) state approaches zero, the characterization overhead should increase accordingly to ensure that the estimation precision remains comparable with the vanishing infidelity itself. On the other hand, the native logical operations on a given QEC code are usually limited, which makes it difficult to directly measure the overlap with the target state. In the first logical-level magic state distillation (MSD) experiment [102], researchers try to use the logical Pauli tomography to reconstruct the final state, but the overhead should be $O(1/\epsilon^2)$ to meet the precision requirement. In Chapter 7, I will discuss a protocol that successfully reduces the cost to $O(1/\epsilon)$ scaling, leveraging a clever trick that is fundamentally connected with the concept of superresolution in quantum metrology [103].

1.4 Structure of the thesis

This thesis contains some of the projects that I was involved in during my time as a Ph.D. student. The overall aim is to reduce the effects of noise on the gate operations through quantum control, or engineering the structure of noise through hardware-efficient encoding, which makes it easy to be corrected by upper-level QEC codes. Both of the schemes utilize a dynamical truncation of the full Hilbert space for either simplification of the pulse optimization or autonomous stabilization of the encoded subspace. Further, a truncation of the Hilbert space or only focusing on a subset of states with certain structures enables the possibility of efficiently characterizing the states via full tomography or fidelity estimation. I will place two examples in the thesis.

Chapter 2 [104] covers a case study about utilizing a novel blockade effect in a Kerr nonlinear resonator for universal high-fidelity operations in the presence of loss. With the Kerr nonlinearity, carefully chosen linear (one-photon) and parametric (two-photon) drives can lead to a destructive interference for the coupling between two nearest Fock states viewed from a displaced frame, which confines the dynamics within the truncated subspace. Arbitrary operations in the blockade subspace can be achieved by controlling the displacement parameter. The reduced dimension from the blockade simplifies the pulse optimization for desired operations, while the improved gate speed boosted by displacement helps to mitigate the effects of the loss on the fidelities. Further, a scheme is proposed to relax for parametric drives while effectively preserving the blockade structure, at the expense of more power consumed.

Chapter 3 [105], 4 [106], and 5 are a series of works on the generalization of the bosonic cat encoding with different platforms. I first start with the pair-cat code, a multimode variant with the ability to continuously correct a photon loss, and discuss the construction of the bias-preserving (BP) operations set with performance analysis. Later, I will talk about how to extend the cat encoding into a long (but finite-dimensional) hyperfine manifold of an atom.

Utilizing the structure of the dark states, we also achieved the autonomous stabilization for the encoded dark spin-cat subspace. Further, BP operations are also designed with the erasure conversion tricks, which are tailored for the noise from the Rydberg decay during the entangling gates. Finally, I will discuss a more generalized spin-cat encoding fitted in the dark-state structure, and highlight the beautiful analogy with the bosonic cat in certain aspects.

Chapter 6 [97] and 7 [107] focus on the characterization of quantum states with certain structures. First, I will provide an example of a tomographic task for quantum states in a multimode resonator, whose maximally allowed photon numbers are bounded. The truncation of the multimode Hilbert space significantly reduces the number of unknown variables to a polynomial dependence on the mode number. Further, our hardware-efficient Wigner measurement approach enables the cost of estimating individual parameters to be independent of the total mode number, which concludes the justification of our efficiency claim. Another example is about the efficient logical fidelity estimation of magic states in the FT regime. Located at the highly symmetrical points in the Hilbert space under Clifford operations, the magic states will support a simplified form of noise after Clifford twirling, while the fidelities stay the same. The simplification makes it possible to estimate the small infidelity ϵ with only $O(1/\epsilon)$ cost to get a precision comparable with ϵ itself, which is in contrast to the $O(1/\epsilon^2)$ scaling provided by a straightforward Pauli tomography [102]. I will provide several circumstances where the $O(1/\epsilon)$ cost is achievable, and a numerical simulation to show the power of our protocol under practical consideration.

CHAPTER 2

UNIVERSAL CONTROL WITH NOVEL PHOTON BLOCKADE IN WEAK-KERR NONLINEAR RESONATORS

2.1 Introduction

Bosonic systems, such as photons in optical or microwave resonators, are a promising platform for quantum information processing. In contrast to qubits, the infinite-dimensional bosonic Hilbert space provides novel ways to encode and robustly process quantum information in a hardware-efficient manner [108]. A challenge however is the need for nonlinear operations [109]. It has been shown that, to achieve universal control in bosonic systems it is necessary and sufficient to have at least one kind of nonlinear operation in addition to linear operations, i.e., unitary evolution under a Hamiltonian linear or quadratic in bosonic raising and lowering operators [110, 111].

In the microwave regime, a typical approach to introduce nonlinearity is to couple the bosonic system directly to a qubit [52, 112–114]. However, if we want to relax the cooling requirements, the resonant frequency of the system needs to become higher. Unfortunately, even in the millimeter wave regime (around 100 GHz), the production of a superconducting qubit with strong nonlinearity will become challenging, while the performance of the state-of-the-art demonstrations is still limited [115]. Similarly, in the optical regime, people may couple a cavity mode with atoms in so-called cavity QED systems. However, due to the low cooperativity in practice, it is again difficult to deterministically prepare any useful non-Gaussian state with this setup while achieving a high fidelity [116, 117].

Given this, it would be ideal to exploit intrinsic nonlinearities in optical or microwave resonators. An extremely common example is a self-Kerr nonlinearity. Examples include micro-ring resonators or photonic crystals with $\chi^{(3)}$ nonlinearities [118, 119] and quantum LC circuits that contain superconducting materials with high kinetic inductance [120, 121].

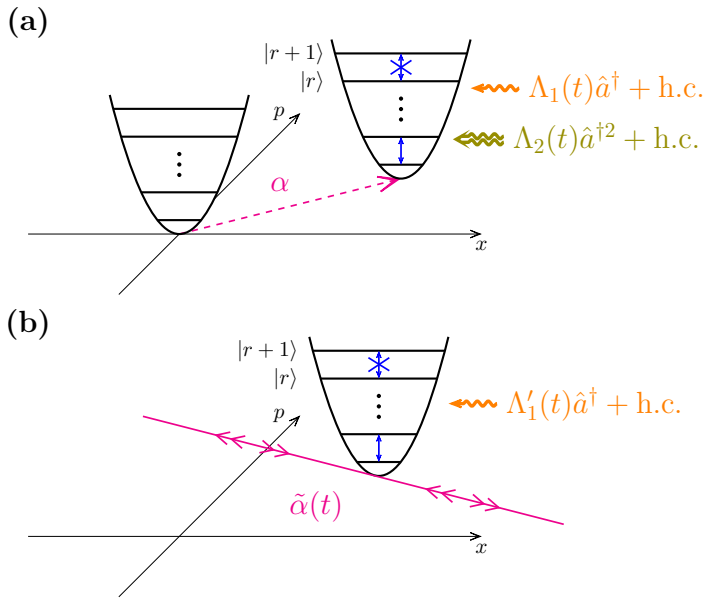

Figure 2.1: Schematic diagram of the photon blockade in the displaced rotating frame with (a) both 1-photon and 2-photon drives, (b) only 1-photon drive with fast oscillation of the displaced rotating frame in phase space.

While in principle these nonlinearities are sufficient for universal control, in practice they are often much weaker than photon loss rates, precluding the ability to achieve high-fidelity nonlinear operations.

While nonlinearities weaker than loss might seem to be of no use for quantum nonlinear operations, recent research suggests that this might not be the case [51]. In this work, the authors propose an intriguing scheme to deterministically prepare a single-photon Fock state using a Kerr nonlinearity, even in cases where this nonlinearity is significantly smaller than the loss rates. This scheme relies on a novel photon blockade phenomenon that requires the displacement of a single bosonic mode together with carefully chosen 1-photon and 2-photon driving amplitudes. Notably, the speed of the operations in the blockade subspace is enhanced by the displacement amplitude, which can be large enough to counteract the effects of loss. A similar idea related to displacement-boosted gates is also presented in Ref. [52].

Here, we generalize this blockade scheme to demonstrate its applications beyond the preparation of single-photon Fock states. We show that using this scheme, one can perform

any unitary operation in a blockaded subspace of Fock states with an arbitrarily chosen dimension. In addition to formal proof in the ideal scenario, we present a gradient-based optimization algorithm to explicitly find the control pulse sequences to implement a desired unitary operation in the blockaded subspace. Another control scheme within a qudit truncated from a resonator mode is discussed in [122], but the need for a coupled qubit in the protocol there simplifies the problem and is beyond the setup that we consider here. Meanwhile, the photon blockade feature viewed in the displaced frame of the Kerr resonators [53] could help to suppress the bit-flip rate with Kerr-cat encoding [123, 124], which is again an interesting application derived from this formalism and will also be used in Chapter 5 for the generalized spin-cat encoding in atomic structures.

In practice, directly implementing the required 2-photon drive can be a challenging task, as this can require additional weak nonlinearities. For example, in certain platforms one could pump an auxiliary mode that interacts nonlinearly with the central mode [125]. Here we show that one can eliminate the need for an explicit 2-photon drive, by instead simply time-modulating the amplitude of the single-photon drive. This is similar in spirit to the operation of double-pumped parametric amplifiers [126].

The optimization algorithm and this modulation scheme are then integrated seamlessly to implement arbitrary operations using only a standard linear, single-photon drive. In practice, however, a larger driving amplitude may result in much stronger input power or the violation of the rotating wave approximation (RWA), while the lack of native 2-photon drive will make those problems more severe. We discuss how these problems limit the achievable fidelities for both protocols with or without the native 2-photon drives, as well as possible methods to mitigate the resulting imperfections. We argue that even with these limitations, both protocols could perform well in experimental platforms where the strength of the single-photon self-Kerr nonlinearity is comparable to or slightly smaller than the loss rate, which is still encouraging for people to further develop clever control schemes for weak-nonlinear

systems and explore future applications.

2.2 System setup

We consider a setup similar to Ref. [51], consisting of a single-mode resonator with self-Kerr nonlinearity, subject to both 1-photon and 2-photon drives. Within the RWA, the Hamiltonian of the system can be written as

$$\hat{H} = \frac{\chi}{2}\hat{a}^{\dagger 2}\hat{a}^2 + \omega_c \hat{a}^\dagger \hat{a} + [\Lambda_1(t)e^{-i\omega_1(t)t}\hat{a}^\dagger + \Lambda_2(t)e^{-i\omega_2(t)t}\hat{a}^{\dagger 2} + \text{h.c.}], \tag{2.1}$$

where χ indicates the strength of the single-photon Kerr nonlinearity, ω_c is the angular frequency of the resonator, and $\Lambda_{1(2)}$ and $\omega_{1(2)}$ are the amplitude and the frequency of the 1(2)-photon drives, respectively. Here, h.c. denotes the Hermitian conjugate of the terms in the bracket. As shown in Ref. [51], by choosing $\omega_2 = 2\omega_1$, going to a frame rotating with $\omega_1(t)$ and then displaced by $\alpha(t)$, we can obtain a blockaded Hamiltonian of the form

$$\hat{H}_{\text{dr}}[\alpha(t)] = \frac{\chi}{2}\hat{a}^{\dagger 2}\hat{a}^2 + \Delta_0 \hat{a}^\dagger \hat{a} + [\chi\alpha(t)\hat{a}^\dagger(\hat{n} - r) + \text{h.c.}], \tag{2.2}$$

where r is an adjustable positive integer determining the dimension of the blockade subspace, and Δ_0 is the detuning term in the displaced rotating frame. To achieve this effective Hamiltonian for a given $\alpha(t)$, Δ_0 and integer r, one needs time-dependent drive amplitudes and frequencies chosen such that

$$\begin{cases} \Lambda_1(t) = \chi\alpha(t)[2|\alpha(t)|^2 - r] - \Delta_0\alpha(t) + i\dot{\alpha}(t), \\ \Lambda_2(t) = -\frac{\chi}{2}\alpha^2(t), \\ \omega_1(t) = \omega_c - \Delta_0 + \frac{1}{t}\int_0^t 2\chi|\alpha(t')|^2\,\mathrm{d}t'. \end{cases} \tag{2.3}$$

The key term in Eq. (2.2) is the nonlinear drive, which has no coupling between $|r\rangle$ and

$|r+1\rangle$ levels. Therefore, the dynamics generated by $\hat{H}_{\rm dr}$ is constrained within the blockade subspace $\mathcal{H}_b$ spanned by $\{|0\rangle, |1\rangle, \ldots, |r\rangle\}$. Also, the amplitude of this nonlinear drive is set by the time-dependent displacement $\alpha(t)$, which serves as a key control parameter in what follows. Previously, it was shown that with a static α, one could use $\hat{H}_{\rm dr}$ to generate the Fock $|1\rangle$ state on a timescale much shorter than $1/\chi$ [51]. Here, we go much further: we show that in fact, by using a time-dependent $\alpha(t)$, one can generate any unitary within the blockade subspace $\mathcal{H}_b$, whose dimension can also be chosen freely.

To account for the effects of photon loss in our system, we use the master equation

$$\frac{\mathrm{d}\hat{\rho}}{\mathrm{d}t} = -i[\hat{H}, \hat{\rho}] + \kappa \mathcal{D}[\hat{a}]\hat{\rho}, \tag{2.4}$$

where $\mathcal{D}[\hat{a}]\hat{\rho} = \hat{a}\hat{\rho}\hat{a}^\dagger - \frac{1}{2}\{\hat{a}^\dagger\hat{a}, \hat{\rho}\}$ is a dissipator that models the photon loss effect and κ is the rate of this process. Note that when transforming to the same displaced rotating frame mentioned earlier, the 1-photon driving amplitude $\Lambda_1(t)$ in Eq. (2.3) now requires the addition of an extra term, namely $i\kappa\alpha(t)/2$, in order to obtain the same Hamiltonian $\hat{H}_{\rm dr}[\alpha(t)]$. However, the form of the dissipator $\mathcal{D}[\hat{a}]$ and the associated loss rate κ remain unaltered in this new frame. Consequently, due to the enhancement of the nonlinear blockade drive by α in $\hat{H}_{\rm dr}$, the operations can be performed on a significantly shorter timescale than $1/\chi$. This, in turn, allows for the mitigation of the impact of photon loss, presenting an opportunity to achieve high-fidelity gates, even when $\chi \ll \kappa$.

2.3 Demonstration of universality

We sketch the proof of universal controllability of our system governed by the Hamiltonian $\hat{H}_{\rm dr}$ shown in Eq. (2.2) here and refer the reader to Appendix A.1 for more details. We first focus on the dynamics within the blockade subspace $\mathcal{H}_b$. Let $\hat{\Pi}_r := \sum_{n=0}^{r} |n\rangle\langle n|$ denote the projector to this $N = r+1$ dimensional subspace. The projection of $\hat{H}_{\rm dr}$ to the blockade

subspace is given by

$$\hat{\Pi}_r \hat{H}_{\rm dr} \hat{\Pi}_r = \hat{H}_{d,0} + \chi {\rm Re}[\alpha(t)] \hat{H}_{c,R} + \chi {\rm Im}[\alpha(t)] H_{c,I}, \tag{2.5}$$

where

$$\begin{cases} \hat{H}_{d,0} = \sum_{n=0}^{r} \left[\chi(n^2 - n)/2 + \Delta_0 n \right] |n\rangle\langle n|, \\ \hat{H}_{c,R} = \sum_{n=0}^{r-1} (n-r)\sqrt{n+1}(|n+1\rangle\langle n| + \text{h.c.}), \\ \hat{H}_{c,I} = i \sum_{n=0}^{r-1} (n-r)\sqrt{n+1}(|n+1\rangle\langle n| - \text{h.c.}). \end{cases} \tag{2.6}$$

In what follows (using the language of quantum control theory, see e.g., [127]) $\hat{H}_{d,0}$ serves as the drift Hamiltonian, while $\hat{H}_{c,R}$ and $\hat{H}_{c,I}$ are the control Hamiltonians. The real and imaginary parts of $\alpha(t)$ are time-dependent functions that can be controlled. Following Ref. [112], we define universal control of a quantum system as the ability to realize any unitary operation $\hat{U}_{\rm tar}$ in the U(N) group via a properly chosen control [$\alpha(t)$ in this case] and evolution time T.

A theorem in Ref. [37] suggests a sufficient condition for the Hamiltonian to make the system universally controllable. It has two requirements. First, the drift Hamiltonian $\hat{H}_{d,0}$ should be diagonal (with eigenvalues denoted as E_k for eigenstates $|k\rangle$) and contain certain type of nonlinearity, specifically, the nearest energy difference $\mu_k := E_k - E_{k+1}$ should satisfy $\mu_0 \neq 0$ and $\mu_k^2 \neq \mu_0^2$ for $k > 0$ (or similarly $\mu_{N-2} \neq 0$ and $\mu_k^2 \neq \mu_{N-2}^2$ for $k < N-2$). Second, one of the control parts $\hat{H}_{c,j}$ should only have couplings between $|k\rangle$ and $|k+1\rangle$ for all $0 \leq k < N-1$. If both these conditions are satisfied, then the generated dynamical Lie group will be U(N) when ${\rm Tr}[\hat{H}_d] \neq 0$ and SU(N) otherwise.

We can easily verify that those requirements for U(N) group (where $N = r+1$ in our case) generation are satisfied with our $\hat{H}_{d,0}$ and $\hat{H}_{c,R}$ as long as $r \neq -\frac{2\Delta_0}{\chi} + 1$. This allows us to fix $\Delta_0 = 0$ for $r \geq 2$ in the rest of the main text. Moreover, the two control degrees of freedom $\hat{H}_{c,R}$ and $\hat{H}_{c,I}$ provide the possibility to do any gate (up to a global phase) in

an arbitrarily fast manner, as these two are also sufficient to generate full SU(N) group (see Appendix A.1). Finally, as r is also adjustable, we can choose any blockade dimension we want. Therefore, unitary operations defined in any finite dimension are in principle achievable.

2.4 Optimal control

The generalized blockade phenomena allow one, in principle, to perform an arbitrary unitary operation in an arbitrarily chosen N-dimensional blockaded subspace in a time much faster than $1/\chi$. The question that we now address is how to design a particular control $\alpha(t)$ to realize a target unitary. In contrast to the conventional setting, where one optimizes the control in the rotating frame of the drive [128], we consider $\hat{H}_{\rm dr}$ defined in the instantaneous displaced rotating frame and optimize $\alpha(t)$, the frame parameter. Consequently, finding an optimal $\alpha(t)$ directly determines the corresponding physical parameters $\Lambda_{1(2)}(t)$ and $\omega_1(t)$ via Eq. (2.3) required for implementing the desired operations in the laboratory. In other words, by hard coding the blockade condition in the evolution through Eq. (2.3), we simplify the optimization task to finding $\alpha(t)$ in the small blockade Hilbert space. In this way, even though the required drive amplitudes $\Lambda_{1(2)}(t)$ can be very large, we do not need to consider large photon number states in the optimization. In this section, we only optimize the controls for a fixed evolution time T in the absence of loss. We further consider the effects of loss and a realistic experimental constraint on total input power (see Appendix A.2); as we show, these additional features lead to there being an optimal choice of gate time T.

Typically, optimal control algorithms such as GRAPE [38, 39] discretize the control pulse $\alpha(t)$, and maximize a figure of merit such as the fidelity by performing gradient-based optimization on the control parameters, i.e., the amplitude of the control at discrete time points. This standard approach would yield a piece-wise constant $\alpha(t)$, something that is highly problematic for our setup: discontinuous jumps in $\alpha(t)$ would require infinite driving

power to implement, as the 1-photon driving amplitude has a term proportional to $i\dot{\alpha}(t)$ [see Eq. (2.3)].

To solve this issue, in contrast to conventional methods, we expand $\alpha(t)$ on a basis of continuous functions. Since it is desirable to be in a non-displaced frame, i.e. $\alpha(0) = \alpha(T) = 0$ in the beginning ($t = 0$) and at the end ($t = T$) of the protocol, we use the following sine-basis ansatz for the control pulse

$$\alpha(t) = \sum_{k=1}^{k_{\max}} \alpha_k \sin\left(\frac{k\pi t}{T}\right). \tag{2.7}$$

Here, $k_{\max}$, the cutoff number for the highest harmonics that we use, is a hyperparameter that is chosen according to the complexity of the task, and α_k are complex-valued optimization variables. In practice, we choose $k_{\max}$ heuristically in the optimization procedure. If the fidelity achieved from the optimal pulses is lower than our target, we increase $k_{\max}$ for better performance. To implement a target unitary operation in N dimensions, we maximize $F_{\mathrm{u}}(\{\alpha_k\}) = \left|\mathrm{Tr}[\hat{U}_{\mathrm{tar}}^{\dagger}\hat{U}(T)]\right|^2/N^2$, where $\hat{U}(T)$ is obtained by solving $\frac{\mathrm{d}}{\mathrm{d}t}\hat{U}(t) = -i\hat{H}_{\mathrm{dr}}[\alpha(t)]\hat{U}(t)$ for $t = T$ with the initial condition $\hat{U}(0) = \hat{I}$. When optimizing F_{u}, we implicitly ignore the irrelevant global phase. Note that the dependence of the objective on $\{\alpha_k\}$ originates from the dependence of $\hat{H}_{\mathrm{dr}}$ on $\alpha(t)$ according to Eq. (2.2). Moreover, while obtaining F_{u} involves solving an ordinary differential equation, it is still differentiable and its gradient with respect to $\{\alpha_k\}$ can be calculated using the chain rule and the adjoint sensitivity method [129]. Therefore, we can use gradient-based optimization to find locally optimal $\{\alpha_k\}$.

To illustrate the universal controllability of our scheme, we consider the problem of implementing the permutation $\hat{U}_{\mathrm{P}}$ or the Fourier transformation $\hat{U}_{\mathrm{FT}}$ in a 3-level blockade

subspace up to a global phase spanned by $\{|m\rangle\}_{m=0}^{2}$, where

$$\hat{U}_{\mathrm{P}} = |2\rangle\langle 0| + |0\rangle\langle 1| + |1\rangle\langle 2|, \tag{2.8a}$$

$$\hat{U}_{\mathrm{FT}} = \frac{1}{\sqrt{3}} \sum_{m,n} e^{i\frac{2\pi mn}{3}} |m\rangle\langle n|. \tag{2.8b}$$

We use the automatic differentiation toolbox of JAX [130], a numerical computing package, to perform the gradient-based optimization. We choose the evolution time $T = 0.2/\chi$. We also fix $\alpha(0) = \alpha(T) = 0$ and $\Delta_0 = 0$ in $\hat{H}_{\mathrm{dr}}$ and find the pulses $\alpha(t)$ that implements the two unitary operations of interest. In both cases, the algorithm finds a solution such that $\left|\mathrm{Tr}[\hat{U}_{\mathrm{tar}}^{\dagger}\hat{U}(T)]\right|^2/N^2 > 1 - 10^{-4}$ (see Fig. 2.2).

2.5 Only 1-photon drive

So far, we have focused on the fundamental questions about controllability by explicitly constructing the 1- and 2-photon drives to achieve fast universal control for weak-Kerr systems in the presence of loss. Here, we move away from the ideal scenario and discuss issues relevant to experimental implementations. One question to address is whether one truly needs a distinct 2-photon drive. While this can be done in some platforms, e.g., superconducting qubits with flux-pumping [111], it can be challenging in other platforms. Here, we present a method that allows our control scheme to be implemented without any explicit independent 2-photon drive. As we show, this idea is intrinsically connected to squeezing by double-pumping a Kerr resonator [126].

In the absence of an independent 2-photon drive, $\Lambda_2(t) = 0$ in Eq. (2.1). Again, we set $\Delta_0 = 0$ and choose the 1-photon driving amplitude such that in a frame rotating with $\omega_1(t)$ and displaced by $\tilde{\alpha}(t)$, we have the same blockade drive. In this frame, the Hamiltonian is

$$\hat{H}'_{\mathrm{dr}}[\tilde{\alpha}(t)] = \frac{\chi}{2}\hat{a}^{\dagger 2}\hat{a}^2 + [\chi\tilde{\alpha}(t)\hat{a}^{\dagger}(\hat{n} - r) + \mathrm{h.c.}] + \left[\frac{\chi}{2}\tilde{\alpha}^2(t)\hat{a}^{\dagger 2} + \mathrm{h.c.}\right], \tag{2.9}$$

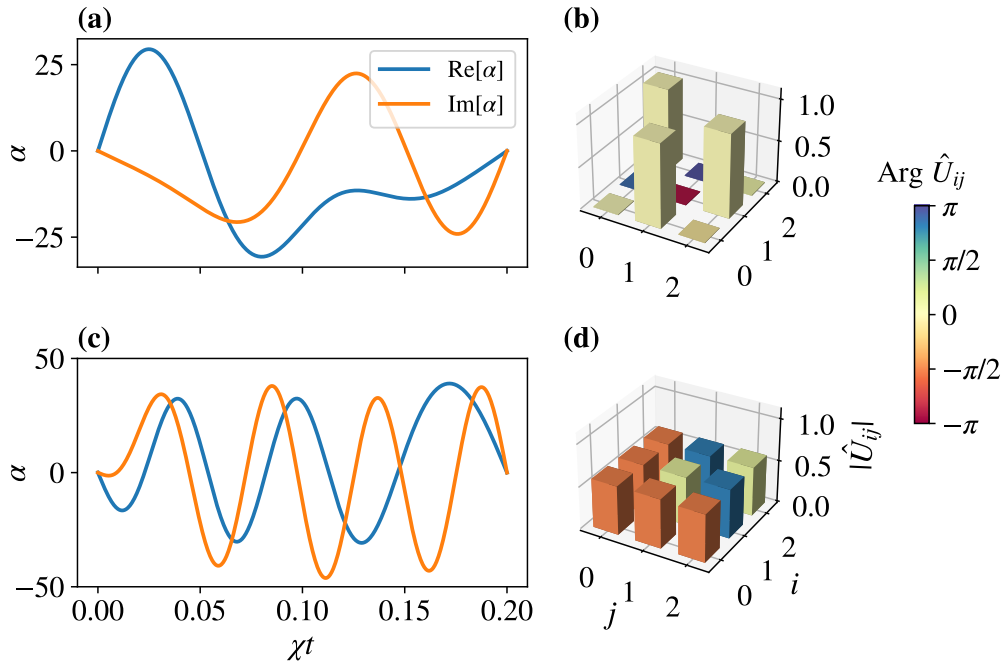

Figure 2.2: Optimal control for implementing arbitrary unitary operations. (a) The optimized $\alpha(t)$ function that implements permutation (2.8a) in a 3-level blockade subspace with $k_{\max} = 5$ harmonics. Total evolution time $T = 0.2/\chi$. (b) The elements of the unitary $\hat{U}(T)$ generated by $\alpha(t)$ that is shown in (a) under the Fock basis, i.e., $\hat{U}_{ij} = \langle i|\,\hat{U}(T)\,|j\rangle$. The height and color of each bar show the absolute value and argument of the corresponding matrix elements. (c) Optimal $\alpha(t)$ for implementing Fourier transform (2.8b) in the same blockade subspace with $k_{\max} = 8$ harmonics and $T = 0.2/\chi$. (d) the matrix elements of $\hat{U}(T)$ generated by $\alpha(t)$ shown in (c).

where the 1-photon driving amplitude is chosen to have the form

$$\Lambda_1'(t) = \chi\tilde{\alpha}(t)[|\tilde{\alpha}(t)|^2 - r] + i\dot{\tilde{\alpha}}(t) + i\kappa\tilde{\alpha}/2. \tag{2.10}$$

While the nonlinear single-photon drive in Eq. (2.9) has the correct form, the induced two-photon drive term in the last line violates the desired blockade condition.

A key observation in our strategy to revive the blockade condition is that the desired fully blockaded Hamiltonian can be written as $\hat{H}_{\rm dr}[\alpha(t)] = \frac{1}{\sqrt{2}}\{\hat{H}'_{\rm dr}[e^{-\frac{i\pi}{4}}\alpha(t)] + \hat{H}'_{\rm dr}[e^{\frac{i\pi}{4}}\alpha(t)]\}$. Intuitively, by alternating the phase of $\tilde{\alpha}(t)$, we can average away and cancel the unwanted 2-photon drive term while retaining the desired nonlinear blockade drive. This observation

combined with the Trotter formula

$$e^{-i\frac{\hat{H}_1+\hat{H}_2}{2}\delta T} = e^{-i\hat{H}_1\delta T/4}e^{-i\hat{H}_2\delta T/2}e^{-i\hat{H}_1\delta T/4} + O[(\delta T)^3] \tag{2.11}$$

suggests that setting $\hat{H}_1(t) = \hat{H}'_{\rm dr}[e^{-\frac{i\pi}{4}}\alpha(t)]$ and $\hat{H}_2(t) = \hat{H}'_{\rm dr}[e^{\frac{i\pi}{4}}\alpha(t)]$ and alternating the evolution for $\delta T/2$ between the two Hamiltonians suppress the errors in violating the blockade condition to $O[(\delta T)^3]$[1]. However, this Trotter scheme requires discretizing $\alpha(t)$ into intervals of length δT and implementing instantaneous displacements between $e^{\pm\frac{i\pi}{4}}\alpha$, which in practice introduces additional complexities.

Inspired by the discrete version of the Trotter formula, we design its continuous counterpart via Magnus expansion (see Appendix A.8)

$$\exp[-i\int_0^{\delta T}\hat{H}(t)\,{\rm d}t] = \mathcal{T}\exp[-i\int_0^{\delta T}\hat{H}(t)\,{\rm d}t] + O[(\delta T)^3], \tag{2.12}$$

where $\mathcal{T}$ is the time-ordering operator and $\hat{H}(t)$ for $t \in [0, \delta T]$ is chosen to be symmetric around $\delta T/2$, i.e., $\hat{H}(t) = \hat{H}(\delta T - t)$. As a result, we construct a new function $\tilde{\alpha}(t)$ that oscillates rapidly with time. Specifically, Eq. (2.12) gives us a recipe for finding $\tilde{\alpha}(t)$ such that the coarse-grained evolution under $\hat{H}'_{\rm dr}[\tilde{\alpha}(t)]$ over an interval of δT is close to that under $\hat{H}_{\rm dr}[\alpha(t)]$. This translates to having the average of $\tilde{\alpha}(t)$ over a δT time interval centered on time t to satisfy

$$\overline{\tilde{\alpha}^2(t)} = 0, \quad \overline{\tilde{\alpha}(t)} = \alpha(t), \tag{2.13}$$

where the overline denotes the coarse-graining time average, and $\alpha(t)$ is the optimal choice of function in $\hat{H}_{\rm dr}$ defined in Eq. (2.2) that generates the desired target unitary operation. To satisfy these constraints, we propose using the ansatz $\tilde{\alpha}(t) = \alpha(t)f(t)$, where $f(t)$ is a

1. We should also note that suppressing errors beyond the third order using the discrete Trotter scheme is not possible due to the non-existence of positive decomposition as shown in Ref. [131]. In other words, higher order suppression requires changing the sign of $\hat{H}'_{\rm dr}$, and that is not possible since the sign of the nonlinearity χ is fixed in a device.

periodic function with period $\delta T = \frac{T}{M}$. Here T is the total evolution time and M is the number of periods during the evolution. We also denote $\omega_r := \frac{2\pi M}{T}$ for further use. Besides, to keep the same structure as the Trotter formula in Eq. (2.12), we further require that

$$f(t) = f(\delta T - t). \tag{2.14}$$

Consequently, the constraints for $\tilde{\alpha}(t)$ in Eq. (2.13) translate to constraints

$$\overline{f(t)} = 1, \quad \overline{f^2(t)} = 0 \tag{2.15}$$

on $f(t)$ over over each period. This ensures that the overall evolution under $H'_{\mathrm{dr}}[\tilde{\alpha}(t)]$ and $\hat{H}_{\mathrm{dr}}[\alpha(t)]$ for time T closely resemble each other, that is

$$\mathcal{T} \exp\left\{-i \int_0^T \hat{H}'_{\mathrm{dr}}[\tilde{\alpha}(t)]\, \mathrm{d}t\right\} \approx \mathcal{T} \exp\left\{-i \int_0^T \hat{H}_{\mathrm{dr}}[\alpha(t)]\, \mathrm{d}t\right\}. \tag{2.16}$$

One possible choice of $f(t)$ is the following

$$f_{\mathrm{dp}}(t) = 1 + i\sqrt{2}\cos(\omega_r t). \tag{2.17}$$

An example of the shape of $\tilde{\alpha}(t)$ after implementing this modulation is shown in Fig. 2.3 (and schematic illustration in Fig. 2.1(b) as well). To obtain physical intuition on the underlying mechanism, we can split $f_{\mathrm{dp}}(t)$ into two parts. The constant part plays the role of Λ_1 in Eq. (2.1), and provides the desired nonlinear blockade drive in the displaced rotating frame. In contrast, the time-dependent term, $i\sqrt{2}\cos(\omega_r t)$, corresponds to the double-pumping scheme with driving frequencies $\omega_1 \pm \omega_r$, as it contributes to $\Lambda'_1(t)$ via $i\dot{\tilde{\alpha}}(t)$. This double-pumping will effectively generate the 2-photon driving Λ_2 in Eq. (2.1) from the Kerr interaction, as a result of 4-wave mixing [126].

The approximation in Eq. (2.16) is valid if the number of periods M is large enough such

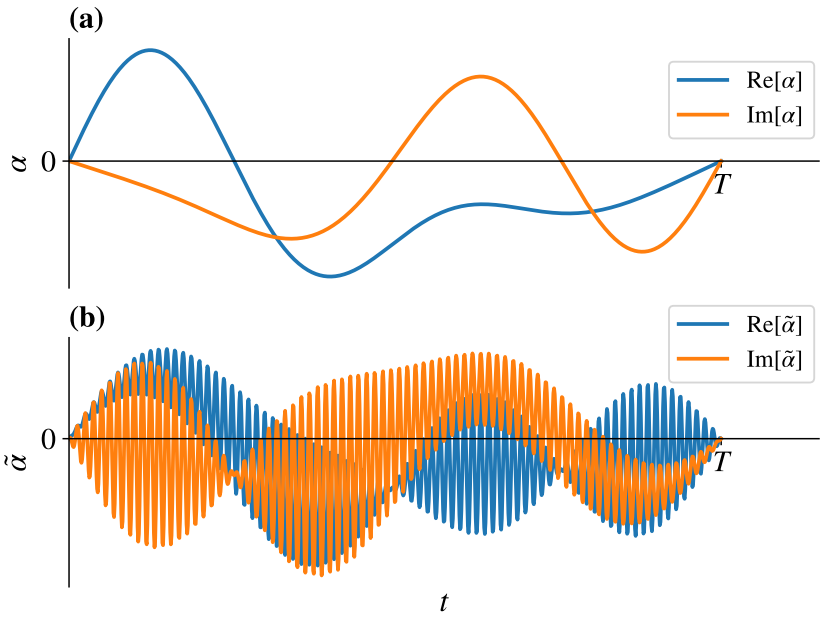

Figure 2.3: Schematic comparison between the (a) original optimal pulse $\alpha(t)$ and (b) the modulated pulse $\tilde{\alpha}(t) = \alpha(t) f_{\mathrm{dp}}(t)$ [see Eq. (2.17)] needed in the absence of a 2-photon drive. In this example, we choose $M = 80$ periods for the modulation.

that (i) the error from Trotter expansion [see Eq. (2.11) and (2.12)] is small, and (ii) within each period of $f(t)$, the control function $\alpha(t)$ varies slowly so that the time average of $\tilde{\alpha}(t)$ deviates little from $\alpha(t)$. In the infinite M limit, the two sides of Eq. (2.16) are identical. The induced infidelity due to the inaccuracy of the Trotter approximation is analyzed in detail in Appendix A.2.

2.6 Discussion

In the previous section, we have demonstrated that in principle a distinct 2-photon drive is unnecessary for universal control and explicitly provided the modified 1-photon pulse design for targeted operations. However, our operations may suffer from coherent errors that come from inaccurate Trotter approximation. This, as well as the incoherent photon loss during gate execution, serve as two sources for gate infidelity.

If the 1-photon driving amplitude $|\Lambda_1|$ (or the input power P_{in}) is unlimited, both error

sources can be sufficiently suppressed. We can increase the amplitude of $\alpha(t)$ so that the gate time T as well as the photon loss probability is reduced. We can also increase the displaced frame oscillating frequency ω_r to reduce Trotter errors. However, if P_{in} is limited, we have to make a trade-off to achieve optimized operation fidelity. In Appendix A.2, we analyze the effect of these errors in detail for the case of preparing a single photon Fock state and show that the state preparation infidelity is given by

$$\epsilon_{\text{tot}} \simeq C(\kappa_e + \kappa_i)T + \frac{C'}{P_{\text{in}}^2 \kappa_e^2 \chi^{10} T^{14}}, \tag{2.18}$$

where $\kappa_{e(i)}$ stands for the external (internal) loss rate. Here the first term captures the effect of loss errors growing with time T, and the second term captures the Trotter errors that scale inverse-polynomially with T. Therefore, there exists an optimal time and external loss rate for a given power and internal loss rate that minimizes the total error, which scales as $\epsilon_{\text{tot}}^{\text{opt}} \propto \kappa_i^{4/5}/(P_{\text{in}}^{2/15}\chi^{2/3})$. For example, suppose we have a niobium nitride resonator [120] with $\omega_c = 2\pi \times 100$ GHz, $\chi = -1$ kHz and an improved $|\chi|/\kappa_i = 0.5$ in the future. Then, to prepare $|1\rangle$ state with 90% fidelity, the required P_{in} is around 30 nW. The corresponding required $\alpha \approx 15$. These numbers suggest that our scheme may work well in the regime where self-Kerr is comparable to or slightly lower than the loss rate. However, if κ_i is increased by a factor of 10, we need an increase of P_{in} by a factor of 10^6 to keep the same $\epsilon_{\text{tot}}^{\text{opt}}$, which indicates that our protocol may not be power-friendly in the large κ_i regime. On the other hand, if we can implement the 2-photon drive directly and also power-efficiently so that we only need to consider the power cost for the 1-photon drive, then the error scaling will be modified as $\epsilon \propto \kappa_i^{5/6}/(P_{\text{in}}^{1/6}\chi^{2/3})$. It will be interesting to look into the power cost for other control protocols, like using both DC and double-frequency pump currents [132], or directly modulating a nonharmonic potential itself [133]. We leave the investigation of other power-saving schemes for universal control as further work.

We also investigate the optimality of the Trotter scheme by comparing it with a direct

optimization scheme. Specifically, we assume that the two-photon drive is absent from the beginning, and directly maximize the fidelity by optimizing $\alpha(t)$ using a more expressive ansatz that utilizes neural networks with periodic activation functions [134]. Our results indicate that this heuristic approach does not yield a better solution (in terms of fidelity and required power) compared to the Trotter scheme. We refer the reader to Appendix A.10 for more details.

Additionally, if our 1-photon drive is too strong, the RWA may not be valid since the frequency ω_c of the resonator is always finite. The counter-rotating (non-RWA) terms like $[\Lambda_1^*(t)e^{i\omega_1(t)t}\hat{a}^\dagger + \text{h.c.}]$ should be considered in Eq. (2.1). In Appendix A.5, we discuss the way to mitigate the dominant effects from non-RWA terms by adjusting the driving amplitude $\Lambda_1(t)$ and frequency $\omega_1(t)$. It is also worth mentioning that there will be no $[\Lambda_1^*(t)e^{i\omega_1(t)t}\hat{a}^\dagger + \text{h.c.}]$ term if on the hardware we can drive both charge and flux simultaneously with properly chosen amplitudes, which provides another way to mitigate non-RWA effects. However, there could still be a fundamental limitation on the fidelity as a function of two dimensionless parameters κ_i/χ and internal quality factor $Q_i := \omega_c/\kappa_i$. For the task of preparing $|1\rangle$, we found a rough lower bound for infidelity as $\epsilon \gtrsim 3\pi(\kappa_i/\chi)^{2/3}/(16Q_i^{1/3})$ even if we can directly implement 2-photon drive or do 1-photon drives on both charge and flux quadratures (see Appendix A.5.3), which indicates that κ_i/χ cannot be too small given finite Q_i. We leave the further improvement of this bound as an open question.

2.7 Conclusion

We show that by making a non-trivial extension of the displacement-induced, weak-Kerr photon blockade of Ref. [51], one can achieve any unitary operation in a blockade subspace of arbitrary dimension. The speed of the operations can be enhanced by using large displacements to overcome the adverse effects of photon loss. Further, this can be implemented using only 1-photon drives, provided that the input power is not a limitation and RWA conditions

are justified. Moreover, from a computational perspective, our work simplifies the task of optimal control in such systems. It reduces the computational overhead for finding control sequences that utilize large photon number states by working in a special instantaneously displaced frame. As a result, our work provides a novel and efficient quantum control protocol for weak nonlinear bosonic systems, which could be helpful for future quantum information processing tasks on suitable platforms.

CHAPTER 3

BIAS-PRESERVING OPERATIONS FOR PAIR-CAT CODES

Quantum information is powerful but fragile due to the presence of noise and imperfections. Quantum error correction can actively correct physical errors and protect the encoded quantum information by introducing redundancy in physical systems. Fault-tolerant design also ensures that errors during the quantum error correction will not compromise the encoded quantum information, which enables us to accomplish quantum tasks as accurately as possible if the error probability of each gate operation on physical qubits is below a certain threshold [29, 30]. For generic depolarization errors, however, fault-tolerant quantum computation often requires a demanding error threshold and resource overhead, which poses a major challenge with the current technology.

One promising approach to overcome this challenge is to design quantum error correction schemes specific for realistic errors in physical devices. For example, when physical systems have a biased-noise structure – one type of error is stronger than all other types of errors [135] – we can design efficient quantum error correction schemes to improve error threshold [136–139] and reduce resource overhead [140]. Hence, seeking a biased-noise structure and preserving the error bias during operations on the physical qubits are highly desirable to make these merits come true. In practice, however, it is non-trivial to preserve the biased-noise structure for all quantum operations. For example, a phase-flip error can be transformed into a bit-flip error and vice versa after a Hadamard gate. Moreover, phase-flip error bias cannot be preserved during the execution of a CX gate for physical qubits encoded in two-level systems [68].

Distinct from two-level physical systems, multi-level systems (such as harmonic oscillators) can encode quantum information with desirable biased-noise structure and bias-preserving quantum operations. For example, we can use harmonic oscillators with cat codes, which encode quantum bits of information using a subspace spanned by two sepa-

rated coherent states $|\, pm\alpha\rangle$ [54]. With a specific choice of computational basis of the cat code, the bit-flip error can be exponentially suppressed by the averaged photon number compared with the phase-flip error, which naturally provides the biased-noise structure [68, 69]. The cat qubits can be stabilized in both driven-dissipative systems [54] and Kerr-nonlinear resonators with 2-photon driving [57]. Both of the stabilization protocols with the biased-noise structure have been experimentally demonstrated [55, 58, 141]. A set of operations that includes CX and Toffoli gates for cat qubits with bias-preserving properties has been proposed in both platforms [68, 69]. Recently, new methods to keep noise bias in Kerr cat qubits suppressing heating-induced leakage [142] and new approaches to realize fast and bias-preserving gates in cat code [143, 144] have also been proposed. Further, cat qubits can be concatenated into repetition code level, on which a universal gate set for quantum computation can be constructed by using fundamental bias-preserving operations on physical qubits. Concatenation of cat qubits with different types of surface codes has also been investigated under practical consideration [59, 145]. In addition, multicomponent cat codes encoded in a single mode can also be used to protect against photon loss errors [54]. However, the corresponding quantum error correction strategy for all the protocols we mentioned above to suppress the effect from photon losses relies on measuring parity $(-1)^{\hat{a}^\dagger \hat{a}}$, which is hard to implement continuously. As a result, extra overhead might be required for those measurements in the middle of the circuits and the following feedback control, which can lead to extra errors and delays.

In our work, we focus on another type of bosonic codes named pair-cat code, which is an important generalization of cat code into multimode bosonic systems [70]. For pair-cat code, any photon loss error happening in either mode can be detected by monitoring the photon number difference between the two modes and we can correct them correspondingly, which enables us to perform continuous error correction against photon loss errors. Different from the parity, the photon number difference is much easier to monitor continuously while

keeping the stabilization on. Moreover, we need fewer averaged photons to get at least the same protection as in the cat code. With all the merits of the pair-cat code, it is natural to ask whether pair-cat code has a similar biased-noise structure and whether we can generalize the methods used to construct bias-preserving operations for cat code [68, 69] into the pair-cat case while keeping the merit of continuous error correction during operations.

In this work, we successfully construct a set of bias-preserving operations for pair-cat codes (including CX and Toffoli, sufficient for universal computation in the repetition code level), which can be compatible with continuous quantum error correction of both bosonic loss and dephasing errors. This chapter is organized as follows. In Sec. 3.1, we will go over the basic encoding scheme of the pair-cat code. In Sec. 3.2, we investigate the construction of the bias-preserving operation set in both driven-dissipative systems and Hamiltonian systems. We summarize our results in Sec. 3.3. In the Appendix B, we summarize some useful properties of pair-cat code itself, including its stabilization, error correction strategy, and optimal error probabilities during the bias-preserving operations we design in the main text.

3.1 Pair-cat code stabilization

The pair-cat code itself with stabilization in the driven-dissipative systems has been proposed in [70]. Here we first summarize basic properties of the code, and then introduce the Hamiltonian protection scheme as a direct generalization of the cat code.

We first mention the encoding of the cat code for further comparison. By focusing on the subspace spanned by two coherent states $\{|\alpha\rangle, |-\alpha\rangle\}$, we introduce the states $|\mathcal{C}_\alpha^\pm\rangle$ with fixed even or odd photon number parity, where

$$|\mathcal{C}_\alpha^\pm\rangle := \frac{|\alpha\rangle \pm |-\alpha\rangle}{\sqrt{\mathcal{N}_\pm}}. \tag{3.1}$$

Here $\mathcal{N}_\pm = 2(1 \pm e^{-2|\alpha|^2})$ is the normalization factor. By encoding $|\mathcal{C}_\alpha^\pm\rangle$ as the eigenstates of X operator of the cat qubit with eigenvalue ± 1, we can see that in the large $|\alpha|$ limit, $|\pm\alpha\rangle$ states serve as the logical $|0\rangle$ and $|1\rangle$ of the code. Since physical relevant errors, like photon loss, gain and dephasing noise only act locally in the phase space, which makes it hard to flip $|\alpha\rangle$ to $|-\alpha\rangle$ and vice versa, the bit-flip error is naturally suppressed with the choice of our encoding. In fact, it is exponentially suppressed as the increase of average photon number in the resonator compared with the phase-flip error.

Then, we consider a system with two bosonic modes and denote them as mode $\hat{a}$ and $\hat{b}$. We introduce the pair-coherent state $|\gamma_\Delta\rangle$ [146], which serves as the basic component in pair-cat code. It is defined as the projection of the identical coherent state in two modes $|\gamma,\gamma\rangle := |\gamma\rangle \otimes |\gamma\rangle$ into a subspace with fixed photon number difference between these modes. Specifically, we have

$$|\gamma_\Delta\rangle = \frac{\hat{P}_\Delta |\gamma,\gamma\rangle}{\sqrt{\mathcal{N}_\Delta}}, \tag{3.2}$$

where $\mathcal{N}_\Delta = e^{-2|\gamma|^2} I_\Delta(2|\gamma|^2)$ is the normalization factor and $I_\Delta(x)$ is the modified Bessel function of the first kind. $\hat{P}_\Delta$ is the projection operator which projects states into a subspace with fixed photon number difference $\hat{\Delta} := \hat{n}_b - \hat{n}_a = \Delta$, which means,

$$\hat{P}_\Delta := \sum_{n=0}^{+\infty} |n, n+\Delta\rangle\langle n, n+\Delta| \qquad (\Delta \geq 0). \tag{3.3}$$

The $\Delta < 0$ case can always be defined similarly by performing a SWAP operation between the two modes. From then on, for simplicity we assume $\Delta \geq 0$ by default in the following discussions if there is no further comment.

Two merits need to be highlighted for the $|\gamma_\Delta\rangle$ state: first, by analogy with the cat code design where $(\hat{a}^2 - \alpha^2)|\pm\alpha\rangle = 0$, here we have

$$(\hat{a}^2\hat{b}^2 - \gamma^4)|\gamma_\Delta\rangle = (\hat{a}^2\hat{b}^2 - \gamma^4)|(i\gamma)_\Delta\rangle = 0. \tag{3.4}$$

Therefore, $\hat{a}^2\hat{b}^2 - \gamma^4$ can dissipatively stabilize the pair-cat code space. We note that the pair-cat has a unique advantage over the cat code, which is that for any number of photon loss in either mode, it can only change the pair-coherent state into another subspace with different Δ. Notice that

$$\hat{a}\hat{P}_\Delta = \hat{P}_{\Delta+1}\hat{a}, \quad \hat{b}\hat{P}_\Delta = \hat{P}_{\Delta-1}\hat{b}, \tag{3.5}$$

we have

$$\hat{a}^k \left|\gamma_\Delta\right\rangle = \gamma^k \sqrt{\frac{\mathcal{N}_{\Delta+k}}{\mathcal{N}_\Delta}} \left|\gamma_{(\Delta+k)}\right\rangle, \quad \hat{b}^l \left|\gamma_\Delta\right\rangle = \gamma^l \sqrt{\frac{\mathcal{N}_{\Delta-l}}{\mathcal{N}_\Delta}} \left|\gamma_{(\Delta-l)}\right\rangle. \tag{3.6}$$

As a result, this type of error syndrome can be easily monitored by measuring $\hat{\Delta}$, and then we could design strategies to correct it correspondingly. However, this method does not work if the system suffers from loss error $\hat{a}\hat{b}$ since Δ does not change after $\hat{a}\hat{b}$ acting on the state. Later, we can see that this will give us an uncorrectable error with our encoding method.

To encode the qubit with the pair-coherent states, we use a generalized "parity" projection operator $\hat{Q}_\pm^{(\Delta)}$ within each Δ-fixed subspace. Before giving the definition of $\hat{Q}_\pm^{(\Delta)}$, we first introduce the projection operator $\hat{P}_\pm^a$ to $\hat{a}$ mode with fixed parity as

$$\hat{P}_\pm^a := \frac{\hat{I} \pm (-1)^{\hat{n}_a}}{2}.$$

Then $\hat{Q}_\pm^{(\Delta)}$ can be defined as

$$\hat{Q}_\pm^{(\Delta)} := \hat{P}_\pm^a \hat{P}_\Delta \qquad (\Delta \geq 0). \tag{3.7}$$

As we pointed out that $\Delta < 0$ case can always be defined by performing a SWAP between two modes, we should use the parity in $\hat{b}$ mode to define $\hat{Q}_\pm^{(\Delta<0)} := \hat{P}_\pm^b \hat{P}_\Delta$.

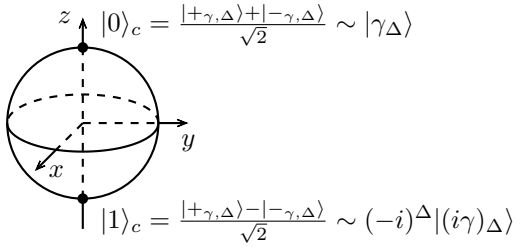

Figure 3.1: Bloch sphere representation of the pair-cat encoding.

Finally, we define our code states as

$$|\pm_{\gamma,\Delta}\rangle := \frac{\hat{Q}_\pm^{(\Delta)} |\gamma,\gamma\rangle}{\sqrt{\mathcal{N}_{\pm,\Delta}}} = \frac{|\gamma_\Delta\rangle \pm (-i)^\Delta |(i\gamma)_\Delta\rangle}{2\sqrt{\mathcal{N}_{\pm,\Delta}/\mathcal{N}_\Delta}}, \tag{3.8}$$

where $\mathcal{N}_{\pm,\Delta} = e^{-2|\gamma|^2}[I_\Delta(2|\gamma|^2) \pm J_\Delta(2|\gamma|^2)]/2$ is also a normalization factor, and $J_\Delta(x)$ is the Bessel function of the first kind. We adopt the convention that the above states are eigenstates of the logical X operator, specifically,

$$|+\rangle_c := |+_{\gamma,\Delta}\rangle\,, \quad |-\rangle_c := |-_{\gamma,\Delta}\rangle\,. \tag{3.9}$$

Note that here we use a different choice of basis compared with Ref. [70], so that the phase-flip error is the dominant type of error in our choice of basis in order to be consistent with the existing literature.

In the large $|\gamma|$ limit, like the cat code, we also have $\left|\langle(i\gamma)_\Delta|\gamma_\Delta\rangle\right|^2 \sim O(e^{-4|\gamma|^2})$, which means these two states are asymptotically orthogonal. As a result,

$$|\pm_{\gamma,\Delta}\rangle \sim \frac{|\gamma_\Delta\rangle \pm (-i)^\Delta |(i\gamma)_\Delta\rangle}{\sqrt{2}} \qquad (|\gamma| \to +\infty). \tag{3.10}$$

Further, the states along Z basis are

$$|0\rangle_c \sim |\gamma_\Delta\rangle\,, \quad |1\rangle_c \sim (-i)^\Delta |(i\gamma)_\Delta\rangle \qquad (|\gamma| \to +\infty). \tag{3.11}$$

On the other hand, in the $|\gamma| \to 0$ limit, we have

$$\begin{cases} |+_{\gamma,\Delta}\rangle \sim (\gamma/|\gamma|)^{\Delta} |n_a = 0, n_b = \Delta\rangle \\ |-_{\gamma,\Delta}\rangle \sim (\gamma/|\gamma|)^{\Delta+2} |n_a = 1, n_b = \Delta + 1\rangle \end{cases} \quad (|\gamma| \to 0), \tag{3.12}$$

which, as indicated later, provides us with one way to prepare code states in a bias-preserving way by adiabatically turning on control parameters.

As mentioned before, the lowest-order uncorrectable loss error is $\hat{a}\hat{b}$. Notice that this error does not cause the code states to go out of the code subspace. We denote $r_\Delta = \sqrt{\mathcal{N}_{-,\Delta}/\mathcal{N}_{+,\Delta}}$. In the large $|\gamma|$ limit, $r_\Delta \simeq 1 - 2e^{-2|\gamma|^2}\cos(\phi)$ where $\phi = 2|\gamma|^2 - \frac{2\Delta+1}{4}\pi$. The projection operator on the code space can be denoted as $\hat{P}_c := |+\rangle_c\langle +| + |-\rangle_c\langle -|$. In the large $|\gamma|$ limit, we have

$$\begin{aligned} \hat{P}_c\hat{a}\hat{b}\hat{P}_c &= \gamma^2 \frac{r_\Delta + r_\Delta^{-1}}{2}\hat{Z}_c + i\gamma^2 \frac{r_\Delta - r_\Delta^{-1}}{2}\hat{Y}_c \\ &\sim \gamma^2\hat{Z}_c - 2i\gamma^2 e^{-2|\gamma|^2}\cos(\phi)\hat{Y}_c. \end{aligned} \tag{3.13}$$

We can see that the Z error is the dominant one while the other errors are exponentially suppressed for large $|\gamma|$.

On the other hand, the error induced by the bosonic dephasing term can always be exponentially suppressed in the large $|\gamma|$ limit. For example,

$$\hat{P}_c\hat{a}^\dagger\hat{a}\hat{P}_c \sim \gamma^2\hat{I}_c + O(\gamma^2 e^{-2|\gamma|^2})\hat{X}_c. \tag{3.14}$$

Therefore, in our work, we will only focus on the effects induced by photon losses and leave the bosonic dephasing error aside.

The pair-cat code can be stabilized in a driven-dissipative system with the jump operator

$\hat{F} = \hat{a}^2\hat{b}^2 - \gamma^4$. The corresponding dynamical equation of motion is

$$\frac{\mathrm{d}\hat{\rho}}{\mathrm{d}t} = \kappa\mathcal{D}[\hat{F}]\hat{\rho} = \kappa(\hat{F}\hat{\rho}\hat{F}^\dagger - \frac{1}{2}\{\hat{F}^\dagger\hat{F}, \hat{\rho}\}), \tag{3.15}$$

where $\{\bullet, \bullet\}$ denotes the anti-commutator, and κ corresponds to the 4-photon dissipation rate. Note that both the photon number difference and the parity are preserved during this evolution because they commute with $\hat{F}$. Since $\hat{F}\left|\gamma_\Delta\right\rangle = \hat{F}\left|(i\gamma)_\Delta\right\rangle = 0$, our code space lies in the decoherence-free subspace of the system. The effective dissipative gap introduced in Ref. [147] inversely relates to the leakage rate out of the steady state subspace under perturbations. In our case, it has exactly the same properties as the energy gap in the Hamiltonian protection scheme that will be introduced later, and as shown in Appendix B.2, it scales as $\Delta_{\mathrm{edg}} \sim O(|\gamma|^6)$.

If we consider the single-photon loss effect of $\hat{a}$ and $\hat{b}$ mode during evolution, the right-hand-side of Eq. (3.15) should be modified with extra terms $\kappa_{1a}\mathcal{D}[\hat{a}]\hat{\rho}$ and $\kappa_{1b}\mathcal{D}[\hat{b}]\hat{\rho}$, where κ_{1a} and κ_{1b} are the single-photon loss rate of the corresponding two modes. For simplicity later we assume κ_{1a} is equal to κ_{1b}, and denote $\kappa_1 = \kappa_{1a} = \kappa_{1b}$. Here we will show how the pair-cat encoding can provide protection against effects from single-photon loss error during gate execution. The property of autonomous error correction of pair-cat code against photon losses is also discussed in [70].

In this work, we suggest that the pair-cat code can also be stabilized by the following Hamiltonian:

$$\hat{H} = -K(\hat{a}^{\dagger 2}\hat{b}^{\dagger 2} - \gamma^{*4})(\hat{a}^2\hat{b}^2 - \gamma^4). \tag{3.16}$$

It is easy to see that both $\left|+_{\gamma,\Delta}\right\rangle$ and $\left|-_{\gamma,\Delta}\right\rangle$ are the most-excited states (suppose $K > 0$) of this Hamiltonian. Since $[\hat{\Delta}, \hat{H}] = [(-1)^{\hat{n}_a}, \hat{H}] = [\hat{\Delta}, (-1)^{\hat{n}_a}] = 0$, this Hamiltonian can be divided into different parts that act on different subspaces with fixed photon-number

difference between two modes and fixed parity:

$$\hat{H} = \sum_{\mu,\Delta} \hat{H}_{\mu,\Delta} = \sum_{\mu,\Delta} \hat{Q}_{\mu}^{(\Delta)} \hat{H} \hat{Q}_{\mu}^{(\Delta)}. \tag{3.17}$$

The energy gap between the code subspace and first-less-excited states is $\sim 8K|\gamma|^6$ in the large $|\gamma|$ limit when Δ is a finite number. A more detailed analysis of this Hamiltonian is performed in Appendix B.2.

We also numerically investigated the possibility of finding lower-order Hamiltonian that has both γ-dependent protection of the code subspace and preserves the photon number difference. Unfortunately, there is no lower-order Hamiltonian that fulfills those criteria. Details can be found in Appendix B.1.

3.2 Construction of bias-preserving gates

The set of bias-preserving operations on cat qubit that does not convert the major errors into the minor errors has been proposed in Ref. [68, 69]. For single-qubit operations, it contains code state preparation, measurement in X basis, single qubit X operation, and rotation along Z axis for arbitrary angle $Z(\theta) := \exp(-i\theta\hat{Z}/2)$. For multi-qubit gates, CX and Toffoli gates can also be performed in a bias-preserving manner, which is not possible for physical qubits in two-level systems. Besides, bias-preserving $ZZ(\theta) := \exp(-i\theta\hat{Z}_1\hat{Z}_2/2)$ gate is also achievable. We denote $\mathcal{S}$ as the set of fundamental bias-preserving operations of the cat code: $\mathcal{S} = \{\mathcal{P}_{|\pm\rangle_c}, \mathcal{M}_X, X, Z(\theta), ZZ(\theta), \text{CX}, \text{Toffoli}\}$. Further, it can also be shown that a universal gate set for fault-tolerant quantum computation can be constructed in the repetition code level by using those bias-preserving operations acting on physical cat qubits.

In this work, we will show that these operations in $\mathcal{S}$ can also be constructed with the pair-cat code in both driven-dissipative protection and Hamiltonian protection schemes, and reveal the similarities between cat code and pair-cat code structures. The construction

Table 3.1: A summary of fundamental bias-preserving operations in cat code and pair-cat code.

	Cat code		Pair-cat code	
	Driven-dissipative scheme [68]	Hamiltonian scheme [69]	Driven-dissipative scheme	Hamiltonian scheme
Stabilization	$\hat{F}_{\text{cat}} = \hat{a}^2 - \alpha^2$	$\hat{H} = -K\hat{F}_{\text{cat}}^\dagger \hat{F}_{\text{cat}}$	$\hat{F}_{\text{pc}} = \hat{a}^2\hat{b}^2 - \gamma^4$	$\hat{H} = -K\hat{F}_{\text{pc}}^\dagger \hat{F}_{\text{pc}}$
Uncorrectable loss error	$\hat{a}$		$\hat{a}\hat{b}$	
$\mathcal{P}_{\lvert+\rangle_c}$	Start at $\lvert 0\rangle$ Evolve to steady state	Start at $\lvert 0\rangle$ $\alpha(t): 0 \to \alpha$	Start at $\lvert 0, \Delta\rangle$ Evolve to steady state	Start at $\lvert 0, \Delta\rangle$ $\gamma(t): 0 \to \gamma$
$\mathcal{M}_X$	Need an ancilla: $\hat{H}_{\text{disp}} = -\chi\lvert e\rangle\langle e\rvert\hat{a}^\dagger\hat{a}$; turn off the protection		Need two ancilla: $\hat{H}_{\text{disp}} = -\chi(\lvert e\rangle_1\langle e\rvert\,\hat{a}^\dagger\hat{a} + \lvert e\rangle_2\langle e\rvert\,\hat{b}^\dagger\hat{b})$; may also need $\hat{\Delta}$ measurement; turn off the protection	
$Z(\theta)$	$\hat{H}_Z = \epsilon_Z(\hat{a}e^{-i\varphi} + \hat{a}^\dagger e^{i\varphi})$		$\hat{H}_Z = \epsilon_Z(\hat{a}\hat{b}e^{-i\varphi} + \hat{a}^\dagger\hat{b}^\dagger e^{i\varphi})$	
$ZZ(\theta)$	$\hat{H}_{ZZ} = \epsilon_{ZZ}(\hat{a}_1\hat{a}_2^\dagger + \hat{a}_1^\dagger\hat{a}_2)$		$\hat{H}_{ZZ} = \epsilon_{ZZ}(\hat{a}_1\hat{b}_1\hat{a}_2^\dagger\hat{b}_2^\dagger + \hat{a}_1^\dagger\hat{b}_1^\dagger\hat{a}_2\hat{b}_2)$	
X	$\alpha(t) = \alpha e^{i\pi\frac{t}{T}}$ together with $\hat{H}_{X,\text{rot}} = -\frac{\pi}{T}\hat{n}$		$\gamma(t) = \gamma e^{i\frac{\pi}{2}\frac{t}{T}}$ together with $\hat{H}_{X,\text{rot}} = -\frac{\pi}{2T}(\hat{n}_a + \hat{n}_b)$	
CX	$\hat{F}_{1,2}$ in Eq. (3.24), with $\hat{H}_{\text{CX,rot}}$ in Eq. (3.25)	Same as Eq. (3.34) with operators for cat code	$\hat{F}_{1,2}$ in Eq. (3.26), with $\hat{H}_{\text{CX,rot}}$ in Eq. (3.27); real-time $\hat{\Delta}_2$ monitoring	$\hat{H}_{\text{CX}}$ in Eq. (3.34); real-time $\hat{\Delta}_2$ monitoring
Toffoli	$\hat{F}_{1,2,3}$ in Eq. (3.29), with $\hat{H}_{\text{Tof,rot}}$ in Eq. (3.30)	Same as Eq. (3.35) with operators for cat code	$\hat{F}_{1,2,3}$ in Eq. (3.31), with $\hat{H}_{\text{Tof,rot}}$ in Eq. (3.32); real-time $\hat{\Delta}_3$ monitoring	$\hat{H}_{\text{Tof}}$ in Eq. (3.35); real-time $\hat{\Delta}_3$ monitoring

of logical gate set on the concatenated code level based on fundamental bias-preserving operations on physical qubits is independent of the specific type of physical qubits we use, which means the results developed using cat code can be adapted to the pair-cat situation.

The biased error in pair-cat code comes from the large distance between $|\gamma_\Delta\rangle$ and $|(i\gamma)_\Delta\rangle$ in the generalized phase space (or "γ-plane", see [70]) and the locality of the physical errors. Therefore, to preserve error bias during the gate operation, $|\gamma|$ should always be kept large.

In terms of the notation, we use subscripts "1", "2", and "3" on each operator to specify which pair-cat qubit it acts on. For the CX gate, we use "1" to denote the control qubit and "2" for the target qubit; while for the Toffoli gate we use "1" and "2" for the two control qubits and "3" for the target qubit. For example, $\hat{\Delta}_2 = \hat{b}_2^\dagger\hat{b}_2 - \hat{a}_2^\dagger\hat{a}_2$ for the CX gate and $\hat{\Delta}_3 = \hat{b}_3^\dagger\hat{b}_3 - \hat{a}_3^\dagger\hat{a}_3$ for the Toffoli gate are both photon number difference operators acting on the target qubits respectively.

3.2.1 Dissipation engineering scheme

In this part, we will show the way to construct bias-preserving operations in $\mathcal{S}$ with driven-dissipative stabilization. We will see how the continuous syndrome ($\hat{\Delta}$) monitoring can help to reduce errors caused by photon loss. We also derive the scaling properties of the error probability during gate operations where optimal gate time is chosen in Appendix B.4 and summarize the results in Table 3.2.

Preparation of $|\pm\rangle_c$ states

In cat code protection with jump operator $\hat{F}_{\text{cat}} = \hat{a}^2 - \alpha^2$, the preparation of $|\mathcal{C}_\alpha^+\rangle$ state can be done by initializing the system at the vacuum state $|0\rangle$, and then just let the system evolve under the Lindblad equation to reach the steady state, which will be the exact code state we want [54, 68]. It is because that, the steady states ρ_∞ of this evolution is a linear superposition of $\{|\mathcal{C}_\alpha^+\rangle\langle\mathcal{C}_\alpha^+|, |\mathcal{C}_\alpha^+\rangle\langle\mathcal{C}_\alpha^-|, |\mathcal{C}_\alpha^-\rangle\langle\mathcal{C}_\alpha^+|, |\mathcal{C}_\alpha^-\rangle\langle\mathcal{C}_\alpha^-|\}$. And, since the parity is preserved during the evolution, the only result will be $\rho_\infty = |\mathcal{C}_\alpha^+\rangle\langle\mathcal{C}_\alpha^+|$ if the initial state is $|0\rangle\langle 0|$ [148]. To prepare $|\mathcal{C}_\alpha^-\rangle$ state, we can either start with Fock $|0\rangle\langle 0|$ state and let the system evolve, or perform $\hat{Z}_c$ operation after the preparation of $|\mathcal{C}_\alpha^+\rangle$ state.

Similarly, in pair-cat code case with jump operator $\hat{F}_{\text{pc}} = \hat{a}^2\hat{b}^2 - \gamma^4$, the space of steady states is spanned by $\{|\mu_{\gamma,\Delta}\rangle\langle\mu'_{\gamma,\Delta'}||\mu,\mu' \in \{+,-\}; \Delta, \Delta' \in \mathbb{Z}\}$. Besides, both the parity of the two modes and the photon number difference are conserved. As a result, if we start with $|0,\Delta\rangle\langle 0,\Delta|$ state and let the system evolve, eventually it will end up at $|+_{\gamma,\Delta}\rangle\langle+_{\gamma,\Delta}|$ state. To prepare $|-_{\gamma,\Delta}\rangle$ state, similarly we can either start with $|1,\Delta+1\rangle$ state and wait for it to reach the steady state, or perform $\hat{Z}_c$ operation, which we will introduce later, on $|+_{\gamma,\Delta}\rangle$ state.

Photon loss errors may happen during the state preparation and the idling time after that. The probability of a single-photon loss in either mode is $p = \kappa_1\bar{n}T$ provided that $p \ll 1$, where κ_1 is the 1-photon loss rate, T is the total time of the process we consider,

and $\bar{n}$ is the average photon number during the whole process. However, as indicated before, only a single-photon loss does not cause a logical error directly. It can be captured by a $\hat{\Delta}$ measurement after the process to determine whether and in which mode the photon loss happens. Then we can apply a recovery operation to correct the error. However, loss errors happen in both modes cannot be identified in this way, which can occur with the probability $p_Z = p^2 \sim O[(\kappa_1 \bar{n} T)^2]$. This corresponds to the Z error probability of the pair-cat code during state preparation and idling process. In the idling part, we have $\bar{n} \sim O(|\gamma|^2)$, and if the time of this part dominates, we can write $p_Z = p^2 \sim O[(\kappa_1 |\gamma|^2 T)^2]$.

Measurement in X basis

In order to distinguish $|+_{\gamma,\Delta}\rangle$ state from $|-_{\gamma,\Delta}\rangle$ state, we can try to check the parity of either mode of the pair-cat code. This can be done in the same way as the cat code case [68]. We could couple $\hat{a}$ mode with an ancilla qubit via dispersive coupling Hamiltonian $\hat{H}_{\text{disp}} = -\chi |e\rangle\langle e| \hat{a}^\dagger \hat{a}$. The ancilla qubit is initialized at $|+\rangle_q$ state where $|\pm\rangle_q = (|g\rangle \pm |e\rangle)/\sqrt{2}$. After time $T = \frac{\pi}{\chi}$, the unitary evolution operator will be $\hat{U} = |g\rangle\langle g| \otimes \hat{I} + |e\rangle\langle e| \otimes e^{i\pi \hat{n}_a}$, and the quantum state will evolve from $|\psi(0)\rangle = |+\rangle_q \otimes (u_0 |+_{\gamma,\Delta}\rangle + u_1 |-_{\gamma,\Delta}\rangle)$ to $|\psi(T)\rangle = u_0 |+\rangle_q \otimes |+_{\gamma,\Delta}\rangle + u_1 |-\rangle_q \otimes |-_{\gamma,\Delta}\rangle$. Finally, we measure the ancilla qubit along X basis. If we get $|+\rangle_q$ state, it is equivalent to saying that we get the $|+_{\gamma,\Delta}\rangle$ by performing the X measurement on pair-cat code.

It is worth mentioning that, during the qubit-dependent rotation of the cavity modes, we have

$$e^{i\theta \hat{n}_a} |\gamma_\Delta\rangle = e^{-i\Delta\theta/2} |(\gamma e^{i\theta/2})_\Delta\rangle \,, \tag{3.18}$$

which means the driven-dissipative stabilization should be turned off during this evolution. However, as indicated in the cat code case [68], turning off the stabilization is not a problem because the only information we need from the measurement is the parity of the state instead of the amplitude γ. Moreover, since the dissipator $\hat{F}$ commutes with the parity, it does not

provide any protection against parity change. As a result, it does not matter whether the dissipative stabilization is on or off during the measurement process.

Note that a single photon loss might change the outcome of the parity measurement. To suppress the loss-induced measurement error to higher order, we can introduce two ancillae and use them to measure both the parity of $\hat{a}$ and $\hat{b}$ mode together. If the outcomes agree with the Δ we fixed for the code space, we can trust the outcomes. Otherwise, we need to perform a measurement of the photon number difference between the two modes immediately after the parity measurement, to check which mode suffers from the photon loss and use the parity of another mode to indicate the generalized parity of the pair-cat code state. However, if both modes suffer from single-photon loss during parity measurement, there is some chance that the parity outcomes are consistent but wrong, or they are inconsistent but cannot be resolved since $\hat{\Delta}$ measurement suggests no loss happened. As a result, the error probability during the measurement process can be suppressed from $O(\kappa_1|\gamma|^2/\chi)$ to $O[(\kappa_1|\gamma|^2/\chi)^2]$ by using the protocol we mentioned here.

$Z(\theta)$ and $ZZ(\theta)$ gates

The $Z(\theta)$ and $ZZ(\theta)$ gates in cat code can be performed by using the following Hamiltonian [54]:

$$\begin{aligned} \hat{H}_Z &= \epsilon_Z(\hat{a}e^{-i\varphi} + \hat{a}^\dagger e^{i\varphi}), \\ \hat{H}_{ZZ} &= \epsilon_{ZZ}(\hat{a}_1\hat{a}_2^\dagger + \hat{a}_1^\dagger\hat{a}_2). \end{aligned} \tag{3.19}$$

By projecting those Hamiltonians into the cat code subspace, we can get the Z and ZZ operators which will generate the $Z(\theta)$ and $ZZ(\theta)$ gates. Here θ can be controlled by the gate time. The validity of this projection can be understood via the quantum Zeno effect, that the dissipation term keeps monitoring the system to prevent the state from leaking out of the code subspace. In Appendix B.4 we have a detailed analysis on the gate error

induced by this leakage and show that the description under quantum Zeno effect will be more accurate with small ϵ_Z or ϵ_{ZZ} and large κ. Since both of the projected Hamiltonians commute with Z error on either cat qubit, these two operations are naturally bias-preserving.

In the pair-cat situation, we can use these Hamiltonian to achieve the two gates [70]:

$$\begin{aligned}\hat{H}_Z &= \epsilon_Z(\hat{a}\hat{b}e^{-i\varphi} + \hat{a}^\dagger\hat{b}^\dagger e^{i\varphi}),\\ \hat{H}_{ZZ} &= \epsilon_{ZZ}(\hat{a}_1\hat{b}_1\hat{a}_2^\dagger\hat{b}_2^\dagger + \hat{a}_1^\dagger\hat{b}_1^\dagger\hat{a}_2\hat{b}_2).\end{aligned} \tag{3.20}$$

Again, by projecting into the code space with $\hat{P}_c = |+\rangle_c\langle+| + |-\rangle_c\langle-|$ while working in the large $|\gamma|$ limit, we have

$$\hat{P}_c\hat{H}_Z\hat{P}_c \sim 2\epsilon_Z\left(\Re[\gamma^2 e^{-i\varphi}]\hat{Z}_c + 2\Im[\gamma^2 e^{-i\varphi}]e^{-2|\gamma|^2}\cos(\phi)\hat{Y}_c\right), \tag{3.21a}$$

$$(\hat{P}_{1c}\hat{P}_{2c})\hat{H}_{ZZ}(\hat{P}_{1c}\hat{P}_{2c}) \sim 2\epsilon_{ZZ}|\gamma|^4(\hat{Z}_{1c}\hat{Z}_{2c} + 4e^{-4|\gamma|^2}\cos^2(\phi)\hat{Y}_{1c}\hat{Y}_{2c}). \tag{3.21b}$$

Here $\phi = 2|\gamma|^2 - \frac{2\Delta+1}{4}\pi$ has been introduced before. We can also see that the $\hat{Y}_c$ and $\hat{Y}_{1c}\hat{Y}_{2c}$ terms are exponentially suppressed so that we can use these Hamiltonians to get $Z(\theta)$ and $ZZ(\theta)$ gates. Besides, in Eq. (3.21a) we can always choose φ so that $\gamma^2 e^{-i\varphi} = |\gamma|^2$. The corresponding gate time to reach θ angle rotation is $t_Z = \frac{\theta}{4|\epsilon_Z||\gamma|^2}$ and $t_{ZZ} = \frac{\theta}{4|\epsilon_{ZZ}||\gamma|^4}$.

X gate

To realize X gate in cat code in a bias-preserving way, one method is to adiabatically change $\alpha(t)$ from α to $-\alpha$ and vice versa, while keeping $|\alpha(t)|$ larger all the time to protect the error bias [68]. After that, $|\mathcal{C}_\alpha^+\rangle$ state will remain as $|\mathcal{C}_\alpha^+\rangle$, while $|\mathcal{C}_\alpha^-\rangle$ changes to $-|\mathcal{C}_\alpha^-\rangle$, which is exactly the outcome of X gate acting on code states.

In pair-cat code case, we can also let $\gamma(t)$ change adiabatically along $\gamma(t) = \gamma e^{i\frac{\pi}{2}\frac{t}{T}}$ from $t = 0$ to T. In this way, $|\gamma_\Delta\rangle$ goes to $|(i\gamma)_\Delta\rangle$ while $|(i\gamma)_\Delta\rangle$ goes to $|(-\gamma)_\Delta\rangle = (-1)^\Delta|\gamma_\Delta\rangle$.

As a result,

$$|+_{\gamma,\Delta}\rangle \to i^{\Delta} |+_{\gamma,\Delta}\rangle , \quad |-_{\gamma,\Delta}\rangle \to i^{\Delta}(- |-_{\gamma,\Delta}\rangle). \tag{3.22}$$

So, equivalently, this is a $e^{i\Delta\pi/2}\hat{X}_c$ operation, while the global phase does not matter.

In order to implement this design in physical systems, we need to engineer the jump operator as $\hat{F} = \hat{a}^2\hat{b}^2 - \gamma^4(t)$. We can also add a Hamiltonian $\hat{H}_{X,\text{rot}} = -\frac{\pi}{2T}(\hat{n}_a + \hat{n}_b)$. It can be checked that

$$\exp(-i\hat{H}_{X,\text{rot}}t)(u_0 |\gamma_\Delta\rangle + u_1 |(i\gamma)_\Delta\rangle) = u_0 |(\gamma e^{i\frac{\pi}{2}\frac{t}{T}})_\Delta\rangle + u_1 |(i\gamma e^{i\frac{\pi}{2}\frac{t}{T}})_\Delta\rangle , \tag{3.23}$$

which means that this state is always annihilated by $\hat{F} = \hat{a}^2\hat{b}^2 - \gamma^4(t)$, so that it is protected by the driven-dissipative stabilization for any time during gate execution. So, with the help of this Hamiltonian, we could relax the requirement of adiabaticity such that $T \to +\infty$ is not needed.

