

TOPICAL REVIEW • OPEN ACCESS

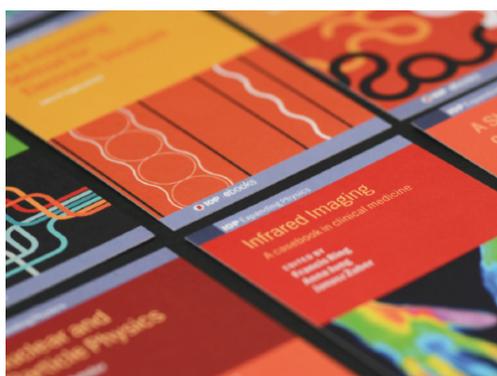
## Towards scalable bosonic quantum error correction

To cite this article: B M Terhal *et al* 2020 *Quantum Sci. Technol.* **5** 043001

View the [article online](#) for updates and enhancements.

### Recent citations

- [Noncommutative graphs based on finite-infinite system couplings: Quantum error correction for a qubit coupled to a coherent field](#)  
G. G. Amosov *et al*
- [Quantum information processing with bosonic qubits in circuit QED](#)  
Atharv Joshi *et al*
- [Bosonic quantum error correction codes in superconducting quantum circuits](#)  
Weizhou Cai *et al*



**IOP | ebooks™**

Bringing together innovative digital publishing with leading authors from the global scientific community.

Start exploring the collection—download the first chapter of every title for free.

# Quantum Science and Technology



## TOPICAL REVIEW

# Towards scalable bosonic quantum error correction

### OPEN ACCESS

RECEIVED  
2 March 2020

REVISED  
1 June 2020

ACCEPTED FOR PUBLICATION  
2 June 2020

PUBLISHED  
9 July 2020

Original content from  
this work may be used  
under the terms of the  
[Creative Commons  
Attribution 4.0 licence](#).

Any further distribution  
of this work must  
maintain attribution to  
the author(s) and the  
title of the work, journal  
citation and DOI.



B M Terhal<sup>1,2,6</sup> , J Conrad<sup>3,4,6</sup>  and C Vuillot<sup>5</sup> 

<sup>1</sup> QuTech, Delft University of Technology, P.O. Box 5046, 2600 GA Delft, The Netherlands

<sup>2</sup> JARA Institute for Quantum Information, Forschungszentrum Juelich, D-52425 Juelich, Germany

<sup>3</sup> Dahlem Center for Complex Quantum Systems, Physics Department, Freie Universität Berlin, Arnimallee 14, 14195 Berlin, Germany

<sup>4</sup> Helmholtz-Zentrum Berlin für Materialien und Energie, Hahn-Meitner-Platz 1, 14109 Berlin, Germany

<sup>5</sup> Inria Paris, 2 rue Simone Iff, 75012 Paris, France

<sup>6</sup> Author to whom any correspondence should be addressed.

E-mail: [bterhal@gmail.com](mailto:bterhal@gmail.com)

**Keywords:** quantum error correction, superconducting qubits, bosonic codes

## Abstract

We review some of the recent efforts in devising and engineering bosonic qubits for superconducting devices, with emphasis on the Gottesman–Kitaev–Preskill (GKP) qubit. We present some new results on decoding repeated GKP error correction using finitely-squeezed GKP ancilla qubits, exhibiting differences with previously studied stochastic error models. We discuss circuit-QED ways to realize CZ gates between GKP qubits and we discuss different scenarios for using GKP and regular qubits as building blocks in a scalable superconducting surface code architecture.

## 1. Introduction

There has been a recent surge in interest in bosonic error correction, both from the experimental as well as from the theoretical side. By bosonic quantum error correction we mean the representation of a qubit as a two-dimensional subspace of an oscillator, a means of performing some error correction on this qubit, as well as a suite of techniques to perform universal computation on the qubit.

We review some of these recent developments and older proposals, with an eye towards integration of the ideas into a scalable (code) architecture. To be concrete, we concentrate on superconducting devices as physical realizations, due to the excellent control and engineerability of strong non-linearities, as described by the formalism of circuit quantum electrodynamics (circuit-QED). For more background, we refer the reader to a recent review of circuit-QED [1] and also the realization of quantum error correction in circuit-QED [2].

Due to the commonality of the quantum optics language, some of our discussion applies more generally to other physical systems realizing oscillators, such as optical modes or mechanical oscillators. Our paper does not aim to be comprehensive in reviewing all possible bosonic codes, but rather seeks to identify some promising approaches and future work to be undertaken, in particular emphasizing scalable bosonic GKP error correction.

A first condition to even consider encoding a qubit into an oscillator is that a high- $Q$  oscillator is available<sup>6</sup>. Examples of such high- $Q$  oscillators are microwave cavity modes, of 3D or co-planar resonators in a frequency range  $f = 3\text{--}10$  GHz where single-photon life-times can be up to  $\tau = 1/\kappa = 1\text{--}10$  ms [3, 4]. Thus, without additional couplings and drives, the native noise model of such microwave cavities is simply photon loss, governed by the cavity decay rate  $\kappa$ .

In order to prepare and manipulate an encoded qubit as prescribed by some code, one induces additional errors which the chosen code should, ideally, also be able to correct. It is thus important to pick a code in which computational manipulations and error corrections are relatively simple and the chosen code can also handle the errors which occur in these processes. As is well known, no finite code can correct all errors, and hence the goal of bosonic quantum error correction is simply to provide a logical qubit which

<sup>6</sup> The  $Q$  of the oscillator captures the number of oscillations until it is fully damped and is given by  $Q = \frac{\omega}{\kappa}$  with  $\omega = 2\pi f$  the angular frequency of the oscillator and  $\kappa$  the oscillator decay rate.

can be used as a building block in a further coding scheme. The repetition or surface codes are the simplest examples of such further encoding steps, using then multiple oscillators.

Even though we can embrace the surface code as the simplest scalable 2D coding scheme [5], variants on who is playing the role of data and ancilla qubit and what additional error correction on these qubits takes place, are important in actually getting the very demanding engineering it, to pan out. If we have learned anything over the past 20 years of Hamiltonian engineering it is that partially-coherent dynamics can be implemented in many quantum systems, while very few to none may allow for the high-precision control and scalability needed for quantum error correction.

Overall, the challenges of efficiently using a bosonic qubit encoding are in (1) keeping the harmonicity of the oscillators as high as possible while temporally coupling to this mode, with high on/off ratio, to create and manipulate non-classical code states, (2) finding a photon number regime in which approximations of engineered Hamiltonians are accurate while the error correcting properties and benefits of the encoding are valid. When using bosonic qubits as basic qubits in a code architecture, it may be advantageous to choose data qubits differently than ancilla qubits and we will give some examples of such choices. The simplest encoding of a single qubit into a bosonic mode can be done using Fock states: the vacuum state represents the logical  $|0\rangle$ , denoted as  $|\bar{0}\rangle$ , and a single-photon state represents the logical  $|1\rangle$ , denoted as  $|\bar{1}\rangle$ .

For superconducting devices one can view the difference between bosonic encoding versus the regular transmon qubit encoding [6] as an interchange between the roles played by the anharmonic and the harmonic oscillator. Using transmon qubits to store information, resonators are used as couplers and for read-out. Using bosonic qubits to store information, anharmonic oscillators can be used for state preparation and couplers generating effective nonlinearities to realize gates. In this review we will refer to systems in which the lowest two energy eigenstates (in the absence of couplers) are used as *regular qubits*: this definition covers a Fock encoding as well as a transmon or a fluxonium qubit.

### 1.1. Preliminaries & notation

Here we collect a few definitions and mathematical identities that are used throughout the paper.

Additionally, useful textbooks for quantum optics and its mathematical description are [7–9]. We use  $\hat{q} = \frac{1}{\sqrt{2}}(a + a^\dagger)$  and  $\hat{p} = \frac{i}{\sqrt{2}}(a^\dagger - a)$ , where  $a$  ( $a^\dagger$ ) are annihilation (creation) operators, so that  $[\hat{q}, \hat{p}] = i\mathbb{1}$  and we sometimes refer to  $\hat{p}$  and  $\hat{q}$  as quadratures. A displacement in phase space is denoted as  $D(\alpha) = \exp(\alpha a^\dagger - \alpha^* a)$  and acts as  $D^\dagger(\alpha)aD(\alpha) = a + \alpha$ , while a coherent state is defined as  $D(\alpha)|0\rangle = |\alpha\rangle = \exp(-|\alpha|^2/2)\sum_{n=0}^{\infty} \frac{\alpha^n}{n!}|n\rangle$ . We have  $e^{i\theta a^\dagger} a e^{-i\theta a^\dagger} = a e^{-i\theta}$  so that  $e^{i\theta a^\dagger} D(\alpha) e^{-i\theta a^\dagger} = D(\alpha e^{i\theta})$ . The following identities hold

$$\exp(-iv\hat{q})|p\rangle = |p - v\rangle, \quad \exp(-iu\hat{p})|q\rangle = |q + u\rangle, \quad |p\rangle = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} dq e^{ipq}|q\rangle, \quad (1)$$

so that

$$\exp(iv\hat{q})\hat{p} \exp(-iv\hat{q}) = \hat{p} - v, \quad \exp(iu\hat{p})\hat{q} \exp(-iu\hat{p}) = \hat{q} + u. \quad (2)$$

A single-mode squeezing transformation is given by  $\exp(-iH_{\text{sq}}t) = \text{Sq}(\xi) = \exp(\frac{1}{2}(\xi^* a^2 - \xi a^{\dagger 2}))$  with Hamiltonian  $H_{\text{sq}} = \mathcal{E} a^{\dagger 2} + \mathcal{E}^* a^2$  with  $\xi = 2i\mathcal{E}t$ . The squeezer enacts the mode transformation  $a_{\text{out}} = \text{Sq}^\dagger(\xi)a\text{Sq}(\xi) = a \cosh(r) - a^\dagger e^{i\theta} \sinh(r)$  with  $r = |\xi|$  and  $\theta = \arg(\xi)$ .

In Lindblad equations we use the notation  $\mathcal{D}(A)(\rho) = A\rho A^\dagger - \frac{1}{2}\{A^\dagger A, \rho\}$  for some operator  $A$ .

It is standard to denote gates acting on a logical qubit subspace with overlines, i.e.  $\overline{\text{CNOT}}$  etc. In order to avoid notation clutter, only in section 2.3.2 we denote logical gates on the GKP codewords without it, i.e. CNOT and  $Z$  instead of  $\overline{\text{CNOT}}$  and  $\overline{Z}$ .

## 2. Bosonic qubits and their components

### 2.1. Early birds & cats and their generalizations

The first bosonic codes were formulated in [10] and designed to protect against photon loss. Of particular interest is a two-mode code with codewords

$$|\bar{0}\rangle = \frac{1}{\sqrt{2}}(|40\rangle + |04\rangle), \quad |\bar{1}\rangle = |22\rangle, \quad (3)$$

with  $|k\rangle$  denoting a Fock state with  $k$  photons. If either  $|\bar{0}\rangle$  or  $|\bar{1}\rangle$  (or both) were hit by the loss of a single photon on any one of the two modes, we can readily see that the resulting states would still be orthogonal. This orthogonality is a prerequisite for being able to correct the photon loss error, but it is not a sufficient condition. To examine the error correction capability of a (bosonic) code, one asks whether a set of

dominant errors satisfies the quantum error correction (QEC) conditions [11] of the code: if this holds (approximately) then there is an (approximate) recovery operation undoing these dominant errors. For a set of errors  $E = \{E_1, \dots, E_k\}$  acting on the encoding of a single qubit, the quantum error conditions are as follows.  $\forall i, j$ , we require

$$\langle \bar{0} | E_i^\dagger E_j | \bar{0} \rangle = \langle \bar{1} | E_i^\dagger E_j | \bar{1} \rangle \text{ logical states are indistinguishable,} \quad (4)$$

$$\langle \bar{0} | E_i^\dagger E_j | \bar{1} \rangle = \langle \bar{1} | E_i^\dagger E_j | \bar{0} \rangle = 0 \text{ orthogonality remains.} \quad (5)$$

To exemplify the use of these conditions, let us first look at the single-mode version of the code in equation (3):

$$|\bar{0}\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |4\rangle), \quad |\bar{1}\rangle = |2\rangle. \quad (6)$$

This code was introduced in [12] as the smallest member of a family of so-called binomial codes, hence its name kitten or ‘baby-binomial’ code. This code and its logical gates has been implemented using a superconducting microwave cavity mode as an oscillator in reference [13], but the life-time of the encoded qubit was comparable to that of a Fock state encoding. One can easily check that for the error set  $E = \{I, \sqrt{\gamma}a\}$ , the QEC conditions in equations (4) and (5) for this code are met. However, these errors are only an approximation of the real noise. A photon loss channel with photon decay rate  $\kappa$  lasting for time  $t$  with  $\gamma \equiv \kappa t \ll 1$ , can be modeled by a superoperator  $\mathcal{N}_\gamma$  with Kraus operators  $E_0 = I - \frac{\gamma}{2}a^\dagger a + O(\gamma^2) \approx e^{-\gamma a^\dagger a/2}$  and  $E_1 = \sqrt{\gamma}a$ , or

$$\mathcal{N}_\gamma(\rho) = E_0 \rho E_0^\dagger + E_1 \rho E_1^\dagger, \quad E_0^\dagger E_0 + E_1^\dagger E_1 = I + O(\gamma^2). \quad (7)$$

For the Kraus operators  $E_0$  and  $E_1$  the QEC conditions in equations (4) and (5) are not quite met. In particular, we have

$$\langle \bar{0} | E_0^\dagger E_0 | \bar{0} \rangle - \langle \bar{1} | E_0^\dagger E_0 | \bar{1} \rangle = O(\gamma^2)(\langle \bar{0} | (a^\dagger a)^2 | \bar{0} \rangle - \langle \bar{1} | (a^\dagger a)^2 | \bar{1} \rangle) \neq 0, \quad (8)$$

as  $|\bar{1}\rangle$  is an eigenstate of  $a^\dagger a$ , while  $|\bar{0}\rangle$  is not. This means that upon the detection of no photon loss (corresponding to  $E_0$ ) the code states undergo an irreversible distortion. The two-mode version of this code, equation (3), improves on this distortion issue as the quantum error correction conditions for the two mode code are met for the error set  $E = \{\sqrt{\gamma}a, \sqrt{\gamma}b, \exp(-\frac{\gamma}{2}(\hat{n}_a + \hat{n}_b)) \approx I - \frac{\gamma}{2}(\hat{n}_a + \hat{n}_b)\}$ . These error operators can be viewed as the three Kraus operators of a process in which there is either photon loss on mode  $a$ , photon loss on mode  $b$ , or no photon loss on either modes. For the states in equation (3) we have no distortion upon not detecting a photon from either modes as  $|\bar{0}\rangle$  and  $|\bar{1}\rangle$  are both eigenstates of  $\exp(-\frac{\gamma}{2}(\hat{n}_a + \hat{n}_b))$  with eigenvalue  $\exp(-2\gamma)$ . As far as we know, this two-mode code is still awaiting experimental realization.

By allowing ourselves code states with higher average photon number, we can correct for more loss errors, as well as gain and dephasing errors. More precisely, reference [12] has introduced families of binomial and so-called cat codes which correct against the set of errors

$E = \{I, a, \dots, a^L, a^\dagger, \dots, a^{\dagger G}, a^\dagger a, \dots, (a^\dagger a)^D\}$  for arbitrary  $L, G$  and  $D$ . For example, the idea behind the binomial codes can be understood as follows. Using the Holstein–Primakoff transformation  $a^\dagger a \rightarrow J_z + J$ , the binomial code words  $|\bar{0}\rangle$  and  $|\bar{1}\rangle$  can be seen as spin-eigenstates of  $J_x = \pm J$  with  $2J = N + 1$ , where one defines  $N = \max(L, G, 2D)$ . Dephasing errors  $(a^\dagger a)^m$ ,  $m = 1, \dots, D$  thus lead to a change in  $J_x$  by at most  $D \leq [J]$ , hence keeping codewords orthogonal. At the same time, protection against photon loss and gain is achieved by using a subspace of sufficiently separated Fock states stabilized by the operator

$\Pi_S = e^{i2\pi a^\dagger a/(S+1)}$  with  $S = L + G$ . For  $S = 1$  this gives the photon parity operator  $\Pi_{S=1} \equiv \Pi_{\text{photon}} = e^{i\pi a^\dagger a}$ : the even-photon codewords in equation (6) are clearly  $+1$  eigenstates of this photon parity operator.

Another family of single-mode codes are the cat codes. A very simple encoding is  $|\bar{0}\rangle \approx |\alpha\rangle$  and  $|\bar{1}\rangle \approx |-\alpha\rangle$  with coherent state  $|\alpha\rangle$ , first proposed in [14, 15]. Since  $|\alpha\rangle$  and  $|-\alpha\rangle$  are not orthogonal, it is more appropriate to define the code states as  $|C_\alpha^\pm\rangle = \frac{1}{\sqrt{N_\pm}}(|\alpha\rangle \pm |-\alpha\rangle)$  with  $N_\pm = 2(1 \pm \exp(-2|\alpha|^2))$ . These states are orthogonal for all  $\alpha$  and we can define  $|\pm\rangle \equiv |C_\alpha^\pm\rangle$ . On this encoding, photon loss induces immediate phase-flip errors since  $a|C_\alpha^\pm\rangle \propto |C_\alpha^\mp\rangle$ . Thus the phase-flip error rate (probability per unit time) is proportional to  $\kappa|\alpha|^2$  with  $\kappa$  the photon loss rate of the encoding mode.

On the other hand, for large enough  $\alpha$ , bit-flips,  $\alpha \leftrightarrow -\alpha$ , can be expected to occur at a much lower error rate as they correspond to a large change of the state in phase space. Particularly interesting is the engineering of Hamiltonians or dissipative processes which have these code states  $|\pm\alpha\rangle$  as degenerate fixed-points, so that there is a ‘macroscopic’ energy barrier to transition between them, leading to a bit flip rate exponentially small in  $|\alpha|^2$ . This design can lead to a qubit for which the noise is biased as phase-flip

errors are more prominent than bit-flip errors. We will discuss this noise-biased qubit in more detail in section 2.2.

The next-level cat encoding was introduced in [15, 16], and is sometimes referred to as the four-legged cat code since its codewords have four blobs in phase space:

$$\begin{aligned} |\bar{0}\rangle &= \frac{1}{\sqrt{N_0}} (|\alpha\rangle + |-\alpha\rangle + |i\alpha\rangle + |-i\alpha\rangle), & |\bar{1}\rangle &= \frac{1}{\sqrt{N_1}} (|\alpha\rangle + |-\alpha\rangle - |i\alpha\rangle - |-i\alpha\rangle), \\ N_b &= 8 e^{-\alpha^2} (\cosh \alpha^2 + (-1)^b \cos \alpha^2), & \alpha &\in \mathbb{R}. \end{aligned} \quad (9)$$

Using the standard identity  $\langle\alpha|\beta\rangle = e^{-(|\alpha|^2+|\beta|^2)/2} e^{\alpha^*\beta}$  one can verify the orthogonality of these two states. As for the kitten code, we can verify that both states are +1 eigenstates of the photon parity operator  $\Pi_{\text{photon}} = e^{i\pi a^\dagger a}$  using that  $\Pi_{\text{photon}}|\alpha\rangle = |-\alpha\rangle$ . The photon parity thus functions as a check operator, taking eigenvalue +1 on the code space and measuring it is a natural way to detect photon loss and perform error correction.

The states  $|\bar{0}\rangle$  and  $|\bar{1}\rangle$ , both having even photon parity, are however distinguished by their photon parity modulo 4, expressed as the  $\pm 1$  eigenvalues of the operator  $\Pi_{\text{photon}}^{1/2} = \exp(i\pi a^\dagger a/2)$ . To measure the photon parity operator via an ancilla qubit, a cavity mode-qubit dispersive interaction  $-\chi Z a^\dagger a/2$  can be used [7]. In circuit-QED the interaction  $Z a^\dagger a$  comes about naturally as the effective interaction between, say, a cavity mode and a linearly-coupled, off-resonant, transmon qubit mode [6]. The measurement of  $\Pi_{\text{photon}}$  then proceeds by preparing the ancilla qubit in  $|+\rangle$ , letting the interaction take place for time  $t = \pi/\chi$  and subsequently measuring the qubit in the  $|\pm\rangle$  basis. Using a transmon qubit and cavity mode, reference [17] has shown that tracking the photon parity by repeated measurements of  $\Pi_{\text{photon}}$  makes for a logical qubit which has a longer life-time than a Fock qubit without error correction in the same cavity mode. This result has essentially been the first demonstration of quantum error correction lengthening the life-time as compared to that of native qubits (transmon and/or Fock encoding) in the hardware.

Before we discuss further generalizations, let us examine the quantum error correction conditions, equations (4) and (5) for this cat code with respect to the set of errors  $\mathbf{E} = \{I, \sqrt{\gamma}a\}$ . One can quickly observe that all conditions are obeyed except  $\langle\bar{0}|a^\dagger a|\bar{0}\rangle \stackrel{?}{=} \langle\bar{1}|a^\dagger a|\bar{1}\rangle$ . Besides the uninteresting case of taking  $\alpha$  very large (so that all  $|\pm\alpha\rangle, |\pm i\alpha\rangle$  are orthogonal), this last condition is *exactly* met at sweet spots given by the equation  $\tan \alpha^2 = -\tanh \alpha^2$ . The smallest sweet-spot at  $|\alpha|^2 = 2.34$  lies close to the number of photons  $\bar{n} = 2$  of the cat code used in the experiment [17].

There are several error channels which impact the performance of the cat code using repeated photon parity measurements. First of all, the code cannot fully correct against the photon loss channel as it cannot correct the distortion Kraus operator  $E_0 = I - \gamma a^\dagger a/2$ . Secondly, two photon-loss events  $\propto a^2$  implement a logical bit-flip  $|\bar{0}\rangle \leftrightarrow |\bar{1}\rangle$ . Thirdly, photon loss in combination with the inevitable Kerr nonlinearity  $\sim (a^\dagger a)^2$  on the cavity mode causes incorrecable dephasing: the Kerr interaction makes the cavity rotation speed depend on the number of photons in the cavity, but this number becomes indeterminate in the presence of photon loss. Last but not least, transmon qubit decay during the qubit controlled- $a^\dagger a$  interaction, is a serious source of *feedback error*. For example, when the qubit decays half-way through the interaction,  $|1\rangle \rightarrow |0\rangle$ , it applies only half the rotation on the cavity mode. The result is that the eigenvalues of  $\Pi_{\text{photon}}^{1/2} = \bar{Z}$  are measured via the qubit measurement, collapsing the logical state.

This last feedback error problem is an important issue for any bosonic qubit, and it has been a central theme in the theory of fault-tolerant computing in general [18]. A disadvantage of the theoretical schemes for fault-tolerant quantum error correction is that they typically require additional hardware resources, such as logical ancilla qubits or (verified) multi-qubit GHZ states. Instead, we may seek hardware-efficient mitigation of the feedback error problem. As an example, reference [19] has addressed the feedback error due to transmon relaxation by drive-engineering the dispersive coupling Hamiltonian to equal  $-\chi(|2\rangle\langle 2| + |1\rangle\langle 1|)a^\dagger a$  and starting the ancilla transmon qubit in the state  $\frac{1}{\sqrt{2}}(|0\rangle + |2\rangle)$ . Transmon qubit decay from  $2 \rightarrow 1$  then commutes with the transmon-cavity interaction and does not cause errors on the cavity mode. The decay does—as in the normal case—affect the reliability of the transmon qubit measurement outcome. All-in all, this has led to an overall factor 5 in improvement of the life-time of the encoded cat qubit [19].

Another way of minimizing feedback errors on a bosonic code is to use a biased-noise ancilla qubit (section 2.2) as an ancilla qubit. As proposed in reference [20], the goal is then to let the strong-noise error channel affect the ancilla qubit measurement, while the low-noise (bit-flip) channel on the ancilla feeds back low-noise to the bosonic code.

Single-mode cat codes with higher-photon numbers can be formulated and form a class of codes [12, 21]. Reference [22] studied the performance of binomial, cat and GKP codes (see section 2.3), against

photon loss, assuming optimal noise-free recovery as permitted by the quantum error correcting conditions. Reference [23] has formulated a general framework of rotation-symmetric codes of which the binomial and cat codes are subclasses: the unifying theme is rotation symmetry of the code states in phase space captured by invariance under the operator  $\Pi_S$ . Another interesting class of bosonic codes uses a three-wave mixing  $\chi^{(2)}$ -interaction, equation (32), as the central element for defining the code and correcting photon loss [24]. Various classes of multi-mode codes against photon loss exist, see for example [22, 25, 26] and references therein.

A challenge in using a bosonic qubit is that some computational manipulations can be more involved than for a regular qubit. For example, on a regular qubit such as a transmon qubit, rotations by an angle  $\theta$  around axes  $X$  or  $Y$  are easily accomplished by temporarily supplying microwave radiation. On a bosonic qubit, these simple single-qubit gates can be non-trivial. An advocated solution in reference [27] is to always use a dual-rail (dr) encoding of a bosonic qubit with  $|\bar{0}\rangle_{\text{dr}} = |\bar{01}\rangle$  and  $|\bar{1}\rangle_{\text{dr}} = |\bar{10}\rangle$ , where  $|\bar{0}\rangle$  and  $|\bar{1}\rangle$  are the states of (an arbitrary) bosonic qubit itself. Having mapped the Bloch-sphere of a qubit onto a two-mode state space, the exponential mode-SWAP operator  $\exp(i\theta\text{SWAP}_{a,b})$  becomes a universal gate to do single-qubit and two-qubit gates, and has been realized in [28]. Here the linear SWAP $_{a,b}$  transformation interchanges two modes  $a$  and  $b$ , i.e. its action on quadrature operators for the modes is given by  $q_a \leftrightarrow q_b$  and  $p_a \leftrightarrow p_b$ . If we envision using a bosonic qubit as a building block qubit in a stabilizer code, it is however not necessary to perform any gate, but rather we can focus on performing CNOT or CZ, Hadamard ( $H$ ) and  $T$  gates possibly using ancilla qubits, see e.g. reference [5].

## 2.2. Noise-biased cat qubit

A method to set up a dissipative process which stabilizes the coherent states  $|\pm\alpha\rangle$  was devised in [15]. The idea is to engineer the Lindblad equation (in a frame rotating at the mode frequency):

$$\dot{\rho} = -i[H_{\text{sq}}, \rho] + \kappa_{2ph}\mathcal{D}(a^2)(\rho) \equiv \mathcal{L}(\rho), \quad (10)$$

with  $H_{\text{sq}} = \mathcal{E}a^{\dagger 2} + \mathcal{E}^*a^2$  where  $\mathcal{E} = i|\mathcal{E}|$  with  $|\mathcal{E}|$  proportional to the strength of a pumped microwave mode acting as a classical field. To understand the fixed points of this evolution— $\rho$  for which  $\mathcal{L}(\rho) = 0$ —we can write the Lindblad equation as

$$\dot{\rho} = -i\left(H_{\text{eff}}\rho - \rho H_{\text{eff}}^{\dagger}\right) + \kappa_{2ph}a^2(\rho)a^{\dagger 2}, \quad (11)$$

with  $H_{\text{eff}} = H_{\text{sq}} - \frac{i\kappa_{2ph}}{2}a^{\dagger 2}a^2$ . We can then use, with  $K = \frac{\kappa_{2ph}}{2}$ :

$$\begin{aligned} -iH_{\text{eff}} &= -Ka^{\dagger 2}a^2 + |\mathcal{E}|\left(a^{\dagger 2} - a^2\right) = -K\tilde{M}_{\alpha}^{\dagger}M_{\alpha} - \frac{|\mathcal{E}|^2}{K}, \\ \text{with } M_{\alpha} &= a^2 - \alpha^2I, \quad \tilde{M}_{\alpha} = a^2 + \alpha^2I, \quad \alpha = \sqrt{\frac{|\mathcal{E}|}{K}}. \end{aligned} \quad (12)$$

This immediately implies that the states  $|\pm\alpha\rangle = \pm\sqrt{|\mathcal{E}|/K}$  are fixed points of the Lindblad evolution, as  $M_{\alpha}|\pm\alpha\rangle = 0$ , and the last term in equation (11) is canceled by the constant  $\frac{-|\mathcal{E}|^2}{K}$  which remains from the first term. Hence, any linear combination of the states  $|\alpha\rangle = \pm\sqrt{|\mathcal{E}|/K}$  is a fixed point of the dynamics.

When the pump inducing the squeezing Hamiltonian  $H_{\text{sq}}$  is off,  $\mathcal{E} = 0$ , we can observe that the Fock states  $|0\rangle = \lim_{\alpha \rightarrow 0}|C_{\alpha}^{+}\rangle$  and  $|1\rangle = \lim_{\alpha \rightarrow 0}|C_{\alpha}^{-}\rangle$  are fixed points, distinguished by their photon parity. Thus when  $\mathcal{E}$  is gradually increased, we can smoothly change from a Fock encoding into the cat  $|C_{\alpha}^{\pm}\rangle$  encoding. Photon loss at rate  $\kappa$ , which can be modeled by introducing an additional term  $\kappa\mathcal{D}(a)(\rho)$  in equation (10), causes phase-flip errors, i.e. flipping between the states  $|C_{\alpha}^{\pm}\rangle$ , but does not interfere with the stabilization itself as  $|\pm\alpha\rangle$  are eigenstates of  $a$  so that  $\mathcal{D}(a)(|\pm\alpha\rangle\langle\pm\alpha|) = 0$ . One can add a drive term  $H_{\text{drive}} = \epsilon(t)a^{\dagger} + \epsilon^*(t)a$  to the Lindblad equation and observe that the annihilation operator  $a$  will generate rotations around the  $Z$ -axis (periodically interchanging  $|C_{\alpha}^{\pm}\rangle$ ). At the same time,  $a^{\dagger}$  in principle leads to a departure from the qubit subspace spanned by  $|\pm\alpha\rangle$  corresponding to leakage. However, due the  $\sim |\alpha|^2$  gap of the Lindbladian, such departure from the eigenvalue-0 manifold is exponentially suppressed and the effect of the driving term can be analyzed by projecting it onto the stabilized subspace. In this subspace it then induces Rabi oscillations around an axis which is exponentially-closely aligned with the  $Z$ -axis, with Rabi frequency  $\Omega \propto |\epsilon||\alpha|$ , experimentally demonstrated in [29].<sup>7</sup> A measurement in the  $X$ -basis can be accomplished by measuring the photon parity through a coupled transmon qubit. The (pumped) squeezing

<sup>7</sup> In references [15, 30] and some other papers a different convention is used, namely  $|C_{\alpha}^{+}\rangle$  (resp.  $|C_{\alpha}^{-}\rangle$ ) is the  $Z$  eigenstate  $|0\rangle$  (resp.  $|1\rangle$ ), thus interchanging what is called  $Z$  and  $X$  here.

interaction and the required two-photon dissipative process have first been experimentally realized in [31]. This was achieved by coupling a 3D storage cavity (at frequency  $\omega_a$ ) via a bridging transmon to a lossy cavity (at different frequency  $\omega_b$ ) and applying a two-tone drive on the lossy cavity so as to set up a process to convert two storage photons to one lossy cavity photon which is subsequently lost (the  $\kappa_{2ph}\mathcal{D}(a^2)$  process). The lossy cavity is driven at pump frequency  $\omega_p = 2\omega_a - \omega_b$  as well as close to its own frequency  $\omega_b$ , generating, through the transmon nonlinearity, an effective degenerate parametric oscillator with resonant terms of the form  $a^{\dagger 2}b + b^{\dagger}a^2$ .

A more recent experimental realization in reference [32] has been able to cleanly generate the desired interactions (via an effective three-wave mixing, see also section 3.3) and observe the exponential decrease of the bit-flip error rate in  $|\alpha|^2$  as well as the linear increase of the phase-flip error rate with  $|\alpha|^2$ .

An alternative, non-dissipative, route towards a noise-biased qubit was first proposed in [33]. Instead of invoking dissipation, the idea is to engineer a Hamiltonian which has  $|\pm\alpha\rangle$  as degenerate eigenstates, using a Kerr nonlinearity and squeezing. The two-photon dissipation is then considered an optional add-on which helps in mitigating leakage, i.e. a departure from the subspace spanned by  $|\pm\alpha\rangle$ . The target Hamiltonian (in the rotating frame of the cavity mode) is

$$H = -Ka^{\dagger 2}a^2 + \mathcal{E}a^{\dagger 2} + \mathcal{E}^*a^2 = -KM_{\alpha}^{\dagger}M_{\alpha} + \frac{|\mathcal{E}|^2}{K},$$

$$M_{\alpha} = a^2 - \alpha^2, \quad \alpha = \sqrt{\frac{|\mathcal{E}|}{K}} e^{i\varphi}, \quad \mathcal{E} = |\mathcal{E}| e^{2i\varphi}. \quad (13)$$

The spectrum of  $H$  has eigenvalues running from  $\frac{|\mathcal{E}|^2}{K}$  downwards as the first term in  $H$  is negative-semi-definite. Omitting the factor  $\frac{|\mathcal{E}|^2}{K}$ , the highest eigenstates are the states  $|\pm\alpha\rangle$  with degenerate zero eigenvalues. We can observe the similarity and difference with equation (12): here we consider a Hermitian matrix and the phase of the pump amplitude  $\mathcal{E}$  is variable and determines the phase of the coherent states which are the zero energy eigenstates. Thus, by adiabatically changing the phase of  $\mathcal{E}$  we move to different zero energy eigenstates, allowing us to transform  $\alpha \rightarrow -\alpha$  and hence realize a  $X$  gate on  $|C_{\alpha}^{\pm}\rangle$ . For the stability of the encoded space it is important to understand the spectrum of  $H$  and the gap below these degenerate zero eigenstates, see the analysis in [20, 33]. To understand this, assume that the phase  $\varphi = 0$  for simplicity. We can displace the Hamiltonian by  $D(\pm\alpha)$  with  $\alpha = \sqrt{|\mathcal{E}|/K}$ . For large  $|\mathcal{E}|/K$ , one can approximate  $D^{\dagger}(\pm\alpha)HD(\pm\alpha) \approx -4K|\alpha|^2 a^{\dagger}a$ , a harmonic oscillator Hamiltonian. This shows that for large  $\alpha$ , the spectrum approximately has the gap  $4K|\alpha|^2$  and the first excited states below  $|\pm\alpha\rangle$  are roughly equal to  $D(\pm\alpha)|1\rangle$ . The so-called ‘Cassianian’ Hamiltonian in equation (13) was first studied in [34]: the surfaces of constant classical energy are described by Cassianian ovals in  $\langle p \rangle$  and  $\langle q \rangle$  with the foci of the ovals at  $\langle q \rangle = \pm\sqrt{|\mathcal{E}|/K}$ . As a quantum system the spectrum is that of an inverted double-well (‘double-oscillator’) with the well maxima at zero energy for the states  $|\pm\alpha\rangle$ . We can consider the effect of driving and several dissipative processes for the Hamiltonian in equation (13). For example, when one includes photon loss  $\kappa\mathcal{D}(\rho)$  in the Lindblad equation and the pump amplitude is sufficiently large, i.e.  $16|\mathcal{E}|^2 > \kappa^2$  [33, 35, 36], the fixed point of the Lindblad equation is the state  $p|\tilde{\alpha}\rangle\langle\tilde{\alpha}| + (1-p)|-\tilde{\alpha}\rangle\langle-\tilde{\alpha}|$  with modified  $\tilde{\alpha}$ ,  $|\tilde{\alpha}|^2 < \frac{|\mathcal{E}|}{K}$ . In this regime the system neatly represent the dissipative storage of a classical bit.

The effect of other sources of noise such as dephasing ( $\kappa_{\text{deph}}\mathcal{D}(a^{\dagger}a)(\rho)$ , see also equation (29)), photon gain ( $\kappa\bar{n}_{\text{therm}}\mathcal{D}(a^{\dagger})(\rho)$ ) due to the coupling with a finite temperature heat bath, as well coupling with baths with other spectral densities are discussed in detail in [15, 20, 33, 37].

Reference [30] has implemented the Kerr-cat Hamiltonian in equation (13) and the corresponding qubit in the resonant mode of a so-called SNAIL element (see section 3.3), coupled to a read-out cavity mode. The fourth-order nonlinearity of the SNAIL element gives the wanted  $-Ka^{\dagger 2}a^2$  term, while one can drive the mode at twice its frequency so as to use the third-order SNAIL term  $\propto a^{\dagger 3} + a^3$  to turn on squeezing. The experiment generated cat states with  $|\alpha|^2 \approx 2.5$  with a dephasing life-time of  $3\mu\text{s}$ , and an enhanced decay life-time of  $105\mu\text{s}$ , and a  $\pi/2$  rotation around the  $Z$ -axis obtained by driving took  $24\text{ ns}$ . The ability to convert the noise-biased qubit to a Fock encoding by turning off the squeezing drive allows to measure Pauli  $X$  via a standard dispersive measurement [30]. One can also measure a noise-biased qubit in the  $X$ -basis by dispersively coupling ( $-\chi a^{\dagger}aZ/2$ ) it to an ancilla qubit to map the photon parity onto the state of the ancilla qubit which is subsequently measured. To realize a (nondestructive) Pauli  $Z$  measurement, distinguishing  $\pm\alpha$ , reference [30] had applied, besides the squeezing drive, a drive at the difference frequency of the SNAIL mode and the read-out cavity mode ( $b$ ) to get a resonant beam-splitting interaction  $\propto a^{\dagger}b + ab^{\dagger}$ . The upshot is that the coherent states  $|\pm\alpha\rangle$  are mapped to corresponding coherent states in the cavity mode which are heterodyne-measured when leaking out of the cavity.

Given that the noise-biased cat qubit is designed to have a low bit-flip error rate, it can function as an ancilla control qubit in the error correction circuit for another code [20] inducing low feedback noise. Assume we have a code which is an eigenspace of a stabilizer  $S = e^{iA}$  and  $S$  is to be measured using the noise-biased cat qubit to detect or correct errors. This requires an interaction of the noise-biased cat qubit and the code of the form  $H_{\text{int}} \propto (a + a^\dagger) \otimes A$  since  $a + a^\dagger \approx Z$  on the noise-biased cat code space (besides some leakage), allowing for a qubit controlled- $S$  operation. For example, for the cat code,  $S = e^{i\pi b^\dagger b} = \Pi_{\text{photon}}$ , requiring a tunable photon–pressure coupling between the two modes of the form  $H_{\text{int}} \propto (a + a^\dagger)b^\dagger b$ . For the GKP code, see section 2.3,  $S$  is a displacement so that  $H_{\text{int}}$  can be chosen to be a tunable beam-splitting interaction of the form  $a^\dagger b + ab^\dagger$ .

It has been argued that, if the noise-bias of this qubit is sufficiently strong, only a classical repetition code [38] might suffice to correct for the dominant phase-flip ( $Z$ ) errors due to photon loss. Crucial in this idea is that the CNOT gate which is needed to measure the  $XX$  checks of this code preserves the noise-bias, that is,  $Z$  errors during the gate do not propagate to become  $X$  errors after the gate. For the Kerr-cat qubit a noise-bias preserving CNOT gate has been proposed in [37]. A similar idea is to use this Kerr-cat qubit as a basic qubit in a surface code architecture in which the  $XXXX$  and  $YYYY$  checks are measured [37, 39]. In this modified form of surface code one gains much more information about  $Z$  errors. It has been shown that when the probability for phase-flip errors and measurement errors is a factor 100 more than that of bit-flip errors within a phenomenological error model, the threshold against  $Z$  errors can be as high as 5% [39]. It is an open question whether such high bias will be feasible in practice as experiments for doing the CZ gate and the noise-bias preserving CNOT gate on these noise-biased qubits are still to come.

### 2.3. The GKP qubit

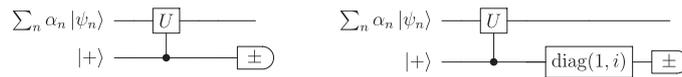
The (square) Gottesman–Kitaev–Preskill (GKP) qubit introduced in reference [40] is defined through two commuting displacement operators, acting as translations in phase space, i.e.  $S_q = \exp(i2\sqrt{\pi}\hat{q})$  and  $S_p = \exp(-i2\sqrt{\pi}\hat{p})$ .<sup>8</sup> The ideal GKP code is the space invariant under these two phase-space translations. As a result, any wave function in  $q$  (resp.  $p$ ) in this space has support on  $q = k\sqrt{\pi}$  (resp.  $p = l\sqrt{\pi}$ ) for integers  $k, l \in \mathbb{Z}$ . The logical operators of the qubit are  $\bar{Z} = \exp(i\sqrt{\pi}\hat{q})$  and  $\bar{X} = \exp(-i\sqrt{\pi}\hat{p})$  with  $\bar{X}\bar{Z} = -\bar{Z}\bar{X}$ . In addition,  $\bar{Y} = i\bar{X}\bar{Z} = \exp(i\pi/2) \exp(-i\sqrt{\pi}\hat{p}) \exp(i\sqrt{\pi}\hat{q}) = \exp(i\sqrt{\pi}(-\hat{p} + \hat{q}))$ . This choice makes the wave function in  $q$  of  $|\bar{0}\rangle$  a sum of delta functions at values of  $q$  which are even multiples of  $\sqrt{\pi}$ , while  $|\bar{1}\rangle$  has uniform support on values of  $q$  which are odd multiples of  $\sqrt{\pi}$ . The ideal code meets the quantum error correction conditions for a continuous set of ‘at most half-logical’ displacements  $\mathbf{E} = \{e^{iu\hat{p}}, e^{iv\hat{q}} : |u|, |v| \leq \sqrt{\pi}/2\}$ , since any products of these shifts maps a  $|\bar{0}\rangle$  onto a state orthogonal to both  $|\bar{1}\rangle$  and  $|\bar{0}\rangle$  (and vice-versa). The set of correctable displacements forms a square Wigner–Seitz or Voronoi cell (containing only one lattice point such that all points in the cell are closer to this point than to another lattice point) in the code lattice generated by the logical phase-space translations.

Naturally, an asymmetric version of the GKP code which corrects more shift errors in  $\hat{q}$  than shift errors in  $\hat{p}$  can also be defined. However, when there is no hardware-based noise asymmetry between  $\hat{p}$  and  $\hat{q}$  this does not seem immediately useful.

In principle, and in theory, to perform quantum error correction the eigenvalues (phases) of the unitary operators  $S_p$  and  $S_q$  are to be measured. Performing such measurements projects the continuum of errors onto (superpositions of) possible displacements, and we perform error correction by choosing a displacement of minimal amplitude which resets these eigenvalues to +1, corresponding to the code space. In section 2.4 we will analyze GKP quantum error correction using encoded GKP ancilla qubits, see figure 7. The advantage of this form of error correction is that it does not suffer from feedback errors induced by a poor ancilla qubit (instead, it suffers feedback errors from a GKP ancilla qubit) and the information gained through measuring the GKP ancilla states is analog rather than binary. The disadvantage is that one needs to prepare GKP ancilla states themselves first.

For this latter task one can perform some form of phase estimation to measure the eigenvalues of the unitary operators  $S_p$  and  $S_q$ . Since the eigenvalues take continuous values, one only ever realizes an approximate estimation of these phases. Phase estimation can readily be executed by coupling the GKP mode repeatedly to a single ancilla qubit via controlled-displacement gates as was proposed and discussed in great detail in reference [41], focusing on a circuit-QED implementation. The idea behind this is simple. To measure the eigenvalue of a unitary operator  $U$  such as the displacements  $S_p$  or  $S_q$ , one can use ancilla qubits applying qubit controlled- $U^k$  gates for  $k = 1, 2, \dots$ . For example, when  $k = 1$ , the circuit on the left in figure 1 has outcome probabilities  $\mathbb{P}(\pm) = \frac{1}{2} (1 \pm \langle \text{Re}(U) \rangle)$ , while the circuit on the right has probabilities  $\mathbb{P}(\pm) = \frac{1}{2} (1 \mp \langle \text{Im}(U) \rangle)$ .

<sup>8</sup> The commutation of  $S_p$  and  $S_q$  can be verified by using the identity  $\exp(A)\exp(B) = \exp(B)\exp(A)\exp([A, B])$  for operators  $A$  and  $B$  whose commutator is proportional to  $I$ .



**Figure 1.** Single round of phase estimation with  $U|\psi_n\rangle = e^{i\theta_n}|\psi_n\rangle$  where the probability for ancilla qubit to be measured in the state  $|\pm\rangle$  equals  $\mathbb{P}(\pm) = \sum_n |\alpha_n|^2 \frac{1}{2}(1 \pm \cos(\theta_n))$  (left) and  $\mathbb{P}(\pm) = \sum_n |\alpha_n|^2 \frac{1}{2}(1 \mp \sin(\theta_n))$  (right). In the applications here  $U$  is a displacement  $S_p$  or  $S_q$ .

In phase-estimation schemes, higher powers  $k > 1$  of  $U^k$  are often used, but applying  $U^k$ , a displacement of strength  $\sim k$ , increases the number of photons in the state by  $\sim k^2$  and does not provide a good approximation of an approximate GKP state [41]. Instead of repeating the phase estimation to collect bits of the phase and then do a final corrective displacement, it is experimentally simpler to opt for immediate feedback on the code state based on each new bit obtained in a round of phase estimation. This is the route taken in the experimental realization of the GKP code in [42], where a small conditional displacement on the GKP qubit is executed depending on the ancilla qubit measurement outcome. In fact, using such immediate feedback the state of the ancilla qubit does not even need to be measured, as the feedback can be done depending on the qubit state itself, followed by an approximate disentangling step [43] or alternatively a qubit reset step (to avoid entropy build-up).

In addition, in reference [42] only the right circuit in figure 1 measuring  $\text{Im}(U)$  is used (instead of measuring both  $\text{Re}(U)$  and  $\text{Im}(U)$ ). If the state to be measured is (approximately) symmetrically centered around the vacuum so that its wavefunction is symmetric under  $q \rightarrow -q$  and  $p \rightarrow -p$ , we have  $\int dp |\psi(p)|^2 \text{Im}(S_p) = 0$  and  $\int dq |\psi(q)|^2 \text{Im}(S_q) = 0$ . This implies that  $\mathbb{P}(\pm) = \frac{1}{2}(1 \pm \langle \text{Im}(U) \rangle) = \frac{1}{2}$ , suggesting that the measurement outcome  $\pm$  can gain a maximal amount of information by weakly projecting onto  $\sin(\theta) \gtrless 0$ , and subsequently shifting the state to the point  $\theta = 0$ . These feedback shifts are realized in [42] by small displacements. Note that if the input state has eigenvalue phase  $\theta$  close to 0, then  $\text{Re}(U)$  is close to 1, implying that not much is learned by doing the measurement with outcomes  $\mathbb{P}(\pm) = \frac{1}{2}(1 \pm \langle \text{Re}(U) \rangle)$ .

We remark that the length of the displacement of the logical  $\bar{Y}$  is  $\sqrt{2}$  larger than that of  $\bar{X}$  and  $\bar{Z}$ . This implies some asymmetry in error correction. Namely, if we correct by measuring  $S_p$  and  $S_q$ , shifts such as  $\exp(-iu\hat{p} + iv\hat{q})$  with  $u^2 \leq \pi/4$  and  $v^2 \leq \pi/4$  can be corrected which, as displacements, are a factor  $\sqrt{2}$  larger than correctable displacements in pure  $\hat{p}$  and  $\hat{q}$  directions. Given a noise model which is rotationally-symmetric in phase space, this does not seem to be an optimal choice. It also implies that logical  $Y$  eigenstates which can flip due to large displacements in pure  $\hat{p}$  and  $\hat{q}$  directions can have shorter lifetimes [42].