CX gate

The idea of implementing a CX gate is similar to the X gate: since $\text{CX} = |0\rangle\langle 0| \otimes \hat{I} + |0\rangle\langle 0| \otimes \hat{X}$, we adiabatically rotate the target mode conditioned on the control mode being in the $|1\rangle$ state. Therefore, in the cat code scheme, the jump operators of these two cat qubits are proposed [68] as

$$\hat{F}_1 = \hat{a}_1^2 - \alpha^2, \quad \hat{F}_2 = \hat{a}_2^2 - \frac{\alpha(\hat{a}_1 + \alpha)}{2} + \frac{\alpha e^{2i\pi t/T}(\hat{a}_1 - \alpha)}{2}, \tag{3.24}$$

where, in the large $|\alpha|$ limit, by fixing control qubit in its code space, we have $\hat{F}_2 \sim |\alpha\rangle\langle\alpha| \otimes (\hat{a}_2^2 - \alpha^2) + |-\alpha\rangle\langle-\alpha| \otimes [\hat{a}_2^2 - \alpha^2(t)]$ where $\alpha(t) = \alpha e^{i\pi t/T}$. So, when $|\alpha|$ is large, if the control qubit is in $|\alpha\rangle$ state which is encoded as $|0\rangle_c$ asymptotically, the state of the target qubit does not change; on the other hand, if the control qubit is in $|-\alpha\rangle$ state, effectively there

will be an X operation acting on the target qubit.

Again, like the X gate construction, we can add a Hamiltonian to generate the conditioned rotation of the target qubit to partially compensate for the error from non-adiabaticity:

$$\begin{aligned}\hat{H}_{\text{CX,rot}} &= \frac{\pi}{2T}\frac{\hat{a}_1-\alpha}{2\alpha}\otimes(\hat{a}_2^\dagger\hat{a}_2-|\alpha|^2)+\text{h.c.}\\ &\sim -\frac{\pi}{T}|-\alpha\rangle\langle-\alpha|\otimes(\hat{a}_2^\dagger\hat{a}_2-|\alpha|^2).\end{aligned} \tag{3.25}$$

To achieve the actual CX operation, we need an extra $Z_1(-\pi|\alpha|^2)$ gate acting on the control qubit [59]. In fact, we can choose $|\alpha|^2$ as an even integer, so that this extra action is not needed.

In the pair-cat code case, we use the following jump operators to stabilize the code states:

$$\hat{F}_1=\hat{a}_1^2\hat{b}_1^2-\gamma^4,\quad \hat{F}_2=\hat{a}_2^2\hat{b}_2^2-\frac{\gamma^2(\hat{a}_1\hat{b}_1+\gamma^2)}{2}+\frac{\gamma^2e^{2i\pi t/T}(\hat{a}_1\hat{b}_1-\gamma^2)}{2}. \tag{3.26}$$

And the Hamiltonian we need for partially compensating the non-adiabatic error is

$$\hat{H}_{\text{CX,rot}}=\frac{\pi}{4T}\frac{\hat{a}_1\hat{b}_1-\gamma^2}{2\gamma^2}\otimes(\hat{a}_2^\dagger\hat{a}_2+\hat{b}_2^\dagger\hat{b}_2-2|\gamma|^2)+\text{h.c.} \tag{3.27}$$

However, the extra phase induced during γ rotation should be taken into consideration, since our effective gate operator now is $\hat{U}\propto|0\rangle_{1c}\langle0|\otimes\hat{I}_{2c}+|1\rangle_{1c}\langle1|\otimes e^{-i\pi(|\gamma|^2-\Delta/2)}\hat{X}_{2c}$. We can always use $Z(\theta)$ gate on the control qubit to correct the induced phase, or choose γ such that $(|\gamma|^2-\Delta/2)$ is an even integer.

Different from cat code where a single-photon loss can cause a phase-flip error on the cat qubit, we seek protocols that for the pair-cat code a single-photon loss during the CX gate execution does not cause logical errors. If we do nothing more than what is discussed above, we will not know when a single-photon loss event might happen, which will lead to a Z type of error on the control qubit if a photon loss occurs on the target qubit. For example, we assume an $\hat{a}_2$ error happens at time t_0 on the target qubit and see what the code states will

become finally. We still consider the large $|\gamma|$ regime where Eq. (3.11) is satisfied, and use $|0_\Delta\rangle_c$ and $|1_\Delta\rangle_c$ to denote the code states defined in specified Δ subspace. Approximately we have

$$\begin{aligned}
&|0_\Delta\rangle_{1c}\,|0_\Delta\rangle_{2c} \to \gamma\,|0_\Delta\rangle_{1c}\,|0_{\Delta+1}\rangle_{2c}\,,\\
&|0_\Delta\rangle_{1c}\,|1_\Delta\rangle_{2c} \to (-1)\gamma\,|0_\Delta\rangle_{1c}\,|1_{\Delta+1}\rangle_{2c}\,,\\
&|1_\Delta\rangle_{1c}\,|0_\Delta\rangle_{2c} \to i^{\Delta+1}\gamma(t_0)\,|1_\Delta\rangle_{1c}\,|1_{\Delta+1}\rangle_{2c}\,,\\
&|1_\Delta\rangle_{1c}\,|1_\Delta\rangle_{2c} \to (-1)i^{\Delta+1}\gamma(t_0)\,|1_\Delta\rangle_{1c}\,|0_{\Delta+1}\rangle_{2c}\,.
\end{aligned} \tag{3.28}$$

Here, we just omit some overall factors that are the same for all the final states in the expressions above. After the evolution, the final states should go through a recovery channel by syndrome ($\hat{\Delta}$) measurement and error correction. We have a more detailed discussion in Appendix B.3 on the recovery strategy based on the outcome of the final Δ we measured. Briefly, the recovery process will map $|0_{\Delta+1}\rangle_{2c}$ to $|0_\Delta\rangle_{2c}$ and map $|1_{\Delta+1}\rangle_{2c}$ to $(-1)\,|1_\Delta\rangle_{2c}$ for the target qubit.

After the recovery, if the control qubit is in $|1_\Delta\rangle_{1c}$, then in addition to the $e^{-i\pi(|\gamma|^2-\Delta/2)}$ phase that will be achieved in the no error case we mentioned above, there will be an extra $\exp[-i\frac{\pi}{2}(1-t_0/T)]$ phase on the final states, since $\gamma(t)=\gamma\exp(i\frac{\pi}{2}\frac{t}{T})$. So, if we do not know what t_0 is, this induced time-dependent phase cannot be corrected.

Indeed, this CX gate is still bias-preserving, since the error induced by single-photon loss is still Z type of error, which is the dominant one. However, it violates one of the proposed merits of pair-cat code that the single-photon loss error in either mode will not cause errors in the code. To solve this issue, one method is to introduce real-time monitoring of photon number difference $\hat{\Delta}_2$ on the target qubit to keep track of the time when the loss error might happen. It is in principle doable since $\hat{\Delta}_2$ commutes with all the generators in the CX gate design and the code states will not be changed during measurement since they are always eigenstates of $\hat{\Delta}_2$, regardless of whether they suffer from loss errors or not. If we know the

specific time that the single-photon loss happens, we can apply a $Z_1(\theta)$ gate on the control qubit to correct the induced phase. Therefore, the leading uncorrectable error will again be suppressed to higher order, which comes from both the inaccuracy of the phase correction due to the finite time interval $\delta\tau$ of two consecutive $\hat{\Delta}$ measurements, and the situation that both $\hat{a}$ and $\hat{b}$ error happen in the same time interval between two $\hat{\Delta}$ measurement. We have a detailed analysis of those gate errors in Appendix B.4. It is worth mentioning that, in the limit that the time interval of two $\hat{\Delta}$ measurements can be ignored, due to the large dissipation gap the optimal CX gate error probability will decrease as γ increases. This is in contrast to the cat code case, where the optimal error probability of CX is independent of the size of the cat states.

We finally discuss the issue that the real-time $\hat{\Delta}$ measurement might not be perfect, which means that the measurement device may misidentify the photon number difference Δ between $\hat{a}$ and $\hat{b}$ modes at the measurement time. In the recovery process, this can be treated as both $\hat{a}$ loss and $\hat{b}$ loss happen in two consecutive $\delta\tau$ intervals, and we will apply an unnecessary $\hat{Z}_c$ operation that will cause a logical failure in the end. To make our protocol fault-tolerant against this error, one simple way is to introduce multiple auxiliary systems that are dispersively coupled with $\hat{a}$ and $\hat{b}$ modes to independently perform $\hat{\Delta}$ measurements at the same time. Then we can do a majority voting to get one "logical" Δ value, so that the effect from measurement error can be suppressed to higher order. Similarly, even if we only have one auxiliary system for $\hat{\Delta}$ monitoring, we can group 3 consecutive $\hat{\Delta}$ measurement results and perform majority voting within those 3 outcomes for a single "logical" Δ value. Therefore, a single measurement error can be correctly identified, while the gate error probability induced by photon losses (as analyzed in Appendix B.4) will keep its scaling but get slightly increased. It is because the effective time interval for two "logical" Δ values will be a little larger. For example, $\hat{a}$ loss and $\hat{b}$ loss happening within $3\delta\tau$ time may be treated as a measurement error and will not be corrected.

Toffoli gate

Since the Toffoli gate is just the Control-CX gate, we can extend the strategy introduced in the construction of the CX gate for the Toffoli case. For the cat code, the jump operators and rotation Hamiltonian have been proposed as [68]

$$\hat{F}_1 = \hat{a}_1^2 - \alpha^2, \quad \hat{F}_2 = \hat{a}_2^2 - \alpha^2, \tag{3.29a}$$

$$\begin{aligned}\hat{F}_3 = \hat{a}_3^2 &- \frac{1}{4}(\hat{a}_1+\alpha)(\hat{a}_2+\alpha) + \frac{1}{4}(\hat{a}_1-\alpha)(\hat{a}_2+\alpha) \\ &+ \frac{1}{4}(\hat{a}_1+\alpha)(\hat{a}_2-\alpha) - \frac{1}{4}e^{2i\pi\frac{t}{T}}(\hat{a}_1-\alpha)(\hat{a}_2-\alpha),\end{aligned} \tag{3.29b}$$

with

$$\hat{H}_{\text{Tof,rot}} = -\frac{\pi}{2T}\left(\frac{\hat{a}_1-\alpha}{2\alpha}\otimes\frac{\hat{a}_2^\dagger-\alpha^*}{2\alpha^*} + \text{h.c.}\right)\otimes(\hat{a}_3^\dagger\hat{a}_3 - |\alpha|^2). \tag{3.30}$$

While, in the pair-cat code case, the jump operators can be chosen as

$$\hat{F}_1 = \hat{a}_1^2\hat{b}_1^2 - \gamma^4, \quad \hat{F}_2 = \hat{a}_2^2\hat{b}_2^2 - \gamma^4, \tag{3.31a}$$

$$\begin{aligned}\hat{F}_3 = \hat{a}_3^2\hat{b}_3^2 &- \frac{1}{4}(\hat{a}_1\hat{b}_1+\gamma^2)(\hat{a}_2\hat{b}_2+\gamma^2) + \frac{1}{4}(\hat{a}_1\hat{b}_1-\gamma^2)(\hat{a}_2\hat{b}_2+\gamma^2) \\ &+ \frac{1}{4}(\hat{a}_1\hat{b}_1+\gamma^2)(\hat{a}_2\hat{b}_2-\gamma^2) - \frac{1}{4}e^{2i\pi\frac{t}{T}}(\hat{a}_1\hat{b}_1-\gamma^2)(\hat{a}_2\hat{b}_2-\gamma^2).\end{aligned} \tag{3.31b}$$

Besides, the Hamiltonian to compensate the non-adiabatic error is

$$\hat{H}_{\text{Tof,rot}} = -\frac{\pi}{4T}\left(\frac{\hat{a}_1\hat{b}_1-\gamma^2}{2\gamma^2}\otimes\frac{\hat{a}_2^\dagger\hat{b}_2^\dagger-\gamma^{*2}}{2\gamma^{*2}} + \text{h.c.}\right)\otimes(\hat{a}_3^\dagger\hat{a}_3 + \hat{b}_3^\dagger\hat{b}_3 - 2|\gamma|^2). \tag{3.32}$$

Some extra work in the CX gate construction should also be done here. The induced phase during the rotation of the target qubit can be corrected by applying both $Z(\theta)$ and $ZZ(\theta)$ gates on the two control qubits, or just use carefully chosen γ such that this phase has no effect. Besides, we need real-time monitoring of $\hat{\Delta}_3$ on the target qubit to correct the error induced by single-photon loss on that qubit.

Table 3.2: The scaling of optimal total Z type of error probability in bias-preserving gates for cat code and pair-cat code (using perfect real-time $\hat{\Delta}$ monitoring with time interval $\delta\tau$ on each pair-cat qubit). Here C and C' are some constant numbers (see Appendix B.4).

	Cat code [59]	Pair-cat code
Z(θ)	$O\left(\frac{1}{\alpha}\sqrt{\frac{\kappa_1}{\kappa}}\right)$	$O\left(\frac{\kappa_1}{\gamma^2}\sqrt{\frac{\delta\tau}{\kappa}}\right)$
ZZ(θ)	$O\left(\frac{1}{\alpha}\sqrt{\frac{\kappa_1}{\kappa}}\right)$	$O\left(\frac{\kappa_1}{\gamma^2}\sqrt{\frac{\delta\tau}{\kappa}}\right)$
X	$O(\kappa_1\alpha^2 T) \xrightarrow{T\to 0} 0$	$O(\kappa_1^2\gamma^4\delta\tau T) \xrightarrow{T\to 0} 0$
CX	$O\left(\sqrt{\frac{\kappa_1}{\kappa}}\right)$	$O\left(\sqrt{\frac{\kappa_1}{\kappa}}\sqrt{\frac{\kappa_1\delta\tau}{\gamma^2}}\sqrt{1+C\kappa\kappa_1\gamma^8(\delta\tau)^2}\right)$
Toffoli	$O\left(\sqrt{\frac{\kappa_1}{\kappa}}\right)$	$O\left(\sqrt{\frac{\kappa_1}{\kappa}}\sqrt{\frac{\kappa_1\delta\tau}{\gamma^2}}\sqrt{1+C'\kappa\kappa_1\gamma^8(\delta\tau)^2}\right)$

3.2.2 Hamiltonian stabilization scheme

We note that in some way Hamiltonian stabilization scheme is similar to the dissipative stabilization scheme. We have already got a sense of such similarity from the structure of stabilization Hamiltonian $\hat{H} = -K\hat{F}^\dagger\hat{F}$ where $\hat{F} = \hat{a}^2\hat{b}^2 - \gamma^4$ is the jump operator we use in the dissipative stabilization scheme. We can make use of such similarities to construct bias-preserving operations in the Hamiltonian stabilization scheme.

Preparation of $|\pm\rangle_c$ states

To prepare $|+\rangle_c$ state of the pair-cat code, we can use a similar method as the state preparation in the Kerr-cat scheme proposed in [69]. Since $|\pm_{\gamma,\Delta}\rangle$ are always the most excited eigenstates of the Hamiltonian shown in Eq. (3.16), and in $|\gamma| \to 0$ limit we have $|+_{\gamma,\Delta}\rangle \sim |0,\Delta\rangle$ and $|-_{\gamma,\Delta}\rangle \sim |1,\Delta+1\rangle$, we can first prepare $|0,\Delta\rangle$ or $|1,\Delta+1\rangle$ and adiabatically increase $\gamma(t)$ from 0 to the final γ we want. Since both $\hat{\Delta}$ and parity are conserved, we will reach the corresponding $|\pm_{\gamma,\Delta}\rangle$ state finally in the adiabatic limit.

Measurement in X basis

This can be done in the same way as proposed in the driven-dissipative scheme, since the protection has to be turned off during the measurement process.

$Z(\theta)$ and $ZZ(\theta)$ gates

We can still use the same Hamiltonian in Eq. (3.20) to generate $Z(\theta)$ and $ZZ(\theta)$ accordingly. It is because that the Hamiltonian in Eq. (3.16) could provide the protection of the code space because of the $O(|\gamma|^6)$ energy gap, and according to Eq. (3.21), within the code space $\hat{H}_Z$ and $\hat{H}_{ZZ}$ serve as the generators of $Z(\theta)$ and $ZZ(\theta)$ gates.

X, CX and Toffoli gates

The ideas for construction of these three bias-preserving operations are quite similar: they all require conditioned adiabatically changing of stabilization parameter $\gamma(t)$ while keeping $|\gamma(t)|$ large all the time, and use another Hamiltonian to actively change the code states to reduce the error from non-adiabaticity due to the finite evolution time.

So, we can use the following Hamiltonian to implement X gate,

$$\hat{H}_X = -K\hat{F}^\dagger\hat{F} + \hat{H}_{X,\text{rot}}, \tag{3.33}$$

where $\hat{F} = \hat{a}^2\hat{b}^2 - \gamma^4(t)$ with $\gamma(t) = \gamma e^{i\frac{\pi}{2}\frac{t}{T}}$ in the first term provides stabilization of the code space and the second term $\hat{H}_{X,\text{rot}} = -\frac{\pi}{2T}(\hat{n}_a + \hat{n}_b)$ can actively change code states according to $\gamma = \gamma(t)$ to compensate the error induced by non-adiabaticity.

For the CX gate, we can use the following $\hat{H}_{\text{CX}}$:

$$\hat{H}_{\text{CX}} = -K(\hat{F}_1^\dagger\hat{F}_1 + \hat{F}_2^\dagger\hat{F}_2) + \hat{H}_{\text{CX,rot}}, \tag{3.34}$$

where $\hat{F}_1$ and $\hat{F}_2$ is defined in Eq. (3.26) to provide stabilization and Hamiltonian $\hat{H}_{\text{CX,rot}}$ is defined in Eq. (3.27) to provide mitigation of non-adiabatic error.

The Toffoli gate can be constructed with $\hat{H}_{\text{Tof}}$:

$$\hat{H}_{\text{Tof}} = -K \sum_{j=1}^{3} \hat{F}_j^\dagger \hat{F}_j + \hat{H}_{\text{Tof,rot}}, \tag{3.35}$$

where $\hat{F}_j$ is defined in Eq. (3.31) and Hamiltonian $\hat{H}_{\text{Tof,rot}}$ is defined in Eq. (3.32).

Same as the driven-dissipative case, the real-time $\hat{\Delta}$ monitoring on target qubits and phase correction on control qubits in both CX and Toffoli gates are also needed here.

3.3 Discussion and conclusion

It is possible to generalize the pair-cat encoding protocol into a multimode multicomponent case in order to correct more photon loss and gain errors [70]. In general, we could stabilize a d level qudit in M modes with jump operator $\hat{F} = (\hat{a}^d)^{\otimes M} - \gamma^{dM}$, and syndromes can be monitored by measuring all the photon number differences between neighboring modes. In this way, any amount of photon loss happening in arbitrary $M-1$ modes can be distinguished, or if $M \geq 3$ then any amount of photon loss or gain happening in $\frac{M-1}{2}$ modes corresponds to a unique syndrome. But there will be a logical error on the qudit if all of the modes suffer from a photon loss together, provided that there is no further encoding on the logical qudit within the d level subspace.

For the multimode pair-cat qubit case ($d = 2$), it is straightforward to achieve the bias-preserving operations from the generalization of the 2-mode pair-cat code, just as the generalization from the cat code to the 2-mode pair-cat. It will be tricky to talk about bias-preserving in cat or pair-cat qudits and their future concatenations, since different single-qudit errors may correspond to different numbers of photon loss or gain which can happen with different probabilities. But, still the continuous monitoring of the syndrome is essential

in gate designs, especially in the generalized control-X gates where only a single-photon loss on the target qudit will induce a time-dependent phase shift on the control qudit. But the continuous syndrome monitoring is hard for multicomponent cat codes with stabilization. We will leave the discussion of qudit properties for further research.

Besides, instead of using continuous syndrome monitoring as we mentioned, we can also try to engineer jump operators $\hat{a}^\dagger \hat{P}_{\Delta+1}$ and $\hat{b}^\dagger \hat{P}_{\Delta-1}$ to achieve the autonomous error correction against single-photon loss [70]. It can give similar scaling results of the gate error probability while further reducing the overhead from feedback control. The details of this proposal are also worth working out in further work.

In summary, we generalize the idea of construction of bias-preserving operations for cat code into pair-cat code to protect against a single-photon loss in either mode during gate operations. Continuous syndrome monitoring plays an essential role in the gate design to suppress errors. The generalization is quite straightforward due to the strong similarity between the two types of codes. Besides, the Hamiltonian protection of the pair-cat code is investigated, and the large energy gap between code space and other states has also been found and numerically verified, which is another interesting feature in the pair-cat code that is worth exploring in the future.

CHAPTER 4

DARK SPIN-CAT ENCODING IN ATOMIC STRUCTURES

4.1 Introduction

In the previous chapter, I have discussed the properties of the cat code and one of its multimode generalizations, pair-cat, with a highlight on their biased noise structures and the protocols to engineer BP operations. It will be interesting to ask if we can still find a structured error pattern with suitable choices of encoding in other physical platforms so that the noise is easy to correct.

In recent years, the neutral atom system has become one of the most promising platforms for quantum simulation and quantum computing, due to the allowed high-fidelity operations and the reconfigurability of the atom arrays. One of the major limitations on the entangling gate fidelities is the decay from the Rydberg levels during gate execution. The fluctuating magnetic field may also cause dephasing of the quantum states, which further reduces the achieved fidelity.

To tackle the Rydberg decay problem, several schemes have been proposed by making the resulting error pattern more structured. One direction is, again, along the introduction of the Pauli bias. In Ref. [81], the two computational basis states are encoded with far-separated magnetic quantum numbers. The bit-flip is unlikely to happen due to such a separation. Another direction is based on the erasure conversion [82]. The location information about which atoms suffer from the error greatly improves the error correction performance and provides a higher error threshold. On the hardware level, people tried to encode the states in a metastable subspace so that a Rydberg decay event could be caught as long as the final state is not in the encoded manifold. Further proposals on the biased erasure have been made, which combine both Pauli bias and the erasure conversion as mentioned above [83]. This will result in a more than 10% error threshold given suitable choices of the error-correcting

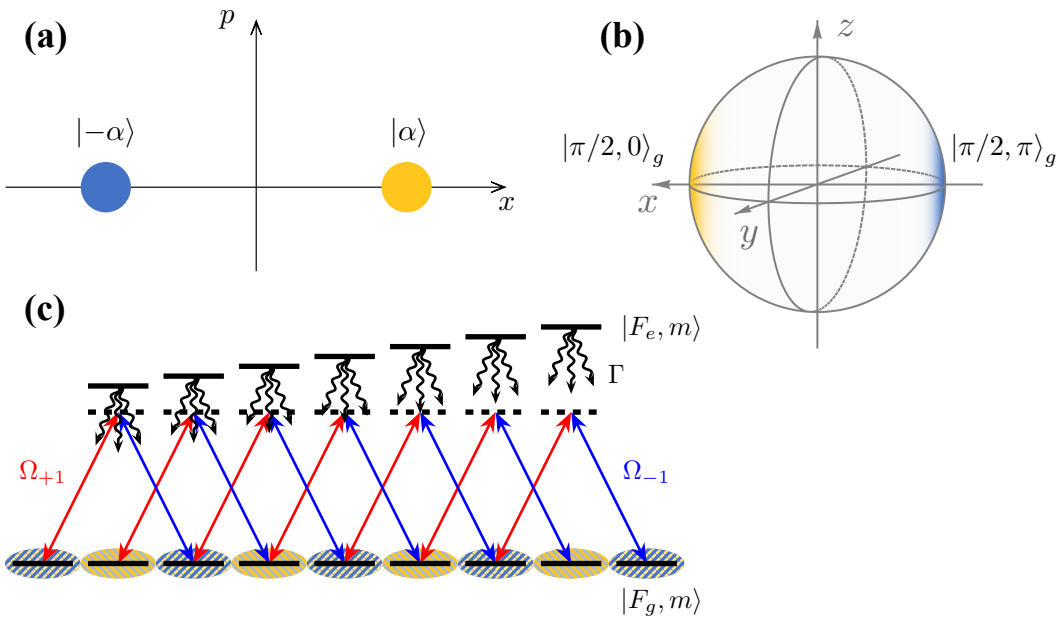

Figure 4.1: Schematic comparison between bosonic cat and dark spin-cat encoding. (a) Bosonic cat encoding, where the code subspace is spanned by two coherent states $|\pm\alpha\rangle$. (b) Dark spin-cat encoding with two antipodal SCSs on the GBS as encoded computational basis states. (c) Dark spin-cat states in the atomic level structures. Two dark states are in the even- and odd-parity subspaces of the ground hyperfine (F_g) manifold.

codes.

In this chapter, I will discuss our proposal for the problem based on the framework of the cat encoding. Consider a hyperfine manifold with a large total angular momentum. This is usually achieved for isotopes with large nuclear spin I, like ^{87}Sr with $I = 9/2$. Now, the phase space of the long hyperfine spin is a sphere, in contrast to a plane as in the bosonic case. Below, I will call it the generalized Bloch sphere (GBS). In our spin-cat protocol, we encode the computational basis of a qubit using two spin-coherent states (SCSs) antipodal with each other on the phase space [see Fig. 4.1(b)]. To dynamically protect the encoded subspace, we apply linear drives that couple the encoded hyperfine manifold with an excited hyperfine manifold. In combination with the spontaneous or engineered decay from the excited levels, the encoded subspace becomes the "dark subspace" if the driving parameters are chosen properly. This provides an autonomous leakage reduction when the state is affected by the

noise. Together with the largest separation of the two encoded SCSs on the GBS, the bit-flip rate is exponentially suppressed with the spin length, and therefore recovers the biased error structure as the bosonic cat counterpart. If the two SCSs are on the equators of the GBS while the polarization of the decay is suitably engineered, then the dephasing error can be further autonomously corrected.

The encoding will be meaningless if we cannot find appropriate protocols for gate execution. Similar to the bosonic cat [149], arbitrary single-qubit operations can be achieved by adiabatically manipulating the location of the two SCSs on the GBS determined by the amplitudes and phases of the linear drives. However, such a protocol in general does not preserve the Pauli error bias, as it will convert phase-flip errors to bit-flip errors. Alternatively, we again developed a set of BP operations including the CZ and CX entangling gates. The CZ implementation fits into the biased erasure framework if the erasure conversion is allowed, while the biased erasure CX is also possible with a further assumption that the Rydberg decay events can be frequently monitored.

The chapter is structured as follows. Sec. 4.2 describes the system we focus on and the structures for the dark subspace that we use to encode a qubit. In Sec. 4.3, I will give an in-depth discussion on the autonomous stabilization of the code space and the resulting biased noise structure on the memory level. In Sec. 4.4, I will explain our protocols to implement BP operations including CZ and CX gates, based on the spin-cat encoding.

4.2 Dark spin-cat encoding

4.2.1 System setup

To start with, we need to specify the system of consideration first. It contains laser drives that couple a ground hyperfine manifold with the total angular momentum F_g and an excited hyperfine manifold with F_e. F_g and F_e here can be both integers or both half-integers. States

in those manifolds with different magnetic quantum numbers are labeled by $|F_{g(e)}, m\rangle$ with $m = -F_{g(e)}, -F_{g(e)}+1, \ldots, F_{g(e)}$, respectively. Moreover, the states in the excited levels may decay into the ground levels through the emission of a photon. The Hamiltonian of the system in the lab frame can be expressed as

$$\hat{H}_{\text{lab}} = \omega_{eg}\hat{P}_e + \delta_e \hat{F}_{e,z} + \delta_g \hat{F}_{g,z} + \frac{1}{2}\sum_{q=0,\pm 1}(e^{-i\omega_q t}\Omega_q \hat{D}_{eg,q} + \text{h.c.}). \tag{4.1}$$

Here we denote $\hat{P}_{g(e)} = \sum_{m=-F_{g(e)}}^{F_{g(e)}} |F_{g(e)}, m\rangle\langle F_{g(e)}, m|$ is the projector onto the ground (excited) manifold, and ω_{eg} is the energy difference between the excited and the ground levels when there is no external field. $\delta_{g(e)} = g_{F,g(e)}\mu_B B_z$ characterizes the energy shift of each level due to an external magnetic field along z-axis, where $g_{F,g(e)}$ is a constant factor that depends on the specific choices of the hyperfine manifold and μ_B is the Bohr magneton. $\hat{F}_{g(e),\xi}$ is the angular momentum operator for each hyperfine manifold with the direction $\xi = x, y, z$. To model the coupling between ground and excited levels via the electric dipole, we use $\hat{D}_{eg,q} = C_{F_g,m;1,q}^{F_e,m+q}|F_e, m+q\rangle\langle F_g, m|$ to indicate the relative coupling strength between different $|F_g, m\rangle \leftrightarrow |F_e, m+q\rangle$ pairs, where $C_{F_g,m;1,q}^{F_e,m+q} = \langle F_g, m; 1, q|F_e, m+q\rangle$ is the Clebsch-Gordan coefficient as a fact of the Wigner-Eckart theorem [150]. $q = 0, \pm 1$ here is determined by the polarization of the lasers, while Ω_q contains the overall driving amplitudes and phases, and ω_q is the frequency of the drives.

If the frequencies of the lasers with different polarizations satisfy the resonant condition $\omega_q - \omega_{q'} = \delta_g(q - q')$, we can find a rotating frame with

$$\hat{U}(t) = \exp\{i[\omega_0 \hat{P}_e + \delta_g(\hat{F}_{g,z} + \hat{F}_{e,z})]t\}, \tag{4.2}$$

so that the resulting Hamiltonian is in a time-independent form

$$\begin{aligned}\hat{H}_{\mathrm{DS}} &= \hat{U}\hat{H}_{\mathrm{lab}}\hat{U}^{\dagger} + i\dot{\hat{U}}\hat{U}^{\dagger} \\ &= \Delta\hat{P}_e + \delta\hat{F}_{e,z} + \frac{1}{2}\sum_{q=0,\pm1}(\Omega_q\hat{D}_{eg,q} + \mathrm{h.c.}),\end{aligned} \tag{4.3}$$

where $\Delta = \omega_{eg} - \omega_0$ and $\delta = \delta_e - \delta_g$. I will use this Hamiltonian for the later discussion.

Then we would like to model how the decay process affects the state of the atom. The dynamics of the atomic state cannot be treated as a unitary process due to the coupling with the photon modes in the environment, and therefore, we need a master equation in general to describe this. If the atom is in vacuum so that the photons can be emitted into any direction with equal probability, and if the emitted photons couple with the atom through the electric dipole moment again, we can model the dynamics as

$$\frac{\mathrm{d}\hat{\rho}}{\mathrm{d}t} = \mathcal{L}_{\mathrm{sp}}\hat{\rho} := -i[\hat{H}_{\mathrm{DS}}, \hat{\rho}] + \Gamma\sum_{q=0,\pm1}\mathcal{D}[\hat{D}_{ge,q}]\hat{\rho}, \tag{4.4}$$

where $\mathcal{L}_{\mathrm{sp}}$ is the full Lindbladian, $\hat{D}_{ge,q} = \hat{D}^{\dagger}_{eg,q}$ is the jump operator indicating that the decay occurs through the dipole coupling with the environmental modes, and $\mathcal{D}[\hat{D}_{ge,q}]\bullet = \hat{D}_{ge,q}\bullet\hat{D}^{\dagger}_{ge,q} - \frac{1}{2}\{\hat{D}^{\dagger}_{ge,q}\hat{D}_{ge,q}, \bullet\}$ is the dissipator correspondingly.

The dissipative process can be modified if the atom is placed in a lossy cavity. Ref. [151] described a situation where the atom couples with two degenerate eigenmodes of the cavity with both polarizations perpendicular to the cavity axis. With a proper choice of the quantized axis, the master equation now becomes

$$\frac{\mathrm{d}\hat{\rho}}{\mathrm{d}t} = \mathcal{L}_{\mathrm{eng}}\hat{\rho} = -i[\hat{H}_{\mathrm{DS}}, \hat{\rho}] + \Gamma\sum_{q=\pm1}\mathcal{D}[\hat{D}_{ge,q}]\hat{\rho}, \tag{4.5}$$

where the jump operator $\hat{D}_{ge,q}$ with the $q = 0$ term is removed from Eq. (4.4). Later I will show how this modification leads to the autonomous correction of the dephasing error under

the spin-cat encoding.

4.2.2 Structures of the dark subspace

After specifying the Hamiltonian and the jump operators of the system, I will now discuss some properties it implies. Let us focus on the Hamiltonian term $\hat{H}_{\mathrm{DS}}$ shown in Eq. (4.3) first. In general, consider a system with d_e excited levels and d_g ground levels. If $d_g > d_e$, and the restriction of the Hamiltonian in the ground subspace is zero (i.e., $\hat{H}_{gg} = \hat{P}_g \hat{H} \hat{P}_g = 0$ with $\hat{P}_g$ the projector), then the system contains at least $(d_g - d_e)$ number of linearly-independent eigenstates in the ground subspace with eigenenergy 0. This claim can be easily justified via the singular value decomposition of the ground-excited coupling term $\hat{H}_{eg} = \hat{P}_e \hat{H} \hat{P}_g$ [152]. Even if we introduce the decay from the excited levels in the system, those states still form a steady-state subspace as they are not affected by the decay. Other states outside this subspace may come into it due to the dissipative process.

This observation motivates us to think whether we can encode a qubit into this "dark" subspace so that if the population leaves the subspace due to the noise, it may autonomously come back under dissipation. Therefore, we will choose $d_g - d_e = 2$ in the system, which corresponds to $F_g - F_e = 1$.

Then we will try to solve the dark states, which only depend on the relative amplitudes and phases of those Ω_q. One simple case comes when $\Omega_{+1} = \Omega_{-1} = 0$ and $\Omega_0 = \Omega$. As illustrated in Fig. 4.2(c), it is clear that the two dark states are $|F_g, \pm F_g\rangle$, which correspond to the SCSs at the north and the south pole of the GBS. In general, we will use the polar angle θ and the azimuthal angle ϕ to label each SCS $|\theta, \phi\rangle_g$ in the F_g-spin, as it can be achieved by the rotation of an SCS starting at the north pole:

$$|\theta, \phi\rangle_g := e^{-i\phi \hat{F}_{g,z}} e^{-i\theta \hat{F}_{g,y}} |F_g, F_g\rangle \,. \tag{4.6}$$

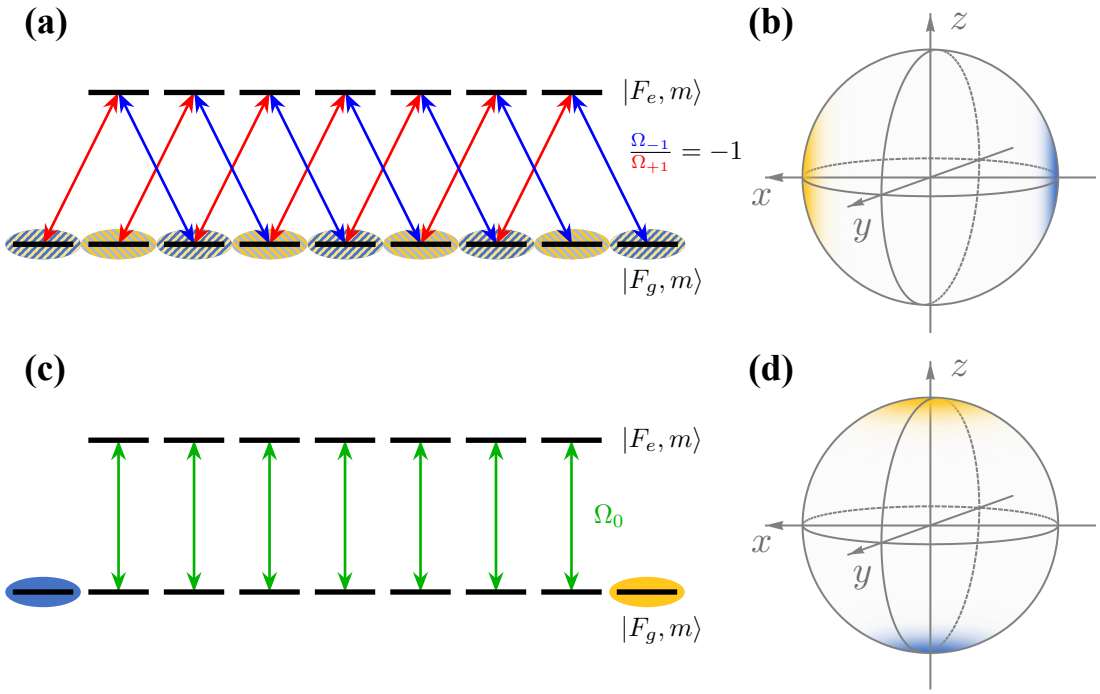

Figure 4.2: Dark state structures with different stabilization drives. (a) Both $q = \pm 1$ drives are applied with $\Omega_{-1}/\Omega_{+1} = -1$. (b) The two dark states in this case are the even and odd superpositions of the two SCSs positioned antipodally along the x-axis. (c) Only the $q = 0$ polarized drive is applied. (d) Clearly the two dark states are $|F_g, \pm F_g\rangle$, which correspond to the north and the south pole on the GBS.

From the physical aspect, the nonzero Ω_0 in the case above corresponds to the π-polarized light. It is linear-polarized with the electric field oscillating along the z direction. Similarly, if we only have a linear-polarized light with the electric field along the x direction, then the two dark states should be the SCSs pointing along the $\pm x$-axis. In this case, $\Omega_{+1} = -\Omega_{-1}$ while $\Omega_0 = 0$. Now the Hilbert space can be separated into two subspace: one spanned by $\{|F_g, 2k - F_g\rangle, |F_e, 2k - F_e\rangle\}$ and $\{|F_g, 2k - F_g + 1\rangle, |F_e, 2k - F_e + 1\rangle\}$ $(k \in \mathbb{N})$ without any Hamiltonian coupling term between them. We can introduce a parity operator

$$\hat{\Pi} := \exp\{-i\pi[(\hat{F}_{g,z} + \hat{F}_{e,z}) + (F_g\hat{P}_g + F_e\hat{P}_e)]\}, \tag{4.7}$$

which commutes with the Hamiltonian in the current situation. In this chapter, I will encode

the code states within the dark subspace described here, with the choice of computational basis

$$
\begin{aligned}
|0\rangle_c &:= e^{-i\frac{\pi}{2}\hat{F}_{g,y}}\,|F_g, F_g\rangle = |\frac{\pi}{2}, 0\rangle_g\,, \\
|1\rangle_c &:= e^{-i\frac{\pi}{2}\hat{F}_{g,y}}\,|F_g, -F_g\rangle = e^{-i\pi F_g}\,|\frac{\pi}{2}, \pi\rangle_g\,,
\end{aligned}
\tag{4.8}
$$

which are at the most separated points on the GBS. A superposition of these two basis states, $|\pm\rangle_c = \frac{|0\rangle_c \pm |1\rangle_c}{\sqrt{2}}$, lies in the even (odd) parity subspace indicated by the ± 1 eigenvalue of the $\hat{\Pi}$ operator. This is also illustrated in Fig. 4.2(a).

In fact, given an arbitrary choice of those Ω_q parameters, the two-dimensional dark subspace is always spanned by two SCSs which in general may not be orthogonal to each other. I will put the demonstration in Appendix C.1. This observation enables us to manipulate the places of the two SCSs on the GBS through adjusting the relative amplitudes and phases of Ω_q, which will be used for single-qubit gate execution. Moreover, in Chapter 5 I will discuss more properties of the generalized spin-cat encoding where the code stabilization comes from this feature. Currently, I will still stick to the case where two dark SCSs are antipodal on the GBS.

4.3 Autonomous stabilization and biased noise structure

4.3.1 Intuitive picture for the biased noise structure

In the previous section, I have discussed the choice of encoding in our protocol named "spin-cat" [85], where we encode the computational basis of a qubit using the two largest separated SCSs on the GBS. Now I will introduce the physical relevant noise in the system and investigate how the dissipative process helps to shape the structures of the error channel or even correct them. For simplicity, in this section below I will fix $\Delta = \delta = 0$ in the Hamiltonian shown in Eq. (4.3), and also choose $(\Omega_{+1}, \Omega_0, \Omega_{-1}) = (\frac{\Omega}{\sqrt{2}}, 0, -\frac{\Omega}{\sqrt{2}})$ so that

Eq. (4.8) will be the desired encoding. The Hamiltonian is now written as

$$\hat{H}_{\mathrm{DS},x} = \frac{\Omega}{2\sqrt{2}}(\hat{D}_{eg,+1} - \hat{D}_{eg,-1} + \mathrm{h.c.}) \tag{4.9}$$

I will also denote $|F_g, m\rangle_x := e^{-i\frac{\pi}{2}\hat{F}_{g,y}}|F_g, m\rangle$ for further use. Therefore, $|0\rangle_c = |F_g, F_g\rangle_x$ and $|1\rangle_c = |F_g, -F_g\rangle_x$.

Similar to the bosonic cat situation, the physical relevant noise like random magnetic field fluctuation $\hat{F}_{g,\xi}$ ($\xi = x, y, z$) acts locally on the phase space (GBS), which only couples states whose magnetic quantum numbers along a certain axis differ by at most 1. This is a general feature of operators as components of a vector, due to the Wigner-Eckart theorem. Since our encoded computational basis states are the largest separated on the GBS, in the $F_g \gg 1$ regime, it is unlikely that one basis state could flip into another one under the noisy process. Rather, it will come to a state close to it on the GBS, and as a result, there will be a high probability that the population of the state will go back to the original place after dissipative stabilization. The bit-flip rate will again be suppressed as the system size (here it is the spin length F_g) increases.

4.3.2 Conserved quantities of the Lindbladian

To justify the intuition described here, we need more careful investigations on the dynamics provided by the full Lindbladian $\mathcal{L}$ as that in Eq. (4.4) or (4.5). Our approach is to solve the conserved quantities associated with $\mathcal{L}$ first, which help us to determine how an arbitrary state ends up in the steady state subspace after evolving sufficiently long time under $\mathcal{L}$ [54]. A conserved quantity $\hat{J}$ is an operator that the overlap between $\hat{J}^\dagger$ and any state evolving under $\mathcal{L}$ will be constant over time, specifically, $\mathrm{Tr}[\hat{J}^\dagger e^{\mathcal{L}t}\hat{\rho}] = \mathrm{Tr}[\hat{J}^\dagger\hat{\rho}]$ applies for arbitrary $\hat{\rho}$. It implies that, a conserved quantity $\hat{J}$ of a Lindbladian $\mathcal{L}$ with Hamiltonian $\hat{H}$ and several

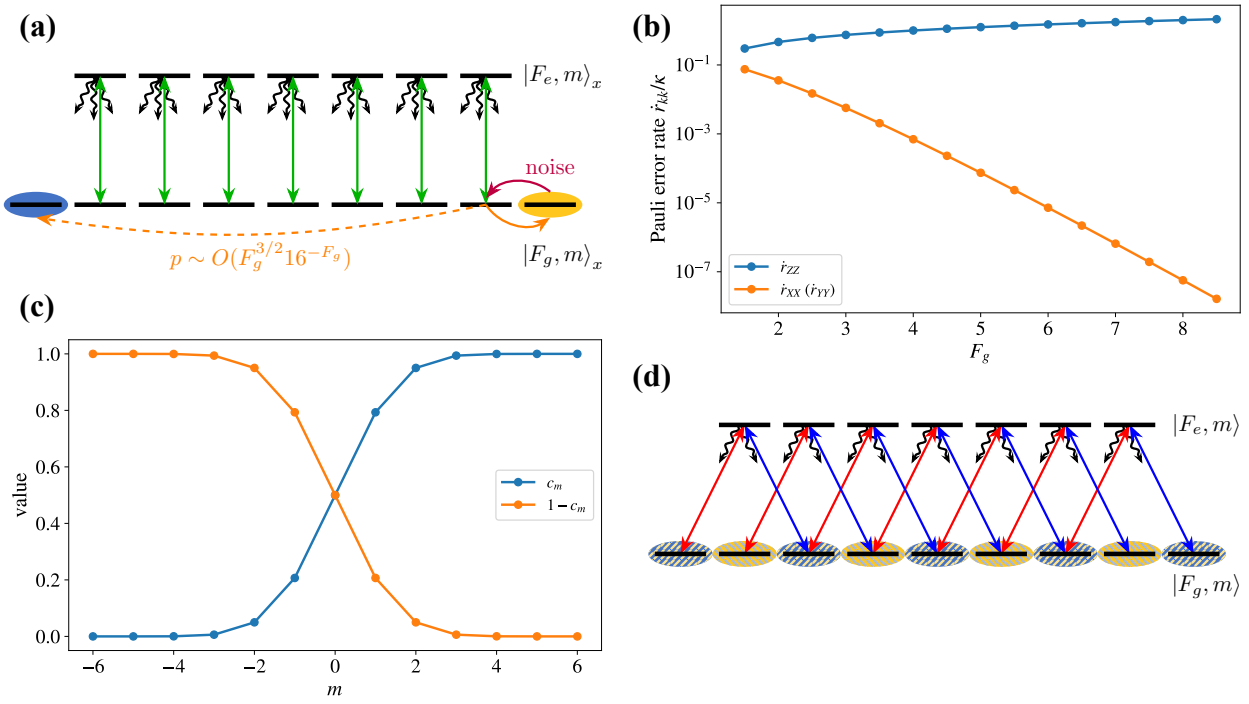

Figure 4.3: Autonomous stabilization and the biased noise structure. (a) Illustration of the bit-flip suppression. If the noise acts locally in the phase space (GBS) like only changing the m quantum number by at most 1, then the fraction of the population to end up at the other SCS after autonomous stabilization will be exponentially suppressed with spin-length F_g. (b) The Pauli error rate $\dot{r}_{kk}$ under perturbative $\kappa\mathcal{D}[\hat{F}_{g,z}]$ noise. The F_g-dependence is plotted using Eq. (4.18). (c) The c_m coefficients (4.13) in conserved quantities $\hat{J}_{00}$ and $\hat{J}_{11}$, when $F_g = 6$. (d) Illustration on the engineered dissipation where only $q = \pm 1$ decays are considered. The parity $\hat{\Pi}$ is a strong symmetry now and will be preserved in the dynamics.

jump operators $\hat{D}_k$ should satisfy the following equation

$$\mathcal{L}^{\ddagger}(\hat{J}) = i[\hat{H}, \hat{J}] + \sum_k (\hat{D}_k^{\dagger} \hat{J} \hat{D}_k - \frac{1}{2}\{\hat{D}_k^{\dagger}\hat{D}_k, \hat{J}\}) = 0. \tag{4.10}$$

Moreover, the number of linear-independent conserved quantities should be equal to the number of linear-independent steady operators ($\hat{\rho}$ such that $\mathcal{L}(\hat{\rho}) = 0$), since they are left or right eigenmatrices of $\mathcal{L}$ with eigenvalue 0 and in general there is a correspondence between them [148]. In our model, the dark subspace is a 2-dimensional Hilbert space, which gives us 4 linear-independent steady states. Therefore, we should also find 4 linear-independent

conserved quantities.

When choosing $\hat{H}_{\mathrm{DS},x}$ in Eq. (4.9) as the Hamiltonian, we can find the exact solutions for all of the 4 conserved quantities associated with $\mathcal{L}_{\mathrm{sp}}$. Here, I just write down the solutions:

$$\begin{cases} \hat{J}_{00} = |F_g, F_g\rangle_x\langle F_g, F_g| + \sum_{m=-F_g+1}^{F_g-1} c_m \Big(|F_g, m\rangle_x\langle F_g, m| + |F_e, m\rangle_x\langle F_e, m| \Big), \\ \hat{J}_{11} = |F_g, -F_g\rangle_x\langle F_g, -F_g| + \sum_{m=-F_g+1}^{F_g-1} c_{-m} \Big(|F_g, m\rangle_x\langle F_g, m| + |F_e, m\rangle_x\langle F_e, m| \Big), \\ \hat{J}_{01} = |F_g, F_g\rangle_x\langle F_g, -F_g|\,, \\ \hat{J}_{10} = |F_g, -F_g\rangle_x\langle F_g, F_g|\,, \end{cases} \tag{4.11}$$

which satisfy $\hat{J}_{11} = \hat{I} - \hat{J}_{00}$ ($\hat{I}$ is the identity operator) and $\hat{J}_{10} = \hat{J}_{01}^{\dagger}$. Also, the convention $\langle p|_c\, \hat{J}_{kl}\, |q\rangle_c = \delta_{pk}\delta_{ql}$ is followed. The coefficient c_m above satisfies the following recurrence relation:

$$\frac{c_{m+1} - c_m}{c_m - c_{m-1}} = \frac{(F_g - m + 1)(F_g - m)}{(F_g + m + 1)(F_g + m)}, \tag{4.12}$$

which can be justified by applying $\mathcal{L}_{\mathrm{sp}}^{\ddagger}(\hat{J}_{00}) = 0$ constraint in Eq. (4.10). With this, c_m can be solved as

$$c_m = \frac{1 + \sum_{m_1=-F_g+1}^{m-1} \Big[\prod_{m_2=-F_g+1}^{m_1} \frac{(F_g - m_2 + 1)(F_g - m_2)}{(F_g + m_2 + 1)(F_g + m_2)} \Big]}{{}_2F_1\big(-2F_g, 1 - 2F_g; 2; 1\big)}. \tag{4.13}$$

The denominator here ${}_2F_1(a, b; c; z)$ is a hypergeometric function, and in our case

$${}_2F_1\big(-2F_g, 1 - 2F_g; 2; 1\big) = \frac{(4F_g)!}{(2F_g + 1)!(2F_g)!} \simeq \frac{16^{F_g}}{F_g^{3/2}\sqrt{8\pi}} \qquad (F_g \gg 1), \tag{4.14}$$

where the approximation is achieved from Stirling's formula $n! \simeq \sqrt{2\pi n}(\frac{n}{e})^n$ under the large n limit.

Then the question comes: how should we understand those conserved quantities? Suppose initially we have a state in $|0\rangle_c = |F_g, F_g\rangle_x$ and then it hops into $|F_g, F_g - 1\rangle_x$ due to the perturbation from the noise $\hat{F}_{g,z}$. The conserved quantities suggest that, after evolving a sufficient long time under $\mathcal{L}_{\text{sp}}$, the population of the resulting state in $|1\rangle_c$ is $\rho_{c,11} = \text{Tr}[\hat{J}_{11}^\dagger |F_g, F_g - 1\rangle_x \langle F_g, F_g - 1|] = c_{-F_g+1} \simeq \sqrt{8\pi} F_g^{3/2} e^{-(4\ln 2)F_g}$, which is suppressed exponentially with F_g. This is similar to the bosonic cat encoding, where the bit-flip rate is exponentially suppressed with the mean photon number.

On the other hand, the coherence between $|0\rangle_c$ and $|1\rangle_c$ can be lost if the state leaves the encoded subspace due to the noise and then comes back under the stabilization. This is because $\hat{J}_{01}$ and $\hat{J}_{10}$ have no overlap with any state outside the encoded subspace. Therefore, the physical noise will be converted into a phase-flip (Z-type) error. In Sec. 4.3.4, I will show how the coherence is preserved under the engineered dissipation if the physical noise itself is a dephasing $\hat{F}_{g,z}$.

4.3.3 Process tomography of the error channel

With the conserved quantities derived above, we can further achieve the effective error rate on the encoded qubit due to the existence of the physical noise. I will use the language of quantum channels to explain this.

Consider a system evolving under the full Lindbladian $\mathcal{L}_{\text{sp}} + \kappa\mathcal{L}_E$, where $\mathcal{L}_E$ describes the noise in the system and κ is its strength. The evolution channel with time t is $\mathcal{E}_{\text{evo}} = e^{(\mathcal{L}_{\text{sp}}+\kappa\mathcal{L}_E)t}$. In the later derivation, we always assume the noise is a perturbative effect, which means κ is much smaller than any energy scales specified from $\mathcal{L}_{\text{sp}}$. Moreover, there always exists a recovery channel $\mathcal{P}_{\text{sp}}^{\infty} = \lim_{t\to+\infty} e^{\mathcal{L}_{\text{sp}}t}$ which maps arbitrary states into the steady state subspace. Also, the outcomes of this channel can be determined from the

conserved quantities, where

$$\mathcal{P}_{\mathrm{sp}}^{\infty}(\hat{\rho}) = \sum_{k,l\in\{0,1\}} |k\rangle_c \langle l| \cdot \mathrm{Tr}[\hat{J}_{kl}^{\dagger}\hat{\rho}]. \tag{4.15}$$

The auxiliary recovery channel allows us to consider the process tomography of the composite channel $\mathcal{P}_{\mathrm{sp}}^{\infty} \circ \mathcal{E}_{\mathrm{evo}}$ on the encoded subspace

$$\mathcal{P}_{\mathrm{sp}}^{\infty} \circ \mathcal{E}_{\mathrm{evo}}(\hat{\rho}_c) = \sum_{k,l} r_{kl} \hat{W}_k \hat{\rho}_c \hat{W}_l^{\dagger}, \tag{4.16}$$

where $\hat{W}_{k(l)} \in \{\hat{I}_c, \hat{X}_c, \hat{Y}_c, \hat{Z}_c\}$ are Pauli operators for the encoded qubit, and r_{kl} is the element of the process matrix. $\hat{\rho}_c$ here is an arbitrary state in the encoded subspace and therefore $\mathcal{L}_{\mathrm{sp}}(\hat{\rho}_c) = 0$.

Given the noise is perturbative, we can further compute the channel above as

$$\begin{aligned} \mathcal{P}_{\mathrm{sp}}^{\infty} \circ \mathcal{E}_{\mathrm{evo}}(\hat{\rho}_c) &= \mathcal{P}_{\mathrm{sp}}^{\infty} \circ \left[e^{\mathcal{L}_{\mathrm{sp}} t} + \kappa \int_0^t e^{\mathcal{L}_{\mathrm{sp}}(t-s)} \mathcal{L}_E e^{\mathcal{L}_{\mathrm{sp}} s} \, \mathrm{d}s + O(\kappa^2) \right](\hat{\rho}_c) \\ &= \hat{\rho}_c + \kappa t \cdot \mathcal{P}_{\mathrm{sp}}^{\infty} \mathcal{L}_E(\hat{\rho}_c) + O(\kappa^2). \end{aligned} \tag{4.17}$$

We can use this formula to derive the effective error rate given the form of the physical noise. For example, when $\mathcal{L}_E = \mathcal{D}[\hat{F}_{g,z}]$, we have

$$\begin{aligned} \dot{r}_{XX} = \dot{r}_{YY} &= \kappa \frac{F_g}{4} c_{-F_g+1} \simeq \kappa \sqrt{\pi/2} F_g^{5/2} e^{-(4\ln 2) F_g}, \\ \dot{r}_{ZZ} &= \kappa \frac{F_g}{4} (1 - c_{-F_g+1}) \simeq \kappa F_g / 4. \end{aligned} \tag{4.18}$$

Similarly, when $\mathcal{L}_E = \mathcal{D}[\hat{F}_{g,x}]$, we have

$$\dot{r}_{XX} = \dot{r}_{YY} = 0, \quad \dot{r}_{ZZ} = \kappa F_g^2. \tag{4.19}$$

Both of the examples illustrate the exponential bit-flip suppression in the dark spin-cat

encoding.

It is worth mentioning that, in practice, there could be F-dependence in κ. Here we consider the noise coming from the random magnetic $\vec{B}$ field fluctuation. The external $\vec{B}$ field couples with the hyperfine manifold through [153]

$$\hat{H}_{B-F} = (g_J \mu_B \vec{J} + g_I \mu_N \vec{I}) \cdot \vec{B}. \tag{4.20}$$

Moreover, we have $\mu_N \ll \mu_B$, which means usually the contribution from the nuclear term is ignorable. If the total electronic angular momentum $J = 0$ (therefore $F = I$), then despite there being no F-dependence in κ, the strength of the noise itself could be sufficiently small. On the other hand, if J itself is nonzero but we have $J \ll I$ to get a long hyperfine spin, then $\hat{H}_{B-F}$ is effectively $\hat{H}_{B-F} \simeq g_F \mu_B \vec{F} \cdot \vec{B}$, where

$$g_F = \frac{F(F+1) + J(J+1) - I(I+1)}{2F(F+1)} g_J. \tag{4.21}$$

If we work on the manifold $F = J + I$, then we have $g_F = J g_J / F$. Therefore, given fixed J, the isotope with longer I will result in the relevant hyperfine manifold with $g_F \sim 1/F$. Since $\kappa \propto g_F^2$ in the incoherent noise model $\kappa \mathcal{D}[\hat{F}_{g,\xi}]$, here κ depends on F through $\kappa \sim 1/F^2$ in this case.

4.3.4 Autonomous dephasing correction with engineered dissipation

Now I will explain how the physical dephasing noise $\hat{F}_{g,z}$ can be autonomously corrected with the engineered dissipation $\mathcal{L}_{\text{eng}}$ in Eq. (4.5). The bit-flip suppression can be argued from the separation of the encoded computational basis on the GBS again, and here I will focus more on the preservation of the coherence between them. Notice that both the Hamiltonian $\hat{H}_{\text{DS},x}$ and the jump operators $\hat{D}_{ge,q}$ ($q = \pm 1$) in $\mathcal{L}_{\text{eng}}$ commute with the parity operator $\hat{\Pi}$, which means the parity now is a strong symmetry [154]. If a state is initially in the

Hilbert subspace with a given parity, then it will still be in this subspace when reaching the steady state under dissipation. Therefore, consider the state $|\pm\rangle_c$ which are eigenstates of $\hat{\Pi}$ (with eigenvalue ± 1). The dephasing noise $\hat{F}_{g,z}$ maps the states into superpositions between $|F_g, F_g - 1\rangle_x$ and $|F_g, -F_g + 1\rangle_x$ but does not modify their parities. As a result, the state will still come back into $|\pm\rangle_c$ under the autonomous stabilization, which indicates that the coherence between computational basis is preserved.

A more rigorous justification again requires the derivation of the conserved quantities. For the Lindbladian $\mathcal{L}_{\text{eng}}$, we have not found the exact expression for all 4 conserved quantities yet. However, in the strong driving regime $\Omega \gg \Gamma$, every dimensionless quantity can be written as a Taylor series expansion with Γ/Ω. We can still explicitly write down the conserved quantities up to the 0-th order (up to $O[(\Gamma/\Omega)^0]$) as follows:

$$\begin{cases} \hat{J}'_{00} = |F_g, F_g\rangle_x\langle F_g, F_g| + \sum_{m=-F_g+1}^{F_g-1} c_m \Big(|F_g, m\rangle_x\langle F_g, m| + |F_e, m\rangle_x\langle F_e, m| \Big) + O(\Gamma/\Omega), \\ \hat{J}'_{11} = \hat{I} - \hat{J}'_{00}, \\ \hat{J}'_{01} = |F_g, F_g\rangle_x\langle F_g, -F_g| + \sum_{m=-F_g+1}^{F_g-1} c_m \Big(|F_g, m\rangle_x\langle F_g, -m| + |F_e, m\rangle_x\langle F_e, -m| \Big) \\ \qquad + O(\Gamma/\Omega), \\ \hat{J}'_{10} = \hat{J}'^{\dagger}_{01} = \hat{\Pi} - \hat{J}'_{01}, \end{cases} \tag{4.22}$$

where the coefficients c_m are same as that in Eq. (4.13).

Similar to the analysis in Sec. 4.3.3, here we can also achieve the effective error rate when the physical noise is $\mathcal{L}_E = \mathcal{D}[\hat{F}_{g,z}]$, using the process tomography method. With the conserved quantities in Eq. (4.22), we can get

$$\dot{r}_{XX} \simeq \kappa \frac{F_g}{2} c_{-F_g+1} \simeq \kappa\sqrt{2\pi} F_g^{5/2} e^{-(4\ln 2)F_g}, \quad \dot{r}_{YY} = \dot{r}_{ZZ} = 0. \tag{4.23}$$

In the $F_g \gg 1$ regime, this confirms the autonomous error correction against physical dephasing noise $\hat{F}_{g,z}$ with the engineered dissipation $\mathcal{L}_{\text{eng}}$.

We also notice that a similar idea has also been experimentally demonstrated in ions recently, where the physical dephasing $\hat{F}_{g,z}$ errors are autonomously corrected using the spin-cat encoding [155, 156]. Here, people use phonon modes to extract the entropy and therefore stabilize the code subspace, instead of the photons used in our work.

4.4 Bias-preserving operations with erasure conversion

As we have already seen the biased noise structure for the dark spin-cat encoding on the memory level, in this section, I will discuss the gate implementation protocols that preserve such error bias during execution. Similar to the bosonic cat and pair-cat examples shown in Chapter 3, for single-qubit control, the set of operations includes the state initialization and measurement in X basis, as well as the X and $Z(\theta)$ operations.

For entangling gates like CZ and CX, however, in the neutral atom platform the implementation usually requires the assistance of the Rydberg interaction, which makes the Rydberg decay a major source of noise. The first proposal for BP CX implementation in neutral atoms was made in Ref. [81], where an ancilla atom is required to avoid the first-order bit-flip during the gate. Meanwhile, the erasure conversion scheme in accompany with metastable encoding is introduced for CZ execution, where the location of the error is determined through the detection of which atom suffers from the Rydberg decay. Further, a fresh atom to substitute the one experienced decay can be initialized in a specific state to avoid a bit-flip in the whole error channel. Such an error is called "biased erasure" and will lead to a high threshold with a suitable choice of error correcting codes, even without a native BP CX gate.

One caveat in the biased erasure CZ implementation is the requirement for a high success probability to catch the decay events; otherwise the threshold may drop quickly [84] as the

erasure conversion rate goes down. In practice, the Rydberg states may also decay into the encoded subspace, which cannot be converted into erasure and may create a bit-flip. This leads to an unfavorable error structure in the system. However, with our spin-cat encoding, even if the population decays into the encoded hyperfine manifold during CZ execution, the bit-flip error is still unlikely to happen due to the separation of the encoded basis state in the magnetic quantum number. Therefore, our CZ design with spin-cat offers a more structured error pattern than those in previous proposals.

Moreover, we also provide a CX design with the biased erasure error structure, based on the spin-cat encoding. As I will show below, our protocol requires a long-time occupation in the Rydberg state and continuous monitoring of the Rydberg decay for the control atom, which may result in a low gate fidelity and difficulties in the control sequences. However, we still believe that our proposal provides an affirmative answer to the theoretical achievability of the biased erasure CX gate in neutral atoms and motivates people to further optimize the design.

4.4.1 Single-qubit bias-preserving operations

To start with, we first consider the implementation of BP operations on a single qubit. The X-basis state preparation and measurement will be BP by definition, as the bit-flip (X) error will trivially act on the prepared states or commute with the measured operator, which leaves no effect on them. Both X and $Z(\theta)$ can be achieved from simple SU(2) rotations on the GBS [85], and therefore we only need to drive the magnetic dipole ($\hat{F}_{g,\xi}$). On the other hand, if such operations are slow in practice due to the weakness of magnetic dipole, we may also consider the adiabatic control over the stabilization drives Ω_q to move the two SCSs on the GBS. As I will show below, we may either swap their places for an X gate, or let them enclose an area to pick up a geometric phase for the $Z(\theta)$ gate.

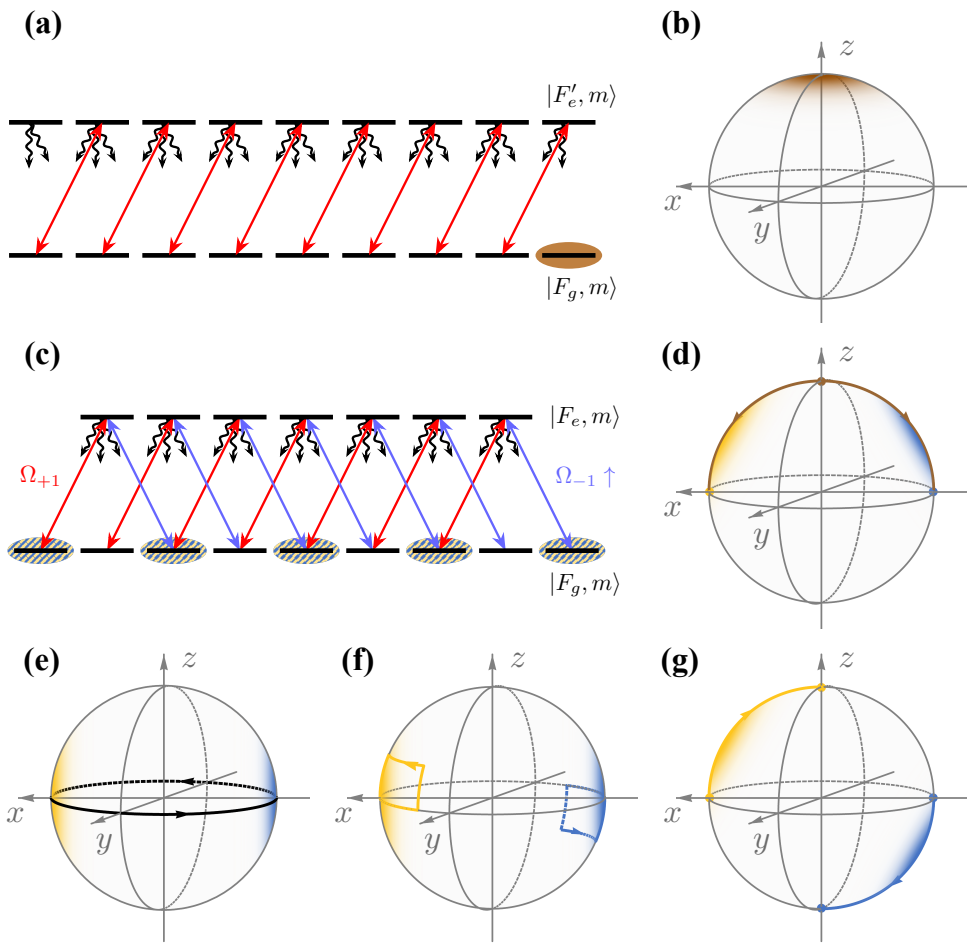

Figure 4.4: Single-qubit bias-preserving operations. (a-d) X-basis state initialization. (a) The state is first initialized in $|F_g, F_g\rangle$ with the GBS representation in (b). (c) Then we gradually increase the amplitude of Ω_{-1} drives, which leads to the trajectories of the SCSs shown in (d). (e-g) The SCS trajectories for (e) X gate, (f) $Z(\theta)$ gate, and (g) coherent state transfer from $|F_g, \pm F_g\rangle_x$ to $|F_g, \pm F_g\rangle$.

X-basis state preparation and measurement

We first recall how we initialize the $|+\rangle_c$ state for the bosonic cat. There, we start with a vacuum state first, and then turn on the dissipative stabilization so that the state will end up in the steady state subspace spanned by $|\pm\alpha\rangle$. Meanwhile, since the parity is preserved during this process, the resulting state should be the even-parity cat $|+\rangle_c \propto (|\alpha\rangle + |-\alpha\rangle)$.

We can do similar things for the dark spin-cat here. First, we will try to initialize the state into $|F_g, -F_g\rangle$ by coupling the ground manifold with an excited hyperfine manifold $F_e' = F_g$

using the Ω_{-1} drive [157]. Under this configuration, clearly $|F_g, -F_g\rangle$ is the only dark state and also a parity eigenstate as well. After this, we will consider another excited manifold with $F_e = F_g - 1$ and turn on the Ω_{-1} drive. $|F_g, -F_g\rangle$ is still one of the dark states up to now. Finally, we adiabatically increase the amplitude of Ω_{+1} drive until $\Omega_{+1} = -\Omega_{-1} = \frac{\Omega}{\sqrt{2}}$. The resulting state should be in the dark state subspace spanned by $|F_g, \pm F_g\rangle_x$, while the parity conservation given the engineered dissipation $\mathcal{L}_{\text{eng}}$ guarantees that output state should be $|+\rangle_c$ if there is no extra noise. Indeed, the parity-breaking decay $\hat{D}_{ge,0}$ in $\mathcal{L}_{\text{sp}}$ may cause phase-flip error on the prepared state, since there is still population on the excited levels during evolution due to the non-adiabaticity. However, such an error will belong to the dephasing (Z) type by default and should be corrected with the next-level QEC codes.

The X-basis measurement can be viewed as the inverse process of the X-basis initialization. With $\Omega_{+1} = -\Omega_{-1}$ in the beginning, the stabilized $|\pm\rangle_c$ states are in the even or odd parity subspace, respectively. We can gradually turn off the Ω_{+1} drive so that the population in the even-parity ($+1$) subspace will come to the $|F_g, -F_g\rangle$ state while that in the odd-parity (-1) subspace will come to $|F_g, 1 - F_g\rangle$. We can further address $|F_g, 1 - F_g\rangle$ without affecting $|F_g, -F_g\rangle$ through a π-polarized light ($\Delta m = 0$) that couples a hyperfine manifold with $F_e = F_g - 1$, and then follow-up controls (like ionization [158]) are allowed to readout its population.

X gate

The X gate requires a coherent exchange between $|F_g, F_g\rangle_x$ and $|F_g, -F_g\rangle_x$. To prevent the bit-flip during gate execution, we expect the two computational basis states to stay largest separated on the GBS. One solution is to implement a Hamiltonian rotation $\hat{H} \propto \hat{F}_{g,z}$, such that

$$e^{-i\phi_Z(t)\hat{F}_{g,z}}(a_0 |0\rangle_c + a_1 |1\rangle_c) = a_0 |\pi/2, \phi_Z(t)\rangle_g + a_1 e^{-i\pi F_g} |\pi/2, \pi + \phi_Z(t)\rangle_g , \qquad (4.24)$$

where $\phi_Z(t) = \pi t/T$. At time $t = T$, the state above will become $e^{i\pi F_g}(a_0 |1\rangle_c + a_1 |0\rangle_c)$, which indicates that we indeed execute an X gate. Meanwhile, the dissipative stabilization drives should be adjusted accordingly so that the subspace spanned by the two SCSs $\{|\pi/2, \phi_Z(t)\rangle_g, |\pi/2, \pi + \phi_Z(t)\rangle_g\}$ is still the instantaneous dark subspace. To do so, we can adjust the phase of $\Omega_{\pm 1}$ drives as

$$\Omega_{+1}(t) = \frac{\Omega e^{i\phi_Z(t)}}{\sqrt{2}}, \quad \Omega_{-1}(t) = -\frac{\Omega e^{-i\phi_Z(t)}}{\sqrt{2}}, \tag{4.25}$$

so that the total electric field for stabilization is still linearly polarized with the direction rotating in the x-y plane.

If the Hamiltonian rotation term is not available or too weak in practice, we may still achieve the desired operation by adiabatically adjusting the phase of $\Omega_{\pm 1}$ drives as described above. However, imperfections will come now due to the non-adiabaticity.

$Z(\alpha)$ gate

Like the X gate implementation, a straightforward way to implement $Z(\alpha) = e^{-i\frac{\alpha}{2}\hat{Z}_c}$ rotation gate is through a Hamiltonian $\hat{H} \propto \hat{F}_{g,x}$:

$$e^{-i\phi_X(t)\hat{F}_{g,x}}(a_0 |0\rangle_c + a_1 |1\rangle_c) = a_0 e^{-i\phi_X(t)F_g} |0\rangle_c + a_1 e^{i\phi_X(t)F_g} |1\rangle_c, \tag{4.26}$$

where $\phi_X(t) = \frac{\alpha}{2F_g}\frac{t}{T}$. In this case, the state does not need to leave the encoded subspace, so there is no need to adjust the stabilization drives.

Alternatively, we can also adiabatically adjust the stabilization drives to implement the $Z(\alpha)$ gate. In this case, we can gradually move the location of the two dark SCSs on the GBS while keeping them antipodal. When their trajectories form two loops after returning to the starting points, the two SCSs will pick up a relative geometric phase. This is equivalent to the action of the desired $Z(\alpha)$ operation. I will defer the detailed derivation to the

Appendix C.2.1.

Coherent state transfer between $|F_g, \pm F_g\rangle_x$ and $|F_g, \pm F_g\rangle$

Before the discussion on the entangling gate, I will also talk about the coherent state transfer from two SCSs $|F_g, \pm F_g\rangle_x$ to another two $|F_g, \pm F_g\rangle$ and vice versa, while keeping the two SCSs largest separated during evolution. Specifically, we want to realize the map between the arbitrary superposition $a_0 |0\rangle_c + a_1 |1\rangle_c = a_0 |F_g, F_g\rangle_x + a_1 |F_g, -F_g\rangle_x$ and $a_0 |F_g, F_g\rangle + a_1 |F_g, -F_g\rangle$. This is not a gate on the encoded subspace; however, in the presence of a constant B_z field, the operation is needed during the entangling operations. As we shall see, this is because that we need to address the population in one SCS into the Rydberg levels during those gates, and it will be much easier to do that in the $|F_g, \pm F_g\rangle$ basis rather than $|F_g, \pm F_g\rangle_x$ due to the frequency selectivity caused by the constant B_z field. I also want to note that this process can be used for Z-basis initialization and measurement, which is different from the X-basis counterpart that we map between $|\pm\rangle_c$ and $\{|F_g, -F_g\rangle, |F_g, 1 - F_g\rangle\}$.

From the definition $|F_g, m\rangle_x = e^{-i\frac{\pi}{2}\hat{F}_{g,y}} |F_g, m\rangle$, we can see that the desired operation can also be achieved with the SU(2) rotation under Hamiltonian $\hat{H} \propto \hat{F}_{g,y}$:

$$e^{-i\phi_Y(t)\hat{F}_{g,y}}(a_0 |0\rangle_c + a_1 |1\rangle_c) = a_0 e^{-i[\frac{\pi}{2}+\phi_Y(t)]\hat{F}_{g,y}} |F_g, F_g\rangle + a_1 e^{-i[\frac{\pi}{2}+\phi_Y(t)]\hat{F}_{g,y}} |F_g, -F_g\rangle , \tag{4.27}$$

where $\phi_Y(t) = -\frac{\pi}{2}\frac{t}{T}$. Regarding the stabilization drives, we need to decrease the amplitude of $\Omega_{\pm 1}$ while increasing Ω_0 accordingly:

$$\Omega_{+1}(t) = -\Omega_{-1}(t) = \frac{\Omega}{\sqrt{2}} \cos[\phi_Y(t)], \quad \Omega_0(t) = \Omega \sin[\phi_Y(t)]. \tag{4.28}$$

In this way, we realize the map from $|F_g, \pm F_g\rangle_x$ to $|F_g, \pm F_g\rangle$. We can also reverse the control sequences for the map from $|F_g, \pm F_g\rangle$ to $|F_g, \pm F_g\rangle_x$.

Similar to the previous cases, the operation can also be implemented by adiabatically adjusting these $\Omega_q(t)$ drives only, without using the Hamiltonian rotation term. In this way, the non-adiabatic error will come again.

4.4.2 Biased erasure CZ gate

Now I will discuss the BP constructions for the entangling operations, and the first one will be the CZ gate. Our design is similar to other Rydberg CZ protocols in existing literatures (see Ref. [159]), where the same pulses are applied to both the control and the target atoms that only one of the computational state (say, $|1\rangle_c$) will couple to the desired Rydberg level $|r\rangle_c$. Due to the Rydberg blockade effect, $|01\rangle_c$ ($|10\rangle_c$) will couple to $|0r\rangle_c$ ($|r0\rangle_c$) with the amplitude Ω^R_{+1}, while $|11\rangle_c$ will couple to $\frac{|1r\rangle_c+|r1\rangle_c}{\sqrt{2}}$ with the amplitude $\sqrt{2}\Omega^R_{+1}$. The difference in the coupling amplitudes enables that, with a careful design of the amplitude and phase in $\Omega^R_{+1}(t)$ as well as the Rydberg detuning Δ_R, the population from $|01\rangle_c$ ($|10\rangle_c$) and $|11\rangle_c$ can simultaneously come back to the encoded subspace, while picking up different phases $e^{i\phi_{01}}$ ($e^{i\phi_{10}} = e^{i\phi_{01}}$) and $e^{i\phi_{11}}$, respectively. If $\phi_{11} - (\phi_{10} + \phi_{01}) = \pi$, then this operation will be equivalent to the CZ gate up to local Z rotation.

When coming to the spin-cat encoding, we can use the polarization selectivity from the driving lasers to couple only one computational state with the Rydberg level. To do so, we may first coherently transfer the state from $|F_g, \pm F_g\rangle_x$ to $|F_g, \pm F_g\rangle$ in both the control and target atoms.[1] As shown in Fig. 4.5, we can find a Rydberg manifold with total angular momentum $F_R = F_g - 1$ and use the σ_+-polarized light to couple the encoded levels with it. Clearly, $|F_g, -F_g\rangle$ couples with the Rydberg level while $|F_g, F_g\rangle$ does not. Therefore, we can perform the CZ entangling operation described above and finally transfer the state back to the $\{|0\rangle_c, |1\rangle_c\}$ subspace.

This CZ design can also be adapted into the biased erasure framework. For example, if

1. As discussed, this step is unnecessary when there is no constant B_z field, but later the polarization of the Rydberg drives should be adjusted accordingly.

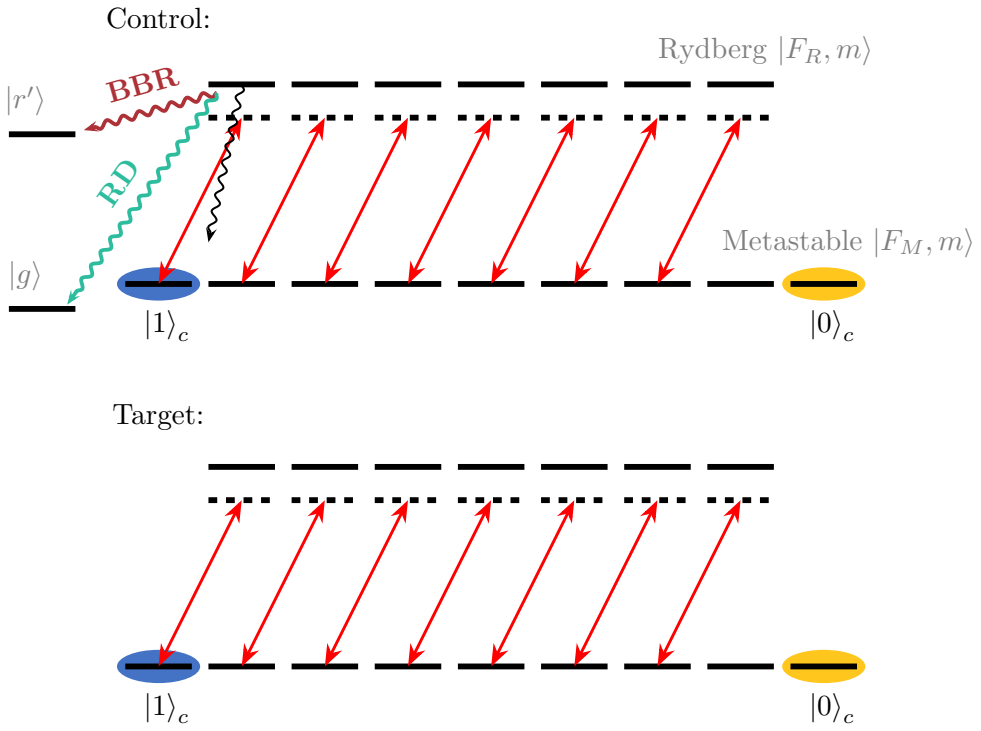

Figure 4.5: Biased erasure CZ implementation with metastable spin-cat encoding. We identify the encoded F_g manifold as the metastable F_M levels to achieve the benefit from erasure conversion. $|1\rangle_c$ state couples with the Rydberg levels while $|0\rangle_c$ does not, because of the selectivity from laser polarization. Even population decays into the F_M manifold and therefore cannot be converted to erasure, it is still unlikely to be a bit-flip due to the large m-separation.

the F_g is chosen to be a metastable hyperfine subspace F_M, then the decay event into the ground levels through the radiative decay (RD) or the near Rydberg levels through blackbody radiation (BBR) may also be caught (see Fig. 4.5 and Ref. [82]). If the Rydberg event is detected in a certain atom, we will know that the population initially comes from $|1\rangle_c$ state, and then prepare a fresh atom in $|1\rangle_c$ state for substitution. In this way, we will know both the location (erasure) and the Pauli type (Z bias) of the error. If the erasure conversion rate is sufficiently large, we can achieve an error threshold approaching 10% with XZZX surface code, even without the aid of the BP CX gate [83].

Moreover, even if the population decays into the encoded hyperfine subspace, the separation of the magnetic quantum number between $|F_M, \pm F_M\rangle$ still makes it difficult to produce

a bit-flip. This is because one decay via dipole coupling will only cause $|\Delta m| \leq 1$, and therefore with high probability the population will come back to $|F_M, -F_M\rangle$ under dissipative stabilization. As a result, in this case, even if the error cannot be converted into erasure, it is still most likely to be a Z error. The error pattern here will be more structured than those in the existing protocols, where a bit-flip may happen if the population decays into the encoded subspace and cannot be detected.

4.4.3 Biased erasure CX gate with decay monitoring

Finally, I would like to discuss a possible implementation for the biased erasure CX gate, which is based on a metastable encoding again with a further assumption that the Rydberg decay into the ground levels can be continuously monitored. The basic idea is that, we want to implement an X gate on the target atom through an effective Hamiltonian rotation $\hat{H} \propto \hat{F}_{g,z}$ (so that the two encoded SCSs are kept to be largest separated during evolution) only when the control atom is in $|1\rangle_c$. In our scheme, the selectivity to the state of control atom is achieved from the Rydberg blockade, while the effective Hamiltonian rotation for the target atom comes from a designed differential AC Stark shift. The detailed protocol is explained below.

Similar to the CZ design, here we also consider F_g to be a metastable hyperfine manifold F_M, and coherently map the state in the control atom from $|F_M, \pm F_M\rangle_x$ to $|F_M, \pm F_M\rangle$ first in order to address the population from the $|0\rangle_c$ state to the Rydberg later. We then do that through a resonant σ_--polarized pulse that couples F_M to a Rydberg manifold $F_R = F_M - 1$. With this, the population in $|F_M, F_M\rangle$ can be fully transferred into the Rydberg state while $|F_M, -F_M\rangle$ is untouched due to the polarization selectivity. After that, we move to the target atom and turn off the stabilization drives. Instead, we implement both σ_+- and σ_--polarized light that couples F_M to a Rydberg $F_R = F_M - 1$ with the same amplitude $\Omega^R_{T,+1} = \Omega^R_{T,-1} = \Omega^R$, but the frequencies of the two drives ($\omega_{RM} \pm \Delta^R$) are

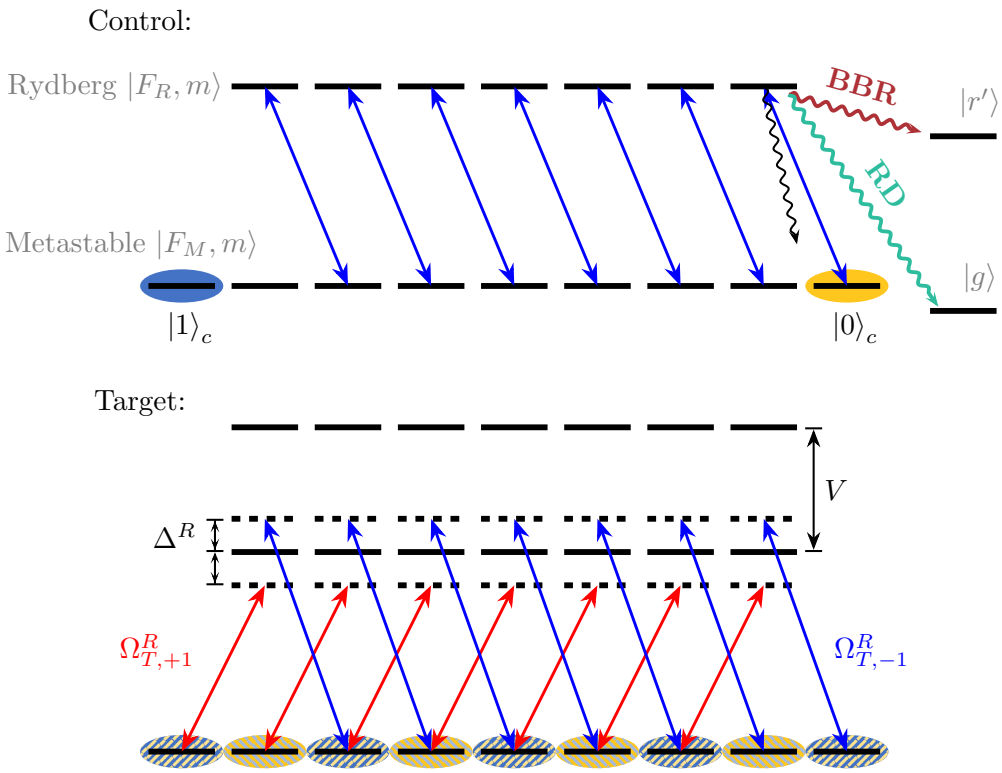

Figure 4.6: Biased erasure CX implementation with metastable spin-cat encoding and decay monitoring. On the control side, $|0\rangle_c$ rather than $|1\rangle_c$ will be addressed to the Rydberg state due to the polarization selectivity. The conditional X gate on the target atom is implemented through a differential AC Stark shift mediated with the Rydberg coupling. The Rydberg decay in the control atom to the ground state $|g\rangle$ should be frequently monitored to avoid the undesired rotation on the target.

detuned from the Rydberg level in a different direction (see Fig. 4.6). If the control atom is initially in $|1\rangle_c$ and therefore it is not addressed to the Rydberg later, then we can write the lab-frame Hamiltonian on the target atom as

$$\hat{H}_{T,\text{lab}} = \omega_{Rg}\hat{P}_R + \frac{1}{2}\sum_{q=\pm1}[e^{-i(\omega_{RM}+q\Delta^R)t}\Omega^R_{T,q}\hat{D}_{RM,q} + \text{h.c.}]. \tag{4.29}$$

Here $\hat{P}_R$ is the projector onto the desired Rydberg hyperfine manifold and $\hat{D}_{RM,q}$ is the dipole coupling between the encoded manifold and the Rydberg manifold with different polarizations. When $\Omega^R \ll \Delta^R$, the $\Omega^R_{T,\pm1}$ drives on the target effectively implement the

$\hat{H} = \mu \hat{F}_{M,z}$ rotation due to the effect of the AC Stark shift, with

$$\mu = \frac{(\Omega^R)^2}{4\Delta^R} \frac{2F_M - 1}{F_M(2F_M + 1)}. \tag{4.30}$$

On the other hand, if the control atom is initially in $|0\rangle_c$ so that it is addressed to the Rydberg level later, then the energy of the Rydberg states used in the target atom will be shifted away due to the Rydberg blockade. In this case, both $\Omega^R_{T,\pm 1}$ drives on the target atom will be far-off-resonant, and therefore the states in the F_M subspace will not rotate anymore. This concludes the part of the conditional-X operation. Finally, we can turn on the stabilization drives on the target atom, while on the control side we coherently transfer the population from the Rydberg to $|F_M, F_M\rangle$ and then map $|F_M, \pm F_M\rangle$ subspace back to the original encoded subspace spanned by $|F_M, \pm F_M\rangle_x$.

A major source of noise in this protocol should be the Rydberg decay in the control atom. This is because, first, the time duration of the gate is expected to be long. In order for the derivation of the AC Stark shift to be valid, we need $\Omega^R \ll \Delta^R$. Therefore, $\mu \propto \frac{(\Omega^R)^2}{\Delta^R} \ll \Omega^R$. Moreover, Δ^R itself is also limited by the Rydberg blockade strength $\Delta^R \ll V$. These facts all indicate that the gate time $T = \frac{\pi}{|\mu|}$ will be longer than other characteristic time set by the inverse of any Hamiltonian parameters, like $\frac{1}{\Omega^R}$. Second, if the control is initially in $|0\rangle_c$, then the full population will stay in the Rydberg level during the whole conditional rotation time T. This is in contrast to the case in the target atom, where the Rydberg levels are only virtually populated. These two features both increase the chance of Rydberg decay in the control atom. Further, if it decays into the ground state, then there will be no Rydberg blockade as expected and the state in the target atom starts to rotate along the z-axis[2]. If we do not know the time that it decays to the ground, we cannot estimate the angle that

2. It is not a problem if the state comes to the nearby Rydberg levels via the black-body radiation, since in this case the Rydberg blockade still applies on the target atom. This is similar to the χ-matching condition used for transmon-cavity (cat) BP CX implementation [61], where the speed of cat rotation will not be affected if transmon decays from $|f\rangle$ to $|e\rangle$.

state in target atom rotates. It will cause bit-flip error on the target atom, which breaks the BP argument.

Similar to the BP CX design in the pair-cat example, here the problem may also be solved with the assumption that we can frequently measure the population on the ground levels to monitor the undesired decay event. If such decay is detected, we can immediately stop the $\Omega_{\pm 1}^{R,T}$ drives on the target atom in order to avoid the extra unwanted rotation. If such checks can be frequent enough, the uncertainty for the extra rotation angle will be limited, and therefore, the bit-flip rate is still bounded. Finally, we can prepare a fresh atom in $|0\rangle_c$ state to replace the control atom in which the Rydberg decay is detected. With the same argument as mentioned in the CZ design, here we also know both the location (erasure) and the Pauli type (dominantly Z error in the control atom) of the error. A detailed error analysis based on this proposal is discussed in Appendix C.2.2. However, if the Rydberg state decays into the encoded metastable manifold and therefore cannot be detected, it may cause bit-flip error in the target atom. As discussed in the early proposal [82, 83], we expect the probability of such a decay event to be small due to a possible minor branching ratio of the related transition.

Indeed, I believe that the current design contains several drawbacks, including the complexity in the control sequence and the limited fidelity caused by the decay in the control atom. Besides, although there is sufficient progress to realize the repetitive readout in the neutral atom platform [160], the continuous monitoring of the Rydberg decay into the ground levels is also believed as a difficult task. I would reiterate that the aim of the proposal is to show the theoretical feasibility of the biased erasure CX implementation, which has not been discussed previously.

It also raises the question about the tradeoff between the improved structuredness of the error pattern and a higher error rate for the operations. In fact, the answer should also depend on the choice of the upper-level QEC codes, which can be further optimized based

on the error structures of the physical qubits to achieve a higher threshold or a reduced qubit overhead [140]. One example in favor of the bias is discovered in Ref. [145], which is based on the Kerr-cat qubits encoded in the XZZX surface code. Given the same physical parameters in the system, the CX operation implemented in the BP manner incurs a moderately higher infidelity compared with that in a non-BP construction (achieved by a CZ gate conjugated with single-qubit rotations on the target qubit). However, the achieved logical error rate can still be lower with the BP CX design, due to a higher error threshold observed in this case. In general, I still believe it is an interesting open problem, and an overall better performance should come as the product of the collaborations between both device physicists and the information theorists, with the improvement from both physical qubits themselves and the QEC code design tailored to the given error patterns.

CHAPTER 5

GENERALIZATION OF DARK SPIN-CAT ENCODING WITH PERFORMANCE ANALYSIS

5.1 Introduction

In the previous chapter, I discussed the dark spin-cat encoding in a large hyperfine manifold of an atom, where the largest separation of the two encoded computational basis states on the GBS provides the exponential bit-flip suppression of the spin-cat qubit on the memory level. The autonomous stabilization of the code subspace and the design of BP operations with erasure conversion are also explained. Despite operating on different platforms, these features all illustrate a great similarity between the dark spin-cat and the cat in a bosonic mode, while during the past decade, the latter has drawn significant attention in the field of continuous-variable QEC from both the theoretical and experimental aspects.

On the other hand, as mentioned in Appendix C.1, the two dark SCSs are not necessarily antipodal with each other. In fact, they can be everywhere on the GBS under the same physical setup, which can be realized by simply adjusting the stabilization drives. Such a generalization offers more flexibility in the choice of the encoded basis that shares the cat-like features, and enables people to analyze them within a much broader framework. In this chapter, I will discuss some physical properties of the generalized spin-cat qubit with a focus on the bit-flip analysis. Several schemes used on the bosonic cat to suppress the bit-flip can also be migrated into here after suitable adaptations. I expect those analogies could not only provide a deeper understanding of the generalized spin-cat model itself but also build a bridge between the bosonic QEC community and the atomic physics community, and therefore motivate people to discover more hardware-efficient schemes in the atomic structure.

The chapter is organized in the following way. In Sec. 5.2, I will specify the general-

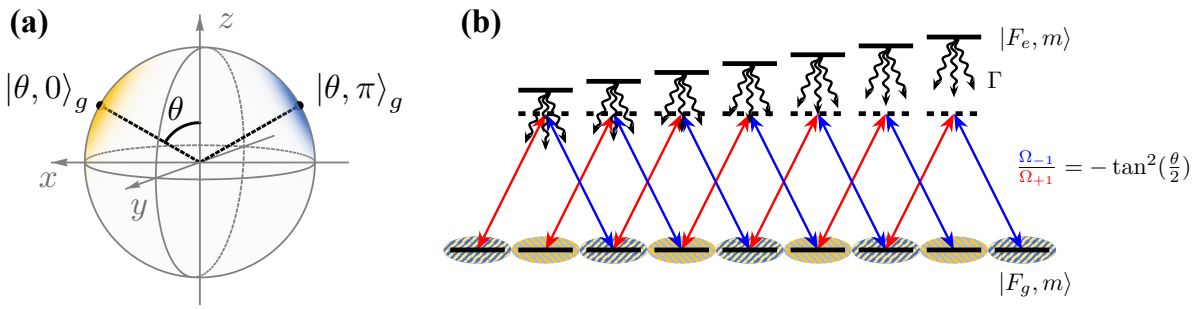

Figure 5.1: Schematic plot on the generalized dark spin-cat configuration. (a) The code subspace is spanned by two SCSs $|\theta,0\rangle_g$ and $|\theta,\pi\rangle_g$, which may not be antipodal on the GBS. (b) The ratio between $\Omega_{\pm1}$ drives to stabilize such a code subspace.

ized spin-cat encoding first and then discuss its analogy with the bosonic cat under the Holstein–Primakoff (H-P) transformation. The similarities are shown not only in the code states themselves but also in the physical stabilization mechanisms. After that, I will focus on the Hamiltonian confinement part first, which, similar to the Kerr-cat Hamiltonian, also provides a double-well feature. The key difference is that in the long spin model the phase space becomes a sphere. In Sec. 5.3 I will use a semiclassical approach named Wentzel–Kramers–Brillouin (WKB) approximation to derive the tunneling rate between the two wells when the state leaves the encoded subspace due to the external perturbation. To suppress the tunneling rate, people working on the Kerr-cat try to introduce a detuning term. When operated at some special points, some excited level pairs between two wells also become degenerate, and therefore the tunneling may not happen. In Sec. 5.4, I will discuss how this trick can be adapted into the spin-cat model. Finally, I will also take the dissipative process into account and perform the full bit-flip analysis on the encoded qubit, which is shown in Sec. 5.5.

5.2 Generalized dark spin-cat and its analogy with bosonic cat

5.2.1 Generalized dark spin-cat encoding

Throughout this chapter, we still consider the physical system with a ground hyperfine manifold F_g and an excited hyperfine manifold $F_e = F_g - 1$, which are coupled with each other through laser drives. We can then achieve the stabilization Hamiltonian $\hat{H}_{\mathrm{DS}}$ in Eq. (4.3) in the rotating frame when the frequencies of the driving lasers are suitably chosen. In general, we can find a degenerate subspace of $\hat{H}_{\mathrm{DS}}$ within the F_g-manifold, and therefore the corresponding eigenvalue is zero. This subspace contains at least two dimensions, and as shown in Appendix C.1, it includes two linear-independent SCSs that may not be orthogonal to each other. If we further consider the decay from the F_e-manifold to the F_g-manifold, then this subspace will also be a steady subspace because any state in it has no occupation on the F_e levels.

The places of the two dark SCSs are determined by the stabilization drives Ω_q only. In general, we can fix them to be $|\theta, 0\rangle_g$ and $|\theta, \pi\rangle_g$, whose overlap $|\langle\theta, 0|\theta, \pi\rangle_g|^2 = (\cos\theta)^{4F_g}$ decays exponentially with F_g. The subspace spanned by these two SCSs can be stabilized if we choose

$$\Omega_{+1} = \Omega\cos^2(\theta/2), \quad \Omega_0 = 0, \quad \Omega_{-1} = -\Omega\sin^2(\theta/2). \tag{5.1}$$

This choice of drives guarantees that $\hat{H}_{\mathrm{DS}}$ commutes with the parity operator $\hat{\Pi}$ defined in Eq. (4.7). Therefore, $\hat{\Pi}$ is still a conserved quantity under the engineered dissipation $\mathcal{L}_{\mathrm{eng}}$ in Eq. (4.5), where only the decay operators $\hat{D}_{ge,\pm1}$ that commute with $\hat{\Pi}$ are included.

Now we can fix the encoding of the generalized spin-cat qubit. To follow the previous

convention, we will encode the X-basis states with determined parity, so that

$$\begin{aligned} |+\rangle_c &:= \frac{|\theta,0\rangle_g + e^{-i\pi F_g}\,|\theta,\pi\rangle_g}{\sqrt{\mathcal{N}_{g,+}}}, \\ |-\rangle_c &:= \frac{|\theta,0\rangle_g - e^{-i\pi F_g}\,|\theta,\pi\rangle_g}{\sqrt{\mathcal{N}_{g,-}}}, \end{aligned} \tag{5.2}$$

where $\mathcal{N}_{g,\pm} = 2[1 \pm (-\cos\theta)^{2F_g}]$ is for normalization. With this choice, the two computational basis states will approximately be the SCSs in the large F_g regime:

$$|0\rangle_c = |\theta,0\rangle_g + O[(\cos\theta)^{2F_g}], \quad |1\rangle_c = e^{-i\pi F_g}\,|\theta,\pi\rangle_g + O[(\cos\theta)^{2F_g}]. \tag{5.3}$$

5.2.2 Holstein-Primakoff transformation

In order to draw a deeper connection with the bosonic cat, here I will introduce the H-P transformation, a technique that maps a long spin to a bosonic mode. Pictorially, although the phase space of a spin is a sphere in general, the neighborhood of the north pole will be flatter when we increase the spin length, and eventually it will become a plane in the infinite length limit.

Following the description above, the H-P transformation first maps $|F_g, F_g\rangle$, the spin state at the north pole, to the ground state $|0\rangle$ in the bosonic mode. Further, the spin operators $\hat{F}_{g,\xi}$ will be transformed into the bosonic operators in the $F_g \gg 1$ limit as follows:

$$\hat{F}_{g,-} \to \sqrt{2F_g}\hat{a}^\dagger, \quad \hat{F}_{g,+} \to \sqrt{2F_g}\hat{a}, \quad \hat{F}_{g,z} \to F_g - \hat{a}^\dagger\hat{a}, \tag{5.4}$$

where $\hat{F}_{g,\pm} = \hat{F}_{g,x} \pm i\hat{F}_{g,y}$. As discussed in Ref. [161], this enables the mapping from an SCS

$$|\theta,\phi\rangle_g = e^{-i\phi F_g}[\cos^2(\theta/2)]^{F_g} \exp[\tan(\theta/2)e^{i\phi}\hat{F}_{g,-}]\,|F_g, F_g\rangle\,, \tag{5.5}$$

to a coherent state $|\alpha\rangle$ in the bosonic mode, where

$$\begin{aligned} |\alpha\rangle &= e^{-|\alpha|^2/2} e^{\alpha \hat{a}^\dagger} |0\rangle , \\ &= \lim_{F_g \to +\infty} \left(1 + \frac{|\alpha|^2}{2F_g}\right)^{-F_g} e^{\alpha \hat{a}^\dagger} |0\rangle . \end{aligned} \tag{5.6}$$

The map from Eq. (5.5) to Eq. (5.6) can be shown with the operator transformation in Eq. (5.4) and the substitution $\alpha = \sqrt{2F_g}\tan(\theta/2)e^{i\phi}$. Therefore, to keep α finite when increasing F_g, we need to make sure $\theta \lesssim O(1/\sqrt{F_g})$. Moreover, there will be an extra $e^{-i\phi F_g}$ phase after mapping to the bosonic coherent state.

Regarding the two SCSs used for encoding, we have the following maps under H-P transformation:

$$|\theta, 0\rangle_g \to |\alpha\rangle , \quad |\theta, \pi\rangle_g \to e^{-i\pi F_g} |-\alpha\rangle \qquad [\text{where } \alpha = \sqrt{2F_g}\tan(\theta/2)]. \tag{5.7}$$

This implies that our encoded states $|\pm\rangle_c$ in Eq. (5.2) will map to $|\pm\rangle_c \propto (|\alpha\rangle \pm e^{-2i\pi F_g} |-\alpha\rangle)$. The extra phase $e^{-2i\pi F_g}$ here comes as the difference in the choice of the parity convention, since $\hat{P}_g \hat{\Pi} \hat{P}_g = e^{-i\pi(\hat{F}_{g,z} + F_g \hat{P}_g)} \to e^{-2i\pi F_g} e^{i\pi \hat{a}^\dagger \hat{a}}$ under the H-P transformation.

Indeed, it looks reasonable for the correspondence between the SCS and the bosonic coherent state. However, more surprising analogies can be discovered when looking into the stabilization mechanism. To illustrate this, I will put a large detuning term $\Delta \gg \Omega, \Gamma$ in $\hat{H}_{\text{DS}}$ so that we can adiabatically eliminate the F_e-levels. I will also fix $\delta = 0$ in $\hat{H}_{\text{DS}}$ for simplicity. The master equation now under $\mathcal{L}_{\text{eng}}$ will be

$$\frac{\mathrm{d}\hat{\rho}}{\mathrm{d}t} = -i[\Delta \hat{P}_e, \hat{\rho}] - \frac{i}{2}[(\Omega_{+1} \hat{D}_{eg,+1} + \Omega_{-1} \hat{D}_{eg,-1} + \text{h.c.}), \hat{\rho}] + \Gamma \sum_{q=\pm 1} \mathcal{D}[\hat{D}_{ge,q}]\hat{\rho}, \tag{5.8}$$

where $\Omega_{\pm 1}$ is given in Eq. (5.1). Then we can use the adiabatic elimination procedure

explained in Ref. [162] to get the effective dynamics restricted to the F_g-manifold:

$$\frac{\mathrm{d}\hat{\rho}_{\mathrm{eff}}}{\mathrm{d}t} = -i[\hat{H}_{\mathrm{eff}}, \hat{\rho}_{\mathrm{eff}}] + \sum_{q=\pm 1} \mathcal{D}[\hat{L}_{\mathrm{eff},q}]\hat{\rho}, \tag{5.9}$$

where

$$\begin{cases} \hat{H}_{\mathrm{NH}} = \Delta \hat{P}_e - \dfrac{i\Gamma}{2} \displaystyle\sum_{q=\pm 1} \hat{D}^{\dagger}_{ge,q} \hat{D}_{ge,q} = (\Delta - \dfrac{i\Gamma}{2})\hat{P}_e - \dfrac{F_g^2 - \hat{F}_{e,z}^2}{F_g(2F_g+1)} \simeq (\Delta - \dfrac{i\Gamma}{2})\hat{P}_e, \\ \hat{H}_{\mathrm{eff}} = -\dfrac{1}{8}(\Omega^*_{+1}\hat{D}^{\dagger}_{eg,+1} + \Omega^*_{-1}\hat{D}^{\dagger}_{eg,-1})[\hat{H}^{-1}_{\mathrm{NH}} + (\hat{H}^{-1}_{\mathrm{NH}})^{\dagger}](\Omega_{+1}\hat{D}_{eg,+1} + \Omega_{-1}\hat{D}_{eg,-1}), \\ \hat{L}_{\mathrm{eff},q} = \dfrac{\sqrt{\Gamma}}{2}\hat{D}_{ge,q}\hat{H}^{-1}_{\mathrm{NH}}(\Omega_{+1}\hat{D}_{eg,+1} + \Omega_{-1}\hat{D}_{eg,-1}). \end{cases} \tag{5.10}$$

The approximation made in $\hat{H}_{\mathrm{NH}}$ comes from the $\Delta \gg \Gamma$ assumption and the fact that we only focus on the levels near the north pole ($\hat{F}_{g(e),z} \simeq F_g$). With this, we can simply write $\hat{H}^{-1}_{\mathrm{NH}} \simeq \frac{1}{\Delta - i\Gamma/2}\hat{P}_e$.

Finally, we need to check the H-P transformation of the effective Hamiltonian $\hat{H}_{\mathrm{eff}}$ and jump operators $\hat{L}_{\mathrm{eff},q}$, respectively. Each relevant term can be computed as

$$\begin{cases} \hat{D}_{ge,-1}\hat{P}_e\hat{D}_{eg,-1} = \dfrac{(F_g + \hat{F}_{g,z} - 1)(F_g + \hat{F}_{g,z})}{2F_g(2F_g+1)} \to \dfrac{2F_g - 1}{2F_g + 1} + O(1/F_g), \\ \hat{D}_{ge,-1}\hat{P}_e\hat{D}_{eg,+1} = \dfrac{(\hat{F}_{g,+})^2}{2F_g(2F_g+1)} \to \dfrac{\hat{a}^2}{2F_g+1}, \\ \hat{D}_{ge,+1}\hat{P}_e\hat{D}_{eg,-1} = (\hat{D}_{ge,-1}\hat{P}_e\hat{D}_{eg,+1})^{\dagger} \to \dfrac{\hat{a}^{\dagger 2}}{2F_g+1}, \\ \hat{D}_{ge,+1}\hat{P}_e\hat{D}_{eg,+1} = \dfrac{(F_g - \hat{F}_{g,z} - 1)(F_g - \hat{F}_{g,z})}{2F_g(2F_g+1)} \to \dfrac{(\hat{a}^{\dagger}\hat{a} - 1)\hat{a}^{\dagger}\hat{a}}{2F_g(2F_g+1)}, \end{cases} \tag{5.11}$$

where we have used the related Clebsch-Gordon coefficients $C_q(m) := C^{F_g-1,m+q}_{F_g,m;1,q}$ during

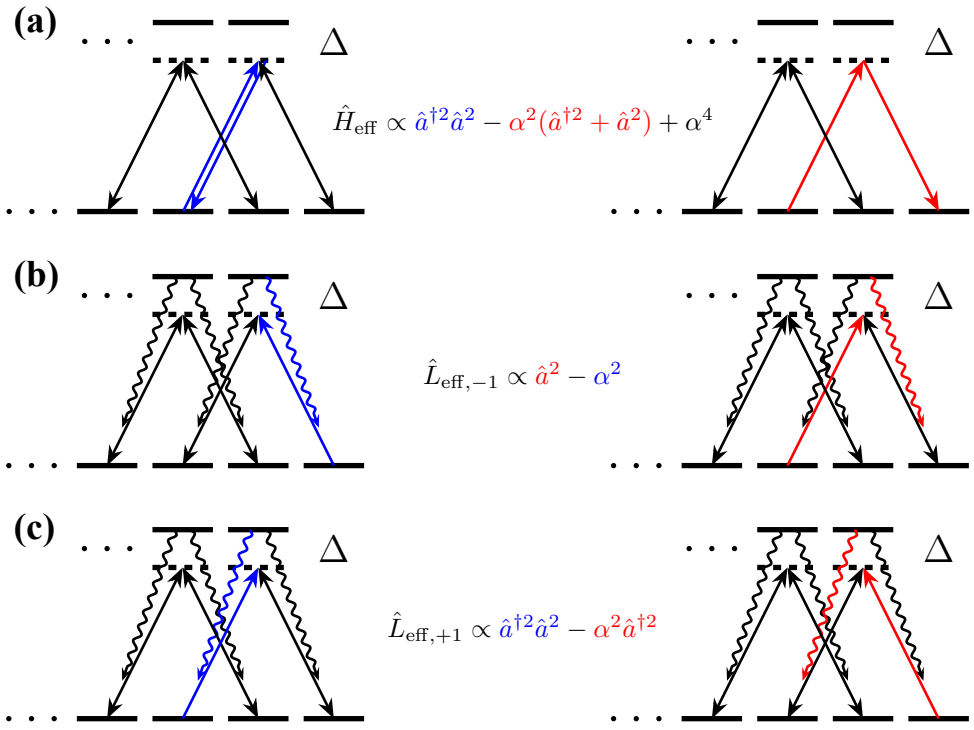

Figure 5.2: The correspondence of the stabilization mechanisms between generalized dark spin-cat and bosonic cat under H-P transformation. (a-c) The illustration of (a) $\hat{H}_{\text{eff}}$, (b) $\hat{L}_{\text{eff},-1}$, and (b) $\hat{L}_{\text{eff},+1}$ respectively. I use blue color to represent m-preserving terms and red for transitions that change m by ± 2.

the calculation. Explicitly, we have

$$C_q(m) = \begin{cases} \sqrt{\dfrac{(F_g+m-1)(F_g+m)}{2F_g(2F_g+1)}} & (q=-1), \\ -\sqrt{\dfrac{(F_g+m)(F_g-m)}{F_g(2F_g+1)}} & (q=0), \\ \sqrt{\dfrac{(F_g-m-1)(F_g-m)}{2F_g(2F_g+1)}} & (q=+1). \end{cases} \tag{5.12}$$

In Eq. (5.11), the first line provides a constant after the H-P transformation; the second and third lines give the parametric drive or dissipation, while the last line is the Kerr nonlinear

term. With these, $\hat{H}_{\text{eff}}$ and $\hat{L}_{\text{eff},q}$ will be transformed into

$$\begin{cases} \hat{H}_{\text{eff}} \to -\dfrac{\Omega^2\Delta}{4\Delta^2+\Gamma^2}\dfrac{\cos^4(\theta/2)}{2F_g(2F_g+1)}[\hat{a}^{\dagger 2}\hat{a}^2-\alpha^2(\hat{a}^{\dagger 2}+\hat{a}^2)+\alpha^4], \\ \hat{L}_{\text{eff},-1} \to \dfrac{\Omega\sqrt{\Gamma}}{2\Delta-i\Gamma}\dfrac{\cos^2(\theta/2)}{2F_g+1}(\hat{a}^2-\alpha^2), \\ \hat{L}_{\text{eff},+1} \to \dfrac{\Omega\sqrt{\Gamma}}{2\Delta-i\Gamma}\dfrac{\cos^2(\theta/2)}{2F_g(2F_g+1)}(\hat{a}^{\dagger 2}\hat{a}^2-\alpha^2\hat{a}^{\dagger 2}), \end{cases} \tag{5.13}$$

where we again choose $\alpha=\sqrt{2F_g}\tan(\theta/2)$. Surprisingly, $\hat{H}_{\text{eff}} \propto (\hat{a}^{\dagger 2}-\alpha^2)(\hat{a}^2-\alpha^2)$ is exactly the Kerr-cat Hamiltonian that provides an energy gap to confine the cat code subspace, while $\hat{L}_{\text{eff},-1} \propto (\hat{a}^2-\alpha^2)$ is the jump operator that dissipatively stabilize the cat. Finally, $\hat{L}_{\text{eff},+1}$ is one order smaller compared with $\hat{L}_{\text{eff},-1}$ in the $F_g \gg 1$ limit. Since $\hat{L}_{\text{eff},+1} \propto \hat{a}^{\dagger 2}(\hat{a}^2-\alpha^2)$, it belongs to a general type of dissipators that also stabilize the cat subspace and preserve the parity.

Coming back to the generalized spin-cat model, the elegant analogy described above indicates that both the dissipation and the energy gap protection like that in the bosonic cat naturally arise in our system by simply applying some laser drives. This is in contrast to the current bosonic cat experiment, where more sophisticated nonlinear devices are engineered to stabilize the code. Therefore, I would expect this model to play a more important role in the field of continuous-variable QEC in the future.

5.3 Semiclassical analysis on the tunneling rate

In this section, I will focus on the properties of the stabilized Hamiltonian $\hat{H}_{\text{DS}}$, especially in the two situations where the stabilization drives are either on resonance ($\Delta=0$) or far off-resonant ($\Delta \gg \Omega$, where F_e-levels can be adiabatically eliminated) separately. However, as we shall see, in both cases $\hat{H}_{\text{DS}}$ still gives a double-well potential in the relevant phase space regime and both the SCSs $|\theta,0\rangle_g$ and $|\theta,\pi\rangle_g$ for encoding are located at the extrema

of each well, respectively.

When a state in one well leaves the code state, it may tunnel into the other well and therefore result in a bit-flip. Here I will use the WKB approach to derive the tunneling rate, which has been used for the same task under the Kerr-cat Hamiltonian [163, 164]. The technical difficulty in the spin-cat setting mainly comes from the presented spherical phase space, which indicates that we need to work with a new set of conjugate variables rather than the conventional position and momentum operators. Luckily, the related problems have been intensively studied by J.L. van Hemmen *et al.* in a series of work (see Ref. [165, 166]), though the final results still seem to contain minor flaws. I will closely follow the derivation there while adapting the generalized spin-cat model into it. The key steps are briefly listed in Sec. 5.3.2. Finally, I will discuss the results and the validity of the WKB approximation.

5.3.1 Generic properties of the stabilization Hamiltonian

Let us first explore the generic structures of the eigenstates and eigen-energies of $\hat{H}_{\rm DS}$. For simplicity, here I also set $\delta = 0$. Now $\hat{H}_{\rm DS}$ can be divided in three parts:

$$\hat{H}_{\rm DS} = \Delta \hat{P}_e + \hat{H}_{eg} + \hat{H}_{ge}, \tag{5.14}$$

where $\hat{H}_{eg} = \hat{P}_e \hat{H}_{\rm DS} \hat{P}_g = \frac{1}{2}\sum_{q=\pm 1} \Omega_q \hat{D}_{eg,q}$ maps F_g-levels to F_e-levels and $\hat{H}_{ge} = \hat{H}_{eg}^\dagger$. Consider an eigenstate $|\psi_k\rangle$ of $\hat{H}_{\rm DS}$ with eigenvalue E_k. $|\psi_k\rangle$ can also be separated into two parts: $|\psi_k\rangle = \hat{P}_g |\psi_k\rangle + \hat{P}_e |\psi_k\rangle := |\psi_k^g\rangle + |\psi_k^e\rangle$, which should satisfy

$$\begin{cases} \Delta |\psi_k^e\rangle + \hat{H}_{eg} |\psi_k^g\rangle = E_k |\psi_k^e\rangle , \\ \hat{H}_{ge} |\psi_k^e\rangle = E_k |\psi_k^g\rangle . \end{cases} \tag{5.15}$$

With this, we can eliminate $|\psi_k^e\rangle$ and achieve

$$\frac{\hat{H}_{ge}\hat{H}_{eg}}{E_k - \Delta}\,|\psi_k^g\rangle = E_k\,|\psi_k^g\rangle\,. \tag{5.16}$$

Therefore, $|\psi_k^g\rangle$ is an eigenstate of $\hat{\mathbf{H}} := \hat{H}_{ge}\hat{H}_{eg}$, and later we can just focus on $\hat{\mathbf{H}}$ itself and its eigenvalues $\mathbf{E}_k = E_k(E_k - \Delta)$. With $\Omega_{\pm 1}$ in Eq. (5.1), we have

$$\begin{aligned}\hat{\mathbf{H}} &= \frac{\Omega^2}{4}\frac{F_g^2[1+\cos^2(\theta)] - \hat{F}_{g,x}^2 - \cos^2(\theta)\hat{F}_{g,y}^2 - (2F_g-1)\cos(\theta)\hat{F}_{g,z}}{2F_g(2F_g+1)} \\ &= \frac{\Omega^2}{4}\frac{F_g^2 - F_g\cos^2(\theta) - \sin^2(\theta)\hat{F}_{g,x}^2 + \cos^2(\theta)\hat{F}_{g,z}^2 - (2F_g-1)\cos(\theta)\hat{F}_{g,z}}{2F_g(2F_g+1)}.\end{aligned} \tag{5.17}$$

When the stabilization drives are on resonance ($\Delta = 0$), both $|\psi_k^g\rangle \pm |\psi_k^e\rangle$ are eigenstates of $\hat{H}_{\mathrm{DS}}$ with eigenvalue $\pm E_k$. The corresponding eigenvalue in $\hat{\mathbf{H}}$ satisfies $\mathbf{E}_k = E_k^2$. Meanwhile, we can verify that $\langle\psi_k^g|\psi_k^g\rangle = \langle\psi_k^e|\psi_k^e\rangle = \frac{1}{2}$.

When $\Delta = 0$, there is also a connection to the two-photon exchange (TPE) Hamiltonian $\hat{H}_{\mathrm{TPE}}$ that is used to confine the bosonic cat [144]:

$$\hat{H}_{\mathrm{TPE}} = (\hat{a}^2 - \alpha^2)\hat{\sigma}_+ + \mathrm{h.c.}, \tag{5.18}$$

where $\hat{\sigma}_+ = |e\rangle\langle g|$ is an operator on the qubit that is coupled with the cavity. The eigenvalue of the TPE Hamiltonian $\hat{H}_{\mathrm{TPE}}$ will be the square root of the corresponding Kerr-Hamiltonian $\hat{H}_K = (\hat{a}^{\dagger 2} - \alpha^{*2})(\hat{a}^2 - \alpha^2)$. The analogy with our spin-cat model can be illustrated by treating the presence of the state in F_g- or F_e-manifold as a gauge qubit. In this way, the $\Omega_{\pm 1}$ drives in $\hat{H}_{eg}$ (or $\hat{H}_{ge}$) will not only modify the state in the spin degree of freedom, but also cause the $|g\rangle - |e\rangle$ transition in the gauge qubit.

On the other hand, if the drives are far off-resonant ($\Delta \gg \Omega$), then the adiabatic elimination treatment in Eq. (5.10) will still apply here. We can find the solutions $E_k \sim O(\frac{\Omega^2}{\Delta}) \ll \Delta$. In this limit, E_k will be proportional to the eigenvalues in $\hat{\mathbf{H}}$: $E_k \simeq -\frac{\mathbf{E}_k}{\Delta}$. As explained in

Sec. 5.2.2, it can also be understood from the effective Hamiltonian perspective $\hat{H}_{\text{eff}} \simeq -\frac{1}{\Delta}\hat{\mathbf{H}}$, which will be more similar to the standard Kerr-cat stabilization.

5.3.2 A brief review of the WKB approach on a spin

Now I will briefly illustrate how to use a semiclassical approach named WKB method to analyze the eigenstate structures of operator $\hat{\mathbf{H}}$ and the difference between two near-degenerate eigenvalues, which will be valid in the $F_g \gg 1$ regime. As mentioned before, the generic WKB approach for a spin system is well studied in existing literature (see Ref. [165]), and here I will apply it to the specific spin-cat configuration. Since $\hat{\mathbf{H}}$ is restricted on the F_g-manifold, in this subsection I will omit the subscript g for simplicity. To keep the convention with the existing work [165], I will rotate the frame so that the two wells are located at $|\frac{\pi}{2} \pm \theta, 0\rangle$. Equivalently speaking, I will focus on

$$\begin{aligned}\hat{\mathbf{H}}_{\text{RN}} &:= \frac{8F(2F+1)}{\Omega^2}\hat{R}_y\left(\frac{\pi}{2}\right)\hat{\mathbf{H}}\hat{R}_y^\dagger\left(\frac{\pi}{2}\right) \\ &= F^2[1+\cos^2(\theta)] - \hat{F}_z^2 - \cos^2(\theta)\hat{F}_y^2 - (2F-1)\cos(\theta)\hat{F}_x.\end{aligned} \tag{5.19}$$

Though neither $\hat{\mathbf{H}}$ nor $\hat{\mathbf{H}}_{\text{RN}}$ has a dimension of energy, I will still treat them as "Hamiltonian" later, since mathematically the dimension can always be adjusted by multiplying a dimensional factor.

To apply the WKB method, the first step we need is to identify a conjugate pair of variables. Unlike the harmonic oscillator where we can simply choose the position and momentum for that, here for a spin (angular momentum) system the classical mechanics suggests that we can choose the following pair [165]:

$$q = f_z, \quad p = -\phi = -\arctan(f_y/f_x), \tag{5.20}$$

where f_ξ ($\xi = x, y, z$) are the components of the classical spin with length $\sqrt{F(F+1)}$. We

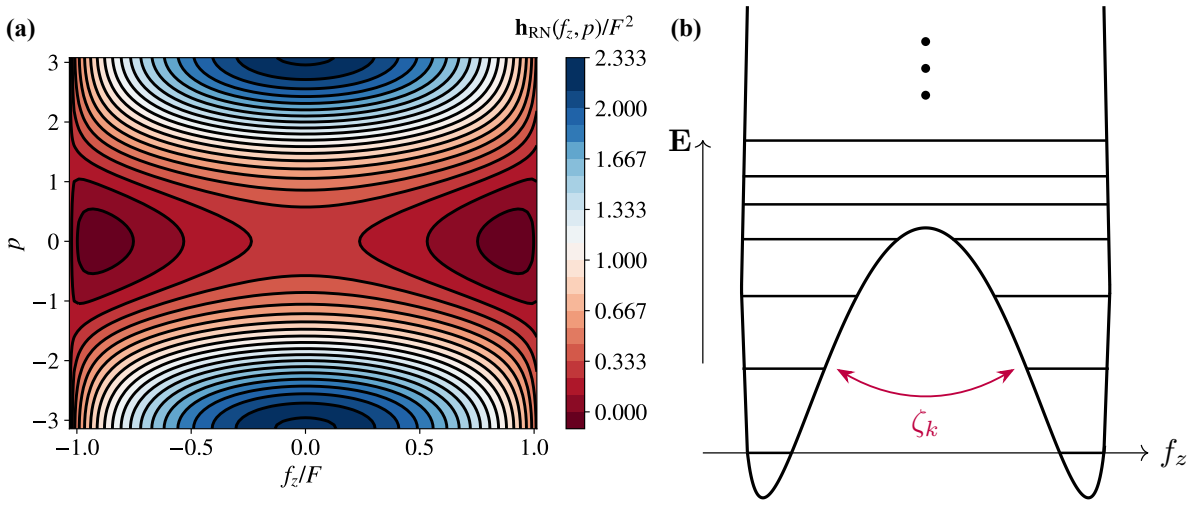

Figure 5.3: Double-well structure in the Hamiltonian stabilization of the generalized spin-cat code subspace. (a) The contour plot of the classical Hamiltonian $\mathbf{h}_{\mathrm{RN}}(f_z,p)$ in the phase space when $F=15$. We can find two wells located symmetrically about the $f_z=0$ line. (b) Schematic plot of the $\hat{\mathbf{H}}_{\mathrm{RN}}$ level structure. The double-well envelope is achieved from $p=0$ and $p=\pi$ intercepts in (a).

can write down the classical Hamiltonian corresponding to $\hat{\mathbf{H}}_{\mathrm{RN}}$ as:

$$\mathbf{h}_{\mathrm{RN}}(f_z,p)=F^2[1+\cos^2(\theta)]-f_z^2-\cos^2(\theta)f_y^2(f_z,p)-(2F-1)\cos(\theta)f_x(f_z,p), \tag{5.21}$$

where

$$\begin{cases} f_x(f_z,p)=b(f_z)\cos(p), \\ f_y(f_z,p)=-b(f_z)\sin(p), \\ b(f_z)=\sqrt{F(F+1)-f_z^2}. \end{cases} \tag{5.22}$$

The double-well feature in $\mathbf{h}_{\mathrm{RN}}(f_z,p)$ can be clearly observed from the contour plot [see Fig. 5.3(a)], with the two minima located at $(f_z=\pm f_{z,\min},p=0)$ that

$$f_{z,\min}=\sqrt{F(F+1)-(F-1/2)^2\cos^2(\theta)}\simeq F\sin(\theta). \tag{5.23}$$

On the other hand, the classical energy at these two points are $\mathbf{E}_{\mathrm{cl,min}}=-F\sin^2(\theta)-$

$\frac{1}{4}\cos^2(\theta)$, which is in contrast with eigen-energy $\mathbf{E}_{\min} = 0$ of the quantum ground state $|\frac{\pi}{2} \pm \theta, 0\rangle$. Similar to a harmonic oscillator, the difference between $\mathbf{E}_{\text{cl,min}}$ and $\mathbf{E}_{\min} = 0$ indicates that there is also a zero-point energy in the quantum ground state here, and I will denote $\mathbf{E}' := \mathbf{E} - \mathbf{E}_{\text{cl,min}}$ to adjust the zero points for further use.

Then let us come back to quantum. Given a state $|\psi\rangle$, an eigenstate of one well with eigenvalue $\mathbf{E}$. We can express its wavefunction with the WKB ansatz:

$$\psi(f_z) := \langle F_g, f_z|\psi\rangle = e^{iS(f_z)} = e^{i[S_0(f_z)+S_1(f_z)+O(1/F)]}, \tag{5.24}$$

where we try to expand $S(f_z)$ as a power series of F so that $S_0 \sim O(F)$, $S_1 \sim O(1)$, and the residue terms can be ignored in the large F limit. In order for $|\psi\rangle$ to be an eigenstate with eigenvalue $\mathbf{E}$, the WKB formalism further suggests that S_0 and S_1 should satisfy

$$\begin{aligned} &- f_z^2 - \cos^2(\theta)b^2(f_z)\sin^2\left(\frac{\mathrm{d}S_0}{\mathrm{d}f_z}\right) - (2F-1)\cos(\theta)b(f_z)\cos\left(\frac{\mathrm{d}S_0}{\mathrm{d}f_z}\right) \\ &= \mathbf{E} - F^2[1+\cos^2(\theta)], \end{aligned} \tag{5.25}$$

and

$$\frac{\mathrm{d}}{\mathrm{d}f_z}\left[\dot{f}_z(f_z)\cdot e^{2iS_1(f_z)}\right] = 0. \tag{5.26}$$

Here Eq. (5.25) is exactly the Hamilton-Jacobi equation in classical mechanics:

$$\mathbf{h}_{\text{RN}}\left(f_z, p = \frac{\mathrm{d}S_0}{\mathrm{d}f_z}\right) = \mathbf{E}, \tag{5.27}$$

with the interpretation of S_0 as the action. Eq. (5.26) corresponds to the conservation of the probability current, where from Hamilton's equation we can get

$$\dot{f}_z = \frac{\partial \mathbf{h}(f_z, p)}{\partial p} = -\cos^2(\theta)b^2(f_z)\sin(2p) + (2F-1)\cos(\theta)b(f_z)\sin(p). \tag{5.28}$$

The f_z dependence further comes from the substitution of $p = p(f_z, \mathbf{E})$ in the equation above. After this, the wavefunction $\psi(f_z)$ can be expressed as

$$\psi(f_z) \simeq \frac{C}{\sqrt{|\dot{f}_z(f_z)|}} e^{iS_0(f_z)}, \tag{5.29}$$

where C is a normalization constant.

In order to determine $S_0(f_z)$, now we need to solve the Hamilton-Jacobi equation (5.25) [or Eq. (5.27) equivalently], which gives

$$\cos(p) = c(f_z) := \frac{(F - \frac{1}{2}) - \sqrt{(F - \frac{1}{2})^2 + \sin^2(\theta) f_z^2 + \mathbf{E} - F[F - \cos^2(\theta)]}}{b(f_z)\cos(\theta)}. \tag{5.30}$$

Classically, it is required that $c(f_z)$ is real and $|c(f_z)| \leq 1$. The first requirement is always satisfied when $\mathbf{E} > F\sin^2(\theta) - \frac{1}{4}$, and as I will show later, the energy of every excited state in $\hat{\mathbf{H}}_{\mathrm{RN}}$ will fulfill this. The second requirement needs more attention. When $\mathbf{E}$ is smaller than the barrier energy $\mathbf{E}_{\mathrm{bar}}$ at $(f_z = 0, p = 0)$, where

$$\mathbf{E}_{\mathrm{bar}} = F^2[1 + \cos^2(\theta)] - (2F - 1)\sqrt{F(F+1)}\cos(\theta) \simeq 4F^2 \sin^4(\theta/2), \tag{5.31}$$

then solving $\cos(p) = c(f_z) = 1$ will give us two turning points $f_{z,\pm}$:

$$f_{z,\pm} = \sqrt{F(F+1) - \left[(F - \frac{1}{2})\cos(\theta) \mp \sqrt{\mathbf{E}'}\right]^2}. \tag{5.32}$$

The classical trajectory will be confined in $f_{z,-} \leq f_z \leq f_{z,+}$ while both $-f_{z,-} \leq f_z \leq f_{z,-}$ and $f_{z,+} \leq f_z \leq \sqrt{F(F+1)}$ are classically forbidden.

In quantum mechanics, however, p can be a complex number in general due to the tunneling effect. This allows us to get a nontrivial wavefunction in both classically allowed and classically forbidden regimes. In the former case, $p = \arccos[c(f_z)]$ is still a real number,

and according to the WKB method, the wavefunction can be written as

$$\psi(f_z) \simeq \frac{2C}{\sqrt{|\dot{f}_z(f_z)|}} \cos\Big[\int_{f_{z,-}}^{f_z} \arccos[c(f)]\,\mathrm{d}f - \frac{\pi}{4}\Big] \qquad f_z \in (f_{z,-}, f_{z,+}). \tag{5.33}$$

In the classically forbidden regime, we have $c(f_z) > 1$. Now p will be a pure imaginary number. We denote $p(f_z) = -ip_I(f_z)$ to make $p_I(f_z) > 0$ when $f_z \in (-f_{z,-}, f_{z,-})$. In this regime, the wavefunction becomes:

$$\psi(f_z) \simeq \frac{C}{\sqrt{|\dot{f}_z(f_z)|}} \exp\Big[-\int_{f_z}^{f_{z,-}} \operatorname{arccosh}[c(f)]\,\mathrm{d}f\Big] \qquad f_z \in (-f_{z,-}, f_{z,-}). \tag{5.34}$$

The difference between the coefficients in Eq. (5.33) and (5.34) comes from the loss of one branch in (5.33) when f_z passes around the boundary point $f_{z,-}$ through analytic continuation [167]. Meanwhile, now $|\dot{f}_z|$ should satisfy

$$|\dot{f}_z(f_z)| = \Big|-\cos^2(\theta)b^2(f_z)\sinh[2p_I(f_z)] + (2F-1)\cos(\theta)b(f_z)\sinh[p_I(f_z)]\Big|, \tag{5.35}$$

which is achieved by the substitution of $p = -ip_I$ in Eq. (5.28).

The last thing we need is to determine the overall coefficient C in the wavefunction. Since most population will be in the classically allowed regime, we can get from Eq. (5.33) that

$$\begin{aligned} 1 &\simeq \int_{f_{z,-}}^{f_{z,+}} |\psi(f)|^2\,\mathrm{d}f \simeq 4|C|^2 \int_{f_{z,-}}^{f_{z,+}} \frac{\mathrm{d}f}{|\dot{f}|}\, \overline{\cos^2\Big[\int_{f_{z,-}}^{f} \arccos[c(f')]\,\mathrm{d}f' - \frac{\pi}{4}\Big]} \\ &\simeq \frac{4|C|^2}{2} \int_{f_{z,-}}^{f_{z,+}} \frac{\mathrm{d}f}{|\dot{f}|} = 2|C|^2 \int \mathrm{d}t = |C|^2 T_{\mathrm{cl}}. \end{aligned} \tag{5.36}$$

Here we have assumed that the cosine term is fast oscillating in the parameter regime and therefore its square is averaged to be 1/2. T_{cl} is the period of the corresponding classical

trajectory, which in practice can be computed via

$$\begin{aligned} T_{\rm cl} &= 2\int_{f_{z,-}}^{f_{z,+}} \frac{\mathrm{d}f}{|\dot{f}|} \\ &= 2\int_{f_{z,-}}^{f_{z,+}} \frac{\mathrm{d}f}{-\cos^2(\theta)b^2(f)\sin[2p(f,\mathbf{E})] + (2F-1)\cos(\theta)b(f)\sin[p(f,\mathbf{E})]}. \end{aligned} \tag{5.37}$$

With this, the coefficient C can be determined as $C = 1/\sqrt{T_{\rm cl}}$.

After achieving the wavefunction of an eigenstate in a single well, we can further derive the energy difference of two near-degenerate eigenstates in the double-well Hamiltonian $\hat{\mathbf{H}}_{\rm RN}$. To do so, we first denote $\psi_\uparrow(f_z)$ as the wavefunction solved in Eq. (5.33) and (5.34), since the corresponding well is located at the $f_z > 0$ regime. Due to the symmetry consideration, $\psi_\downarrow(f_z) = \psi_\uparrow(-f_z)$ should be an eigenstate of the well at $f_z < 0$ regime. Therefore, if we view the double-well Hamiltonian as a whole, then the corresponding eigenstates should be an even or odd superposition between $\psi_{\uparrow/\downarrow}(f_z)$:

$$\psi_\pm(f_z) \simeq \frac{\psi_\uparrow(f_z) \pm \psi_\downarrow(f_z)}{\sqrt{2}}, \tag{5.38}$$

with the corresponding eigenvalue $\mathbf{E}_\pm$. The energy difference $\mathbf{E}_{\rm dif} = |\mathbf{E}_+ - \mathbf{E}_-|$ can also be interpreted as twice of the tunneling rate between $\psi_{\uparrow/\downarrow}(f_z)$, the corresponding eigenstates in each well separately.

To evaluate $\mathbf{E}_{\rm dif}$, here I adopt a trick used in Ref. [165, 167], with details shown in Appendix D.1.3. We can finally get

$$\mathbf{E}_{\rm dif} \simeq 2|\dot{f}_z(f_z=0)| \cdot \psi_\uparrow^2(0) \simeq \frac{2}{T_{\rm cl}} \exp\Big[-2\int_0^{f_{z,-}} \mathrm{arccosh}[c(f)]\,\mathrm{d}f\Big]. \tag{5.39}$$

The result here is different from that in the existing literature by a prefactor [165], and the correctness of Eq. (5.39) is justified from the numerical study. Due to the F-dependence in $f_{z,-} \sim O(F)$, the formula also reflects the generic exponential suppression of the tunneling

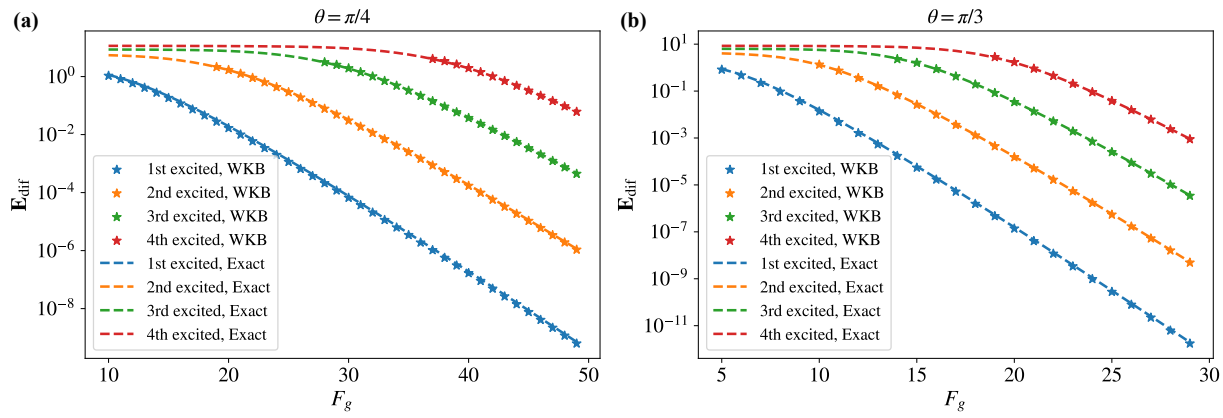

Figure 5.4: WKB results on the estimation of tunneling rates, when (a) $\theta = \pi/4$ and (b) $\theta = \pi/3$. The exact $\mathbf{E}_{\text{dif}}$ numbers are plotted with the dashed lines while the WKB results (5.39) are shown with star points. Here I use the true $\bar{\mathbf{E}} = \frac{\mathbf{E}_+ + \mathbf{E}_-}{2}$ in the WKB formula to justify its correctness. In practice, approximated $\mathbf{E}$ numbers can be achieved as mentioned in Sec. 5.3.3, which should be used to avoid the exact diagonalization.

rate with the spin length F.

5.3.3 Discussion on the results

After deriving the exact tunneling rate formula between two wells using the WKB approximation, I will now move on to discuss under what circumstances the approximation is valid. As mentioned, the eigen-energy $\mathbf{E}$ of the corresponding states should be lower than the barrier $\mathbf{E}_{\text{bar}} \simeq 4F^2 \sin^4(\theta/2)$, which raises the question of how many pairs of eigenstates meet this requirement.

To investigate the problem, we need to extract the eigen-energies of the low-excited eigenstates in $\hat{\mathbf{H}}_{\text{RN}}$. Recall that, when dealing with the Kerr-cat [168] or pair-cat Hamiltonian (see Chapter 3 or Appendix B.2), what we do is to go to the displaced frame so that one well is located at the origin. Here what we can do is to further rotate $\hat{\mathbf{H}}_{\text{RN}}$ (or $\hat{\mathbf{H}}$ equivalently)

such that one well is located at the north pole of the GBS. This will give us

$$\begin{aligned}\hat{R}_y^\dagger\left(\frac{\pi}{2}-\theta\right)\hat{\mathbf{H}}_{\mathrm{RN}}\hat{R}_y\left(\frac{\pi}{2}-\theta\right) = \cos^2(\theta)(\hat{F}_z-F)(\hat{F}_z-F+1)-\sin^2(\theta)(\hat{F}_z^2-F^2)\\ +\Big[\frac{\sin(2\theta)}{2}\hat{F}_-(\hat{F}_z-F)+\mathrm{h.c.}\Big].\end{aligned} \tag{5.40}$$

The diagonal term with state $|F_g, F_g-k\rangle$ is $2Fk\sin^2(\theta)+k[k\cos(2\theta)-\cos^2(\theta)]$, which is approximately $2Fk\sin^2(\theta)$ when $k\sim O(1)\ll F$. In the same regime, however, the off-diagonal term between $|F_g, F_g-k\rangle$ and $|F_g, F_g-(k+1)\rangle$ only scales as $O(\sqrt{F})$. Therefore, now we may use the leading order $2Fk\sin^2(\theta)$ in the diagonal term to approximate the true eigen-energies[1], which give a rough estimation on the number of excited eigenstate pairs below the energy barrier as $n_{\mathrm{crit}}\sim\frac{F}{2}\tan^2(\theta/2)$. A more detailed derivation requires us to solve

$$2Fn_{\mathrm{crit}}\sin^2(\theta)+n_{\mathrm{crit}}[n_{\mathrm{crit}}\cos(2\theta)-\cos^2(\theta)]=4F^2\sin^4(\theta/2), \tag{5.41}$$

which leads to

$$n_{\mathrm{crit}}\simeq\frac{F\tan^2(\theta/2)}{1+\sqrt{1+\frac{\cos(2\theta)}{4\cos^4(\theta/2)}}}. \tag{5.42}$$

Another problem comes when we consider the energy difference between a pair of near-degenerate eigenstates in the physical Hamiltonian $\hat{H}_{\mathrm{DS}}$, if the stabilization drives are on resonance ($\Delta=0$). The eigen-energies $E_\pm$ in $\hat{H}_{\mathrm{DS}}$ are now proportional to the square root of the corresponding eigenvalues $\mathbf{E}_\pm$ in $\hat{\mathbf{H}}_{\mathrm{RN}}$. This allows us to further derive that

$$\begin{aligned}E_{\mathrm{dif}}:=|E_+-E_-| &= \frac{\Omega}{\sqrt{8F(2F+1)}}|\sqrt{\mathbf{E}_+}-\sqrt{\mathbf{E}_-}| = \frac{\Omega}{\sqrt{8F(2F+1)}}\frac{\mathbf{E}_{\mathrm{dif}}}{2\sqrt{\bar{\mathbf{E}}}}\\ &= \frac{\Omega^2}{8F(2F+1)}\frac{\mathbf{E}_{\mathrm{dif}}}{2\bar{E}},\end{aligned} \tag{5.43}$$

where we have used $\mathbf{E}_\pm=\bar{\mathbf{E}}\pm\frac{\mathbf{E}_{\mathrm{dif}}}{2}$. In fact, the $O(\bar{\mathbf{E}}^{-1/2})$ dependence in E_{dif} is also a

1. This also justifies the claim that the energy $\mathbf{E}$ of every excited states in $\hat{\mathbf{H}}_{\mathrm{RN}}$ satisfies $\mathbf{E}>F\sin^2(\theta)-\frac{1}{4}$, so that $c(f_z)$ in Eq. (5.30) is always real.

feature in the bosonic cat confinement with the TPE Hamiltonian (5.18) (see Ref. [144]).

5.4 Tunneling rate reduction with occasional degeneracy

5.4.1 Occasional degeneracy with Kerr-cat Hamiltonian

In the previous section, I have used the WKB approach to analyze the tunneling rate between the two wells. In fact, people working on the Kerr-cat have found one method to suppress that, which relies on the introduction of a detuning term $\delta_a \hat{a}^\dagger \hat{a}$. If δ_a is chosen at some special points, then we can receive several degenerate level pairs between the two wells simultaneously [123, 124].

The idea behind it is the novel photon blockade viewed from the displaced frame, which is essentially the same as that used in Chapter 2. Consider the Hamiltonian $\hat{H}_{K,\delta}$ that

$$\hat{H}_{K,\delta} = -K(\hat{a}^{\dagger 2} - \alpha^{*2})(\hat{a}^2 - \alpha^2) + \delta_a \hat{a}^\dagger \hat{a}. \tag{5.44}$$

If viewed from the displaced frame $\hat{D}(\pm\alpha)$, then the Hamiltonian will become

$$\begin{aligned} \hat{D}^\dagger(\pm\alpha)\hat{H}_{K,\delta}\hat{D}(\pm\alpha) = & -K\hat{a}^{\dagger 2}\hat{a}^2 + (\delta_a - 4|\alpha|^2 K)\hat{a}^\dagger \hat{a} + \delta_a |\alpha|^2 \\ & \pm [\alpha \hat{a}^\dagger (\delta_a - 2K\hat{n}) + \text{h.c.}]. \end{aligned} \tag{5.45}$$

The displaced Hamiltonian only couples the nearest Fock states, and the coupling between $|n_0\rangle$ and $|n_0+1\rangle$ will be zero if we choose $\delta_a = 2Kn_0$. Therefore, the full Hilbert space can be divided into two parts without coupling in between. For the eigenstates within the subspace spanned by $\{|0\rangle, |1\rangle, \ldots, |n_0\rangle\}$, their corresponding eigenvalues E should be solved by diagonalizing the Hamiltonian restricted to the blockade subspace. Alternatively, E should be the zeros of the characteristic polynomial of the matrix:

$$\det(\hat{P}_{\leq n_0}\hat{D}^\dagger(\pm\alpha)\hat{H}_{K,\delta}\hat{D}(\pm\alpha)\hat{P}_{\leq n_0} - E \cdot \hat{P}_{\leq n_0}) = 0, \tag{5.46}$$

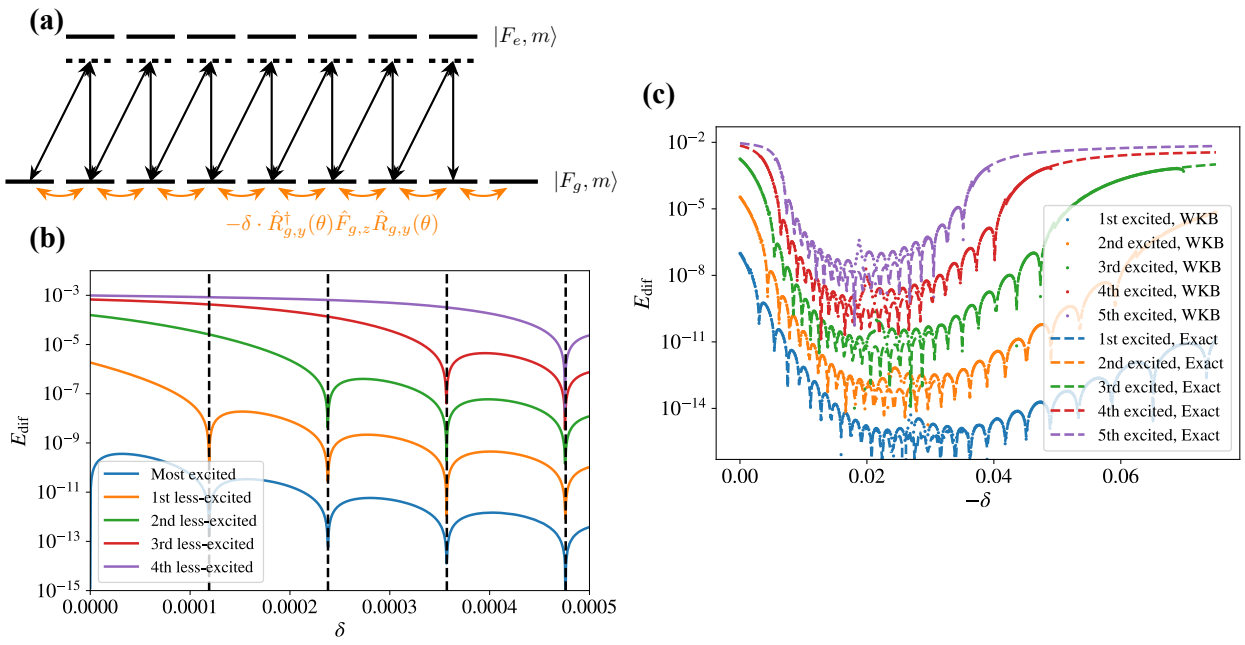

Figure 5.5: Occasional degeneracies with the introduction of F_g-level detuning. (a) Illustration of the blockade in the rotated frame. We will find $\Omega_{-1} = 0$ in this frame, which implies that only the nearest levels in the F_g manifold will be coupled after adiabatic elimination. $\hat{F}_{g,z}$ after rotation will also couple the nearest levels in a different way (it also shifts the levels but is not shown here), which means we can adjust δ to cancel certain transitions. (b) E_{dif} computed with $F_g = 10$, $\Omega = 1$, $\Delta = 5$ and $\theta = \pi/3$. Black dashed lines are references for $\delta = \dfrac{\Omega^2 \cos(\theta)}{4\Delta F_g(2F_g+1)} m_0$ with integer multiple m_0. (c) WKB results with resonant drives. The system is chosen as $F_g = 15$, $\Omega = 1$, and $\theta = \pi/3$ while I only look into the positive energy branch. Dashed lines are exact while dots are achieved with Eq. (5.61) and exact $\bar{E}$ input. Some mismatch in the middle comes from the singularity in numerical integration.

where $\hat{P}_{\leq n_0} := \sum_{k=0}^{n_0} |k\rangle\langle k|$. Further, if we have two tridiagonal matrices with the same diagonal terms but their off-diagonal elements are different by a minus sign, then they will have the same corresponding characteristic polynomial, and therefore the same set of eigenvalues. As shown in Eq. (5.45), this applies to the blockade Hamiltonians viewed in the two different displaced frames $\hat{D}(\pm\alpha)$, which implies that $\hat{H}_{K,\delta}$ should contain $(n_0 + 1)$ degenerate energy pairs when $\delta_a = 2Kn_0$.

5.4.2 Generalized spin-cat with far off-resonant drives

When it comes to the spin-cat, similar features can be observed in the far off-resonant regime ($\Delta \gg \Omega$), if we have a detuning term on the F_g-manifold. To achieve this, we can choose the frequencies of each stabilization drive to satisfy $\omega_q - \omega_{q'} = \delta_e(q - q')$. Then, from the lab Hamiltonian $\hat{H}_{\text{lab}}$ in Eq. (4.1), we can get one in a rotating frame

$$\hat{H}'_{\text{DS}} = \Delta \hat{P}_e - \delta \hat{F}_{g,z} + \frac{1}{2} \sum_{q=0,\pm 1} (\Omega_q \hat{D}_{eg,q} + \text{h.c.}), \tag{5.47}$$

where $\delta = \delta_e - \delta_g$. If we further have $\Delta \gg \Omega, \delta$, then the F_e-levels can be eliminated [162]. The effective Hamiltonian left now is:

$$\begin{aligned} \hat{H}_{\text{eff},\delta} &= -\frac{1}{\Delta} \hat{\mathbf{H}} - \delta \hat{F}_{g,z} \\ &= -\frac{\Omega^2}{4\Delta} \frac{-\sin^2(\theta) \hat{F}^2_{g,x} + \cos^2(\theta) \hat{F}^2_{g,z} - (2F_g - 1)\cos(\theta) \hat{F}_{g,z}}{2F_g(2F_g + 1)} - \delta \hat{F}_{g,z} \end{aligned} \tag{5.48}$$

where Ω_q is chosen as Eq. (5.1), and some constant terms are omitted.

In analogy with the displaced frame, here we can go to the rotated frame $\hat{R}_{g,y}(\pm\theta) = e^{\mp i\theta \hat{F}_{g,y}}$ in order to observe the blockade effect. From Fig. 5.5(a), it is straightforward to see that $\hat{R}^\dagger_{g,y}(\pm\theta) \hat{\mathbf{H}} \hat{R}_{g,y}(\pm\theta)$ only couples $|F_g, m\rangle$ with the nearest levels $|F_g, m \pm 1\rangle$, due to the absence of the Ω_{-1} drives viewed from here. Besides, $\hat{R}^\dagger_{g,y}(\pm\theta) \hat{F}_{g,z} \hat{R}_{g,y}(\pm\theta) = \cos(\theta) \hat{F}_{g,z} \mp \sin(\theta) \hat{F}_{g,x}$ only couples the nearest levels as well. These two features suggest that, if we want to reproduce the blockade effect, we can adjust δ here to exactly cancel out the coupling between two nearest levels, like $|F_g, F_g - m_0\rangle$ and $|F_g, F_g - (m_0 + 1)\rangle$ for

example. More rigorously, we have

$$\hat{R}^{\dagger}_{g,y}(\pm\theta)\hat{H}_{\text{eff},\delta}\hat{R}_{g,y}(\pm\theta) = -\frac{\Omega^2}{4\Delta}\frac{\cos(2\theta)\hat{F}^2_{g,z}-(2F_g-1)\cos^2(\theta)\hat{F}_{g,z}}{2F_g(2F_g+1)}-\delta\cos(\theta)\hat{F}_{g,z} \\ \pm\Big[\frac{\Omega^2\sin(2\theta)}{8\Delta}\frac{\hat{F}_{g,-}(\hat{F}_{g,z}-F_g)}{2F_g(2F_g+1)}+\frac{\delta\sin(\theta)}{2}\hat{F}_{g,-}+\text{h.c.}\Big] \tag{5.49}$$

Therefore, we can choose $\delta=\frac{\Omega^2\cos(\theta)}{4\Delta F_g(2F_g+1)}m_0$ with integer m_0 to achieve the desired destructive interference, and the effective Hamiltonian in the rotated frame becomes block-diagonal in the two subspaces spanned by $\{|F_g,F_g-m_0\rangle,\ldots,|F_g,F_g\rangle\}$ and its complementary subspace. Then, as the Kerr-cat situation, we can still use the characteristic polynomial argument to show that $\hat{H}_{\text{eff},\delta}$ contains (m_0+1) pairs of degenerate eigen-energies. However, the allowed m_0 need to be bounded by $m_0\leq F_g-1$ to avoid that the achieved eigenstates with the same eigenvalue from the two blockade Hamiltonian $\hat{P}_{\leq m_0}\hat{R}^{\dagger}_{g,y}(\pm\theta)\hat{H}_{\text{eff},\delta}\hat{R}_{g,y}(\pm\theta)\hat{P}_{\leq m_0}$ are actually the same state, where now $\hat{P}_{\leq m_0}=\sum_{k=0}^{m_0}|F_g,F_g-k\rangle\langle F_g,F_g-k|$.

5.4.3 Generalized spin-cat with resonant drives: A semiclassical analysis

However, if the stabilization drives are close to resonance, then we may not be able to achieve several degenerate energy pairs simultaneously by simply adjusting the detuning δ on the F_g-levels. To illustrate this, I will choose $\Delta=0$ and then focus on the eigenstate $|\psi_k\rangle=|\psi^g_k\rangle+|\psi^e_k\rangle$ of $\hat{H}'_{\text{DS}}$ in Eq. (5.47) with eigenvalue E_k. In this way, $|\psi^g_k\rangle$ should satisfy

$$\hat{\mathbf{H}}'|\psi^g_k\rangle=E_k^2|\psi^g_k\rangle, \quad \text{where} \quad \hat{\mathbf{H}}'=\hat{\mathbf{H}}-E_k\delta\hat{F}_{g,z}. \tag{5.50}$$

Suppose now we want to reproduce the blockade feature, and for simplicity, let us assume that the blockade subspace only contains two dimensions. In this case, we should have $E_k\delta=-\frac{\Omega^2\cos(\theta)}{4F_g(2F_g+1)}$. Meanwhile, now we can also diagonalize $\hat{R}^{\dagger}_{g,y}(\pm\theta)\hat{\mathbf{H}}'\hat{R}_{g,y}(\pm\theta)$ in the

two dimensional blockade subspace, which gives

$$E_{k,\pm}^2 = \frac{\Omega^2}{4}\frac{(2F_g-1)[1+\cos^2(\theta)] \pm \sqrt{(2F_g-1)^2\sin^4(\theta)+4\cos^2(\theta)}}{8F_g(2F_g+1)}. \tag{5.51}$$

Unfortunately, we cannot get the two pairs of degenerate eigen-energies simultaneously since they correspond to different detuning δ. For example, if we choose $\delta = -\frac{\Omega^2\cos(\theta)}{4F_g(2F_g+1)\sqrt{E_{k,+}^2}}$, then although the first-excited levels in each of the two wells get degenerate, the two ground levels do not and will result in unwanted tunneling in between.

On the other hand, we can still perform the semiclassical analysis when $\Delta = 0$, and investigate how the tunneling rate varies with the introduction of the F_g-level detuning term $\delta\hat{F}_{g,z}$. To do so, here I still rotate and rescale $\hat{\mathbf{H}}'$ in Eq. (5.50) such that

$$\begin{aligned}\hat{\mathbf{H}}'_{\mathrm{RN}} &:= \frac{8F(2F+1)}{\Omega^2}R_y(\frac{\pi}{2})\hat{\mathbf{H}}'R_y^\dagger(\frac{\pi}{2}) \\ &= F^2[1+\cos^2(\theta)] - \hat{F}_z^2 - \cos^2(\theta)\hat{F}_y^2 - [(2F-1)\cos(\theta)+\gamma]\hat{F}_x,\end{aligned} \tag{5.52}$$

where $\gamma = \frac{8F(2F+1)}{\Omega^2}E\delta$ is a dimensionless parameter. The corresponding classical Hamiltonian now becomes

$$\mathbf{h}'_{\mathrm{RN}}(f_z,p) = F^2[1+\cos^2(\theta)] - f_z^2 - \cos^2(\theta)f_y^2(f_z,p) - [(2F-1)\cos(\theta)+\gamma]f_x(f_z,p). \tag{5.53}$$

Within a wide range of γ, $\mathbf{h}'_{\mathrm{RN}}(f_z,p)$ still processes the double-well structure. The two minima may still be located at $(f_z = \pm f'_{z,\min}, p=0)$ that

$$f'_{z,\min} = \sqrt{F(F+1)-\beta^2/4}, \tag{5.54}$$

where $\beta := (2F-1)\cos(\theta)+\gamma$. The corresponding minimum energy is now $\mathbf{E}'_{\mathrm{cl,min}} = F^2\cos^2(\theta) - F - \frac{\beta^2}{4}$.