A ‘hexagonal’ GKP qubit has also been defined in [40] by choosing two phase-space lattice translations which are not orthogonal such that all three logical operators  $\bar{X}$ ,  $\bar{Y}$  and  $\bar{Z}$  have the same length as phase-space translation vectors. For this choice we take as stabilizers  $\exp(i\xi(\sqrt{3}\hat{q} - \hat{p})/2)$  and  $\exp(i\xi\hat{p})$  with  $\xi = 2\sqrt{2\pi/\sqrt{3}}$ , generating a hexagonal lattice in phase space. Again the logical operators are half-stabilizers, forming the vectors generating a hexagonal lattice. The correctable displacements now form a hexagonal Wigner–Seitz cell. This cell is larger in volume than the square Wigner–Seitz cell in the square GKP lattice. If we assume that displacement errors occur according to a stochastic Gaussian model as in equation (24), it implies that the hexagonal code can correct a larger probability volume of errors.

If we were to choose stabilizers  $S_q = \exp(i\sqrt{2\pi}\hat{q})$  and  $S_p = \exp(-i\sqrt{2\pi}\hat{p})$ , there would be no additional commuting displacement operators, implying that the  $+1$  eigenspace  $S_p$  and  $S_q$  is one-dimensional. This eigenstate, also called the sensor state  $|\psi_{\text{sensor}}\rangle$  in [44], is a uniform sum of delta function at  $q = k\sqrt{2\pi}$  with  $k \in \mathbb{Z}$  (and similarly a uniform sum of delta functions at  $p = l\sqrt{2\pi}$  with  $l \in \mathbb{Z}$ ). The sensor state is interesting in allowing one to simultaneously estimate the complex and real part of the amplitude  $\alpha$  of a displacement  $D(\alpha)$ , by performing phase-estimation for  $S_q$  and  $S_p$  on  $D(\alpha)|\psi_{\text{sensor}}\rangle$  [44].

We will uniquely focus on the square GKP code in the remainder of this review, although most points apply with small variation to the hexagonal code.

### 2.3.1. Approximate GKP states

Any physical GKP code state will occupy a finite volume in phase space and will have a finite number of photons. In principle, an infinite number of approximations to the perfect GKP code states exist, but some are more useful than other’s and here we will mention four. Reference [40] introduced a form of approximate GKP state obtained by applying a Gaussian superposition of displacements, characterized by a

‘squeezing’ parameter  $\Delta > 0$  to a perfect state:

$$|\bar{\psi}\rangle = \mathfrak{E}|\bar{\psi}_{\text{ideal}}\rangle, \quad \mathfrak{E} \equiv \frac{1}{\sqrt{\pi\Delta^2}} \int_{\mathbb{R}^2} dudv \exp\left(-\frac{u^2+v^2}{2\Delta^2}\right) \exp(-iu\hat{p} + iv\hat{q}). \quad (14)$$

For this model wavefunction it holds that  $\bar{n} \approx \frac{1}{2\Delta^2} - \frac{1}{2}$  [40, 41]. One can perform the Gaussian phase-space integral in equation (14) and—neglecting contributions  $O(\Delta^4) \ll 1$ , see e.g. [45]—one gets a different approximation using an operator  $\mathfrak{D}$ :

$$\mathfrak{E} \approx 2\sqrt{\pi\Delta^2}\mathfrak{D}, \quad \mathfrak{D} \equiv \exp(-\Delta^2\hat{n}). \quad (15)$$

The envelope operator  $\mathfrak{D}$  has approximately the same effect as the ‘no loss’ Kraus operator of a photon loss channel  $\mathcal{N}_\gamma$ , equation (7), with  $\gamma = 2\Delta^2$ . Another approximation, valid for small  $\Delta$  is

$$\mathfrak{E}|\bar{0}\rangle \approx \mathfrak{F}|\bar{0}\rangle \propto \int_{\mathbb{R}} dq \sum_{k \in \mathbb{Z}} \underbrace{e^{-2\Delta^2\pi k^2}}_{\text{envelope}} \underbrace{e^{-\frac{1}{2\Delta^2}(q-2k\sqrt{\pi})^2}}_{\text{comb}} |q\rangle, \quad (16)$$

$$\mathfrak{E}|\bar{1}\rangle \approx \mathfrak{F}|\bar{1}\rangle \propto \int_{\mathbb{R}} dq \sum_{k \in \mathbb{Z}} \underbrace{e^{-\frac{1}{2}\Delta^2\pi k^2}}_{\text{envelope}} \underbrace{e^{-\frac{1}{2\Delta^2}(q-k\sqrt{\pi})^2}}_{\text{comb}} |q\rangle. \quad (17)$$

The state  $\mathfrak{F}|\bar{0}\rangle$  can be interpreted as the result of preparing a squeezed state  $\frac{1}{\pi^{1/4}} \int dq \exp(-q^2/2\Delta^2)|q\rangle$  to which one applies a Gaussian-enveloped coherent sum over stabilizer translations, enacting  $\rho \rightarrow \sum_{k,l \in \mathbb{Z}} e^{-2\Delta^2\pi(k^2+l^2)} S_p^k \rho S_p^{-l}$ . The result is a state which is both an approximate eigenstate of  $S_q$  (and  $\bar{Z}$ ) due to squeezing, as well as an approximate eigenstate of the translation  $S_p$ . Note that unlike  $\mathfrak{E}$  and  $\mathfrak{D}$ , approximation  $\mathfrak{F}$  has an asymmetry in  $p$  and  $q$ . The three approximations  $\mathfrak{D}$ ,  $\mathfrak{E}$ ,  $\mathfrak{F}$  have been discussed and shown to fit a standard form in [46]. In addition, the normalization of these approximate forms can be computed and expressed in terms of theta functions, see e.g. appendix A for the  $\mathfrak{D}$ -approximation.

In equation (B.2) we will see a fourth, von-Mises or reverse-Villain, approximation using a cosine function to represent the periodicity in the wave-function comb. This reverse-Villain approximation has been used in [47, 48]. All these approximate states  $\mathfrak{E}|\bar{0}\rangle$  and  $\mathfrak{E}|\bar{1}\rangle$  (or  $\mathfrak{D}|\bar{0}\rangle$  and  $\mathfrak{F}|\bar{0}\rangle$  etc) are +1 eigenstates of the photon parity operator  $e^{i\pi a^\dagger a}$  as they are invariant under  $q \rightarrow -q$  and  $p \rightarrow -p$ , implying that they only have support on even photon number states. In appendix A we show how to get exact Fock state amplitudes for the approximation  $\mathfrak{D}|\bar{0}\rangle$ —which for this purpose has the simplest form—and this turns out to involve  $n$ -the order derivatives of theta functions. We show in appendix A that the photon number distribution of these GKP states, as well as the sensor state, is following a thermal distribution [22] (see figures A1 and A2), with interesting oscillations on top.

One can propose various measures of state quality or fidelity besides the characterization of the state in terms of  $\Delta$ . For example, when we measure  $\hat{q}$  to infer  $\bar{Z}$  on a state, all outcomes in which  $q$  is closer to an even multiple of  $\sqrt{\pi}$  are interpreted as outcome  $\bar{Z} = 1$  and vice-versa. For a state  $\int dq \psi(q)|q\rangle$ , the probability for this outcome is then

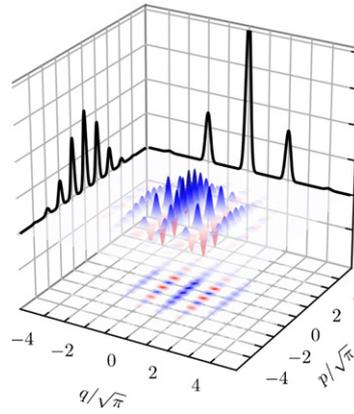
$$\mathbb{P}(\bar{Z} = (-1)^b) = \int_{I_b} dq |\psi(q)|^2, \quad I_b = \left\{ q \mid \exists k \in \mathbb{Z}, -\frac{\sqrt{\pi}}{2} \leq q + (2k+b)\sqrt{\pi} \leq \frac{\sqrt{\pi}}{2} \right\}. \quad (18)$$

If we apply this to the form  $\mathfrak{E}|\bar{0}\rangle$ , the error probability  $\mathbb{P}(\bar{Z} = -1) < \frac{2\Delta}{\pi} \exp(-\pi/4\Delta^2)$ . Since a perfect (homodyne) measurement of  $\hat{q}$  is practically not possible,  $\mathbb{P}(\bar{Z} = -1)$  only provides a lower bound on the logical error probability of an approximate state  $|\bar{0}\rangle$ . We can also examine the expectation value for  $\bar{Z}$  on the approximate form  $\mathfrak{F}|\bar{0}\rangle$  (for simplicity) which equals

$$\frac{\langle \bar{0} | \mathfrak{F}^\dagger \bar{Z} \mathfrak{F} | \bar{0} \rangle}{\langle \bar{0} | \mathfrak{F}^\dagger \mathfrak{F} | \bar{0} \rangle} \approx \frac{\sum_{k \in \mathbb{Z}} e^{-4\Delta^2\pi k^2} \int_{\mathbb{R}} dq e^{i\sqrt{\pi}q} e^{-\frac{1}{\Delta^2}(q-2k\sqrt{\pi})^2}}{\sum_{k \in \mathbb{Z}} e^{-4\Delta^2\pi k^2} \int_{\mathbb{R}} dq e^{-\frac{1}{\Delta^2}(q-2k\sqrt{\pi})^2}} = e^{-\pi\Delta^2/4}, \quad (19)$$

and similarly  $\frac{\langle \bar{1} | \mathfrak{F}^\dagger \bar{Z} \mathfrak{F} | \bar{1} \rangle}{\langle \bar{1} | \mathfrak{F}^\dagger \mathfrak{F} | \bar{1} \rangle} \approx -e^{-\pi\Delta^2/4}$ , showing that the expectation decays exponentially in  $\Delta^2$  towards 0. In the approximation in equation (19) we have assumed that  $\Delta$  is small enough so that the peaks at different  $k$  do not overlap, giving an easy expression for the probability distribution over  $q$  of the approximate GKP state. We further discuss the logical  $\bar{Z}$  or  $\bar{X}$  measurement of a GKP qubit in section 3.2.

It has become common to describe the quality of a GKP state in terms of an amount of squeezing expressed in dB. For a regular squeezed state (squeezed along  $q$ ) one has variances  $\text{Var}(q) = \frac{\Delta^2}{2}$ ,  $\text{Var}(p) = \frac{1}{2\Delta^2}$  as the vacuum (or coherent state) has  $\text{Var}(q) = \text{Var}(p) = \frac{1}{2}$  with  $\Delta = e^{-|\xi|} < 1$ . The



**Figure 2.** Wigner function of the state  $\mathfrak{F}(\bar{0})$  at  $\Delta = 0.3$ , and the reduced probability distributions over  $q$  and  $p$  in black. Unlike the  $\mathfrak{E}$ - and  $\mathfrak{D}$ -approximation, the  $\mathfrak{F}$ -approximation has a clear asymmetry with respect to  $p$  and  $q$ . Since the Wigner function has a grid-like periodic structure in phase space, the GKP states are also referred to as grid states.

convention which is used in the literature for denoting the dB of squeezing of an approximate GKP state is  $\#dB = -10 \log_{10} \Delta^2$ , see e.g. [45].

We can view a GKP state as being ‘squeezed’ in both  $p$  and  $q$  and interpret this squeezing as the extent in which the state is an eigenstate of a unitary operator such as  $S_p$  or  $S_q$ . Since a quantum state may not fit one of the standard GKP approximations, a measure of the effective squeezing is useful in expressing the quality of the state. Since we are interested in modular values of  $\hat{q}$  and  $\hat{p}$ , it is appropriate to use the Holevo phase variance (or the variance of periodic variables such as phases used in circular statistics) to express this squeezing, i.e. one can define [44, 49]:

$$\Delta_{p/q} = \sqrt{\frac{1}{2\pi} \ln \left( \frac{1}{|\text{Tr } S_{p/q} \rho|^2} \right)}. \quad (20)$$

Note that this measure does not express a logical error rate, e.g. the completely mixed state inside the perfect code space has  $\Delta_p = \Delta_q = 0$ .

### 2.3.2. Logical gates

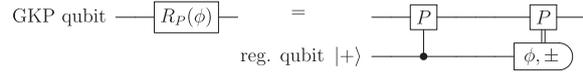
An appealing feature of the GKP code is that all logical Clifford transformations are Gaussian quantum operations, realizable by optical elements [40, 45] which enact linear transformations on the operators  $\hat{p}$  and  $\hat{q}$  in the Heisenberg picture. Important gates such as the CNOT and S gate do however involve two-mode, respectively single-mode squeezing: the experimental realization of such squeezing transformations is typical through pumped optical *non-linearities*. Such elements are relatively straightforward to obtain for optical fields which travel through nonlinear  $\chi^{(2)}$  or  $\chi^{(3)}$  materials, while for superconducting devices these elements are engineered through the use of Josephson junctions. In contrast, passive linear optical elements—beam-splitters and phase-shifters in optics language—are readily available in circuit-QED by linear capacitive or inductive (fixed) circuit couplings.

In section 3 we will discuss the engineered non-linearities in superconducting hardware which can be activated by microwave drives or activated by flux-drives, while here we discuss the logical gates for the GKP code at a formal level.

As unitary displacement operators,  $Z$  and  $X$  are not self-inverse, i.e.  $X \neq X^\dagger$ . On a perfect, completely shift-invariant code state  $X$  acts identically to  $X^\dagger$ , but on a finite-photon number state, see e.g. the wave function in figure 2, it does not: a shift to the left or right moves the envelope away from the center. The Hadamard gate has Heisenberg action  $\hat{p} \rightarrow -\hat{q}$  and  $\hat{q} \rightarrow \hat{p}$  so that  $H^\dagger X H = Z$ ,  $H^\dagger Z H = X^\dagger$  and  $H^\dagger Y H = -Y$ . The Hadamard gate corresponds to a phase-space rotation by an angle  $\pi/2$ , i.e. we can choose  $\text{Had} \equiv \exp(i\frac{\pi}{2} a^\dagger a)$ , and note again that  $\text{Had} \neq \text{Had}^{-1}$ . A Had gate could be done by a quarter-cycle waiting in the self-evolution of the oscillator (so comes for free).

A disadvantage of using such quarter-cycle waiting Hadamard gate in a GKP surface code architecture is discussed in section 4. The alternative is to use single-qubit rotations around the logical  $X$ ,  $Y$  or  $Z$  axes to compose a Hadamard gate.

For the GKP code these rotations around logical axes,  $R_p(\phi) \equiv \exp(-i\phi P/2)$  with logical Pauli  $P = X, Y, Z$  are not natural as the logical Pauli, which is a displacement, sits in the exponent. Note also that



**Figure 3.** Performing a single-qubit gate  $R_P(\phi)$  with  $P = X, Y, Z$  on a perfect GKP qubit via a regular ancilla qubit, requiring a qubit controlled-displacement. The measurement is in the basis  $|\phi, \pm\rangle = \frac{1}{\sqrt{2}}(e^{i\phi/2}|+\rangle \pm e^{-i\phi/2}|-\rangle)$  and upon outcome  $-1$ ,  $P$  is applied.

this gate  $R_P(\phi)$  is only unitary when acting on a subspace for which  $P^2 = I$ . However, one can perform  $R_P(\phi)$ , using a controlled-displacement coupling with a regular qubit and a regular qubit rotation, as shown in figure 3, and realized in [42, 50]. This circuit applies  $R_P(\phi) \equiv \exp(-i\phi P/2)$  on the space of states for which  $P^2 = I$  but we can examine its effect more generally. Imagine applying the circuit in figure 3 with  $P = Y$  and  $\phi = \pi/2$ . Upon outcome  $\pm$ , the Kraus operator action on the GKP qubit equals  $A_+ = \cos(\phi/2)I - i \sin(\phi/2)P$  resp.  $A_- = -i \sin(\phi/2)I + \cos(\phi/2)P$ . On the perfect code subspace where  $P^2 = I$ ,  $A_+$  acts as a unitary and equals  $R_P(\phi)$ , while  $A_-$  can be converted to  $R_P(\phi)$  by the additional  $\pi$ -rotation  $P$ . However, on a finitely-squeezed GKP state, these Kraus operators are not unitary and their action leads to the envelope of the GKP state to be no longer centered around the vacuum. However, one can apply a displacement  $P^{-1/2}$  [42] to approximately re-center the GKP state.

A single-qubit gate such as the  $T = R_Z(\pi/4)$  gate can be done in this manner as well. The  $S$  gate with action  $S^\dagger X S = -Y$  and  $S^\dagger Z S = Z$  can be realized by the transformation  $\hat{q} \rightarrow \hat{q}$ ,  $\hat{p} \rightarrow \hat{p} - \hat{q}$  corresponding to  $S = \exp(-i\hat{q}^2/2)$ .<sup>9</sup> The  $S$  gate can thus be implemented by means of pump-activated squeezing, see section 3, or by using an ancilla qubit as in the circuit in figure 3. Alternative methods for performing a  $T$  gate via magic state preparation or using a cubic phase gate  $V_\gamma = \exp(i\gamma\hat{q}^3)$  exist [40]. For example, one can create a  $+1$  eigenstate of the Hadamard gate  $\text{Had} = \exp(i\frac{\pi}{2}a^\dagger a)$  by starting with a vacuum state, which is already a  $+1$  eigenstate of  $\text{Had}$ , and measuring  $S_p$  and  $S_q$  without photon-number changing feedback [51].

When using GKP qubits as basic qubits in a surface code, see section 4, we note that  $T$  and  $S$  gates are not needed for error correction: their only use is to prepare magic GKP ancilla qubits to be grown into the surface code-encoded magic states using GKP CZ and CNOT gates or parity check measurements, see e.g. [52] and references therein.

The CNOT gate can be realized by the Heisenberg action  $\hat{q}_c \rightarrow \hat{q}_c$ ,  $\hat{p}_c \rightarrow \hat{p}_c - \hat{p}_t$ ,  $\hat{q}_t \rightarrow \hat{q}_c + \hat{q}_t$  and  $\hat{p}_t \rightarrow \hat{p}_t$ . This gate is also called the SUM gate in [40] and  $\text{SUM}(g)$  with  $g = 1$  in [45]. We see that  $\text{CNOT} = \exp(-i\hat{p}_t\hat{q}_c)$  by using equation (2) with  $v = \hat{p}_t$  and  $u = \hat{q}_c$ . The inverse CNOT has action  $\hat{q}_c \rightarrow \hat{q}_c$ ,  $\hat{p}_c \rightarrow \hat{p}_c + \hat{p}_t$ ,  $\hat{q}_t \rightarrow \hat{q}_t - \hat{q}_c$  and  $\hat{p}_t \rightarrow \hat{p}_t$ .

We define the action of the CZ gate as  $\text{Had}_t \text{CNOT} \text{Had}_t^\dagger$  where  $\text{Had}_t$  is a Hadamard gate on the target mode. That is, it enacts the transformation  $\hat{q}_t \rightarrow \hat{q}_t$ ,  $\hat{p}_t \rightarrow \hat{p}_t - \hat{q}_c$ ,  $\hat{q}_c \rightarrow \hat{q}_c$ ,  $\hat{p}_c \rightarrow \hat{p}_c - \hat{q}_t$ , or  $\text{CZ} = \exp(-i\hat{q}_t\hat{q}_c)$ . If either oscillator is a state where  $q$  is an even multiple of  $\sqrt{\pi}$ , then CZ acts as  $\exp(-i\pi 2k) = 1$ . If both oscillators are in a state where  $q$  is an odd multiple of  $\sqrt{\pi}$ , then CZ acts as  $\exp(-i\pi(2n+1)(2k+1)) = -1$  for  $n, k \in \mathbb{Z}$ .

Sections 3.3 and 3.4 will discuss how the GKP CZ gate between two GKP modes can be executed using a three-wave or four-wave mixing element. There is however another circuit to perform a CNOT gate which uses a sequence of beam-splitters and some single-mode squeezing [41, 45] which can be more useful in some circumstances, see figure 4. For the CNOT gate the mode transformation on control ( $c$ ) and target ( $t$ ) mode equals

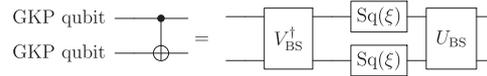
$$\begin{pmatrix} a_c^{\text{out}} \\ a_t^{\text{out}} \end{pmatrix} = A \begin{pmatrix} a_c \\ a_t \end{pmatrix} + B \begin{pmatrix} a_c^\dagger \\ a_t^\dagger \end{pmatrix}, \quad (21)$$

with

$$A = \begin{pmatrix} 1 & -\frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix}. \quad (22)$$

By the Bloch–Messiah decomposition [53] the singular value decompositions are  $A = U D_A V^\dagger$  and  $B = U D_B V^T$  with unitary matrices  $U$  and  $V$ . For the CNOT gate the singular values are degenerate:  $D_A = \text{diag}(\frac{\sqrt{5}}{2}, \frac{\sqrt{5}}{2})$  and  $D_B = \text{diag}(\frac{1}{2}, \frac{1}{2})$ , implying that the beam-splitting transformations  $U$  and  $V$  are not

<sup>9</sup> Perhaps the simplest way to derive this identity is to calculate  $\exp(i\hat{q}^2/2)\hat{p} \exp(-i\hat{q}^2/2) = \sum_{n=0}^{\infty} \frac{1}{n!} (\text{ad}_{i\hat{q}^2/2})^n(\hat{p})$ , with  $(\text{ad}_A)^0(B) = B$ ,  $\text{ad}_A(B) = [A, B]$ ,  $(\text{ad}_A)^2(B) = [A, [A, B]]$  etc. We can use that  $[\hat{q}^2, \hat{p}] = 2i\hat{q}$  and higher-order commutators are zero, leading to  $\exp(i\hat{q}^2/2)\hat{p} \exp(-i\hat{q}^2/2) = \hat{p} - \hat{q}$ .



**Figure 4.** The realization of a CNOT via 50:50 beam-splitters, i.e.  $V_{BS}$  and  $U_{BS}$  defined in equation (23), and single-mode squeezing  $Sq(\xi)$  with  $\xi \approx -0.4812$ .

unique. Reference [53] notes that taking 50:50 beamsplitters with

$$U = U_{BS} \equiv \frac{1}{\sqrt{2}} \begin{pmatrix} ie^{i\theta} & ie^{-i\theta} \\ -e^{i\theta} & e^{-i\theta} \end{pmatrix}, \quad V = V_{BS} \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} U^*, \quad (23)$$

with  $\theta = \frac{1}{2} \sin^{-1}(2/\sqrt{5})$  can be chosen (while [45] makes a different choice). We see that the single-mode squeezing represented by the diagonal matrices corresponds to a squeezer  $Sq(\xi)$  with  $\xi = -\cosh^{-1}(\sqrt{5}/2)$ .

It is clear that logical gates are not unique as physical operations as they only have to perform the right action on the code space. Reference [45] has discussed how logical gates propagate or amplify errors on the approximate GKP code states. Keeping the (average) number of photons in an approximate GKP state low by centering the state symmetrically around the vacuum, emerges as a good overall strategy to minimize the propagation of errors and the effect of the inaccurate action of gates.