When $\delta < 0$, then if we focus on the eigenstates of $\hat{H}'_{\rm DS}$ with positive energies, we will have $\gamma < 0$. The distance between two minima $\pm f'_{z,\rm min}$ will keep increasing with $|\gamma|$ until $|\gamma|$ reaches $(2F-1)\cos(\theta)$. The corresponding tunneling rate will be suppressed accordingly as the two wells become more separated. On the other hand, if we consider the eigenstates with energies in the negative branch, then γ will become positive. The tunneling rate will be increased due to the reduction of the distance between $\pm f'_{z,\rm min}$.

We have also observed several occasional degenerate points between a specific pair of excited states in Fig. 5.5(c), when looking into the $\gamma < 0$ regime and varying δ. In fact, this can still be captured by the WKB analysis. The Hamilton-Jacobi equation we need to solve now becomes

$$\mathbf{h}'_{\rm RN}\left(f_z, p = \frac{\mathrm{d}S_0}{\mathrm{d}f_z}\right) = \mathbf{E}, \tag{5.55}$$

which leads to the solution

$$\cos(p) = c(f_z) := \frac{\frac{\beta}{2\cos(\theta)} - \sqrt{\sin^2(\theta) f_z^2 + \mathbf{E}_\beta}}{b(f_z)\cos(\theta)}, \tag{5.56}$$

where we have denoted $\mathbf{E}_\beta := (\frac{\beta}{2\cos(\theta)})^2 + \mathbf{E} - F[F - \cos^2(\theta)]$. As $|\beta|$ becomes smaller, E_β could be negative in general. In this case, if $f_z \in (-f_{z,0}, f_{z,0})$ with $f_{z,0} = \frac{\sqrt{-\mathbf{E}_\beta}}{\sin(\theta)}$, then $c(f_z)$ itself is a complex number, and therefore p will contain both nonzero real and imaginary parts. On the other hand, solving $c(f_z) = 1$ will give us two classical turning points $f_{z,\pm}$ as before:

$$f_{z,\pm} = \sqrt{F(F+1) - \left(\frac{\beta}{2} \mp \sqrt{\mathbf{E}''}\right)^2}, \tag{5.57}$$

where $\mathbf{E}'' = \mathbf{E} - \mathbf{E}'_{\rm cl,min}$.

Now we can use the WKB method to analyze the wavefunction of an eigenstate in the upper well, which contains different structures within different f_z regimes. Since p is a complex number in general, I will denote $p(f_z) = p_R(f_z) - ip_I(f_z)$ for further use. When f_z is in the classically allowed regime $f_z \in (f_{z,-}, f_{z,+})$, the wavefunction should have the

same structure as that in Eq. (5.33). When $f_z \in (f_{z,0}, f_{z,-})$, $c(f_z)$ is still a real number but $c(f_z) > 1$. Then, $p(f_z)$ becomes pure imaginary, and the wavefunction will be in the same form as Eq. (5.34), an exponential decay pattern.

When $f_z \in (-f_{z,0}, f_{z,0})$, however, $c(f_z)$ itself becomes a complex number. We can solve both $p_R(f_z)$ and $p_I(f_z)$ from

$$\cosh(p_I)\cos(p_R) + i\sinh(p_I)\sin(p_R) = \frac{\frac{\beta}{2\cos(\theta)} + i\sqrt{|\sin^2(\theta) f_z^2 + \mathbf{E}_\beta|}}{b(f_z)\cos(\theta)}. \tag{5.58}$$

The wavefunction $\psi(f_z)$ in this regime can be expressed as

$$\psi(f_z) \simeq \frac{2C}{\sqrt{|\dot{f}_z(f_z)|}} \exp\Big[-\int_{f_z}^{f_{z,-}} p_I(f)\,\mathrm{d}f\Big] \cdot \cos\Big[\chi(f_z) - \int_{f_z}^{f_{z,0}} p_R(f)\,\mathrm{d}f\Big] \quad f_z \in (-f_{z,0}, f_{z,0}). \tag{5.59}$$

As f_z decreases from $f_{z,0}$, $\psi(f_z)$ will start to oscillate with damping. Due to Eq. (5.26), the phase $\chi(f_z)$ here can be determined by tracking the change of complex phase in $\dot{f}_z(f_z)$ when f_z varies from $f_{z,0} + \epsilon$ to 0 (ϵ is a positive infinitesimal number), which gives

$$\chi(f_z) = -\frac{\arg[i\dot{f}_z(f_z)]}{2} = -\frac{\arg\{-i \cdot \sinh[p_I(f_z) + ip_R(f_z)]\}}{2}. \tag{5.60}$$

After achieving the wavefunction $\psi(f_z)$ in all the relevant regimes, we can use a similar trick as that in Eq. (D.14) to compute the energy difference between two near-degenerate levels in $\hat{H}'_{\mathrm{DS}}$. The only difference is that here $\hat{\mathbf{H}}'_{\mathrm{RN}}$ itself also depends on the eigen-energy E of $\hat{H}'_{\mathrm{DS}}$. With a detailed derivation, we can finally achieve that

$$E_{\mathrm{dif}} \simeq \frac{\Omega^2}{8F(2F+1)} \frac{\frac{4}{T_{\mathrm{cl}}} \exp[-2\int_0^{f_{z,-}} p_I(f)\,\mathrm{d}f] \cos[2\int_0^{f_{z,0}} p_R(f)\,\mathrm{d}f]}{2\bar{E} + \delta\,\langle \hat{F}_x \rangle}. \tag{5.61}$$

Here $\langle \hat{F}_x \rangle$ is the expectation value of $\hat{F}_x$ with the state $\psi_\uparrow(f_z)$, which can also be estimated

in a semiclassical manner

$$\langle \hat{F}_x \rangle \simeq \frac{\int f_x(f_z, p)\, \mathrm{d}t}{T_{\mathrm{cl}}} = \frac{1}{T_{\mathrm{cl}}/2} \int_{f_{z,-}}^{f_{z,+}} \frac{b(f)\cos[p(f)]}{\dot{f}(f)}\, \mathrm{d}f. \tag{5.62}$$

The agreement between the WKB estimation and the exact numerical results for E_{dif} is also compared in Fig. 5.5(c). In the positive energy branch, we can confirm that the occasional degeneracy is also captured by the semiclassical approach via the cosine term in Eq. (5.61), which comes from the damped oscillation wavefunction near $f_z = 0$ [see Eq. (5.59)].

5.5 Bit-flip analysis under dissipation

The tunneling rate itself from the Hamiltonian part is not sufficient to determine the bit-flip rate of the generalized spin-cat qubit. The population that leaves the encoded subspace due to the external noise cannot be brought back with the Hamiltonian evolution only. Therefore, we also need to take the dissipative process into account for the full bit-flip analysis. A similar study for the bosonic Kerr cat was first performed in Ref. [142] and then a more thorough work [164]. Here I will apply these analysis to our generalized spin-cat model. In this section, I assume the $\Omega_{\pm 1}$ drives are both on resonance ($\Delta = 0$), and the F_g-level detuning $\delta = 0$ as well for simplicity. The two SCSs $|\theta, 0\rangle_g$ and $|\theta, \pi\rangle_g$ in the F_g levels still form the dark subspace under dissipation. Otherwise, when $\delta \neq 0$, any eigenstate of $\hat{H}_{\mathrm{DS}}$ may not be the dark state due to its participation in F_e levels. We may need colored dissipation to stabilize a pair of eigenstates in this situation [123, 142]. In the later derivation, I will focus on the $\Omega \gg \Gamma$ regime where the F_e-levels cannot be eliminated directly.

Let us start with a review of the level structures in $\hat{H}_{\mathrm{DS}}$ (5.14) first. When $\Delta = 0$, then as we have discussed in Sec. 5.3.1, the eigen-energies always come with $\pm E_k$ pairs with the corresponding eigenstates $|\psi_k^g\rangle \pm |\psi_k^e\rangle$ respectively. Therefore, we can separate the whole spectrum into the positive branch and the negative branch. Below we denote $|\psi_k\rangle$ as

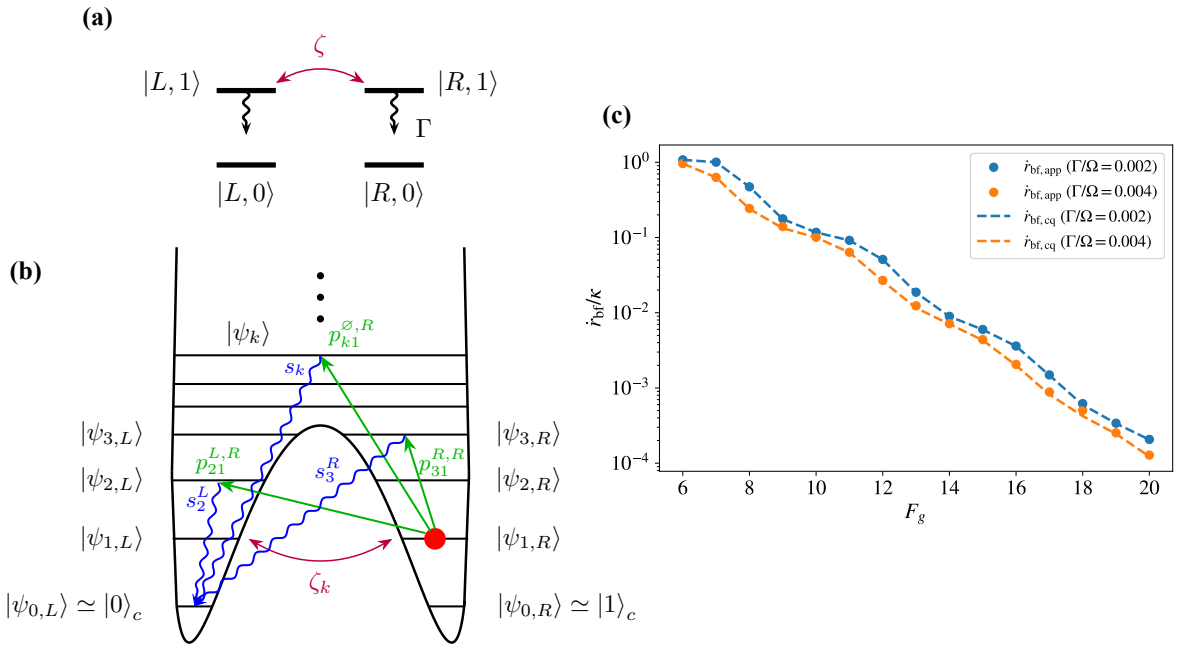

Figure 5.6: Bit-flip analysis with both Hamiltonian tunneling and autonomous stabilization. (a) The simplified 4-level model. (b) Illustration of the bit-flip mechanism. A state at $|\psi_{k,R}\rangle$ may finally come to $|0\rangle_c$ by first jumping to another $|\psi_{l,\iota'}\rangle$ ($|\tilde{\psi}_{l,\iota'}\rangle$) and then coming to $|0\rangle_c$. Or it may tunnel to $|\psi_{k,L}\rangle$ for the first jump. The negative branch ($|\tilde{\psi}_{l,\iota'}\rangle$) is omitted in the plot due to the symmetry. (c) Comparison of bit-flip rate $\dot{r}_{\rm bf}$ achieved from the simplified model (5.80) and the conserved quantities (5.81). Here I fix $\theta = \pi/3$ while varying F_g and Γ/Ω.

eigenstates with positive energy and $|\tilde{\psi}_k\rangle = (\hat{I} - 2\hat{P}_e)\,|\psi_k\rangle$ as the corresponding one with negative energy. In each branch, when E_k is below the barrier, we can find states $|\psi^g_{k,L(R)}\rangle$ in each well with a small tunneling rate ζ_k in between, which has been computed through the WKB analysis. However, when the energy is above the barrier, we may not be able to find pairs of close-to-degenerate levels. The wavefunctions will be supported in both of the wells.

To clarify the notation, we use $|\psi^g_{k,\pm}\rangle$ ($k \leq n_{\rm crit}$) to denote the two near-degenerate level pairs with even/odd parities. Then we denote $\bar{E}_k = \frac{E_{k,+}+E_{k,-}}{2}$ as the average of energies, and $\zeta_k = \frac{E_{k,+}-E_{k,-}}{2}$ is the tunneling rate. We also denote $|\psi_{k,L(R)}\rangle = \frac{|\psi_{k,+}\rangle \pm |\psi_{k,-}\rangle}{\sqrt{2}}$, where $|\psi^g_{k,L(R)}\rangle = \hat{P}_g\,|\psi_{k,L(R)}\rangle = \frac{|\psi^g_{k,+}\rangle \pm |\psi^g_{k,-}\rangle}{\sqrt{2}}$ will mostly locate in one of the double wells. The

two states $|\psi_{k,L(R)}\rangle$ with $k = 0$ are the encoded computational basis states $|0\rangle_c$ and $|1\rangle_c$ [Eq. (5.3)] correspondingly. When $k > n_{\rm crit}$, then each eigenstate $|\psi_k\rangle$ will be treated individually. The Hamiltonian can now be written in the form that

$$\hat{H}_{\rm DS} = \sum_{k\leq n_{\rm crit}} [\bar{E}_k(\hat{I}_k - \hat{\tilde{I}}_k) + \zeta_k(\hat{X}_k - \hat{\tilde{X}}_k)] + \sum_{k>n_{\rm crit}} E_k(|\psi_k\rangle\langle\psi_k| - |\tilde{\psi}_k\rangle\langle\tilde{\psi}_k|), \tag{5.63}$$

where

$$\hat{I}_k = |\psi_{k,L}\rangle\langle\psi_{k,L}| + |\psi_{k,R}\rangle\langle\psi_{k,R}|, \quad \hat{\tilde{I}}_k = |\tilde{\psi}_{k,L}\rangle\langle\tilde{\psi}_{k,L}| + |\tilde{\psi}_{k,R}\rangle\langle\tilde{\psi}_{k,R}|, \tag{5.64}$$

and

$$\hat{X}_k = |\psi_{k,L}\rangle\langle\psi_{k,R}| + |\psi_{k,R}\rangle\langle\psi_{k,L}|, \quad \hat{\tilde{X}}_k = |\tilde{\psi}_{k,L}\rangle\langle\tilde{\psi}_{k,R}| + |\tilde{\psi}_{k,R}\rangle\langle\tilde{\psi}_{k,L}|. \tag{5.65}$$

I also introduce

$$\hat{\tilde{H}}_{\rm DS} = \sum_{k\leq n_{\rm crit}} \bar{E}_k(\hat{I}_k - \hat{\tilde{I}}_k) + \sum_{k>n_{\rm crit}} E_k(|\psi_k\rangle\langle\psi_k| - |\tilde{\psi}_k\rangle\langle\tilde{\psi}_k|), \tag{5.66}$$

which will be used later.

Then let us go to the rotating frame under Hamiltonian $\hat{\tilde{H}}_{\rm DS}$. When $\Omega \gg \Gamma$, then with the rotating wave approximation (RWA), only the population on each eigenstate and the coherence between each level pair $|\psi_{k,L}\rangle, |\psi_{k,R}\rangle$ (and $|\tilde{\psi}_{k,L}\rangle, |\tilde{\psi}_{k,R}\rangle$) when $k \leq n_{\rm crit}$ will contribute to the dynamics.

To get a taste of the interplay between coherent tunneling and the dissipative process, we can first consider a simplified 4-level system that contains $\{|L,0\rangle, |L,1\rangle, |R,0\rangle, |R,1\rangle\}$. There will be a coherent coupling between $|L,1\rangle$ and $|R,1\rangle$ with rate ζ, and $|L(R),1\rangle$ can decay into $|L(R),0\rangle$ at a rate of Γ. The full dynamics can be described with the following

master equation:

$$\frac{\mathrm{d}\hat{\rho}}{\mathrm{d}t} = -i\zeta[(|L,1\rangle\langle R,1| + |R,1\rangle\langle L,1|),\hat{\rho}] + \Gamma\mathcal{D}[|L,0\rangle\langle L,1|]\hat{\rho} + \Gamma\mathcal{D}[|R,0\rangle\langle R,1|]\hat{\rho}. \tag{5.67}$$

Then, if the state is initially in $|R,1\rangle$, then the final population of the state in $|L,0\rangle$ will be

$$p_f = \int_0^{+\infty} e^{-\Gamma t}\sin^2(\zeta t)\Gamma\,\mathrm{d}t = \frac{1}{2}\left(1 - \frac{1}{1+(2\zeta/\Gamma)^2}\right). \tag{5.68}$$

Similarly, the final population at $|R,0\rangle$ will be $(1-p_f)$. We can also understand it as, due to the coherent tunneling between $|R,1\rangle$ and $|L,1\rangle$ the chance that the quantum jump starting at $|L,1\rangle$ will be p_f, and the chance to start at $|R,1\rangle$ will be $(1-p_f)$ respectively.

On the other hand, tunneling itself is not the only mechanism that leads to the bit-flip error. The dissipative operator $\hat{D}_{ge,q}$ may also cause a direct quantum jump from states in one well to states in the other well. Due to the structural similarity between $|\tilde{\psi}_{l,\iota'}\rangle$ and $|\psi_{l,\iota'}\rangle$, we can verify that when $l > 0$ we have

$$|\,\langle\tilde{\psi}_{l,\iota'}|\,\hat{D}_{ge,q}\,|\psi_{k,\iota}\rangle\,|^2 = |\,\langle\psi_{l,\iota'}|\,\hat{D}_{ge,q}\,|\psi_{k,\iota}\rangle\,|^2, \tag{5.69}$$

where $\iota(\iota') \in \{L,R\}$ when $k(l) \le n_{\mathrm{crit}}$ and $\iota(\iota') = \varnothing$ if $k(l) > n_{\mathrm{crit}}$. This means that the possibility for a state $|\psi_{k,\iota}\rangle$ to jump into the corresponding state pair $|\psi_{l,\iota'}\rangle$, $|\tilde{\psi}_{l,\iota'}\rangle$ in the positive and negative branches is equal.

With this, I will use $p_{lk}^{\iota',\iota}$ to denote the proportion that a state $|\psi_{k,\iota}\rangle$ will hop to $|\psi_{l,\iota'}\rangle$ and $|\tilde{\psi}_{l,\iota'}\rangle$ $(l > 0)$ in one round of the jump:

$$p_{lk}^{\iota',\iota} = \frac{\sum_q |\,\langle\psi_{l,\iota'}|\,\hat{D}_{ge,q}\,|\psi_{k,\iota}\rangle\,|^2 + \sum_q |\,\langle\tilde{\psi}_{l,\iota'}|\,\hat{D}_{ge,q}\,|\psi_{k,\iota}\rangle\,|^2}{\sum_q \langle\psi_{k,\iota}|\,\hat{D}_{ge,q}^{\dagger}\hat{D}_{ge,q}\,|\psi_{k,\iota}\rangle} \qquad (k,l>0). \tag{5.70}$$

When $l = 0$, $|\psi_{0,\iota'}\rangle$ will be in the F_g-levels only. Therefore $|\psi_{0,\iota'}\rangle$ and $|\tilde{\psi}_{0,\iota'}\rangle$ are in fact the

same. In this case, we should denote

$$p_{0k}^{\iota',\iota}=\frac{\sum_q|\left\langle\psi_{0,\iota'}\right|\hat{D}_{ge,q}\left|\psi_{k,\iota}\right\rangle|^2}{\sum_q\left\langle\psi_{k,\iota}\right|\hat{D}_{ge,q}^{\dagger}\hat{D}_{ge,q}\left|\psi_{k,\iota}\right\rangle}. \tag{5.71}$$

Finally, since $|\psi_{0,L(R)}\rangle$ are both the dark states, we have $p_{l0}^{\iota',\iota}=\delta_{l0}\delta_{\iota\iota'}$. Notice that all of the proportions should be summed to unity, which indicates the following expression

$$\sum_{0\leq l\leq n_{\text{crit}}}(p_{lk}^{R,\iota}+p_{lk}^{L,\iota})+\sum_{l>n_{\text{crit}}}p_{lk}^{\varnothing,\iota}=1. \tag{5.72}$$

Let us denote s_k^{ι} as the population that the state $|\psi_{k,\iota}\rangle$ (or $|\tilde{\psi}_{k,\iota}\rangle$) will finally end up in $|0\rangle_c$ ($|\psi_{0,L}\rangle$). We should consider what contributes to s_k^{ι} in total, which contains both the coherent tunneling and a direct jump. Take s_k^R as an example. Notice that, due to the tunneling between $|\psi_{k,L}\rangle$ and $|\psi_{k,R}\rangle$, there will be p_f possibility that the first jump happens from $|\psi_{k,L}\rangle$ and $(1-p_f)$ possibility that it happens from $|\psi_{k,R}\rangle$. The state may jump into any possible $|\psi_{l,\iota'}\rangle$ (or $|\tilde{\psi}_{l,\iota'}\rangle$) with the probability determined by $p_{lk}^{\iota',\iota}$. Finally, each $|\psi_{l,\iota'}\rangle$ (or $|\tilde{\psi}_{l,\iota'}\rangle$) will end up with a population $s_l^{\iota'}$ in the $|0\rangle_c$ state.

Given the argument above, we can achieve an iterative expression for s_k^R in the following form:

$$\begin{aligned}s_k^R=&\frac{1}{2}\left[1+\frac{1}{1+(2\zeta_k/\Gamma_k)^2}\right]\cdot\Big[\sum_{0\leq l\leq n_{\text{crit}}}(p_{lk}^{R,R}s_l^R+p_{lk}^{L,R}s_l^L)+\sum_{l>n_{\text{crit}}}p_{lk}^{\varnothing,R}s_l\Big]\\&+\frac{1}{2}\left[1-\frac{1}{1+(2\zeta_k/\Gamma_k)^2}\right]\cdot\Big[\sum_{0\leq l\leq n_{\text{crit}}}(p_{lk}^{R,L}s_l^R+p_{lk}^{L,L}s_l^L)+\sum_{l>n_{\text{crit}}}p_{lk}^{\varnothing,L}s_l\Big].\end{aligned} \tag{5.73}$$

Here we have taken the contribution from both the positive branch and the negative branch into account. Also, $\Gamma_k=\frac{\Gamma_{k,L}+\Gamma_{k,R}}{2}$ is the averaged total rate for quantum jumps starting at

$|\psi_{k,R}\rangle$ or $|\psi_{k,L}\rangle$, where

$$\Gamma_{k,\iota} = \Gamma \sum_q \langle \psi_{k,\iota}| \hat{D}^\dagger_{ge,q} \hat{D}_{ge,q} |\psi_{k,\iota}\rangle \qquad (k > 0). \tag{5.74}$$

When $k = 0$, both $|\psi_{0,R}\rangle$ and $|\psi_{0,L}\rangle$ are the degenerate steady states, and there is no tunneling between them. In this case, we just set $\zeta_0 = 0$.

Due to the symmetry between the two wells, we have $s^R_k = 1 - s^L_k$. We also believe that $p^{R,R}_{lk} = p^{L,L}_{lk}$ and $p^{R,L}_{lk} = p^{L,R}_{lk}$ for the same reason. Meanwhile, it is straightforward that $s^R_0 = 0$ and $s^L_0 = 1$, since s^ι_k is defined as how much the population ends up in $|0\rangle_c$ ($|\psi_{0,L}\rangle$) when the state is initially in $|\psi_{k,\iota}\rangle$, and both $|\psi_{0,L(R)}\rangle$ are steady states.

To further simplify the analysis, we assume $s_l = 1/2$ when $l > n_{\text{crit}}$. This assumption comes as the wavefunction $|\psi^g_l\rangle$ is not solely located in any of the wells. Together with Eq. (5.72), the expression of s^R_k in Eq. (5.73) now can be written as

$$s^R_k - 1/2 = \frac{1}{1 + (2\zeta_k/\Gamma_k)^2} \sum_{0 \le l \le n_{\text{crit}}} (p^{R,R}_{lk} - p^{L,R}_{lk})(s^R_l - 1/2). \tag{5.75}$$

We can also express the above equation in the following compact form

$$\vec{s}' = (DP^{\text{T}}) \cdot \vec{s}', \tag{5.76}$$

where $\vec{s}'$ is an $(n_{\text{crit}} + 1)$ dimensional vector with each element $(s^R_k - 1/2)$, D is a diagonal matrix with $D_{kk} = \frac{1}{1+(2\zeta_k/\Gamma_k)^2}$, and P satisfies $P_{lk} = p^{R,R}_{lk} - p^{L,R}_{lk}$. P here can be understood as the difference between the population transfer matrices that a state stays in one well or goes to another well after a round of quantum jump.

With this, s^R_k can be solved by finding the right eigenvector of (DP^{T}) with the eigenvalue 1. It requires us to further examine the structure of the matrix (DP^{T}). Notice that $P_{00} = 1$

and $P_{l0} = 0$ for $l > 0$. We can write (DP^{T}) as the following form where

$$DP^{\mathrm{T}} = \begin{pmatrix} 1 & 0 \\ 0 & D_{>0} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \vec{w} & P_{>0}^{\mathrm{T}} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ D_{>0}\vec{w} & D_{>0}P_{>0}^{\mathrm{T}} \end{pmatrix}. \tag{5.77}$$

$\vec{w}$ here is an n_{crit}-dimensional vector that $w_k = P_{0k}$. Based on this, $\vec{s}'$ can be achieved as

$$\vec{s}' = -\frac{1}{2} \begin{pmatrix} 1 \\ (1 - D_{>0}P_{>0}^{\mathrm{T}})^{-1} D_{>0}\vec{w} \end{pmatrix}. \tag{5.78}$$

Here $\vec{s}'$ is normalized as $s_0' = -1/2$, which corresponds to $s_0^R = 0$. Therefore, the n_{crit} dimensional $\vec{s}^R$ with elements s_k^R $(1 \leq k \leq n_{\mathrm{crit}})$ satisfies

$$\vec{s}^R = \frac{1}{2}[\vec{1} - (1 - D_{>0}P_{>0}^{\mathrm{T}})^{-1} D_{>0}\vec{w}] = \frac{1}{2} \left[\vec{1} - \sum_{n=0}^{+\infty} (D_{>0}P_{>0}^{\mathrm{T}})^n D_{>0}\vec{w} \right]. \tag{5.79}$$

$\vec{1}$ above is the vector where every element is 1. Remember that the D matrix reflects the coherent tunneling effect while the P matrix is only determined by the quantum jumps. As a result, $(D_{>0}P_{>0}^{\mathrm{T}})^n$ here indicates the higher-order process with the interplay between tunneling and dissipation.

After achieving s_k^R, we can finally start to derive the bit-flip rate in the presence of noise. For example, let us consider the incoherent noise $\kappa\mathcal{D}[\hat{F}_{g,\xi}]$ again. Here we assume $\kappa \ll \Gamma$ so that we can treat the noise perturbatively. Suggested by Eq. (4.17), the leading-order contribution from the noise can be estimated by first acting the noise on the states in the code subspace and then checking how it comes back to the code subspace through autonomous stabilization.

To evaluate the bit-flip rate, we can apply the noise on the state $|1\rangle_c$ and see how the population ends up on $|0\rangle_c$. The (normalized) state will become $\frac{\hat{F}_{g,\xi}|1\rangle_c}{\sqrt{\langle 1|_c \hat{F}_{g,\xi}^\dagger \hat{F}_{g,\xi}|1\rangle_c}}$ after one

jump from the noise, which happens in a rate of $\kappa \langle 1|_c \hat{F}^\dagger_{g,\xi} \hat{F}_{g,\xi} |1\rangle_c$. Since we have also assumed $\Gamma \ll \Omega$, here it will be sufficient to consider the contribution from population on each Hamiltonian eigenstate only. Therefore, we can estimate the increasing rate of $|0\rangle_c$ population through

$$\dot{r}_{\mathrm{bf,app}} = \kappa \Bigg[|\langle 0|_c \hat{F}_{g,\xi} |1\rangle_c|^2 + 2 \sum_{0<l\leq n_{\mathrm{crit}}} (s_l^R |\langle \psi_{l,R}| \hat{F}_{g,\xi} |1\rangle_c|^2 + s_l^L |\langle \psi_{l,L}| \hat{F}_{g,\xi} |1\rangle_c|^2) + \sum_{l>n_{\mathrm{crit}}} |\langle \psi_l| \hat{F}_{g,\xi} |1\rangle_c|^2 \Bigg]. \tag{5.80}$$

I would like to remark that here I implicitly use the fact that $\hat{F}_{g,\xi} |1\rangle_c$ is only in the F_g-manifold so that $|\langle \psi_{l,\iota}| \hat{F}_{g,\xi} |1\rangle_c|^2 = |\langle \tilde{\psi}_{l,\iota}| \hat{F}_{g,\xi} |1\rangle_c|^2$, as well as the assumption that $s_l = \frac{1}{2}$ when $l > n_{\mathrm{crit}}$.

On the other hand, the bit-flip rate in the perturbative regime can also be evaluated through the conserved quantities of the full Lindbladian. As explained in Sec. 4.3, since we can find a 2-dimensional dark Hilbert subspace in the system, there should be 4 linear-independent conserved quantities $\hat{J}_{kl}$ ($k, l \in \{0, 1\}$) which in general can be computed numerically. Further, they can be normalized through $\langle p|_c \hat{J}_{kl} |q\rangle_c = \delta_{pk}\delta_{ql}$. If we again initialize the quantum state in $|1\rangle_c$, then the $|0\rangle_c$ population increasing rate can also be estimated to the leading order as

$$\dot{r}_{\mathrm{bf,cq}} = \kappa \,\mathrm{Tr}\left[\hat{J}^\dagger_{00} \mathcal{D}[\hat{F}_{g,\xi}](|1\rangle_c\langle 1|)\right]. \tag{5.81}$$

Finally, we can compare the bit-flip rate r_{bf} achieved from the simple model (5.80) and that from the conserved quantities (5.81). In both situations, we assume that the dissipative stabilization comes from spontaneous emission so that all 3 polarizations $q = 0, \pm 1$ in the jump operators $\hat{D}_{ge,q}$ are allowed. In this case, $\Gamma_{k,\iota}$ in Eq. (5.74) will always be $\Gamma_{k,\iota} = \frac{\Gamma}{2}$. The noise model we use is still $\kappa\mathcal{D}[\hat{F}_{g,z}]$. In order for the approximated formula (5.80)

to be valid, we assumed $\kappa \ll \Gamma \ll \Omega$ in the numerical study. The tunneling rate ζ_k is achieved from exact diagonalization of the Hamiltonian, which is still less costly compared with diagonalization of the full Lindbladian due to the reduction in the dimension of the matrices.

CHAPTER 6

EFFICIENT WIGNER TOMOGRAPHY WITHIN A TRUNCATED MULTIMODE HILBERT SPACE

6.1 Introduction

Previously, we have discussed several examples of how the Hilbert space truncation helps in both quantum control and error correction aspects. Starting from this chapter, I will consider the problems of quantum state characterization. After the evolution of a quantum system or the execution of a quantum circuit, sometimes we hope to know the full information about the output state. However, it is usually not easy, as the dimension of the full Hilbert space grows exponentially with the number of parties. Meanwhile, the resulting state may not be pure due to the ubiquitous coupling with the environment. For an M-partite quantum system with d levels in each party, we need $(d^{2M}-1)$ number of real parameters to determine a generic mixed state. When M is large, the huge number of parameters will prevent us from any brute-force attempt. Such a challenge could push us to think more about the structures of the quantum states so that the state characterization tasks may become feasible.

In this chapter, I will present a case study where the efficiency of the state characterization comes from the truncation of the multi-party Hilbert space. Specifically, we consider the tomographic task for a quantum state in M resonator modes with bounded maximum total photon number N. The constraint on the maximum total photon number provides a cutoff on the dimension of the relevant Hilbert space, which now reduces to $\binom{M+N}{N}$, a polynomial dependence on the mode number M in the large M limit. Such a state can be prepared using a multimode photon blockade scheme, and with this, the high-fidelity state preparation for the $N=1$ case (i.e., the $|W\rangle$ state) has been experimentally demonstrated [169]. Further, we have observed a hardware-efficient approach so that the overhead to estimate each density matrix element in the Fock basis with ϵ-precision is independent of M. The reduction of the

dimension and the efficiency to estimate each element together result in a poly(M) overhead to reconstruct the full density matrix or estimate the fidelity with a given state up to a certain precision. We name the whole protocol DEMESST, the abbreviation of the Direct Extraction of density Matrix Elements from Subspace Sampling Tomography method.

I will continue the discussion with the following structure. In Sec. 6.2, I will explain the physics of the device we are considering, which contains a multimode cavity dispersively coupled with a transmon. Our aim is to efficiently reconstruct the full density matrix of the multimode state with bounded total photons in this cavity. In Sec. 6.3, I will explain our efficient multimode Wigner tomography scheme with detailed overhead analysis. Later, in Sec. 6.4, I will talk about the experimental protocol we adopt to demonstrate our theoretical proposal with bounded total photon number $N = 1$. Finally, in Sec. 6.5, I will present a numerical simulation result on the overhead requirement to estimate the fidelity of an M-mode $|W\rangle$ state up to a given precision. We also compare it with a traditional method where the dimension cutoff only applies for each mode individually.

6.2 Device setup

To start with, I want to specify the device in our minds. We consider a multimode resonator dispersively coupled with a transmon qubit [169, 170]. The Hamiltonian of the transmon-cavity coupled system can be modeled as

$$\hat{H} = \sum_m \omega_{a,m} \hat{a}_m^\dagger \hat{a}_m + (\omega_q \hat{c}^\dagger \hat{c} - \frac{E_C}{2} \hat{c}^\dagger \hat{c}^\dagger \hat{c} \hat{c}) - \sum_m g_m (\hat{a}_m^\dagger - \hat{a}_m)(\hat{c}^\dagger - \hat{c}). \tag{6.1}$$

Here, $\hat{a}_m$ and $\hat{c}$ represent the cavity modes and the transmon qubit. $\omega_{a,m}$ and ω_q are their resonant frequencies respectively. E_C, the charging energy of the transmon, characterizes the energy shift of the transmon levels due to the nonlinearity. Finally, g is the coupling strength between the transmon and each cavity mode.

In the dispersive coupling regime where the coupling strength g is much less compared with qubit-cavity detuning $|\omega_q - \omega_{a,m}|$ (i.e., $g \ll |\omega_q - \omega_{a,m}|$), The Hamiltonian of the system can be simplified as

$$\begin{aligned}\hat{H} \simeq & \sum_m \tilde{\omega}_{a,m} \hat{\tilde{a}}_m^\dagger \hat{\tilde{a}}_m + \left[\tilde{\omega}_q |e\rangle\langle e| + (2\tilde{\omega}_q - E_C)|f\rangle\langle f|\right] \\ & + \sum_m \hat{\tilde{a}}_m^\dagger \hat{\tilde{a}}_m (\chi_m^{ge} |e\rangle\langle e| + \chi_m^{gf} |f\rangle\langle f|),\end{aligned} \tag{6.2}$$

where the tilde in the equation above indicates that all the modes are hybridized, and I will ignore this notation later for simplicity. The Hamiltonian is achieved by diagonalizing the linear coupling part in Eq. (6.1) and then applying RWA to ignore the off-resonant terms. The transmon is truncated into 3 levels (up to $|f\rangle$) for further convenience. Finally, the induced nonlinear terms (like self- and cross-Kerr) among cavity modes due to the hybridization with transmon are ignored since they are usually smaller than other parameters shown in Eq. (6.2).

The dispersive shift terms χ_m^{ge} and χ_m^{gf} in Eq. (6.2) can be understood as the transition frequencies between levels in the transmon are affected due to the photon occupation in the cavity modes. For example, the occupation of n photons in the m-th mode shifts the transmon $|g\rangle - |e\rangle$ transition frequency by $n \cdot \chi_m^{ge}$ compared with the case that all cavity modes are in the vacuum state. On the other hand, the resonant frequencies of the cavity modes also depend on the transmon state. We will keep using these observations in the design of the hardware control schemes later.

6.3 Efficient Wigner tomography in a multimode truncated subspace: theoretical framework

In this section, I will discuss our theoretical proposal for an efficient tomography (DEMESST) method, which works for the subspace in a multimode system with a bounded total photon

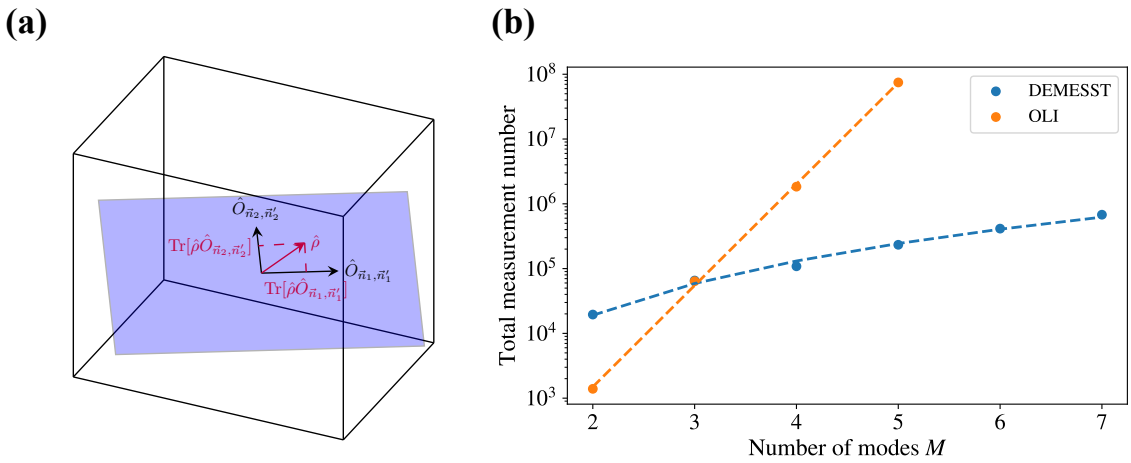

Figure 6.1: Schematic plot on the DEMESST protocol and its performance. (a) The illustration that DEMESST focuses on a truncated subspace while estimating each relevant density matrix element individually. (b) Comparison of the resource overhead to reconstruct $|W\rangle$ state in $N = 1$ subspace, with both DEMESST and OLI method. To achieve desired reconstruction fidelity, the resource required by DEMESST is fitted with poly(M) while that needed from OLI grows exponentially. Details are discussed in Sec. 6.5.

number. The protocol requires us to measure Wigner functions of the unknown quantum state at randomly sampled points. Therefore, I will first discuss the properties of the Wigner function itself and its generalizations that can be measured in practice. Then, I will explain how to estimate each density matrix element under the Fock basis using the Wigner measurement. Together with an intermediate vacuum-state projection trick, we can show that the overhead to estimate each element up to a constant precision will be independent of the total mode number M. Since the dimension of the relevant subspace only scales with poly(M), we can finally show that the total cost for the full density matrix reconstruction is poly(M) as well, which justifies our efficiency claim.

6.3.1 Wigner function and its generalization

Wigner function is one of the most important functions in the field of quantum optics, which serves as the quasiprobability distribution of the state of interest in the phase space. The

Wigner function of a single-mode state ρ is defined as [171]

$$W_\rho(\alpha) = 2\,\mathrm{Tr}[\hat{\rho}\hat{D}(\alpha)e^{i\pi\hat{a}^\dagger\hat{a}}\hat{D}^\dagger(\alpha)], \tag{6.3}$$

where $\hat{D}(\alpha) = e^{\alpha\hat{a}^\dagger - \alpha^*\hat{a}}$ is the displacement operator. Therefore, $W_\rho(\alpha)$ is proportional to the expectation value of the parity operator with the state ρ displaced by amplitude $-\alpha$. Similarly, we can define the Wigner function for any operator $\hat{O}$ with finite Frobenius norm (F-norm for short) $||\hat{O}||_F = \sqrt{\mathrm{Tr}[\hat{O}^\dagger\hat{O}]} < +\infty$ by simply substituting $\hat{\rho}$ with the operator $\hat{O}$ in Eq. (6.3).

The Wigner function of an M-mode state ρ can be simply generalized as

$$W_\rho(\vec{\alpha}) = 2^M\,\mathrm{Tr}[\hat{\rho}\hat{D}(\vec{\alpha})e^{i\pi\sum_{m=1}^M\hat{a}_m^\dagger\hat{a}_m}\hat{D}^\dagger(\vec{\alpha})], \tag{6.4}$$

where $\hat{D}(\vec{\alpha}) = \prod_{m=1}^M\hat{D}_m(\alpha_m) = \prod_{m=1}^M e^{\alpha_m\hat{a}_m^\dagger - \alpha_m^*\hat{a}_m}$ is the displacement operator in the multimode system. However, for our current device setup (6.2), it is difficult to measure joint parity $e^{i\pi\sum_{m=1}^M\hat{a}_m^\dagger\hat{a}_m}$ of relevant modes in a straightforward way. It is because the dispersive cavity-qubit coupling rate χ_m^{ge} are not equal for different modes, and therefore the relevant modes may not be able to pick up a π angle rotation simultaneously when the qubit is in $|e\rangle$ level. To resolve this issue, here we consider a generalized version of the Wigner function introduced in [172]:

$$W_\rho(\vec{\alpha},\vec{\theta}) = \frac{2^M\,\mathrm{Tr}[\hat{\rho}\hat{D}(\vec{\alpha})e^{i\sum_{m=1}^M\theta_m\hat{a}_m^\dagger\hat{a}_m}\hat{D}^\dagger(\vec{\alpha})]}{\prod_{m=1}^M[1+i\cot(\theta_m/2)]}, \tag{6.5}$$

where θ_m may not be equal to π in each mode. We also denote a normalized one

$$\tilde{W}_\rho(\vec{\alpha},\vec{\theta}) = \mathrm{Tr}[\hat{\rho}\hat{D}(\vec{\alpha})e^{i\sum_{m=1}^M\theta_m\hat{a}_m^\dagger\hat{a}_m}\hat{D}^\dagger(\vec{\alpha})], \tag{6.6}$$

where $\tilde{W}_\rho(\vec{\alpha},\vec{\theta})$ can be a complex number now with $|\tilde{W}_\rho(\vec{\alpha},\vec{\theta})| \leq 1$, since $e^{i\sum_{m=1}^M\theta_m\hat{a}_m^\dagger\hat{a}_m}$ is

no longer Hermitian and the modulus of its eigenvalues will always be 1. We can also verify that $\tilde{W}_\rho(\vec{\alpha},\vec{\theta}) = \tilde{W}^*_{\rho^\dagger}(\vec{\alpha},-\vec{\theta})$. As we will show in Sec. 6.4, the generalized $\tilde{W}_\rho(\vec{\alpha},\vec{\theta})$ function is more relevant to what we can measure in experiments. But now, let us keep the theoretical discussion, and in the next part we will demonstrate that the generalized Wigner function of quantum states plays a similar role as the normal Wigner function when estimating the expectation values with the desired observables.

6.3.2 Expectation value estimation with Wigner sampling

In this part, we discuss the Wigner sampling method in detail for estimation of the expectation value with any finite F-norm operator $\hat{O}$, and then analyze the overhead in order to reach a certain accuracy.

To begin with, we recall the expression that reflects the relationship between expectation values and Wigner functions:

$$\mathrm{Tr}[\hat{\rho}\hat{O}] = \int \frac{\mathrm{d}^{2M}\vec{\alpha}}{\pi^M}\, W_\rho(\vec{\alpha}) W_O(\vec{\alpha}). \tag{6.7}$$

Therefore, the overlap between two operators can be computed through integrating the product of their Wigner functions over the phase space. Moreover, a similar expression can be achieved for the generalized Wigner function as well:

$$\mathrm{Tr}[\hat{\rho}\hat{O}] = \int \frac{\mathrm{d}^{2M}\vec{\alpha}}{\pi^M}\, W_\rho(\vec{\alpha},-\vec{\theta}) W_O(\vec{\alpha},\vec{\theta}) = C_M \int \mathrm{d}^{2M}\vec{\alpha}\; \tilde{W}_\rho(\vec{\alpha},-\vec{\theta})\tilde{W}_O(\vec{\alpha},\vec{\theta}), \tag{6.8}$$

where $C_M = \prod_{m=1}^{M}[2(1-\cos\theta_m)/\pi]$ here is a constant factor.

Those equations above have motivated people to develop a Wigner-based approach for direct fidelity estimation between experimentally prepared state $\hat{\rho}$ and the desired pure state

$\hat{\sigma}$ [98]. For example, Eq. (6.8) can also be written as

$$\begin{aligned} F(\hat{\rho},\hat{\sigma}) = \mathrm{Tr}[\hat{\rho}\hat{\sigma}] &= C_M \int \mathrm{d}^{2M}\vec{\alpha}\ |\tilde{W}_\sigma(\vec{\alpha},\vec{\theta})|^2 \frac{\tilde{W}_\rho(\vec{\alpha},-\vec{\theta})}{\tilde{W}_\sigma(\vec{\alpha},-\vec{\theta})} \\ &= \int \mathrm{d}^{2M}\vec{\alpha}\ p_{W^2}(\vec{\alpha}) \frac{\mathrm{Re}[e^{i\phi(\vec{\alpha})}\tilde{W}_\rho(\vec{\alpha},-\vec{\theta})]}{|\tilde{W}_\sigma(\vec{\alpha},-\vec{\theta})|}, \end{aligned} \tag{6.9}$$

where $p_{W^2}(\vec{\alpha}) = C_M|\tilde{W}_\sigma(\vec{\alpha},\vec{\theta})|^2$ is a non-negative function and $\phi(\vec{\alpha}) = \arg\tilde{W}_\sigma(\vec{\alpha},\vec{\theta}) = -\arg\tilde{W}_\sigma(\vec{\alpha},-\vec{\theta})$ is the complex phase of $\tilde{W}_\sigma(\vec{\alpha},\vec{\theta})$. Here $p_{W^2}(\vec{\alpha})$ can serve as a probability distribution function since it satisfies $\int \mathrm{d}^{2M}\vec{\alpha}\ p_{W^2}(\vec{\alpha}) = \mathrm{Tr}[\hat{\sigma}^2] = 1$. We can use it to generate a random set of vectors $\{\vec{\alpha}^{(k)}\}$. We then measure $\mathrm{Re}[e^{i\phi(\vec{\alpha}^{(k)})}\tilde{W}_\rho(\vec{\alpha}^{(k)},-\vec{\theta})]$ at those $\vec{\alpha}^{(k)}$ points and finally calculate the average of $\frac{\mathrm{Re}[e^{i\phi(\vec{\alpha}^{(k)})}\tilde{W}_\rho(\vec{\alpha}^{(k)},-\vec{\theta})]}{|\tilde{W}_\sigma(\vec{\alpha}^{(k)},-\vec{\theta})|}$ to get prediction of $F(\hat{\rho},\hat{\sigma})$. Such a sampling method based on $p_{W^2}(\vec{\alpha}) \propto |\tilde{W}_\sigma(\vec{\alpha},\vec{\theta})|^2$ is named as W^2 method. In Sec. 6.4, we will explicitly show that $\mathrm{Re}[e^{i\phi(\vec{\alpha}^{(k)})}\tilde{W}_\rho(\vec{\alpha}^{(k)},-\vec{\theta})]$ is the expectation value of a Bernoulli random variable $A_j^{(k)} \in \{1,-1\}$ from each measurement on our device. I will denote $e(\vec{\alpha}) := \mathrm{Re}[e^{i\phi(\vec{\alpha})}\tilde{W}_\rho(\vec{\alpha},-\vec{\theta})]$ for future convenience. As a result, $\mathbb{E}_j[A_j^{(k)}] = e(\vec{\alpha}^{(k)})$.

In general, we do have the freedom to choose any probability distribution function $p(\vec{\alpha})$ we want to generate sampling points $\{\vec{\alpha}^{(k)}\}$, but we need to adjust the way of post-processing the data. Since

$$\mathrm{Tr}[\hat{\rho}\hat{O}] = C_M \int \mathrm{d}^{2M}\vec{\alpha}\ p(\vec{\alpha}) \frac{|\tilde{W}_O(\vec{\alpha},\vec{\theta})|}{p(\vec{\alpha})} e(\vec{\alpha}), \tag{6.10}$$

now we need to calculate the average of $C_M \frac{|\tilde{W}_O(\vec{\alpha}^{(k)},\vec{\theta})|}{p(\vec{\alpha}^{(k)})} A_j^{(k)}$ over all sampling points and measurements per point for an estimation of $\mathrm{Tr}[\hat{\rho}\hat{O}]$. However, an optimal choice of $p(\vec{\alpha})$ may exist given specific settings. For example, if we only measure once per sampling vector

$\vec{\alpha}^{(k)}$, the variance of the estimator $\frac{1}{N_{\rm spl}}\sum_k C_M \frac{|\tilde{W}_O(\vec{\alpha}^{(k)},\vec{\theta})|}{p(\vec{\alpha}^{(k)})}A^{(k)}$ will be

$$\begin{aligned}
\text{Var} &= \frac{C_M^2}{N_{\rm spl}}\int \mathrm{d}^{2M}\vec{\alpha}\; p(\vec{\alpha})\frac{|\tilde{W}_O(\vec{\alpha},\vec{\theta})|^2}{p^2(\vec{\alpha})} - \frac{(\mathrm{Tr}[\hat{\rho}\hat{O}])^2}{N_{\rm spl}} \\
&\geq \frac{[C_M\int \mathrm{d}^{2M}\vec{\alpha}\; |\tilde{W}_O(\vec{\alpha},\vec{\theta})|]^2}{N_{\rm spl}\int \mathrm{d}^{2M}\vec{\alpha}\; p(\vec{\alpha})} - \frac{(\mathrm{Tr}[\hat{\rho}\hat{O}])^2}{N_{\rm spl}} \\
&= \frac{1}{N_{\rm spl}}\left[C_M\int \mathrm{d}^{2M}\vec{\alpha}\; |\tilde{W}_O(\vec{\alpha},\vec{\theta})|\right]^2 - \frac{(\mathrm{Tr}[\hat{\rho}\hat{O}])^2}{N_{\rm spl}}.
\end{aligned} \tag{6.11}$$

Here we have used the Cauchy-Schwarz inequality. The minimum variance is only achieved when $p(\vec{\alpha}) \propto |\tilde{W}_O(\vec{\alpha},\vec{\theta})|$, which we call $|W|$ sampling method.

We also point out that, for W^2 method where $p(\vec{\alpha}) = C_M|\tilde{W}_O(\vec{\alpha},\vec{\theta})|^2$, the integral shown in the first line of Eq. (6.11) will be divergent. Therefore, Ref. [98, 173] have proposed to choose a cutoff so that we can discard some sampling vectors $\vec{\alpha}^{(k)}$ that make the denominator of $\frac{A_j^{(k)}}{|\tilde{W}_O(\vec{\alpha}^{(k)},\vec{\theta})|}$ too small. This cutoff procedure may lead to bias in the estimation of $\mathrm{Tr}[\hat{\rho}\hat{O}]$ and makes the error analysis more complicated. The detailed analysis of the cutoff effect under the multimode setting is beyond the scope of our current work.

Instead, we focus on the $|W|$ method. The sampling distribution function satisfies

$$p_{|W|}(\vec{\alpha}) = \frac{|\tilde{W}_O(\vec{\alpha},\vec{\theta})|}{Z_O}, \tag{6.12}$$

with the normalization factor

$$Z_O = \int \mathrm{d}^{2M}\vec{\alpha}\; |\tilde{W}_O(\vec{\alpha},\vec{\theta})|. \tag{6.13}$$

In the $|W|$ method, we need to average $C_M Z_O A_j^{(k)}$ over all sampling vectors $\vec{\alpha}^{(k)}$ and all Bernoulli outcomes $A_j^{(k)}$ per sampling vector. If we sample $N_{\rm spl}$ number of $\vec{\alpha}^{(k)}$ vectors and measure $N_{\rm ms}$ times at each sampling vector, then the variance of the estimator

$C_M Z_O \frac{\sum_{k=1}^{N_{\rm spl}} \sum_{j=1}^{N_{\rm ms}} A_j^{(k)}}{N_{\rm spl} N_{\rm ms}}$ will be

$$\begin{aligned} \mathrm{Var} = & \frac{C_M^2 Z_O^2}{N_{\rm spl}} \Big[\int \mathrm{d}^{2M} \vec{\alpha}\ p_{|W|}(\vec{\alpha}) e^2(\vec{\alpha}) - \frac{(\mathrm{Tr}[\hat{\rho}\hat{O}])^2}{C_M^2 Z_O^2} \Big] \\ & + \frac{C_M^2 Z_O^2}{N_{\rm spl} N_{\rm ms}} \Big[1 - \int \mathrm{d}^{2M} \vec{\alpha}\ p_{|W|}(\vec{\alpha}) e^2(\vec{\alpha}) \Big]. \end{aligned} \tag{6.14}$$

Since $\frac{\mathrm{Tr}[\hat{\rho}\hat{O}]}{C_M Z_O} = \int \mathrm{d}^{2M} \vec{\alpha}\ p_{|W|}(\vec{\alpha}) e(\vec{\alpha})$, we know from the Cauchy-Schwarz inequality that the first line in Eq. (6.14) is no less than zero. Therefore, when fixing the total number of measurements $N_{\rm tot} = N_{\rm spl} N_{\rm ms}$, the choice of $N_{\rm ms} = 1$ will lead to the smallest variance.

From now let us fix $N_{\rm ms} = 1$, where we perform one measurement per $\vec{\alpha}^{(k)}$. We can use Hoeffding's inequality to get a lower bound for the sampling resource $N_{\rm spl}$ in order to satisfy

$$P\left(\left| \frac{C_M Z_O}{N_{\rm spl}} \sum_{k=1}^{N_{\rm spl}} A^{(k)} - \mathrm{Tr}[\hat{\rho}\hat{O}] \right| \geq \epsilon_1 \right) \leq \delta_1, \tag{6.15}$$

which gives

$$N_{\rm spl} \geq \lceil \frac{2 C_M^2 Z_O^2}{\epsilon_1^2} \ln(2/\delta_1) \rceil. \tag{6.16}$$

We can see that in general $N_{\rm spl} \propto (C_M Z_O)^2$. In the next step, we will try to analyze the properties of $C_M Z_O$ for the specific operators of our interest. Especially, we focus on how it depends on the total mode number M. If we can find out an approach to make it independent of M, and if we only have poly(M) number of observables to measure, then the total state reconstruction overhead will still be poly(M).

6.3.3 Density matrix reconstruction

In this part, we discuss our DEMESST scheme for density matrix reconstruction of an unknown state, which is based on the $|W|$ sampling method to estimate each element of the

density operator under the Fock basis. Specifically, we consider a system with M modes and maximum total photon number N among those modes, which restricts the dimension of the Hilbert space for state reconstruction to be $\binom{M+N}{N}$. We focus on the scaling analysis of sampling overhead in terms of mode number M, in the limit where M is much larger than $2N$. We will show that the sampling overhead of our method scales as poly(M) with bounded photon number N, which justifies the efficiency claim as the mode number increases.

Matrix element estimation with M-mode Wigner measurement

Let us consider the expectation value of the unknown state $\hat{\rho}$ with Fock-basis operator $\hat{O}$ that belongs to one of the following:

$$\begin{aligned}
&\hat{O}_{\vec{n},\vec{n}} = |\vec{n}\rangle\langle\vec{n}|, \\
&\hat{O}^R_{\vec{n},\vec{n}'} = \frac{|\vec{n}\rangle\langle\vec{n}'| + |\vec{n}'\rangle\langle\vec{n}|}{\sqrt{2}} \quad (\vec{n} \neq \vec{n}'), \\
&\hat{O}^I_{\vec{n},\vec{n}'} = i\frac{|\vec{n}\rangle\langle\vec{n}'| - |\vec{n}'\rangle\langle\vec{n}|}{\sqrt{2}} \quad (\vec{n} \neq \vec{n}').
\end{aligned} \tag{6.17}$$

Here $|\vec{n}\rangle = \bigotimes_{m=1}^{M} |n_m\rangle$ and $|\vec{n}'\rangle = \bigotimes_{m=1}^{M} |n'_m\rangle$ are multimode Fock basis that satisfy the bounded total photon condition $\sum_{m=1}^{M} n_m \leq N$ and $\sum_{m=1}^{M} n'_m \leq N$. The operators $\hat{O}$ are chosen to be Hermitian $\hat{O} = \hat{O}^\dagger$ and normalized as $\mathrm{Tr}[\hat{O}^\dagger\hat{O}] = 1$. For a quantum system with M modes and maximum total photon number N, we have $\binom{M+N}{N}^2$ number of those operators, which scales as poly(M).

One essential observation is that, in the situation where $M > 2N$, for any $|\vec{n}\rangle\langle\vec{n}'|$ operator there are at least $(M - 2N)$ modes that support the ground state projector $|0\rangle\langle 0|$. We use the set $S_{(\vec{n},\vec{n}')} = \{m | n_m = n'_m = 0\}$ (we use S for short later) to include the indexes of these trivial modes. We also use $\bar{S} = \{1, 2, \ldots, M\} \backslash S$ to denote the set of nontrivial modes. With

this, the operator shown in Eq. (6.17) can always be decomposed as

$$\hat{O} = \Big(\bigotimes_{m\in S} |0\rangle\langle 0|_m \Big) \otimes \hat{O}_{\bar{S}}, \tag{6.18}$$

where $\hat{O}_{\bar{S}}$ is supported on those nontrivial modes with indexes $m \in \bar{S}$. Note that the total number of elements in $\bar{S}$ is no greater than $2N$, and it is independent of M as well.

Similarly, we can decompose the generalized Wigner function of such an operator $\hat{O}$ as

$$\tilde{W}_O(\vec{\alpha}, \vec{\theta}) = \Big(\prod_{m\in S} \tilde{W}_{|0\rangle\langle 0|}(\alpha_m, \theta_m) \Big) \cdot \tilde{W}_{O_{\bar{S}}}(\vec{\alpha}_{\bar{S}}, \vec{\theta}_{\bar{S}}), \tag{6.19}$$

where $\vec{\alpha}_{\bar{S}}$, $\vec{\theta}_{\bar{S}}$ are the restriction of $\vec{\alpha}$, $\vec{\theta}$ in the nontrivial modes $m \in \bar{S}$. Again, $\tilde{W}_{O_{\bar{S}}}(\vec{\alpha}_{\bar{S}}, \vec{\theta}_{\bar{S}})$ does not depend on M.

Now let us consider the sampling overhead to get a precise estimation of $\mathrm{Tr}[\hat{\rho}\hat{O}]$ where $\hat{O}$ satisfies those properties mentioned above. One method is to sample according to the M-mode function $p(\vec{\alpha}) = |\tilde{W}_O(\vec{\alpha}, \vec{\theta})|/Z_O$ directly. According to Eq. (6.16), the key quantity we should focus on is $C_M Z_O$, which satisfies

$$\begin{aligned} C_M Z_O &= \Big[\prod_{m=1}^{M} \frac{2(1-\cos\theta_m)}{\pi} \Big] \int \mathrm{d}^{2M}\vec{\alpha}\ |\tilde{W}_O(\vec{\alpha}, \vec{\theta})| \\ &= \prod_{m\in S} \Big[\frac{2(1-\cos\theta_m)}{\pi} \int \mathrm{d}^2\vec{\alpha}\ |\tilde{W}_{|0\rangle\langle 0|}(\alpha_m, \theta_m)| \Big] \cdot C_{\bar{S}} \int \mathrm{d}^{2|\bar{S}|}\vec{\alpha}_{\bar{S}}\ |\tilde{W}_{O_{\bar{S}}}(\vec{\alpha}_{\bar{S}}, \vec{\theta}_{\bar{S}})| \\ &= 2^{M-|\bar{S}|} C_{\bar{S}} Z_{O_{\bar{S}}}. \end{aligned} \tag{6.20}$$

Here $C_{\bar{S}} = \prod_{m\in\bar{S}}[2(1-\cos\theta_m)/\pi]$. $\hat{O}_{\bar{S}}$ is only supported on at most $2N$ modes, and it is independent of M when $M > 2N$. Therefore, the only M-dependence in $C_M Z_O$ comes from the 2^M factor, which is still unfavorable since it grows exponentially as mode number M increases.

Matrix element estimation with ground state verification

One way to resolve this issue is to consider $\hat{O}_{\bar{S}}$, rather than the full operator $\hat{O}$ itself. We introduce projection operator $\hat{P}_S = \bigotimes_{m\in S}|0\rangle\langle 0|_m$ and denote $\hat{\rho}_{\bar{S}} = \mathrm{Tr}_S[\hat{\rho}\hat{P}_S]$. Here $\mathrm{Tr}_S[\bullet]$ means the partial trace over all modes with indexes $m \in S$. It is easy to show that

$$\mathrm{Tr}[\hat{\rho}\hat{O}] = \mathrm{Tr}[\hat{\rho}_{\bar{S}}\hat{O}_{\bar{S}}], \tag{6.21}$$

where both $\hat{\rho}_{\bar{S}}$ and $\hat{O}_{\bar{S}}$ are both supported on modes in $\bar{S}$ that contains at most $2N$ elements. We can do $|W|$ sampling according to $\hat{O}_{\bar{S}}$ as follows:

$$\mathrm{Tr}[\hat{\rho}_{\bar{S}}\hat{O}_{\bar{S}}] = C_{\bar{S}}Z_{O_{\bar{S}}}\int \mathrm{d}^{2|\bar{S}|}\vec{\alpha}_{\bar{S}}\; p_{|W|}(\vec{\alpha}_{\bar{S}})\; \mathrm{Re}[e^{i\phi(\vec{\alpha}_{\bar{S}})}\tilde{W}_{\rho_{\bar{S}}}(\vec{\alpha}_{\bar{S}}, -\vec{\theta}_{\bar{S}})]. \tag{6.22}$$

In Sec. 6.4, I will explain how to use the $|f\rangle$ level of the transmon qubit in our device to effectively restrict the state of interest as $\hat{\rho}_{\bar{S}}$. Then we can get Bernoulli outcomes $A_j^{(k)} \in \{1, -1\}$ again through transmon readout in the end, whose expectation values are exactly $\mathrm{Re}[e^{i\phi(\vec{\alpha}_{\bar{S}}^{(k)})}\tilde{W}_{\rho_{\bar{S}}}(\vec{\alpha}_{\bar{S}}^{(k)}, -\vec{\theta}_{\bar{S}})]$.

We can also use Hoeffding's inequality to estimate the sampling overhead. If we only measure once per sampling vector $\vec{\alpha}_{\bar{S}}^{(k)}$, we will have

$$P\left(\left|\frac{C_{\bar{S}}Z_{O_{\bar{S}}}}{N_{\mathrm{spl}}}\sum_{k=1}^{N_{\mathrm{spl}}}A^{(k)} - \mathrm{Tr}[\hat{\rho}_{\bar{S}}\hat{O}_{\bar{S}}]\right| \geq \epsilon_2\right) \leq \delta_2, \tag{6.23}$$

when

$$N_{\mathrm{spl}} \geq \lceil \frac{2C_{\bar{S}}^2 Z_{O_{\bar{S}}}^2}{\epsilon_2^2}\ln(2/\delta_2)\rceil. \tag{6.24}$$

It is worth pointing out again that, with bounded N, $C_{\bar{S}}Z_{O_{\bar{S}}}$ is independent of mode number

M when $M > 2N$. Therefore, $N_{\rm spl}$ estimated in (6.24) scales as

$$N_{\rm spl} \sim \mathcal{O}_M\Big(\frac{f(N)}{\epsilon_2^2}\ln(2/\delta_2)\Big), \tag{6.25}$$

where $\mathcal{O}_M$ indicates that we only focus on the scaling over M in the large-M limit, while $f(N) = 2C_{\bar{S}}^2 Z_{O_{\bar{S}}}^2$ is some functions that only depend on N and specific choice of operator in Eq. (6.17). We also introduce $f_{\rm max}(N)$ to represent the maximum $f(N)$ value among all those $\binom{M+N}{N}^2$ operators.

Full state reconstruction overhead

We finally talk about the reconstructed density matrix $\hat{\rho}_{\rm est}$. By performing the expectation value estimation of the unknown state ρ with all those $\binom{M+N}{N}^2$ operators in (6.17), we can achieve

$$P(B) \geq 1 - \binom{M+N}{N}^2 \delta_2 \tag{6.26}$$

with a total number of sampling instances

$$N_{\rm tot} \sim \mathcal{O}_M\left[\binom{M+N}{N}^2 \frac{f_{\rm max}(N)}{\epsilon_2^2}\ln(2/\delta_2)\right], \tag{6.27}$$

where B requires all the conditions below:

$$\begin{aligned}
&|\,\langle\vec{n}|\,(\hat{\rho}_{\rm est}-\hat{\rho})\,|\vec{n}\rangle\,| \leq \epsilon_2,\\
&|{\rm Re}[\langle\vec{n}|\,(\hat{\rho}_{\rm est}-\hat{\rho})\,|\vec{n}'\rangle]| \leq \epsilon_2/\sqrt{2} \quad (\vec{n}\neq\vec{n}'),\\
&|{\rm Im}[\langle\vec{n}|\,(\hat{\rho}_{\rm est}-\hat{\rho})\,|\vec{n}'\rangle]| \leq \epsilon_2/\sqrt{2} \quad (\vec{n}\neq\vec{n}').
\end{aligned} \tag{6.28}$$

If B is satisfied, we will have

$$||\hat{\rho}_{\rm est}-\hat{\rho}||_F = \sqrt{\sum_{\vec{n},\vec{n}'} |\,\langle\vec{n}|\,(\hat{\rho}_{\rm est}-\hat{\rho})\,|\vec{n}'\rangle\,|^2} \leq \binom{M+N}{N}\epsilon_2. \tag{6.29}$$

Also, if $\hat{\rho}$ is a pure state, we will also have

$$\begin{aligned} |F(\hat{\rho}_{\text{est}}, \hat{\rho}) - 1| &= |\operatorname{Tr}[(\hat{\rho}_{\text{est}} - \hat{\rho})\hat{\rho}] \leq \epsilon_2 \cdot \sum_{\vec{n},\vec{n}'} |\langle n| \hat{\rho} |\vec{n}'\rangle| \\ &\leq \epsilon_2 \cdot \sqrt{\sum_{\vec{n},\vec{n}'} |\langle n| \hat{\rho} |\vec{n}'\rangle|^2} \cdot \binom{M+N}{N} \\ &= \epsilon_2 \cdot \sqrt{\operatorname{Tr}[\hat{\rho}^2]} \cdot \binom{M+N}{N} = \binom{M+N}{N} \epsilon_2. \end{aligned} \tag{6.30}$$

In summary, by choosing $\epsilon_2 = \epsilon / \binom{M+N}{N}$ and $\delta_2 = \delta / \binom{M+N}{N}^2$, our sampling method will require the amount of sampling instances

$$N_{\text{tot}} \sim \mathcal{O}_M \left[\binom{M+N}{N}^4 \frac{f_{\max}(N)}{\epsilon^2} \log \left(2 \binom{M+N}{N}^2 / \delta \right) \right] \tag{6.31}$$

to achieve

$$P(||\hat{\rho}_{\text{est}} - \hat{\rho}||_F \leq \epsilon) \geq 1 - \delta, \tag{6.32}$$

even if we only have one qubit measurement per sampling instance. With the same amount of sampling, we can also achieve

$$P(|F(\hat{\rho}_{\text{est}}, \hat{\rho}) - 1| \leq \epsilon) \geq 1 - \delta, \tag{6.33}$$

where $\hat{\rho}$ is a pure state.

6.4 Experimental schemes for demonstration

In this section, we mainly focus on the experimental protocols that allow us to measure the quantities related to the generalized Wigner function $\tilde{W}_\rho(\vec{\alpha}, \vec{\theta})$. We first discuss about the method to estimate $\text{Re}[e^{i\phi(\vec{\alpha})} \tilde{W}_\rho(\vec{\alpha}, -\vec{\theta})]$ for a generic M-mode state $\hat{\rho}$, and then we talk about how to use transmon $|f\rangle$ for subsystem tomography to achieve the Wigner values for

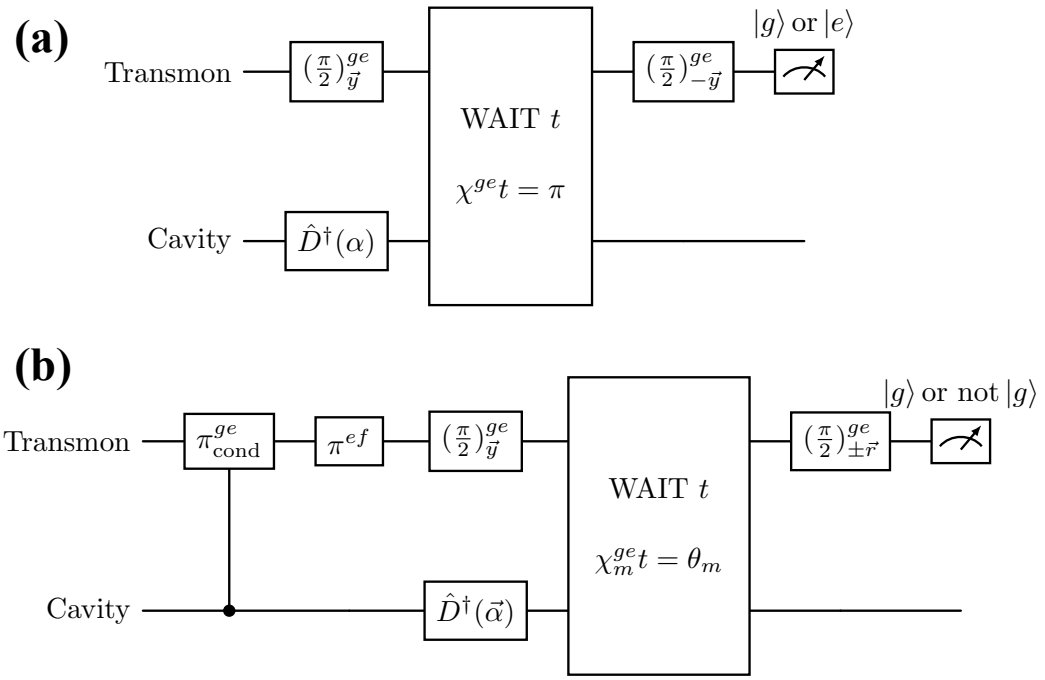

Figure 6.2: Pulse sequences for Wigner measurements. (a) Single-mode Wigner measurement. (b) Multimode generalized Wigner measurement with the $|f\rangle$-level projection trick. In the actual experiment [97], we reverse the order of cavity displacement and the conditioned transmon $|g\rangle$-$|e\rangle$ π-pulses (which should be replaced by a more complicated frequency combs) for a better performance in practice.

projected state $\hat{\rho}_{\bar{S}}$, which is similar to the idea used in Ref. [174]. For simplicity, in later derivation we assume the cavity is initialized as a pure state $|\psi\rangle$, but the argument will apply for a generic density operator $\hat{\rho}$ which can always be decomposed as $\hat{\rho} = \sum_i c_i |\psi_i\rangle\langle\psi_i|$ and understood as an ensemble average from a set of pure states $\{|\psi_i\rangle\langle\psi_i|\}$ with probability c_i.

First, we consider the generalized Wigner function of an M-mode state. The cavity-qubit composite system is initialized as $|\psi\rangle|g\rangle$. In the beginning, we give the qubit a short $\pi/2$ pulse and each mode a short linear drive separately such that the qubit part becomes $\exp(-i\frac{\pi}{4}\hat{\sigma}_y)|g\rangle = \frac{|g\rangle+|e\rangle}{\sqrt{2}}$ and the cavity part becomes $|\psi_D\rangle = \hat{D}(-\vec{\alpha})|\psi\rangle$. Then, we simply wait for time t. Due to the cavity-qubit interaction Hamiltonian $\hat{H}_{\text{int}} = \sum_m \chi_m \hat{a}_m^\dagger \hat{a}_m |e\rangle\langle e|$, the cavity modes will be entangled with the qubit as $\frac{1}{\sqrt{2}}[|\psi_D\rangle|g\rangle + e^{-i\sum_m \theta_m \hat{a}_m^\dagger \hat{a}_m}|\psi_D\rangle|e\rangle]$ where $\theta_m = \chi_m t$. Ideally, we can choose any time t as long as none of the θ_m is an integer

multiple of 2π. However, in practice we prefer those waiting time t to make each θ_m modulo 2π closer to π. Finally, we will give the qubit another $\pi/2$ pulse, but the phase of the pulse will be different from the previous one. Specifically, we consider the qubit rotation along $\vec{r} = -\sin\phi\ \vec{e}_x - \cos\phi\ \vec{e}_y$. By applying this $\exp(-i\frac{\pi}{4}\vec{r}\cdot\hat{\vec{\sigma}})$ operation, finally the cavity-qubit entangled state $|\Psi\rangle$ will be

$$\begin{aligned} |\Psi\rangle = & \frac{|\psi_D\rangle + e^{i\phi}e^{-i\sum_m \theta_m \hat{a}_m^\dagger \hat{a}_m}|\psi_D\rangle}{2}|g\rangle \\ & + \frac{-e^{-i\phi}|\psi_D\rangle + e^{-i\sum_m \theta_m \hat{a}_m^\dagger \hat{a}_m}|\psi_D\rangle}{2}|e\rangle\,. \end{aligned} \tag{6.34}$$

Therefore, when we perform the qubit readout, the final probability of achieving $|g\rangle$ is

$$\begin{aligned} P_g &= \frac{1+\mathrm{Re}\{e^{i\phi}\,\mathrm{Tr}[\hat{D}^\dagger(\vec{\alpha})|\psi\rangle\langle\psi|\hat{D}(\vec{\alpha})e^{-i\sum_m \theta_m \hat{a}_m^\dagger \hat{a}_m}]\}}{2} \\ &= \frac{1}{2}\{1+\mathrm{Re}[e^{i\phi}\tilde{W}_{|\psi\rangle\langle\psi|}(\vec{\alpha},-\vec{\theta})]\}. \end{aligned} \tag{6.35}$$

Therefore, if we record $A = 1$ if we get $|g\rangle$ in qubit readout and $A = -1$ otherwise, then the expectation value of A will be exactly $\mathrm{Re}[e^{i\phi}\tilde{W}_\rho(\vec{\alpha},-\vec{\theta})]$. The derivations above apply for any ϕ, but as we mentioned before, the choice of ϕ depends on the operator $\hat{O}$ and the sampling vector $\vec{\alpha}$.

We finally explain our protocol to measure the generalized Wigner function for a projected state $\hat{\rho}_{\bar{S}}$, which is defined in Sec. 6.3.3. In our experiment (see Ref. [97]), we make use of the multimode blockade technique mentioned in Ref. [169] to make sure the maximally populated total photon number is $N = 1$. For density matrix reconstruction, the Hilbert space under consideration is spanned by $\{|\vec{n}\rangle\,|\,\sum_{m=1}^M n_m \le 1\}$. In this case, the projected operator $\hat{O}_{\bar{S}}$ introduced in Eq. (6.18) is supported on at most 2 modes. Because of the dispersive coupling between the cavity and the qubit, the resonant frequency of the qubit is dependent on the state of the cavity. For example, the qubit frequency will become $\omega_q + \chi_m$ when the cavity is in the Fock state $|1\rangle_m \otimes |\vec{0}\rangle_{\mathrm{els}}$. Therefore, we first give the qubit several long π pulse between

$|g\rangle - |e\rangle$ transition with frequency $\omega_q + \chi_m$ for $m \in S$, such that the qubit that coupled with $(\hat{I} - \hat{P}_S)|\psi\rangle$ cavity part will transfer from $|g\rangle$ to $|e\rangle$, while the part that coupled with $\hat{P}_S|\psi\rangle$ will stay in $|g\rangle$. Then we give the qubit a short π pulse for $|e\rangle - |f\rangle$ transition. After those steps, the cavity-qubit composite state will become

$$|\Psi\rangle = \hat{P}_S|\psi\rangle|g\rangle + (\hat{I} - \hat{P}_S)|\psi\rangle|f\rangle. \tag{6.36}$$

Finally, we can use a similar procedure as we mentioned before when focusing on the Wigner value measurement of a generic M-mode state $\hat{\rho}$. However, we only need to drive those modes with index $m \in \bar{S}$ such that those modes are displaced by $\hat{D}(-\vec{\alpha}_{\bar{S}})$. The short $\pi/2$ pulses with different phases are still applied between the qubit $|g\rangle - |e\rangle$ transition. The probability of receiving a $|g\rangle$ outcome in the final qubit readout is

$$P_{g,\phi} = \frac{\mathrm{Tr}[\hat{\rho}_{\bar{S}}] + \mathrm{Re}[e^{i\phi}\tilde{W}_{\rho_{\bar{S}}}(\vec{\alpha}_{\bar{S}}, -\vec{\theta}_{\bar{S}})]}{2}, \tag{6.37}$$

where $\hat{\rho}_{\bar{S}} = \mathrm{Tr}_S[\hat{P}_S|\psi\rangle\langle\psi|]$ in current situation.

However, if we still assign $A = 1$ if we get $|g\rangle$ and $A = -1$ otherwise, the expectation value of A will not be the desired quantity we want. To solve this issue, recall that we have the freedom to choose the phase of the second qubit $\pi/2$ pulse. If we choose the second qubit $\pi/2$ rotation to be along $-\vec{r}$ instead of $\vec{r}$ (or equivalently choose $(\phi + \pi)$ instead of ϕ, then the $|g\rangle$ outcome probability will be

$$P_{g,(\phi+\pi)} = \frac{\mathrm{Tr}[\hat{\rho}_{\bar{S}}] - \mathrm{Re}[e^{i\phi}\tilde{W}_{\rho_{\bar{S}}}(\vec{\alpha}_{\bar{S}}, -\vec{\theta}_{\bar{S}})]}{2}. \tag{6.38}$$

Therefore, one thing we can do is to generate a random Bernoulli variable s first. There is 50% probability that $s = 1$ and 50% probability that $s = -1$. If we get $s = 1$ we will choose the second qubit $\pi/2$ rotation to be along $\vec{r}$, otherwise it will be $-\vec{r}$ instead. In both cases,

we assign $A = 1$ if the qubit measurement outcome is $|g\rangle$, or we assign $A = -1$ if it is not $|g\rangle$. The expectation value of sA will be exactly $\mathrm{Re}[e^{i\phi}\tilde{W}_{\rho_{\bar{S}}}(\vec{\alpha}_{\bar{S}}, -\vec{\theta}_{\bar{S}})]$. Another merit of this procedure is that it will work even when it is hard to distinguish $|e\rangle$ and $|f\rangle$ levels in qubit readout, as long as we can distinguish $|g\rangle$ outcomes from others.

In our experiments [97], we did not use the trick of random number s generation, since it is not necessary to only perform a single-shot measurement per sampling vector $\vec{\alpha}_{\bar{S}}$. We tried to repeat the experiment 10 times while the rotation of the second qubit $\pi/2$ pulse is along $\vec{r}$, and 10 times while the rotation is along $-\vec{r}$. Finally, we subtract the averaged probability of achieving outcome $|g\rangle$ between the two situations to get an estimation for $\mathrm{Re}[e^{i\phi}\tilde{W}_{\rho_{\bar{S}}}(\vec{\alpha}_{\bar{S}}, -\vec{\theta}_{\bar{S}})]$.

6.5 Numerical simulations on the overhead estimation

In this section, we present simulations of the infidelity and F-norm matrix distance vs. point number for both the DEMESST tomography method and an optimized linear inversion (OLI) method, whose details are explained below. Briefly, in the OLI method we neither apply the assumption on the collective truncation nor implement the ground state verification trick. In the simulations, we use the two methods to reconstruct an ideal W state $|W\rangle = \frac{1}{\sqrt{M}}\sum_{m=1}^{M}|1_m\rangle$, where $|1_m\rangle$ represents the multimode Fock state with a single photon in the m-th mode. For DEMESST, we simulate up to $M = 7$ modes, while for OLI we consider up to $M = 5$ modes. We vary the number of randomly chosen mode displacement vectors $\{\vec{\alpha}\}$, and observe the fidelity and the F-norm distance between the reconstructed density matrices and the ideal one, $|W\rangle\langle W|$. The results are shown in Fig. 6.3. For each set of randomly sampled $\vec{\alpha}$ vectors, we perform the state reconstruction 96 times while modeling readout errors to produce the error bars.

We now discuss how to reconstruct the W state in a single run of the simulation. In DEMESST, we have $(M+1)^2$ basis operators [defined in Eq. (6.17)] for which we want

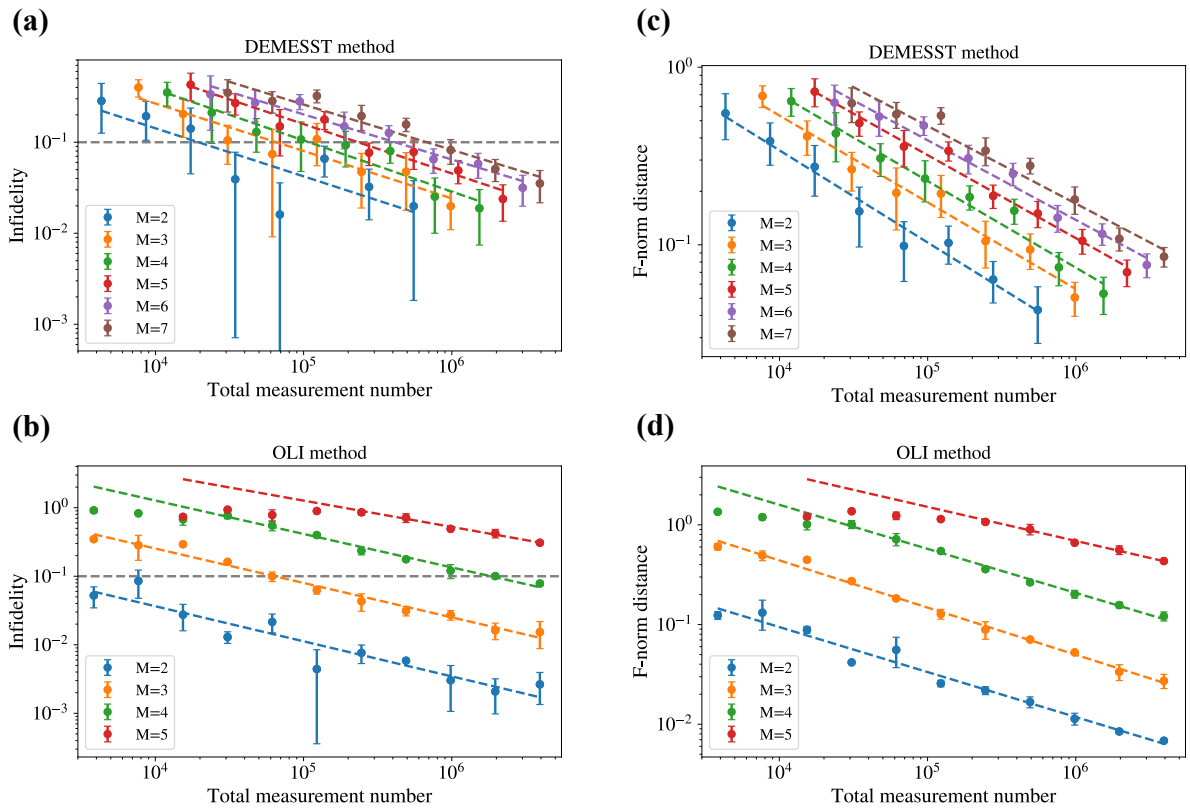

Figure 6.3: Resource overhead simulation to reconstruct $|W\rangle$ state with both DEMESST and OLI methods. (a-b) Simulated reconstruction infidelity with (a) DEMESST and (b) OLI methods, while varying the total measurement number and the mode number M. The gray dashed lines are the references of 0.1 infidelity. (c-d) Simulated F-norm distance between reconstructed density matrix and the ideal one with (c) DEMESST and (d) OLI methods. In all 4 plots, the dependence on the total measurement number x is fitted with a polynomial function $y = ax^b$, where y is either the infidelity or the F-norm distance. We use the fitted lines to estimate the resource requirements shown in Fig. 6.1(b).

to estimate expectation values. Following Eq. (6.21), we project $(M-2)$ modes out of the sampling problem. Then, we only need to sample 2-mode $\vec{\alpha}_{\bar{S}}$ vectors and perform our generalized Wigner measurements in those remaining 2 modes. For each basis operator, we randomly sample 48–6144 $\vec{\alpha}_{\bar{S}}$ vectors according to the probability density function $p(\vec{\alpha}_{\bar{S}})$ [defined in Eq. (6.12)]. For each $\vec{\alpha}_{\bar{S}}$ vector, we simulate a generalized Wigner measurement with 10 averages. As explained in Sec. 6.4, we sample a Bernoulli variable 10 times with probability $P_{g,\phi}$ to get the outcome $A = 1$, and another 10 times with probability $P_{g,(\phi+\pi)}$ to get the outcome $A = 1$ [see Eq. (6.37) and Eq. (6.38)]. We use these results to estimate

the expectation value for the basis operator. Repeating for all basis operators gives us the reconstructed density matrix. Finally, since it is possible for the reconstructed state to be not physical, we further utilize the algorithm introduced in Ref. [175] to find a maximally likely choice of density matrix that satisfies physicality constraints. We calculate the fidelity and F-norm distance of this physical density matrix with respect to the ideal $|W\rangle\langle W|$ state. The results are shown in Fig 6.3(a) and (c). The Total measurement number is $(M+1)^2$ basis operators, times the number of sampled $\vec{\alpha}_{\bar{S}}$ vectors per basis operator, times 10 measurements per $\vec{\alpha}_{\bar{S}}$ vector.

In practice, we have imperfect qubit readout. We model this in the simulation in the following way: if the Bernoulli outcome is $A = 1$ (qubit in $|g\rangle$), the readout voltage V is generated by a normal distribution with mean V_g and variance σ_g^2. If instead $A = -1$ (qubit in $|e\rangle$ or $|f\rangle$), V is generated by a normal distribution with mean V_e and variance σ_e^2 (in our device, $|e\rangle$ and $|f\rangle$ were difficult to distinguish). Then, for each measurement, instead of recording a Bernoulli outcome $A = \pm 1$, we actually record a continuous value $\frac{2V-(V_g+V_e)}{V_g-V_e}$ that is normalized to get a final unbiased estimation. The readout error increases the variance in the Wigner measurement and therefore the total sampling overhead. Based on experimental data, we choose $\frac{\sigma_g}{|V_g-V_e|} = 0.37$ and $\frac{\sigma_e}{|V_g-V_e|} = 0.35$ for our simulations. If we choose $V = \frac{V_g+V_e}{2}$ as the threshold to distinguish $|g\rangle$ or non-$|g\rangle$ outcomes, this can also be understood as the qubit in $|g\rangle$ having an 8% chance to be misidentified as non-$|g\rangle$, while the qubit in $|e\rangle$ or $|f\rangle$ having a 9% chance to be misidentified as $|g\rangle$.

Continuing to the OLI simulations, we follow the protocol introduced in Ref. [169] to reconstruct the state by measuring the real part of generalized M-mode Wigner functions $\mathrm{Re}[\tilde{W}_\rho(\vec{\alpha},\vec{\theta})] = \mathrm{Tr}[\hat{\rho}\hat{D}(\vec{\alpha})\cos(\sum_{m=1}^{M}\theta_m\hat{a}_m^\dagger\hat{a}_m)\hat{D}^\dagger(\vec{\alpha})]$. We reconstruct the state within the 2^M dimensional subspace spanned by $\{|0\rangle, |1\rangle\}^{\otimes M}$ through a matrix inversion described in Ref. [176]. To find a set of displacement vectors $\vec{\alpha}$, for each mode (with given θ_m), we generate a dataset $\{\alpha\}_m$ which produces minimal condition number [177] in single-mode

Wigner tomography. We construct 384–393216 $\vec{\alpha}$ vectors, where each component α_m of the vector is randomly chosen from its corresponding dataset $\{\alpha\}_m$. To simulate the Wigner measurement, we again sample a Bernoulli variable 10 times with probability P_g [defined in Eq. (6.35)] to get $A = 1$ for $\phi = 0$, and another 10 times with P_g for $\phi = \pi$. We also model the readout error as described above. Finally, we use linear inversion to obtain the reconstructed density matrix and apply the algorithm in Ref. [175] to get a physical result. The OLI results for the fidelity and F-norm distance with respect to the ideal $|W\rangle\langle W|$ state are shown in Fig 6.3(b) and (d). The total measurement number is defined as the number of sampled $\vec{\alpha}$ vectors times 10 averages per $\vec{\alpha}$ vector.

Finally, we use the intersection with 0.1 (90% reconstruction fidelity) in Fig 6.3(a) and (b) to generate the values plotted in Fig 6.1(b). We can see that the OLI method requires fewer measurements than DEMESST for 2 modes, but DEMESST has a lower sampling requirement for more than 3 modes. This effect becomes increasingly apparent for larger M. Based on the theoretical results from 6.3.3 and the discussion here, DEMESST scales polynomially with M, while OLI scales exponentially with M. Similarly, in the F-norm distance plots in Fig 6.3(c) and (d), we again see that the DEMESST method performs increasingly more efficiently as the mode number increases.

CHAPTER 7

EFFICIENT LOGICAL FIDELITY ESTIMATION FOR MAGIC STATES WITH $O(1/\epsilon)$ COST

7.1 Introduction

Other than previous chapters where we all have a specific physical system in mind (either resonators or atoms), now let us move to the fault-tolerant (FT) regime. As explained in Chapter 1, to protect the fragile quantum information from the noise, one way is to encode it into a redundant amount of physical qubits under the quantum error correction (QEC) formalism. An error that happened during one physical operation may not directly alter the logical information but create syndromes that can be detected, followed by a recovery map to correct the error.

However, the protection also makes it difficult to perform logical operations arbitrarily. For example, universal logical operations cannot be simply achieved via transversal operations on physical qubits. Based on the choice of the QEC code, some logical operations are relatively easier to implement than others. The FT logical Clifford operations, which map Pauli to Pauli via conjugation, can be implemented transversally on the 2D color code, or through lattice surgery on the 2D surface code. Similarly, FT logical Pauli Z and X measurements are also allowed to be transversal on any Calderbank-Shor-Steane (CSS) code. So, in the FT regime, people usually assume that noiseless Clifford gates and Pauli measurements are free to use, and I will keep this assumption throughout this chapter.

Unfortunately, Clifford gates are not sufficient for universal operations. To solve the problem, one strategy is to prepare a certain type of logical states called magic states, which introduce non-Clifford operations in the circuit via state injection. A number of high-fidelity magic state preparation protocols have been proposed, including the recent resource-efficient proposal named cultivation. The preparation of high-fidelity logical magic states has been

achieved across a variety of platforms and QEC codes. Further, the fidelity of the magic states can be consistently improved via magic state distillation (MSD) [178], which has also been experimentally demonstrated on both physical [179] and recently logical levels [102].

To estimate the performance of the non-Clifford operations in the circuit, one way is to characterize the logical fidelity of the distilled magic states. However, as the infidelity ϵ reduces with the hardware development, it usually requires more overhead to characterize such a small infidelity that the estimation accuracy can be as precise as ϵ itself. One naive thought will be, if the standard deviation (std) of the estimated fidelity is of the order $O(\epsilon)$, we need $N \sim O(1/\epsilon^2)$ copies of magic states if each of the measurement outcomes is independent and identically distributed (i.i.d.). It indeed becomes a problem in a recent experiment demonstrating 5-to-1 MSD with [[7,1,3]] logical qubits, where logical Pauli tomography is used for fidelity characterization [102].

In this chapter, I will provide an alternative protocol we observed where only $N \sim O(1/\epsilon)$ magic states are required for ϵ precision. This protocol consumes two copies at the same time for a Bell measurement. Though seemingly irrelevant, here I would like to point out the deep connection between our work and the underlying merits of superresolution [103], an important concept in quantum metrology that improves the precision of estimating the small distance between two incoherent light sources. Both protocols rely on a modification of the measurement basis of the conventional choice in order to avoid an uninformative fluctuation in the measurement outcomes under the no-signal limit. Further, I will discuss the structure of certain magic states where only one copy is sufficient for fidelity estimation with $O(1/\epsilon)$ cost.

7.2 Efficient logical fidelity estimation for magic $|SH\rangle$ state

To illustrate our protocol, I first start with the $|SH\rangle$ magic state, a state that fits the 5-to-1 distillation scheme [178] for infidelity suppression as recently demonstrated on the

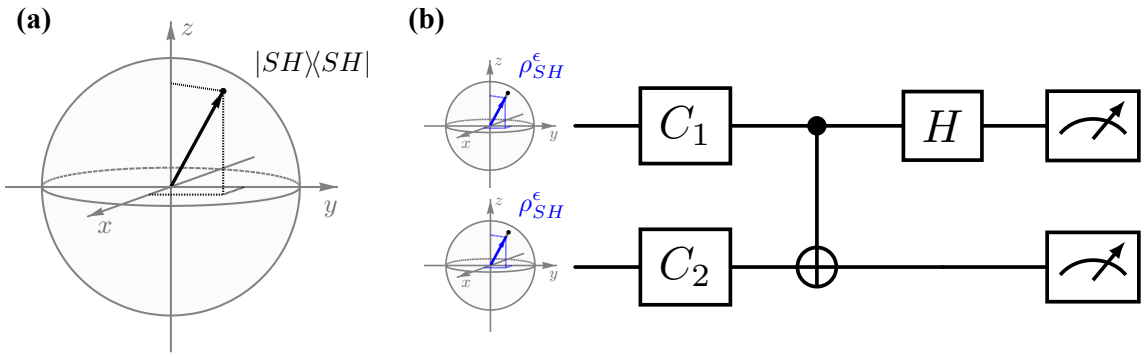

Figure 7.1: Illustration of the efficient fidelity characterization protocol through Bell measurements. (a) Bloch sphere representation for $|SH\rangle$ magic state, with the coordinate $(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$. (b) The Bell measurement circuit. Here C_1 and C_2 are randomly chosen from $G_{SH} = \{I, (SH), (SH)^\dagger\}$ for twirling purpose.

logical level [102]. The corresponding density matrix of the $|SH\rangle = \cos(\theta_{SH}/2)\,|0\rangle + e^{i\pi/4}\sin(\theta_{SH}/2)\,|1\rangle$ state with $\theta_{SH} = \arccos(1/\sqrt{3})$ can be expressed as

$$|SH\rangle\langle SH| = \frac{1}{2}\left(I + \frac{X+Y+Z}{\sqrt{3}}\right), \tag{7.1}$$

which indicates that $|SH\rangle$ is an eigenstate of the Clifford operator $S \cdot H$, since $S \cdot H$ simply permutes the Pauli operators as follows:

$$(SH)\cdot X\cdot (SH)^\dagger = Z, \quad (SH)\cdot Y\cdot (SH)^\dagger = X, \quad (SH)\cdot Z\cdot (SH)^\dagger = Y. \tag{7.2}$$

Now we are given a noisy copy of $|SH\rangle$ state and want to know its small infidelity ϵ ($\epsilon \ll 1$). We first assume that the noisy state is in the depolarization form, with the density matrix

$$\rho_{SH}^{\epsilon} = (1-\epsilon)|SH\rangle\langle SH| + \epsilon|SH^\perp\rangle\langle SH^\perp| = \frac{1}{2}\left[I + \frac{1-2\epsilon}{\sqrt{3}}(X+Y+Z)\right], \tag{7.3}$$

where $|SH^\perp\rangle = \sin(\theta_{SH}/2)\,|0\rangle - e^{i\pi/4}\cos(\theta_{SH}/2)\,|1\rangle$ is orthogonal to $|SH\rangle$. To estimate ϵ, one straightforward approach will be to perform logical Pauli measurements and then

reconstruct the density matrix based on the outcomes. For ρ_{SH}^{ϵ} state, Pauli measurements will always lead to a binomial distribution (± 1 outcomes) with the probability $p_{\pm 1}(\epsilon) = \frac{1}{2} \pm \frac{1-2\epsilon}{2\sqrt{3}}$, regardless of which Pauli operator we choose. The expectation value will be $\mu = \frac{1-2\epsilon}{\sqrt{3}}$, and therefore we can use the measured one ($\hat{\mu}$) to estimate ϵ by $\hat{\epsilon} = \frac{1-\sqrt{3}\hat{\mu}}{2}$. It is easy to calculate the std for the $\hat{\epsilon}$ estimated in this manner, which gives $\mathrm{std}[\hat{\epsilon}] \simeq \frac{1}{\sqrt{2N}}$ (here N is the number of magic state copies). Clearly, to make $\mathrm{std}[\hat{\epsilon}]$ comparable with ϵ, it requires $N \sim O(1/\epsilon^2)$ costs. In fact, we can calculate the classical Fisher information associated with such a binomial distribution, which gives

$$\mathcal{I}(\epsilon) = \frac{1}{p_{+1}(\epsilon)} \left[\frac{\mathrm{d}p_{+1}(\epsilon)}{\mathrm{d}\epsilon}\right]^2 + \frac{1}{p_{-1}(\epsilon)} \left[\frac{\mathrm{d}p_{-1}(\epsilon)}{\mathrm{d}\epsilon}\right]^2 = \frac{2}{1+2\epsilon-2\epsilon^2} \simeq 2. \tag{7.4}$$

This provides a lower bond for the std of any unbiased estimator of ϵ due to the Cramér-Rao bound $\mathrm{std}[\hat{\epsilon}] \geq 1/\sqrt{N \cdot \mathcal{I}(\epsilon)}$, which results in the same $N \sim O(1/\epsilon^2)$ scaling to meet the precision requirement.

Things become different when we have two copies of ρ_{SH}^{ϵ}. With the assumption that Clifford operations are free in the FT regime, now we are allowed to perform the Bell measurement over the two copies [see Fig 7.1(b) for the circuit]. Notice that, $(\rho_{SH}^{\epsilon})^{\otimes 2}$ can be expressed as

$$\begin{aligned}(\rho_{SH}^{\epsilon})^{\otimes 2} = (1-\epsilon)^2 |SH\rangle\langle SH|^{\otimes 2} + \epsilon^2 |SH^{\perp}\rangle\langle SH^{\perp}| \\ + \epsilon(1-\epsilon)\left(|\Psi_{SH}^{+}\rangle\langle\Psi_{SH}^{+}| + |\Psi_{SH}^{-}\rangle\langle\Psi_{SH}^{-}|\right),\end{aligned} \tag{7.5}$$

where $|\Psi_{SH}^{\pm}\rangle = \frac{1}{\sqrt{2}}\left(|SH\rangle\,|SH^{\perp}\rangle \pm |SH^{\perp}\rangle\,|SH\rangle\right)$ is symmetric (antisymmetric) under qubit exchange. Bell measurement, on the other hand, enables us to get the population of the state inside the antisymmetric subspace, as the probability to achieve $(-1,-1)$ outcome is $p_S = \epsilon(1-\epsilon)$. We can estimate ϵ with the proportion of singlet count in all the outcomes ($\hat{p}_S$) by $\hat{\epsilon} = \frac{1-\sqrt{1-4\hat{p}_S}}{2} \simeq \hat{p}_S$, which gives $\mathrm{std}[\hat{\epsilon}] \simeq \sqrt{\frac{p_S}{N/2}} \simeq \sqrt{2\epsilon/N}$ (one Bell measurement

needs two copies of magic states). The $\sqrt{\epsilon}$ scaling in the std leads to a vanishing fluctuation to decide if the state is singlet or triplet in the noiseless limit, which reduces the magic state overhead to $N \sim O(1/\epsilon)$ if we want std$[\hat{\epsilon}]$ to be comparable with ϵ itself.

There are two important features in our protocol that are worth highlighting. The first is that the noisy state ρ_{SH}^{ϵ} is in the depolarization form, which only contains one unknown parameter in the density matrix. This assumption is valid due to the allowed Clifford subgroup twirling. An arbitrary state can be twirled into this form by randomly applying one of the operators in the Clifford subgroup $G_{SH} = \{I, (SH), (SH)^{\dagger}\}$ with equal (1/3) probability. Meanwhile, this twirling process will not affect the fidelity of the noisy copy with respect to $|SH\rangle$, since $|SH\rangle$ is an eigenstate of any element in G_{SH}. Second, there exists a stabilizer state (an eigenstate of n commuting logical Pauli operators, where n is the number of logical qubits) which has no overlap with the noiseless state but $O(\epsilon)$ overlap when the state becomes noisy. Since Clifford operations and Pauli measurements are free in the FT regime, we can get a binomial distribution where the probability of the nontrivial outcome is equal to the overlap itself. This gives the $O(\sqrt{\epsilon})$ dependence in the std for ϵ estimation. In the example here, $|\Psi_{SH}^{-}\rangle = \frac{1}{\sqrt{2}}(|0\rangle|1\rangle - |1\rangle|0\rangle)$ is the desired stabilizer state (eigenstate of $-ZZ$ and $-XX$) that has no overlap with $|SH\rangle^{\otimes 2}$. The probability of the Bell measurement with $(-1,-1)$ outcome is exactly the overlap with $|\Psi_{SH}^{-}\rangle$, which finally results in the efficiency claim for the fidelity estimation protocol we proposed.

7.3 Efficient fidelity estimation for multi-qubit magic states with Bell measurement

Now we expect to generalize our protocol so that the fidelity of a multi-qubit magic state can be efficiently estimated. As discussed, two key ingredients in our protocol are Clifford subgroup twirling and the ability to find a stabilizer state that has $O(\epsilon)$ overlap with the noisy input state. Here, I will elaborate on how these features can be generalized to the

multi-qubit situations. I will first analyze the generic form of a state twirled by a Clifford subgroup, specify the types of magic states that our protocol applies for, and then discuss how to extend the Bell measurement protocol into the multi-qubit cases, which allows us to efficiently estimate the fidelity of the state with at most $O(\epsilon^2)$ bias. In the small ϵ limit, such a bias will be negligible when compared with the expected order $O(\epsilon)$ fluctuations in ϵ estimation.

7.3.1 Clifford subgroup twirling

As has been explained, the Clifford subgroup twirling aims to simplify the noisy form of the density matrix without affecting the overlap with the ideal state. Therefore, for a target magic state, we can choose those Clifford operators such that the magic state is their co-eigenstate. Since every Clifford operator is unitary, the modulus of the corresponding eigenvalue for each operator will always be 1. It is straightforward to verify that the set of allowed operators has the structure of a group, specifically, it is a subgroup of the Clifford group. We define "Clifford subgroup twirling" as randomly picking an element in the Clifford subgroup with equal probability and applying it to the noisy state. Given an arbitrary state ρ, the resulting state ρ_{tw} after twirling by a Clifford subgroup $\mathcal{G}$ is

$$\rho_{\text{tw}} = \frac{1}{|\mathcal{G}|} \sum_{g \in \mathcal{G}} g \rho g^\dagger, \tag{7.6}$$

where $|\mathcal{G}|$ is the the number of elements in $\mathcal{G}$.

We can further use the group representation theory to explore the structure of ρ_{tw}. The full Hilbert space $\mathscr{H}$ can be separated into several invariant (irreducible) subspaces $\mathscr{H}_i$ (i.e., $\mathscr{H} = \bigoplus_i \mathscr{H}_i$), such that a state initially in one subspace will not leave it when acted by any operator in $\mathcal{G}$. Meanwhile, each $\mathscr{H}_i$ cannot be further separated. The operations of the group elements, when restricted in each subspace $\mathscr{H}_i$, form an irreducible representation

of the group $\mathcal{G}$. We then use (α, s) to label these subspaces, where α indicates different inequivalent irreducible representations and s labels different subspaces with the equivalent irreducible representation. We also denote $\Pi_{(\alpha,s)}$ as the projector onto each irreducible subspace $\mathscr{H}_{(\alpha,s)}$.

Given a twirled state ρ_{tw} as in Eq. (7.6), we can show that for any $g \in \mathcal{G}$,

$$g\rho_{\text{tw}} = \frac{1}{|\mathcal{G}|}\sum_{h\in\mathcal{G}}(gh)\rho h^\dagger = \frac{1}{|\mathcal{G}|}\sum_{h'\in\mathcal{G}} h'\rho(g^\dagger h')^\dagger = \rho_{\text{tw}} g, \tag{7.7}$$

which indicates that Schur's lemma in the group representation theory applies here []. Schur's lemma leads to a great simplification for the structure in ρ_{tw}, as

$$\begin{cases} \Pi_{(\alpha,s)}\rho_{\text{tw}}\Pi_{(\alpha,s)} = \lambda_{(\alpha,s)}\Pi_{(\alpha,s)}; \\ \Pi_{(\alpha,s)}\rho_{\text{tw}}\Pi_{(\alpha',s')} = 0 \qquad (\alpha \neq \alpha'). \end{cases} \tag{7.8}$$

The restriction of ρ_{tw} in each subspace $\mathscr{H}_{(\alpha,s)}$ will be proportional to the identity, and the coherence between two inequivalent irreducible sectors will be zero. Therefore, ρ_{tw} can be expressed as

$$\rho_{\text{tw}} = \sum_{(\alpha,s)} \lambda_{(\alpha,s)}\Pi_{(\alpha,s)} + \sum_{\alpha} \rho^{(\alpha)}_{\text{tw,off}}, \tag{7.9}$$

where $\rho^{(\alpha)}_{\text{tw,off}}$ contains off-diagonal terms between different subspaces with equivalent irreducible representations.

Given a quantum state $|\psi\rangle$, we can find the associated Clifford subgroup $\mathcal{G}_\psi$ where $|\psi\rangle$ is a co-eigenstate of every element in $\mathcal{G}_\psi$. This means that the irreducible subspace $\mathscr{H}_{|\psi\rangle}$ containing $|\psi\rangle$ only has dimension 1 ($\mathscr{H}_{(0,0)} = \text{span}\{|\psi\rangle\}$). If we cannot find another $|\psi'\rangle$ such that $|\psi\rangle$ and $|\psi'\rangle$ are both eigenstates with the same eigenvalue for every element in $\mathcal{G}_\psi$, then there is no other equivalent irreducible representation of $\mathcal{G}_\psi$ as its representation on $\mathscr{H}_{(0,0)}$. According to Eq. (7.8), given an arbitrary state ρ_{tw} twirled by $\mathcal{G}_\psi$, there will not

be any coherence between $|\psi\rangle$ and any other states orthogonal to it. In the following, we will only focus on the fidelity estimation with the state $|\psi\rangle$ that meets the above requirement. Under this condition, an arbitrary state twirled by $\mathcal{G}_\psi$ will be

$$\rho_{\psi,\mathrm{tw}} = (1-\epsilon)|\psi\rangle\langle\psi| + \sum_{(\alpha,s),\alpha>0} \frac{\epsilon_{(\alpha,s)}}{\dim[\mathcal{H}_{(\alpha,s)}]}\Pi_{(\alpha,s)} + \sum_{\alpha>0} \rho^{(\alpha)}_{\psi,\mathrm{tw,off}}, \tag{7.10}$$

where the summation in the last two terms covers all (α, s) pairs except $(0, 0)$, which corresponds to the subspace $\mathscr{H}_{(0,0)} = \mathrm{span}\{|\psi\rangle\}$. Moreover, the infidelity ϵ satisfies

$$\epsilon = \sum_{(\alpha,s),\alpha>0} \epsilon_{(\alpha,s)}. \tag{7.11}$$

We should note that not every state satisfies the condition, as it may not be an eigenstate of any Clifford operator other than the trivial identity. However, for the commonly used magic state like the $|T\rangle = \frac{|0\rangle+e^{i\pi/4}|1\rangle}{\sqrt{2}}$ state, the above mentioned $|SH\rangle$ state and several multi-qubit magic states we will show later, we can indeed find a subgroup of Clifford operators to fulfill the requirement. The magic state will be the only co-eigenstate of all elements in the subgroup with the given eigenvalues it corresponds to.

7.3.2 Multi-qubit Bell measurement

As previously explained, the state $|\psi\rangle$ that we want to estimate the fidelity with should satisfy the following condition: starting with an arbitrary noisy state, the resulting state after twirling by the Clifford subgroup $\mathcal{G}_\psi$ associated with $|\psi\rangle$ contains no coherence between $|\psi\rangle$ and any other state orthogonal to it. With the diagonalization of $\rho_{\psi,\mathrm{tw}}$ shown in Eq. (7.10), we have

$$\rho_{\psi,\mathrm{tw}} = (1-\epsilon)|\psi\rangle\langle\psi| + \sum_k \epsilon_k |\psi_k^\perp\rangle\langle\psi_k^\perp|, \tag{7.12}$$

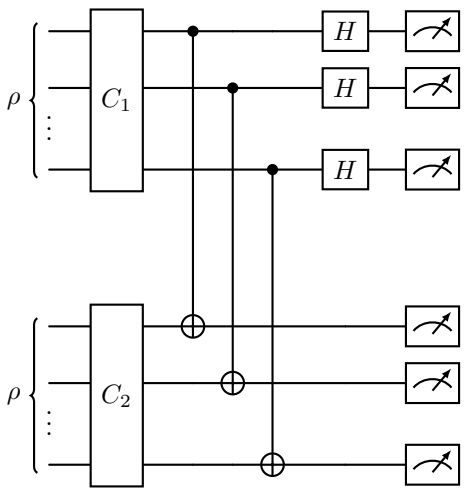

Figure 7.2: Circuit for transversal Bell measurements. Here C_1 and C_2 are n-qubit Clifford gates used for Clifford subgroup twirling.

where $\epsilon = \sum_k \epsilon_k$ is the infidelity of the noisy state and $|\psi_k^\perp\rangle$ are some states orthogonal to $|\psi\rangle$. I will show that, given two identical copies of the noisy $\rho_{\psi,\mathrm{tw}}$, we can still use the Bell measurement scheme to efficiently calibrate ϵ with an unknown bias up to $O(\epsilon^2)$ order.

To illustrate the idea, we consider the population of $\rho_{\psi,\mathrm{tw}}^{\otimes 2}$ in the antisymmetric subspace under the exchange of the two n-qubit states. Similar to Eq. (7.5), we can also write $\rho_{\psi,\mathrm{tw}}^{\otimes 2}$ as

$$\begin{aligned}\rho_{\psi,\mathrm{tw}}^{\otimes 2} = (1-\epsilon)^2 |\psi\rangle\langle\psi|^{\otimes 2} + \sum_k \epsilon_k (1-\epsilon) \left(|\Psi_k^+\rangle\langle\Psi_k^+| + |\Psi_k^-\rangle\langle\Psi_k^-| \right) \\ + \sum_k \epsilon_k^2 |\psi_k^\perp\rangle\langle\psi_k^\perp| + \sum_{k<l} \epsilon_k \epsilon_l \left(|\Psi_{k,l}^+\rangle\langle\Psi_{k,l}^+| + |\Psi_{k,l}^-\rangle\langle\Psi_{k,l}^-| \right),\end{aligned} \tag{7.13}$$

where

$$\begin{cases} |\Psi_k^\pm\rangle = \dfrac{1}{\sqrt{2}} \left(|\psi\rangle\, |\psi_k^\perp\rangle \pm |\psi_k^\perp\rangle\, |\psi\rangle \right), \\ |\Psi_{k,l}^\pm\rangle = \dfrac{1}{\sqrt{2}} \left(|\psi_k^\perp\rangle\, |\psi_l^\perp\rangle \pm |\psi_l^\perp\rangle\, |\psi_k^\perp\rangle \right). \end{cases} \tag{7.14}$$

Clearly, $|\Psi_k^+\rangle$ and $|\Psi_{k,l}^+\rangle$ are symmetric under exchange while $|\Psi_k^-\rangle$ and $|\Psi_{k,l}^-\rangle$ are antisym-

metric. The population of $\rho_{\psi,\mathrm{tw}}^{\otimes 2}$ in the antisymmetric subspace can be computed as

$$p_A = \sum_k \epsilon_k(1-\epsilon) + \sum_{k<l} \epsilon_k \epsilon_l = \epsilon - \epsilon^2 + \sum_{k<l} \epsilon_k \epsilon_l = \epsilon - \frac{\epsilon^2 + \sum_k \epsilon_k^2}{2}. \tag{7.15}$$

Therefore, p_A can be bounded as $\epsilon - \epsilon^2 \leq p_A < \epsilon - \epsilon^2/2$. Even if we use p_A itself as an estimator for ϵ, the bias will be at most ϵ^2, which will be negligible when the infidelity ϵ itself is small.

The next question will be how to measure the population of $\rho_{\psi,\mathrm{tw}}^{\otimes 2}$ in the antisymmetric subspace. This can be achieved with transversal Bell measurements on each pair of qubits (see Fig 7.2) and counting the singlet outcomes. If we get an even (odd) number of singlet outcomes, then the full state is projected into the symmetric (antisymmetric) subspace. This can be understood as, since the swap operation between two n-qubit states SWAP_n can be decomposed as $\mathrm{SWAP}_n = \mathrm{SWAP}^{\otimes n}$, the even (odd) count of the singlet outcomes indicates the full projected state is the eigenstate of the SWAP_n operator with eigenvalue $+1$ (-1). Therefore, the estimation of the antisymmetric subspace population p_A is equivalent to the estimation of the probability of getting an odd number of singlet outcomes in the transversal Bell measurements.

The discussions above allow us to summarize our protocols into the following theorem, with the proof deferred to the Appendix E.

Theorem 7.1. *Given a state $|\psi\rangle$ that can be twirled into the form of Eq. (7.10) via the associated Clifford subgroup, and a constant factor $0 < r < 1$. When infidelity ϵ of the noisy state is smaller than a constant, then by performing the Bell measurement on two copies of twirled states, it only requires $N \sim O[\frac{1}{\epsilon}\log(1/\delta)]$ total copies such that the probability of the deviation between ϵ and a specific estimator $\hat{\epsilon}$ satisfies $P(|\hat{\epsilon} - \epsilon| > r\epsilon) < \delta$.*

7.4 Efficient protocols with one-copy magic state

7.4.1 Generic formalism

In the previous section, I have discussed our protocol to efficiently characterize the fidelity of the magic states with the Bell measurement, provided the states can be twirled into the form of Eq. (7.10) using a Clifford subgroup. This protocol requires manipulating two copies of the states at the same time. However, when it comes to a multi-qubit magic state, sometimes a single copy at one time is sufficient for the $O(1/\epsilon)$ efficiency claim.

To explain the idea, we can take a look at the twirled noisy state Eq. (7.10) again. For an irreducible subspace $\mathcal{H}_{(\alpha,s)}$, if we can find a stabilizer state $|S_{(\alpha,s)}\rangle$ inside it, then the small overlap $\frac{\epsilon_{(\alpha,s)}}{\dim[\mathcal{H}_{(\alpha,s)}]}$ between the twirled state $\rho_{\psi,\text{tw}}$ and $|S_{(\alpha,s)}\rangle\langle S_{(\alpha,s)}|$ can be efficiently characterized through a Clifford circuit. This is because for every stabilizer state we can use Clifford operations to convert it into $|0\rangle^{\otimes n}$, and then the Pauli Z measurement will be sufficient to get the overlap. Such an $O[\epsilon_{(\alpha,s)}]$ expectation value also implies an $O[\sqrt{\epsilon_{(\alpha,s)}}]$ std in the estimation rather than $O(1)$, so that our efficiency argument in previous sections also applies here. Moreover, if every $\epsilon_{(\alpha,s)}$ can be characterized efficiently, then we can compute the total infidelity ϵ since it is the summation of all the $\epsilon_{(\alpha,s)}$. The total sampling cost also needs to be summed together; however, we may still make the efficiency claim on ϵ characterization since the overhead to estimate each $\epsilon_{(\alpha,s)}$ is efficient. To make it rigorous, we formulate this intuition into the following theorem, while deferring the proof to Appendix E.

Theorem 7.2. *Given a state $|\psi\rangle$ that can be twirled into the form of Eq.* (7.10) *via the associated Clifford subgroup, and a constant factor $0 < r < 1$. If for every irreducible subspace $\mathcal{H}_{(\alpha,s)}$ (other than $\mathcal{H}_{(0,0)} = \text{span}\{|\psi\rangle\}$) we can find a stabilizer state $|S_{(\alpha,s)}\rangle$ in it, then we can use $N \sim O[\frac{1}{\epsilon}\log(1/\delta)]$ total copies such that the infidelity ϵ can be approximated by an estimator $\hat{\epsilon}$ with the probability $P(|\hat{\epsilon} - \epsilon| > r\epsilon) < \delta$, and each time only one copy of noisy state is consumed.*

Although the condition that every $\mathcal{H}_{(\alpha,s)}$ contains a stabilizer state looks demanding, below I will show that both the two-qubit $|\mathrm{CZ}\rangle$ and the three-qubit $|\mathrm{CCZ}\rangle$ magic states meet the requirement and therefore support the efficient characterization of their infidelities. Before diving into those specific examples, I will also discuss another perspective on the Clifford subgroup twirling, which will be helpful in the calculation of a specific twirled magic state.

7.4.2 Clifford subgroup twirling: another perspective

In Sec. 7.3.1, we have used Schur's lemma to investigate the generic structure of a state twirled by a Clifford subgroup. Now I will try to revisit the Clifford subgroup twirling from the viewpoint of how Clifford maps Pauli operators to Pauli operators via conjugation (i.e., maps a Pauli operator P to $CPC^\dagger$ where C is a Clifford operator). As I will show later, the new perspective provides a simple way to calculate the forms of the twirled states for specific examples.

I start with the discussion on how the Clifford operations act on the Hermitian Pauli operators via conjugation. The set of n-qubit Hermitian Pauli operators $\mathcal{P}_n^H$ is constituted by the tensor product of single-qubit Pauli $\{I, X, Y, Z\}$, with an overall prefactor ± 1. We notice that, first, the Hermiticity of those Pauli operators is preserved under the map. Second, since the Clifford operators are invertible, we cannot map two different Pauli operators to the same one. As a result, the action of a Clifford operation via conjugation only permutes the elements in $\mathcal{P}_n^H$.

A Clifford subgroup $\mathcal{G}$ can lead to a partition of $\mathcal{P}_n^H$ into several equivalent classes. Given $P, P' \in \mathcal{P}_n^H$, we define them to be equivalent if and only if there exists $g \in \mathcal{G}$ such that $P' = gPg^\dagger$. We denote this equivalence relation as $P \sim P'$. Clearly, $\mathcal{P}_n^H$ can be written as the union of several disjoint equivalent classes $\mathcal{K}_\alpha$ (i.e., $\mathcal{P}_n^H = \bigcup_\alpha \mathcal{K}_\alpha$), where all the elements in each $\mathcal{K}_\alpha$ are equivalent to each other under $\mathcal{G}$. Finally, since each $\mathcal{K}_\alpha$ is closed

by the action of any $g \in \mathcal{G}$ via conjugation, such an action simply results in a permutation of elements in each $\mathcal{K}_\alpha$.

Now we focus on the effects of twirling. For $P_1, P_2 \in \mathcal{K}_\alpha$, there exists $h \in \mathcal{G}$ such that $P_2 = hP_1h^\dagger$. This leads to

$$\frac{\sum_{g\in\mathcal{G}} gP_2g^\dagger}{|\mathcal{G}|} = \frac{\sum_{g\in\mathcal{G}}(gh)P_1(gh)^\dagger}{|\mathcal{G}|} = \frac{\sum_{g\in\mathcal{G}} gP_1g^\dagger}{|\mathcal{G}|} = \frac{1}{|\mathcal{G}|}\sum_{g\in\mathcal{G}} g\frac{\sum_{P\in\mathcal{K}_\alpha} P}{|\mathcal{K}_\alpha|}g^\dagger = \frac{\sum_{P\in\mathcal{K}_\alpha} P}{|\mathcal{K}_\alpha|}, \tag{7.16}$$

where $|\mathcal{K}_\alpha|$ is the number of elements in $|\mathcal{K}_\alpha|$. The last equal sign comes as the action of g via conjugation permutes the elements in $\mathcal{K}_\alpha$. Since arbitrary ρ can be expressed as a linear superposition of Pauli operators, the twirled state ρ_{tw} will have the following simplified form that

$$\rho_{\text{tw}} = \frac{\sum_{g\in\mathcal{G}} g\rho g^\dagger}{|\mathcal{G}|} = \sum_\alpha c_\alpha \frac{\sum_{P\in\mathcal{K}_\alpha} P}{|\mathcal{K}_\alpha|}, \tag{7.17}$$

which is now a linear combination of $(\sum_{P\in\mathcal{K}_\alpha} P)$, the summation of every element in the same class. I will use this property to investigate the generic form of the twirled states when estimating the fidelity with magic $|\text{CZ}\rangle$ and $|\text{CCZ}\rangle$ states below.

7.4.3 Case study: magic $|\text{CZ}\rangle$ state

We first focus on the fidelity estimation of the $|\text{CZ}\rangle$ logical magic state, which is used by IBM for demonstrating the break-even fidelity of the prepared magic state encoded in the [[4,1,2]] error-detecting code [180]. The $|\text{CZ}\rangle$ state is defined as

$$|\text{CZ}\rangle = \frac{|00\rangle + |01\rangle + |10\rangle}{\sqrt{3}}, \tag{7.18}$$

which is an eigenstate of the Clifford operator CZ with eigenvalue 1. We can easily verify that it is also an eigenstate of both Clifford operators SWAP and $X_2 \cdot \text{CX}_{1,2}$ (I denote qubit j as the control and k as the target with the notation $\text{CX}_{j,k}$). Those operators generate a

Clifford subgroup $\mathcal{G}_{\rm CZ} = \langle {\rm CZ}, {\rm SWAP}, X_2 \cdot {\rm CX}_{1,2} \rangle$ where the $|{\rm CZ}\rangle$ state is a co-eigenstate of every element in $\mathcal{G}_{\rm CZ}$. We can use this subgroup to twirl the noisy state without affecting its overlap with the pure $|{\rm CZ}\rangle$.

To investigate the generic form of a state twirled by the group $\mathcal{G}_{\rm CZ}$, as explained previously, we consider how $\mathcal{G}_{\rm CZ}$ leads to a partition of all 2-qubit Hermitian Pauli operators into different equivalent classes. Here I list them as

$$\begin{aligned}
&\mathcal{K}^{\rm CZ}_{0,+} = \{II\}, \quad \mathcal{K}^{\rm CZ}_{1,+} = \{IZ, ZI, -ZZ\}, \quad \mathcal{K}^{\rm CZ}_{2,+} = \{IX, XI, XZ, ZX, XX, YY\}; \\
&\mathcal{K}^{\rm CZ}_{m,-} = \{-P | P \in \mathcal{K}^{\rm CZ}_{m,+}\} \quad (m = 0, 1, 2); \\
&\mathcal{K}^{\rm CZ}_{3} = \{\pm YI, \pm IY, \pm XY, \pm YX, \pm YZ, \pm ZY\}.
\end{aligned} \tag{7.19}$$

It is easy to check that $\sum_{P \in \mathcal{K}^{\rm CZ}_{m,+}} P = -\sum_{P \in \mathcal{K}^{\rm CZ}_{m,-}} P$ for $m = 0, 1, 2$ and $\sum_{P \in \mathcal{K}^{\rm CZ}_{3}} P = 0$. According to Eq. (7.17), any 2-qubit state twirled by $\mathcal{G}_{\rm CZ}$ can be written as

$$\rho_{\rm CZ,tw} = \frac{\sum_{g \in \mathcal{G}_{\rm CZ}} g\rho g^\dagger}{|\mathcal{G}_{\rm CZ}|} = \frac{1}{4} I + \frac{c_1}{3} \sum_{P \in \mathcal{K}^{\rm CZ}_{1,+}} P + \frac{c_2}{6} \sum_{P \in \mathcal{K}^{\rm CZ}_{2,+}} P. \tag{7.20}$$

With the substitution that $c_1 = \frac{1}{4} - \epsilon_1$ and $c_2 = 1 - \epsilon_1 - \frac{3\epsilon_2}{2}$, $\rho_{\rm CZ,tw}$ now has the form that

$$\rho_{\rm CZ,tw} = (1 - \epsilon_1 - \epsilon_2) |{\rm CZ}\rangle\langle{\rm CZ}| + \epsilon_1 |11\rangle\langle 11| + \epsilon_2 \frac{I - |{\rm CZ}\rangle\langle{\rm CZ}| - |11\rangle\langle 11|}{2}, \tag{7.21}$$

which has the desired structure as suggested by Eq. (7.10). The infidelity of the twirled state is $\epsilon = \epsilon_1 + \epsilon_2$.

Finally, both ϵ_1 and ϵ_2 can be efficiently calibrated by measuring the overlap with certain stabilizer states. To estimate ϵ_1, we measure the overlap with the $|11\rangle$ state. This can be achieved by performing logical Z measurements on both qubits and counting the $(-1, -1)$ outcomes. The probability of such an outcome is exactly ϵ_1 itself. Similarly, we can perform the Bell measurement to get the overlap with the singlet state $|\Psi^-\rangle = \frac{1}{\sqrt{2}}(|0\rangle |1\rangle - |1\rangle |0\rangle)$.

The probability of this outcome will be $\epsilon_2/2$. The conditions for Theorem 7.2 are satisfied, and therefore we can consume one copy of $|\mathrm{CZ}\rangle$ magic state each time to efficiently calibrate its logical fidelity with a total $O(1/\epsilon)$ overhead.

7.4.4 Case study: magic $|\mathrm{CCZ}\rangle$ state

$|\mathrm{CCZ}\rangle$ state is another multi-qubit magic state that is widely considered as the non-Clifford resource in the FT quantum computing. Defined as

$$|\mathrm{CCZ}\rangle = \mathrm{CCZ}\,|+\rangle^{\otimes 3} = \frac{1}{2}(|0\rangle\,|0\rangle\,|+\rangle + |0\rangle\,|1\rangle\,|+\rangle + |1\rangle\,|0\rangle\,|+\rangle + |1\rangle\,|1\rangle\,|-\rangle), \tag{7.22}$$

the $|\mathrm{CCZ}\rangle$ state can be converted into two $|T\rangle$ magic states given another $|T\rangle$ catalyst [181]. To find the efficient protocols for $|\mathrm{CCZ}\rangle$ fidelity estimation, again we first look for the generic form of a noisy $|\mathrm{CCZ}\rangle$ state after Clifford subgroup twirling, and then discuss the choice of the stabilizer states which lead to the efficiency claim by measuring the overlap with them.

To begin with, we need to find the Clifford operators for which $|\mathrm{CCZ}\rangle$ is an eigenstate. Since $|+\rangle^{\otimes 3}$ is an eigenstate for both X_i on any qubit and $\mathrm{CX}_{j,k}$ on any two qubits, $|\mathrm{CCZ}\rangle$ is an eigenstate for both $\mathrm{CCZ}\cdot X_i\cdot\mathrm{CCZ}$ and $\mathrm{CCZ}\cdot\mathrm{CX}_{j,k}\cdot\mathrm{CCZ}$. We can further verify that these two operators are still Clifford, since

$$\begin{cases} \mathrm{CCZ}\cdot X_1\cdot\mathrm{CCZ} = X_1\cdot\mathrm{CZ}_{2,3}, \\ \mathrm{CCZ}\cdot\mathrm{CX}_{1,2}\cdot\mathrm{CCZ} = \mathrm{CZ}_{1,3}\cdot\mathrm{CX}_{1,2}, \end{cases} \tag{7.23}$$

and the CCZ operation is permutation-invariant. We will use the Clifford subgroup $\mathcal{G}_{\mathrm{CCZ}} = \langle \mathrm{CCZ}\cdot X_i\cdot\mathrm{CCZ}, \mathrm{CCZ}\cdot\mathrm{CX}_{j,k}\cdot\mathrm{CCZ}\rangle$ (for $i,j,k\in\{1,2,3\}$) to twirl the noisy state.

Next, we need to find the generic form of a noisy state after $\mathcal{G}_{\mathrm{CCZ}}$ twirling. To simply the discussion, I will first present the result from twirling by $\mathcal{G}_{XXX} = \langle X_i, \mathrm{CX}_{j,k}\rangle$ (for $i,j,k\in\{1,2,3\}$) first, with is only different from our desired solution by a CCZ conjugation.

$\mathcal{G}_{XXX}$ leaves a partition of the 3-qubit Hermitian Pauli operators in the following way. First, III itself forms an equivalent class $\mathcal{K}_{0,+}^{XXX} = \{III\}$. Then, since $\mathrm{CX}_{j,k} \in \mathcal{G}_{XXX}$, all the tensor product of Pauli X (and identity I) with the overall prefactor $+1$ are in the same equivalent class $\mathcal{K}_{1,+}^{XXX} = \{X_1, X_2, X_3, X_1X_2, X_1X_3, X_2X_3, X_1X_2X_3\}$. Finally, if a Pauli operator P contains Y_i or Z_i on any qubit, then $-P$ and P are in the same equivalent class due to the action of $X_i \in \mathcal{G}_{XXX}$ via conjugation. The summation of every element in that class will be zero. Therefore, as indicated by Eq. (7.17), any 3-qubit state twirled by $\mathcal{G}_{XXX}$ becomes

$$\rho_{XXX,\mathrm{tw}} = \frac{1}{8}I + \frac{c_1}{7} \sum_{P \in \mathcal{K}_{1,+}^{XXX}} P. \tag{7.24}$$

With the substitution $c_1 = \frac{7}{8} - \epsilon$, we have

$$\rho_{XXX,\mathrm{tw}} = (1-\epsilon)|+++\rangle\langle+++| + \epsilon\frac{I - |+++\rangle\langle+++|}{2^3 - 1}. \tag{7.25}$$

This is equivalent to say that, for any noisy $|\mathrm{CCZ}\rangle$ state twirled by $\mathcal{G}_{\mathrm{CCZ}}$, it can be written as

$$\rho_{\mathrm{CCZ,tw}} = (1-\epsilon)|\mathrm{CCZ}\rangle\langle\mathrm{CCZ}| + \epsilon\frac{I - |\mathrm{CCZ}\rangle\langle\mathrm{CCZ}|}{2^3 - 1}, \tag{7.26}$$

where ϵ is the infidelity that we need to estimate.

Finally, we need to find the desired stabilizer states that have no overlap with the noiseless $|\mathrm{CCZ}\rangle$ state. From Eq. (7.22), we can clearly see that $|0\rangle|0\rangle|-\rangle$, $|0\rangle|1\rangle|-\rangle$, $|1\rangle|0\rangle|-\rangle$ and $|1\rangle|1\rangle|+\rangle$ meet the requirements, which correspond to 4 different outcomes from the Pauli ZZX measurement. The total probability to get any one of them in the ZZX measurement is $\frac{4\epsilon}{7}$ given a noisy $\rho_{\mathrm{CCZ,tw}}$ state, and, as analyzed before, we can use one copy at each time to efficiently estimate ϵ with $O(1/\epsilon)$ overhead in order to achieve a precision comparable with ϵ itself.

7.5 Numerical study with realistic experimental setup

In the sections above, I have mainly focused on the formalism of our efficient fidelity estimation protocols. Now I will discuss how our protocol performs given a realistic experimental setup. As mentioned, recently QuEra has demonstrated the 5-to-1 MSD protocol by encoding a logical qubit into a [[7,1,3]] color code, where all the logical Clifford operations can be executed transversally [102]. Moreover, such transversal operations are also feasible on their hardware platform due to the reconfigurability of the atom arrays. Therefore, what they did is that they first used a non-FT circuit to prepare 5 copies of logical $|SH\rangle$ magic states on the [[7,1,3]] code, and then executed the MSD circuit on the logical level, where only transversal Clifford operations are required. Finally, the logical Pauli tomography was performed on the successfully distilled magic states, while the data post-selection was also applied based on the measured syndrome information on all 5 logical qubits in order to get a higher reported fidelity.

In our simulations, we adopt the same circuit as that in Ref. [102] for both the non-FT state preparation and the MSD process. However, instead of the logical Pauli tomography, here for the fidelity estimation we consider the Clifford subgroup G_{SH} twirling on two copies of successfully distilled magic states followed by a logical Bell measurement between them. For coding simplicity, in our simulations we require that the two distilled copies need to be successful at the same time, which reduces the overall success rate in the simulation. In practice, however, this will be unnecessary since we can always store the successfully distilled state until the second one comes. Therefore, when counting the consumed distilled states in the simulation, we still keep the assumption that the states can be stored, in order to show the overhead that our protocol can offer at best. For comparison purposes, we also simulate the performance of the fidelity estimation with the logical tomography method. See Fig 7.3 (a) and (b) for the fidelity characterization circuits, respectively.

We consider a full circuit-level noise model in the simulations, which contains a noisy

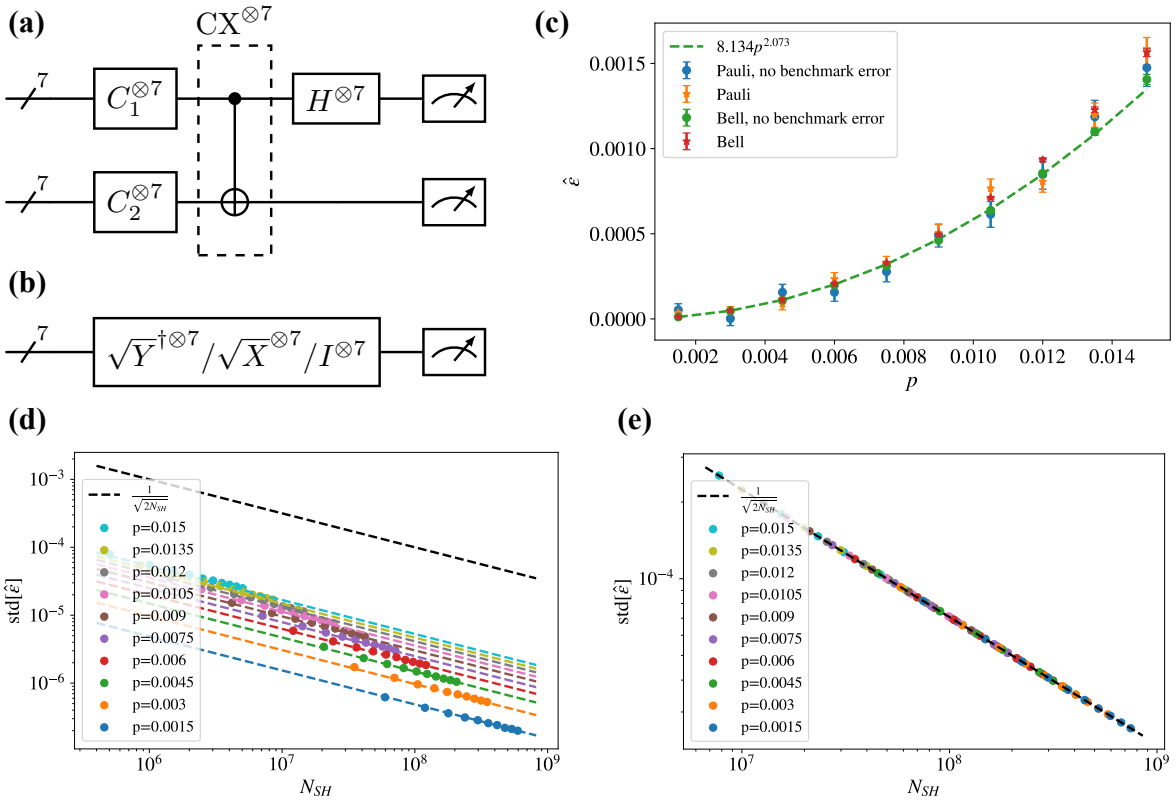

Figure 7.3: Numerical simulation on $|SH\rangle$ magic state fidelity characterization overhead. (a-b) Fidelity characterization circuit on the [[7,1,3]] logical magic $|SH\rangle$ state using (a) Bell measurement and (b) logical Pauli tomography method. $C_{1(2)}$ are single-qubit Clifford gates randomly chosen from G_{SH}. I omit to plot the noisy magic state preparation and distillation circuits used in the simulation since they are the same as [102]. (c) The estimated infidelity of the logical magic state after distillation. The green and blue dots are simulated with noiseless characterization (a-b) circuits, and the green dashed line is achieved by fitting the green dots. (d-e) The standard deviation of the estimator $\hat{\epsilon}$ when varying N_{SH} and p, using (d) Bell measurement and (e) Pauli tomography method. All colored dashed lines in (d) are references of $\sqrt{2\hat{\epsilon}/N_{SH}}$ where $\hat{\epsilon}$ is shown as the green dots in (c). Black dashed lines in (d-e) are references to $1/\sqrt{2N_{SH}}$. Up to $p = 0.015$, we observe the std$[\hat{\epsilon}]$ with the Bell measurement scheme are much below this reference line. While in the Pauli tomography scheme, all std$[\hat{\epsilon}]$ points are located along this line, as it is insensitive to the change of noisy strength p.

channel associated with every physical state preparation and measurement and every single-qubit and entangling operation on physical qubits. The overall strength of the noise is characterized by a single parameter p. Specifically, we have:

- Single-qubit gates are followed by a depolarizing channel with average fidelity $1 - p/5$.

- CZ gates are followed by a 2-qubit depolarizing channel with average fidelity $1-p$.
- Qubit initialization is followed by a depolarizing channel with average fidelity $1-p/2$.
- Each qubit measurement suffers a classical bit-flip with probability $p/2$.

Here, the convention of the channel fidelity is kept with Ref. [182].

To simulate the noisy circuit that includes magic state inputs, we adopt a two-stage method described in Ref. [102]. The depolarizing noise channel can be simulated by randomly inserting Pauli errors in the circuit. We use the Stim package to simulate these processes and see how the noise affects the final syndrome and logical measurement outcomes, while the noiseless MSD on the logical level is simulated separately. Finally, we apply the error-detecting strategy to post-select the results, where the full instance will be discarded if a nontrivial syndrome on any logical qubit is detected. This is because we expect that the initially prepared logical magic states contain an order of $O(p)$ logical infidelity, which will be suppressed into $O(p^2)$ after a successful distillation. On the other hand, since the QEC code we use has a distance $d = 3$, and every logical Clifford is transversal, we expect the fidelity characterization part will only add $O(p^3)$ extra errors when using the error detection, which will be negligible compared with the desired $O(p^2)$ truth when p is small.

The simulation results are shown in Fig 7.3. We first plot how the estimated magic state infidelity varies with p, and confirm the $O(p^2)$ scaling as the desired outcome from the MSD. We also compare with the results where the fidelity characterization part of the circuit is noiseless, where little deviation is observed. After this, the standard deviations on the fidelity estimators $\hat{\epsilon}$ are plotted with both the Pauli tomography and the Bell measurement methods. Even when we vary p up to 1.5×10^{-2}, the Pauli tomography method still keeps $\mathrm{std}[\hat{\epsilon}] \simeq \frac{1}{\sqrt{2N}}$, which seems independent with p itself. For the Bell measurement scheme, we also confirm the $\mathrm{std}[\hat{\epsilon}] \simeq \sqrt{2\epsilon/N}$ scaling with different p values. As a result, if we set a target precision $\mathrm{std}[\hat{\epsilon}] = r\epsilon$ with a constant factor r, we need the $N \sim O(1/\epsilon^2)$ overhead for

the Pauli tomography method and $N \sim O(1/\epsilon)$ overhead for the Bell measurement method as expected.

CHAPTER 8

OUTLOOK

It has been a long journey before we arrive here, where we have seen a variety of approaches to truncate the Hilbert space of the related quantum systems and their applications from different aspects. Indeed, it is a tool to simplify the problem at hand, and most people will more or less use it in their own research. Looking ahead, there are some directions that I think are worth people looking into, based on my own research experience:

First, an engineered truncation could break the symmetry of the system and improve the controllability. We have seen sufficient evidence where the cutoff comes from either destructive interference (Chapter 2) or the large detuning introduced between certain transitions (examples including resonator modes like Ref. [169] and Chapter 6; as well as the atomic inner structures, see Ref. [183]). It will be interesting to further look into the dynamics in this kind of blockaded systems, and see if we can find some common features that will guide us to find more robust or optimal (time, fidelity, etc) control sequences. On the other hand, the population leakage to the outer space should also be taken care of when the truncation itself is not perfect.

The above direction may sound too technical, but the following one should be of more practical interest. It is the co-design between hardware-level error structures and upper-level QEC codes. We have discussed different types of error structures, including Pauli bias and erasures, but there still lacks an overall merit on the tradeoff between the structuredness and the gate fidelity itself. People should optimize the design of QEC codes and the decoding strategies to fully exploit the benefits from error structures on the physical qubits, where neural network optimization could be a powerful tool in this. The logical error rate should be an important index for comparison, but we should also consider the difficulties in performing logical operations on the designed codes. On the hardware aspects, it will be interesting if we can find other kinds of error structures in the physical qubits, but it is hard for me to

envision what it looks like at the current time.

When talking about the hardware-efficient encoding of qudits, the definitions of Pauli error weight or code distance in the upper-level QEC code may need modification. It is because even two Pauli errors have the same type and weight, they may come from the physical noise with different orders in their probability. Take the three-legged bosonic cat as an example, where the qutrit computational basis is encoded by three coherent states with $\pm\frac{2\pi}{3}$ rotational symmetry. Then, although Z and Z^2 errors are both supported on one qutrit, they may come from physical noise $\hat{a}$ and $\hat{a}^2$ accordingly, with a probability $O(p)$ and $O(p^2)$. The concept of the "bias-preserving" should also be adapted in the new framework to take care of this issue.

Then we may come to the many-body physics, where we should find richer behaviors in dynamics compared with most of the problems I discussed in this thesis. Together with the driven-dissipative process, we may still observe a decoherence-free subspace where we may encode quantum information in it. Further, we should check how the noise possibly affects the states in the encoded subspace and see if it can be corrected autonomously. Indeed, we need some luck in the discovery. But magic can happen sometimes. Theoretical guidance will always be helpful in this case.

A collective truncation may also lead to nontrivial dynamics in the many-body setting. However, the connectivity requirement could be more demanding than the local interaction. For example, in Chapter 6, our device does contain a transmon that couples with all the modes. Another approach to quickly build long-range correlation is through the mid-circuit measurement (MCM) and feedforward (FF) control. In the past years, we have observed several works discussing what quantum states can be prepared in constant depth given this framework [184–186]. I expect to see if some gadgets here can be generalized from the circuit level to the physical level, similar to the generalization from the standard QEC protocol (which also requires MCM and FF control to correct errors) to the hardware-efficient

encoding.

Regarding the state characterization project, I would say that the collective truncation is somehow a very special case, as it is contained in the framework of MPS/MPO. On the other hand, the state or gate fidelity characterization on the logical level should be an important direction to pursue. Though syndrome information is only used for postselection in the protocol described in Chapter 7, in general there is evidence that such extra information accumulated during stabilizer measurements could be helpful in error characterization of the logical states [187]. I expect to see if it can also overcome certain limitations in some characterization tasks where only encoded information is used [188].

Those are some of the near-term generalizations that I can come up with, based on the topics I discussed in the thesis. I sincerely hope that you, the reader of this lengthy thesis, could still find something useful or inspiring in it. May the work spark your thoughts at some point!

APPENDIX A

APPENDICES FOR CHAPTER 2

A.1 Detailed universality proof

In this part, we show that the Hamiltonian $\hat{H}_{\rm dr}$ [see Eq. (2.2)] can generate arbitrary unitary operations in U(N) within the $N = r + 1$ dimensional blockade subspace $\mathcal{H}_b$. Since r is an adjustable parameter, our protocol allows for "universal control" in bosonic systems with Kerr nonlinearities in the sense that any unitary operator with any chosen dimension is realizable in the system.

In quantum control theory [127], a Hamiltonian defined in an N dimensional Hilbert space spanned by $\{|k\rangle\}_{k=0}^{N-1}$ contains a drift part $\hat{H}_d$ and several control parts $\hat{H}_{c,j}$. It can be expressed as

$$\hat{H}(t) = \hat{H}_d + \sum_j v_j(t)\hat{H}_{c,j}. \tag{A.1}$$

Here, we consider the drift and one of the control parts with the following form

$$\begin{aligned} \hat{H}_d &= \sum_{k=0}^{N-1} E_k|k\rangle\langle k|, \\ \hat{H}_{c,1} &= \sum_{k=0}^{N-2} d_k(|k+1\rangle\langle k| + |k\rangle\langle k+1|). \end{aligned} \tag{A.2}$$

Here $d_k \in \mathbb{R}$ and $d_k \neq 0$. The "universal control" [112] is named as the ability to realize any unitary operation $\hat{U}_{\rm tar}$ in $U(N)$ with properly chosen $v_j(t)$ and evolution time T, such that

$$\hat{U}_{\rm tar} = \hat{U}(T) = \mathcal{T}\exp\left[-i\int_0^T \hat{H}(t')\,{\rm d}t'\right]. \tag{A.3}$$

A Theorem in Ref. [37] suggests a sufficient condition for the choice of E_k and d_k to make the system universally controllable. We first repeat the theorem here and check that

our Hamiltonian $\hat{H}_{\rm dr}$ restricted in blockade subspace [$\hat{\Pi}_r \hat{H}_{\rm dr} \hat{\Pi}_r$, see Eq. (2.5)] satisfies those criteria.

Theorem A.1 (Ref. [37]). *Denote $\mu_k = E_k - E_{k+1}$. If $\mu_0 \neq 0$ and $\mu_k^2 \neq \mu_0^2$ for $k > 0$ (or similarly if $\mu_{N-2} \neq 0$ and $\mu_k^2 \neq \mu_{N-2}^2$ for $k < N-2$), then the dynamical Lie group of the system $\hat{H}(t)$ defined in Eq.* (A.1) *and* (A.2) *is at least SU(N). Further, if* $\mathrm{Tr}[\hat{H}_d] \neq 0$, *the dynamical Lie group is U(N).*

We notice that $\hat{\Pi}_r \hat{H}_{\rm dr} \hat{\Pi}_r$ contains the drift part $\hat{H}_{d,0}$ and two control parts $\hat{H}_{c,R}$ and $\hat{H}_{c,I}$. It is easy to verify that $\hat{H}_{c,R}$ meets the requirement for the control part. Then we should check the nonlinear condition for $\hat{H}_{d,0}$. In this case, we have $E_k = \chi(k^2 - k)/2 + \Delta_0 k$, which gives the nearest energy difference $\mu_k = -\chi k - \Delta_0$. To match the condition in Theorem A.1, we need to make sure $\mu_0^2 \neq \mu_{N-2}^2$, which leads to $r \neq -\frac{2\Delta_0}{\chi} + 1$. Further, since μ_k is monotonic in k, we cannot find both k_1, k_2 ($0 < k_1, k_2 < N-2$) such that $\mu_0^2 = \mu_{k_1}^2$ and $\mu_{N-2}^2 = \mu_{k_2}^2$. This concludes the proof of the universal controllability of our system, even if we can fix $\mathrm{Im}[\alpha(t)] = 0$ all the time.

Moreover, the ability to control $\mathrm{Im}[\alpha(t)]$ provides us with an additional degree of freedom to control the system, which in principle makes it possible to perform operations arbitrarily fast. To demonstrate this, we first show that the nested commutators between $i\hat{H}_{c,R}$ and $i\hat{H}_{c,I}$ are sufficient to form a complete basis of the Lie algebra associated with SU(N). The reason is that $[i\hat{H}_{c,R}, i\hat{H}_{c,I}] = \sum_{n=0}^{r} -2[3n^2 - (4r+1)n + r^2] i|n\rangle\langle n|$ is diagonal in Fock basis, has zero trace, and fulfills the anharmonicity requirement as that for the drift term $\hat{H}_d$ in Theorem A.1. Further, as proved in Theorem A.1, $[i\hat{H}_{c,R}, i\hat{H}_{c,I}]$ and $i\hat{H}_{c,R}$ are sufficient to generate a set of complete basis of Lie algebra $\mathfrak{su}(N)$. In practice, the overall phase for unitary operations has no physical meaning, which makes SU(N) group sufficient for universality. Second, since $\chi \mathrm{Re}[\alpha(t)] \hat{H}_{c,R} + \chi \mathrm{Im}[\alpha(t)] \hat{H}_{c,I}$ is sufficient to achieve the desired unitary, we can simply increase the amplitude of $\alpha(t)$ to do things arbitrarily fast. In the short time limit, the drift term $\hat{H}_{d,0}$ does not contribute to the dynamics and therefore does

not impose a speed constraint.

A.2 Basic error analysis

In this section, we analyze the errors in implementing a desired operation using our protocol where only 1-photon drive is allowed. Notice that, even if all the driving parameters are perfectly implemented, there are errors stemming from photon loss and inaccuracies of the Trotter design in Eq. (2.16). Reducing the total time of the protocol T helps to mitigate the photon loss errors, but at the same time increases the required input power, which could be limited in practice. This limited power also introduces coherent errors in the Trotter approximation of the unitary operator of interest. Therefore, there is a trade-off between these two sources of error, which we discuss in detail in the following.

For simplicity, we only consider the error in state preparation tasks characterized by infidelity

$$\epsilon = 1 - \langle \psi_{\text{tar}} | \, \hat{\rho}(T) \, | \psi_{\text{tar}} \rangle \, , \tag{A.4}$$

where $|\psi_{\text{tar}}\rangle$ is a pure state that we are interested in preparing and $\hat{\rho}(T)$ is the state of the system at the end of the evolution. Here, $\hat{\rho}(T)$ is the state obtained after evolving for time T under the Lindblad equation

$$\frac{\mathrm{d}\hat{\rho}}{\mathrm{d}t} = -i[\hat{H}'_{\text{dr}}, \hat{\rho}] + \kappa \mathcal{D}[\hat{a}]\hat{\rho}, \tag{A.5}$$

where $\hat{H}'_{\text{dr}}$ is introduced in Eq. (2.9).

To study the scaling of error ϵ with relevant physical parameters in the problem, we focus our attention on the case of single-photon state preparation when the blockade subsystem has only 2 dimensions and $\alpha(t)$ is constant, as originally proposed in Ref. [51]. Therefore, in this case we set $|\psi_{\text{tar}}\rangle = |1\rangle$, and $\hat{\rho}(0) = |0\rangle\langle 0|$. To prepare the target state $|1\rangle$, we choose $\Delta_0 = 0$ and $\alpha(t) = \alpha$, a constant that is assumed to be a real number for simplicity.

Therefore, the effective Hamiltonian in the blockade subspace is

$$\hat{H}_{\mathrm{dr}}^{\mathrm{qb}} = -\chi\alpha\hat{\sigma}_x. \tag{A.6}$$

Here $\hat{\sigma}_x = |0\rangle\langle 1| + |1\rangle\langle 0|$. Moreover, in this subspace the photon loss dissipator $\mathcal{D}[\hat{a}]$ in Eq. (A.5) is now simply modified to $\mathcal{D}[\hat{\sigma}_-]$, where $\hat{\sigma}_- = |0\rangle\langle 1|$. Note that in the ideal case, when there is no loss or Trotter error, we can perfectly prepare $|1\rangle$ by evolving $|0\rangle$ under $\hat{H}_{\mathrm{dr}}^{\mathrm{qb}}$ in Eq. (A.6) for time $T = \frac{\pi}{2\chi\alpha}$. For simplicity, here we ignore the part of the dynamics where $\alpha(t)$ increases from 0 to α in the beginning and decreases to 0 in the end [51], while only discussing errors with a constant $\alpha(t)$. In fact, using a time-dependent $\alpha(t)$ will introduce additional imperfection for the Trotter approximation. However, as discussed in Appendix A.3, given the assumption that $\alpha(t)$ varies slowly over time, these additional errors are higher-order effects compared with the major Trotter error scaling we focus on later.

To find the scaling properties of ϵ, we treat the photon loss (ϵ_{loss}) and Trotter (ϵ_{tt}) errors independently.

First, for the photon loss error, we assume that the Hamiltonian $\hat{H}_{\mathrm{dr}}^{\mathrm{qb}}$ in Eq. (A.6) is perfectly implemented so that the quantum state is confined to the blockade subspace all the time. If $\chi\alpha \gg \kappa$, we can treat the dissipative term $\kappa\mathcal{D}[\sigma_-]$ perturbatively. So, to the lowest order, the error from the photon loss process satisfies $\epsilon_{\mathrm{loss}} = c_1\kappa T$, where the coefficient c_1 in general depends on the function $\alpha(t)$. In our case, we find that $c_1 = \frac{3}{8}$ (see Appendix A.9).

Next, to estimate the error from the Trotter approximation, we ignore the photon loss process and assume that the system undergoes unitary evolution with Hamiltonian $\hat{H}'_{\mathrm{dr}}$ shown in Eq. (2.9). For simplicity, we use the discrete Trotter formula to illustrate analysis, but the results apply to the proposed continuous version as well (see Appendix A.8). Let $\hat{U}_{\mathrm{tar}} = e^{-i\frac{\hat{H}_1+\hat{H}_2}{2}\delta T}$ and $\hat{U} = e^{-i\hat{H}_1\delta T/4}e^{-i\hat{H}_2\delta T/2}e^{-i\hat{H}_1\delta T/4}$. Using the discrete version of

Trotter formula [189], we have

$$\hat{U}_{\rm tar}^{-1/2}\hat{U}\hat{U}_{\rm tar}^{-1/2} = \exp\{(\delta T)^3\hat{A}_3 + O[(\delta T)^5]\}, \tag{A.7}$$

where $\hat{A}_3 = i[[\hat{H}_1, \hat{H}_2], \hat{H}_1 + 2\hat{H}_2]/192$ is anti-Hermitian. Further, recall the form of $\hat{H}'_{\rm dr}[\tilde{\alpha}]$. Since we work in the regime that $|\alpha| \gg \sqrt{N}$ (N is the dimension of the blockade subspace), the dominant contribution to $\hat{A}_3$ comes from the $(\frac{\chi}{2}\alpha_{1,2}^2\hat{a}^{\dagger 2} + \text{h.c.})$ terms in the Hamiltonian $\hat{H}_1$ and $\hat{H}_2$, and therefore the matrix elements of $\hat{A}_3$ that we focus on scale as $O[(\chi\alpha^2)^3]$. Finally, as mentioned before, we need to sequentially apply $M = T/\delta T$ repetitions of the Trotter operation $\hat{U}$ to approximately achieve our desired unitary $\hat{U}_{\rm tar}^M$, where $|\psi_{\rm tar}\rangle\langle\psi_{\rm tar}| = \hat{U}_{\rm tar}^M\hat{\rho}(0)\hat{U}_{\rm tar}^{\dagger M}$. Therefore, we find

$$\epsilon_{\rm tt} = 1 - \langle\psi_{\rm tar}|\,\hat{U}^M\hat{\rho}(0)\hat{U}^{\dagger M}\,|\psi_{\rm tar}\rangle = O[(M(\chi\alpha^2\delta T)^3)^2]. \tag{A.8}$$

Recalling the relations $T \propto 1/(\chi\alpha)$ and $M = T/\delta T$, we finally obtain $\epsilon_{\rm tt} = c_2/[M^4(\chi T)^6]$, where c_2 is again the coefficient that depends on specific form of our pulse design and can be estimated numerically (see Appendix A.9). Note that in our numerical work $f(t)$ function is chosen as

$$f(t) = \frac{i\pi}{2}e^{-i2\pi|s(t)-1/2|}, \tag{A.9}$$

where $s(t) = t/\delta T - \lfloor t/\delta T\rfloor$. This choice of $f(t)$ also satisfies the requirements shown in Eq. (2.14) and (2.15), and also $|f(t)|$ is constant over time. In this situation, the reference displaced frame [characterized by $\tilde{\alpha}(t)$] oscillates back and forth along a semi-circle. In practice, it can be implemented with 1-photon drives using two interleaved tones with frequencies $(\omega_1 \pm \omega_r)$.