### 2.3.3. Noise on a GKP qubit

A simple numerically convenient noise channel, playing the role of depolarizing channel for an oscillator, is the independent Gaussian displacement channel  $\mathcal{N}(\rho)$  with standard deviation  $\sigma_0$ :

$$\mathcal{N}(\rho) = \int_{-\infty}^{\infty} du \int_{-\infty}^{\infty} dv \mathbb{P}_{\sigma_0}(u) \mathbb{P}_{\sigma_0}(v) e^{iu\hat{p}+iv\hat{q}} \rho e^{-iu\hat{p}-iv\hat{q}}. \quad (24)$$

Here  $\rho$  is a single-mode density matrix and  $\mathbb{P}_{\sigma_0}(x)$  the Gaussian probability density function with mean zero and variance  $\sigma_0^2$ , i.e.  $\mathbb{P}_{\sigma_0}(x) = (2\pi\sigma_0^2)^{-1/2} e^{-x^2/2\sigma_0^2}$ . This channel does not naturally correspond to physical sources of noise, but (1) one can convert photon loss via amplification to this channel [22], (2) one can ‘displacement twirl’ noise so that the effective channel is that of probabilistic mixture of displacements [54]. The exact displacement twirl is not a physical operation as it requires large displacements, so this type of modeling should be considered less justified than in the qubit Pauli case when we use a depolarizing noise model through a Pauli twirling approximation.

It is thus of interest to study how realistic noise affects the approximate GKP states beyond this toy model. We will explore the question of stochastic Gaussian displacement noise versus coherent finite-squeezing error during quantum error correction in the next section 2.4. In this section we describe the interesting effect of photon loss on a GKP qubit using Wigner function dynamics [42], and mention some literature discussing other sources of noise.

An oscillator state undergoing photon loss at rate  $\kappa$  can be described, in a rotating frame at its resonant frequency, using a Lindblad equation  $\dot{\rho} = \kappa \mathcal{D}(a)(\rho)$  using the density matrix  $\rho$ . Here we assume that the thermal environment which induces this photon loss is at zero temperature, hence there are no photon gain processes. Alternatively, and conveniently, one describes this dynamics through differential equations using phase-space probability distributions such as the Wigner function. The Wigner function  $W(q, p, t) \equiv \frac{1}{2\pi} \int dx e^{-ipx} \langle q + \frac{x}{2} | \rho(t) | q - \frac{x}{2} \rangle$  for the photon loss dynamics can be shown to obey a two-dimensional Fokker–Planck equation, see [9, 42, 55]

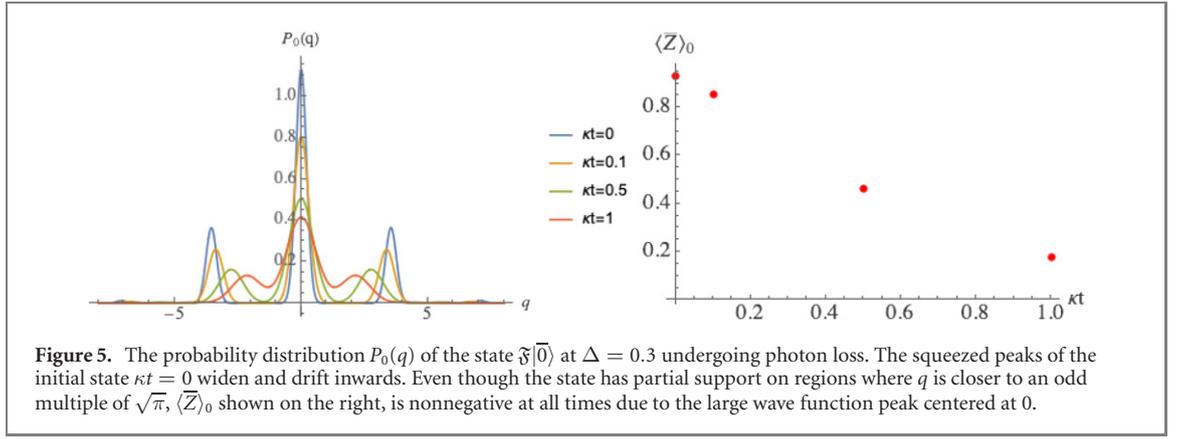
$$\frac{\partial W(q, p, t)}{\partial t} = \frac{\kappa}{2} \left( \frac{\partial}{\partial q} (qW(q, p)) + \frac{\partial}{\partial p} (pW(q, p, t)) + \frac{1}{2} \left( \frac{\partial^2 W(q, p, t)}{\partial^2 p} + \frac{\partial^2 W(q, p, t)}{\partial^2 q} \right) \right).$$

This Fokker–Planck equation describes a process of diffusion—a spread in the variance of the variables  $p$  and  $q$  to the vacuum noise variance equal to  $1/2$ —and drift, i.e. the mean values of  $p$  and  $q$  flow towards 0. Instead of considering the Wigner function dynamics, we can integrate over, say,  $p$  and consider the corresponding Fokker–Planck equation for the probability distribution  $P(q, t) = \int_{\mathbb{R}} dp W(q, p, t)$ , which has the solution:

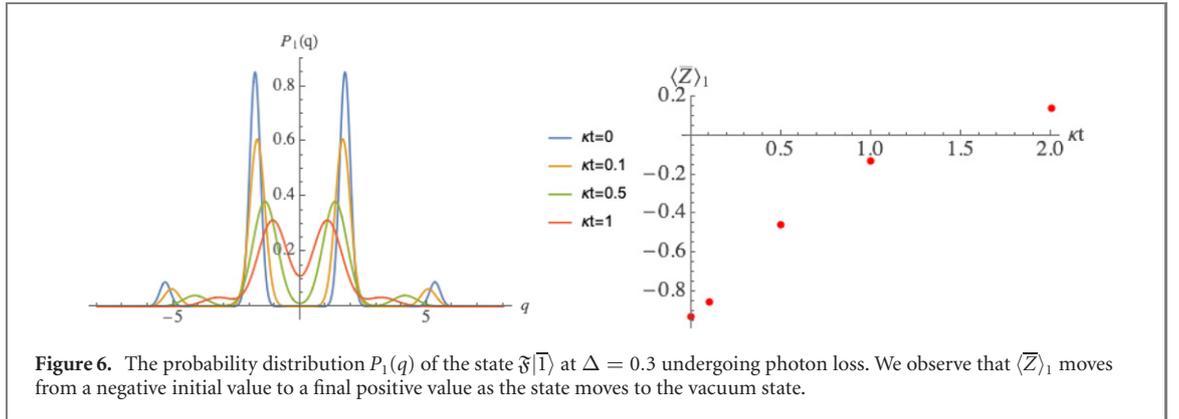
$$P(q, t) = \int dq' P_{\text{trans}}(q, t|q', 0) P(q', t = 0),$$

$$P_{\text{trans}}(q, t|q', 0) = \sqrt{\frac{1}{2\pi\sigma^2(t)}} \exp\left(-\frac{(q - q' e^{-\kappa t/2})^2}{2\sigma^2(t)}\right),$$

$$\sigma^2(t) = \frac{1}{2}(1 - \exp(-\kappa t)). \quad (25)$$



**Figure 5.** The probability distribution  $P_0(q)$  of the state  $\mathfrak{F}|\bar{0}\rangle$  at  $\Delta = 0.3$  undergoing photon loss. The squeezed peaks of the initial state  $\kappa t = 0$  widen and drift inwards. Even though the state has partial support on regions where  $q$  is closer to an odd multiple of  $\sqrt{\pi}$ ,  $\langle \bar{Z} \rangle_0$  shown on the right, is nonnegative at all times due to the large wave function peak centered at 0.



**Figure 6.** The probability distribution  $P_1(q)$  of the state  $\mathfrak{F}|\bar{1}\rangle$  at  $\Delta = 0.3$  undergoing photon loss. We observe that  $\langle \bar{Z} \rangle_1$  moves from a negative initial value to a final positive value as the state moves to the vacuum state.

In figure 5 we plot the effect of photon loss of a normalized state  $\mathfrak{F}|\bar{0}\rangle$  with  $\Delta = 0.3$  for  $\kappa t = 0.1, 0.5$  and 1.

We can consider the expectation of a stabilizer or logical  $\bar{Z}$  over time, i.e. we consider

$\text{Tr} e^{i\alpha\hat{q}}\rho(t) = \int dq P(q, t) e^{i\alpha q}$  with  $\alpha = \sqrt{\pi}$  or  $2\sqrt{\pi}$ . Using Gaussian integration and equation (25) this gives

$$\text{Tr} e^{i\alpha\hat{q}}\rho(t) = \left( \sqrt{\frac{1}{2\pi\sigma^2(t)}} \int dq e^{-\frac{q^2}{2\sigma^2(t)}} e^{i\alpha q} \right) \text{Tr} e^{i\alpha(t)\hat{q}}\rho(0) = e^{-\frac{1}{4}(\alpha^2(0) - \alpha^2(t))} \text{Tr} e^{i\alpha(t)\hat{q}}\rho(0), \quad (26)$$

with  $\alpha(t) = \alpha e^{-\kappa t/2}$ . On the right-hand side, we see an exponential decrease as well as a direct dependence on the expectation value of a displacement operator with exponentially shrinking shift on the initial state. When the initial state  $\rho(0)$  is invariant under  $q \rightarrow -q$ , we can replace  $\text{Tr} e^{i\alpha(t)\hat{q}}\rho(0)$  by  $\text{Tr} \cos(\alpha(t)\hat{q})\rho(0)$ . Thus when symmetrically centering the state in phase-space the phases of the stabilizer or logical  $\bar{Z}$  never become complex. In addition, when the initial state is an approximate logical  $|\bar{0}\rangle$  such as  $\mathfrak{F}|\bar{0}\rangle$ , the expectation value  $\langle \bar{Z}(t) \rangle \geq 0$  at all times as shown for a few points in figure 5 on the right. This is interesting as it shows that  $|\bar{0}\rangle$  ‘never looks more like a  $|\bar{1}\rangle$  than a  $|\bar{0}\rangle$ ’ under photon loss. The state  $\mathfrak{F}|\bar{1}\rangle$  whose decay is plotted in figure 6 starts at  $\bar{Z}(t=0) < 0$  and eventually, for large enough  $t$ ,  $\bar{Z}(t) > 0$  as the final state is the vacuum centered around  $q = 0$ . This asymmetry in its effect on  $|\bar{0}\rangle$  versus  $|\bar{1}\rangle$  is reminiscent of a logical amplitude-damping channel.

Now assume that the initial state is displaced away from its centered location by, say, a stabilizer shift  $S_p^m$  which does not affect its initial eigenvalue for  $\bar{Z}$ . Using equation (26) we get

$$\langle \bar{Z}(t) \rangle = e^{2\pi i m e^{-\kappa t/2}} e^{-\frac{1}{4}(\alpha^2(0) - \alpha^2(t))} \text{Tr} e^{i\alpha(t)\hat{q}}\rho(0), \quad (27)$$

which shows that the expectation value can now become complex, *but is not faster decaying in its absolute value*. When  $m$  is large, we see that the additional phase changes rapidly in time, so that the expectation can rapidly change from positive to negative. However, if we know  $m$  and  $\kappa$  and it is the only source of noise, this phase change can be treated as a systematic error. Note that if we had applied an arbitrary but known displacement  $e^{i\mu\hat{p}}$  on the initial state, the effect would have been similar.

Going beyond photon loss, other sources of inaccuracy and error could also readily be described using dynamics of the Wigner function. A Lindblad equation dynamics of an  $n$ -mode system for which the Hamiltonian is quadratic in creation and annihilation operators (beam-splitting, squeezing etc) or *linear*

(driving terms  $\sim a + a^\dagger$  enacting displacements) while the dissipator models photon loss or photon gain, can be mapped to a Fokker–Planck equation of a general solvable form:

$$\begin{aligned} \frac{\partial}{\partial t} W(q_1, p_1, \dots, q_n, p_n) &= \left( -\vec{\nabla} \cdot (A\vec{x}) + \frac{1}{2} \vec{\nabla} \cdot D\vec{\nabla} \right) W(q_1, p_1, \dots, q_n, p_n), \\ \vec{x}^T &= (q_1, p_1, \dots, q_n, p_n), \quad \vec{\nabla}^T = \left( \frac{\partial}{\partial q_1}, \frac{\partial}{\partial p_1}, \dots, \frac{\partial}{\partial q_n}, \frac{\partial}{\partial p_n} \right), \end{aligned} \quad (28)$$

with constant  $2n \times 2n$  matrices  $A$  and  $D$ . This general behavior follows from the fact that every term in a Lindblad equation which is linear in  $a$  or  $a^\dagger$  (e.g.  $a\rho$ ), gives rise to a first-order derivative in the differential equation for the Wigner function (plus a term which is linear in  $\hat{p}$  and  $\hat{q}$ ) [9, 55], so that terms quadratic in  $a$  and  $a^\dagger$  (e.g.  $a^\dagger a\rho$ ) gives second-order derivatives. The Gaussian Green's function for equation (28) can be readily given, basically forming a multi-dimensional analog of equation (25), see [9]. All these Gaussian processes keep an initially nonnegative Wigner function nonnegative and hence are simulatable by stochastic means.

On the other hand, nonlinear elements such as a self-Kerr nonlinearity  $-Ka^\dagger a^2$  lead to third-order derivatives in the differential equation for the Wigner function, as well as terms in which  $A$  is not constant (corresponding to a so-called nonlinear Fokker–Planck equation): the upshot is that the Wigner function can become negative and non-classical during the dynamics and attempts at classical stochastic simulation will suffer from the sign problem. As an example, reference [56] discusses Wigner function dynamics for a single oscillator in the presence of a self-Kerr nonlinearity and dissipation.

Dephasing, meaning the application of a rotation  $e^{i\theta a^\dagger a}$  with unknown  $\theta$  is a possible error mechanism as it rotates the quadratures  $\hat{p}$  and  $\hat{q}$  into each other. Dephasing can come about, for example, from an interplay of photon loss and a Kerr nonlinearity, or a fluctuating mode frequency. In a simple stochastic model the angle  $\theta$  is drawn from a distribution  $\mathbb{P}(\theta)$  with mean  $\langle \theta \rangle = 0$  and some moments  $\langle \theta^k \rangle$ . For small higher-order moments  $\langle \theta^k \rangle \ll 1$  for  $k > 2$ , we can expand

$$\mathcal{N}_{\text{deph}, \langle \theta^2 \rangle}(\rho) = \int d\theta \mathbb{P}(\theta) e^{i\theta a^\dagger a} \rho e^{-i\theta a^\dagger a} \approx \rho + \langle \theta^2 \rangle a^\dagger \rho a - \frac{1}{2} \langle \theta^2 \rangle (a^\dagger \rho + \rho a^\dagger a) + O(\langle \theta^3 \rangle). \quad (29)$$

This is a dephasing channel which corresponds to the dynamics of a Lindblad equation  $\dot{\rho} = \kappa_{\text{deph}} \mathcal{D}(a^\dagger a)\rho$  for a short time with  $\kappa_{\text{deph}} t = \langle \theta^2 \rangle \ll 1$ . The fixed point of this equation is any mixture of Fock states  $|n\rangle\langle n|$ ; when the initial state is  $\sum_n c_n |n\rangle$  the channel maps it onto  $\sum_n |c_n|^2 |n\rangle\langle n|$ . In appendix A we evaluate the photon number distribution of such fully-dephased  $\mathfrak{D}|\bar{0}\rangle$  and  $\mathfrak{D}|\bar{1}\rangle$ . We prove that the photon number distribution is asymptotically thermal, independent of the logical state. Hence complete dephasing seems to wash out much of distinction between the two logical GKP states.

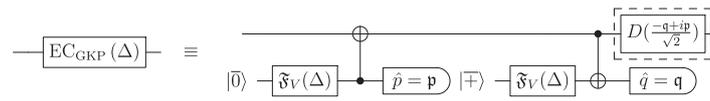
Reference [22] has discussed the detrimental effect of a Kerr nonlinearity on a variety of single-mode bosonic codes. Numerical simulations of several sources of inaccuracies on GKP state preparation using an ancilla qubit were also discussed in e.g. [42, 44, 45, 50, 57] using Lindblad equation dynamics.

## 2.4. Repeated GKP error correction and decoding: finite squeezing

In this section we examine the effect of (coherent) finite-squeezing errors on repeated GKP error correction using GKP ancilla's. This is follow-up work from reference [48] in which a similar problem was examined using a stochastic Gaussian displacement error model, equation (24), applied to GKP ancilla and data qubits as a proxy for finite-squeezing errors. The goal of this section is to understand whether there are crucial differences between finite-squeezing coherent errors and the Gaussian displacement error model and try to develop a dedicated, computationally-efficient, decoder with good performance.

The dynamics to be analyzed is the repeated execution of the quantum circuit in figure 7 on a single GKP input state  $\mathfrak{F}(\Delta)|\bar{\psi}\rangle$  for  $m = 1, \dots, M$  cycles. We remark that a variant of such 'Steane error correction' exists: in [58] the authors observed that applying a beam-splitter between GKP ancilla and GKP data qubit followed by squeezing on the GKP data qubit, can also perform error correction. Reference [59] has analyzed the repeated execution of this variant of error correction in more detail.

A clear difference between a stochastic error model and the finite-squeezing model is that in the former entropy build-up is possible, while in the latter the state conditioned on the measurement outcomes in figure 7 is pure at all times. One can invoke displacement twirling as a method to convert a coherent noise model in which one applies a superposition of displacements to a stochastic mixture of displacements. For example, displacement twirling a finitely-squeezed state  $\mathfrak{E}|\bar{\psi}\rangle$  with some  $\Delta$  gives a perfect state  $|\bar{\psi}\rangle$  subject to the Gaussian displacement channel with  $\Delta^2 = 2\sigma_0^2$  [48]. After such stochasticification of the noise on a GKP ancilla, one can then represent the feedback error (a shift in one of the quadratures) induced by the ancilla in the circuit in figure 7 effectively as an incoming stochastic shift error on the data qubit. The



**Figure 7.** A single round of fault-tolerant GKP error correction for both logical  $X$  and  $Z$  errors. Each measurement is a perfect homodyne measurement of  $\hat{q}$  or  $\hat{p}$  with outcomes  $q$  and  $p$  respectively. The finitely-squeezed ancilla states are modeled as approximate GKP states, using a slightly-different small- $\Delta$  approximation than  $\mathfrak{E}$ ,  $\mathfrak{D}$ ,  $\mathfrak{F}$  in equations (14), (15) and (17), given in equation (B.2) and denoted as  $\mathfrak{F}_V$ . We show that the optional corrective displacement (dashed box) keeps the state at low average photon number  $\bar{n}$  in figure 9.

stochastic shift of the other quadrature of the ancilla then causes a measurement error of the same strength. On this basis, reference [48] stochastically modeled finite-squeezing errors as incoming stochastic displacement errors on the GKP data qubit and measurement errors. Another difference in the models is that in the finite-squeezing error model the measurement outcomes carry non-modular information about the measured quadrature. This can be exploited to recenter the state by choosing a corrective displacement immediately after a single round of error correction, while such corrective displacements would have no effect in the stochastic model.

We will represent the GKP wave-function in the  $q$ -basis with  $\psi_m(q)$ , the wavefunction after  $m$  rounds of error correction. We will sometimes omit the normalization of states when these normalizations play no role.

We will now analyze the time evolution without the corrective displacement in the dashed box in figure 9. A single round of quantum error correction shown in figure 7 with measurement outcomes  $p_m, q_m$  gives  $\psi_m(q) = \int dq' G(q \leftarrow q' | q_m, p_m) \psi_{m-1}(q')$  with Green's function

$$G(q \leftarrow q' | q_m, p_m) = \int dq'' G_+(q \leftarrow q'' | q_m) G_0(q'' \leftarrow q' | p_m) = \psi^+(q - q_m) \psi^0(q - q') e^{-ip_m(q - q')}, \quad (30)$$

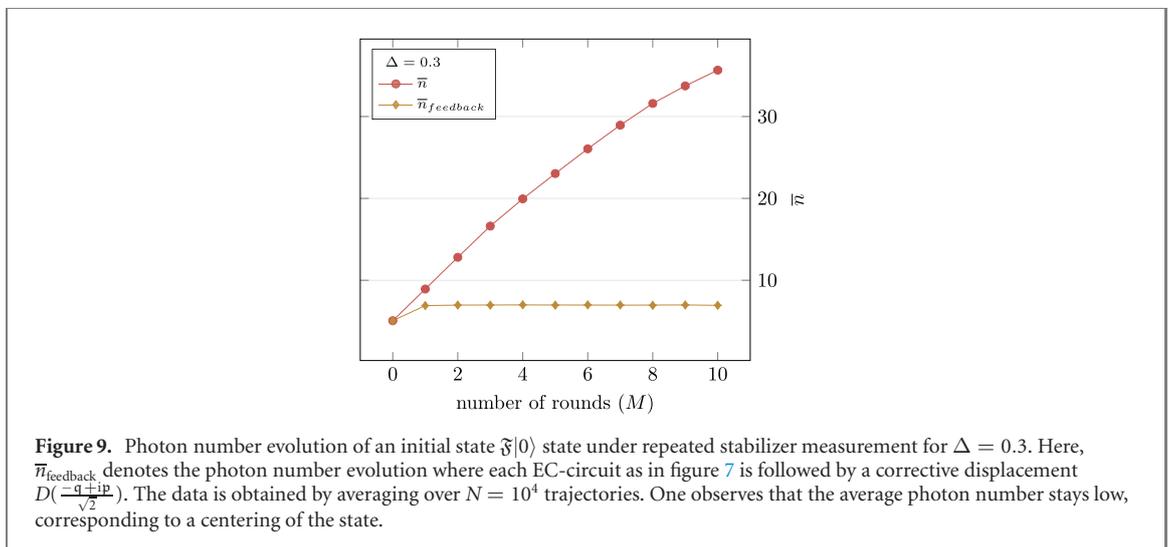
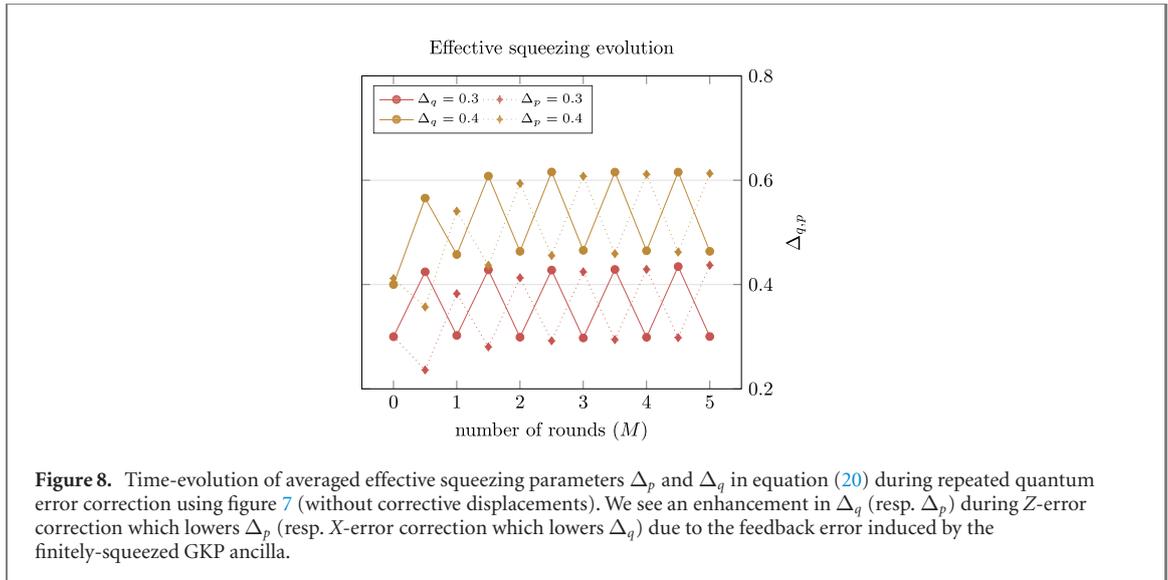
using  $G_0(q'' \leftarrow q' | p_m) = \psi^0(q'' - q') e^{-ip_m(q'' - q')}$  ( $Z$ -error correction) followed by  $G_+(q \leftarrow q'' | q_m) = \delta(q'' - q) \psi^+(q - q_m)$  ( $X$ -error correction). To understand this Green's function, observe that in the limit  $\Delta \rightarrow 0$ , the wavefunction  $\psi^+(q)$  has uniform support on  $q = k\sqrt{\pi}$ , with  $k \in \mathbb{Z}$ , so that the outgoing wave function is supported solely on  $q = q_m + k\sqrt{\pi}$ , hence the code state sits in the perfect code space with a known shift on top. However, before this, the interaction with the imperfect  $|\bar{0}\rangle$  ancilla for the  $Z$ -error correction applies a convolution to the incoming wavefunction. If the ancilla is perfect ( $\Delta \rightarrow 0$ ), this convolution amounts to applying superpositions of stabilizer shifts  $S_p^k$  with  $2k\sqrt{\pi} = q - q'$ , each with a phase which depends on  $p_m$ , on the incoming wave-function. If we assume that all wavefunctions are of the form  $\mathfrak{F}|\bar{0}\rangle$ , i.e. sums of Gaussians, the convolution leads again to a sum of Gaussians and can be exactly evaluated, that is, one has

$$\begin{aligned} & \int dq' \psi^0(q - q') e^{-ip_m(q - q')} \psi^0(q') \\ & \propto \sum_{k, k' \in \mathbb{Z}} e^{-2\Delta^2 \pi(k^2 + k'^2)} \int dq' e^{-ip_m(q - q')} e^{-\frac{1}{2\Delta^2} [(q' - 2k'\sqrt{\pi})^2 + (q - q' - 2k\sqrt{\pi})^2]} \\ & = \sqrt{\pi\Delta^2} e^{-\Delta^2 p_m^2 / 4} \sum_{k, k' \in \mathbb{Z}} e^{-2\Delta^2 \pi(k^2 + k'^2)} e^{-\frac{(q - 2\sqrt{\pi}(k + k'))^2}{4\Delta^2}} e^{-i\frac{p_m}{2}(q + 2\sqrt{\pi}(k - k'))}. \end{aligned} \quad (31)$$

What we observe is that the convolution broadens the peaks and they acquire phases which depend on the location of the peaks and the outcome  $p_m$ . The convolution step, which corrects shifts in  $p$ , thus introduces a feedback error in the form of peak broadening for the  $q$  variable.