Since errors from the loss and Trotter approximation are independent, to the lowest order

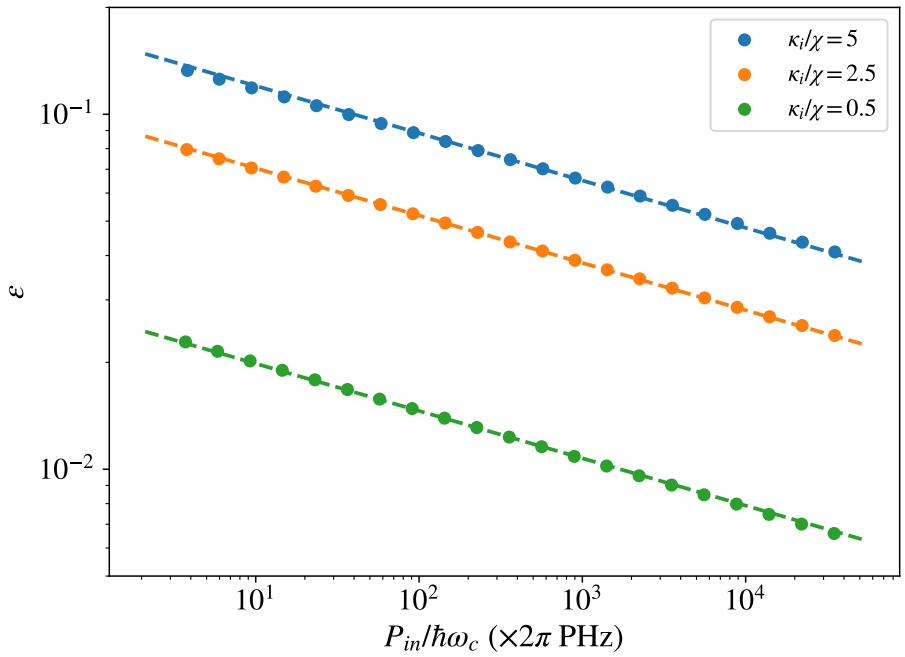

Figure A.1: Numerical results of the infidelity of Fock state $|1\rangle$ preparation with given input power and corresponding optimal choice of gate time. The choice of Kerr parameter is $\chi = 2\pi \times 3$ kHz. The dynamics is calculated with the assumption that RWA is valid. Different colors indicate different κ_i/χ ratios. The dashed lines are references for $\epsilon \propto P_{\text{in}}^{-2/15}$ scaling.

they can be added directly to find the total infidelity as

$$\epsilon_{\text{tot}} = \epsilon_{\text{loss}} + \epsilon_{\text{tt}} = c_1 \kappa T + \frac{c_2}{M^4 (\chi T)^6}. \tag{A.10}$$

If the driving amplitude is unbounded, we can use an arbitrarily short time T with arbitrary fast oscillating $\tilde{\alpha}(t)$ ($M \to \infty$) such that both ϵ_{loss} and ϵ_{tt} are suppressed to zero. However, in practice there are possible limitations that prevent us from doing that. Here we consider a key limitation, namely the input power P_{in} that can be applied to the system, and evaluate the optimal ϵ_{tot} that we can achieve under this constraint.

First, we assume that $\chi|\alpha|^2 \ll \omega_c$, which can be checked self-consistently since otherwise RWA will not apply, as discussed in Appendix A.5. This assumption allows us to

approximately write down the input power as

$$P_{\text{in}} \simeq \frac{|\Lambda_1|^2}{\kappa_e}\hbar\omega_c, \tag{A.11}$$

where κ_e is the external loss rate induced by coupling to the driving port. The total loss κ is composed of $\kappa = \kappa_e + \kappa_i$, where κ_i is the internal loss rate of the resonator. We notice from Eq. (2.10) that, $\Lambda_1 \sim O(\chi\tilde{\alpha}^3) + i\dot{\tilde{\alpha}}$. We then need to estimate the contribution from $\dot{\tilde{\alpha}}$. In our design we have $\dot{\tilde{\alpha}} \sim O(\alpha M/T)$. Besides, since we are only considering the lowest-order errors from the Trotter expansion, the Trotter error should be small. Therefore, we should have $\chi\alpha^2\delta T \ll 1$, or equivalently $\chi\alpha^2 \ll \omega_r$. This implies that $M \gg \chi\alpha^2 T$, and as a result, $i\dot{\tilde{\alpha}}$ is the dominating term in Λ_1. So, we can write P_{in} as

$$P_{\text{in}} \sim O\left(\frac{M^2\alpha^2}{\kappa_e T^2}\right) := \frac{c_3 M^2}{\kappa_e \chi^2 T^4}, \tag{A.12}$$

where $\hbar\omega_c$ and the proportionality constant originating from $f(t)$ are both absorbed in c_3. Using Eq. (A.12) to eliminate M in Eq. (A.10), we obtain

$$\epsilon_{\text{tot}} = c_1(\kappa_e + \kappa_i)T + \frac{c_2 c_3^2}{P_{\text{in}}^2 \kappa_e^2 \chi^{10} T^{14}}. \tag{A.13}$$

As a result, we find that the scaling of T^{opt}, namely the optimal time that minimizes ϵ_{tot}, and the corresponding optimal error $\epsilon_{\text{tot}}^{\text{opt}}$ is given by

$$\begin{aligned} T^{\text{opt}} &= \left(\frac{14 c_2 c_3^2}{c_1}\right)^{1/15} \cdot \frac{1}{P_{\text{in}}^{2/15}\chi^{2/3}\kappa_e^{2/15}(\kappa_e+\kappa_i)^{1/15}}, \\ \epsilon_{\text{tot}}^{\text{opt}} &= 15\left(\frac{c_1^{14} c_2 c_3^2}{14^{14}}\right)^{1/15} \cdot \left(\frac{1}{P_{\text{in}}^{2/15}\chi^{2/3}} \frac{(\kappa_e+\kappa_i)^{14/15}}{\kappa_e^{2/15}}\right). \end{aligned} \tag{A.14}$$

It is also natural to assume that κ_i is a fixed property of the device while κ_e is adjustable

to further optimize $\epsilon_{\text{tot}}^{\text{opt}}$. We can find the optimal ratio of $\kappa_e/\kappa_i = \frac{1}{6}$ that minimizes $\epsilon_{\text{tot}}^{\text{opt}}$ while keeping χ, κ_i and input power P_{in} fixed and using the optimal T^{opt}. This gives

$$\epsilon_{\text{tot}}^{\text{opt}} = \frac{15(c_1^{14}c_2c_3^2)^{1/15}}{2^{14/15}\cdot 6^{4/5}} \cdot \left(\frac{\kappa_i^{4/5}}{P_{\text{in}}^{2/15}\chi^{2/3}}\right). \tag{A.15}$$

We also numerically investigate the infidelity ϵ as prescribed by Eq. (A.4) as a function of input power. To find the optimal time for the protocol in our numerical simulations, we first need to determine the constants in Eq. (A.13). More details can be found in Appendix A.9. We first obtain c_2 by varying M and T in the simulations and extracting the corresponding proportionality constant in Eq. (A.10). We then obtain c_3 by using the explicit form of $f(t)$ in Eq. A.9, which results in $c_3 \simeq \frac{\pi^6}{4}\hbar\omega_c$. Therefore, together with $c_1 = \frac{3}{8}$ from perturbative analysis, we can then find T^{opt} for a given power, which subsequently determines the required M using Eq. (A.12). In this way, both the optimal protocol time T^{opt} and $\tilde{\alpha}(t)$ are fully determined. Here we also use the optimal choice of $\kappa_e = \frac{1}{6}\kappa_i$.

In Fig. A.1, we observe that the numerically obtained infidelities agree well with the estimated $\epsilon \propto P_{\text{in}}^{-2/15}$ scaling in a wide range of P_{in} values.

Finally, in the numerical simulation, we have assumed the validity of RWA so that the dynamics is irrelevant with ω_c after going to the rotating frame. We notice that if the required α or M (equivalently ω_r) is so large that Λ_1 is comparable with the frequency of the resonator ω_c, then those off-resonant terms which has been ignored under RWA in the beginning may lead to non-negligible effects on the dynamics. In Appendix A.5, we use scaling analysis to briefly discuss those effects and strategies to partially compensate them by adjusting $\Lambda_1(t)$ and $\omega_1(t)$, or driving both charge and flux quadratures together rather than only one of them on hardware.

A.3 Trotter errors with time-dependent $\alpha(t)$

In this section, we will also discuss the error scaling properties of our protocol, but specifically focus on the Trotter error in the case that $\alpha(t)$ designed by the optimal control algorithm is a slowly varying function that depends on time t. Here, by "slowly varied", we mean the time derivative of $\alpha(t)$ scales as $\dot{\alpha}(t) \sim O(\alpha/T)$, and in general $\mathrm{d}^n\alpha(t)/\,\mathrm{d}t^n \sim O(\alpha/T^n)$. To be more precise, we will calculate the difference between both sides of Eq. (2.16) in detail, where on the left-hand side is the evolution we can achieve with only 1-photon drives, and on the right-hand side is the target unitary operation that we want to achieve. Notice that, unlike the constant α case in the main text where $T \propto 1/(\chi\alpha)$, here we do not have such a property rigorously. But, it is still reasonable to assume that T roughly scales as $O(1/\chi\alpha)$ since the dominant part in $\hat{H}_{\mathrm{dr}}$ is still $[\chi\alpha(t)\hat{a}^\dagger(\hat{n}-r)+\mathrm{h.c.}]$ if α is large. We will keep this assumption in the following derivation.

First, let us calculate the difference of unitary operation within each time slice $t \in [k\delta T, (k+1)\delta T]$ where k is an integer and $0 \le k < M$. We denote

$$\begin{cases} \hat{U}_k = \mathcal{T}\exp\left\{-i\int_{k\delta T}^{(k+1)\delta T} \hat{H}'_{\mathrm{dr}}[\tilde{\alpha}(t)]\,\mathrm{d}t\right\}, \\ \hat{U}_{\mathrm{tar},k} = \mathcal{T}\exp\left\{-i\int_{k\delta T}^{(k+1)\delta T} \hat{H}_{\mathrm{dr}}[\alpha(t)]\,\mathrm{d}t\right\}. \end{cases} \tag{A.16}$$

Recall that $\tilde{\alpha}(t) = \alpha(t)f(t)$, where $f(t)$ is a periodic function with period δT, scales as $O(1)$, and satisfies Eq. (2.14) and 2.15. Without loss of generality, we can just focus on the $k=0$ case and calculate $\hat{U}_0\hat{U}^\dagger_{\mathrm{tar},0}$, since $\hat{U}_k\hat{U}^\dagger_{\mathrm{tar},k}$ with any k has a similar structure as the situation

with $k = 0$. By analogy with the Magnus expansion, we have

$$\begin{aligned}\hat{U}_0\hat{U}^{\dagger}_{\mathrm{tar},0} = \exp\Bigg(& -i\int_0^{\delta T} \mathrm{d}t\ [\hat{H}'_{\mathrm{dr}}(t) - \hat{H}_{\mathrm{dr}}(t)] \\ & -\frac{1}{2}\int_0^{\delta T}\mathrm{d}t_2\int_0^{t_2}\mathrm{d}t_1\{[\hat{H}'_{\mathrm{dr}}(t_2),\hat{H}'_{\mathrm{dr}}(t_1)] - [\hat{H}_{\mathrm{dr}}(t_2),\hat{H}_{\mathrm{dr}}(t_1)]\} \\ & +\frac{1}{2}\int_0^{\delta T}\mathrm{d}t_2\int_0^{\delta T}\mathrm{d}t_1[\hat{H}'_{\mathrm{dr}}(t_2),\hat{H}_{\mathrm{dr}}(t_1)] + \hat{R}_3\Bigg).\end{aligned} \tag{A.17}$$

It is easy to see that $\hat{R}_3$ is anti-Hermitian and $\hat{R}_3 \sim O[(\chi\alpha^2\delta T)^3]$. As we discussed in the main text, if $\alpha(t)$ is a constant within $t \in [0, \delta T]$, the first and the second order terms written explicitly in Eq. (A.17) are zero. But here we want to talk about the more general case that $\alpha(t)$ is time-dependent. Even in this situation, we will show that the contributions of the error from the first two orders are small compared with the third-order term.

We can write $\alpha(t)$ as

$$\alpha(t) = \alpha(t_0) + \dot{\alpha}(t_0)(t-t_0) + \frac{1}{2}\ddot{\alpha}(t_0)(t-t_0)^2 + O[(t-t_0)^3], \tag{A.18}$$

where $t_0 = \frac{\delta T}{2}$. Since we assumed that $\alpha(t)$ is a slowly varied function, we have $\dot{\alpha} \sim O(\alpha/T)$ and $\ddot{\alpha} \sim O(\alpha/T^2)$. As a result, based on the structure of $f(t)$, for the first-order term we have

$$\begin{aligned}\int_0^{\delta T}\mathrm{d}t\ \chi[\tilde{\alpha}(t) - \alpha(t)] &\sim O\left[\frac{\chi\alpha(\delta T)^3}{T^2}\right], \\ \int_0^{\delta T}\mathrm{d}t\ \chi\tilde{\alpha}^2(t) &\sim O\left[\frac{\chi\alpha^2(\delta T)^3}{T^2}\right].\end{aligned} \tag{A.19}$$

Since we have assumed that $T \sim O(1/\chi\alpha)$, both of the two terms are smaller compared with $O[(\chi\alpha^2\delta T)^3]$ when α is large.

A similar analysis can be performed for the second-order terms. The dominant part lies

in $[\hat{H}'_{\rm dr}(t_2), \hat{H}'_{\rm dr}(t_1)]$ since it contains $O[(\chi\alpha^2)^2]$ coefficients. We find that

$$\int_0^{\delta T} \mathrm{d}t_2 \int_0^{t_2} \mathrm{d}t_1 \; \chi^2[\tilde{\alpha}^2(t_2)\tilde{\alpha}^{*2}(t_1) - \tilde{\alpha}^{*2}(t_2)\tilde{\alpha}^2(t_1)] \sim O\left(\frac{\chi^2\alpha^4(\delta T)^3}{T}\right), \tag{A.20}$$

which is also a higher order term compared with $O[(\chi\alpha^2\delta T)^3]$. So, the dominant scaling of Trotter error does not change even if $\alpha(t)$ is a slowly varied function introduced in Eq. (2.7).

A.4 Errors from inaccurate control

In this part, we will investigate the extra infidelity induced by inaccurate control of $\Lambda_1(t)$. As shown in Appendix A.2, $|\Lambda_1| \sim O(\alpha M/T)$ can be strong enough so that a deviation for a small portion of it may cause a huge impact. Here we consider the problem that how the infidelity changes if the actual driving pulse we use is $(1+\eta)\Lambda_1(t)$ where η is a small dimensionless number.

We first briefly derive the contribution of the extra $\eta\Lambda_1(t)$ to the dynamics. Notice that, if we go to the same displaced rotating frame as mentioned in the main text, now the Hamiltonian should be written as

$$\hat{H}_\eta = \hat{H}'_{\rm dr}[\tilde{\alpha}(t)] + [\eta\Lambda_1(t)\hat{a}^\dagger + \text{h.c.}], \tag{A.21}$$

where $\hat{H}'_{\rm dr}[\tilde{\alpha}(t)]$ has been defined in Eq. (2.9) and $\Lambda_1(t) = \chi\tilde{\alpha}(t)[|\tilde{\alpha}(t)|^2 - r] + i\kappa\tilde{\alpha}(t)/2 + i\dot{\tilde{\alpha}}(t)$ (as in Eq. (2.10)). We can again perform a frame transformation such that $\hat{a} \to \hat{a} + \beta(t)$ where $\beta(t)$ satisfies

$$\dot{\beta}(t) + \frac{\kappa}{2}\beta(t) = -i\eta\Lambda_1(t). \tag{A.22}$$

Further, due to the specific form of $\Lambda_1(t)$, we can write $\beta(t) = \beta_1(t) + \beta_2(t)$ where $\beta_1(t) = \eta\tilde{\alpha}(t)$ and $\beta_2(t) = -i\eta\int_0^t e^{-\kappa(t-t')/2}\chi\tilde{\alpha}(t')[|\tilde{\alpha}(t')|^2 - r]\,\mathrm{d}t'$. It is worth mentioning that, although $i\dot{\tilde{\alpha}}$ dominates in Λ_1, its contribution to β is much smaller compared with the

contribution from the $O(\chi\alpha^3)$ term, provided that $\alpha \gg 1$.

Now the Hamiltonian in the new frame should be

$$\begin{aligned}\hat{H}'_\eta = &\frac{\chi}{2}(\hat{a}^\dagger + \beta^*)^2(\hat{a} + \beta)^2 + \{\chi\tilde{\alpha}(\hat{a}^\dagger + \beta^*)[(\hat{a}^\dagger + \beta^*)(\hat{a} + \beta) - r] + \text{h.c.}\} \\ &+ [\frac{\chi}{2}\tilde{\alpha}^2(\hat{a}^\dagger + \beta^*)^2 + \text{h.c.}].\end{aligned} \tag{A.23}$$

Clearly $\hat{H}'_\eta$ deviates from the ideal $\hat{H}'_{\text{dr}}[\tilde{\alpha}(t)]$ as β increases from zero. Besides, in general $\beta(T) \neq 0$, which suggests that we need an extra displacement operation $\hat{D}[\beta(T)]$ to let the states go back to the non-displaced frame. But in practice, we do not know what the $\beta(T)$ is and will not actively apply this operation, which may lead to extra errors.

Finally, we try to numerically consider a specific example, which is the Fock $|1\rangle$ state preparation with constant α as mentioned in Appendix A.2. We assume that, T and M are chosen optimally such that for a given target infidelity ϵ_0 we have $c_1\kappa T = \frac{14}{15}\epsilon_0$ and $\frac{c_2}{M^4(\chi T)^6} = \frac{1}{15}\epsilon_0$. However, during the constant α evolution time, the actual 1-photon driving amplitude we implement is $(1+\eta)\Lambda_1(t)$. From Fig. A.2 we can see how the fidelity $F = 1-\epsilon$ decreases as the increase of inaccuracy η. We notice that for a smaller ϵ_0 expected, we need a larger α. However, larger α will result in larger deviation β. The fidelity will be more sensitive to η with larger α, and therefore, when ϵ_0 is smaller, it decays faster as η grows.

A.5 Dynamics beyond rotating wave approximation

A.5.1 Overview

In the main text, we did not talk about any property related to the resonator frequency ω_c, since it simply disappears after we go to the rotating frame and perform the rotating wave approximation (RWA). In practice, however, the validity of RWA will set a limitation on our current protocol. If we can do either charge or flux drive but not both, then the driving

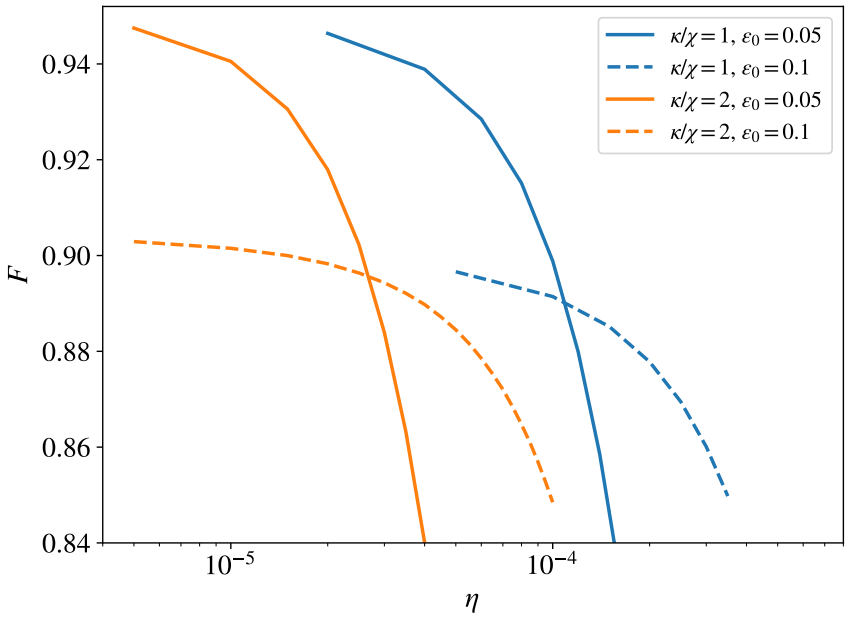

Figure A.2: Fidelity changes due to the amplitude deviation with the ideal pulses for the Fock $|1\rangle$ state preparation task, as explained in Appendix A.4.

amplitude Λ_1 cannot be too strong, which further limits the maximum $|\alpha|$ and the oscillation frequency ω_r for $\tilde{\alpha}(t)$ that is used to suppress the infidelity. In this section, we will discuss the effects of those off-resonant terms in the aspect of scaling, and we also provide some strategies to partially compensate for them.

Before we start to talk about the specific problem, we will briefly explain the analysis strategy we are going to use. Rather than going to the displaced rotating frame we mentioned in the main text, which is treated as the 0-th order reference frame here, we try to find another reference frame with higher order corrections on the original one and compare the difference between Hamiltonian there and the one ($\hat{H}'_{\rm dr}[\tilde{\alpha}(t)]$) we desire to engineer. To make sure the dynamics induced by the difference between the two Hamiltonians are negligible, we need higher-order corrections on the 1-photon driving amplitude $\Lambda_1(t)$ and frequency $\omega_1(t)$, together with several assumptions on the relationship among some relevant parameters (see Eq. (A.31) for example).

As we shall see, there are two kinds of terms in the difference between the desired and

actual Hamiltonian. One kind includes terms that have a fast-oscillating coefficient, which is usually proportional to $e^{ik\omega_c t}$ or $e^{ik\omega_r t}$ with nonzero integer k; while another kind contains those terms without fast-oscillating factors, which are called "slowly varied" terms. Besides, we also call terms with coefficients $e^{ik\omega_r t}$ "slowly varied" when they are compared with $e^{ik\omega_c t}$ (or we call "slowly varied" compared with ω_c for short).

For both cases, the dynamics induced by those terms is negligible if the amplitudes (absolute value) of their coefficients are too small such that the integration of the amplitudes over gate time T is far less than 1. For example, terms with slowly varied coefficient $c(t)$ or the fast oscillating $c(t)e^{i\omega t}$ can be both ignored if $\int_0^T |c(t)|\,\mathrm{d}t \ll 1$. However, for terms with fast oscillating coefficients that do not belong to the former situation, we need to consider some effective Hamiltonian corresponding to those terms. For example, the dynamics generated by a Hamiltonian like $\hat{H} = \hat{A}e^{i\omega t} + \hat{A}^\dagger e^{-i\omega t}$ can be approximated by that generated via an effective Hamiltonian $\hat{H}_{\text{eff}} = [\hat{A}, \hat{A}^\dagger]/\omega$, provided that ω is much larger than some specific matrix norm of $\hat{A}$. This can be simply demonstrated via the Magnus expansion. In this way, we will focus on the scaling of the coefficients for those slowly varied effective counterparts and see if the time integration of those is far less than 1 and therefore negligible as well. Similarly, remember that in the parameter regime we want to work in, the coherent error induced by $(\frac{\chi}{2}\tilde{\alpha}^2\hat{a}^{\dagger 2}+\text{h.c.})$ is negligible. In general, for terms that oscillate in the form of $f^2(t)$ [where $f(t)$ is a periodic function with frequency ω_r and satisfies Eq. (2.14) and (2.15)], their contribution to the dynamics can be ignored even when their amplitudes scale as $O(\chi\alpha^2)$. We will also make use of this fact later. These discussions provide support for the following derivations.

Our scaling analysis also relies on some assumptions for the ideal blockade Hamiltonian. We assume the dimension of the blockade subspace is small, such that in the scaling analysis the matrix elements of $\hat{a}$ and $\hat{a}^\dagger$ are of the order $O(1)$ when dynamics is restricted to the blockade subspace. Also, we assume that the control function $\alpha(t)$ is slowly varied in time,

which (as mentioned in Appendix A.3) implies that the time derivative of $\alpha(t)$ scales as $\dot{\alpha}(t) \sim O(\alpha/T)$, and similarly $\mathrm{d}^n\alpha(t)/\,\mathrm{d}t^n \sim O(\alpha/T^n)$.

A.5.2 Detailed scaling analysis on the dynamics

In general, the Hamiltonian in the lab frame without RWA should be written as

$$\hat{H} = \frac{\chi}{12}(\hat{a}^\dagger + \hat{a})^4 + (\omega_c - \chi)\hat{a}^\dagger\hat{a} + [\Lambda_1(t)e^{-i\omega_1(t)t} + \Lambda_1^*(t)e^{i\omega_1(t)t}](\hat{a}^\dagger + \hat{a}). \tag{A.24}$$

From then on we denote $\phi_d(t) := \omega_1(t)t - \omega_c t$ and as shown in the main text we have $\dot{\phi}_d(t) \sim O(\chi\alpha^2)$. If we still go to the displaced rotating frame shown in the main text, the Hamiltonian will become

$$\begin{aligned}
\hat{H}_{\mathrm{NR}} = {} & \frac{\chi}{2}(\hat{a}^\dagger + \tilde{\alpha}^*)^2(\hat{a} + \tilde{\alpha})^2 - \dot{\phi}_d(t)(\hat{a}^\dagger + \tilde{\alpha}^*)(\hat{a} + \tilde{\alpha}) + [(\Lambda_1(t) - i\dot{\tilde{\alpha}} - i\kappa\tilde{\alpha}/2)\hat{a}^\dagger + \mathrm{h.c.}] \\
& + (\Lambda_1^*(t)e^{2i\phi_d(t)}e^{2i\omega_c t}\hat{a}^\dagger + \mathrm{h.c.}) + \\
& \left\{\left[\frac{\chi}{3}(\hat{a}^\dagger + \tilde{\alpha}^*)^3(\hat{a} + \tilde{\alpha}) + \frac{\chi}{2}(\hat{a}^\dagger + \tilde{\alpha}^*)^2\right]e^{2i\phi_d(t)}e^{2i\omega_c t} + \mathrm{h.c.}\right\} \\
& + \left[\frac{\chi}{12}(\hat{a}^\dagger + \tilde{\alpha}^*)^4 e^{4i\phi_d(t)}e^{4i\omega_c t} + \mathrm{h.c.}\right].
\end{aligned} \tag{A.25}$$

It can be easily seen that we can recover the Hamiltonian under RWA by throwing away all the fast oscillating terms related to $e^{i\omega_c t}$. Therefore, as proposed earlier we can choose $\Lambda_1^{[0]}$ and $\phi_d^{[0]}$ to achieve the desired Hamiltonian under RWA by

$$\begin{aligned}
\Lambda_1^{[0]} &= \dot{\phi}_d^{[0]}\tilde{\alpha}^{[0]} - \chi\tilde{\alpha}^{[0]}(|\tilde{\alpha}^{[0]}|^2 + r) + i\dot{\tilde{\alpha}}^{[0]} + i\kappa\tilde{\alpha}^{[0]}/2, \\
\dot{\phi}_d^{[0]} &= 2\chi|\tilde{\alpha}^{[0]}|^2.
\end{aligned} \tag{A.26}$$

where $\tilde{\alpha}^{[0]}$ is the $\tilde{\alpha}$ function we designed in the main text.

Now we want to focus on those fast oscillating terms. One way is to go to a slightly

different displaced frame to absorb some large components. Note that $\Lambda_1 \sim i\dot{\tilde{\alpha}} \sim \alpha\omega_r \gg \chi\alpha^3$, which means $(\Lambda_1^*(t)e^{2i\phi_d(t)}e^{2i\omega_c t}\hat{a}^\dagger + \text{h.c.})$ has the largest amplitude and we want to deal with it first. Basically, we want to choose $\tilde{\alpha} = \tilde{\alpha}^{[0]} + \tilde{\alpha}^{[1]}$ such that the amplitude of the linear $\hat{a}^\dagger$ term can be mostly absorbed in $\tilde{\alpha}^{[1]}$. We first introduce $\tilde{\alpha}^{[1']}$ tentatively such that

$$\begin{aligned} i\dot{\tilde{\alpha}}^{[1']} + i\kappa\tilde{\alpha}^{[1']}/2 = [\Lambda_1^{*[0]} + \chi\tilde{\alpha}^{*[0]}(|\tilde{\alpha}^{[0]}|^2 + 1)]e^{2i\phi_d^{[0]}(t)}e^{2i\omega_c t} \\ + \frac{\chi}{3}(\tilde{\alpha}^{*[0]})^3 e^{4i\phi_d^{[0]}(t)}e^{4i\omega_c t} + \frac{\chi}{3}(\tilde{\alpha}^{[0]})^3 e^{-2i\phi_d^{[0]}(t)}e^{-2i\omega_c t}. \end{aligned} \tag{A.27}$$

For simplicity, we denote the right-hand-side (RHS) of the equation above as $g(t)$.

The solution for $\tilde{\alpha}^{[1']}$ can be written down explicitly as

$$\tilde{\alpha}^{[1']}(t) = -i\int_0^t e^{-\kappa(t-t')/2}g(t')\,\mathrm{d}t'. \tag{A.28}$$

Via the technique of integration by parts, we have the following identity

$$\begin{aligned} \int_0^t h(t')e^{i\omega t'}\,\mathrm{d}t' = \sum_{n=0}^{k}\left(\frac{i}{\omega}\right)^{n+1}[h^{(n)}(0) - h^{(n)}(t)e^{i\omega t}] \\ + \left(\frac{i}{\omega}\right)^{k+1}\int_0^t h^{(k+1)}(t')e^{i\omega t'}\,\mathrm{d}t', \end{aligned} \tag{A.29}$$

where $h^{(n)}(t) = \frac{\mathrm{d}^n}{\mathrm{d}t^n}h(t)$. In Eq. (A.29) we have implicitly assumed that $h^{(n)}(t)$ is a continuous function for $n \le k$. Notice that if for any m, $\frac{1}{\omega}|\frac{\mathrm{d}}{\mathrm{d}t}\ln h^{(m)}(t)| = \frac{1}{\omega}\frac{|h^{(m+1)}(t)|}{|h^{(m)}(t)|} \ll 1$, then the summation in the first line of Eq. (A.29) can be convergent quickly and the residual terms can be ignored. In our situation here when calculating $\tilde{\alpha}^{[1']}(t)$, we need to assume $\omega_r/\omega_c \ll 1$ since $\tilde{\alpha}^{[0]}(t)$ and $\Lambda_1^{[0]}(t)$ contains frequency components related to ω_r. In fact, as mentioned in former sections we always assume that $\kappa \ll \chi\alpha \ll \dot{\phi}_d \sim \chi\alpha^2 \ll \omega_r$ to guarantee high-fidelity operations, which means the assumption of $\omega_r/\omega_c \ll 1$ implies $\kappa/\omega_c \ll 1$ and $\chi\alpha^2/\omega_c \ll 1$. Later on, we simply ignore contributions from the factor $e^{-\kappa(t-t')/2}$ since we want to work in the regime that $\kappa t \le \kappa T \ll 1$ to achieve high fidelity. Also, it varies

slowly over time, so it will not affect any scaling analysis when taking time derivatives of this.

With the trick mentioned above, we can write $\tilde{\alpha}^{[1']}$ in the following form:

$$\tilde{\alpha}^{[1']} = \tilde{\alpha}^{[1',0]} + \tilde{\alpha}^{[1',1]}e^{2i\omega_c t} + \tilde{\alpha}^{[1',2]}e^{4i\omega_c t} + \tilde{\alpha}^{[1',-1]}e^{-2i\omega_c t}. \tag{A.30}$$

We first analyze $\tilde{\alpha}^{[1',1]}$ since $\Lambda_1^{*[0]}$ dominates in $g(t)$. For the lowest order term ($n = 0$) in Eq. (A.29), notice that $\Lambda_1^{[0]}$ contains $i\dot{\tilde{\alpha}}^{[0]}$ term where $\tilde{\alpha}^{[0]}(t) = \alpha(t)f(t)$, we have the leading contribution in $\tilde{\alpha}^{[1',1]}$ as $\tilde{\alpha}^{[1',1]} \simeq \frac{i\alpha^*\dot{f}^*}{2\omega_c}e^{2i\phi_d^{[0]}} \sim O(\frac{\alpha\omega_r}{\omega_c})$. It also contains $O(\frac{\chi\alpha^3}{\omega_c})$ terms that come from those $O(\chi\alpha^3)$ parts in $g(t)$ and those implicitly in $\Lambda_1^{[0]}$. Also, since we assumed $\alpha(t)$ varies slowly that $\dot{\alpha}(t) \sim O(\alpha/T) \sim O(\chi\alpha^2)$, it gives a correction to the estimation of $\tilde{\alpha}^{[1',1]}$ with $O(\frac{\dot{\alpha}f}{\omega_c}) \sim O(\frac{\chi\alpha^2}{\omega_c})$. Besides, the second order term ($n = 1$) in Eq. (A.29) can give a contribution of $O(\frac{\alpha\ddot{f}}{\omega_c^2}) \sim O(\frac{\alpha\omega_r^2}{\omega_c^2})$. Using similar argument, we can show that $\tilde{\alpha}^{[1',2]} \sim O(\frac{\chi\alpha^3}{\omega_c})$ and $\tilde{\alpha}^{[1',-1]} \sim O(\frac{\chi\alpha^3}{\omega_c})$. The analysis for $\tilde{\alpha}^{[1',0]}$ is a little different. As mentioned in the main text, in practice we start with $\alpha(0) = 0$ while $\dot{\alpha} \sim O(\alpha/T) \sim O(\chi\alpha^2)$. Therefore, we have $\Lambda_1^{[0]}(0) \sim O(\chi\alpha^2)$ so that $\tilde{\alpha}^{[1',0]} \sim O(\frac{\chi\alpha^2}{\omega_c})$. Other terms from $g(t)$ will only cause higher-order corrections.

Later we want to show how we can partially compensate for the effects of those fast-oscillating terms, and even with this technique, we still need further assumptions that

$$\frac{\chi\alpha^3}{\omega_c} \ll 1, \quad \frac{\alpha\omega_r^2}{\omega_c^2} \ll 1, \tag{A.31}$$

so that the residual effects could be small.

We focus on the change of the Hamiltonian under the new displaced frame with $\tilde{\alpha} = \tilde{\alpha}^{[0]} + \tilde{\alpha}^{[1']}$. We want first to check those extra terms (in comparison with $\hat{H}'_{\rm dr}[\tilde{\alpha}(t)]$) which are induced by $\tilde{\alpha}^{[1']}$ and do not contain fast oscillating factor (factor $e^{ik\omega_c t}$ where $|k| \in \mathbb{N}_+$). We try to show the dominant effect from each term in Eq. (A.25), and evaluate their contributions

under the assumptions of Eq. (A.31).

For the corrections with cubic $\hat{a}$ or $\hat{a}^\dagger$ terms (including $\hat{a}^{\dagger 3}$, $\hat{a}^{\dagger 2}\hat{a}$ and their Hermitian conjugate), their coefficients will be of $O(\chi\tilde{\alpha}^{[1']})$, whose amplitudes scale at most $O(\chi\alpha\frac{\omega_r}{\omega_c})$. Even without the fast-oscillating factor, the integration of their amplitudes over time will give us $\chi\alpha T\frac{\omega_r}{\omega_c} \sim \frac{\omega_r}{\omega_c} \ll 1$ (where we have used $T \sim \frac{1}{\chi\alpha}$), which means the contribution from those correction terms to the dynamics are small.

For the correction induced by slowly varied (in comparison with ω_c) $\hat{a}^\dagger\hat{a}$ terms, it can come from either $2\chi|\tilde{\alpha}|^2\hat{a}^\dagger\hat{a}$ or $(\chi\tilde{\alpha}^{*2}e^{2i\phi_d^{[0]}}e^{2i\omega_c t} + \text{c.c.})\hat{a}^\dagger\hat{a}$, where c.c. stands for "complex conjugate". The former will lead extra slowly varied $\hat{a}^\dagger\hat{a}$ terms with coefficients like $(2\chi\tilde{\alpha}^{[0]}\tilde{\alpha}^{*[1',0]} + \text{c.c.}) \sim O(\frac{1}{T}\frac{\chi\alpha^2}{\omega_c})$ and $2\chi|\tilde{\alpha}^{[1',1]}|^2 \sim O(\frac{1}{T}\frac{\alpha\omega_r^2}{\omega_c^2})$, while the coefficient of the dominant correction from the later will be $(2\chi\tilde{\alpha}^{*[0]}\tilde{\alpha}^{*[1',1]}e^{2i\phi_d^{[0]}} + \text{c.c.}) \sim O(\chi\alpha^2\frac{\omega_r}{\omega_c})$. We introduce the $\phi_d^{[1]}(t)$ function and make its time derivative $\dot{\phi}_d^{[1]}(t)$ equal to the summation of all the slowly varied coefficients of the correction terms, as

$$\begin{aligned}\dot{\phi}_d^{[1]}(t) = 2\chi\Big\{ &\sum_{k=-1}^{2} |\tilde{\alpha}^{[1',k]}|^2 + (\tilde{\alpha}^{[0]}\tilde{\alpha}^{*[1',0]} + \text{c.c.}) \\ &+ \{[(\tilde{\alpha}^{*[0]} + \tilde{\alpha}^{*[1',0]})\tilde{\alpha}^{*[1',1]} + \tilde{\alpha}^{*[1',2]}\tilde{\alpha}^{*[1',-1]}]e^{2i\phi_d^{[0]}} + \text{c.c.}\}\Big\}.\end{aligned} \tag{A.32}$$

Apparently we still have $\dot{\phi}_d^{[1]} \sim O(\chi\alpha^2\frac{\omega_r}{\omega_c})$. This term may not be ignored, and later we will show that we can compensate for that by changing the frequency $\omega_1(t)$ of the linear drive.

Then we discuss the corrections for slowly varied (in comparison with ω_c) $\hat{a}^{\dagger 2}$ terms which come from $\frac{\chi}{2}(\tilde{\alpha} + \tilde{\alpha}^* e^{2i\phi_d^{[0]}}e^{2i\omega_c t})^2\hat{a}^{\dagger 2}$. Those corrections are

$$\begin{aligned}\frac{\chi}{2}\{&(2\tilde{\alpha}^{[0]} + \tilde{\alpha}^{[1',0]} + \tilde{\alpha}^{*[1',1]}e^{2i\phi_d^{[0]}})(\tilde{\alpha}^{[1',0]} + \tilde{\alpha}^{*[1',1]}e^{2i\phi_d^{[0]}}) \\ &+ 2[\tilde{\alpha}^{*[1',1]} + (\tilde{\alpha}^{*[0]} + \tilde{\alpha}^{*[1',0]})e^{2i\phi_d^{[0]}}](\tilde{\alpha}^{*[1',-1]} + \tilde{\alpha}^{*[1',2]}e^{2i\phi_d^{[0]}})\}\hat{a}^{\dagger 2}.\end{aligned} \tag{A.33}$$

The dominant coefficient of the correction will be $\chi\tilde{\alpha}^{[0]}\tilde{\alpha}^{*[1',1]}e^{2i\phi_d^{[0]}} \sim O(\frac{\chi\alpha^2\omega_r}{\omega_c})$. Although

the amplitude of this term might be large, we notice that $\tilde{\alpha}^{[1',1]} \simeq \frac{i\alpha^* \dot{f}^*}{2\omega_c} e^{2i\phi_d^{[0]}}$ with corrections of the order $O(\frac{\chi\alpha^2}{\omega_c})$ and $O(\frac{\alpha\omega_r^2}{\omega_c^2})$. As a result, $\chi\tilde{\alpha}^{[0]}\tilde{\alpha}^{*[1',1]}e^{2i\phi_d^{[0]}} \simeq -i\frac{\chi\alpha^2 f\dot{f}}{2\omega_c}$. We notice that $\frac{f\dot{f}}{\omega_c}$ is a periodic function with period $\delta T = \frac{2\pi}{\omega_r}$. By averaging within each period, we have $\overline{f\dot{f}} = \int_0^{\delta T} f\dot{f}\,\mathrm{d}t = 0$. So, $f\dot{f}$ is a linear superposition of fast oscillating $e^{ik\omega_r t}$ functions with nonzero integer k. We can now use the effective Hamiltonian argument we brought up previously. The commutator between this and the 0-th order $[\chi(\tilde{\alpha}^{[0]})^2\hat{a}^{\dagger 2} + \mathrm{h.c.}]$ will give an effective term with amplitude scaled as $\frac{1}{\omega_r}(\frac{\chi\alpha^2\omega_r}{\omega_c})(\chi\alpha^2) \sim \frac{\chi\alpha^3}{\omega_c}\frac{1}{T}$, which is negligible after integration over time T under our assumptions Eq. (A.31). However, this trick does not apply to higher order correction terms of $\tilde{\alpha}^{*[1',1]}$ like $-\frac{\alpha\ddot{f}}{4\omega_c^2} \sim O(\frac{\alpha\omega_r^2}{\omega_c^2})$, since it may lead to slowly varied corrections (compared with ω_r) in $\chi\tilde{\alpha}^{[0]}\tilde{\alpha}^{*[1',1]}e^{2i\phi_d^{[0]}}$. But with the second assumption in Eq. (A.31), after integration over time we still have $\chi\alpha\frac{\alpha\omega_r^2}{\omega_c^2}T \ll 1$, which means the contribution from the higher order correction of $\tilde{\alpha}^{*[1',1]}$ can still be ignored. We can similarly argue that other terms in Eq. (A.33) can be ignored since the amplitudes of their coefficients are also too small.

Now we start the procedure of partial compensation. We adjust the 1-photon driving frequency such that $\phi_d = \phi_d^{[0]} + \phi_d^{[1]}$. Therefore, we choose a new correction $\tilde{\alpha}^{[1]}$ for the displaced frame which satisfies

$$\begin{aligned} i\dot{\tilde{\alpha}}^{[1]} + i\kappa\tilde{\alpha}^{[1]}/2 = {} & [\Lambda_1^{*[0]} + \chi\tilde{\alpha}^{*[0]}(|\tilde{\alpha}^{[0]}|^2 + 1)]e^{2i[\phi_d^{[0]}(t)+\phi_d^{[1]}(t)]}e^{2i\omega_c t} \\ & + \frac{\chi}{3}(\tilde{\alpha}^{*[0]})^3 e^{4i[\phi_d^{[0]}(t)+\phi_d^{[1]}(t)]}e^{4i\omega_c t} + \frac{\chi}{3}(\tilde{\alpha}^{[0]})^3 e^{-2i[\phi_d^{[0]}(t)+\phi_d^{[1]}(t)]}e^{-2i\omega_c t}. \end{aligned} \tag{A.34}$$

The only difference between here and Eq. (A.27) is that we use $e^{2ki[\phi_d^{[0]}(t)+\phi_d^{[1]}(t)]}$ instead of $e^{2ki\phi_d^{[0]}(t)}$. We denote

$$\tilde{\alpha}^{[1]} = \tilde{\alpha}^{[1,0]} + \tilde{\alpha}^{[1,1]}e^{2i\omega_c t} + \tilde{\alpha}^{[1,2]}e^{4i\omega_c t} + \tilde{\alpha}^{[1,-1]}e^{-2i\omega_c t}. \tag{A.35}$$

Then, from Eq. (A.29) we can find that

$$\tilde{\alpha}^{[1,1]} = \tilde{\alpha}^{[1',1]} e^{2i\phi_d^{[1]}(t)} + O(\frac{\Lambda_1}{\omega_c}\frac{\dot{\phi}_d^{[1]}}{\omega_c}) = \tilde{\alpha}^{[1',1]} e^{2i\phi_d^{[1]}(t)} + O(\frac{\chi\alpha^3\omega_r^2}{\omega_c^3}). \tag{A.36}$$

Similarly, we can also find that the difference between $\tilde{\alpha}^{[1,k]}$ and $\tilde{\alpha}^{[1',k]} e^{2ik\phi_d^{[1]}(t)}$ (for $k = -1, 0, 2$) are even smaller compared with $O(\frac{\chi\alpha^3\omega_r^2}{\omega_c^3})$ in the scaling aspect. The difference between $\tilde{\alpha}^{[1]}$ and $\tilde{\alpha}^{[1']}$ can lead to extra non-fast-oscillating (compared with ω_c) terms. For example, the extra coefficient induced in this way for $\hat{a}^\dagger\hat{a}$ will be at most the same scaling as $[2\chi\tilde{\alpha}^{*[0]}(\tilde{\alpha}^{*[1,1]} e^{2i(\phi_d^{[0]}+\phi_d^{[1]})} - \tilde{\alpha}^{*[1',1]} e^{2i\phi_d^{[0]}}) + \text{c.c.}] \sim O(\frac{\chi\alpha^3\omega_r^2}{T\omega_c^3})$, which will be far less that 1 and therefore can be ignored after integration over time. We can use the same way to argue that the extra effect from the non-fast-oscillating correction of $\hat{a}^{\dagger 2}$ term can be ignored, due to the similarity in the structure of $\tilde{\alpha}^{[1]}$ and $\tilde{\alpha}^{[1']}$.

The coefficient of slowly varied correction (in comparison with ω_c) for $\hat{a}^\dagger$ term will be at most $O(\chi\alpha^2\tilde{\alpha}^{[1,1]}) \sim O(\chi\alpha^3\frac{\omega_r}{\omega_c})$. To deal with it, we introduce a correction for $\Lambda_1 = \Lambda_1^{[0]} + \Lambda_1^{[1]}$, where

$$\begin{aligned}
&\Lambda_1^{[1]} = \\
&-\chi[(|\tilde{\alpha}^{[0]} + \tilde{\alpha}^{[1,0]}|^2 + |\tilde{\alpha}^{[1,1]}|^2 + |\tilde{\alpha}^{[1,2]}|^2 + |\tilde{\alpha}^{[1,-1]}|^2)(\tilde{\alpha}^{[0]} + \tilde{\alpha}^{[1,0]} + \tilde{\alpha}^{*[1,1]} e^{2i\phi_d}) - |\tilde{\alpha}^{[0]}|^2\tilde{\alpha}^{[0]}] \\
&-\chi[\tilde{\alpha}^{[1,1]}(\tilde{\alpha}^{*[0]} + \tilde{\alpha}^{*[1,0]}) + (\tilde{\alpha}^{[0]} + \tilde{\alpha}^{[1,0]})\tilde{\alpha}^{*[1,-1]} + \tilde{\alpha}^{[1,2]}\tilde{\alpha}^{*[1,1]}](\tilde{\alpha}^{[1,-1]} + \tilde{\alpha}^{*[1,2]} e^{2i\phi_d}) \\
&-\chi[\tilde{\alpha}^{*[1,1]}(\tilde{\alpha}^{[0]} + \tilde{\alpha}^{[1,0]}) + (\tilde{\alpha}^{*[0]} + \tilde{\alpha}^{*[1,0]})\tilde{\alpha}^{[1,-1]} + \tilde{\alpha}^{*[1,2]}\tilde{\alpha}^{[1,1]}][\tilde{\alpha}^{[1,1]} + (\tilde{\alpha}^{*[0]} + \tilde{\alpha}^{*[1,0]}) e^{2i\phi_d}] \\
&-\chi[(\tilde{\alpha}^{[0]} + \tilde{\alpha}^{[1,0]})\tilde{\alpha}^{*[1,2]} + \tilde{\alpha}^{[1,-1]}\tilde{\alpha}^{*[1,1]}](\tilde{\alpha}^{[1,2]} + \tilde{\alpha}^{*[1,-1]} e^{2i\phi_d}) + \chi\tilde{\alpha}^{*[1,1]} e^{2i\phi_d} \\
&+[\dot{\phi}_d^{[0]}\tilde{\alpha}^{[1,0]} + \dot{\phi}_d^{[1]}(\tilde{\alpha}^{[0]} + \tilde{\alpha}^{[1,0]})] \\
&-\chi[\tilde{\alpha}^{*[1,2]}(\tilde{\alpha}^{*[0]} + \tilde{\alpha}^{*[1,0]})^2 + 2\tilde{\alpha}^{*[1,2]}\tilde{\alpha}^{*[1,1]}\tilde{\alpha}^{*[1,-1]} + (\tilde{\alpha}^{*[1,1]})^2(\tilde{\alpha}^{*[0]} + \tilde{\alpha}^{*[1,0]})] e^{4i\phi_d} \\
&-\chi[2\tilde{\alpha}^{[1,-1]}(\tilde{\alpha}^{[0]} + \tilde{\alpha}^{[1,0]})\tilde{\alpha}^{[1,2]} + \tilde{\alpha}^{[1,-1]}(\tilde{\alpha}^{[1,1]})^2 + (\tilde{\alpha}^{[0]} + \tilde{\alpha}^{[1,0]})^2\tilde{\alpha}^{[1,1]}] e^{-2i\phi_d}.
\end{aligned} \tag{A.37}$$

Here $\Lambda_1^{[1]} \sim O(\chi\alpha^3\frac{\omega_r}{\omega_c})$ is chosen to fully absorb those non-fast-oscillating corrections. It is still worth to mention that after this change there will be extra $\Lambda_1^{*[1]}e^{2i\phi_d(t)}e^{2i\omega_c t}$ and some other fast-oscillating terms with amplitude scaled at most $O(\chi\alpha^3\frac{\omega_r}{\omega_c})$ in the new Hamiltonian, since all others with larger amplitudes have been absorbed due to Eq. (A.34).

Further, we can introduce a correction $\tilde{\alpha}^{[2]} \sim O(\chi\alpha^3\frac{\omega_r}{\omega_c^2}) \ll 1$ and repeat all the steps we did (including $\phi_d^{[2]}$ and $\Lambda_1^{[2]}$ compensation) so that the contribution from $\tilde{\alpha}^{[2]}$ are too small and we can ignore this. After k rounds of all those corrections, terms $\hat{a}^\dagger$ with coefficients oscillating in the speed of ω_c will be at most $O[\alpha\omega_r(\frac{\chi\alpha^2}{\omega_c})^k]$, which will be small compared with $\frac{1}{T}$ for $k \geq 3$ under the assumptions in Eq. (A.31).

Finally, we talk about the effects from those fast-oscillating terms (under ω_c) of $\hat{a}^{\dagger 2}$, $\hat{a}^\dagger\hat{a}$ and $\hat{a}^2$ in Eq. (A.25), whose coefficients are at most $O(\chi\alpha^2)$. We denote the summation of them as $\hat{H}_{\mathrm{NR},2}(t)$, which can be written as

$$\begin{aligned}
&\hat{H}_{\mathrm{NR},2}(t) = \\
&\left\{\left[\frac{\chi}{2}[(2|\tilde{\alpha}^{[0]}|^2+1)e^{2i\phi_d}e^{2i\omega_c t} + (\tilde{\alpha}^{*[0]})^2 e^{4i\phi_d}e^{4i\omega_c t}] + O(\frac{\chi\alpha^2\omega_r}{\omega_c}e^{2ik\omega_c t})\right]\hat{a}^{\dagger 2} + \text{h.c.}\right\} \\
&+ \left[[\chi(\tilde{\alpha}^{*[0]})^2 e^{2i\phi_d}e^{2i\omega_c t} + \text{c.c.}] + O(\frac{\chi\alpha^2\omega_r}{\omega_c}e^{2ik\omega_c t})\right]\hat{a}^\dagger\hat{a}.
\end{aligned} \tag{A.38}$$

We denote $\hat{U}_{\mathrm{NR},2}(t) := \mathcal{T}\exp(-i\int_0^t \hat{H}_{\mathrm{NR},2}(t')\,\mathrm{d}t')$, then for any Hamiltonian $\hat{H}$ we can perform a frame transformation to get a new one $\hat{H}' = \hat{U}_{\mathrm{NR},2}^\dagger\hat{H}\hat{U}_{\mathrm{NR},2} + i\dot{\hat{U}}_{\mathrm{NR},2}^\dagger\hat{U}_{\mathrm{NR},2} = \hat{U}_{\mathrm{NR},2}^\dagger(\hat{H} - \hat{H}_{\mathrm{NR},2})\hat{U}_{\mathrm{NR},2}$. We can choose $\hat{H}$ as the Hamiltonian after going to the displaced rotating frame with all rounds of corrections. To achieve $\hat{H}'$, we can replace $\hat{a}$ with $\hat{a}(t) := \hat{U}_{\mathrm{NR},2}^\dagger(t)\hat{a}\hat{U}_{\mathrm{NR},2}(t)$ and $\hat{a}^\dagger$ with $\hat{a}^\dagger(t) := \hat{U}_{\mathrm{NR},2}^\dagger(t)\hat{a}^\dagger\hat{U}_{\mathrm{NR},2}(t)$ in $(\hat{H} - \hat{H}_{\mathrm{NR},2})$ (also if we consider loss we need to change $\mathcal{D}[\hat{a}]$ into $\mathcal{D}[\hat{a}(t)]$). From the definition of $\hat{a}(t)$ and $\hat{a}^\dagger(t)$, we

have

$$\frac{\mathrm{d}\hat{a}(t)}{\mathrm{d}t} = (-i)[\chi(2|\tilde{\alpha}^{[0]}|^2+1)e^{2i\phi_d}e^{2i\omega_c t} + \chi(\tilde{\alpha}^{*[0]})^2 e^{4i\phi_d}e^{4i\omega_c t} + O(\frac{\chi\alpha^2\omega_r}{\omega_c}e^{2ik\omega_c t})]\hat{a}^\dagger(t) - i\Big[[\chi(\tilde{\alpha}^{*[0]})^2 e^{2i\phi_d}e^{2i\omega_c t} + \text{c.c.}] + O(\frac{\chi\alpha^2\omega_r}{\omega_c}e^{2ik\omega_c t})\Big]\hat{a}(t), \tag{A.39}$$

and also $\dot{\hat{a}}^\dagger(t) = [\dot{\hat{a}}(t)]^\dagger$ is also equal to a linear superposition of $\hat{a}^\dagger(t)$ and $\hat{a}(t)$. From scaling analysis we know that $\frac{(\chi\alpha^2)^2}{\omega_c}T \sim \frac{\chi\alpha^3}{\omega_c} \ll 1$, which allows us to use Dyson series expansion to solve Eq. (A.39) and only keep the lower order outcomes. Together with Eq. (A.29), we can achieve the solution for $\hat{a}(t)$ and $\hat{a}^\dagger(t)$ approximately with dominant terms in the following:

$$\begin{aligned}\hat{a}(t) \simeq \Big\{1 &- \frac{\chi}{2\omega_c}[(\tilde{\alpha}^{*[0]})^2 e^{2i\phi_d}e^{2i\omega_c t} - \text{c.c.}] \\ &+ \frac{i\chi^2}{\omega_c}\int_0^t \Big[2\Big(|\tilde{\alpha}^{[0]}(t')|^2 + \frac{1}{2}\Big)^2 + \frac{|\tilde{\alpha}^{[0]}(t')|^4}{4} + O\Big(\frac{\alpha^4\omega_r}{\omega_c}\Big)\Big]\,\mathrm{d}t'\Big\}\hat{a} + \\ \Big\{ &- \Big[\frac{\chi(2|\tilde{\alpha}^{[0]}|^2+1)}{2\omega_c}e^{2i\phi_d}e^{2i\omega_c t} + \frac{\chi(\tilde{\alpha}^{*[0]})^2}{4\omega_c}e^{4i\phi_d}e^{4i\omega_c t}\Big] \\ &+ \frac{i\chi^2}{\omega_c}\int_0^t \Big[[\tilde{\alpha}^{[0]}(t')]^2(2|\tilde{\alpha}^{[0]}(t')|^2+1) + O\Big(\frac{\alpha^4\omega_r}{\omega_c}\Big)\Big]\,\mathrm{d}t'\Big\}\hat{a}^\dagger.\end{aligned} \tag{A.40}$$

We can see that the difference between $\hat{a}(t)$ and $\hat{a}$ is linear in $\hat{a}$ and $\hat{a}^\dagger$ whose coefficients are at most $O(\frac{\chi^2\alpha^4 T}{\omega_c})$, which are far less than 1. Let us focus on the terms without $e^{ik\omega_c t}$ factors first. Notice that the leading-order terms come from the time integration of some polynomials of $\tilde{\alpha}^{[0]}$ and $\tilde{\alpha}^{*[0]}$ with at most fourth order, which is denoted as $p(\tilde{\alpha}^{[0]}, \tilde{\alpha}^{*[0]})$ later in general. Also $\tilde{\alpha}^{[0]}(t) = \alpha(t)f(t)$ where $f(t)$ is a periodical function that can be written as a Fourier series with basis $e^{ik\omega_r t}$ ($k \in \mathbb{Z}$). Due to the time integration in $\frac{\chi^2}{\omega_c}\int_0^t p(\tilde{\alpha}^{[0]}, \tilde{\alpha}^{*[0]})\,\mathrm{d}t'$, only the coefficient of $e^{ik\omega_r t}$ with $k = 0$ scales as $O(\frac{\chi^2\alpha^4 T}{\omega_c}) \sim O(\frac{\chi\alpha^3}{\omega_c})$, while terms that oscillate as $e^{ik\omega_r t}$ (with nonzero k) only have amplitudes that scale as at most $O(\frac{\chi^2\alpha^4}{\omega_r\omega_c})$. Therefore, for the corrections that come from $\frac{\chi}{2}(\tilde{\alpha}^{[0]})^2([\hat{a}^\dagger(t)]^2 - \hat{a}^{\dagger 2})$, they contain terms with amplitude $O(\chi\alpha^2 \cdot \frac{\chi\alpha^3}{\omega_c}) \ll O(\chi\alpha^2)$ but oscillate in the form of $f^2(t)$, whose effect is ignorable as we

argued at the beginning of the section. They also contain slowly varied part with coefficients at most $O(\chi\alpha^2 \cdot \frac{\chi^2\alpha^4}{\omega_c\omega_r}) \sim O(\frac{\chi\alpha^2}{\omega_r} \cdot \frac{\chi\alpha^3}{\omega_c} \cdot \frac{1}{T})$, which are again ignorable after integration over time T. We can similarly demonstrate that the corrections coming from $\chi\tilde{\alpha}^{[0]}\tilde{\alpha}^{*[1,1]}e^{2i\phi_d}[\hat{a}^\dagger(t)]^2$ and other high-order terms are negligible.

Besides, after going to the reference frame with $\hat{U}_{\mathrm{NR},2}$, the fast-oscillating (under ω_c) $\hat{a}^{\dagger 2}$ or $\hat{a}^\dagger\hat{a}$ terms left have amplitudes at most $O(\frac{\chi^2\alpha^4}{\omega_c}) \sim \frac{\chi\alpha^3}{\omega_c} \cdot \frac{1}{T})$, which are sufficiently small. Finally, since we work in a small-dimensional blockade subspace with only several excitations as assumed, as well as $\frac{\chi^2\alpha^4 T}{\omega_c} \ll 1$, the corrections due to the frame transformation $\hat{U}^\dagger_{\mathrm{NR},2}(T)$ on the final states are negligible. This concludes all the evidence that under assumptions in Eq. (A.31), the dynamics affected by terms beyond RWA in Eq. (A.25) could be ignorable, in the aspect of scaling analysis.

A.5.3 Both charge and flux drives

Things will be different if we can drive both charge and flux simultaneously. In this case, our 1-photon driving term can be implemented as

$$\begin{aligned} &\mathrm{Re}[\Lambda_1(t)e^{-i\omega_1(t)t}](\hat{a}+\hat{a}^\dagger) + \mathrm{Re}[i\Lambda_1(t)e^{-i\omega_1(t)t}] \cdot i(\hat{a}-\hat{a}^\dagger) \\ =\ &\Lambda_1(t)e^{-i\omega_1(t)t}\hat{a}^\dagger + \Lambda_1^*(t)e^{i\omega_1(t)t}\hat{a}. \end{aligned} \tag{A.41}$$

In this way, our leading non-RWA term in Eq. (A.24) is gone. As a result, the improved hardware controllability can also help to mitigate the non-RWA effects.

A.5.4 Rough lower bound on operation infidelity

We want to point out that the requirement for $\frac{\chi\alpha^3}{\omega_c} \ll 1$ is actually necessary since without it the $\hat{H}_{\mathrm{NR},2}$ term will affect the dynamics a lot. Even if we can directly implement 2-photon drive, or we can do both charge and flux drives, the requirement is still there due to the difference between the original Kerr term $\chi(\hat{a}+\hat{a}^\dagger)^4/12$ and the simplified one $\chi\hat{a}^{\dagger 2}\hat{a}^2/2$

after RWA. We can use the $\frac{\chi\alpha^3}{\omega_c} \ll 1$ condition to derive a lower bound on the infidelity of Fock $|1\rangle$ state preparation using constant α. Since $\alpha < (\omega_c/\chi)^{1/3}$, we have

$$T > \frac{\pi}{2(\omega_c\chi^2)^{1/3}}. \tag{A.42}$$

So, even in the case that power is not a constraint, and we are allowed to design κ_e as small as possible, we still have

$$\epsilon > c_1\kappa_i T > \frac{3\pi(\kappa_i/\chi)}{16(\omega_c/\chi)^{1/3}} = \frac{3\pi(\kappa_i/\chi)^{2/3}}{16Q_i^{1/3}}, \tag{A.43}$$

where we have already put $c_1 = \frac{3}{8}$ in it, and $Q_i = \omega_c/\kappa_i$ is the internal quality factor. We denote $\epsilon_{\min} := \frac{3\pi(\kappa_i/\chi)^{2/3}}{16Q_i^{1/3}}$, which serves as a lower bound of the infidelity that can be achieved.

A.6 Rough lower bound for input power needed

Here we consider some extreme cases to provide a rough lower bound of the input power needed for our protocol. We assume that we could directly implement 2-photon drives so that the infidelity only comes from photon loss that $\epsilon = c_1(\kappa_e + \kappa_i)T$. Besides, we further assume that the 2-photon drives can be implemented in some power-efficient manner such that we only need to consider the power consumed by 1-photon drives. However, from Eq. (2.3) we still have $\Lambda_1 \simeq 2\chi\alpha^3$, and therefore the input power we need scales as $P_{\text{in}} \propto \alpha^6$. To prepare $|1\rangle$ state with constant α, the power-dependent infidelity satisfies the following equation

$$\epsilon = c_1\frac{\pi}{2}\frac{\kappa_e + \kappa_i}{\kappa_e^{1/6}}\left(\frac{4\hbar\omega_c}{P_{\text{in}}\chi^4}\right)^{1/6}. \tag{A.44}$$

We can choose $\frac{\kappa_e}{\kappa_i} = \frac{1}{5}$ to further minimize ϵ when κ_i is fixed, and therefore achieve the

Table A.1: Performance analysis of our protocol in photonic crystals designed in [119], which includes the power needed to achieve 90% fidelity of Fock state $|1\rangle$ as well as the self-consistency check for RWA requirements.

Cavity design	Nonlinear materials	$\omega_c/(2\pi)$	$\lvert\chi\rvert/(2\pi)$	$\kappa_i/(2\pi)$	$\epsilon_{\min}$	$P_{\text{in},0.9}$	$(\omega_r/\omega_c)_{0.9}$	$(\alpha\omega_r^2/\omega_c^2)_{0.9}$	$(\chi\alpha^3/\omega_c)_{0.9}$
Tip	PIC	521 THz	3.5 THz	0.52 GHz	$\chi \gg \kappa_i$				
	ITO	255 THz	1.9 GHz	0.26 GHz	1.5×10^{-3}	1.3×10^{-5} W	5.8×10^{-5}	3.3×10^{-9}	7.1×10^{-6}
Bridge	GaAs	283 THz	2.7 MHz	0.19 GHz	0.088	4.3×10^{8} W	0.46	112	1.3
	Ge	94.6 THz	9.9 kHz	63 MHz	1.76	$\epsilon_{\min} > 1$			

error scaling $\epsilon \propto \kappa_i^{5/6}/(P_{\text{in}}^{1/6}\chi^{2/3})$. Equivalently, the rough lower bound of the input power P_{in} for a given infidelity ϵ satisfies

$$P_{\text{in}} = \frac{4(3\pi c_1)^6}{5^5}\hbar\omega_c\frac{\kappa_i^5}{\chi^4\epsilon^6}. \tag{A.45}$$

A.7 Feasibility on actual platforms

In this section, we try to consider parameters from different kinds of experimental platforms and check if it is feasible to use our protocol to achieve some high-fidelity operations on those devices. For simplicity, we still consider the Fock $|1\rangle$ state preparation task. We will check the requirements for RWA and the input power needed in order to achieve a certain fidelity derived under RWA. It is worth mentioning that, although our RWA requirements are derived under assumptions that $\alpha(t)$ is slowly varied with $\alpha(0) = \alpha(T) = 0$ and $f(t)$ itself is analytic, in this section we still focus on the method with constant α and $f(t)$ defined in Eq. (A.9) to calculate relevant parameters. More physical $\alpha(t)$ and $f(t)$ functions will lead to a change of those c_i coefficients instead of the scaling properties, so we can still use those ideal choices to get a taste of the feasibility of our protocol on different platforms.

A.7.1 Optical ring resonators

In this part, we talk about the possibility of implementing our protocol on optical ring resonators with $\chi^{(3)}$ nonlinearity. We first estimate the self-Kerr χ value with experimentally achievable platforms, and then check the RWA requirements by calculating $\epsilon_{\min}$, which is a rough lower bound for infidelity achieved by perturbative analysis together with parts of the RWA requirements (see Appendix A.5.4).

The self-Kerr parameter χ in optical ring resonators with $\chi^{(3)}$ nonlinearity can be calculated by [190]:

$$\chi = -\frac{\hbar\omega_c^2 c n_2}{2n^2 A_{\text{eff}} L}, \tag{A.46}$$

where ω_c is the frequency of the mode, n and n_2 are the refractive index and nonlinear refractive index correspondingly, and A_{eff} and L are the effective area and length of the waveguide. In the system proposed in Ref. [191], the authors there assumed $\frac{\omega_c c n_2}{n^2 A_{\text{eff}} L} = 3.7 \times 10^{20}$ $\text{W}^{-1}\text{s}^{-2}$ without pointing out which kind of materials they use nor those geometrical parameters of the resonators. Besides, the authors hope to achieve an internal loss rate $\kappa_i = 2\pi \times 25$ MHz, and they mentioned that this assumption is reasonable since the corresponding intrinsic quality factor $Q_i \sim 8 \times 10^6$ is achievable with the silicon-nitride platform. This discussion indicates that they have chosen $\omega_c \approx 2\pi \times 200$ THz. With those parameters, the corresponding $\chi = -2\pi \times 3.90$ Hz. However, when putting these numbers into Eq. (A.43), we can find out that $\epsilon_{\min} = 102 \gg 1$, which indicates that our perturbative analysis should fall and RWA will be significantly violated if we want to use our current protocol to prepare a Fock $|1\rangle$ state.

A similar result is achieved using parameters from another literature [192]. In their simulation, the parameters are chosen as $\gamma_{\text{NL}} := \frac{n_2 \omega_c}{c A_{\text{eff}}} = 1$ $\text{W}^{-1}\text{s}^{-2}$, together with $\omega_c = 2\pi \times 193$ THz, $n = 1.7$, and $L = 400$ µm. With these numbers, we can find that $\chi = -2\pi \times 0.79$ Hz. The intrinsic Q-factor is chosen as 2×10^6, which corresponds to $\kappa_i = 2\pi \times 96.5$ MHz. This gives $\epsilon_{\min} = 1151$, which again indicates that our protocol does not work.

A.7.2 Photonic crystals

Photonic crystals may perform better to increase Kerr nonlinearities with smart design by suppression of effective mode volume. Our discussion in this part builds on the work from Ref. [119], where the authors discussed the χ and κ values that can be possibly achieved using their novel design of nanocavities with ultra-small mode volumes. It is claimed to be promising for achieving a "single-photon Kerr nonlinearity" ($\chi \sim \kappa$) regime. In that work, χ is determined by

$$\chi = -\frac{3\chi^{(3)}\hbar\omega_c^2}{2\epsilon_0 n^4}\frac{V_M}{V_{\text{eff}}^2}, \tag{A.47}$$

where V_{eff} is the effective mode volume, and V_M is another parameter related to cavity design with the unit of volume.

The authors there provided two types of cavity design: tip design and bridge design. With the tip cavity design there, in simulations people could achieve $\frac{QV_M}{V_{\text{eff}}^2} \approx 2\times 10^7\lambda^{-3}$ and quality factor $Q \approx 10^6$. The λ here is the wavelength of the mode. Besides, in that work, those authors assumed that the cavity radiation loss is much larger than material loss, so that the same quality factor applies to cavities with the same design but different materials. Here we simply follow this assumption to derive κ_i.

For organic materials like J aggregate (PIC), it has $|\chi^{(3)}|/n^4 \sim 1.1 \times 10^{-15}\text{m}^2/\text{V}^2$ at $\lambda = 575$ nm, which gives $\chi = -2\pi \times 3.5$ THz and $\kappa_i = 2\pi \times 0.52$ GHz. In this case, χ is several orders of magnitude larger than κ_i, so there will be no need to use our protocol to achieve universal control.

Inorganic materials like indium tin oxide (ITO) has $|\chi^{(3)}|/n^4 \sim 2.12 \times 10^{-17}\text{m}^2/\text{V}^2$ at $\lambda = 1175$ nm. It gives $\chi = -2\pi \times 1.9$ GHz with $\kappa_i = 2\pi \times 0.26$ GHz. The parameters still lie in the strong Kerr-nonlinearity regime as $|\chi|/\kappa_i > 1$, and we only need input power $P_{\text{in}} = 1.3 \times 10^{-5}$ W to achieve 90% fidelity of a single photon state with our protocol.

Similarly, for their bridge cavity design they have $\frac{QV_M}{V_{\text{eff}}^2} \approx 5.5 \times 10^7\lambda^{-3}$ and quality factor $Q \approx 1.5 \times 10^6$, which in general requires a small $|\chi^{(3)}|/n^4$ to achieve the same Kerr

nonlinearity. But, as they mentioned, for usual semiconductor materials like gallium arsenide (GaAs) with $|\chi^{(3)}|/n^4 \sim 0.97 \times 10^{-20}\mathrm{m}^2/\mathrm{V}^2$ at $\lambda = 1.06$ µm, we can get $\chi = -2\pi \times 2.7$ MHz and $\kappa_i = 2\pi \times 0.19$ GHz. In this case, if we want to achieve 90% fidelity at the optimal point we derived in Appendix A.2 under RWA, in fact we can see that $\alpha\omega_r^2/\omega_c^2 = 112 \gg 1$, which means the RWA is actually violated. If people can get a better device with κ_i 10 times smaller (so that $\kappa_i = 2\pi \times 19$ MHz and $\kappa_i/\chi = 7.0$) while other parameters do not change, it is promising to pass the RWA requirement with $\omega_r/\omega_c = 1.5 \times 10^{-3}$, $\chi\alpha^3/\omega_c = 1.3 \times 10^{-3}$ and $\alpha\omega_r^2/\omega_c^2 = 1.1 \times 10^{-4}$ if we want to achieve 90% fidelity. The power needed will be $P = 4.3 \times 10^2$ W.

For bridge cavity design using Germanium (Ge) as nonlinear materials with $|\chi^{(3)}|/n^4 \sim 0.86 \times 10^{-20}\mathrm{m}^2/\mathrm{V}^2$ at $\lambda = 3.17$ µm, we have $\chi = -2\pi \times 9.9$ kHz and $\kappa_i = 2\pi \times 63$ MHz. However, we find that $\epsilon_{\min} = 1.76 > 1$, which rules out the possibility of using our protocol due to the violation of RWA.

A.8 Derivation of continuous version formulae

In this section, we will use the Magnus expansion method to derive some "continuous version" formulae including Eq. (2.12) and (A.7).

We first use Magnus expansion to compute $\hat{U} = \mathcal{T}\exp[-i\int_0^{\delta T}\hat{H}(t)\,\mathrm{d}t]$ up to fourth order explicitly with $\hat{H}(t) = \hat{H}(\delta T - t)$. Notice that

$$\hat{U} = \exp\left\{\sum_{m=1}^{4}(-i)^m\hat{\Omega}_m + O[(\delta T)^5]\right\}, \tag{A.48}$$

where

$$\begin{aligned}
\hat{\Omega}_1 &= \int_0^{\delta T} \hat{H}_1 \, \mathrm{d}t_1, \quad \hat{\Omega}_2 = \frac{1}{2} \int \mathrm{d}^2 t \; [\hat{H}_1, \hat{H}_2], \\
\hat{\Omega}_3 &= \frac{1}{6} \int \mathrm{d}^3 t \; ([\hat{H}_1, [\hat{H}_2, \hat{H}_3]] + [\hat{H}_3, [\hat{H}_2, \hat{H}_1]]), \\
\hat{\Omega}_4 &= \frac{1}{12} \int \mathrm{d}^4 t \; ([[[\hat{H}_1, \hat{H}_2], \hat{H}_3], \hat{H}_4] + [\hat{H}_1, [[\hat{H}_2, \hat{H}_3], \hat{H}_4]] \\
&\quad + [\hat{H}_1, [\hat{H}_2, [\hat{H}_3, \hat{H}_4]]] + [\hat{H}_2, [\hat{H}_3, [\hat{H}_4, \hat{H}_1]]]).
\end{aligned} \tag{A.49}$$

Here we use abbreviations for $\hat{H}_k := \hat{H}(t_k)$ and

$$\int \mathrm{d}^k t := \int_{0 \le t_k \le t_{k-1} \le \cdots \le t_1 \le \delta T} \cdots \int \mathrm{d}t_1 \ldots \mathrm{d}t_k. \tag{A.50}$$

Making use of the property $\hat{H}(t) = \hat{H}(\delta T - t)$, we can further achieve:

$$\hat{\Omega}_2 = \frac{1}{2} \int \mathrm{d}^2 t \; [\hat{H}_2, \hat{H}_1] = -\hat{\Omega}_2 = 0 \tag{A.51}$$

$$\int \mathrm{d}^4 t \; [[[\hat{H}_1, \hat{H}_2], \hat{H}_3], \hat{H}_4] = \int \mathrm{d}^4 t \; [[[\hat{H}_4, \hat{H}_3], \hat{H}_2], \hat{H}_1] = -\int \mathrm{d}^4 t \; [\hat{H}_1, [\hat{H}_2, [\hat{H}_3, \hat{H}_4]]] \tag{A.52}$$

Then, from Jacobi identity we have

$$\begin{aligned}
\int \mathrm{d}^4 t \; [\hat{H}_1, [[\hat{H}_2, \hat{H}_3], \hat{H}_4]] &= \int \mathrm{d}^4 t \; -[\hat{H}_4, [\hat{H}_1, [\hat{H}_2, \hat{H}_3]]] - [[\hat{H}_2, \hat{H}_3], [\hat{H}_4, \hat{H}_1]] \\
&= \int \mathrm{d}^4 t \; -[\hat{H}_1, [\hat{H}_4, [\hat{H}_3, \hat{H}_2]]] - [[\hat{H}_2, \hat{H}_3], [\hat{H}_4, \hat{H}_1]] \\
&= \int \mathrm{d}^4 t \; -[\hat{H}_1, [[\hat{H}_2, \hat{H}_3], \hat{H}_4]] - [[\hat{H}_2, \hat{H}_3], [\hat{H}_4, \hat{H}_1]] \\
&= -\frac{1}{2} \int \mathrm{d}^4 t \; [[\hat{H}_2, \hat{H}_3], [\hat{H}_4, \hat{H}_1]].
\end{aligned} \tag{A.53}$$

Similarly, we have

$$\int \mathrm{d}^4 t\ [\hat{H}_2, [\hat{H}_3, [\hat{H}_4, \hat{H}_1]]] = -\frac{1}{2}\int \mathrm{d}^4 t\ [[\hat{H}_4, \hat{H}_1], [\hat{H}_2, \hat{H}_3]], \tag{A.54}$$

which will lead to $\hat{\Omega}_4 = 0$.

In summary, we have

$$\begin{aligned} \hat{U} &= \exp\left\{-i\hat{\Omega}_1 + i\hat{\Omega}_3 + O[(\delta T)^5]\right\}, \\ \hat{U}_{\mathrm{tar}} &= \exp\left[-i\int_0^{\delta T} \hat{H}(t)\,\mathrm{d}t\right] = \exp(-i\hat{\Omega}_1). \end{aligned} \tag{A.55}$$

Since $\hat{\Omega}_3$ is in the order of $O[(\delta T)^3]$, it is easy to see that

$$\hat{U} - \hat{U}_{\mathrm{tar}} = O[(\delta T)^3], \tag{A.56}$$

which gives the proof of Eq. (2.12).

Similarly, with the Baker–Campbell–Hausdorff (BCH) formula, we can find that

$$\hat{U}_{\mathrm{tar}}^{-1/2}\hat{U}\hat{U}_{\mathrm{tar}}^{-1/2} = \exp\{i\hat{\Omega}_3 + O[(\delta T)^5]\}, \tag{A.57}$$

which gives the proof of the "continuous version" correspondence of Eq. (A.7).

A.9 Estimation of undetermined factors in Eq. (A.13)

In this section, we will show how we estimate the coefficients c_1, c_2 and c_3 in Eq. (A.13) when focusing on a specific task described in Appendix A.2, which is to prepare Fock state $|1\rangle$ with constant α and oscillation function $f(t)$ defined in Eq. (A.9).

A.9.1 Estimation of c_1

In the estimation of c_1, we can assume that the photon blockade is perfectly achieved with the ideal Hamiltonian $\hat{H}_{\mathrm{dr}}^{\mathrm{qb}}$ in Eq. (A.6), so we only need to take care of the infidelity induced by photon loss. We can start with the following Lindblad equation with operators restricted in the two-level blockade subspace:

$$\frac{\mathrm{d}\hat{\rho}}{\mathrm{d}t} = -i[-\chi\alpha\hat{\sigma}_x, \hat{\rho}] + \kappa\mathcal{D}[\hat{\sigma}_-]\hat{\rho}. \tag{A.58}$$

The state is initialized as $\hat{\rho}(0) = |0\rangle\langle 0|$. To prepare the $|1\rangle$ state, we require the evolution to last for $T = \frac{\pi}{2\chi\alpha}$. Since we work in the regime that $\kappa \ll \chi\alpha$, we will treat the loss as a perturbation. We first go to the interaction picture, and then denote $\hat{U}_0(t) = e^{i\chi\alpha t\hat{\sigma}_x}$ as well as $\hat{\rho}_I(t) = \hat{U}_0^\dagger(t)\hat{\rho}(t)\hat{U}_0(t)$. Therefore, the evolution of $\hat{\rho}_I(t)$ satisfies

$$\frac{\mathrm{d}\hat{\rho}_I(t)}{\mathrm{d}t} = \kappa\mathcal{D}[\hat{U}_0^\dagger(t)\hat{\sigma}_-\hat{U}_0(t)]\hat{\rho}_I(t). \tag{A.59}$$

The final state $\hat{\rho}_I(T)$ up to first order in κ is

$$\hat{\rho}_I(T) = \hat{\rho}_I(0) + \kappa\int_0^T \mathcal{D}[\hat{U}_0^\dagger(t)\hat{\sigma}_-\hat{U}_0(t)]\hat{\rho}_I(0)\,\mathrm{d}t. \tag{A.60}$$

Recall our definition for infidelity ϵ in Eq. (A.4), we have

$$\begin{aligned}
\epsilon &= 1 - \langle 1|\,\hat{\rho}(T)\,|1\rangle = 1 - \langle 0|\,\hat{\rho}_I(T)\,|0\rangle \\
&= -\kappa\,\langle 0|\left[\int_0^T \mathcal{D}[\hat{U}_0^\dagger(t)\hat{\sigma}_-\hat{U}_0(t)]\hat{\rho}_I(0)\,\mathrm{d}t\right]|0\rangle \\
&= \frac{3}{8}\kappa T,
\end{aligned} \tag{A.61}$$

where $T = \frac{\pi}{2\chi\alpha}$. This result directly indicates that $c_1 = \frac{3}{8}$ in the protocol we considered.

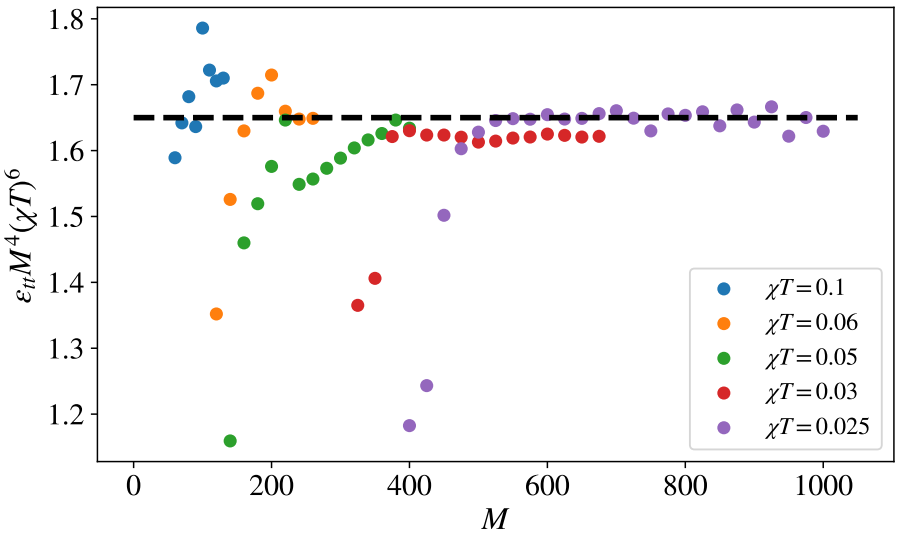

Figure A.3: Determine $c_2 = \epsilon_{\mathrm{tt}} M^4 (\chi T)^6$ numerically by varying M and T. The black dashed line is the reference for $c_2 = 1.65$.

A.9.2 Estimation of c_2

Notice that c_2 is introduced as a factor in the expression of Trotter error $\epsilon_{\mathrm{tt}} = c_2/[M^4(\chi T)^6]$, so we can numerically estimate it by calculating the infidelity of the state preparation under Hamiltonian $\hat{H}'_{\mathrm{dr}}[\tilde{\alpha}(t)]$ evolution while varying M and T. We show the corresponding plot in Fig. A.3 where the data points on it are chosen to satisfy $\epsilon_{\mathrm{tt}} < 0.2$, and find out that in the small χT (to reduce infidelity coming from photon loss) but large M (to reduce ϵ_{tt}) regime we have $c_2 \approx 1.65$. Since we mainly care about the scaling behavior in our protocol instead of exact numbers, we will just use $c_2 \approx 1.65$ as a rough estimation when it is needed.

A.9.3 Estimation of c_3

The derivation of c_3 is straightforward, since we just need to put the exact form of $\tilde{\alpha}(t) = \alpha f(t)$ into Eq. (2.9) and then calculate the power. As we mentioned in Appendix A.2 that $i\dot{\tilde{\alpha}}$ is the dominant term in Λ_1, and due to our specific design of $f(t)$ in Eq. (A.9), we have

$$|\Lambda_1| \simeq |\dot{\tilde{\alpha}}(t)| = |\alpha \dot{f}(t)| = \frac{\pi^2 M \alpha}{T} = \frac{\pi^3 M}{2\chi T^2}, \tag{A.62}$$

in which we have used $\alpha = \frac{\pi}{2\chi T}$. Therefore, we can finally achieve

$$P_{\text{in}} = \frac{|\Lambda_1|^2}{\kappa_e}\hbar\omega_c = \frac{\pi^6}{4}\hbar\omega_c\frac{M^2}{\kappa_e\chi^2 T^4}, \tag{A.63}$$

which directly gives us that $c_3 = \frac{\pi^6}{4}\hbar\omega_c$.

A.10 Generation of control sequence with neural network

In this part, we will discuss the way to use a neural network to generate control pulses in order to solve the imperfect blockade situation due to the assumption that we do not have the direct 2-photon drive in the experiment.

We still first go to the rotating frame and then the displaced frame, without making a constraint between $\alpha(t)$ and $\Lambda_1(t)$. Besides, we can choose $\omega_1(t)$ such that there is no detuning term $(\hat{a}^\dagger\hat{a})$ in the new Hamiltonian. Therefore, we will end up with the following Hamiltonian:

$$\hat{H}_{\text{dr}}^{\text{new}} = \frac{\chi}{2}\hat{a}^{\dagger 2}\hat{a}^2 + \{\chi\alpha^{\text{new}}(t)\hat{a}^\dagger\hat{n} + \frac{\chi}{2}[\alpha^{\text{new}}(t)]^2\hat{a}^{\dagger 2} + \Lambda_1^{\text{new}}(t)\hat{a}^\dagger + \text{h.c.}\}. \tag{A.64}$$

In this case, both $\alpha^{\text{new}}(t)$ and $\Lambda_1^{\text{new}}(t)$ are controllable. We note that, although r does not come into $\hat{H}_{\text{dr}}^{\text{new}}$ explicitly, we still try to achieve the effective photon blockade during evolution by optimizing the fidelity of the unitary operations or state preparation tasks within the subspace spanned by $\{|0\rangle, \ldots, |r\rangle\}$ that we want to block, while adding a penalty function to penalize the average population leakage out of the blockade subspace during evolution. Specifically, we can choose $g_{\text{u}} = \frac{1}{(r+1)T}\int_0^T \text{d}t\,\text{Tr}[\hat{U}^\dagger(t)(\hat{I} - \hat{\Pi}_r)\hat{U}(t)\hat{\Pi}_r]$ as the penalty function in the unitary operation task, or $g_{\text{st}} = \frac{1}{T}\int_0^T \text{d}t\,\langle\psi(t)|\,(\hat{I} - \hat{\Pi}_r)\,|\psi(t)\rangle$ for the state preparation task, where as in the main text $\hat{\Pi}_r = \sum_{n=0}^r |n\rangle\langle n|$ is the projection operator onto the blockade subspace. Here we can understand the dynamics as that, the states that

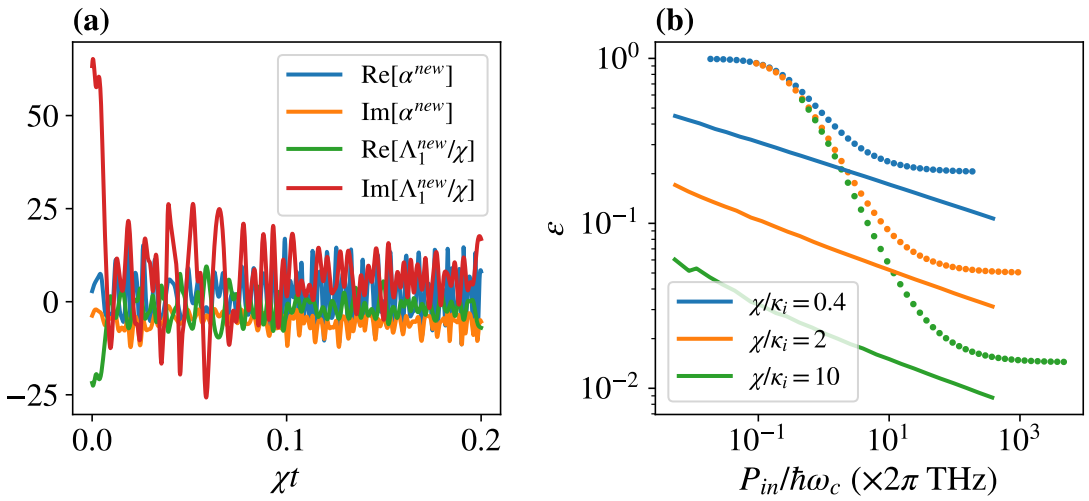

Figure A.4: Control pulses optimization with the neural network approach. (a) The outcome of pulse sequences $\alpha^{\text{new}}(t)$ and $\Lambda_1^{\text{new}}(t)$ from the neural network with fixed $\chi T = 0.2$. (b) The comparison of maximum power needed and infidelities of state preparation between pulses from neural network optimization (shown with dots) with fixed total time T and Trotter-based design (shown with solid lines) with optimized T and M.

initially lie in the blockade subspace will always evolve within it with small leakage, by keeping track of the evolution with a proper time-dependent displaced frame $\alpha^{\text{new}}(t)$. This also provides another advantage that we do not have to take a large dimension cutoff, which is always associated with numerical computation in bosonic systems, since we just have to focus on the dynamics confined in a small subspace with slight leakage.

To encode the variables $\alpha^{\text{new}}(t)$ and $\Lambda_1^{\text{new}}(t)$ in the optimization problem, we use sinusoidal representation neural network with three layers that takes t as an input and returns the real and imaginary part of the corresponding control parameter by $\Phi^{(t)} = \mathbf{W}_2 \sin(\mathbf{W}_1 \sin(\mathbf{W}_0 t + \mathbf{b}_0) + \mathbf{b}_1) + \mathbf{b}_2$, where $\mathbf{W}_0 \in \mathbb{R}^{20\times 1}$, $\mathbf{W}_1 \in \mathbb{R}^{20\times 20}$, and $\mathbf{W}_2 \in \mathbb{R}^{2\times 20}$, and $\mathbf{b}_i$ are vectors with a dimension consistent with their corresponding $\mathbf{W}_i$ [134]. We then maximize $F_{\text{task}}(\boldsymbol{\alpha}) - g(\boldsymbol{\alpha})$ by finding an appropriate set of $\mathbf{W}_i$ and $\mathbf{b}_i$ that specifies the controls, where $F_{\text{task}}(\boldsymbol{\alpha}) = \left|\text{Tr}[\hat{U}_{\text{tar}}^\dagger \hat{U}(T)]\right|^2/(r+1)^2$ for unitary operations and $F_{\text{task}}(\boldsymbol{\alpha}) = \left|\langle\psi_{\text{tar}}|\psi(T)\rangle\right|^2$ for state preparation tasks. To do this, we first fix the total evolution time χT and perform gradient-based optimization until the convergence criterion is

satisfied. After finding $\alpha^{\text{new}}(t)$ and $\Lambda_1^{\text{new}}(t)$ at the end of the optimization, we compute the infidelity by taking the photon loss $(\kappa_i + \kappa_e)\mathcal{D}[\hat{a}]$ into account, and also the input power needed. We fix κ_i and vary κ_e to find infidelity as a function of the required power.

Note that, as opposed to the Trotter-based method with the sine-basis ansatz of Eq. (2.7), we do not impose $\alpha(0) = \alpha(T) = 0$ for the neural network ansatz to allow for more flexibility. These potential non-zero values at the beginning and end of the protocol can be realized by fast displacements. It is also possible to impose a constraint on these values by adding another penalty term to the cost function.

To compare this method with the Trotter scheme, we consider the single-photon Fock $|1\rangle$ state preparation task. Unlike the protocol examined in Appendix A.2, where the total time T can be optimized analytically based on χ, κ_i, and the input power allowed, here we fix T without any optimization attempt, and vary κ_i to compare the infidelity as a function of the power for the pulse sequence generated by the neural network with those obtained in Appendix A.2 based on analytical intuition. We show an example of the outcome in Fig. A.4.

We believe that the inferior performance of the neural network is due to the complexity of finding highly oscillatory functions (as also seen in the Trotter-based scheme) required for relaxing the $\Lambda_2(t)$ requirement. Solving ordinary differential equations with highly oscillatory functions is computationally costly and increases the optimization time. Therefore, even though, in principle, a highly expressive neural network is capable of representing the solutions found using the Trotter scheme, it is difficult to find those (or better) solutions directly using our optimization techniques. It is interesting to devise more efficient optimization methods using prior information from the success of the Trotter-based scheme. We leave these questions for future work.

APPENDIX B

APPENDICES FOR CHAPTER 3

B.1 Lower order Hamiltonian stabilization of pair-cat code

In the main part, we have shown that the Hamiltonian $\hat{H} = -K(\hat{a}^{\dagger 2}\hat{b}^{\dagger 2} - \gamma^{*4})(\hat{a}^2\hat{b}^2 - \gamma^4)$ can stabilize the pair-cat code space. Here we seek for Hamiltonian with a lower maximum order N that can also provide γ-dependent stabilization of the pair-cat code space. Specifically, we hope to find the Hamiltonian

$$\hat{H} = \sum_{m+n+p+q \leq N} f_{mn,pq} \hat{a}^{\dagger m} \hat{a}^n \hat{b}^{\dagger p} \hat{b}^q, \tag{B.1}$$

such that both $|\gamma_{\Delta_0}\rangle$ and $|(i\gamma)_{\Delta_0}\rangle$ are eigenstates of $\hat{H}$ with the same eigen-energy. Without loss of generality, we can specify the eigen-energy as 0. So, we require

$$\hat{H}\,|\gamma_{\Delta_0}\rangle = \hat{H}\,|(i\gamma)_{\Delta_0}\rangle = 0. \tag{B.2}$$

To make use of this requirement, we need to find a set of linearly independent states such that $\hat{H}\,|\gamma_{\Delta_0}\rangle$ and $\hat{H}\,|(i\gamma)_{\Delta_0}\rangle$ can be written as a linear superposition of them. We notice that $\{\hat{P}_\Delta \hat{a}^{\dagger n}\hat{b}^{\dagger m}\,|\gamma,\gamma\rangle\,|n,m \in \mathbb{N}, \Delta \in \mathbb{Z}\}$ are linearly dependent due to the following identity:

$$\hat{P}_\Delta \hat{a}^{\dagger n}\hat{b}^{\dagger m}\,|\gamma,\gamma\rangle = \hat{P}_\Delta \hat{a}^{\dagger (n+1)}\hat{b}^{\dagger (m-1)}\,|\gamma,\gamma\rangle + \frac{\Delta + n - m + 1}{\gamma}\hat{P}_\Delta \hat{a}^{\dagger n}\hat{b}^{\dagger (m-1)}\,|\gamma,\gamma\rangle\,. \tag{B.3}$$

With this recursive formula, every $\hat{P}_\Delta \hat{a}^{\dagger n}\hat{b}^{\dagger m}\,|\gamma,\gamma\rangle$ can be written as a linear superposition of $\{\hat{P}_\Delta \hat{a}^{\dagger k}\,|\gamma,\gamma\rangle\,|k \leq n+m\}$. As a result, we can write $\hat{H}\,|\gamma_{\Delta_0}\rangle$ as a linear superposition of $\hat{P}_\Delta \hat{a}^{\dagger k}\,|\gamma,\gamma\rangle$ with different k and Δ, and Eq. (B.2) requires all the coefficients are 0. Similarly, we can write $\hat{H}\,|(i\gamma)_{\Delta_0}\rangle$ under $\{\hat{P}_\Delta \hat{a}^{\dagger k}\,|i\gamma,i\gamma\rangle\,|k \in \mathbb{N}; \Delta \in \mathbb{Z}\}$ to get another set

of linear equations on $f_{mn,pq}$.

Besides, the Hermiticity of $\hat{H}$ requires that

$$f_{mn,pq} = f^*_{nm,qp}. \tag{B.4}$$

Eq. (B.2) and (B.4) form a set of linear equations of the real and imaginary part of $f_{mn,pq}$, and we hope to find all of its solutions with numerical help. In general, the solutions for $f_{mn,pq}$ should be a function of γ and Δ_0. However, even for one specific Δ_0 we cannot find any γ-dependent solution up to $N = 6$.

Here we call a set of solutions γ-independent if (with a proper overall factor since the set of linear equations is homogeneous) all of $f_{mn,pq}$ can be written as independent of γ. For example, when $\Delta_0 = 0$ we can easily find that $\hat{\Delta} = \hat{b}^\dagger\hat{b} - \hat{a}^\dagger\hat{a}$ is one solution since $\hat{\Delta}\,|\gamma_0\rangle = \hat{\Delta}\,|(i\gamma)_0\rangle = 0$, but all the non-zero coefficients of $\hat{a}^{\dagger m}\hat{a}^n\hat{b}^{\dagger p}\hat{b}^q$ in $\hat{\Delta}$ is either -1 or 1, which are both independent of γ. So, we call this solution γ-independent. Other solutions that do not satisfy the γ-independent criteria are regarded as γ-dependent.

If we further restrict $\hat{H}$ to commute with all the $\hat{\Delta}$, which means the photon number difference of two modes is a conserved quantity, then all the non-zero $f_{mn,pq}$ should satisfy $n + p = m + q$, which means the maximum order in $\hat{H}$ should be an even number. So, our Hamiltonian $\hat{H} = -K(\hat{a}^{\dagger 2}\hat{b}^{\dagger 2} - \gamma^{*4})(\hat{a}^2\hat{b}^2 - \gamma^4)$ with $N = 8$ is the one with lowest order that could provide both nontrivial γ-dependent protection and photon number difference conservation properties that we can find out.

B.2 Structure of the stabilization Hamiltonian

In this appendix, we briefly investigate the eigenstates and eigen-energies of the stabilization Hamiltonian $\hat{H} = -K(\hat{a}^{\dagger 2}\hat{b}^{\dagger 2} - \gamma^{*4})(\hat{a}^2\hat{b}^2 - \gamma^4)$ defined in Eq. (3.16) in large $|\gamma|$ limit.

The first strategy is to perform a displacement operation on the two modes. We denote

$\hat{D}(\gamma,\gamma) = \hat{D}(\gamma) \otimes \hat{D}(\gamma)$ where $\hat{D}(\gamma)$ is the displacement operator. So, in the displaced frame of the two modes, the Hamiltonian is

$$\hat{H}_{\text{disp},\gamma} = \hat{D}(\gamma,\gamma)\hat{H}\hat{D}^{\dagger}(\gamma,\gamma) = -8K|\gamma|^6 \left(\frac{\hat{a}^{\dagger}+\hat{b}^{\dagger}}{\sqrt{2}}\right)\left(\frac{\hat{a}+\hat{b}}{\sqrt{2}}\right) + O(|\gamma|^5). \tag{B.5}$$

We can only keep the first term in $\hat{H}_{\text{disp},\gamma}$ if we just focus on the subspace where in the displaced frame the matrix elements of $\hat{a}$, $\hat{b}$ are far less than γ.

Besides, we can denote $\hat{A} = (\hat{a}+\hat{b})/\sqrt{2}$ and $\hat{B} = (\hat{a}-\hat{b})/\sqrt{2}$, which serve as two new independent modes.

We define $|\psi\rangle$ (with unit norm) as the "asymptotic eigenstate" of an operator $\hat{O}(\gamma)$ if in the large $|\gamma|$ limit $|\psi\rangle$ is parallel with $\hat{O}|\psi\rangle$, or the norm of $\hat{O}|\psi\rangle$ goes to zero in that limit. In this case, if we denote the states $|(n,m)\rangle$ as

$$|(n,m)\rangle := \frac{\hat{A}^{\dagger n}\hat{B}^{\dagger m}}{\sqrt{n!m!}}|0,0\rangle, \tag{B.6}$$

they will be the asymptotic eigenstates of $\hat{H}_{\text{disp},\gamma}$ in the large $|\gamma|$ limit as long as they satisfy either of the following two conditions. The one is $n \geq 1$ and $n, m \ll |\gamma|^2$, while the other one is $n = 0$ and $m = 0, 1$. So, in both of the two cases, $\hat{D}(\gamma,\gamma)|(n,m)\rangle$ are the asymptotic eigenstates of $\hat{H}$. The derivation is also valid for states $\hat{D}(i\gamma, i\gamma)|(n,m)\rangle$ which are also asymptotic eigenstates of $\hat{H}$ as long as n, m satisfy the criteria we just mentioned.

It can also be shown that $|(0,m)\rangle$ are not asymptotic eigenstates of $\hat{H}_{\text{disp},\gamma}$ when $m \geq 2$. We can calculate the angle between the two states $|(0,m)\rangle$ and $\hat{H}_{\text{disp},\gamma}|(0,m)\rangle$, and find out that they are not parallel but actually perpendicular to each other under large $|\gamma|$ limit when $m \geq 2$ because of the lower order corrections in Eq. (B.5).

In fact, this protocol can be generalized by using $\hat{D}(\gamma_1,\gamma_2)$ to find asymptotic eigenstates where $\gamma_1^2\gamma_2^2 = \gamma^4$. The dominant part of the displaced Hamiltonian can be transformed into the form of a single oscillator via Gaussian operations. The energy spacing of the new mode

is $4K|\gamma_1|^2|\gamma_2|^2(|\gamma_1|^2+|\gamma_2|^2)$, which is no less than $8K|\gamma|^6$ due to the constraint between γ_1 and γ_2.

To safely claim that the Hamiltonian in Eq. (3.16) can provide a protection of code space with the $8K|\gamma|^6$ energy gap, we also perform the exact diagonalization of the Hamiltonian with numerical help. We first separate $\hat{H}$ into different subspaces with fixed parity and photon number difference. Specifically, we focus on $\hat{H}_{\mu,\Delta}$ that mentioned in Eq. (3.17), and then numerically calculate the energy gap between $|\mu_{\gamma,\Delta}\rangle$ state and the "first-less-excited" eigenstate of $\hat{H}_{\mu,\Delta}$. In Fig. B.1(a), we can see that in the $\Delta=0$ case we do have $8K|\gamma|^6$ protection of the code space in the large $|\gamma|$ limit.

In general, it is hard to write down the explicit form of all the asymptotic eigenstates of $\hat{H}_{\mu,\Delta}$, but we can see that the state $|\psi_{e1,\mu,\Delta}\rangle$, which can be written as

$$|\psi_{e1,\mu,\Delta}\rangle \simeq \frac{\hat{Q}_{\mu}^{(\Delta)}(\hat{a}^\dagger+\hat{b}^\dagger-2\gamma^*)\,|\gamma,\gamma\rangle}{\sqrt{2\mathcal{N}_{e1,\mu,\Delta}}}, \tag{B.7}$$

is the asymptotic eigenstate of $\hat{H}_{\mu,\Delta}$ with eigen-energy $E_{e1,\mu,\Delta}\simeq -8K|\gamma|^6$. Here $\mathcal{N}_{e1,\mu,\Delta}$ is a normalization factor.

We can also numerically investigate the difference among $E_{e1,\mu,\Delta}$ for different μ and Δ. Like the cat code case, $|E_{e1,+,\Delta}-E_{e1,-,\Delta}|$ is suppressed exponentially with $|\gamma|^2$ due to the exponentially suppressed overlap between any two states with large separation in the γ-plane, which means that the two "first-less-excited" states with the same Δ and different parity are approximately degenerate in large $|\gamma|$ limit. Besides, $|E_{e1,\mu,\Delta}-E_{e1,\mu,0}|\sim O(\Delta^2|\gamma|^2)$ in the large $|\gamma|$ limit with finite Δ, which is a small correction compared with the $O(|\gamma|^6)$ gap. These facts together indicate that we do have the $O(|\gamma|^6)$ energy gap to protect the code space.

Those scaling results may change when considering another limit with finite γ but focusing on the subspaces with Δ as large as possible. However, since typically we prefer to choose

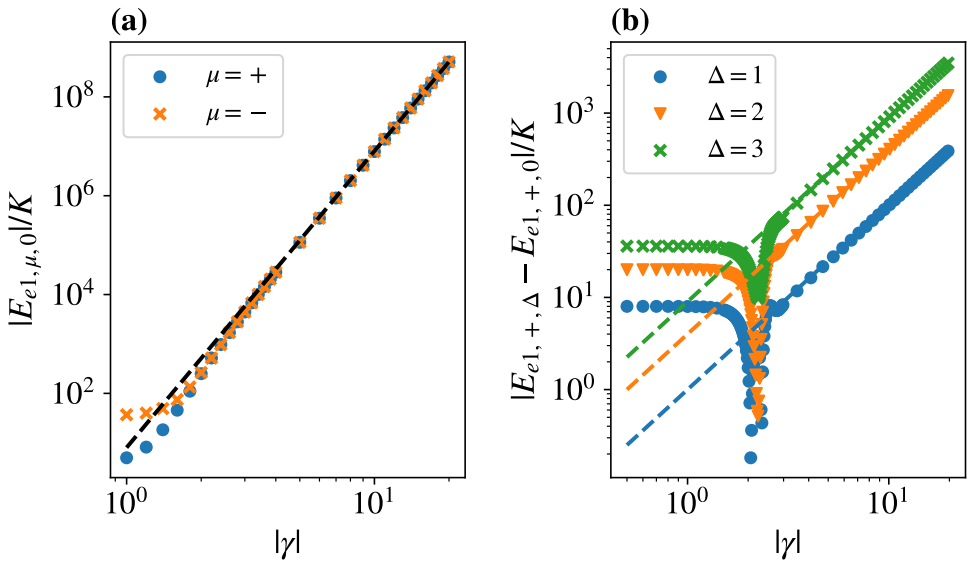

Figure B.1: Properties of the energy spectrum in the pair-cat stabilization Hamiltonian. (a) Energy gap $|E_{e1,\mu,0}|$ between code state $|\mu_{\gamma,\Delta}\rangle$ and the "first-less-excited" eigenstate in $\mu = +$ subspace and $\mu = -$ subspace. Dashed line corresponds to $|E_{e1,\mu,0}|/K = 8|\gamma|^6$. (b) Energy difference between $E_{e1,\mu,\Delta}$ and $E_{e1,\mu,0}$ with different Δ. Dashed lines from bottom to top correspond to $|E_{e1,+,\Delta} - E_{e1,+,0}|/K = \Delta^2|\gamma|^2$ for $\Delta = 1$, 2 and 3, respectively.

$\Delta = 0$ as the code space and the evolution is ideally Δ-preserving, it is difficult for our states to go to the large Δ regime. So, we do not discuss this regime further but just point out this issue.

B.3 Quantum error correction strategies of pair-cat code against photon loss

In this appendix, we will talk about the quantum error correction properties of the pair-cat code where noise only comes from photon loss. In the first part, we consider the system evolves under infinite strength of the dissipative stabilization while suffering from photon loss. We will talk about the recovery strategy and calculate the remaining error after the recovery process. In the second part, we consider the pair-cat code evolves with no stabilization but only photon loss, which means we can simulate the dynamics using a lossy bosonic channel

(LBC). We will compare the results from the two cases. For simplicity, in this appendix we fix $\Delta = 0$ for our code space.

B.3.1 Lossy process with dissipative stabilization

In this part, we consider the situation where the dissipative stabilization has been turned on during the lossy process. Specifically, the evolution channel can be written as

$$\mathcal{E}^D_{\text{evo}}\rho = e^{\mathcal{L}t}\rho = e^{\kappa t\mathcal{L}_D + \kappa_1 t\mathcal{L}_E}\rho, \tag{B.8}$$

where $\mathcal{L}_D = \mathcal{D}[\hat{a}^2\hat{b}^2 - \gamma^4]$ and $\mathcal{L}_E = \mathcal{D}[\hat{a}] + \mathcal{D}[\hat{b}]$. To simplify the derivation, we will consider the extreme situation where $\kappa \to +\infty$, such that for any $t > 0$ we have $e^{\kappa t\mathcal{L}_D}\rho = \mathcal{P}\rho = \hat{P}_D\rho\hat{P}_D$, where $\hat{P}_D = \sum_{\mu,\Delta}|\mu_{\gamma,\Delta}\rangle\langle\mu_{\gamma,\Delta}|$ is the projection operator for the subspace stabilized by the dissipator $\mathcal{D}[\hat{a}^2\hat{b}^2 - \gamma^4]$. Then, because of the following identity [193]:

$$\begin{aligned} e^{(A+B)t} &= e^{At} + \int_0^t \mathrm{d}s\; e^{(A+B)(t-s)}Be^{As} \\ &= e^{At} + \sum_{k=1}^{N}\iint_{\sum_i \tau_i \le t}(\prod_{i=1}^{k}\mathrm{d}\tau_i)e^{A(t-\sum_i\tau_i)}\prod_{i=1}^{k}(Be^{A\tau_i}) \\ &\quad + \iint_{\sum_i \tau_i \le t}(\prod_{i=1}^{N+1}\mathrm{d}\tau_i)e^{(A+B)(t-\sum_i\tau_i)}\prod_{i=1}^{N+1}(Be^{A\tau_i}), \end{aligned} \tag{B.9}$$

we have

$$\mathcal{E}^D_{\text{evo}} = \mathcal{P} + \sum_{k=1}^{N}\frac{(\kappa_1 t)^k}{k!}(\mathcal{P}\mathcal{L}_E\mathcal{P})^k + O[(\kappa_1 t)^{N+1}]. \tag{B.10}$$

Now let us discuss the recovery process. As mentioned in the main text, we should first measure $\hat{\Delta}$ of the final states to extract the syndrome and then decide which operation we should apply. One intuitive way is to assume all the loss errors happen only in one mode, since other loss errors that lead to the same Δ correspond to a higher order of κ_1, which are

less likely to happen in the case that $\kappa_1\gamma^2 t \ll 1$. So, if the final $\Delta > 0$ we will assume that loss only happens in mode $\hat{a}$, and if $\Delta < 0$ we assume that loss only happens in mode $\hat{b}$.

We notice that the code states $|\mu_{\gamma,\Delta}\rangle$ defined in Eq. (3.8) satisfy

$$\hat{a}^k\hat{b}^l\,|\mu_{\gamma,0}\rangle = \gamma^{k+l}\sqrt{\frac{\mathcal{N}_{\mu',k-l}}{\mathcal{N}_{\mu,0}}}\,|\mu'_{\gamma,k-l}\rangle\,, \tag{B.11}$$

where

$$\mu' = \mu\cdot(-1)^{\max(k,l)}. \tag{B.12}$$

Here we use $\mu = +1$ to indicate even $(+)$ parity and $\mu = -1$ for odd $(-)$ parity. It is also worth to point out here that for $\Delta < 0$ case $\mathcal{N}_{\mu,\Delta} := \mathcal{N}_{\mu,|\Delta|}$.

As a result, if the final $\Delta = k - l$ is odd, we will assume either k or l is zero and another is odd, so according to Eq. (B.12) μ' is different from μ and recover operation should be able to restore the parity of the states; and if Δ is even, we assume μ' does not change from μ. It is easy to see that $\hat{a}\hat{b}$ will result in an uncorrectable error under this strategy, because in this case the final $\Delta = 0$ is even and we will assume $\mu' = \mu$. However, according to Eq. (B.12) μ' is different from μ since $k = l = 1$, which means our assumption is wrong and it will cause an error even after the recovery process. Besides, the amplitude γ does not change after evolution due to the strong stabilization we use.

In summary, the recovery channel $\mathcal{R}^D$ can be chosen as a set of Δ-dependent unitary operations $\hat{R}_\Delta = \sum_{\mu'=\pm}|\mu''_{\gamma,0}\rangle\langle\mu'_{\gamma,\Delta}|$ that map $|\mu'_{\gamma,\Delta}\rangle$ to $|\mu''_{\gamma,0}\rangle$, where $\mu'' = \mu'\cdot(-1)^{|\Delta|}$. So, we can write

$$\mathcal{R}^D\rho = \sum_\Delta \hat{R}_\Delta\rho\hat{R}_\Delta^\dagger. \tag{B.13}$$

Finally we can investigate the effect of whole process $\mathcal{E}^D_{\text{tot}} = \mathcal{R}^D\circ\mathcal{E}^D_{\text{evo}}$ acting on ρ that lies in the code space spanned by $|\pm\rangle_c = |\pm_{\gamma,0}\rangle$. With numerical help, we can calculate the coefficients of the process tomography under Pauli basis to indicate the error probability

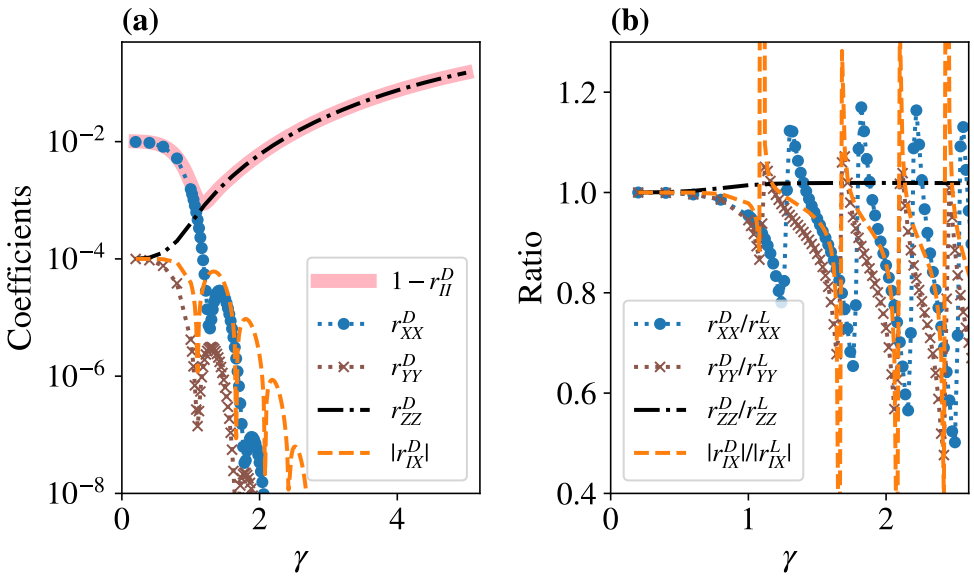

Figure B.2: Error analysis of a pair-cat memory qubit when varying its size γ. (a) The coefficients of the process tomography of $\mathcal{E}^D_{\text{tot}}$ with $1-e^{-\kappa_1 t}=2\%$. (b) Comparison between $|r^D_{jk}|$ and $|r^L_{jk}|$ with $\epsilon=1-e^{-\kappa_1 t}=2\%$.

after the recovery channel. Specifically, we have

$$\mathcal{E}^D_{\text{tot}}\rho=\sum_{j,k\in\{I,X,Y,Z\}} r^D_{jk}\hat{W}_j\rho\hat{W}^\dagger_k, \tag{B.14}$$

where $\hat{W}_j\in\{\hat{I}_c,\hat{X}_c,\hat{Y}_c,\hat{Z}_c\}$. In Fig. B.2(a), from the diagonal term r^D_{jj} we can clearly see the bias structure of the noise and find the local optimal γ value to suppress the bit-flip error, or the total error itself. For the off-diagonal term of r^D_{jk}, numerically we can see that only $r^D_{IX(XI)}$ and $r^D_{YZ(ZY)}$ are non-zero, and we can also find $|r^D_{IX}|=|r^D_{YZ}|$. In the small γ regime, we have $|r^D_{IX}|\ll r^D_{XX}$, while when γ is large we have $|r^D_{IX}|\ll r^D_{ZZ}$.

Here we provide a simple explanation for the origin of those error terms. $\hat{X}_c$ term comes from the slight difference of the photon loss error probability between $|+\rangle_c$ and $|-\rangle_c$ as shown in Eq. (3.14), which decreases exponentially as the increase of $|\gamma|^2$. $\hat{Z}_c$ and $\hat{Y}_c$ terms come from the uncorrectable photon loss $\hat{a}\hat{b}$, and according to Eq. (3.13) Z error will increase as $|\gamma|$ becomes larger, while the $\hat{Y}_c$ term will be suppressed exponentially.

In the $\gamma \to 0$ limit, we have $|+\rangle_c = |0,0\rangle$ and $|-\rangle_c = |1,1\rangle$ under Fock basis. $\hat{P}_D$ projects into the subspace spanned by those Fock states where either $\hat{a}$ or $\hat{b}$ mode is populated with at most one photon. We can keep the terms up to $O[(\kappa_1 t)^2]$ in $\mathcal{E}^D_{\text{evo}}$ and then derive $\mathcal{E}^D_{\text{tot}}$. It is easy to see that $r^D_{XX} \sim O(\kappa_1 t)$ while r^D_{ZZ}, r^D_{YY} and $|r^D_{IX(YZ)}|$ scales as $O[(\kappa_1 t)^2]$, which indicates that X error is the dominant one in the $\gamma \to 0$ limit.

B.3.2 Lossy bosonic channels

In this part, we consider another situation where there is no stabilization but only photon loss during evolution. Therefore, we can treat the evolution process as an LBC, whose effects on cat code have been discussed in Ref. [194]. We will first introduce what the LBC is, and then develop the recovery protocol and finally estimate the remaining errors. We consider that both $\hat{a}$ mode and $\hat{b}$ mode suffer independently from an LBC which can be written in the following form [71]:

$$\mathcal{E}^L_{\text{evo}}\rho = \sum_{k,l=0}^{+\infty} \hat{E}^a_k \hat{E}^b_l \rho \hat{E}^{a\dagger}_k \hat{E}^{b\dagger}_l, \tag{B.15}$$

where $\hat{E}^a_k := (k!)^{-1/2}\epsilon_a^{k/2}(1-\epsilon_a)^{\hat{n}_a/2}\hat{a}^k$ and $\hat{E}^b_l$ can be written in the same way for the $\hat{b}$ mode. For simplicity we assume $\epsilon_a = \epsilon_b = \epsilon$. In cavities with photon loss rate κ_1 and evolution time t, we have $\epsilon = 1 - e^{-\kappa_1 t}$. Then we will fix $\Delta = 0$ as our code space, and the recovery strategy is based on the final Δ we get after the lossy channel. After being applied by the Kraus operators of LBC, the code state $|\mu_{\gamma,\Delta}\rangle$ will become:

$$\hat{E}^a_k \hat{E}^b_l |\mu_{\gamma,0}\rangle = \frac{(\gamma\sqrt{\epsilon})^{k+l}}{\sqrt{k!l!}} e^{-\epsilon|\gamma|^2} \sqrt{\frac{\mathcal{N}^{\gamma'}_{\mu',k-l}}{\mathcal{N}^{\gamma}_{\mu,0}}} |\mu'_{\gamma',k-l}\rangle\,, \tag{B.16}$$

where $\mathcal{N}^{\gamma}_{\mu,\Delta}$ is value of the normalization function $\mathcal{N}_{\mu,\Delta}$ at γ. Besides, we have $\gamma' = \gamma\sqrt{1-\epsilon}$ and μ' is the same as that in Eq. (B.12).

For the recovery process after evolution, we still first measure $\hat{\Delta}$ and then apply the

Δ-dependent unitary operation $\hat{R}_\Delta = \sum_{\mu'=\pm} |\mu''_{\gamma,0}\rangle\langle\mu'_{\gamma',\Delta}|$ where $\mu'' = \mu' \cdot (-1)^{|\Delta|}$, which is similar to the strategy we mentioned in the strong dissipative stabilization case except for the fact that without stabilization we have $\gamma' \neq \gamma$. Here we can explicitly write down the matrix representation of $\mathcal{E}^L_{\text{tot}} = \mathcal{R}^L \circ \mathcal{E}^L_{\text{evo}}$ under $\rho = (\rho_{++}\ \rho_{+-}\ \rho_{-+}\ \rho_{--})^{\text{T}}$ where ρ is written under the $|\pm\rangle_c = |\pm_{\gamma,0}\rangle$ basis. Denote $\mathcal{M}_{k,l} = (\epsilon|\gamma|^2)^{k+l} e^{-2\epsilon|\gamma|^2}/(k!l!)$, we have

$$\mathcal{E}^L_{\text{tot}} = \sum_{k,l=0}^{+\infty} \mathcal{M}_{k,l} \cdot \begin{pmatrix} \delta_0(k,l)\frac{\mathcal{N}^{\gamma'}_{\mu_+,k-l}}{\mathcal{N}^{\gamma}_{+,0}} & 0 & 0 & \delta_1(k,l)\frac{\mathcal{N}^{\gamma'}_{\mu_-,k-l}}{\mathcal{N}^{\gamma}_{-,0}} \\ 0 & \delta_0(k,l)\sqrt{\frac{\mathcal{N}^{\gamma'}_{+,k-l}\mathcal{N}^{\gamma'}_{-,k-l}}{\mathcal{N}^{\gamma}_{+,0}\mathcal{N}^{\gamma}_{-,0}}} & \delta_1(k,l)\sqrt{\frac{\mathcal{N}^{\gamma'}_{+,k-l}\mathcal{N}^{\gamma'}_{-,k-l}}{\mathcal{N}^{\gamma}_{+,0}\mathcal{N}^{\gamma}_{-,0}}} & 0 \\ 0 & \delta_1(k,l)\sqrt{\frac{\mathcal{N}^{\gamma'}_{+,k-l}\mathcal{N}^{\gamma'}_{-,k-l}}{\mathcal{N}^{\gamma}_{+,0}\mathcal{N}^{\gamma}_{-,0}}} & \delta_0(k,l)\sqrt{\frac{\mathcal{N}^{\gamma'}_{+,k-l}\mathcal{N}^{\gamma'}_{-,k-l}}{\mathcal{N}^{\gamma}_{+,0}\mathcal{N}^{\gamma}_{-,0}}} & 0 \\ \delta_1(k,l)\frac{\mathcal{N}^{\gamma'}_{\mu_+,k-l}}{\mathcal{N}^{\gamma}_{+,0}} & 0 & 0 & \delta_0(k,l)\frac{\mathcal{N}^{\gamma'}_{\mu_-,k-l}}{\mathcal{N}^{\gamma}_{-,0}} \end{pmatrix}. \tag{B.17}$$

Here, $\mu_\pm := (\pm 1) \cdot (-1)^{\max(k,l)}$, and $\delta_q(k,l)$ satisfies

$$\delta_q(k,l) = \begin{cases} 1 & \text{if } [\min(k,l) - q] \text{ is even,} \\ 0 & \text{if } [\min(k,l) - q] \text{ is odd.} \end{cases} \tag{B.18}$$

We can also find the coefficients of process tomography and denote them as r^L_{jk}. With the explicit form of $\mathcal{E}^L_{\text{tot}}$ in Eq. (B.17), it is easy to check that $r^L_{IY} = r^L_{IZ} = r^L_{XY} = r^L_{XZ} = 0$ and $|r^L_{IX}| = |r^L_{YZ}|$. In Fig. B.2(b), we compare the results between r^D_{jk} and r^L_{jk} under $\epsilon = 1 - e^{-\kappa_1 t} = 2\%$. We can see that r_{ZZ} is approximately the same between the two situations. However, since the local minima in $r^D_{XX}(\gamma)$ and $r^L_{XX}(\gamma)$ correspond to slightly different γ, the r^D_{XX}/r^L_{XX} curves fluctuate near 1. This argument also works for the r_{YY}

and r_{IX} cases.

B.4 Perturbative analysis of gate errors

In this appendix, we will discuss the scaling of the Z error probability of the pair-cat code during gate operations in the dissipative stabilization scheme. This type of error can be induced by both photon loss and the leakage out of the protected code space in the middle of the gate execution. Since photon loss errors will not cause leakage out of the stabilized subspace, we will treat the two effects separately. Our analysis is based on the methods introduced in [59] where gate errors are investigated for cat code. There, the adiabatic elimination method [162] has been used in order to achieve effective dynamical equations in the stabilized subspace. Due to the similarities between cat code and pair-cat code, we can also use similar strategies to derive the error probability of the gates we construct for the pair-cat code. Therefore, in the following discussion we will only mention those key ingredients in the derivation to achieve Z error probability and properties specialized for pair-cat code, while detailed reasoning for each step of the derivation can be found in [59] that focuses on the cat code counterpart. We highlight that for the pair-cat code, all the gate error probabilities can be achieved to scale at least linearly in single-photon loss rate κ_1, which works better compared with the cat code where error probability mainly scales as $\sqrt{\kappa_1}$ [59, 68]. To simplify the notation, we will assume γ to be a positive real number.

B.4.1 $Z(\theta)$ gate

The $Z(\theta)$ gate can be implemented via the 2-mode squeezing Hamiltonian $\hat{H}_Z$ in Eq. (3.20). The $\hat{a}^\dagger\hat{b}^\dagger$ term will cause leakage out of the code space via

$$\hat{a}^\dagger\hat{b}^\dagger\left|\mu_{\gamma,\Delta}\right\rangle \simeq \gamma^2\left|\mu'_{\gamma,\Delta}\right\rangle + \sqrt{2}\gamma\left|\psi_{e1,\mu',\Delta}\right\rangle + O(1), \tag{B.19}$$

where $\mu' \neq \mu$. It means the parity of the two modes is changed. Further, this leakage can be recovered back to the code space via the stabilization dissipator $\kappa\mathcal{D}[\hat{a}^2\hat{b}^2 - \gamma^4]$, since

$$(\hat{a}^2\hat{b}^2 - \gamma^4)\,|\psi_{e1,\mu',\Delta}\rangle \simeq 2\sqrt{2}\gamma^3\,|\mu'_{\gamma,\Delta}\rangle\,. \tag{B.20}$$

This whole process together will give us a Z error on the pair-cat code. The effective error rate can be estimated via adiabatic elimination, which will give us $4(\epsilon_Z\sqrt{2}\gamma)^2/(8\kappa\gamma^6) = \epsilon_Z^2/(\kappa\gamma^4)$.

Z error probability induced by photon loss is simple to analyze. Only a single-photon loss in one mode will not cause errors in the pair-cat code since it can be detected by $\hat{\Delta}$ measurement at the end of the gate operation and corrected via a recovery channel. Z error will be induced if both $\hat{a}$ and $\hat{b}$ errors happens. Therefore, the combined error probability is $(\kappa_1\gamma^2 T)^2$, and the total dephasing error probability during the $Z(\theta)$ gate is

$$p_Z = \frac{\epsilon_Z^2}{\kappa\gamma^4}T + (\kappa_1\gamma^2 T)^2. \tag{B.21}$$

Recall that $\epsilon_Z = \theta/(4\gamma^2 T)$, we can find the optimal time T to minimize p_Z scales as $T^{\text{opt}} = (\theta^2/32\kappa\kappa_1^2\gamma^{12})^{1/3}$, and the corresponding $p_Z^{\text{opt}} \sim O\big[(\kappa_1/\kappa)^{2/3}/\gamma^4\big]$.

On the other hand, this scaling can be changed by using real-time monitoring of $\hat{\Delta}$ as we introduced in the design of CX gate. Suppose we keep measuring $\hat{\Delta}$ at a time interval of $\delta\tau$, then only the case that both $\hat{a}$ and $\hat{b}$ errors happen within the $\delta\tau$ interval will cause Z error, otherwise we will know exactly that the code does suffer from $\hat{a}\hat{b}$ loss instead of nothing happens. In this case, the Z error probability induced by loss is $(\kappa_1\gamma^2\delta\tau)^2(T/\delta\tau) = \kappa_1^2\gamma^4\delta\tau T$, and the optimal $p_Z^{\text{opt}} \sim O(\kappa_1\sqrt{\delta\tau/\kappa}/\gamma^2)$, which is linear in κ_1.

B.4.2 $ZZ(\theta)$ gate

Like the $Z(\theta)$ gate, the Hamiltonian $\hat{H}_{ZZ}$ in Eq. (3.20) that we use to construct $ZZ(\theta)$ gate will also cause leakage out of the code space, which will induce Z type of errors. For example, the $\hat{a}_1\hat{b}_1\hat{a}_2^\dagger\hat{b}_2^\dagger$ term couples the code state of the second qubit with states out of the code space while changing the parities of both pair-cat qubits at a rate of $\sqrt{2}\epsilon_{ZZ}\gamma^3$, which is the dominant leakage rate from this term. Besides, $\hat{a}_1^\dagger\hat{b}_1^\dagger\hat{a}_2\hat{b}_2$ term has a similar effect but causes leakage in the first qubit. These two channels, after going back to the code space due to the dissipative stabilization, will cause a Z_1Z_2 error on the pair-cat qubits at a rate of $2\times 4(\sqrt{2}\epsilon_{ZZ}\gamma^3)^2/(8\kappa\gamma^6) = 2\epsilon_{ZZ}^2/\kappa$. So,

$$p_{Z_1Z_2} = 2\epsilon_{ZZ}^2 T/\kappa = \frac{\theta^2}{8\kappa\gamma^8 T}. \tag{B.22}$$

The photon-loss induced dephasing error can be analyzed in the same way as the former case in the $Z(\theta)$ case. If there is no real-time $\hat{\Delta}$ monitoring, then $p_{Z_1} = p_{Z_2} = (\kappa_1\gamma^2 T)^2$, and the optimal total error probability $p = p_{Z_1} + p_{Z_2} + p_{Z_1Z_2}$ scales as $O\big[(\kappa_1/\kappa)^{2/3}/\gamma^4\big]$. If we have this real-time $\hat{\Delta}$ monitoring with time interval $\delta\tau$, then $p_{Z_1} = p_{Z_2} = \kappa_1^2\gamma^4\delta\tau T$ and optimal p scales as $O(\kappa_1\sqrt{\delta\tau/\kappa}/\gamma^2)$.

B.4.3 X gate

The X gate is implemented by changing stabilization parameter γ with respect to time that $\gamma(t) = \gamma e^{i\frac{\pi}{2}\frac{t}{T}}$ and use an extra Hamiltonian $\hat{H}_{X,\text{rot}} = -\frac{\pi}{2T}(\hat{n}_a + \hat{n}_b)$ to compensate the error induced by non-adiabaticity. If there is no loss happening during this gate execution, the X gate can be implemented perfectly, and there is no term to cause leakage out of the code space. So, the only source that can induce dephasing error is from photon loss.

Compared with the idling case where γ stays as a constant, a single-photon loss at time t_0 in either mode will induce an extra global phase $e^{i\frac{\pi}{2}\frac{t_0}{T}}$. Unlike in the CX gate case we

mentioned in the main text where this induced phase on the target qubit does cause a Z type rotation of the control qubit, here this is just a global phase on a single qubit that does not matter at all. As a result, Z error probability induced by photon loss during X gate has no difference compared with other gates: if there is no real-time $\hat{\Delta}$ monitoring, $p_Z = (\kappa_1\gamma^2 T)^2$; if we have such monitoring, $p_Z = \kappa_1^2\gamma^4\delta\tau T$. So, to reduce p_Z, we should choose T to be as small as possible.

B.4.4 CX gate

The CX gate can be implemented by changing stabilization parameter $\gamma(t)$ of the target qubit conditioned on the states of the control qubit via jump operators $\hat{F}_{1,2}$ defined in Eq. (3.26). We also apply another Hamiltonian $\hat{H}_{\text{CX,rot}}$ in Eq. (3.27) to reduce the error induced by non-adiabaticity. However, this extra Hamiltonian cannot fully compensate it, because other than the desired conditional rotation $\hat{H}' = -\frac{\pi}{2T}|1\rangle_{1c}\langle 1|\otimes(\hat{a}_2^\dagger\hat{a}_2+\hat{b}_2^\dagger\hat{b}_2-2|\gamma|^2)$, the $\hat{a}_1^\dagger\hat{b}_1^\dagger(\hat{a}_2^\dagger\hat{a}_2+\hat{b}_2^\dagger\hat{b}_2-2|\gamma|^2)$ term can also cause excitation of both control and target qubits. Other than Eq. (B.19), we also have

$$(\hat{a}^\dagger\hat{a}+\hat{b}^\dagger\hat{b}-2|\gamma|^2)\,|\mu_{\gamma,\Delta}\rangle \simeq \sqrt{2}\gamma\,|\psi_{e1,\mu,\Delta}\rangle\,, \tag{B.23}$$

where the parity of the target qubit does not change in this process. But, as discussed before, the parity of the control qubit will change, which effectively causes a Z operation on it.

We follow the method in Ref. [59] by going to the rotating frame according to $\hat{H}'$. In this frame, $\hat{a}_2\hat{b}_2$ will be transformed to

$$\hat{a}_2\hat{b}_2 \to |0\rangle_{1c}\langle 0|\otimes\hat{a}_2\hat{b}_2 + |1\rangle_{1c}\langle 1|\otimes\hat{a}_2\hat{b}_2 e^{i\pi t/T}. \tag{B.24}$$

This will result that, if the both control and target qubit state get excited together as we mentioned and then decay back to the code space via $\hat{F}_1$ and $\hat{F}_2$, effectively there will be

a $Z_1 Z_1(-\frac{2\pi t}{T}) = Z_1(\pi\frac{T-2t}{T})$ error acting on the control qubit. It is because in the rotating frame, going back to the code space of the target qubit via $\hat{F}_2$ will cause a $Z_1(-\frac{2\pi t}{T})$ operation on the control qubit. The corresponding effective error rate is $\frac{\pi^2}{64\kappa\gamma^6 T^2}$. This $Z_1(\theta)$ error does not change when going back to the original frame, and after averaging with total time, the Z_1 error probability induced from the non-adiabaticity is $\frac{\pi^2}{128\kappa\gamma^6 T}$. It is also worth to mention that in the derivation of the error probability we have ignored the situation that the $\hat{a}_1\hat{b}_1$ term in $\hat{F}_2$ can also help the control qubit to decay back to the code space, but this term will not cause the change of the scaling property of the result we achieved, as discussed in [59].

Then let us discuss the error induced by photon loss. We assume to have real-time $\hat{\Delta}$ monitoring on both control and target qubits. The photon loss on control qubit does not affect the phase of the target qubit state, so it will just give a Z_1 error with probability $\kappa_1^2\gamma^4\delta\tau T$. Things will become different when loss errors happen on the target qubit. As discussed in the main text, the photon loss on the target qubit will induce a time-dependent phase shift on the control qubit; therefore we want to use real-time $\hat{\Delta}$ measurement to monitor when the loss happens and correct it with an extra $Z_1(\theta)$ gate.

In practice, however, there are two relevant processes to induce Z type of errors because of photon loss on the target qubit. One comes from that, even only a single-photon loss happens, due to the finite $\delta\tau$ time of two Δ measurements, we can only correct the extra phase for $|1\rangle_{1c}$ state of the control qubit up to small deviation ranging within $[-\frac{\delta\theta}{2}, \frac{\delta\theta}{2}]$, where $\delta\theta = \frac{\pi}{2}\frac{\delta\tau}{T}$. This effect on average will give a Z_1 error probability as $\frac{\kappa_1\gamma^2(\delta\tau)^2\pi^2}{96T}$. The other comes from that both $\hat{a}_2$ and $\hat{b}_2$ happen to the target qubit within $\delta\tau$ time, which not only causes a Z_2 error on target qubit but also induces an extra $e^{i\pi t/T}$ phase on $|1\rangle_{1c}$ state of the control qubit. Using the same method we did in Eq. (3.28), we found that this error has the form of $Z_1 Z_1(-\frac{\pi t}{T})Z_2 = Z_1(\pi\frac{T-t}{T})Z_2$ with error rate $\kappa_1^2\gamma^4\delta\tau$, which on average gives both Z_2 error and $Z_1 Z_2$ error a probability of $\kappa_1^2\gamma^4\delta\tau T/2$.

In summary, we have

$$\begin{aligned} p_{Z_1} &= \kappa_1^2\gamma^4\delta\tau T + \frac{\pi^2}{128\kappa\gamma^6 T} + \frac{\kappa_1\gamma^2(\delta\tau)^2\pi^2}{96T}, \\ p_{Z_2} &= p_{Z_1Z_2} = \frac{\kappa_1^2\gamma^4\delta\tau T}{2}. \end{aligned} \tag{B.25}$$

By choosing the optimal time T to minimize the total error probability $p = p_{Z_1} + p_{Z_2} + p_{Z_1Z_2}$, we find that $p \sim O\Big(\sqrt{\frac{\kappa_1}{\kappa}}\sqrt{\frac{\kappa_1\delta\tau}{\gamma^2}}\sqrt{1 + C\kappa\kappa_1\gamma^8(\delta\tau)^2}\Big)$, where $C = \frac{4}{3}$ is a constant number. Therefore, the total error probability scales between $O(\kappa_1)$ and $O(\kappa_1^{3/2})$.

B.4.5 Toffoli gate

The error properties in the Toffoli gate are similar to those in the CX gate, and we can use the same method to analyze them. For convenience we denote $CZ(\theta) := \exp(-i\theta|11\rangle\langle 11|)$. Recall the Hamiltonian defined in Eq. (3.32) which is used to compensate the non-adiabatic error; it could cause joint excitation of 1, 3 states or 2, 3 states out of their code spaces, together with a parity change in either qubit 1 or qubit 2 that gets excited. Again, by going to the rotating frame according to $\hat{H}' = -\frac{\pi}{2T}|1\rangle_{1c}\langle 1| \otimes |1\rangle_{2c}\langle 1| \otimes (\hat{a}_3^\dagger\hat{a}_3 + \hat{b}_3^\dagger\hat{b}_3 - 2|\gamma|^2)$, we realize that the two control qubit states will gain a phase if they are in $|1\rangle_{1c}|1\rangle_{2c}$ when the target qubit state decays back to the code space due to the dissipator $\hat{F}_3$ defined in Eq. (3.31b). This is similar to the effect caused by Eq. (B.24) in the CX gate, but here we need to focus on the transformation of $\hat{a}_3\hat{b}_3$ into the rotating frame instead. By using adiabatic elimination method, the effective error will be either $\frac{I-Z_1}{2}Z_2CZ_{1,2}(-\frac{2\pi t}{T})$ or $\frac{I-Z_2}{2}Z_1CZ_{1,2}(-\frac{2\pi t}{T})$ with the same error rate $\pi^2/(64\kappa\gamma^6T^2)$. Going back to the original frame will not cause a change in the error forms. After averaging over time, we can see the non-adiabaticity will give an error probability of $\frac{\pi^2}{256\kappa\gamma^6 T}$ for all Z_1, Z_2, and Z_1Z_2 types of errors. Similar to the derivation in the CX case, we also ignore the contribution that the excitations of two control qubits can decay back via $\hat{F}_3$ since this effect does not change the scaling of the error probability

we derived.

Then we discuss about Z type of error induced by photon loss. We again assume to have real-time $\hat{\Delta}$ monitoring of all three pair-cat qubits. The loss happens in either of the two control qubits will not cause errors in qubits that do not suffer from photon loss, so only both $\hat{a}_i$ and $\hat{b}_i$ $(i = 1, 2)$ happen in the same time interval of $\hat{\Delta}$ measurement will cause a Z_i error. The corresponding error probability is again $\kappa_1^2\gamma^4\delta\tau T$.

Similar to the CX gate, the loss errors in the target qubit will also cause two effects. First, a single-photon loss of either $\hat{a}_3$ or $\hat{b}_3$ will induce a phase on $|1\rangle_{1c}|1\rangle_{2c}$ state, and due to the finite time duration between each $\hat{\Delta}$ measurement, this phase can only be corrected up to a small deviation ranging between $[-\frac{\delta\theta}{2}, \frac{\delta\theta}{2}]$ where $\delta\theta = \frac{\pi}{2}\frac{\delta\tau}{T}$. This effect will on average give an error probability of $\frac{\kappa_1\gamma^2(\delta\tau)^2\pi^2}{384T}$ for Z_1, Z_2, and Z_1Z_2 type of errors. The second effect comes from both $\hat{a}_3$ and $\hat{b}_3$ happening in the same $\hat{\Delta}$ measurement interval. It can induce an effective $CZ_{1,2}(\pi\frac{T-t}{T})Z_3$ error with error rate $\kappa_1^2\gamma^4\delta\tau$. By averaging over time, this will give an error probability of $\frac{5}{8}\kappa_1^2\gamma^4\delta\tau T$ for Z_3 error and $\frac{1}{8}\kappa_1^2\gamma^4\delta\tau T$ for Z_1Z_3, Z_2Z_3, and $Z_1Z_2Z_3$ error.

In summary, the error probability for all types of errors can be listed as

$$\begin{aligned}
p_{Z_1} &= p_{Z_2} = \kappa_1^2\gamma^4\delta\tau T + \frac{\pi^2}{256\kappa\gamma^6 T} + \frac{\kappa_1\gamma^2(\delta\tau)^2\pi^2}{384T}, \\
p_{Z_1Z_2} &= \frac{\pi^2}{256\kappa\gamma^6 T} + \frac{\kappa_1\gamma^2(\delta\tau)^2\pi^2}{384T}, \\
p_{Z_3} &= \frac{5}{8}\kappa_1^2\gamma^4\delta\tau T, \\
p_{Z_1Z_3} &= p_{Z_2Z_3} = p_{Z_1Z_2Z_3} = \frac{1}{8}\kappa_1^2\gamma^4\delta\tau T.
\end{aligned} \tag{B.26}$$

To minimize the total error probability p by using optimal choice of T, we can see that p again scales as $p \sim O\Big(\sqrt{\frac{\kappa_1}{\kappa}}\sqrt{\frac{\kappa_1\delta\tau}{\gamma^2}}\sqrt{1 + C'\kappa\kappa_1\gamma^8(\delta\tau)^2}\Big)$ where $C' = \frac{2}{3}$ is another constant number. We can see that p here shares a similar scaling property to that in the case of the CX gate.

APPENDIX C

APPENDICES FOR CHAPTER 4

C.1 Dark states with generic stabilization drives

In this section, I will discuss the structures of dark subspace under generic stabilization drives $\Omega_{\pm1,0}$. Since in general the stabilization Hamiltonian $\hat{H}_{\mathrm{DS}}$ in Eq. (4.3) can be decomposed as $\hat{H}_{\mathrm{DS}} = \hat{H}_{ee} + \hat{H}_{eg} + \hat{H}_{ge}$, and the jump operators $\hat{D}_{ge,q}$ only maps the states in excited manifold to the ground manifold, then it will be sufficient for a state $|\mathrm{DS}\rangle_g$ in the ground manifold to be a dark state if

$$\hat{H}_{eg} |\mathrm{DS}\rangle_g = 0. \tag{C.1}$$

From Fig. C.1, it will be easy to see that, when $\Omega_{+1} = 0$, then $|F_g, -F_g\rangle$ will be a dark state. Similarly, $|F_g, F_g\rangle$ will be a dark state when $\Omega_{-1} = 0$. In the following, I will consider different frame transformations through the rotations, so that in the new frame we can find either of $\Omega_{\pm1} = 0$. With this, we can locate one dark state in the new frame, which will still be an SCS after going back to the original frame.

Then let us look into how the rotation transformation will affect the coupling Hamiltonian $\hat{H}_{eg}$. First, to simplify the notation, we can write $\hat{H}_{eg} = \frac{\vec{\Omega}}{2} \cdot \vec{\hat{D}}_{eg}$, where $\vec{\Omega} = (\Omega_{+1}\ \Omega_0\ \Omega_{-1})^{\mathrm{T}}$ and $\vec{\hat{D}}_{eg} = (\hat{D}_{eg,+1}\ \hat{D}_{eg,0}\ \hat{D}_{eg,-1})^{\mathrm{T}}$, We also introduce the rotation operator on the whole system

$$\hat{R}_\xi(\theta) = e^{-i(\hat{F}_{g,\xi} + \hat{F}_{e,\xi})\theta}, \qquad \text{where} \quad \xi = x, y, z. \tag{C.2}$$

When acting on the Hamiltonian $\hat{H}_{eg}$, the transformation of the operators can be equivalently expressed as the transformation of the driving vector Ω. Specifically,

$$\hat{R}_\xi(\theta)\hat{H}_{eg}\hat{R}_\xi^\dagger(\theta) = \frac{1}{2}\vec{\Omega} \cdot [\hat{R}_\xi(\theta)\vec{\hat{D}}_{eg}\hat{R}_\xi^\dagger(\theta)] = \frac{1}{2}\vec{\Omega}_{\xi,\theta} \cdot \vec{\hat{D}}_{eg}. \tag{C.3}$$

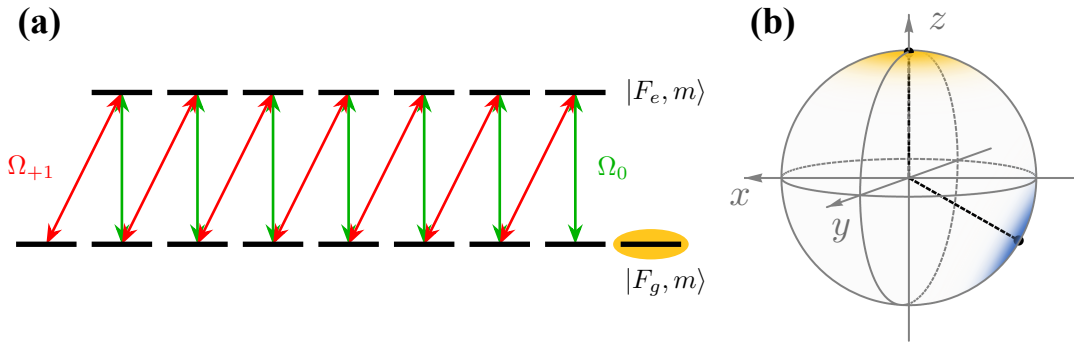

Figure C.1: The dark subspace is always spanned by two SCSs that may not be antipodal on the GBS. (a) Dark state structures without the Ω_{-1} drive. Then $|F_g, F_g\rangle$ will be a dark state, which corresponds to the north pole of the GBS in (b). As proved here, we can always find two different rotation transformations to make $\Omega_{-1} = 0$ in the new frame.

Here, we explicitly write down the outcomes from y- or z-rotation as

$$\vec{\Omega}_{y,\theta} = \begin{pmatrix} \frac{1+\cos(\theta)}{2} & -\frac{\sin(\theta)}{\sqrt{2}} & \frac{1-\cos(\theta)}{2} \\ \frac{\sin(\theta)}{\sqrt{2}} & \cos(\theta) & -\frac{\sin(\theta)}{\sqrt{2}} \\ \frac{1-\cos(\theta)}{2} & \frac{\sin(\theta)}{\sqrt{2}} & \frac{1+\cos(\theta)}{2} \end{pmatrix} \begin{pmatrix} \Omega_{+1} \\ \Omega_0 \\ \Omega_{-1} \end{pmatrix}, \quad \vec{\Omega}_{z,\phi} = \begin{pmatrix} e^{-i\phi} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & e^{i\phi} \end{pmatrix} \begin{pmatrix} \Omega_{+1} \\ \Omega_0 \\ \Omega_{-1} \end{pmatrix}. \tag{C.4}$$

Now we consider a frame transformation that makes $\Omega_{-1} = 0$. To do so, we split the process into two steps. First, we try to find a frame where $\Omega_0' = 0$. As indicated from Eq. (C.1), a global complex phase in $\hat{H}_{eg}$ does not affect the solution for $|\mathrm{DS}\rangle_g$. Therefore, we assume Ω_0 is a real number, and denote the complex number $\Omega_{\pm 1}$ as $\Omega_{\pm 1} = \Omega_{\pm 1}^R + i\Omega_{\pm 1}^I$. To make $\Omega_0 = 0$, we first rotate the system with $\hat{R}_z(\phi_a)$ and then $\hat{R}_y(\theta_a)$. We need to find the (θ_a, ϕ_a) pairs to make

$$\frac{\sin(\theta_a)}{\sqrt{2}} e^{-i\phi_a}\Omega_{+1} + \cos(\theta_a)\Omega_0 - \frac{\sin(\theta_a)}{\sqrt{2}} e^{i\phi_a}\Omega_{-1} = 0. \tag{C.5}$$

The left-hand-side expression in the equation above is generally a complex number, so we

can split it as

$$\begin{cases} \dfrac{\sin(\theta_a)}{\sqrt{2}}\mathrm{Re}[e^{-i\phi_a}\Omega_{+1} - e^{i\phi_a}\Omega_{-1}] + \cos(\theta_a)\Omega_0 = 0, & \text{(C.6a)} \\ [-\Omega^R_{+1}\sin(\phi_a) + \Omega^I_{+1}\cos(\phi_a)] - [\Omega^R_{-1}\sin(\phi_a) + \Omega^I_{-1}\cos(\phi_a)] = 0. & \text{(C.6b)} \end{cases}$$

With this, we can first determine ϕ_a through Eq. (C.6b), which gives

$$\phi_a = \arctan\frac{\Omega^I_{+1} - \Omega^I_{-1}}{\Omega^R_{+1} + \Omega^R_{-1}}. \tag{C.7}$$

Then, θ_a can also be determined through Eq. (C.6a) accordingly:

$$\theta_a = \arctan\frac{-\sqrt{2}\Omega_0}{\mathrm{Re}[e^{-i\phi_a}\Omega_{+1} - e^{i\phi_a}\Omega_{-1}]}. \tag{C.8}$$

In the new frame with driving vector $\vec{\Omega}'$, only the components $\Omega'_{\pm1}$ could be nonzero. Let us perform the second transformation by first applying $\hat{R}_z(-\phi_b)$ and then $\hat{R}_y(-\theta_b)$. We expect $\Omega''_{-1} = 0$ in this frame. Then, the (θ_b, ϕ_b) pairs should satisfy

$$\frac{1-\cos(\theta_b)}{2}e^{-i\phi_b}\Omega'_{+1} + \frac{1+\cos(\theta_b)}{2}e^{i\phi_b}\Omega'_{-1} = 0. \tag{C.9}$$

Here $\Omega'_{\pm1}$ are still complex numbers, and we can denote them as $\Omega'_{\pm1} = |\Omega'_{\pm1}|e^{i\beta'_{\pm1}}$. We first choose ϕ_b to satisfy

$$e^{i(\beta'_{+1}-\phi_b)} = -e^{i(\beta'_{-1}+\phi_b)}. \tag{C.10}$$

It will be sufficient if

$$\phi_b = \frac{\beta'_{+1} - \beta'_{-1} + \pi}{2}. \tag{C.11}$$

Finally, θ_b needs to be determined from

$$\frac{1-\cos(\theta_b)}{2}|\Omega'_{+1}| - \frac{1+\cos(\theta_b)}{2}|\Omega'_{-1}| = 0, \tag{C.12}$$

which leaves two different solutions for us:

$$\theta_b = \pm \arccos \frac{|\Omega'_{+1}| - |\Omega'_{-1}|}{|\Omega'_{+1}| + |\Omega'_{-1}|}. \tag{C.13}$$

As a result, we can find two different rotations such that $\Omega''_{-1} = 0$ in the new frame. It means that there are two different SCSs as the dark states in the original frame.

One extreme case comes when both Ω'_0 and Ω'_{-1} (or similarly Ω'_{+1} here) become zero after the first set of transformations under (θ_a, ϕ_a). The dark subspace now are spanned by $|F_g, F_g\rangle$ and $|F_g, F_g - 1\rangle$. In fact, since

$$|F_g, F_g - 1\rangle = \lim_{\theta \to 0} \frac{|\theta, 0\rangle_g - e^{i\pi F_g} |\theta, \pi\rangle_g}{\theta \sqrt{2F_g}}, \tag{C.14}$$

we can still interpret the dark subspace as it is spanned by two SCSs $|\theta, 0\rangle_g$ and $|\theta, \pi\rangle_g$, in the limit that $\theta \to 0$.

C.2 Universal gate set on the physical qubit level

C.2.1 Single-qubit holonomic control

In this part, I will discuss how to implement the universal single-qubit gate set through holonomic control. The gates in the encoded subspace are achieved through the adiabatic variation of the stabilization drives, which also modifies the place of the two dark SCSs on the GBS accordingly. The idea here follows a similar design in the bosonic cat counterpart, which was first proposed in Ref. [149]. We also omit the subscript g here since in the adiabatic limit we only need to care about the properties of the dark subspace, which lies in the F_g manifold only.

$Z(\alpha)$ gate

To implement the Z rotation gate, we can adiabatically move the two SCSs while keeping them antipodal. The dark state subspace is now spanned by $|\theta_0(t), \phi_0(t)\rangle$ and $|\theta_1(t), \phi_1(t)\rangle$, with $\theta_1 = \pi - \theta_0$ and $\phi_1 = \pi + \phi_0$. Assume we work in the adiabatic limit, then the quantum state $\hat{\rho}(t)$ will always be in the instantaneous dark subspace with the following generic form:

$$\begin{aligned}\hat{\rho}(t) = a_{00}(t)|\theta_0(t), \phi_0(t)\rangle\langle\theta_0(t), \phi_0(t)| + a_{01}(t)|\theta_0(t), \phi_0(t)\rangle\langle\theta_1(t), \phi_1(t)| \\ + a_{10}(t)|\theta_1(t), \phi_1(t)\rangle\langle\theta_0(t), \phi_0(t)| + a_{11}(t)|\theta_1(t), \phi_1(t)\rangle\langle\theta_1(t), \phi_1(t)|.\end{aligned} \tag{C.15}$$

On the other hand, the master equation restricted to the dark subspace will give:

$$\hat{P}(t)\frac{\mathrm{d}\hat{\rho}(t)}{\mathrm{d}t}\hat{P}(t) = \hat{P}(t)\mathcal{L}(t)[\hat{\rho}(t)]\hat{P}(t) = 0, \tag{C.16}$$

where $\hat{P}(t) = |\theta_0(t), \phi_0(t)\rangle\langle\theta_0(t), \phi_0(t)| + |\theta_1(t), \phi_1(t)\rangle\langle\theta_1(t), \phi_1(t)|$ is the projector to the stabilized subspace. Notice that

$$\frac{\mathrm{d}\,|\theta(t), \phi(t)\rangle}{\mathrm{d}t} = \{-i\dot{\phi}\hat{F}_z - i\dot{\theta}[\cos(\phi)\hat{F}_y - \sin(\phi)\hat{F}_x]\}\,|\theta(t), \phi(t)\rangle\,, \tag{C.17}$$

we can get the following expression from Eq. (C.16):

$$\begin{cases}\dot{a}_{00} = \dot{a}_{11} = 0, \\ \dot{a}_{01}/a_{01} = -2iF\dot{\phi}_0\cos(\theta_0), \\ a_{10} = a_{01}^*.\end{cases} \tag{C.18}$$

Therefore, the population a_{00} and a_{11} for each SCS is unchanged during evolution. Also, when the path of each SCS forms a loop L on the GBS, then we have

$$a_{01,\text{final}} = a_{01}(0)\exp\left(-i\oint_L \mathrm{d}\alpha\right), \tag{C.19}$$

where

$$\oint_L \mathrm{d}\alpha = \oint_L 2F\cos(\theta_0)\,\mathrm{d}\phi_0 = \iint -2F\frac{\partial\cos(\theta_0)}{\partial\theta_0}\,\mathrm{d}\theta_0\,\mathrm{d}\phi_0 = \iint 2F\,\mathrm{d}\Omega_S = 2F\Omega_S. \quad \text{(C.20)}$$

With this, we successfully demonstrate the implementation of the $Z(\alpha)$ gate with the phase $\alpha = 2F\Omega_S$ proportional to the solid angle Ω_S enclosed by the trajectory loop on the GBS.

$X(\alpha)$ gate

Now we briefly present how to construct a single-qubit $X(\alpha)$ gate by adiabatically changing the driving parameters $\Omega_{\pm1}(t)$. We start with $\Omega_{+1} = -\Omega_{-1} = \Omega$, such that the code subspace is spanned by $\{|F_g, F_g\rangle_x, |F_g, -F_g\rangle_x\}$, i.e. two SCSs pointing along the x-axis on the GBS. We first adiabatically reduce the amplitude Ω_{+1} to 0 in order to collide the two SCSs at the south pole of the generalized Bloch sphere. Then, we separate them along another longitudinal line by adiabatically increasing Ω_{+1} to $\Omega e^{2i\alpha}$, and finally changing Ω_{+1} to Ω without affecting its amplitude. The location change for the two SCSs during the whole process is illustrated in Fig C.2(a).