In figure 8 we plot the effective squeezing parameters  $\Delta_p, \Delta_q$  of the state, equation (20), after rounds of  $Z$  and  $X$ -error correction.

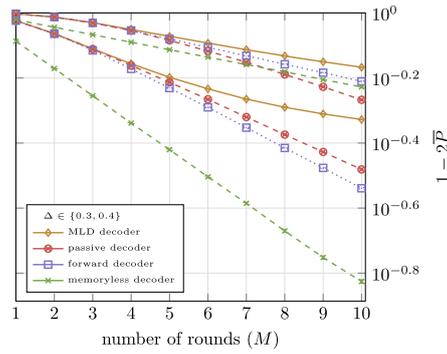
Represented as a state evolution, the stabilizer measurements of  $S_p$  and  $S_q$  with outcomes  $q_m, p_m$  effectively map an incoming state  $|\phi\rangle$  to  $\psi^+(\hat{p} + p_m)|\phi\rangle$  and  $\psi^+(\hat{q} - q_m)|\phi\rangle$ . The finite envelope of the approximate GKP states has the effect that the outgoing states are dominantly supported around  $\hat{p} = -p_m, \hat{q} = q_m$ . We understand this gain of non-modular information in one quadrature as a reflection of the loss of modular information (i.e. in terms of the increase in  $\Delta_{q/p}$ ) in its conjugate. In figure 9 we observe, that a displacement about  $\alpha = \frac{-q_m + ip_m}{\sqrt{2}}$  indeed contributes to a recentering of the state.



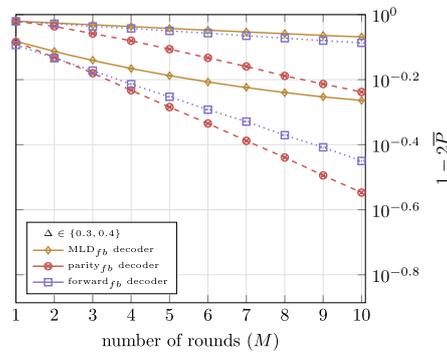
What is noteworthy about the error correcting dynamics in equations (30) and (31) is that the support of the outgoing wavefunction lies within the support of the incoming wavefunction ( $\psi^0(q)$ ) plus the support of the ancilla wavefunction (here also  $\psi^0(q)$ ) since  $\text{Supp}(f \star g) \subset \text{Supp}(f) + \text{Supp}(g)$  for a convolution of two functions  $f$  and  $g$ . The X-error correction step *multiplies* the convoluted wavefunction by  $\psi^+(q - q_m)$  which cannot extend its support. Within its support the outgoing wavefunction can have changed amplitudes, depending on the outcomes  $p_m$  and  $q_m$ . These arguments are relevant as the GKP state  $\mathfrak{F}|\bar{0}\rangle$  has support which is concentrated around even multiples of  $\sqrt{\pi}$  overlapping by an amount exponentially suppressed in  $\Delta^2$  with the support of  $\mathfrak{F}|\bar{1}\rangle$ . When one uses code states  $\psi^0(q)$  and  $\psi^1(q)$  whose supports have negligible overlap—which is the case for  $\mathfrak{F}|\bar{0}\rangle$  and  $\mathfrak{F}|\bar{1}\rangle$  for sufficiently small  $\Delta$ —it implies that, no matter what the measurement outcomes, 0 will largely remain a 0 and 1 will largely remain a 1 in the error correction rounds.

It implies that the picture of stochastic error feedback by using a finite-squeezed  $|\bar{\pm}\rangle$  ancilla is inadequate as in this picture the support of the wavefunction gets shifted around by such feedback error, while here we observe that instead amplitudes get changed *within the support*. We see some of this behavior in the performance of the maximum likelihood decoder versus a passive decoder in figure 10: a passive decoder which decides that 0 stays a 0 does surprisingly well, which may be understandable if we consider that the support of the  $\psi_0(q)$  wavefunction can never grow by QEC (of course, in principle, the wavefunctions have support everywhere albeit exponentially-suppressed with  $1/\Delta^2$ ).

We have formulated a classical ‘forward’ decoder, see details in appendix B, and compared its performance with an optimal, density matrix decoding method (maximum likelihood decoding) as well as a so-called passive decoder, see the numerical results without active displacement feedback in figure 10. The



**Figure 10.** Logistical error rate  $\bar{P}$  for decoding the circuit in figure 7 (without the corrective displacement) for different decoders and squeezing parameters  $\Delta = 0.3$  (top data) and  $0.4$  (bottom data) with the number of stabilizer measurement rounds  $M = 1, \dots, 10$ . For  $\Delta = 0.3$  the mean performance of the forward decoder outperforms passive decoding, while the opposite is true for  $\Delta = 0.4$ . We have observed that the single sample logical error rates strongly fluctuate at  $O(10^{-2}) - O(10^{-1})$  around the averages, suggesting rather chaotic behavior. The data are obtained by sampling over  $N = 5 \times 10^4$  trajectories of measurement outcomes. Error bars denote the standard deviations of the averaged logistical error rates. The memoryless decoder applies a different corrective displacement after each QEC round described in appendix B.



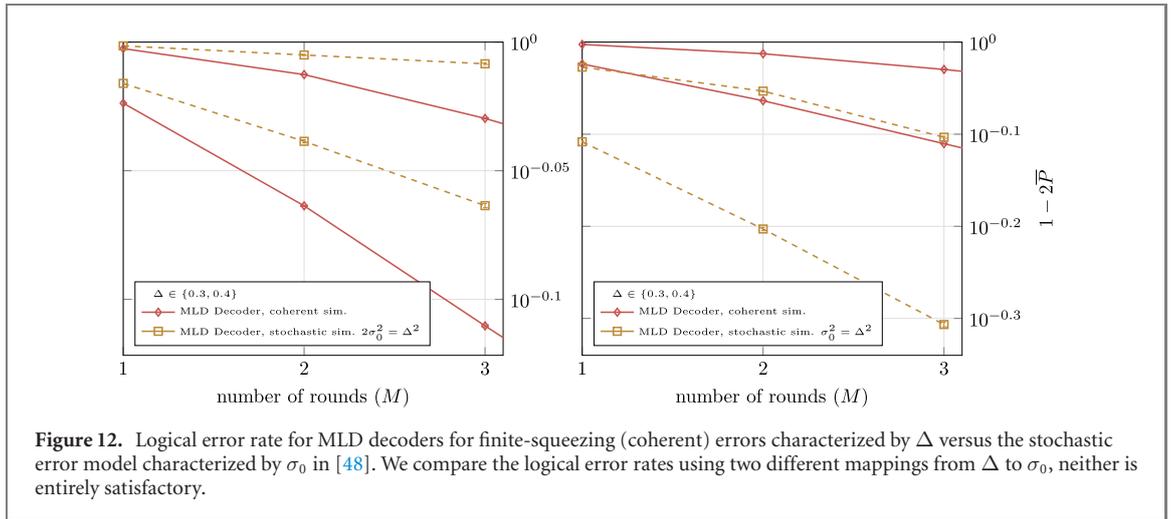
**Figure 11.** Logistical error rate  $\bar{P}$  for decoding the circuit with active feedback as in figure 7. We observe an overall improvement in the logistical error rate as compared to the performance of the circuits without immediate feedback.

passive decoder functions as an important sanity check: this decoder throws away all the measurement data, including the final measurement and *simply always decides that the outcome is 0* when the input state to the rounds of error correction is an approximate GKP 0 state. Since we do not necessarily know the input state, we note that this decoder is of little practical value. By comparing the performance of this decoder with the MLD decoder we learn to what extent the quantum error correction circuits are preserving quantum information irrespective of the measurement outcomes and to what extent the measurement outcomes provide the proper logical information for correction. This decoder is similar to the passive decoder in the stochastic model of [48] in which none of the error information in each round is used and only the last perfect measurement determines whether the state is identified as 0 or 1. However, in [48] this passive decoder clearly performed worse for these values for  $\Delta$ .

We compare the decoder performance in figure 11 to a ‘memoryless’ decoder using the measurement of the current QEC round for immediate logical feedback, see appendix B, clearly showing worse performance with this strategy. In figure 11 we show the better performance of the MLD and the (feedback-adapted) forward decoder using active feedback which minimizes photon number. We compare their performance with a ‘parity’ decoder which similarly applies the corrective displacement, but then applies a final logical correction (or not) when the sum of all applied shifts is closer to an odd (or even) multiple of  $\sqrt{\pi}$ .

We observe that for small  $\Delta$  and number of EC rounds  $M \leq 3$ , the passive decoder performs comparably to the MLD decoder, consistent with the intuition given earlier in this section. At larger  $M$  it performs worse at  $\Delta = 0.3$ . Figures 10 and 11 show the average logistical error rates obtained from  $N = 5 \times 10^4$  samples for all decoders. We note (not shown in the figures) that there are large fluctuations of  $O(10^{-2}) - O(10^{-1})$  per run around the average logistical error rate.

Last, to study the difference between coherent and stochastic errors, we plot  $\bar{P}_{\text{MLD}}^{\text{stoch}}(\sigma_0)$  and  $\bar{P}_{\text{MLD}}(\Delta)$  using the identification  $\Delta^2 = 2\sigma_0^2$  and  $\Delta^2 = \sigma_0^2$  as a function of  $M$  and  $\Delta$ , see figure 12. The simulation for



**Figure 12.** Logical error rate for MLD decoders for finite-squeezing (coherent) errors characterized by  $\Delta$  versus the stochastic error model characterized by  $\sigma_0$  in [48]. We compare the logical error rates using two different mappings from  $\Delta$  to  $\sigma_0$ , neither is entirely satisfactory.

the MLD decoder based on a stochastic error model is implemented following [48]. We observe, that the conversion  $\Delta^2 = 2\sigma_0^2$  underestimates the logical error, while it is overestimated for  $\Delta^2 = \sigma_0^2$ .

The simulation and data are accessible at [https://github.com/JonCYeh/GKP\\_EC\\_Sim](https://github.com/JonCYeh/GKP_EC_Sim).

### 3. Circuit-QED realizations of GKP qubit components

In this section we review and discuss schemes for state preparation, logical gates and quantum error correction for GKP qubits in circuit-QED. In circuit-QED a natural candidate for a bosonic GKP encoding is a resonant mode of a 3D microwave cavity, having low loss rate. Multiple GKP qubits are then stored in multiple low-loss 3D cavities: an engineering platform for multiple coupled cavities—multi-layer microwave-integrated quantum circuits (MMIQC) [60]—is under development.

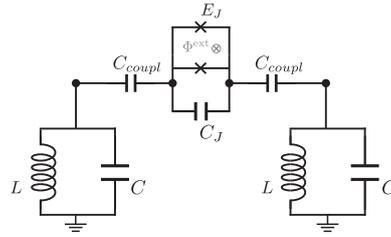
The coupling between cavity modes of different cavities can be mediated by dipolar ‘antenna’ couplings between the electric field of the cavity mode and that of an inserted planar chip in the cavity wall hosting a coupler mode. The idea is then to activate two-mode gates such as the CZ gate by applying microwave drives or flux-modulation outside of the 3D cavities on the coupler mode. As a simple circuit example one can take the electric circuit in figure 13 in which the two LC oscillators correspond to the cavity modes: two superconducting islands, each protruding into one cavity and coupling to the electric field of the resonant mode, are connected by a Josephson junction (so-called bridge configuration): such setup can generate the  $\Phi^4$ -interaction for a CZ gate as will be described in section 3.4, while more involved circuits could be used to engineer an effective  $\Phi^3$ -interaction for the same purpose, see section 3.3.

#### 3.1. Coupling with regular qubits

To prepare a logical GKP state or to realize single-qubit gates, one can employ an interaction with a regular qubit in which the state of regular qubit controls the application of a displacement on the GKP mode as in figure 1. If the regular qubit is realized by an anharmonic oscillator such as a transmon qubit, this then requires the engineering of a tunable qubit controlled-displacement interaction of the form  $b^\dagger b \hat{q}_\theta$  with  $\hat{q}_\theta = \cos(\theta)\hat{q} + \sin(\theta)\hat{p}$  so that  $b^\dagger b$  acts as Pauli  $Z$  in the regular qubit subspace.

As mentioned earlier, a common interaction between an off-resonantly coupled transmon qubit and cavity mode is the dispersive or cross-Kerr interaction of the form  $-\chi a^\dagger a b^\dagger b \rightarrow -\chi a^\dagger a Z/2$ . This interaction realizes a qubit controlled-rotation which can be converted, in principle, to a qubit controlled-displacement, using additional displacements and qubit-flips as follows. Since  $e^{i\theta a^\dagger a Z} D(\alpha) e^{-i\theta a^\dagger a Z} = D(\alpha e^{i\theta Z})$ , choosing  $\theta = \pi/2$  gives the displacement  $D(\alpha)$  when  $Z = 1$  and the displacement  $D(-\alpha)$  when  $Z = -1$ . Thus this sequence of gates does what is needed. To realize  $e^{i\theta a^\dagger a Z}$ , one can simply conjugate the interaction  $e^{-i\theta a^\dagger a Z}$  by  $\pi$ -bit-flips on the qubit, so that we can do all 3 gates in the decomposition of  $D(\alpha e^{i\theta Z})$ .

Note that the strength of the controlled-displacement  $\alpha$  only depends on the strength of the uncontrolled displacement (which can easily be made very strong in  $O(10)$  ns). If the entire controlled-displacement is to be done in, say 50 ns, it requires two rotations each with time  $t = \pi/\chi = 20$  ns, or  $\frac{\chi}{2\pi} = 25$  MHz. This realization thus requires making  $\chi$  tunable, i.e. when the transmon qubit is to be measured or prepared, it is important for the cross-Kerr interaction to be ‘off’, as it induces rotations on the GKP grid state which dephase the state in the Fock basis. However, there is a limit to the on–off ratio of  $\chi$  obtained by flux-tuning of the transmon qubit, in particular if  $\chi$  is flux-tuned to be



**Figure 13.** Electric circuit representing a typical capacitive coupling between two LC oscillators, modeling cavity modes, and a (possibly flux-tunable) transmon qubit.

stronger, then the resonator becomes more anharmonic as well, see [61] and equation (37).

Instead of needing a tunable interaction, reference [42] realized a qubit controlled-displacement in  $1.2 \mu\text{s}$ , using a very weak dispersive coupling  $\frac{\chi}{2\pi} = 28 \text{ kHz}$  between transmon qubit and cavity mode. This weak coupling obviates the need for a tunable interaction, but would make for a very slow gate when using the method described in the previous paragraph. The idea in [42] is to realize a qubit controlled-displacement by temporally displacing the cavity mode to states with  $|\beta|^2 = 320$  photons, so that even a small qubit-induced cavity rotation can have a large effect. The scheme is best understood by using a displacement frame, see appendix C.1, on a cavity-driven dispersive shift Hamiltonian  $H = -\frac{\chi}{2} a^\dagger a Z + i\mathcal{E}(t)a^\dagger - i\mathcal{E}^*(t)a$ . The displacement frame shift shows that the dynamics is due to an effective Hamiltonian

$$\tilde{H}(t) = -\frac{\chi}{2} a^\dagger a Z - \frac{\chi}{2} |\beta(t)|^2 Z - \frac{\chi}{2} (a\beta^*(t) + a^\dagger \beta(t)) Z, \quad (32)$$

where  $\beta(t) = \int_0^t dt' \mathcal{E}(t')$ . The effect of the last term in this Hamiltonian after a time  $T$  (taking  $\beta = \beta^*$  for simplicity) is the qubit controlled-displacement  $\frac{\chi}{\sqrt{2}} \int_0^T dt \int_0^t dt' \mathcal{E}(t') |\hat{q} Z$ . In order to cancel the qubit controlled-rotation (first term) the qubit state is flipped midway in the interval  $T$ , requiring that the cavity displacement direction is also inverted midway, i.e.  $\beta \rightarrow -\beta$ . We see that in this realization the applied displacement power and the dispersive shift  $\chi$  together determine the strength of the qubit controlled-displacement.

We can ask how to improve on the execution of the qubit controlled-displacement gate and the subsequent qubit measurement, where improvement means a faster as well as more reliable execution. As for the realization in [42] one may worry that the large displacements of the state in phase-space during the execution of the gate lead to errors on the GKP state. Even though the cavity has a long lifetime (single-photon life-time  $T = 245 \mu\text{s}$  in [42]), the logical operator of a displaced GKP state takes on a complex oscillatory value in time due to photon loss, see equation (27). It is desirable to shorten the duration of the transmon qubit measurement (700 ns in [42]), but it is hard to make the dispersive read-out of the qubit via a read-out resonator very fast. For example, the measurement pulse followed by active read-out resonator depletion is  $O(600)$  ns in [62] and  $O(250)$  ns (including resonator occupancy) in [63]. Replacing the qubit measurement by feedback and disentangling [43] requiring other controlled-displacements can only lead to a shorter overall preparation time if the duration of such controlled-displacement can be shortened from what was achieved in [42]. Note that the replacement of measurement by coherent interactions could also be done for the GKP qubit rotation in figure 3.

Instead of a transmon qubit as ancilla, one may consider a different qubit such as fluxonium [64], again dispersively coupled to the 3D cavity mode. Advantages of a fluxonium qubit are its long coherence and larger anharmonicity leading to lower leakage [65], equally fast-single qubit gate operations ( $O(10)$  ns) as well as potentially very fast and powerful qubit measurement, see e.g. the GrAl-based fluxonium qubit in [66–68]. In addition, flux-tuning fluxonium may give a strongly-tunable dispersive shift  $\chi$  [64, 69], without the unwanted side-effect of strengthening the cavity anharmonicity.

Another proposal is to use a noise-biased cat qubit to measure the stabilizer displacements of a GKP qubit [20] using a tunable beam-splitter interaction between the two cavity modes of the form  $H = g(t)ab^\dagger + g^*(t)a^\dagger b$ , as argued in section 2.2. To use the interaction, we thus imagine first preparing the cat qubit in  $|C_\alpha^+\rangle$  (by starting in the vacuum state  $|0\rangle$  and turning on the pump), then activate the tunable beam-splitter, and measure the noise-biased cat in the  $|C_\alpha^\pm\rangle$  basis or employ feedback and disentangling via qubit controlled-displacements [43].

### 3.2. Logical GKP measurement

How does one determine whether a GKP state is  $|\bar{0}\rangle$  or  $|\bar{1}\rangle$ , that is, realize a logical  $\bar{Z}$ -measurement? Such logical measurement may be completely destructive, but is desired to have high-fidelity, hence be fault-tolerant in its implementation, meaning that the outcome is insensitive to imperfections in the state. Even though one can measure a logical displacement, i.e.  $\bar{X}$ ,  $\bar{Z}$  or  $\bar{Y}$ , using a single ancilla qubit as was done in references [42, 50], such measurement has an intrinsic probability of error on an approximate GKP state. For example, measuring  $\bar{Z}$  on  $|\bar{0}\rangle$  does not give outcome 1, since the state is not a perfect eigenstate of  $\bar{Z}$  but obeys equation (19). If we assume that this measurement circuit is otherwise perfect and is applied to  $\mathfrak{F}|\bar{b}\rangle$  with  $b = 0, 1$ , the probability for the ancilla qubit to be measured as  $\pm$  equals  $\mathbb{P}(\pm|\bar{b}) = \frac{1}{4} \frac{\langle \bar{b}|\mathfrak{F}^\dagger(I \pm \bar{Z})(I \pm \bar{Z})\mathfrak{F}|\bar{b}\rangle}{\langle \bar{b}|\mathfrak{F}^\dagger\mathfrak{F}|\bar{b}\rangle} \approx \frac{1}{2}(1 \pm (-1)^b e^{-\pi\Delta^2/4})$ , using equation (19). The upshot is that the ancilla measurement is flipped with symmetric error probability  $q = \frac{1}{2}(1 - e^{-\pi\Delta^2/4})$  which goes to 0 when  $\Delta \rightarrow 0$ . At, say,  $\Delta = 0.3$ , this readout error probability  $q$  is about 3.4% and much larger than the probability for an incorrect  $\bar{Z}$ -outcome through the ideal homodyne measurement given in equation (18). Some repetition of the controlled-displacement circuit with the ancilla qubit and taking a majority vote of the answers could bring down the error probability  $q$  at the price of more time and possibly additional feedback error.

A target for future work could be to achieve an improved logical GKP qubit measurement by releasing the GKP state from a superconducting cavity via a switch-release mechanism [70] (taking  $O(1) \mu\text{s}$  in time in [70]) into a transmission line and then enact phase-sensitive amplification (e.g. squeezing) so as to measure one quadrature, say  $\hat{q}$ , with no further added noise. After calibration of the measurement using  $\Delta$ -squeezed displaced states and their targeted measurement outcomes, the measurement could proceed by determining whether the amplified signal corresponds to a  $\hat{q}$  which is closer to an even ( $\rightarrow$ outcome 0) or odd multiple ( $\rightarrow$ outcome 1) of  $\sqrt{\pi}$ . Photon loss in this process may be expected to be a dominant source of noise. To get an estimate of the error rate in the presence of photon loss, we can compute equation (18) for a state at  $\Delta = 0.3$  undergoing photon loss as in equation (25) with  $\kappa t = 0.1$  (so that a coherent state loses  $1 - \exp(-\kappa t) \approx 10\%$  of its intensity), giving  $\mathbb{P}(\bar{Z} = 1|\mathfrak{F}|\bar{0}\rangle, \kappa t = 0.1) = 99.5\%$ . For  $\kappa t = 0.5$  this measurement success probability is already down to 80%.

### 3.3. GKP CZ gate via three-wave mixing

In this section we describe how one could realize the CZ interaction between two GKP modes via a three-wave mixing element which is activated by applying a (strong) microwave pump tone to a coupler mode, see e.g. [71]. An example of pure three-wave mixing used for broadband parametric amplification is the Josephson-ring modulator circuit [72, 73].