Note that, in the absence of the Ω_0 drive, one of the dark states lies in the subspace spanned by $\{|F_g, -F_g + 2k\rangle_z\}$ (even parity), while another one is in the subspace spanned by $\{|F_g, -F_g + (2k+1)\rangle_z\}$ (odd parity). Therefore, during the first step where we adiabatically turn off Ω_{+1}, the state $|+\rangle_c$ connects to $|F_g, -F_g\rangle_z$, while $|-\rangle_c$ is connected to $|F_g, -F_g + 1\rangle_z$. Following the convention in Ref. [149], we denote this mapping as $\hat{S}_0^\dagger$. Due to the rotational symmetry in the system, the operation for the second step is $\hat{S}_\alpha$, which is related to $\hat{S}_0$ by a rotation operator $\hat{S}_\alpha = \hat{R}_z(\alpha)\hat{S}_0\hat{R}_z^\dagger(\alpha)$. The third step is simply a rotation $\hat{R}_z^\dagger(\alpha)$. The whole process leads to an operation $\hat{R}_z^\dagger(\alpha)\hat{S}_\alpha\hat{S}_0^\dagger = \hat{S}_0\hat{R}_z^\dagger(\alpha)\hat{S}_0^\dagger$ on the dark-state subspace. Therefore, $|+\rangle_c$ will pick up a $e^{-i\alpha F_g}$ phase while $|-\rangle_c$ will get a phase of $e^{-i\alpha(F_g-1)}$. As a result, an arbitrary superposition of the two code states will become the following state

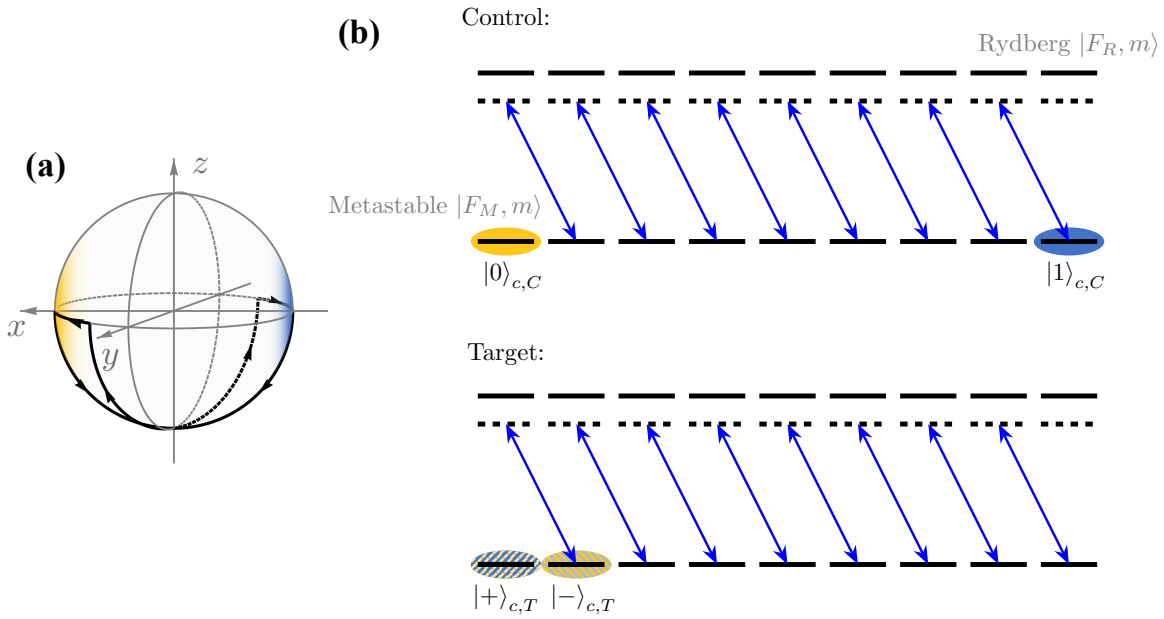

Figure C.2: Native non-bias-preserving operations for the spin-cat code. (a) The adiabatic paths for the two SCSs on the GBS to implement the $X(\alpha)$ gate. (b) The code states and the entangling pulses in the middle of the native implementation for the non-BP CX gate.

after the whole process

$$c_+ |+\rangle_c + c_- |-\rangle_c \longrightarrow e^{-i\alpha F_g} \left(c_+ |+\rangle_c + c_- e^{i\alpha} |-\rangle_c \right), \tag{C.21}$$

which is exactly the $X(\alpha)$ gate. This operation, together with the $Z(\alpha)$ gate, provides arbitrary 1-qubit rotation on the dark spin-cat encoding.

The $X(\alpha)$ gate in general does not preserve the Pauli error bias. It will convert part of the Z error before the gate to the Y error when α is not an integer multiple of π. This can also be viewed from the trajectories of the two SCSs during the gate execution, as their collision will lead to the population exchange between computational basis states and convert phase-flip error to bit-flip. As a result, the CX implementation involving the use of such $X(\alpha)$ gate does not preserve the Pauli error bias either.

C.2.2 A native implementation for non-bias-preserving CX gate

With arbitrary 1-qubit rotations and the CZ entangling gate, we can achieve universal operations on the physical encoding level, including the CX operations as $\mathrm{CX} = Y_T(\frac{\pi}{2}) \cdot \mathrm{CZ} \cdot Y_T(-\frac{\pi}{2})$. However, we can further simplify the CX control sequence in a more native manner using the collision-and-expansion trick we introduced in the $X(\alpha)$ gate construction. Here we also assume to use the metastable encoding. As a result, later we will change the notation from F_g to F_M for the encoded hyperfine manifold.

We briefly explain the idea as follows. The first step is again converting the encoding from Fock states in x basis to Fock in z basis, but for the control and target atom the protocol is different. For the control atom, the conversion protocol can be the same as that used for the control atom in bias-preserving CX or CZ construction shown in the main text, i.e., adiabatically transfer $|0(1)\rangle_c \propto |F_M, \pm F_M\rangle_x$ to $|F_M, \mp F_M\rangle_z$ while always keeping them antipodal during the evolution. On the other hand, for the target atom we can use the collision trick that converts $|+\rangle_c$ to $|F_M, -F_M\rangle_z$ and $|-\rangle_c$ to $|F_M, -F_M + 1\rangle_z$ by adiabatically turning off the Ω_{+1} drive.

After the conversion above, we need to perform the entangling operation. Here we consider a Rydberg manifold with $F_R = F_M$. Consider using a σ_- polarized drive to address the state in both atoms from the encoded manifold to the Rydberg manifold. Due to the selectivity that comes from the polarization, $|F_M, -F_M\rangle_z$ will not couple with the Rydberg levels, but $|F_M, F_M\rangle_z$ and $|F_M, -F_M + 1\rangle_z$ will. Therefore, one can use the standard CZ control sequence [159] so that both $|F_M, -F_M\rangle_{z,C} \otimes |F_M, -F_M + 1\rangle_{z,T}$ and $|F_M, F_M\rangle_{z,C} \otimes |F_M, -F_M\rangle_{z,T}$ will pick up a phase $e^{i\phi}$, while $|F_M, F_M\rangle_{z,C} \otimes |F_M, -F_M + 1\rangle_{z,T}$ will get a phase $e^{i(2\phi-\pi)}$. This is equivalent to a CZ operation up to single-qubit phase rotations. Finally, we need to map the encoded subspaces from the Fock z basis back to the Fock x basis. We can choose a different trajectory compared with that used in the first step to account for the required extra 1-qubit gates. For the control atom, we still need to keep

the two SCSs antipodal, which is similar to the idea used in $Z(\alpha)$ gate construction. On the other hand, for the target atom, we can follow the 2nd and 3rd steps used in $X(\alpha)$ gate design. At the end of the protocol, only $|1\rangle_{c,C} \otimes |-\rangle_{c,T}$ will get a relative π phase, which results in a CX gate.

Finally, we would like to point out that this simplified version of CX execution can still be adapted to the "CZ biased-erasure" framework [83] (despite here CX itself is not BP), provided that the error from 1-qubit control is negligible and the Rydberg decay during entangling operations can be detected. If the decay is detected on the control atom, we prepare a fresh $|1\rangle_c$, and if it is on the target atom, we prepare a fresh $|-\rangle_c$. In this way, we again have the knowledge about both the location and the type of the error (for the control atom, the error is Z type, while for the target it is X type). This noise pattern is equivalent to the $\mathrm{CX} = Y_T(\frac{\pi}{2}) \cdot \mathrm{CZ} \cdot Y_T(-\frac{\pi}{2})$ construction proposed in [83] where the error during CZ in the middle is biased-erasure such that both the control and target will have Z type of error after erasure conversion. Z error on the target will be converted to X error after $Y_T(\frac{\pi}{2})$ rotation, which leads to the same error structure for the CX construction as we proposed here.

APPENDIX D

APPENDICES FOR CHAPTER 5

D.1 Miscellaneous derivations in the WKB analysis

D.1.1 Conjugate variables in the classical spin dynamics

In this part, I will try to show that $(f_z, -\phi)$ in Eq. (5.20) is a pair of conjugate variables in the classical dynamics. To do so, we consider the angular momentum l_i written as

$$l_i = \epsilon_{ijk} x_j p_k, \tag{D.1}$$

where the ϵ_{ijk} is the Levi-Civita symbol, $x_i \in \{x, y, z\}$ and $p_i \in \{p_x, p_y, p_z\}$.

To verify l_z and $-\phi = \arctan(l_y/l_x)$ is indeed a conjugate pair, we should compute the classical Poisson bracket:

$$\begin{aligned} \{l_z, -\phi\}_{\text{cl}} &= \sum_i \frac{\partial l_z}{\partial x_i}\frac{\partial(-\phi)}{\partial p_i} - \frac{\partial l_z}{\partial p_i}\frac{\partial(-\phi)}{\partial x_i} \\ &= p_y \frac{\partial(-\phi)}{\partial p_x} - p_x \frac{\partial(-\phi)}{\partial p_y} - x\frac{\partial(-\phi)}{\partial y} + y\frac{\partial(-\phi)}{\partial x}. \end{aligned} \tag{D.2}$$

On the other hand, from the definition of ϕ, we have

$$\frac{\partial\phi}{\partial p_x} = \frac{z l_x}{l_x^2 + l_y^2}, \quad \frac{\partial\phi}{\partial p_y} = \frac{z l_y}{l_x^2 + l_y^2}, \quad \frac{\partial\phi}{\partial y} = -\frac{p_z l_y}{l_x^2 + l_y^2}, \quad \frac{\partial\phi}{\partial x} = -\frac{p_z l_x}{l_x^2 + l_y^2}. \tag{D.3}$$

With this, we can confirm that $\{l_z, -\phi\}_{\text{cl}} = 1$, and therefore it is a conjugate pair we need. It can be generalized into any spin dynamics like the $(f_z, -\phi)$ pair we meet here, where f is the total angular momentum of a hyperfine manifold.

D.1.2 Derivation of Eq. (5.25) *and* (5.26)

Now I will use WKB to investigate the wavefunctions $\psi(f_z) = \langle F, f_z|\psi\rangle$ [see Eq. (5.24)] of the eigenstates in $\hat{\mathbf{H}}_{\mathrm{RN}}$ (5.19). The derivation follows Ref. [165, 166], and I place it here for completeness.

To begin with, we should check how spin operators act on the wavefunctions $\psi(f_z)$. Notice that

$$\hat{F}_{\pm}\,|F, f_z\rangle = b\left[\sqrt{f_z(f_z \pm 1)}\right]|F, f_z \pm 1\rangle\,, \tag{D.4}$$

where $b(f_z) = \sqrt{F(F+1) - f_z^2}$ is already defined in Eq. (5.22). When acting on the wavefunctions, we have

$$\begin{cases} (\hat{F}_{+}\psi)(f_z) = b\left[\sqrt{f_z(f_z-1)}\right]\psi(f_z-1) = b\left[\sqrt{f_z(f_z-1)}\right]e^{-\mathrm{d}_{f_z}}\psi(f_z), \\ (\hat{F}_{-}\psi)(f_z) = b\left[\sqrt{f_z(f_z+1)}\right]\psi(f_z+1) = b\left[\sqrt{f_z(f_z+1)}\right]e^{\mathrm{d}_{f_z}}\psi(f_z). \end{cases} \tag{D.5}$$

On the other hand, if $|\psi\rangle$ is an eigenstate of $\hat{\mathbf{H}}_{\mathrm{RN}}$ with eigenvalue $\mathbf{E}$, then we should have

$$(\hat{\mathbf{H}}_{\mathrm{RN}}\psi)(f_z) = \mathbf{E}\psi(f_z), \tag{D.6}$$

which implies

$$\psi^{-1}(f_z)\cdot(\hat{\mathbf{H}}_{\mathrm{RN}}\psi)(f_z) = \mathbf{E}. \tag{D.7}$$

Given the WKB ansatz $\psi(f_z) = \exp[iS(f_z)]$ (5.24), the equations shown above are sufficient to derive formulae like Eq. (5.25) and (5.26). However, here are several tricks that can be used in the derivations. First, functions $b\left[\sqrt{f_z(f_z \pm 1)}\right]$ can be Taylor expanded as

$$b\left[\sqrt{f_z(f_z \pm 1)}\right] = b(f_z) \pm \frac{b'(f_z)}{2} + O(1/F). \tag{D.8}$$

With this, we can compute terms like

$$\begin{aligned}\psi^{-1}\hat{F}_x\psi &= \frac{e^{-iS(f_z)}}{2}\left\{\left[b(f_z)-\frac{b'(f_z)}{2}\right]e^{iS(f_z-1)}+\left[b(f_z)+\frac{b'(f_z)}{2}\right]e^{iS(f_z+1)}\right\}. \\ &= b(f_z)e^{iS''(f_z)/2}\cos[S'(f_z)]+\frac{ib'(f_z)e^{iS''(f_z)/2}}{2}\sin[S'(f_z)].\end{aligned} \tag{D.9}$$

Here we have omitted the $O(1/F)$ term, while utilizing the Taylor expansion

$$S(f_z\pm 1)=S(f_z)\pm S'(f_z)+\frac{S''(f_z)}{2}+o(1/F). \tag{D.10}$$

To further simplify Eq. (D.9), we can apply the scaling analysis. We can expand $S = S_0 + S_1 + O(1/F)$, where $S_0 \sim O(F)$ and $S_1 \sim O(1)$. Meanwhile, d_{f_z} should scale as $\mathrm{d}_{f_z} \sim O(1/F)$, which explains why the residual terms in Eq. (D.10) are higher-ordered than $1/F$. With this, we can re-write Eq. (D.9) while keeping up to $O(1)$ terms as

$$\begin{aligned}\psi^{-1}\hat{F}_x\psi &= b(f_z)\cos(S_0')+\left\{\frac{ib(f_z)}{2}[\sin(S_0')]'+\left[\frac{ib'(f_z)}{2}-b(f_z)S_1'\right]\sin(S_0')\right\} \\ &= b(f_z)\cos(S_0')+\frac{i}{2}e^{-2iS_1}\left[b(f_z)\sin(S_0')e^{2iS_1}\right]'.\end{aligned} \tag{D.11}$$

The first term in the equation above scales as $O(F)$ while the second term is in $O(1)$. Similarly, we can compute $\psi^{-1}\hat{F}_y^2\psi$ while keeping up to $O(F)$ order:

$$\begin{aligned}\psi^{-1}\hat{F}_y^2\psi &= b^2(f_z)\sin^2(S_0')-\left\{\frac{ib^2(f_z)}{2}[\sin(2S_0')]'+b(f_z)\left[ib'(f_z)-b(f_z)S_1'\right]\sin(2S_0')\right\} \\ &= b^2(f_z)\sin^2(S_0')-\frac{i}{2}e^{-2iS_1}\left[b^2(f_z)\sin(2S_0')e^{2iS_1}\right]'.\end{aligned} \tag{D.12}$$

The first term scales as $O(F^2)$, while the second term is in $O(F)$.

Finally, we should put those expressions into Eq. (D.7) and perform order counting. Assume $\mathbf{E}$ in general scales as $O(F^2)$ as well. Then, $O(F^2)$ terms in Eq. (D.7) will form the

classical Hamilton-Jacobi equation shown in Eq. (5.25), while the $O(F)$ terms will form the probability current conservation formula shown in Eq. (5.26).

D.1.3 Tunneling rate derivation

Derivation of Eq. (5.39)

Now I start to derive the eigen-energy difference of two near-degenerate eigenstates in $\hat{\mathbf{H}}_{\mathrm{RN}}$, using the wavefunctions we derived through the WKB method. I will use a trick developed in Ref. [165, 167], which requires us to focus on the wavefunctions $\psi_{\pm}(f_z)$ introduced in Eq. (5.38). Notice that

$$\hat{\mathbf{H}}_{\mathrm{RN}}\psi_{\pm}(f_z) = \mathbf{E}_{\pm}\psi_{\pm}(f_z), \tag{D.13}$$

which implies

$$\begin{aligned}
&\sum_{f_z=0}^{F}\Big[\psi_{-}(f_z)\hat{\mathbf{H}}_{\mathrm{RN}}\psi_{+}(f_z) - \psi_{+}(f_z)\hat{\mathbf{H}}_{\mathrm{RN}}\psi_{-}(f_z)\Big] \\
=\ &\mathbf{E}_{+}\sum_{f_z=0}^{F}\psi_{-}(f_z)\psi_{+}(f_z) - \mathbf{E}_{-}\sum_{f_z=0}^{F}\psi_{+}(f_z)\psi_{-}(f_z).
\end{aligned} \tag{D.14}$$

The summation here only covers the $f_z \geq 0$ regime. Therefore, the second line in Eq. (D.14) will be approximately $\frac{\mathbf{E}_{+}-\mathbf{E}_{-}}{2}$, since the overlap between $\psi_{+}(f_z)$ and $\psi_{-}(f_z)$ in this regime mainly comes from $\psi_{\uparrow}(f_z)$. On the other hand, from Eq. (D.5) we can compute the first line of Eq. (D.14), which gives

$$\begin{aligned}
&\sum_{f_z=0}^{F}\Big[\psi_{-}(f_z)\hat{\mathbf{H}}_{\mathrm{RN}}\psi_{+}(f_z) - \psi_{+}(f_z)\hat{\mathbf{H}}_{\mathrm{RN}}\psi_{-}(f_z)\Big] \\
=\ &b(0)\frac{\cos^2(\theta)}{4}[b(\sqrt{2})\psi_{+}(0)\psi_{-}(2) + 2b(0)\psi_{+}(1)\psi_{-}(1)] \\
&-\frac{(2F-1)\cos(\theta)}{2}b(0)\psi_{+}(0)\psi_{-}(1).
\end{aligned} \tag{D.15}$$

Due to the continuity of the functions, the expression above can be further simplified. With the explicit expressions of the wavefunction $\psi_\uparrow(f_z)$ near $f_z = 0$ [see Eq. (5.34)], we have

$$\begin{cases} b(\sqrt{2}) \simeq b(0), \\ \psi_-(1) = \sqrt{2}\psi_\uparrow(0)\sinh\left[\int_0^1 p_I(f)\,\mathrm{d}f\right] \simeq \sqrt{2}\psi_\uparrow(0)\sinh[p_I(0)], \\ \psi_-(2) \simeq \sqrt{2}\psi_\uparrow(0)\sinh[2p_I(0)], \\ \psi_+(1) \simeq \sqrt{2}\psi_\uparrow(0)\cosh[p_I(0)]. \end{cases} \tag{D.16}$$

With these approximations, we can get a simplified expression for $\mathbf{E}_{\mathrm{dif}}$, where

$$\begin{aligned} \mathbf{E}_{\mathrm{dif}} &= 2\cdot\Big|b^2(0)\cos^2(\theta)\sinh[2p_I(0)] - b(0)(2F-1)\cos(\theta)\sinh[p_I(0)]\Big|\cdot\psi_\uparrow^2(0) \\ &= 2|\dot{f}_z(f_z=0)|\cdot\psi_\uparrow^2(0) \simeq \frac{2}{T_{\mathrm{cl}}}\exp\Big[-2\int_0^{f_{z,-}} p_I(f)\,\mathrm{d}f\Big], \end{aligned} \tag{D.17}$$

which is exactly Eq. (5.39) shown in the main text.

Derivation of Eq. (5.61)

Now I continue to derive Eq. (5.61), where the Hamiltonian $\hat{\mathbf{H}}'_{\mathrm{RN}}$ (5.52) itself contains the E-dependence. For two near-degenerate eigen-energies $E_\pm$ of $\hat{H}'_{\mathrm{DS}}$ (5.47) with $\Delta = 0$, we have

$$\hat{\mathbf{H}}'_{\mathrm{RN}}(E_\pm) = \hat{\mathbf{H}}'_{\mathrm{RN}}(\bar{E}) \mp \frac{8F(2F+1)}{\Omega^2}\frac{E_+ - E_-}{2}\delta\hat{F}_x. \tag{D.18}$$

It is worth to remind here that, $\mathbf{E}_\pm$, the eigenvalue of $\hat{\mathbf{H}}'_{\mathrm{RN}}(E_\pm)$, relate to $E_\pm$ through $\mathbf{E}_\pm = \frac{8F(2F+1)}{\Omega^2}E_\pm^2$.

Then we can apply this trick similar to Eq. (D.14), which means to consider

$$\begin{aligned}
&\sum_{f_z=0}^{F}\Big[\psi_-(f_z)\hat{\mathbf{H}}'_{\mathrm{RN}}(\bar{E})\psi_+(f_z)-\psi_+(f_z)\hat{\mathbf{H}}'_{\mathrm{RN}}(\bar{E})\psi_-(f_z)\Big]\\
&-\frac{8F(2F+1)}{\Omega^2}\frac{E_+-E_-}{2}\delta\sum_{f_z=0}^{F}\Big[\psi_-(f_z)\hat{F}_x\psi_+(f_z)+\psi_+(f_z)\hat{F}_x\psi_-(f_z)\Big]\\
=&\,(\mathbf{E}_+-\mathbf{E}_-)\sum_{f_z=0}^{F}\psi_+(f_z)\psi_-(f_z).
\end{aligned}\tag{D.19}$$

The third line in Eq. (D.19) will be again $\frac{\mathbf{E}_+-\mathbf{E}_-}{2}$ as I explained before, while the summation in the second line will give $\langle\hat{F}_x\rangle$, the expectation value of $\hat{F}_x$ with state $\psi(f_z)$. The first line also leaves the following expression again:

$$\begin{aligned}
&\sum_{f_z=0}^{F}\Big[\psi_-(f_z)\hat{\mathbf{H}}'_{\mathrm{RN}}(\bar{E})\psi_+(f_z)-\psi_+(f_z)\hat{\mathbf{H}}'_{\mathrm{RN}}(\bar{E})\psi_-(f_z)\Big]\\
=&\,b(0)\frac{\cos^2(\theta)}{4}[b(\sqrt{2})\psi_+(0)\psi_-(2)+2b(0)\psi_+(1)\psi_-(1)]-\frac{\beta(\bar{E})}{2}b(0)\psi_+(0)\psi_-(1).
\end{aligned}\tag{D.20}$$

However, now we should use the wavefunction $\psi(f_z)$ in the form of Eq. (5.59) near $f_z=0$, and denote ψ_0 shown below for further use, where

$$\psi_0=\frac{Ce^{i\chi(0)}}{\sqrt{|\dot{f}_z(0)|}}\exp\Big[-\int_0^{f_{z,-}}p_I(f)\,\mathrm{d}f\Big]\exp\Big[-i\int_0^{f_{z,0}}p_R(f)\,\mathrm{d}f\Big].\tag{D.21}$$

With this, we have

$$\begin{cases}
\psi_+(0)\simeq\sqrt{2}\psi_0+\text{c.c.},\\
\psi_-(1)\simeq\sqrt{2}\psi_0\sinh[p_I(0)+ip_R(0)]+\text{c.c.},\\
\psi_-(2)\simeq\sqrt{2}\psi_0\sinh\{2[p_I(0)+ip_R(0)]\}+\text{c.c.},\\
\psi_+(1)\simeq\sqrt{2}\psi_0\cosh[p_I(0)+ip_R(0)]+\text{c.c.}
\end{cases}\tag{D.22}$$

When putting Eq. (D.22) directly into Eq. (D.20), we can find exactly

$$\begin{aligned}
&\sum_{f_z=0}^{F}\Big[\psi_-(f_z)\hat{\mathbf{H}}'_{\mathrm{RN}}(\bar{E})\psi_+(f_z)-\psi_+(f_z)\hat{\mathbf{H}}'_{\mathrm{RN}}(\bar{E})\psi_-(f_z)\Big]\simeq -i\psi_0^2\dot{f}_z(0)+\text{c.c.}\\
=&-\frac{2}{T_{\mathrm{cl}}}\exp\Big[-2\int_0^{f_{z,-}}p_I(f)\,\mathrm{d}f\Big]\cos\Big[2\int_0^{f_{z,0}}p_R(f)\,\mathrm{d}f\Big],
\end{aligned}\tag{D.23}$$

where we have used the definition of $\chi(f_z)$ (5.60) in the above derivation that

$$\begin{cases}
\dot{f}_z=-ie^{-2i\chi}|\dot{f}_z|,\\
\sinh(p_I+ip_R)=ie^{-2i\chi}|\sinh(p_I+ip_R)|,\\
\sinh(p_I-ip_R)=-e^{4i\chi}\sinh(p_I+ip_R).
\end{cases}\tag{D.24}$$

We can finally achieve Eq. (5.61) by substituting Eq. (D.23) into Eq. (D.19).

APPENDIX E

APPENDICES FOR CHAPTER 7

E.1 Proof of the theorems

Lemma E.1 (Chernoff bounds [195]). *Let $X_1, \ldots, X_n$ be independent Bernoulli random variables taking values in $\{0, 1\}$, such that $P(X_i = 1) = p_i$. denote $X = \sum_{i=1}^{n} X_i$ and $\mu = \sum_{i=1}^{n} p_i$ is the expectation value of X. Then the following Chernoff bounds hold:*

$$P[X \geq (1+\delta)\mu] \leq \left(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right)^{\mu} \leq e^{-\frac{\delta^2\mu}{2+\delta}} \qquad (\delta > 0); \tag{E.1}$$

$$P[X \leq (1-\delta)\mu] \leq \left(\frac{e^{-\delta}}{(1-\delta)^{(1-\delta)}}\right)^{\mu} \leq e^{-\frac{\delta^2\mu}{2}} \qquad (0 < \delta < 1); \tag{E.2}$$

$$P(|X-\mu| \geq \delta\mu) \leq 2e^{-\frac{\delta^2\mu}{3}} \qquad (0 < \delta < 1). \tag{E.3}$$

Eq. (E.2), (E.3) and the first inequality in Eq. (E.1) are all proved in Ref. [195]. Here I want to show the second inequality in Eq. (E.1), which is equivalent to show $\delta - (1+\delta)\log(1+\delta) \leq -\frac{\delta^2}{2+\delta}$, or $\frac{2\delta}{2+\delta} \leq \log(1+\delta)$ when $\delta > 0$. Denote $f(\delta) = \log(1+\delta) - \frac{2\delta}{2+\delta}$. We can show that $f(0) = 0$ and $f'(\delta) = \frac{1}{1+\delta} - \frac{4}{(2+\delta)^2} = \frac{\delta^2}{(1+\delta)(2+\delta)^2}$. When $\delta > 0$, we will have $f'(\delta) > 0$ and therefore $f(\delta) \geq 0$ as well, which concludes the proof of the desired inequality.

E.1.1 Proof of the Theorem. 7.1

In each run of the n-qubit Bell measurement experiments, we can assign a random variable $X_i = 1$ if we get an odd number of singlet outcomes, while $X_i = 0$ otherwise. Then each $X_i \in \{0, 1\}$ will be the i.i.d. Bernoulli random variable with the expectation value $\mathbb{E}[X_i] = p_A$. As discussed in Chapter 7, p_A satisfies $\epsilon - \epsilon^2 \leq p_A < \epsilon - \epsilon^2/2$.

We first consider the number of Bell measurements N_m that can bound the estimation on p_A. Here we denote $\hat{p}_A = \frac{\sum_{i=1}^{N_m} X_i}{N_m}$, and choose $r' = r/2$ such that $0 < r' < 1$. From

Eq. (E.3), we know that the following inequality holds

$$P(|\hat{p}_A - p_A| \geq r' p_A) \leq \delta, \tag{E.4}$$

if

$$N_m \geq \frac{3}{r'^2 p_A} \log(2/\delta). \tag{E.5}$$

Then let us choose an estimator of ϵ by

$$\hat{\epsilon} = \frac{2}{3}\left(1 - \sqrt{1 - 3\hat{p}_A}\right). \tag{E.6}$$

With this choice, we have $\hat{\epsilon} - 3\hat{\epsilon}^2/4 = \hat{p}_A$. I will try to show that, when ϵ is smaller than a constant, then $|\hat{p}_A - p_A| < r' p_A$ will directly imply $|\hat{\epsilon} - \epsilon| < r\epsilon$.

We first assume that $\epsilon < \frac{1}{3(1+r')}$, which will guarantee $(1+r')p_A < (1+r')\epsilon < \frac{1}{3}$. Therefore, when $\hat{p}_A < (1+r')p_A$, Eq. (E.6) will always give a valid $\hat{\epsilon}$ estimation with $\hat{\epsilon} \leq \frac{2}{3}$.

Then, I will illustrate that, when ϵ is smaller than a constant, then $\hat{p}_A < (1+r')p_A$ will imply $\hat{\epsilon} < (1+r)\epsilon$. Since $(1+r)\epsilon < \frac{1+r}{3(1+r')} < \frac{2}{3}$, it is equivalent to show

$$\hat{p}_A = \hat{\epsilon} - 3\hat{\epsilon}^2/4 < (1+r)\epsilon - 3(1+r)^2\epsilon^2/4. \tag{E.7}$$

On the other hand, $\hat{p}_A < (1+r')p_A < (1+r')(\epsilon - \epsilon^2/2)$. As long as ϵ also satisfies

$$0 < \epsilon < \frac{4(r-r')}{3(1+r)^2 - 2(1+r')} = \frac{2r}{3r^2 + 5r + 1}, \tag{E.8}$$

then $\hat{p}_A < (1+r')p_A$ will give $\hat{\epsilon} < (1+r)\epsilon$. Meanwhile, when $0 < r < 1$, we always have $\frac{2r}{3r^2+5r+1} < \frac{2}{3(2+r)}$. With this, Eq. (E.8) will also guarantee $\epsilon < \frac{1}{3(1+r')}$.

Similarly, when $r \in (0,1)$, Eq. (E.8) also implies

$$0 < \epsilon < \frac{2r}{3r^2+5r+1} < \frac{2r}{r(4-3r)+1} = \frac{4(r-r')}{4(1-r')-3(1-r)^2}. \tag{E.9}$$

Therefore, $(1-r')p_A < \hat{p}_A$ gives

$$(1-r)\epsilon - 3(1-r)^2\epsilon^2/4 < (1-r')(\epsilon-\epsilon^2) < (1-r')p_A < \hat{p}_A = \hat{\epsilon} - 3\hat{\epsilon}^2/4. \tag{E.10}$$

Since both $(1-r)\epsilon$ and $\hat{\epsilon}$ are smaller than $\frac{2}{3}$, the inequality above will also imply $(1-r)\epsilon < \hat{\epsilon}$.

From the above discussion, we understand that, when Eq. (E.8) holds, then $|\hat{p}_A - p_A| < r'p_A$ will give $|\hat{\epsilon}-\epsilon| < r\epsilon$. Equivalently, $P(|\hat{\epsilon}-\epsilon| \geq r\epsilon) \leq P(|\hat{p}_A - p_A| \geq r'p_A)$. On the other hand, when $0 < \epsilon < \frac{1}{3}$, we have $\frac{1}{p_A} \leq \frac{1}{\epsilon-\epsilon^2} < \frac{3}{2\epsilon}$. Therefore, when choosing

$$N = 2N_m \geq \frac{36}{r^2\epsilon}\log(2/\delta) > \frac{6}{r'^2 p_A}\log(2/\delta), \tag{E.11}$$

we will have

$$P(|\hat{\epsilon}-\epsilon| \geq r\epsilon) \leq P(|\hat{p}_A - p_A| \geq r'p_A) \leq \delta. \tag{E.12}$$

Here we use $N = 2N_m$ since one Bell measurement consumes two copies of the unknown states. This concludes the proof on the $N \sim O[\frac{1}{\epsilon}\log(1/\delta)]$ overhead scaling.

E.1.2 Proof of the Theorem. 7.2

In order to get a precise estimation on ϵ, here our approach is to estimate each $\epsilon_{(\alpha,s)}$ and then sum them up. Here we denote k as the number of irreducible representations other than $\mathcal{H}_{(0,0)}$, or equivalently there are k different $\epsilon_{(\alpha,s)}$ variables to estimate. For the simplicity of the notation, here we use i to label each (α,s), and denote $d_i = \dim[\mathcal{H}_{(\alpha,s)}]$ to be the dimension of each irreducible representation.

We notice that, when measuring the overlap with each stabilizer state $|S_i\rangle$ in $\mathcal{H}_i$, we will

get a series of i.i.d. Bernoulli variables $X_j^{(i)} \in \{0,1\}$ with the expectation value $\mathbb{E}_j[X_j^{(i)}] = \frac{\epsilon_i}{d_i}$. Therefore, each ϵ_i can be estimated by $\hat{\epsilon}_i = d_i \frac{\sum_{j=1}^{N_i} X_j^{(i)}}{N_i}$. With this, we can achieve $\hat{\epsilon} = \sum_{i=1}^{k} \hat{\epsilon}_i$ in the end.

From Eq. (E.1) and (E.2), we will know that if N_i satisfies

$$N_i > \frac{kd_i(2k+r)}{r^2\epsilon} \log(2k/\delta) \geq \left\{ \frac{2(\epsilon_i/d_i)}{[r\epsilon/(kd_i)]^2} + \frac{1}{r\epsilon/(kd_i)} \right\} \log(2k/\delta) > \frac{2(\epsilon_i/d_i)}{[r\epsilon/(kd_i)]^2} \log(2k/\delta), \tag{E.13}$$

then the following two inequalities will both hold:

$$\begin{cases} P\left(\hat{\epsilon}_i \geq \epsilon_i + \dfrac{r\epsilon}{k}\right) = P\left(\dfrac{\hat{\epsilon}_i}{d_i} \geq \dfrac{\epsilon_i}{d_i} + \dfrac{r\epsilon}{kd_i}\right) < \dfrac{\delta}{2k}, & \text{(E.14)} \\ P\left(\hat{\epsilon}_i \leq \epsilon_i - \dfrac{r\epsilon}{k}\right) = P\left(\dfrac{\hat{\epsilon}_i}{d_i} \leq \dfrac{\epsilon_i}{d_i} - \dfrac{r\epsilon}{kd_i}\right) < \dfrac{\delta}{2k}. & \text{(E.15)} \end{cases}$$

Here we want to point out that we have used $\epsilon_i \leq \epsilon$ in Eq. (E.13), while Eq. (E.15) will be trivially correct if $\epsilon_i \leq r\epsilon/k$. As a result, when $N_i > \frac{kd_i(2k+r)}{r^2\epsilon} \log(2k/\delta)$, we have

$$P\left(|\hat{\epsilon}_i - \epsilon_i| < \frac{r\epsilon}{k}\right) \geq 1 - \frac{\delta}{k}. \tag{E.16}$$

Further, if Eq. (E.13) holds for every N_i, then we will get

$$P\left(|\hat{\epsilon} - \epsilon| < r\epsilon\right) \geq P\left(\forall i : |\hat{\epsilon}_i - \epsilon_i| < \frac{r\epsilon}{k}\right) \geq 1 - \delta. \tag{E.17}$$

The total cost will be

$$N = \sum_{i=1}^{k} N_i > \frac{k(d-1)(2k+r)}{r^2\epsilon} \log(2k/\delta). \tag{E.18}$$

Here $d = 2^n = 1 + \sum_{i=1}^{k} d_i$ is the dimension of the n-qubit Hilbert space. The $N \sim O[\frac{1}{\epsilon} \log(1/\delta)]$ overhead scaling is justified again.

BIBLIOGRAPHY

[1] E. Schrödinger, "Die gegenwärtige Situation in der Quantenmechanik", Naturwissenschaften **23**, 807 (1935).

[2] A. Einstein, B. Podolsky, and N. Rosen, "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?", Physical Review **47**, 777 (1935).

[3] R. P. Feynman, "Simulating physics with computers", International Journal of Theoretical Physics **21**, 467 (1982).

[4] P. Shor, "Algorithms for quantum computation: discrete logarithms and factoring", in Proceedings 35th Annual Symposium on Foundations of Computer Science (1994), pp. 124–134.

[5] N. A. Gershenfeld and I. L. Chuang, "Bulk Spin-Resonance Quantum Computation", Science **275**, 350 (1997).

[6] I. L. Chuang, L. M. K. Vandersypen, X. Zhou, D. W. Leung, and S. Lloyd, "Experimental realization of a quantum algorithm", Nature **393**, 143 (1998).

[7] L. M. K. Vandersypen, M. Steffen, G. Breyta, C. S. Yannoni, M. H. Sherwood, and I. L. Chuang, "Experimental realization of Shor's quantum factoring algorithm using nuclear magnetic resonance", Nature **414**, 883 (2001).

[8] J. I. Cirac and P. Zoller, "Quantum Computations with Cold Trapped Ions", Physical Review Letters **74**, 4091 (1995).

[9] C. J. Ballance, T. P. Harty, N. M. Linke, M. A. Sepiol, and D. M. Lucas, "High-Fidelity Quantum Logic Gates Using Trapped-Ion Hyperfine Qubits", Physical Review Letters **117**, 060504 (2016).

[10] C. D. Bruzewicz, J. Chiaverini, R. McConnell, and J. M. Sage, "Trapped-ion quantum computing: Progress and challenges", Applied Physics Reviews **6**, 021314 (2019).

[11] S. A. Moses, C. H. Baldwin, M. S. Allman, R. Ancona, L. Ascarrunz, C. Barnes, J. Bartolotta, B. Bjork, P. Blanchard, M. Bohn, J. G. Bohnet, N. C. Brown, N. Q. Burdick, W. C. Burton, S. L. Campbell, et al., "A Race-Track Trapped-Ion Quantum Processor", Physical Review X **13**, 041052 (2023).

[12] M. Iqbal, N. Tantivasadakarn, R. Verresen, S. L. Campbell, J. M. Dreiling, C. Figgatt, J. P. Gaebler, J. Johansen, M. Mills, S. A. Moses, J. M. Pino, A. Ransford, M. Rowe, P. Siegfried, R. P. Stutz, et al., "Non-Abelian topological order and anyons on a trapped-ion processor", Nature **626**, 505 (2024).

[13] D. Jaksch, J. I. Cirac, P. Zoller, S. L. Rolston, R. Côté, and M. D. Lukin, "Fast Quantum Gates for Neutral Atoms", Physical Review Letters **85**, 2208 (2000).

[14] 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. Vuletić, and M. D. Lukin, "High-fidelity parallel entangling gates on a neutral-atom quantum computer", Nature **622**, 268 (2023).

[15] 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. Bonilla Ataides, N. Maskara, I. Cong, X. Gao, P. Sales Rodriguez, et al., "Logical quantum processor based on reconfigurable atom arrays", Nature **626**, 58 (2024).

[16] H. J. Manetsch, G. Nomura, E. Bataille, K. H. Leung, X. Lv, and M. Endres, *A tweezer array with 6100 highly coherent atomic qubits*, arXiv:2403.12021 [quant-ph].

[17] J. Koch, T. M. Yu, J. Gambetta, A. A. Houck, D. I. Schuster, J. Majer, A. Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, "Charge-insensitive qubit design derived from the Cooper pair box", Physical Review A **76**, 042319 (2007).

[18] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, B. Burkett, Y. Chen, Z. Chen, B. Chiaro, R. Collins, et al., "Quantum supremacy using a programmable superconducting processor", Nature **574**, 505 (2019).

[19] A. Blais, A. L. Grimsmo, S. M. Girvin, and A. Wallraff, "Circuit quantum electrodynamics", Reviews of Modern Physics **93**, 025005 (2021).

[20] Google Quantum AI and Collaborators, "Quantum error correction below the surface code threshold", Nature **638**, 920 (2025).

[21] G. Waldherr, Y. Wang, S. Zaiser, M. Jamali, T. Schulte-Herbrüggen, H. Abe, T. Ohshima, J. Isoya, J. F. Du, P. Neumann, and J. Wrachtrup, "Quantum error correction in a solid-state hybrid spin register", Nature **506**, 204 (2014).

[22] M. K. Bhaskar, R. Riedinger, B. Machielse, D. S. Levonian, C. T. Nguyen, E. N. Knall, H. Park, D. Englund, M. Lončar, D. D. Sukachev, and M. D. Lukin, "Experimental demonstration of memory-enhanced quantum communication", Nature **580**, 60 (2020).

[23] M. H. Abobeih, Y. Wang, J. Randall, S. J. H. Loenen, C. E. Bradley, M. Markham, D. J. Twitchen, B. M. Terhal, and T. H. Taminiau, "Fault-tolerant operation of a logical qubit in a diamond quantum processor", Nature **606**, 884 (2022).

[24] C. M. Knaut, A. Suleymanzade, Y.-C. Wei, D. R. Assumpcao, P.-J. Stas, Y. Q. Huan, B. Machielse, E. N. Knall, M. Sutula, G. Baranes, N. Sinclair, C. De-Eknamkul, D. S. Levonian, M. K. Bhaskar, H. Park, et al., "Entanglement of nanophotonic quantum memory nodes in a telecom network", Nature **629**, 573 (2024).

[25] H.-S. Zhong, H. Wang, Y.-H. Deng, M.-C. Chen, L.-C. Peng, Y.-H. Luo, J. Qin, D. Wu, X. Ding, Y. Hu, P. Hu, X.-Y. Yang, W.-J. Zhang, H. Li, Y. Li, et al., "Quantum computational advantage using photons", Science **370**, 1460 (2020).

[26] S. G. J. Philips, M. T. Mądzik, S. V. Amitonov, S. L. de Snoo, M. Russ, N. Kalhor, C. Volk, W. I. L. Lawrie, D. Brousse, L. Tryputen, B. P. Wuetz, A. Sammak, M. Veldhorst, G. Scappucci, and L. M. K. Vandersypen, "Universal control of a six-qubit quantum processor in silicon", Nature **609**, 919 (2022).

[27] PsiQuantum team, "A manufacturable platform for photonic quantum computing", Nature **641**, 876 (2025).

[28] D. P. DiVincenzo, "The Physical Implementation of Quantum Computation", Fortschritte der Physik **48**, 771 (2000).

[29] P. Shor, "Fault-tolerant quantum computation", in Proceedings of 37th Conference on Foundations of Computer Science (1996), pp. 56–65.

[30] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, 2010).

[31] D. Gottesman, *An Introduction to Quantum Error Correction and Fault-Tolerant Quantum Computation*, arXiv:0904.2557 [quant-ph].

[32] D. Aharonov and M. Ben-Or, *Fault Tolerant Quantum Computation with Constant Error*, 1996, arXiv:quant-ph/9611025.

[33] E. Knill, R. Laflamme, and W. H. Zurek, "Resilient Quantum Computation", Science **279**, 342 (1998).

[34] A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, "Surface codes: Towards practical large-scale quantum computation", Physical Review A **86**, 032324 (2012).

[35] J. Preskill, "Quantum Computing in the NISQ era and beyond", Quantum **2**, 79 (2018).

[36] V. Ramakrishna, M. V. Salapaka, M. Dahleh, H. Rabitz, and A. Peirce, "Controllability of molecular systems", Physical Review A **51**, 960 (1995).

[37] S. G. Schirmer, H. Fu, and A. I. Solomon, "Complete controllability of quantum systems", Physical Review A **63**, 063410 (2001).

[38] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, and S. J. Glaser, "Optimal control of coupled spin dynamics: design of NMR pulse sequences by gradient ascent algorithms", Journal of Magnetic Resonance **172**, 296 (2005).

[39] S. Machnes, U. Sander, S. J. Glaser, P. de Fouquières, A. Gruslys, S. Schirmer, and T. Schulte-Herbrüggen, "Comparing, optimizing, and benchmarking quantum-control algorithms in a unifying programming framework", Physical Review A **84**, 022305 (2011).

[40] N. Leung, M. Abdelhafez, J. Koch, and D. Schuster, "Speedup for quantum optimal control from automatic differentiation based on graphics processing units", Physical Review A **95**, 042318 (2017).

[41] J. T. Merrill and K. R. Brown, in *Quantum Information and Computation for Chemistry* (John Wiley & Sons, Ltd, 2014), pp. 241–294.

[42] C. Fromonteil, D. Bluvstein, and H. Pichler, "Protocols for Rydberg Entangling Gates Featuring Robustness against Quasistatic Errors", PRX Quantum **4**, 020335 (2023).

[43] F. Motzoi, J. M. Gambetta, P. Rebentrost, and F. K. Wilhelm, "Simple Pulses for Elimination of Leakage in Weakly Nonlinear Qubits", Physical Review Letters **103**, 110501 (2009).

[44] A. A. Clerk, M. H. Devoret, S. M. Girvin, F. Marquardt, and R. J. Schoelkopf, "Introduction to quantum noise, measurement, and amplification", Reviews of Modern Physics **82**, 1155 (2010).

[45] L. Viola and S. Lloyd, "Dynamical suppression of decoherence in two-state quantum systems", Physical Review A **58**, 2733 (1998).

[46] G. S. Uhrig, "Keeping a Quantum Bit Alive by Optimized π-Pulse Sequences", Physical Review Letters **98**, 100504 (2007).

[47] M. Abdelhafez, D. I. Schuster, and J. Koch, "Gradient-based optimal control of open quantum systems using quantum trajectories and automatic differentiation", Physical Review A **99**, 052327 (2019).

[48] V. V. Sivak, A. Eickbusch, H. Liu, B. Royer, I. Tsioutsios, and M. H. Devoret, "Model-Free Quantum Control with Reinforcement Learning", Physical Review X **12**, 011059 (2022).

[49] L. Ding, M. Hays, Y. Sung, B. Kannan, J. An, A. Di Paolo, A. H. Karamlou, T. M. Hazard, K. Azar, D. K. Kim, B. M. Niedzielski, A. Melville, M. E. Schwartz, J. L. Yoder, T. P. Orlando, et al., "High-Fidelity, Frequency-Flexible Two-Qubit Fluxonium Gates with a Transmon Coupler", Physical Review X **13**, 031035 (2023).

[50] V. V. Sivak, A. Eickbusch, B. Royer, S. Singh, I. Tsioutsios, S. Ganjam, A. Miano, B. L. Brock, A. Z. Ding, L. Frunzio, S. M. Girvin, R. J. Schoelkopf, and M. H. Devoret, "Real-time quantum error correction beyond break-even", Nature **616**, 50 (2023).

[51] A. Lingenfelter, D. Roberts, and A. A. Clerk, "Unconditional Fock state generation using arbitrarily weak photonic nonlinearities", Science Advances **7**, eabj1916 (2021).

[52] A. Eickbusch, V. Sivak, A. Z. Ding, S. S. Elder, S. R. Jha, J. Venkatraman, B. Royer, S. M. Girvin, R. J. Schoelkopf, and M. H. Devoret, "Fast universal control of an oscillator with weak dispersive coupling to a qubit", Nature Physics **18**, 1464 (2022).

[53] D. Roberts and A. A. Clerk, "Driven-Dissipative Quantum Kerr Resonators: New Exact Solutions, Photon Blockade and Quantum Bistability", Physical Review X **10**, 021022 (2020).

[54] M. Mirrahimi, Z. Leghtas, V. V. Albert, S. Touzard, R. J. Schoelkopf, L. Jiang, and M. H. Devoret, "Dynamically protected cat-qubits: a new paradigm for universal quantum computation", New Journal of Physics **16**, 045014 (2014).

[55] Z. Leghtas, S. Touzard, I. M. Pop, A. Kou, B. Vlastakis, A. Petrenko, K. M. Sliwa, A. Narla, S. Shankar, M. J. Hatridge, M. Reagor, L. Frunzio, R. J. Schoelkopf, M. Mirrahimi, and M. H. Devoret, "Confining the state of light to a quantum manifold by engineered two-photon loss", Science **347**, 853 (2015).

[56] N. Ofek, A. Petrenko, R. Heeres, P. Reinhold, Z. Leghtas, B. Vlastakis, Y. Liu, L. Frunzio, S. M. Girvin, L. Jiang, M. Mirrahimi, M. H. Devoret, and R. J. Schoelkopf, "Extending the lifetime of a quantum bit with error correction in superconducting circuits", Nature **536**, 441 (2016).

[57] S. Puri, S. Boutin, and A. Blais, "Engineering the quantum states of light in a Kerr-nonlinear resonator by two-photon driving", npj Quantum Information **3**, 18 (2017).

[58] A. Grimm, N. E. Frattini, S. Puri, S. O. Mundhada, S. Touzard, M. Mirrahimi, S. M. Girvin, S. Shankar, and M. H. Devoret, "Stabilization and operation of a Kerr-cat qubit", Nature **584**, 205 (2020).

[59] 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, et al., "Building a Fault-Tolerant Quantum Computer Using Concatenated Cat Codes", PRX Quantum **3**, 010329 (2022).

[60] U. Réglade, A. Bocquet, R. Gautier, J. Cohen, A. Marquet, E. Albertinale, N. Pankratova, M. Hallén, F. Rautschke, L.-A. Sellem, P. Rouchon, A. Sarlette, M. Mirrahimi, P. Campagne-Ibarcq, R. Lescanne, et al., "Quantum control of a cat qubit with bit-flip times exceeding ten seconds", Nature **629**, 778 (2024).

[61] H. Putterman, K. Noh, C. T. Hann, G. S. MacCabe, S. Aghaeimeibodi, R. N. Patel, M. Lee, W. M. Jones, H. Moradinejad, R. Rodriguez, N. Mahuli, J. Rose, J. C. Owens, H. Levine, E. Rosenfeld, et al., "Hardware-efficient quantum error correction via concatenated bosonic qubits", Nature **638**, 927 (2025).

[62] D. Gottesman, A. Kitaev, and J. Preskill, "Encoding a qubit in an oscillator", Physical Review A **64**, 012310 (2001).

[63] B. Royer, S. Singh, and S. M. Girvin, "Stabilization of Finite-Energy Gottesman-Kitaev-Preskill States", Physical Review Letters **125**, 260509 (2020).

[64] A. L. Grimsmo and S. Puri, "Quantum Error Correction with the Gottesman-Kitaev-Preskill Code", PRX Quantum **2**, 020101 (2021).

[65] B. L. Brock, S. Singh, A. Eickbusch, V. V. Sivak, A. Z. Ding, L. Frunzio, S. M. Girvin, and M. H. Devoret, "Quantum error correction of qudits beyond break-even", Nature **641**, 612 (2025).

[66] A. L. Grimsmo, J. Combes, and B. Q. Baragiola, "Quantum Computing with Rotation-Symmetric Bosonic Codes", Physical Review X **10**, 011058 (2020).

[67] 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, "Performance and structure of single-mode bosonic codes", Physical Review A **97**, 032346 (2018).

[68] J. Guillaud and M. Mirrahimi, "Repetition Cat Qubits for Fault-Tolerant Quantum Computation", Physical Review X **9**, 041053 (2019).

[69] S. Puri, L. St-Jean, J. A. Gross, A. Grimm, N. E. Frattini, P. S. Iyer, A. Krishna, S. Touzard, L. Jiang, A. Blais, S. T. Flammia, and S. M. Girvin, "Bias-preserving gates with stabilized cat qubits", Science Advances **6**, eaay5901 (2020).

[70] V. V. Albert, S. O. Mundhada, A. Grimm, S. Touzard, M. H. Devoret, and L. Jiang, "Pair-cat codes: autonomous error-correction with low-order nonlinearity", Quantum Science and Technology **4**, 035007 (2019).

[71] I. L. Chuang, D. W. Leung, and Y. Yamamoto, "Bosonic quantum codes for amplitude damping", Physical Review A **56**, 1114 (1997).

[72] M. H. Michael, M. Silveri, R. T. Brierley, V. V. Albert, J. Salmilehto, L. Jiang, and S. M. Girvin, "New Class of Quantum Error-Correcting Codes for a Bosonic Mode", Physical Review X **6**, 031006 (2016).

[73] M. Y. Niu, I. L. Chuang, and J. H. Shapiro, "Hardware-efficient bosonic quantum error-correcting codes based on symmetry operators", Physical Review A **97**, 032323 (2018).

[74] A. Kubica, A. Haim, Y. Vaknin, H. Levine, F. Brandão, and A. Retzker, "Erasure Qubits: Overcoming the T_1 Limit in Superconducting Circuits", Physical Review X **13**, 041022 (2023).

[75] J. D. Teoh, P. Winkel, H. K. Babla, B. J. Chapman, J. Claes, S. J. de Graaf, J. W. O. Garmon, W. D. Kalfus, Y. Lu, A. Maiti, K. Sahay, N. Thakur, T. Tsunoda, S. H. Xue, L. Frunzio, et al., "Dual-rail encoding with superconducting cavities", Proceedings of the National Academy of Sciences **120**, e2221736120 (2023).

[76] H. Levine, A. Haim, J. S. C. Hung, N. Alidoust, M. Kalaee, L. DeLorenzo, E. A. Wollack, P. Arrangoiz-Arriola, A. Khalajhedayati, R. Sanil, H. Moradinejad, Y. Vaknin, A. Kubica, D. Hover, S. Aghaeimeibodi, et al., "Demonstrating a Long-Coherence Dual-Rail Erasure Qubit Using Tunable Transmons", Physical Review X **14**, 011051 (2024).

[77] K. S. Chou, T. Shemma, H. McCarrick, T.-C. Chien, J. D. Teoh, P. Winkel, A. Anderson, J. Chen, J. C. Curtis, S. J. de Graaf, J. W. O. Garmon, B. Gudlewski, W. D. Kalfus, T. Keen, N. Khedkar, et al., "A superconducting dual-rail cavity qubit with erasure-detected logical measurements", Nature Physics **20**, 1454 (2024).

[78] S. J. de Graaf, S. H. Xue, B. J. Chapman, J. D. Teoh, T. Tsunoda, P. Winkel, J. W. O. Garmon, K. M. Chang, L. Frunzio, S. Puri, and R. J. Schoelkopf, "A mid-circuit erasure check on a dual-rail cavity qubit using the joint-photon number-splitting regime of circuit QED", npj Quantum Information **11**, 1 (2025).

[79] M. Grassl, T. Beth, and T. Pellizzari, "Codes for the quantum erasure channel", Physical Review A **56**, 33 (1997).

[80] N. Mehta, J. D. Teoh, T. Noh, A. Agrawal, A. Anderson, B. Birdsall, A. Brahmbhatt, W. Byrd, M. Cacioppo, A. Cabrera, L. Carroll, J. Chen, T.-C. Chien, R. Chamberlain, J. C. Curtis, et al., *Bias-preserving and error-detectable entangling operations in a superconducting dual-rail system*, arXiv:2503.10935 [quant-ph].

[81] I. Cong, H. Levine, A. Keesling, D. Bluvstein, S.-T. Wang, and M. D. Lukin, "Hardware-efficient, fault-tolerant quantum computation with rydberg atoms", Physical Review X **12**, 021049 (2022).

[82] Y. Wu, S. Kolkowitz, S. Puri, and J. D. Thompson, "Erasure conversion for fault-tolerant quantum computing in alkaline earth Rydberg atom arrays", Nature Communications **13**, 4657 (2022).

[83] K. Sahay, J. Jin, J. Claes, J. D. Thompson, and S. Puri, "High-Threshold Codes for Neutral-Atom Qubits with Biased Erasure Errors", Physical Review X **13**, 041013 (2023).

[84] G. Baranes, M. Cain, J. P. B. Ataides, D. Bluvstein, J. Sinclair, V. Vuletic, H. Zhou, and M. D. Lukin, *Leveraging Atom Loss Errors in Fault Tolerant Quantum Algorithms*, arXiv:2502.20558 [quant-ph].

[85] S. Omanakuttan, V. Buchemmavari, J. A. Gross, I. H. Deutsch, and M. Marvian, "Fault-Tolerant Quantum Computation Using Large Spin-Cat Codes", PRX Quantum **5**, 020355 (2024).

[86] Z. Hou, H.-S. Zhong, Y. Tian, D. Dong, B. Qi, L. Li, Y. Wang, F. Nori, G.-Y. Xiang, C.-F. Li, and G.-C. Guo, "Full reconstruction of a 14-qubit state within four hours", New Journal of Physics **18**, 083036 (2016).

[87] F. Huszár and N. M. T. Houlsby, "Adaptive Bayesian quantum tomography", Physical Review A **85**, 052120 (2012).

[88] G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo, "Neural-network quantum state tomography", Nature Physics **14**, 447 (2018).

[89] Y. Quek, S. Fort, and H. K. Ng, "Adaptive quantum state tomography with neural networks", npj Quantum Information **7**, 105 (2021).

[90] D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, and J. Eisert, "Quantum State Tomography via Compressed Sensing", Physical Review Letters **105**, 150401 (2010).

[91] S. T. Flammia, D. Gross, Y.-K. Liu, and J. Eisert, "Quantum tomography via compressed sensing: error bounds, sample complexity and efficient estimators", New Journal of Physics **14**, 095022 (2012).

[92] D. Perez-Garcia, F. Verstraete, M. M. Wolf, and J. I. Cirac, *Matrix Product State Representations*, arXiv:quant-ph/0608197.

[93] M. Cramer, M. B. Plenio, S. T. Flammia, R. Somma, D. Gross, S. D. Bartlett, O. Landon-Cardinal, D. Poulin, and Y.-K. Liu, "Efficient quantum state tomography", Nature Communications **1**, 149 (2010).

[94] B. P. Lanyon, C. Maier, M. Holzäpfel, T. Baumgratz, C. Hempel, P. Jurcevic, I. Dhand, A. S. Buyskikh, A. J. Daley, M. Cramer, M. B. Plenio, R. Blatt, and C. F. Roos, "Efficient tomography of a quantum many-body system", Nature Physics **13**, 1158 (2017).

[95] T. Baumgratz, A. Nüßeler, M. Cramer, and M. B. Plenio, "A scalable maximum likelihood method for quantum state tomography", New Journal of Physics **15**, 125004 (2013).

[96] T. Baumgratz, D. Gross, M. Cramer, and M. B. Plenio, "Scalable Reconstruction of Density Matrices", Physical Review Letters **111**, 020401 (2013).

[97] K. He, M. Yuan, Y. Wong, S. Chakram, A. Seif, L. Jiang, and D. I. Schuster, "Efficient multimode Wigner tomography", Nature Communications **15**, 4138 (2024).

[98] M. P. da Silva, O. Landon-Cardinal, and D. Poulin, "Practical Characterization of Quantum Devices without Tomography", Physical Review Letters **107**, 210404 (2011).

[99] Y. Wong, M. Yuan, K. He, S. Chakram, A. Seif, D. I. Schuster, and L. Jiang, *Efficient quantum tomography of a polynomial subspace*, arXiv:2503.00255 [quant-ph].

[100] A. Erhard, H. Poulsen Nautrup, M. Meth, L. Postler, R. Stricker, M. Stadler, V. Negnevitsky, M. Ringbauer, P. Schindler, H. J. Briegel, R. Blatt, N. Friis, and T. Monz, "Entangling logical qubits with lattice surgery", Nature **589**, 220 (2021).

[101] K. Mayer, C. Ryan-Anderson, N. Brown, E. Durso-Sabina, C. H. Baldwin, D. Hayes, J. M. Dreiling, C. Foltz, J. P. Gaebler, T. M. Gatterman, J. A. Gerber, K. Gilmore, D. Gresh, N. Hewitt, C. V. Horst, et al., *Benchmarking logical three-qubit quantum Fourier transform encoded in the Steane code on a trapped-ion quantum computer*, arXiv:2404.08616 [quant-ph].

[102] P. Sales Rodriguez, J. M. Robinson, P. N. Jepsen, Z. He, C. Duckering, C. Zhao, K.-H. Wu, J. Campo, K. Bagnall, M. Kwon, T. Karolyshyn, P. Weinberg, M. Cain, S. J. Evered, A. A. Geim, et al., "Experimental demonstration of logical magic state distillation", Nature, 10.1038/s41586-025-09367-3 (2025).

[103] M. Tsang, R. Nair, and X.-M. Lu, "Quantum Theory of Superresolution for Two Incoherent Optical Point Sources", Physical Review X **6**, 031033 (2016).

[104] M. Yuan, A. Seif, A. Lingenfelter, D. I. Schuster, A. A. Clerk, and L. Jiang, "Universal control in bosonic systems with weak Kerr nonlinearities", Physical Review A **111**, 032606 (2025).

[105] M. Yuan, Q. Xu, and L. Jiang, "Construction of bias-preserving operations for pair-cat codes", Physical Review A **106**, 062422 (2022).

[106] A. Kruckenhauser, M. Yuan, H. Zheng, M. Mamaev, P. Zeng, X. Mao, Q. Xu, T. V. Zache, L. Jiang, R. van Bijnen, and P. Zoller, "Dark Spin-Cat States as Biased Qubits", Physical Review Letters **135**, 020601 (2025).

[107] S.-u. Lee, M. Yuan, S. Chen, K. Tsubouchi, and L. Jiang, *Efficient benchmarking of logical magic state*, arXiv:2505.09687 [quant-ph].

[108] A. Joshi, K. Noh, and Y. Y. Gao, "Quantum information processing with bosonic qubits in circuit QED", Quantum Science and Technology **6**, 033001 (2021).

[109] P. Kok, W. J. Munro, K. Nemoto, T. C. Ralph, J. P. Dowling, and G. J. Milburn, "Linear optical quantum computing with photonic qubits", Reviews of Modern Physics **79**, 135 (2007).

[110] S. Lloyd and S. L. Braunstein, "Quantum computation over continuous variables", Physical Review Letters **82**, 1784 (1999).

[111] A. M. Eriksson, T. Sépulcre, M. Kervinen, T. Hillmann, M. Kudra, S. Dupouy, Y. Lu, M. Khanahmadi, J. Yang, C. Castillo-Moreno, P. Delsing, and S. Gasparinetti, "Universal control of a bosonic mode via drive-activated native cubic interactions", Nature Communications **15**, 2512 (2024).

[112] S. Krastanov, V. V. Albert, C. Shen, C.-L. Zou, R. W. Heeres, B. Vlastakis, R. J. Schoelkopf, and L. Jiang, "Universal control of an oscillator with dispersive coupling to a qubit", Physical Review A **92**, 040303(R) (2015).

[113] S. Hacohen-Gourgy, L. S. Martin, E. Flurin, V. V. Ramasesh, K. B. Whaley, and I. Siddiqi, "Quantum dynamics of simultaneously measured non-commuting observables", Nature **538**, 491 (2016).

[114] A. A. Diringer, E. Blumenthal, A. Grinberg, L. Jiang, and S. Hacohen-Gourgy, "Conditional-NOT displacement: fast multioscillator control with a single qubit", Physical Review X **14**, 011055 (2024).

[115] A. Anferov, F. Wan, S. P. Harvey, J. Simon, and D. I. Schuster, "Millimeter-Wave Superconducting Qubit", PRX Quantum **6**, 020336 (2025).

[116] J. Hastrup and U. L. Andersen, "Protocol for generating optical Gottesman-Kitaev-Preskill states with cavity QED", Physical Review Letters **128**, 170503 (2022).

[117] R. Dahan, G. Baranes, A. Gorlach, R. Ruimy, N. Rivera, and I. Kaminer, "Creation of optical cat and GKP states using shaped free electrons", Physical Review X **13**, 031001 (2023).

[118] Z. Vernon and J. E. Sipe, "Spontaneous four-wave mixing in lossy microring resonators", Physical Review A **91**, 053802 (2015).

[119] H. Choi, M. Heuck, and D. Englund, "Self-similar nanocavity design with ultrasmall mode volume for single-photon nonlinearities", Physical Review Letters **118**, 223605 (2017).

[120] A. Anferov, A. Suleymanzade, A. Oriani, J. Simon, and D. I. Schuster, "Millimeter-wave four-wave mixing via kinetic inductance for quantum devices", Physical Review Applied **13**, 024056 (2020).

[121] M. Xu, R. Cheng, Y. Wu, G. Liu, and H. X. Tang, "Magnetic field-resilient quantum-limited parametric amplifier", PRX Quantum **4**, 010322 (2023).

[122] Y. Liu, J. Sinanan-Singh, M. T. Kearney, G. Mintzer, and I. L. Chuang, "Constructing qudits from infinite-dimensional oscillators by coupling to qubits", Physical Review A **104**, 032605 (2021).

[123] D. Ruiz, R. Gautier, J. Guillaud, and M. Mirrahimi, "Two-photon driven Kerr quantum oscillator with multiple spectral degeneracies", Physical Review A **107**, 042407 (2023).

[124] J. Venkatraman, R. G. Cortiñas, N. E. Frattini, X. Xiao, and M. H. Devoret, "A driven Kerr oscillator with two-fold degeneracies for qubit protection", Proceedings of the National Academy of Sciences **121**, e2311241121 (2024).

[125] A. W. Bruch, X. Liu, J. B. Surya, C.-L. Zou, and H. X. Tang, "On-chip $\chi^{(2)}$ microring optical parametric oscillator", Optica **6**, 1361 (2019).

[126] A. Kamal, A. Marblestone, and M. Devoret, "Signal-to-pump back action and self-oscillation in double-pump Josephson parametric amplifier", Physical Review B **79**, 184301 (2009).

[127] U. Boscain, M. Sigalotti, and D. Sugny, "Introduction to the pontryagin maximum principle for quantum optimal control", PRX Quantum **2**, 030203 (2021).

[128] S. Ashhab, F. Yoshihara, T. Fuse, N. Yamamoto, A. Lupascu, and K. Semba, "Speed limits for two-qubit gates with weakly anharmonic qubits", Physical Review A **105**, 042614 (2022).

[129] L. S. Pontryagin, *Mathematical Theory of Optimal Processes* (Routledge, 1987).

[130] J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin, G. Necula, A. Paszke, J. VanderPlas, S. Wanderman-Milne, and Q. Zhang, *JAX: composable transformations of Python+NumPy programs*, 2018.

[131] M. Suzuki, "General theory of fractal path integrals with applications to many-body theories and statistical physics", Journal of Mathematical Physics **32**, 400 (1991).

[132] D. J. Parker, M. Savytskyi, W. Vine, A. Laucht, T. Duty, A. Morello, A. L. Grimsmo, and J. J. Pla, "Degenerate parametric amplification via three-wave mixing using kinetic inductance", Physical Review Applied **17**, 034064 (2022).

[133] P. T. Grochowski, H. Pichler, C. A. Regal, and O. Romero-Isart, *Quantum control of continuous systems via nonharmonic potential modulation*, arXiv:`2311.16819 [quant-ph]`.

[134] V. Sitzmann, J. Martel, A. Bergman, D. Lindell, and G. Wetzstein, "Implicit neural representations with periodic activation functions", Advances in Neural Information Processing Systems **33**, 7462 (2020).

[135] P. Aliferis, F. Brito, D. P. DiVincenzo, J. Preskill, M. Steffen, and B. M. Terhal, "Fault-tolerant computing with biased-noise superconducting qubits: a case study", New Journal of Physics **11**, 013061 (2009).

[136] P. Aliferis and J. Preskill, "Fault-tolerant quantum computation against biased noise", Physical Review A **78**, 052331 (2008).

[137] M. Li, D. Miller, M. Newman, Y. Wu, and K. R. Brown, "2D compass codes", Physical Review X **9**, 021041 (2019).

[138] D. K. Tuckett, S. D. Bartlett, S. T. Flammia, and B. J. Brown, "Fault-tolerant thresholds for the surface code in excess of 5% under biased noise", Physical Review Letters **124**, 130501 (2020).

[139] J. P. Bonilla Ataides, D. K. Tuckett, S. D. Bartlett, S. T. Flammia, and B. J. Brown, "The XZZX surface code", Nature Communications **12**, 2172 (2021).

[140] Q. Xu, N. Mannucci, A. Seif, A. Kubica, S. T. Flammia, and L. Jiang, "Tailored XZZX codes for biased noise", Physical Review Research **5**, 013035 (2023).

[141] R. Lescanne, M. Villiers, T. Peronnin, A. Sarlette, M. Delbecq, B. Huard, T. Kontos, M. Mirrahimi, and Z. Leghtas, "Exponential suppression of bit-flips in a qubit encoded in an oscillator", Nature Physics **16**, 509 (2020).

[142] H. Putterman, J. Iverson, Q. Xu, L. Jiang, O. Painter, F. G. S. L. Brandão, and K. Noh, "Stabilizing a Bosonic Qubit Using Colored Dissipation", Physical Review Letters **128**, 110502 (2022).

[143] Q. Xu, J. K. Iverson, F. G. S. L. Brandão, and L. Jiang, "Engineering fast bias-preserving gates on stabilized cat qubits", Physical Review Research **4**, 013082 (2022).

[144] R. Gautier, A. Sarlette, and M. Mirrahimi, "Combined Dissipative and Hamiltonian Confinement of Cat Qubits", PRX Quantum **3**, 020339 (2022).

[145] A. S. Darmawan, B. J. Brown, A. L. Grimsmo, D. K. Tuckett, and S. Puri, "Practical Quantum Error Correction with the XZZX Code and Kerr-Cat Qubits", PRX Quantum **2**, 030345 (2021).

[146] G. S. Agarwal, "Nonclassical statistics of fields in pair coherent states", J. Opt. Soc. Am. B **5**, 1940 (1988).

[147] V. V. Albert, "Lindbladians with multiple steady states: theory and applications", Ph.D. thesis (Yale University, 2017).

[148] V. V. Albert and L. Jiang, "Symmetries and conserved quantities in Lindblad master equations", Physical Review A **89**, 022118 (2014).

[149] V. V. Albert, C. Shu, S. Krastanov, C. Shen, R.-B. Liu, Z.-B. Yang, R. J. Schoelkopf, M. Mirrahimi, M. H. Devoret, and L. Jiang, "Holonomic Quantum Control with Continuous Variable Systems", Physical Review Letters **116**, 140502 (2016).

[150] J. J. Sakurai and J. Napolitano, *Modern Quantum Mechanics*, 2nd ed. (Cambridge University Press, 2017).

[151] A. Piñeiro Orioli, J. K. Thompson, and A. M. Rey, "Emergent Dark States from Superradiant Dynamics in Multilevel Atoms in a Cavity", Physical Review X **12**, 011054 (2022).

[152] J. R. Morris and B. W. Shore, "Reduction of degenerate two-level excitation to independent two-state systems", Physical Review A **27**, 906 (1983).

[153] C. J. Foot, *Atomic Physics* (Oxford University Press, 2004).

[154] B. Buča and T. Prosen, "A note on symmetry reductions of the Lindblad equation: transport in constrained open spin chains", New Journal of Physics **14**, 073007 (2012).

[155] K. DeBry, N. Meister, A. V. Martinez, C. D. Bruzewicz, X. Shi, D. Reens, R. McConnell, I. L. Chuang, and J. Chiaverini, *Error correction of a logical qubit encoded in a single atomic ion*, arXiv:2503.13908 [quant-ph].

[156] Y. Li, Q. Mei, Q.-X. Jie, W. Cai, Y. Li, Z. Liu, Z.-J. Chen, Z. Xie, X. Cheng, X. Zhao, Z. Luo, M. Zhang, X.-B. Zou, C.-L. Zou, Y. Lin, et al., *Beating the break-even point with autonomous quantum error correction*, arXiv:2504.16746 [quant-ph].

[157] Y. A. Yang, W.-T. Luo, J.-L. Zhang, S.-Z. Wang, C.-L. Zou, T. Xia, and Z.-T. Lu, "Minute-scale Schrödinger-cat state of spin-5/2 atoms", Nature Photonics **19**, 89 (2025).

[158] S. Ma, A. P. Burgers, G. Liu, J. Wilson, B. Zhang, and J. D. Thompson, "Universal Gate Operations on Nuclear Spin Qubits in an Optical Tweezer Array of ^{171}Yb Atoms", Physical Review X **12**, 021028 (2022).

[159] 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, "Parallel Implementation of High-Fidelity Multiqubit Gates with Neutral Atoms", Physical Review Letters **123**, 170503 (2019).

[160] W. Huie, L. Li, N. Chen, X. Hu, Z. Jia, W. K. C. Sun, and J. P. Covey, "Repetitive Readout and Real-Time Control of Nuclear Spin Qubits in ^{171}Yb Atoms", PRX Quantum **4**, 030337 (2023).

[161] J. M. Radcliffe, "Some properties of coherent spin states", Journal of Physics A: General Physics **4**, 313 (1971).

[162] F. Reiter and A. S. Sørensen, "Effective operator formalism for open quantum systems", Physical Review A **85**, 032111 (2012).

[163] M. Marthaler and M. I. Dykman, "Quantum interference in the classically forbidden region: A parametric oscillator", Physical Review A **76**, 010102 (2007).

[164] Q. Su, R. G. Cortiñas, J. Venkatraman, and S. Puri, *Unraveling the switching dynamics in a quantum double-well potential*, arXiv:2501.00209 [quant-ph].

[165] J. L. Van Hemmen and A. Sütö, in *Quantum Tunneling of Magnetization — QTM '94*, edited by L. Gunther and B. Barbara (Springer Netherlands, 1995), pp. 19–57.

[166] J. L. van Hemmen and A. Sütő, "WKB for quantum spins", Physica A: Statistical Mechanics and its Applications **321**, 493 (2003).

[167] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics: Non-Relativistic Theory* (Pergamon, 1977).

[168] S. Puri, A. Grimm, P. Campagne-Ibarcq, A. Eickbusch, K. Noh, G. Roberts, L. Jiang, M. Mirrahimi, M. H. Devoret, and S. M. Girvin, "Stabilized Cat in a Driven Nonlinear Cavity: A Fault-Tolerant Error Syndrome Detector", Physical Review X **9**, 041009 (2019).

[169] S. Chakram, K. He, A. V. Dixit, A. E. Oriani, R. K. Naik, N. Leung, H. Kwon, W.-L. Ma, L. Jiang, and D. I. Schuster, "Multimode photon blockade", Nature Physics **18**, 879 (2022).

[170] S. Chakram, A. E. Oriani, R. K. Naik, A. V. Dixit, K. He, A. Agrawal, H. Kwon, and D. I. Schuster, "Seamless High-Q Microwave Cavities for Multimode Circuit Quantum Electrodynamics", Physical Review Letters **127**, 107701 (2021).

[171] K. E. Cahill and R. J. Glauber, "Density Operators and Quasiprobability Distributions", Physical Review **177**, 1882 (1969).

[172] K. E. Cahill and R. J. Glauber, "Ordered Expansions in Boson Amplitude Operators", Physical Review **177**, 1857 (1969).

[173] S. T. Flammia and Y.-K. Liu, "Direct Fidelity Estimation from Few Pauli Measurements", Physical Review Letters **106**, 230501 (2011).

[174] J. M. Gertler, S. van Geldern, S. Shirol, L. Jiang, and C. Wang, "Experimental Realization and Characterization of Stabilized Pair-Coherent States", PRX Quantum **4**, 020319 (2023).

[175] J. A. Smolin, J. M. Gambetta, and G. Smith, "Efficient Method for Computing the Maximum-Likelihood Quantum State from Measurements with Additive Gaussian Noise", Physical Review Letters **108**, 070502 (2012).

[176] C. Shen, R. W. Heeres, P. Reinhold, L. Jiang, Y.-K. Liu, R. J. Schoelkopf, and L. Jiang, "Optimized tomography of continuous variable systems using excitation counting", Physical Review A **94**, 052327 (2016).

[177] P. Reinhold, "Controlling Error-Correctable Bosonic Qubits", Ph.D. thesis (Yale University, 2019).

[178] S. Bravyi and A. Kitaev, "Universal quantum computation with ideal Clifford gates and noisy ancillas", Physical Review A **71**, 022316 (2005).

[179] A. M. Souza, J. Zhang, C. A. Ryan, and R. Laflamme, "Experimental magic state distillation for fault-tolerant quantum computing", Nature Communications **2**, 169 (2011).

[180] R. S. Gupta, N. Sundaresan, T. Alexander, C. J. Wood, S. T. Merkel, M. B. Healy, M. Hillenbrand, T. Jochym-O'Connor, J. R. Wootton, T. J. Yoder, A. W. Cross, M. Takita, and B. J. Brown, "Encoding a magic state with beyond break-even fidelity", Nature **625**, 259 (2024).

[181] C. Gidney and A. G. Fowler, "Efficient magic state factories with a catalyzed $|CCZ\rangle$ to $2\,|T\rangle$ transformation", Quantum **3**, 135 (2019).

[182] E. Magesan, J. M. Gambetta, and J. Emerson, "Scalable and Robust Randomized Benchmarking of Quantum Processes", Physical Review Letters **106**, 180504 (2011).

[183] A. Kruckenhauser, R. van Bijnen, T. V. Zache, M. Di Liberto, and P. Zoller, "High-dimensional SO(4)-symmetric Rydberg manifolds for quantum simulation", Quantum Science and Technology **8**, 015020 (2022).

[184] K. C. Smith, E. Crane, N. Wiebe, and S. M. Girvin, "Deterministic Constant-Depth Preparation of the AKLT State on a Quantum Processor Using Fusion Measurements", PRX Quantum **4**, 020315 (2023).

[185] H. Buhrman, M. Folkertsma, B. Loff, and N. M. P. Neumann, "State preparation by shallow circuits using feed forward", Quantum **8**, 1552 (2024).

[186] K. C. Smith, A. Khan, B. K. Clark, S. M. Girvin, and T.-C. Wei, "Constant-Depth Preparation of Matrix Product States with Adaptive Quantum Circuits", PRX Quantum **5**, 030344 (2024).

[187] T. Wagner, H. Kampermann, D. Bruß, and M. Kliesch, "Learning Logical Pauli Noise in Quantum Error Correction", Physical Review Letters **130**, 200601 (2023).

[188] S. Chen, Y. Liu, M. Otten, A. Seif, B. Fefferman, and L. Jiang, "The learnability of Pauli noise", Nature Communications **14**, 52 (2023).

[189] M. Suzuki, "Decomposition formulas of exponential operators and lie exponentials with some applications to quantum mechanics and statistical physics", Journal of Mathematical Physics **26**, 601 (1985).

[190] U. B. Hoff, B. M. Nielsen, and U. L. Andersen, "Integrated source of broadband quadrature squeezed light", Optics Express **23**, 12013 (2015).

[191] S. Mittal, G. Moille, K. Srinivasan, Y. K. Chembo, and M. Hafezi, "Topological frequency combs and nested temporal solitons", Nature Physics **17**, 1169 (2021).

[192] Z. Vernon, N. Quesada, M. Liscidini, B. Morrison, M. Menotti, K. Tan, and J. Sipe, "Scalable squeezed-light source for continuous-variable quantum sampling", Physical Review Applied **12**, 064024 (2019).

[193] J. Lebreuilly, K. Noh, C.-H. Wang, S. M. Girvin, and L. Jiang, *Autonomous quantum error correction and quantum computation*, arXiv:2103.05007 [quant-ph].

[194] L. Li, C.-L. Zou, V. V. Albert, S. Muralidharan, S. M. Girvin, and L. Jiang, "Cat codes with optimal decoherence suppression for a lossy bosonic channel", Physical Review Letters **119**, 030502 (2017).

[195] M. Mitzenmacher and E. Upfal, *Probability and Computing: Randomized Algorithms and Probabilistic Analysis* (Cambridge University Press, 2005).