An example of the use of parametrically-activated three-wave mixing is the experiment in [74]: flux-modulation through a coupling Josephson junction (instead of microwave driving) is used to activate a  $a^\dagger b$  coupling between a logical (co-planar microwave) resonator whose state is to be manipulated and an ancilla (co-planar microwave) resonator.

For simplicity, we here assume that the following non-degenerate three-wave mixing Hamiltonian is available:

$$H_{\chi^{(2)}} = \omega_a \left( a^\dagger a + \frac{1}{2} \right) + \omega_b \left( b^\dagger b + \frac{1}{2} \right) + \omega_c \left( c^\dagger c + \frac{1}{2} \right) + \chi^{(2)} (a + a^\dagger) (b + b^\dagger) (c + c^\dagger). \quad (33)$$

Here  $a, b$  are annihilation operators for two GKP oscillators while  $c$  is the annihilation of the pump oscillator. We assume that all frequencies  $\omega_a, \omega_b$  and  $\omega_c$  are sufficiently detuned, so that the  $\chi^{(2)}$  interaction between the modes will approximately time-average away in the rotating wave approximation (RWA) in the absence of any active driving, see the discussion in appendix C. Moving to the rotating frame of the GKP oscillators we have

$$\tilde{H}_{\chi^{(2)}} = \omega_c \left( c^\dagger c + \frac{1}{2} \right) + \chi^{(2)} (ab e^{-i(\omega_a + \omega_b)t} + ab^\dagger e^{-i(\omega_a - \omega_b)t} + \text{h.c.}) (c + c^\dagger). \quad (34)$$

Since there are two time-dependencies involved, we can make all  $\chi^{(2)}$ -interactions resonant by driving the pump mode  $c$  with a two-tone drive, namely at  $\omega_p = \omega_a + \omega_b$  and  $\omega_p = \omega_a - \omega_b$ . Both pump tones will need to be of equal amplitude to get equal contributions from beam-splitting ( $a^\dagger b + ab^\dagger$ ) as well as two-mode squeezing ( $ab + a^\dagger b^\dagger$ ). Assuming that the pump mode is a (fairly) harmonic mode which can be strongly driven, we replace the operator  $c$  by its classical time-dependent expectation value  $\langle c(t) \rangle = \mathcal{E} (e^{-i(\omega_a + \omega_b)t} + e^{-i(\omega_a - \omega_b)t})$  with, say,  $\mathcal{E} \in \mathbb{R}$ . Making a rotating-wave-approximation, appendix C, gives the generating interaction of the CZ gate between modes  $a$  and  $b$ :

$$H_{\text{CZ}} = 2\chi^{(2)} \mathcal{E} \hat{q}_a \hat{q}_b. \quad (35)$$

Changing the phase of the pump tone ( $\mathcal{E}$ ) allows one to realize  $CZ^{-1}$ . If we want to do a GKP CNOT gate via two-tone pump, we cannot start with the interaction in equation (33), but a Hadamard or single-qubit  $R_Y(\pi/2)$  would be required to convert  $\hat{q} \rightarrow \hat{p}$ .

Instead of applying two simultaneous pump tones to get a CZ (and with extra rotations, a CNOT), we could also decompose the CNOT circuit as in figure 4, i.e. a sequence of beam-splitters and single-mode squeezers. When we drive the pump mode at the difference frequency of the modes,  $\omega_p = \omega_a - \omega_b$  in the nondegenerate three-wave mixing Hamiltonian in equation (33), we realize a beam-splitter interaction as the pump photon assists in converting one mode- $a$  photon to a mode- $b$  photon.

Single-mode squeezing can be activated by using a degenerate version of the three-wave mixing element  $\hat{q}_a \hat{q}_b \hat{q}_c$  in equation (33) with a Hamiltonian proportional to  $\hat{q}_a^2 \hat{q}_c$  (or similarly  $\hat{q}_b^2 \hat{q}_c$ ). By applying a pump tone at frequency  $\omega_p = 2\omega_a$ , one activates a squeezing Hamiltonian  $H_{sq}$  on mode  $a$ : we down-convert one pump photon into two mode- $a$  photons and vice-versa.

For superconducting devices the only native non-linear circuit element that we have at our disposal are Josephson junctions which—in their simplest use, without externally applied fluxes—realize a  $U(\Phi) = -E_J \cos(2\pi\Phi/\Phi_0)$  potential interaction. Here the flux variable  $\Phi$  can be expanded as a linear combination of the  $q$  quadratures of the bosonic modes which participate in the junction<sup>10</sup>. Usually, if  $E_J$  is large ( $\gg E_C$ ), we expand this cosine potential around its potential minimum  $\Phi = 0$ , obtaining only interactions which are symmetric under  $\Phi \rightarrow -\Phi$  such as  $\Phi^4$  (while absorbing the  $\Phi^2$ -terms in the quadratic part of the Hamiltonian).

It is clear from the discussion above that it would be desirable to engineer a  $\Phi^3$ -interaction where  $2\pi\Phi/\Phi_0 = \alpha q_a + \beta q_b + \gamma q_c$ . When the three modes have sufficiently different frequencies, we can observe that all terms in this  $\Phi^3$ -interaction, except those proportional to  $a^\dagger a$ ,  $b^\dagger b$  or  $c^\dagger c$ , average out in time, hence the interaction is ‘off’ in the absence of active driving of one of the modes, not inducing any nonlinearity on the modes in this off-state. At the same time, by choosing the pump mode drive frequencies appropriately, we can activate, with the same interaction element, either a squeezer for mode  $a$ , a squeezer for mode  $b$ , a beam-splitter between modes  $a$  and  $b$  or/and a two-mode squeezer between modes  $a$  and  $b$ .

Besides the Josephson ring modulator, another three-wave mixing element, called an SNAIL, has been proposed in [77]: it uses a superconducting (SQUID-like) loop containing an asymmetric array of a few Josephson junctions and external flux is applied through the loop. The effective potential induced by this SNAIL is of the form  $U(\Phi) = c_2\Phi^2 + c_3\Phi^3 + c_4\Phi^4$  with  $c_3 \gg c_4 \neq 0$  where  $\Phi$  is the flux variable expanded around its potential minimum, determined by the external flux  $\Phi_{ext}$ .

A recent paper [78] discusses the circuit-QED engineering required to realize a universal set of gates for continuous-variable computation using GKP states. By flux-modulating the SNAIL, one can activate some of the terms in the three-wave mixing Hamiltonian in equation (33), mimicking the effect of microwave driving of the pump mode. The authors in [78] then use this activation to show for example how to realize an interaction  $\hat{q}^3$ , required to enact the cubic phase gate  $V_\gamma = e^{i\gamma\hat{q}^3}$ .

Another use of a  $\Phi^3$ -coupling for GKP state preparation has been proposed in [49]. In this paper the aim is to produce a tunable opto-mechanical coupling of the form  $b^\dagger b \hat{q}_{GKP}$  between a GKP mode and (harmonic) ancilla mode ( $b$ ) which is initially prepared in a coherent state. Such coupling can be used to prepare the GKP mode into a logical state starting from a vacuum state, similar as the preparation via regular qubits discussed in section 3.1. The idea here is that the frequency of the ancilla oscillator is shifted depending on the value for  $\hat{q}$  of the GKP mode, leading to a  $\hat{q}$ -dependent rotation of the coherent state of the ancilla mode. When the interaction time is chosen so that all  $q = k\sqrt{\pi}$  for  $k \in \mathbb{Z}$  lead to the same rotation of the coherent state, measuring the coherent amplitude realizes an approximate modular measurement of  $\hat{q}$ , resolving the value of  $q \bmod \sqrt{\pi}$ . Such modular measurement of  $\hat{q}$  is equivalent to measuring the eigenvalue phases of  $S_q = \exp(i2\sqrt{\pi}\hat{q})$ . A possible advantage of this method over the coupling with regular qubits is that one gets more information per ancilla mode measurement than 1 bit. In this proposal an externally-applied flux is modulated around a value for which there is an effective third-order  $\Phi^3$ -coupling between the two oscillators while the  $\Phi^4$ -coupling vanishes at this flux setting. Choosing the GKP oscillator at much lower frequency ( $\sim 0.5$  GHz) than the ancilla oscillator ( $\sim 10$  GHz)

<sup>10</sup> A side comment: in circuit-QED we cannot passively get interactions where some cosine potential depends on the  $\hat{q}$ s of some subset of modes and another cosine potential depends on the  $\hat{p}$ s of a subset of modes, since all the variables in the circuit Lagrangian which enter the potential energy—such as the Josephson junction cosine potential energy—commute when promoted to quantum operators, i.e. they will never be conjugated variables. The upshot of this is that it is very hard (see an attempt at [75]) to entirely passively Hamiltonian engineer, say, the toric code checks on a collection of bosonic modes as the essential property of such stabilizer checks is that they either act as  $\hat{q}$  and  $\hat{p}$  on a single mode. An exception would be the simultaneous use of Josephson junction elements ( $\cos(2\pi\Phi/\Phi_0)$ ) and so-called phase-slip elements ( $\cos(\pi Q/e)$ ) for conjugate variables flux  $\Phi$  and charge  $Q$ . Reference [76] shows that one can passively engineer an effective GKP Hamiltonian using a gyrator element  $G/2(\Phi_1 \dots \Phi_2 - \Phi_2 \dots \Phi_1)$  in the Lagrangian of the circuit.

creates an asymmetry so that a term like  $b^\dagger b \hat{q}_{\text{GKP}}$  dominates in the  $\Phi^3$ -interaction and the term is made resonant via flux modulation.

### 3.4. Use of four-wave mixing?

We comment on the use of a  $\Phi^4$ -interaction for realizing the GKP CZ gate. The set-up we have in mind is modeled by the electric circuit in figure 13. Applying circuit-quantization to this circuit leads to a Hamiltonian with three active modes. Due to the coupling between the LC oscillators, each described as a single mode, some hybridization will happen between the bare cavity modes and the transmon coupler mode, and so we will associate annihilation operators  $a$ ,  $b$  and  $c$  with these dressed modes. Due to this hybridization the three dressed modes with annihilation operators  $a$ ,  $b$  and  $c$  will partake in the Josephson junction. This means that for the flux-variable operator  $\hat{\Phi}$  across the Josephson-junction branch, we can write  $2\pi\hat{\Phi}/\Phi_0 = \alpha\hat{q}_a + \beta\hat{q}_b + \gamma\hat{q}_c$  with dimensionless  $\alpha, \beta, \gamma$  modeling the participation of the effective modes in the Josephson junction [61]. Expanding the cosine potential up to fourth-order, and diagonalizing the linear interactions of the Hamiltonian (quadratic in creation and annihilation operators) thus gives rise to three dressed eigenmodes at frequencies  $\tilde{\omega}_a, \tilde{\omega}_b$  and  $\tilde{\omega}_c$ , and we have the Hamiltonian:

$$H = \tilde{\omega}_a \left( a^\dagger a + \frac{1}{2} \right) + \tilde{\omega}_b \left( b^\dagger b + \frac{1}{2} \right) + \tilde{\omega}_c \left( c^\dagger c + \frac{1}{2} \right) - \frac{E_J}{4!} (\alpha\hat{q}_a + \beta\hat{q}_b + \gamma\hat{q}_c)^4. \quad (36)$$

As in the discussion on three-wave mixing we assume that all frequencies  $\tilde{\omega}_a, \tilde{\omega}_b, \tilde{\omega}_c$  are sufficiently different (detuned). If there is no active driving (or flux-modulation), a full RWA approximation, whose accuracy depends on the amount of detuning, will leave only energy-conserving self-Kerr and cross-Kerr terms. In other words, in the off-state, the Hamiltonian is approximately

$$H_{\text{off}} \approx \omega_a \left( a^\dagger a + \frac{1}{2} \right) + \omega_b \left( b^\dagger b + \frac{1}{2} \right) + \omega_c \left( c^\dagger c + \frac{1}{2} \right) - \frac{1}{2} (\chi_{aa}(a^\dagger a)^2 + \chi_{bb}(b^\dagger b)^2 + \chi_{cc}(c^\dagger c)^2) \\ - \chi_{ab} a^\dagger a b^\dagger b - \chi_{ac} a^\dagger a c^\dagger c - \chi_{bc} b^\dagger b c^\dagger c, \quad (37)$$

where  $\chi_{ii'} = 2\sqrt{\chi_{ii}\chi_{i'i}}$  [61]. Here  $\omega_a = \tilde{\omega}_a - \frac{E_J\alpha^4}{24}$  (and similarly for  $\omega_b$  and  $\omega_c$ ) due to rewriting the excitation-conserving terms in  $\hat{q}_a^4$  as a quadratic term  $\propto a^\dagger a$  and the self-Kerr term  $\propto (a^\dagger a)^2$  and  $\chi_{aa} = \frac{E_J\alpha^4}{12}$ . Here we clearly see the advantage of a pure three-wave mixing element over a four-wave mixing element: in the off-state the four-wave mixing element induces unwanted Kerr and cross-Kerr anharmonicities on the GKP storage modes  $a$  and  $b$ .

In the off-state, mode  $c$  is (ideally) in its vacuum state, hence the cross-Kerr interaction with this mode does not contribute. However, if this mode were driven these corrections are relevant and they induce additional cavity rotations. Let us now indeed discuss the effect of applying a drive on mode  $c$ . For this, we expand the fourth-order term in equation (36) which becomes

$$H_{\chi^{(3)}} = -\frac{E_J}{4!} \left[ \binom{4}{2} \times 2\alpha\beta\gamma^2 \hat{q}_a \hat{q}_b \hat{q}_c^2 + \binom{4}{2} \alpha^2 \gamma^2 \hat{q}_a^2 \hat{q}_c^2 + \dots \right]. \quad (38)$$

We can apply a two-tone drive on mode  $c$  at frequency  $\frac{\omega_a + \omega_b}{2}$  and  $\frac{\omega_a - \omega_b}{2}$  with equal amplitudes  $\mathcal{E}$ . Replacing  $\hat{q}_c$  by its time-dependent expectation  $\langle \hat{q}_c(t) \rangle = \frac{\mathcal{E}}{\sqrt{2}} (e^{i\tau\frac{\omega_a + \omega_b}{2}} + e^{i\tau\frac{\omega_a - \omega_b}{2}} + \text{h.c.})$  in equation (38) and going to the rotating frame of all modes  $a$  and  $b$ , we find that a term like  $\hat{q}_a \hat{q}_b \hat{q}_c^2$  leads to a time-independent resonant term proportional to  $\hat{q}_a \hat{q}_b$ . This can be seen as follows. First, note that the signal  $\langle q_c(t) \rangle^2$  only contains frequencies  $\omega_a, \omega_b, \omega_a + \omega_b$  and  $\omega_a - \omega_b$ , all of equal strength. The frequency  $\omega_a + \omega_b$  matches two-mode squeezing ( $ab + \text{h.c.}$ ), while  $\omega_a - \omega_b$  matches beamsplitting ( $ab^\dagger + \text{h.c.}$ ). Besides this, we throw out all time-dependent terms (RWA). In particular we have

- Terms without any  $\hat{q}_c$  are only leading to self-Kerr and cross-Kerr for modes  $a$  and  $b$ .
- Terms with a single  $\hat{q}_c$  or  $\hat{q}_c^3$  are not frequency-matched to become time-independent (as  $\omega_a$  and  $\omega_b$  are sufficiently different).
- Terms  $\hat{q}_c^4$  leads to self-Kerr for mode  $c$ .
- Terms with  $\hat{q}_c^2$  lead to cross-Kerr between modes  $c$  and  $a$  or  $c$  and  $b$ . Note that from a term such as  $\hat{q}_a^2 \hat{q}_c^2$  there is a contribution proportional to  $\mathcal{E}^2 a^\dagger a$  (and similarly  $\hat{q}_b^2 \hat{q}_c^2$  gives  $\mathcal{E}^2 b^\dagger b$ ).

We could also realize the CNOT gate using the beam-splitter and squeezer sequence in figure 4, i.e. we chose a single-tone pump at  $\omega_p = (\omega_a - \omega_b)/2$  for the beam-splitter to let two pump photons assist in converting one mode- $a$  photon to one mode- $b$  photon. Single-mode squeezing can be realized by using the interaction  $\hat{q}_a^2 \hat{q}_c^2$  in equation (38), i.e. we should take  $\omega_p = \omega_a$  so that two pump photons are converted into

two mode- $a$  photons. However, note that this also make the unwanted interaction  $b^\dagger b(a^\dagger c + ac^\dagger)$  (which comes from the  $\hat{q}_b^2 \hat{q}_a \hat{q}_c$  term) resonant, which makes this scheme unattractive.

An important parameter measuring the quality of the CZ gate via this  $\Phi^4$ -interaction is the relative strength of the unwanted Kerr and cross-Kerr terms versus the strength of the two-tone pump-activated wanted interactions  $\hat{q}_a \hat{q}_b$ . In part, this relies on the error contributions due to the rotating wave approximation which should be better quantified theoretically (appendix C). Another contributing factor is the relative strength of the participation parameters  $\alpha, \beta, \gamma$  and the pump strength, namely, without error contribution from the RWA, one has

$$\frac{\|H_{\text{CZ}}(a, b)\|}{\|H_{\text{cross-Kerr}}(a, b)\|} \propto \frac{\mathcal{E}^2 \gamma^2}{\alpha \beta}; \quad (39)$$

Hence, a large  $\gamma^2/\alpha\beta$  is desirable but a large  $\gamma$  also makes mode  $c$  more anharmonic as  $\chi_{cc} \propto \gamma^4$  and this again severely restricts the pump power  $\mathcal{E}$ . These conflicting constraints may make this scheme less suitable in practice.

#### 4. Prospects for a GKP-surface code architecture

In this section we would like to provide a perspective of what it would take to build a surface code architecture based on GKP qubits, point out the challenges in this approach, as well as contrast it with existing efforts to engineer a similar architecture using transmon qubits [79], see section 4.1.

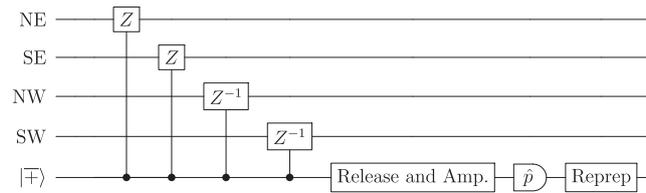
We can partially use the results in reference [48] as a starting point for such GKP-surface code architecture. In this code architecture, there are two layers of protection. On the one hand, each GKP qubit is either stabilized or error corrected individually, reducing a continuous set of (displacement) errors to a mostly discrete set of GKP qubit Pauli errors. On the other hand, the surface code layer is there to suppress the logical error rate of a GKP qubit to values which decrease exponentially with the side length of the surface code lattice.

In reference [48] GKP error correction takes place with ancilla GKP modes using the circuits in figure 7. Note that these circuits can also be implemented via CZ gates, but will then require Hadamard or  $R_Y(\pi/2)$  rotations on the GKP data mode. Interspersed with this GKP error correction, parity checks of the surface code, shown in figure 15, are to be measured in QEC cycles. These circuits are similar as for a regular surface code, except that the underlying qubits are GKP qubits encoded in oscillators, see figure 14 for a Z-check. We use the fact that GKP logical operators are not self-inverse as displacements—and as displacements they obey  $[\bar{X}_1 \bar{X}_2, \bar{Z}_1^{-1} \bar{Z}_2] = 0$ —to measure checks which mutually commute on the entire oscillator space<sup>11</sup>. In this figure the measurement of the GKP ancilla is shown as the release and amplification of the cavity state, followed by a quadrature measurement, as discussed in section 3.2. Such measurement would give useful analog information, but the usefulness of this analog information is challenged by losing photons in the step, nor has it been experimentally realized.

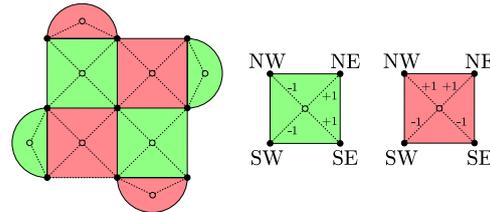
We call this GKP-surface code architecture **Only-SurfaceCode-GKP-Ancilla** and **All-Regular-Qubit-Ancilla**. This set-up would require each GKP qubit oscillator to have CZ capability with 5 other GKP oscillators, namely 4 GKP ancilla oscillators for the surface code and 1 GKP ancilla oscillator for its own error correction.

Reference [48] used a model of Gaussian stochastic displacement noise, equation (24), as an effective, numerically-simulatable, error model for this architecture. The noise channel acts in the different locations: (1) on each GKP qubit prior to GKP error correction and a round of surface-code parity check measurements, (2) prior to the homodyne measurement in figure 7 and (3) prior to the homodyne measurement of the ancilla GKP in the surface code check. Taking the standard deviations from these Gaussian channels to be equal, a threshold standard deviation  $\sigma_{0c} \approx 0.243$  for the toric code was found. Note that this model includes all sources of errors, including finite squeezing and feedback errors, albeit stochastically. Using the conversion  $\Delta^2 = 2\sigma_0^2$ , this gives a threshold of  $\Delta = 0.34$  or 9.3 dB, but the data in figure 12 show this conversion is somewhat too optimistic: using squeezed states with  $\Delta$  gives an error rate which is somewhat worse than a stochastic model with  $\sigma_0 = \Delta/\sqrt{2}$ , so a worst-case threshold estimate using  $\Delta^2 = \sigma_0^2$  would be 12.3 dB. Reference [81] considered a variation on this stochastic noise model—explicitly including error feedback—and applied this noise to a concatenation of the GKP qubit with the surface code. Both [48, 81] used minimum weight matching decoders to find thresholds. Reference [48] identified the defects and the distance function between them following the associated compact-QED

<sup>11</sup> One expects worse error behavior when one measures non-commuting checks outside the ideal code space; for example, this also happens when regular qubits leak [80]. In addition, the maximum likelihood decoding analysis in [48] relies on this choice to map onto a compact-QED model with proper lattice versions of rotations and divergences.



**Figure 14.** A  $Z$ -check parity measurement for the surface-GKP code in figure 15, on oscillators labelled NE–SW (northeast to southwest of ancilla qubit) using CZ and  $CZ^{-1}$  gates as defined in section 2.3.2.  $|+\rangle$  is an approximate  $+1$  eigenstate of  $\bar{X}$  and the GKP stabilizers. Release and amplification (Amp) followed by measurement of the quadrature  $\hat{p}$  is a way to do a logical  $\bar{X}$  measurement and reprep is a unit standing for the reparation of the GKP ancilla state.



**Figure 15.** Distance-3 rotated surface code in its standard surface-17 layout with green  $Z$ -checks and red  $X$ -checks: black filled circles are data qubits, open circles are ancilla qubits, dashed lines are two-qubit interactions. The  $\pm 1$  patterns on each check denotes the use of inverses as in equation (40), so that all checks commute as displacements.

model closely in order to approach exact minimum-weight decoding. Different from this, reference [81] identified the positions where the surface code check outcomes change as defects, but altered the distance function between these defects based on GKP error information.

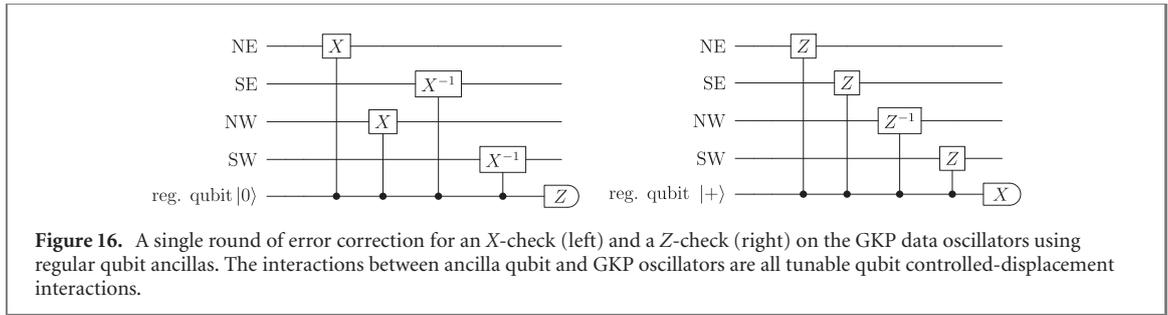
We add two more observations about this scheme. First, when we use stabilizer error correction, such as surface code error correction, on bosonic codes, we need to implement parity check operators which sometimes act like a logical  $X$  on a bosonic qubit, and sometimes like a logical  $Z$ . For GKP qubits this translates into the ability to perform CZ gates as well as CNOT gates. For standard (transmon) qubits, the switch between CZ and CNOT is easily achieved by applying a layer of Hadamard gates between a parity  $X$ -cycle and a parity  $Z$ -cycle. For a GKP qubit encoded into an oscillator with frequency  $f$ , such Hadamard gate seems simple: it constitutes waiting for time  $t = 1/(4f)$ . But since all data qubits have to undergo this Hadamard gate, it implies that the resonant frequencies of the data qubit oscillators (resonant modes of identical 3D cavities) should all be identical, which seems like a narrow target to aim at (although the difference between a simulation-based predicted 3D cavity frequency and the measured frequency can be less than 0.1% [3]).

As alternative to the Hadamard gate one can use  $R_Y(\pi/2)$  gate and  $R_Y(-\pi/2)$  gate, using a regular qubit as in figure 3, to toggle back and forth between  $Z$  and  $X$  error correction, but it costs a lot more hassle and time than doing a  $R_Y(\pi/2)$  on a transmon qubit in  $O(10)$  ns. A second observation is that the use of parametrically-driven three-wave or four-wave mixing as discussed in sections 3.3 and 3.4 could allow for the simultaneous execution of the CZ and  $CZ^{-1}$  gates needed to do a surface-code parity check measurement, as the activation of the CZ or  $(CZ^{-1})$  gate only requires the application of a pump tone to the coupler between each data oscillator and ancilla oscillator (4 couplers in total). The coupling strength of these CZ couplers may not be equally strong, hence the duration of these four pump drives can vary, but an advantage of only driving the coupler mode (instead of the GKP mode) is to enable the simultaneous execution of these commuting gates. Another way of looking at the simultaneously-executed parity check is to observe that a green  $Z$ -check in figure 15 on oscillator NE, SE, NW, SW corresponds to

$$\hat{G} = \hat{q}_{NE} + \hat{q}_{SE} - \hat{q}_{NW} - \hat{q}_{SW}. \quad (40)$$

An interaction Hamiltonian  $H_G = -\hat{q}_A \hat{G}$  applied for time  $t$  has the effect that  $\hat{p}_A \rightarrow \hat{p}_A - \hat{G}t$ , using equation (2) (while for all data oscillators  $i = NE, SE, NW, SW$  participating in the check, we have  $\hat{p}_i \rightarrow \hat{p}_i \pm t\hat{q}_A$ ). Taking  $t = 1$ , we see that by measuring the ancilla quadrature  $\hat{p}_A$ , we measure  $\hat{G}$  modulo even multiples of  $\sqrt{\pi}$  (as  $|+\rangle$  has sharp peaks at  $p_A$  being even multiples of  $\sqrt{\pi}$ ). Thus, if one of the oscillators undergoes a  $\sqrt{\pi}$  shift in  $q$ , the measurement will detect this.

Besides the Steane error correction in figure 7, one can also imagine a more hardware-efficient form of GKP error correction via stabilization using a regular qubit as discussed in section 3.1, regularly



interspersed with parity check measurements for the surface code which, for example, do use a GKP ancilla. The advantage here is that one does not need to prepare and couple the ancillary GKP qubit as in figure 7 (which again requires a regular qubit). In particular, (ancilla) GKP state preparation is time-consuming ( $60 \mu\text{s}$  in [42]) due to requiring slow controlled-displacement gates and slow qubit measurement, and during this process photon loss is affecting the GKP qubit. At the same time, we keep the GKP ancilla for the possibly-less frequent surface code QEC cycle in order to get still analog error information. We refer to this intermediate scheme as **Only-SurfaceCode-GKP-Ancilla**.

Another choice is to use regular qubits to extract both GKP and surface code error information, see the circuits in figure 16. We refer to this scheme as **All-Regular-Qubit-Ancilla**. An advantage of this scheme is that no Hadamards or  $R_Y(\pi/2)$  rotations are needed on GKP modes and tunable controlled-displacement gates are used *throughout*.

The **All-Regular-Qubit-Ancilla** architecture can however be less tolerant towards errors: it might be hard to get below threshold for the surface code, when all error information is obtained through qubits, giving 1 bit of information at the time. We can provide arguments for this by using a simple error model in which we assume that GKP error correction generates an effective phenomenological error model in each surface code QEC cycle and we assume that the surface code QEC cycle is otherwise perfect. We model the effect of GKP error correction as stabilizing an approximate GKP qubit of the form, say,  $\mathfrak{F}|\bar{b}\rangle$  at some  $\Delta$ , besides having a logical error  $b \rightarrow -b$  on top with probability  $p$ . Effectively then, the approximate GKP code states coming into a perfect surface-code parity check circuit as in figure 16 will flip the regular qubit ancilla with some effective error probability  $q$  which depends on  $\Delta$ . We thus map our error model onto a known (phenomenological) surface code error model in which there is an incoming error with probability  $p$  in each QEC round and a measurement error with probability  $q(\Delta)$ . For this model, figure 3 in reference [82] shows the numerically-found below-threshold region and for  $q = p$  the threshold is optimally 3.3% [83]. Reference [82] does not investigate the below-threshold region and its shape for low  $p$ , but it certainly lies within the below-threshold region for the repetition code which is conjectured to have a below-threshold region given by  $H_2(p) + H_2(q) \leq 1$  with  $H_2(p)$  the binary entropy [84].

To estimate  $q(\Delta)$  we can write the probability to measure the ancilla qubit in the X-basis as

$$\mathbb{P}(\pm) = \frac{1}{2} \frac{(\langle 0|\mathfrak{F}^\dagger\mathfrak{F}|0\rangle)^4 \pm \frac{1}{2}[(\langle 0|\mathfrak{F}^\dagger Z\mathfrak{F}|0\rangle)^4 + (\langle 0|\mathfrak{F}^\dagger Z^\dagger\mathfrak{F}|0\rangle)^4]}{(\langle 0|\mathfrak{F}^\dagger\mathfrak{F}|0\rangle)^4},$$

which, using equation (19), approximately gives  $\mathbb{P}(\pm) \approx \frac{1}{2}(1 \pm e^{-\pi\Delta^2})$  and thus  $q(\Delta) = \frac{1}{2}(1 - e^{-\pi\Delta^2})$ . At  $\Delta = 0.3$ , we already have  $q(\Delta) = 12\%$  while  $\Delta = 0.15$  just suffices to get  $q(\Delta) = 3.4\%$ .

The frequency of doing the surface code error correction could also be adapted to the logical decay rate of the stabilized GKP state, e.g.  $275 \mu\text{s}$  in [42], so that the logical qubit error rate between surface code QEC cycles is at least less than 3.3%. It is an open question how to analyze the noise threshold for the **All-Regular-Qubit-Ancilla** architecture for a more elaborate error model.

In this **All-Regular-Qubit-Ancilla** architecture the workhorse is the controlled-displacement gate with the regular qubit and the regular qubit preparation and measurement. The desiderata for these regular qubits are clearly (1) ability to enact a fast and accurate tunable controlled-displacement with a 3D cavity mode, (2) low leakage to higher excited states, (3) long  $T_1$  and  $T_2$ , beyond  $100 \mu\text{s}$ , and (4) fast measurement below  $O(100)$  ns, (5) fast preparation of  $|+\rangle$  and single-qubit gates ( $O(10)$  ns). At first sight, this seems like a wishlist for any good qubit, however it is not necessary to have a high-quality two-qubit gate between these qubits, which is a nontrivial component for the surface code with transmon qubits. Furthermore, the frequency of the 3D cavity GKP modes can be taken to be far different than those of the ancillary qubits and their coupled read-out resonators, possibly leading to easier frequency control and less frequency crowding than in an architecture with only one type of device qubit such as the surface code with transmon qubits [79, 85].

#### 4.1. Comparison: Fock qubit surface code and transmon qubit surface code

Given that we imagine using high- $Q$  3D cavities for qubit storage, we can ask how to compare a GKP encoding with a simple Fock encoding in a surface code architecture, omitting any additional error correction. CZ gates between a 3D-cavity encoded Fock qubit (mode  $a$ ) and an ancilla transmon qubit can be realized by a dispersive coupling  $-\chi a^\dagger a Z/2$ , allowing for the execution of the  $Z$ -check measurement. Similar as for the controlled-displacement in the GKP encoding, tunability of this interaction, for example, by using an intermediate frequency-tunable resonator to vary the coupling strength, is important. This type of parity check, using 1 transmon qubit to read out 4 Fock qubits, is the reverse of using one bosonic mode to read out the parity of four coupled transmon qubits as realized in reference [86]. For the  $X$ -check measurement one requires a CNOT gate with transmon qubit as target, which can be realized by performing a CZ followed and preceded by Hadamard or  $R_Y(\pi/2)$ . Again, similar as in the GKP encoding, these simple single-qubit gates require the use of an ancillary qubit, but arbitrary cavity manipulations through such coupled transmon qubit have been demonstrated in [87], albeit of rather long,  $O(1)$   $\mu$ s, duration, and having some, inevitable, leakage towards the state  $|2\rangle$  or higher.

An engineering effort for making a surface code architecture using 2D transmon qubits is underway at e.g. Google, IBM, TU Delft and ETH Zürich. Besides using an optimized decoder [88], the crucial numbers which determine whether such architecture will be ‘below threshold’ are the quality, leakage [80], time-duration and cross-talk of the two-qubit gate and the duration (and cross-talk) of the qubit measurement versus the dephasing and relaxation time of the qubits. Flux-tunable transmon qubits have recently achieved very good numbers for their two-qubit gates: reference [89] reports on a 99.1% CZ fidelity of 40 ns duration and low leakage 0.1%, while Google’s supremacy experiments [90] have shown the performance of ISWAP-like two-qubit gates on a 54-qubit Sycamore chip with an average error rate of 0.62% and duration  $O(10)$  ns. It is an open question how much further transmon performance numbers, including measurement duration, can be pushed beyond their current values. The use of different superconducting materials [91] can provide new opportunities to lengthen  $T_2$  and  $T_1$  times. Note however that an enhanced  $T_1$  also leads to an enhanced duration of leakage. Frequency crowding and limits on highly-accurate frequency targeting, in particular for non-flux tunable transmons, leading to spurious cross-talk couplings is another challenge in realizing the surface code.

We thus believe that there is plenty of room and, in fact, necessity for developing an alternative surface code architecture in which a data qubit, such as a Fock or GKP qubit, is encoded in a very harmonic mode of high- $Q$  (3D) cavity, while transmon qubits or their next-generation versions such as fluxonium or noise-biased cat qubits, are used as ancilla qubits. If a pump-activated CZ or controlled-displacement gate has high on/off ratio, one expects that spurious couplings between the 3D cavity data modes, due to common coupling to the ancilla qubits, would be well suppressed.

## Acknowledgments

We thank Alessandro Ciani, David DiVincenzo, Ioan Pop and Daniel Weigand for useful feedback and discussions and some help with the figures and the numerics. We acknowledge support from the European Research Council (EQEC, ERC Consolidator Grant No. 682726). CV and BMT acknowledge support from a QuantERA grant for the QCDA consortium. This research was supported in part by Perimeter Institute for Theoretical Physics. Research at Perimeter Institute is supported by the Government of Canada through Industry Canada and by the Province of Ontario through the Ministry of Economic Development & Innovation.

## Appendix A. Fock state representation of GKP grid states

In this appendix we examine the Fock coefficients of approximate GKP states, using the  $\mathcal{D}$ -approximation, and the sensor grid state [44]. We study the asymptotic behavior of these Fock coefficients, showing that the photon number distribution trends along a geometric or thermal distribution. It turns out that these Fock coefficients relate to some interesting nontrivial mathematics.

The theta function with rational parameters  $(a, b) \in \mathbb{Q}^2$ , adopting the notations of [46, 92], is given by

$$\vartheta \begin{bmatrix} a \\ b \end{bmatrix} (z, \tau) = \sum_{k \in \mathbb{Z}} \exp(\pi i \tau (k + a)^2 + 2\pi i (k + a)(z + b)), \quad (\text{A.1})$$

where  $(z, \tau) \in \mathbb{C}$  and  $\text{Im}(\tau) > 0$  ensuring absolute convergence of the series. Some common shorthands are the following

$$\vartheta \begin{bmatrix} 0 \\ 0 \end{bmatrix} (0, ix) = \theta_3(x) = \sum_{k \in \mathbb{Z}} \exp(-\pi x k^2), \tag{A.2}$$

$$\vartheta \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} (0, ix) = \theta_2(x) = \sum_{k \in \mathbb{Z}} \exp(-\pi x(k + 1/2)^2), \tag{A.3}$$

$$\vartheta \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} (0, ix) = \theta_4(x) = \sum_{k \in \mathbb{Z}} \exp(-\pi x k^2 + \pi i k). \tag{A.4}$$

The multidimensional generalization with rational vectors  $(\vec{a}, \vec{b}) \in (\mathbb{Q}^m)^2$  is given by

$$\vartheta \begin{bmatrix} \vec{a} \\ \vec{b} \end{bmatrix} (\vec{z}, \Omega) = \sum_{\vec{k} \in \mathbb{Z}^m} \exp \left[ \pi i (\vec{k} + \vec{a})^T \Omega (\vec{k} + \vec{a}) + 2\pi i (\vec{k} + \vec{a})^T (\vec{z} + \vec{b}) \right], \tag{A.5}$$

where  $\vec{z} \in \mathbb{C}^m$  is a complex vector,  $\Omega \in \mathbb{C}^{m \times m}$  is a complex matrix and  $\text{Im}(\Omega)$  is positive definite which ensures the absolute convergence of the series.

First, it is important to note that one can properly normalize the approximate GKP state  $|j_{\text{approx}}\rangle$ , using the  $\mathfrak{D}$ -approximation defined in equation (15), for the two logical states  $j = 0, 1$ :

$$|j_{\text{approx}}\rangle = \frac{1}{\sqrt{N(\Delta, j)}} \mathfrak{D} |\vec{j}_{\text{ideal}}\rangle \tag{A.6}$$

$$= \frac{1}{\sqrt{N(\Delta, j)}} \exp(-\Delta^2 \hat{n}) \sum_{k \in \mathbb{Z}} |q = (2k + j)\sqrt{\pi}\rangle \tag{A.7}$$

with

$$N(\Delta, j) = \frac{1}{\sqrt{\pi(1-u^2)}} \vartheta \begin{bmatrix} j \\ 2 \\ \vec{1} \\ 0 \end{bmatrix} (\vec{0}, \Omega), \tag{A.8}$$

where

$$u = \exp(-2\Delta^2), \tag{A.9}$$

$$\Omega = \frac{2i}{1-u^2} \begin{pmatrix} 1+u^2 & -2u \\ -2u & 1+u^2 \end{pmatrix}, \tag{A.10}$$

and  $\vec{0}$  and  $\vec{1}$  are the all-zeros and all-ones vectors respectively. Note that there is a bijection between  $\Delta \in (0, \infty)$  and  $u \in [0, 1)$  so we write either  $N(\Delta, j)$  or  $N(u, j)$  depending on convenience.

In order to obtain this expression one can use the position representation of Fock states in terms of Hermite functions  $\Psi_n(q)$  (or Hermite polynomials  $H_n(q)$ ),

$$\langle n|q\rangle = \Psi_n(q) = \frac{1}{\sqrt{2^n \sqrt{\pi} n!}} \exp\left(-\frac{q^2}{2}\right) H_n(q), \tag{A.11}$$

and the so-called Mehler's Hermite polynomial formula:

$$\sum_{n \in \mathbb{N}} u^n \Psi_n(x) \Psi_n(y) = \frac{1}{\sqrt{\pi(1-u^2)}} \exp\left(\frac{(1+u^2)(x^2+y^2) - 4uxy}{2(1-u^2)}\right). \tag{A.12}$$

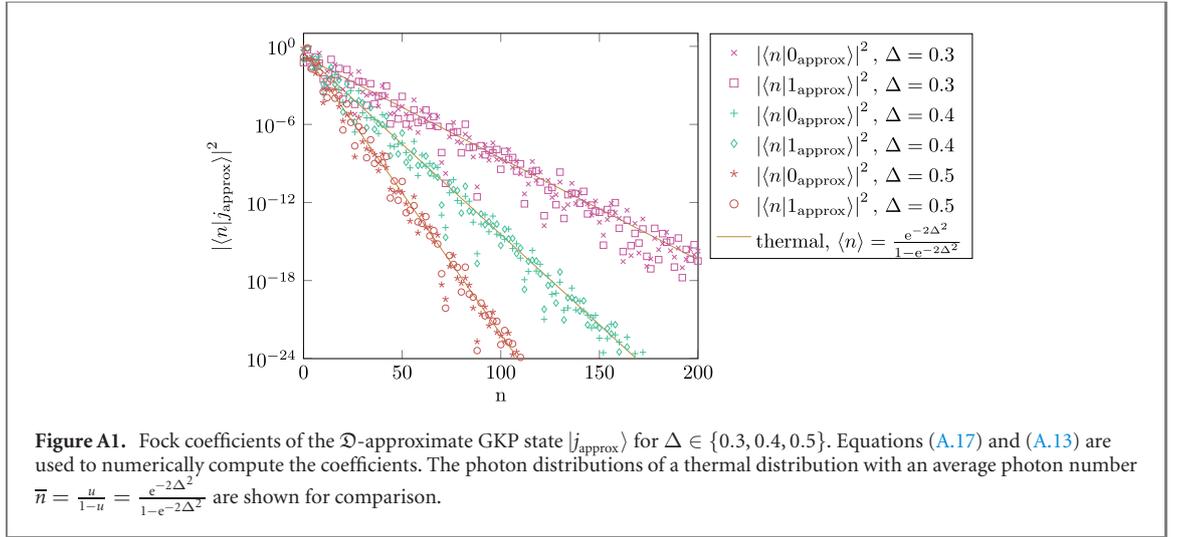
It is possible to rewrite equation (A.8) in the form given in reference [46]

$$N(u, j) = \frac{1}{2\sqrt{\pi}(1+u)} \left( \vartheta \begin{bmatrix} j \\ 2 \\ 0 \end{bmatrix} (0, i8\sigma^2) \vartheta \begin{bmatrix} 0 \\ 0 \end{bmatrix} (0, i\sigma^2/2) + \vartheta \begin{bmatrix} j+1 \\ 2 \\ 0 \end{bmatrix} (0, i8\sigma^2) \vartheta \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} (0, i\sigma^2/2) \right). \tag{A.13}$$

This expression can be recovered using

$$\Omega = i \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 2\sigma^2 & 0 \\ 0 & \frac{1}{2\sigma^2} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \tag{A.14}$$

$$2\sigma^2 = \tanh(\Delta^2) = \frac{1-u}{1+u}, \tag{A.15}$$



and the modular transformation

$$\vartheta \begin{bmatrix} j \\ 2 \\ 0 \end{bmatrix} (0, i/x) = \sqrt{x} \vartheta \begin{bmatrix} 0 \\ j \\ 2 \end{bmatrix} (0, ix). \quad (\text{A.16})$$

We can turn to the Fock coefficients

$$c_n(j) = \langle n|j_{\text{approx}}\rangle = \frac{1}{\sqrt{N(\Delta, j)}} e^{-\Delta^2 n} \sum_{k \in \mathbb{Z}} \Psi_n((2k+j)\sqrt{\pi}). \quad (\text{A.17})$$

As the parity of the Hermite function  $\Psi_n(q)$  is that of the parity of  $n$ , the coefficients  $c_n(j)$  vanish for odd  $n$ . We can use that for  $q \geq 0$ , see [93],

$$\Psi_{2n}(q) = \sqrt{\frac{4^n}{\sqrt{\pi}(2n)!}} \frac{d^n}{dz^n} \left( \frac{1}{\sqrt{1+z}} \exp\left(-\frac{q^2}{2} \cdot \frac{1-z}{1+z}\right) \right) \Bigg|_{z=0}, \quad (\text{A.18})$$

with  $q = (2k+j)\sqrt{\pi}$ . This implies that we have

$$c_{2n}(j) = \sqrt{\frac{4^n}{N(u, j)\sqrt{\pi}(2n)!}} u^n \frac{d^n}{dz^n} \left[ \frac{1}{\sqrt{1+z}} \theta_{3-j} \left( 2 \frac{1-z}{1+z} \right) \right] \Bigg|_{z=0}. \quad (\text{A.19})$$

This gives a somewhat concise expression although it is not directly useful. In order to numerically evaluate the coefficients for example, equation (A.17) is more convenient as the Hermite functions,  $\Psi_n(q)$ , have support essentially within  $[-2\sqrt{n}, 2\sqrt{n}]$  and are easily computed recursively using

$$\Psi_0(q) = \pi^{-1/4} e^{-\frac{q^2}{2}}, \quad (\text{A.20})$$

$$\Psi_n(q) = q \sqrt{\frac{2}{n}} \Psi_{n-1}(q) - \sqrt{\frac{n-1}{n}} \Psi_{n-2}(q). \quad (\text{A.21})$$

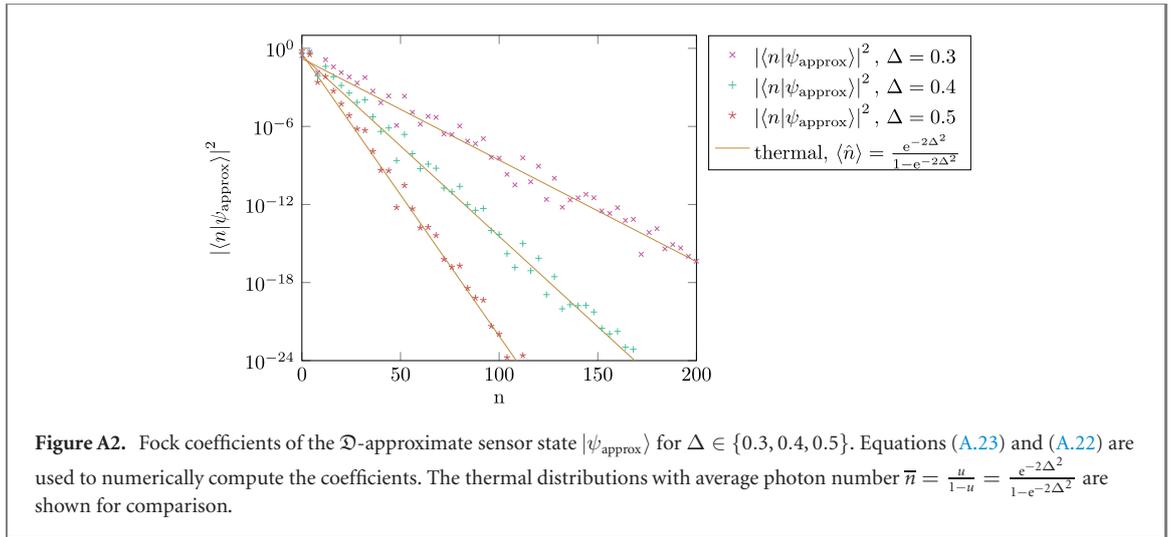
This is used together with equation (A.13) to plot the coefficients in figure A1.

The case of the Fock representation of the so-called sensor state [44] which is an approximate eigenstate of  $e^{i\sqrt{2\pi}\hat{q}}$  and  $e^{i\sqrt{2\pi}\hat{p}}$  is very similar. For the perfect sensor state  $|\psi_{\text{ideal}}\rangle$  the wavefunction in  $q$  is a sum over  $\delta$ -functions at integer multiples of  $\sqrt{2\pi}$ , and similarly the wavefunction in  $p$  is a sum over  $\delta$ -functions at integer multiples of  $\sqrt{2\pi}$ .

The approximate state is  $|\psi_{\text{approx}}\rangle = \frac{1}{\sqrt{N(\Delta)}} \mathfrak{D} |\psi_{\text{ideal}}\rangle$  with normalization

$$N(\Delta) = \frac{1}{\sqrt{\pi(1-u^2)}} \vartheta \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \left( \vec{0}, \frac{1}{2}\Omega \right), \quad (\text{A.22})$$

where  $u$  and  $\Omega$  are the same as defined in equations (A.9) and (A.10). This can be obtained in a similar way and we can also write  $N(u)$  instead of  $N(\Delta)$  when convenient. The sensor state  $|\psi_{\text{approx}}\rangle$  is an eigenstate of



$\exp(i\pi a^\dagger a/2)$ , hence the photon number is 0 mod 4 for this state [44]. We now have  $c_n = \langle n|\psi_{\text{approx}}\rangle$  with  $c_{2n+1} = 0$  and

$$c_{2n} = \frac{1}{\sqrt{N(\Delta)}} e^{-2\Delta^2 n} \sum_{k \in \mathbb{Z}} \Psi_{2n}(k\sqrt{2\pi}). \quad (\text{A.23})$$

Reference [93] derives an expression analogous to equation (A.19) for the sum on the right-hand side:

$$\sum_{k \in \mathbb{Z}} \Psi_{2n}(k\sqrt{2\pi}) = \theta_3(1) \sqrt{\frac{4^n \Phi^n}{\sqrt{\pi}(2n)!}} d(n/2), \quad n \text{ even} \quad (\text{A.24})$$

$$\sum_{k \in \mathbb{Z}} \Psi_{2n}(k\sqrt{2\pi}) = 0, \quad n \text{ odd},$$

with  $\Phi = \frac{\Gamma(\frac{1}{4})^8}{128\pi^4}$ , with  $\Gamma(\cdot)$  the Euler gamma function, and  $\{d(n)\}_{n=0}^\infty$  is a particular integer sequence studied in [93] which is directly related to the derivatives of  $\theta_3(x)$  at  $x = 1$  by

$$\frac{1}{\sqrt{1+z}} \theta_3\left(\frac{1-z}{1+z}\right) = \theta_3(1) \sum_{n=0}^\infty \frac{d(n)}{(2n)!} \Phi^n z^{2n}. \quad (\text{A.25})$$

We show the Fock coefficients for the sensor state in figure A2. Note that the sign of the integers  $d(n)$  is the same as the sign of  $c_{4n}$  through equations (A.23) and (A.24). We can write the sensor state  $|\psi_{\text{ideal}}\rangle \propto \sum_{t,s \in \mathbb{Z}} e^{-i\sqrt{2\pi}\hat{p}t} e^{i\sqrt{2\pi}\hat{q}s} |0\rangle$  where  $|0\rangle$  is the vacuum state. Using that  $e^{i\sqrt{2\pi}\hat{q}} = D(i\sqrt{\pi})$ ,  $e^{-i\sqrt{2\pi}\hat{p}} = D(\sqrt{\pi})$  and  $D(\alpha)|\beta\rangle = e^{i\text{Im}(\alpha\beta^*)} |\alpha + \beta\rangle$ ,  $\langle n|\alpha\rangle = e^{-\frac{|a|^2}{2}} \frac{\alpha^n}{\sqrt{n!}}$ , we have

$$\text{sign}(d(n)) = \text{sign}(c_{4n}) = \text{sign}\left(\sum_{t \in \mathbb{Z}, s \in \mathbb{Z}: z=t+is} e^{-\frac{1}{2}\pi|z|^2} z^{4n} (-1)^{ts}\right). \quad (\text{A.26})$$

We want now to derive the asymptotic behavior of the coefficients. This can be done by expressing the normalization of the states by an equality which has to hold for all  $\Delta \in (0, \infty)$  or equivalently all  $u \in [0, 1)$ . In the case of the sensor state,  $\sum_{n \in \mathbb{N}} |c_{4n}|^2 = 1$  implies:

$$N(\Delta) = \frac{\theta_3(1)^2}{\sqrt{\pi}} \sum_{n \in \mathbb{N}} \frac{(2\sqrt{\Phi}u)^{4n}}{(4n)!} d^2(n), \quad (\text{A.27})$$

which is of the form:

$$N(u) = \frac{A}{\sqrt{\pi}} \sum_{n \in \mathbb{N}} a_{4n} u^{4n}, \quad (\text{A.28})$$

where the constant  $A$  and the sequence  $a_{4n}$ , both independent of  $u$ , are defined by equation (A.27). Similarly for the approximate GKP states we have:

$$N(u, j) = \frac{1}{\sqrt{\pi}} \sum_{n \in \mathbb{N}} \frac{(2u)^{2n}}{(2n)!} \left( \frac{d^n}{dz^n} \left[ \frac{1}{\sqrt{1+z}} \theta_{3-j}\left(2\frac{1-z}{1+z}\right) \right] \Big|_{z=0} \right)^2, \quad (\text{A.29})$$

which is of the same form:

$$N(u, j) = \frac{B(j)}{\sqrt{\pi}} \sum_{n \in \mathbb{N}} b_{2n}(j) u^{2n}, \quad (\text{A.30})$$

for some other constant  $B(j)$  and sequence  $b_{2n}(j)$ , also both independent of  $u$ , defined by equation (A.29).

Seen as complex functions of  $u \in \mathbb{C}$ ,  $N(u, j)$  and  $N(u)$  are both analytic and have a convergence radius of 1. Their behavior for  $u \rightarrow 1$  can be obtained (see also [46]) using equation (A.13) and the fact that

$$\theta_{3-j}(x) \underset{x \rightarrow 0}{\sim} \frac{1}{\sqrt{x}}, \quad \theta_4(x) \underset{x \rightarrow 0}{\sim} \frac{2 e^{-\pi/4x}}{\sqrt{x}}. \quad (\text{A.31})$$

This gives

$$\frac{A}{\sqrt{\pi}} \sum_{n \in \mathbb{N}} a_{4n} u^{4n} = N(u) \underset{u \rightarrow 1}{\sim} \frac{1}{2\sqrt{\pi}(1-u)}, \quad (\text{A.32})$$

$$\frac{B(j)}{\sqrt{\pi}} \sum_{n \in \mathbb{N}} b_{2n}(j) u^{2n} = N(u, j) \underset{u \rightarrow 1}{\sim} \frac{1}{2\sqrt{\pi}(1-u)}. \quad (\text{A.33})$$

We then apply a transfer theorem, see [94], to deduce the asymptotic behavior of the  $a_{4n}$  and  $b_{2n}(j)$  sequences. More precisely they both converge to some finite value

$$a_{4n} \underset{n \rightarrow \infty}{\sim} \frac{1}{2A}, \quad b_{2n}(j) \underset{n \rightarrow \infty}{\sim} \frac{1}{2B(j)}. \quad (\text{A.34})$$

In turn this gives the asymptotic behavior of the Fock coefficients:

$$|\langle 4n | \psi_{\text{approx}} \rangle|^2 \underset{n \rightarrow \infty}{\sim} \frac{u^{4n}}{2\sqrt{\pi}N(u)}, \quad |\langle 2n | j_{\text{approx}} \rangle|^2 \underset{n \rightarrow \infty}{\sim} \frac{u^{2n}}{2\sqrt{\pi}N(u, j)}. \quad (\text{A.35})$$

In both cases the coefficients are asymptotically equivalent to a geometric or thermal distribution which is usually parametrized by the average photon number  $\bar{n}$  as follows

$$p_n^{\text{thermal}} = \frac{1}{\bar{n} + 1} \left( \frac{\bar{n}}{\bar{n} + 1} \right)^n. \quad (\text{A.36})$$

We can therefore deduce the average photon number of the equivalent asymptotic thermal distribution to the approximate GKP and sensor states using

$$u = \frac{\bar{n}}{\bar{n} + 1} \Rightarrow \bar{n} = \frac{u}{1-u} = \frac{e^{-2\Delta^2}}{1 - e^{-2\Delta^2}}. \quad (\text{A.37})$$

These thermal distributions for the different  $\Delta$ 's considered are also shown in figures A1 and A2. One can see a good agreement of the general trend although oscillations of the order of the probabilities themselves persist. Note also that for small  $\Delta$ ,  $\bar{n} \sim \frac{1}{2\Delta^2}$ , consistent with approximate expressions derived in other literature (see main text).

## Appendix B. Decoders for repeated GKP error correction with finite squeezing

For computational, simulation efficiency (as well as our formulation of a classical decoder) we use the approximation  $\mathfrak{F}_V$  for the GKP ancillas. This form of the states can be viewed as applying a reverse Villain approximation to the approximate  $\mathfrak{F}$ -states in equation (17). The reverse Villain approximation, which is tight for  $a \geq 4\pi^2$ , reads

$$\sum_{n \in \mathbb{Z}} e^{-an^2 + bn} \approx e^{-\frac{a}{2\pi^2}} e^{\frac{b^2}{4a} + \frac{a}{2\pi^2} \cos(\frac{\pi b}{a})}. \quad (\text{B.1})$$

Using equation (B.1) and  $\frac{\Delta^2}{\Delta^4 + 1} \approx \Delta^2$ , we then have

$$\begin{aligned} \mathfrak{F}_V|\bar{0}\rangle &= \int_{\mathbb{R}} dq \psi^0(q)|q\rangle, \quad \psi^0(q) \propto e^{-\frac{\Delta^2}{2}q^2} e^{\frac{1}{\pi\Delta^2} \cos(\sqrt{\pi}q)}, \\ \mathfrak{F}_V|\bar{+}\rangle &= \int_{\mathbb{R}} dq \psi^+(q)|q\rangle, \quad \psi^+(q) \propto e^{-\frac{\Delta^2}{2}q^2} e^{\frac{1}{4\pi\Delta^2} \cos(2\sqrt{\pi}q)}. \end{aligned} \quad (\text{B.2})$$

Using these ancillas, the Green's function for  $M$  rounds of error correction (*without* active feedback between rounds), with outcomes denoted as  $M$ -dimensional vectors  $\vec{p}$  and  $\vec{q}$ , can be written as

$$G^M(q_{\text{out}} \leftarrow q_{\text{in}} | \vec{q}, \vec{p}) = \int_{\mathbb{R}^{M+1}} d\vec{q} \exp(-S[\vec{q} | \vec{q}, \vec{p}]) \delta(q_{\text{out}} - q_M) \delta(q_{\text{in}} - q_0), \quad (\text{B.3})$$

with one-dimensional 'action'  $S[\vec{q} | \vec{q}, \vec{p}]$ :

$$\begin{aligned} S[\vec{q} | \vec{q}, \vec{p}] &= \sum_{m=1}^M \frac{\Delta^2}{2} (q_m - q_{m-1})^2 - \frac{1}{\Delta^2 \pi} \cos \sqrt{\pi} (q_m - q_{m-1}) - \sum_{m=1}^M \frac{1}{4\pi \Delta^2} \cos 2\sqrt{\pi} (q_m - q_m) \\ &+ \frac{\Delta^2}{2} (q_m - q_m)^2 + i \sum_{m=1}^M p_m (q_m - q_{m-1}). \end{aligned} \quad (\text{B.4})$$

We can readily interpret the first line in the last equation as a kinetic energy term  $T$ , generating dynamics in the position variable, while the second line is a potential energy term  $-U$ , pinning the position variable to the measured values  $q_m$ . We note that in the potential energy, the  $\cos 2\sqrt{\pi} (q_m - q_m)$  term is dominant when  $\Delta$  is small and hence we can omit the wide parabolic potential proportional to  $\Delta^2$ . Similarly, for the kinetic energy, when  $\Delta$  is small, expanding  $\cos \sqrt{\pi} (q_m - q_{m-1})$  gives heavy-mass terms with mass  $\sim \frac{1}{\Delta^2}$  and the light-mass quadratic term  $\sim \Delta^2$  contributes little. We see that aside from the imaginary term on the last line, the action in this approximation is the same as in the Hamiltonian developed for stochastic noise in the reverse Villain approximation in reference [48] identifying  $\sigma_0^2 = \Delta^2$  with  $\sigma_0$  the standard deviation in the Gaussian displacement model in equation (24). Note however that here the dynamics occurs at the level of wave functions, whereas the description in reference [48] took place at the level of the probabilities (for shift errors). The interesting difference lies in the pure imaginary term making the action  $S$  complex. If the path integral is approximated by taking a single 'classical' path which minimized  $\text{Re}(S)$  then this phase factor only gives an additional phase to this path. However, the pure imaginary term contributes a phase to each path so that the total sum of paths  $q_{\text{in}} \rightarrow q_{\text{out}}$  can be different, due to interference, than the case in which these phases are absent. Note that these terms come from the  $Z$ -error correction step with outcome  $p_m$  which puts a feedback error on the data oscillator.

We formulate a decoder which will provide an approximate tracking of the dynamics of the wave function in  $q$  and  $p$  of the encoded state which is classical. This approximation can be viewed, to some extent, as making a classical approximation, i.e. selecting a single optimal classical path of a quantum path integral. We believe that similar ideas could be applied to simplified tracking of the Wigner function for the purpose of decoding: this may be of interest when we want to study the effects of a fuller noise model, which includes, say, photon loss on the cavity mode during ancilla GKP state preparation and photon loss on the ancilla prior to measurement.

The use of a classical approximation<sup>12</sup> to this path-integral is to determine the outgoing wave function without fully calculating the entire evolution by executing the  $M - 1$ -dimensional integral. In this classical approximation we would have

$$\psi_{\text{out}}(q_{\text{out}}) \approx \int dq_{\text{in}} \exp(-S[\vec{q}_{\text{opt}} | \vec{q}, \vec{p}]) \psi_{\text{in}}(q_{\text{in}}). \quad (\text{B.5})$$

We imagine doing a final measurement of  $\hat{q}$  on the outgoing wave function  $\psi_{\text{out}}$  and will be interested in evaluating equation (18). For the classical approximation we have

$$\mathbb{P}(\bar{Z} = (-1)^b) = \int_{I_b} dq |\psi_{\text{out}}(q)|^2 = \int_{I_b} dq \int dq_{\text{in}} \int q'_{\text{in}} e^{-S[\vec{q}_{\text{opt}}(q \leftarrow q_{\text{in}}) | \vec{q}, \vec{p}] - S^*[\vec{q}_{\text{opt}}(q'_{\text{in}} \leftarrow q) | \vec{q}, \vec{p}]} \psi_{\text{in}}(q_{\text{in}}) \psi_{\text{in}}^*(q'_{\text{in}}), \quad (\text{B.6})$$

where we have explicitly indicated how the optimal path depends on the initial position  $q_{\text{in}}$  and the final position  $q$ . We note that the action in equation (B.4) has the property  $S^*[q_0, \dots, q_M | \vec{q}, \vec{p}] = S[q_M, \dots, q_0 | \vec{q}, \vec{p}]$ , i.e. it is the action of the time-reversed path. Thus when  $q'_{\text{in}} = q_{\text{in}}$ , we have  $e^{-S[\vec{q}_{\text{opt}}(q \leftarrow q_{\text{in}}) | \vec{q}, \vec{p}] - S^*[\vec{q}_{\text{opt}}(q'_{\text{in}} \leftarrow q) | \vec{q}, \vec{p}]} = e^{-2\text{Re}(S[\vec{q}_{\text{opt}}(q \leftarrow q_{\text{in}}) | \vec{q}, \vec{p}])}$ . In words: when we take a path along a closed loop from  $q_{\text{in}} \rightarrow q \rightarrow q_{\text{in}}$ , there is no phase accumulation. However, in equation (B.6) there are certainly contributions when  $q_{\text{in}} \neq q'_{\text{in}}$ . For the purpose of developing an efficient classical decoder, we apply a stochastic approximation to equation (B.6), keeping only the diagonal terms i.e.

$$\mathbb{P}_{\text{class}}(b | \psi_{\text{in}}, \vec{q}, \vec{p}) = N \int_{I_b} dq \int dq_{\text{in}} e^{-2\text{Re}(S[\vec{q}_{\text{opt}}(q \leftarrow q_{\text{in}}) | \vec{q}, \vec{p}])} |\psi_{\text{in}}(q_{\text{in}})|^2, \quad (\text{B.7})$$

<sup>12</sup> One can also define a semi-classical approximation by expanding around the classical path and performing the Gaussian integral, but we have not numerically explored these variations on decoding.

where  $N$  is a normalization to make  $\mathbb{P}_{\text{class}}()$  a probability. Since this normalization does not play a role in the use of this expression in decoding, we do not need to determine it. To evaluate equation (B.7), one can generate a  $q$  uniformly at random in the interval  $I_b$  and  $q_{\text{in}}$  is generated according to  $|\psi_{\text{in}}|^2$ . Given  $q$  and  $q_{\text{in}}$ , if we can evaluate the classical path  $\vec{q}_{\text{opt}}$  between these points and hence compute the weight  $e^{-2\text{Re}(S[\vec{q}_{\text{opt}}(q \leftarrow q_{\text{in}})|\vec{q}, \vec{p}])}$  corresponding to this path, then we can stochastically estimate equation (B.7).

However, as was observed in reference [48] this classical-path approach is not computationally simple as the dynamics of the  $q$ -variable can be chaotic due to it taking place in a random potential induced by the measurement outcomes  $\vec{q}, \vec{p}$ . Hence, instead of sampling  $q_{\text{in}}$  and the endpoint  $q$ , we will sample  $q_{\text{in}}$  from  $|\psi_{\text{in}}(q_{\text{in}})|^2$  and then apply a forward minimization technique, similar as in reference [48], on the function  $\text{Re}(S)[.]$  to find an approximately optimal path  $\vec{q}_{\text{opt}}$  given  $q_{\text{in}}$ , leading to a final value for  $q_{\text{out}}$ . We then determine whether  $q_{\text{out}}$  lies in  $I_b$  and repeat to gather statistics. In fact, we observe numerically that this strategy shows the same performance as fixing an initial  $q_{\text{in}} = \underset{q}{\text{argmax}} |\psi_{\text{in}}(q)|^2$  to which we apply the forward minimization. The probability to land in the  $I_b$  interval is denoted as  $\mathbb{P}_{\text{forward}}(b|\psi_{\text{in}}, \vec{q}, \vec{p})$ .

To adapt this forward decoder to the corrective displacement in each round as shown in figure 7, the action in equation (B.4) is modified by substituting  $q_m \rightarrow q_m + \mathbf{q}_m$  for each round (and the addition of a term  $i\mathbf{q}_m \mathbf{p}_m$  which is irrelevant for this decoder).

The memoryless decoder presented in figure 10 in the main text is implemented by decomposing each measurement outcome  $\mathbf{q}_m = l_m^q \sqrt{\pi} + n_m^q 2\sqrt{\pi} + \epsilon_m^q$  and  $\mathbf{p}_m = l_m^p \sqrt{\pi} + n_m^p 2\sqrt{\pi} + \epsilon_m^p$  with  $l_m^{q/p} \in \mathbb{Z}_2$  (yes/no logical shift)  $n_m^{q/p} \in \mathbb{Z}$  (number of stabilizer shifts),  $|\epsilon_m^{q/p}| < \frac{\sqrt{\pi}}{2}$  (minimal shift error) and applying a corrective shift of  $\alpha_m = \frac{-\delta_{q,m} + i\delta_{p,m}}{\sqrt{2}}$  with  $\delta_{q/p,m} = n_m^{q/p} 2\sqrt{\pi} + \epsilon_m^{q/p}$  after the QEC round. Note that this is not the same correction as in figure 7. This correction is motivated by a stochastic-shift error model and ensures that we keep the photon number low by applying an appropriate number of stabilizer displacements while correcting the perceived error.

Let us now further discuss how we test and evaluate our decoders.

*Decoding of repeated GKP error correction.* The experimentalist implementing the repeated rounds of error correction learns the value of  $\vec{p}, \vec{q}$  and the outcome of the final perfect homodyne measurement of  $\hat{q}$ , but in general she should make decoding decisions *without knowing the input state*<sup>13</sup>. However, she can play the game in which she assumes that the input wave state is either  $\psi_{\text{in}}^0(q)$  (or  $\psi_{\text{in}}^1(q)$ ) and then determine whether the dynamics—given her measurement data  $\vec{q}, \vec{p}$ —would lead to a read-out of  $\bar{Z} = 1$  or  $\bar{Z} = -1$ . Naturally this may not be a deterministic process, so she can calculate the probabilities  $\mathbb{P}(0|\psi_{\text{in}}^0, \vec{q}, \vec{p})$  ( $\bar{Z} = 1$ ) and  $\mathbb{P}(1|\psi_{\text{in}}^0, \vec{q}, \vec{p})$  ( $\bar{Z} = -1$ ). This defines a maximum-likelihood full-density matrix decoder. When using this decoder, the experimentalist evaluates the (normalized) Green's function in equation (B.3), given her measurement data. She then flips the final experimentally measured logical outcome bit (modeled as a perfect homodyne measurement of  $\hat{q}$ ) whenever  $\mathbb{P}(1|\psi_{\text{in}}^0, \vec{q}, \vec{p}) > \mathbb{P}(0|\psi_{\text{in}}^0, \vec{q}, \vec{p})$ . As logical error rate estimate for this decoder we take

$$\bar{P}_{\text{MLD}} = \int d\vec{q} \int d\vec{p} \mathbb{P}(\vec{q}, \vec{p}) \min(\mathbb{P}(0|\psi_{\text{in}}^0, \vec{q}, \vec{p}), \mathbb{P}(1|\psi_{\text{in}}^0, \vec{q}, \vec{p})). \quad (\text{B.8})$$

This estimate assumes that the error rate is (roughly) the same when the input wavefunction is  $\psi_{\text{in}}^1(q)$  and that a superposition of such inputs behaves similarly, without logical interference.

For systems of many modes, this decoding method will not typically be efficient (even for a single mode it comes down to a sizable computational effort), hence the goal of decoding is to infer errors without tracking the entire wavefunction. As we argued before, it turns out to be most computationally efficient to use a forward strategy and estimate  $\mathbb{P}_{\text{forward}}(0|\psi_{\text{in}}^0, \vec{q}, \vec{p})$ . With this probability in hand, let  $f(\vec{q}, \vec{p}) = \underset{b}{\text{argmin}} (\mathbb{P}_{\text{forward}}(b|\psi_{\text{in}}^0, \vec{q}, \vec{p}))$  be the indicator bit whether to flip. The logical error estimate that we consider is then

$$\bar{P}_{\text{forward}} = \int d\vec{q} \int d\vec{p} \mathbb{P}(\vec{q}, \vec{p}) [\delta_{f(\vec{q}, \vec{p}), 0} \mathbb{P}(1|\psi_{\text{in}}^0, \vec{q}, \vec{p}) + \delta_{f(\vec{q}, \vec{p}), 1} \mathbb{P}(0|\psi_{\text{in}}^0, \vec{q}, \vec{p})]. \quad (\text{B.9})$$

Again, in using only this estimate, it is assumed (and not *a priori* given) that the evolution of the input state  $\psi_{\text{in}}^1(q)$  or an arbitrary superposition will have the same logical error rate, and that assuming a different input state would not lead the decoder to different conclusions. The logical error probability for the passive decoder which does not use any syndrome measurement data, but declares that output state is the same as input, is defined as

$$\bar{P}_{\text{passive}} = \int d\vec{q} \int d\vec{p} \mathbb{P}(\vec{q}, \vec{p}) \mathbb{P}(1|\psi_{\text{in}}^0, \vec{q}, \vec{p}). \quad (\text{B.10})$$

<sup>13</sup> The reason is that this state may be part of a complicated quantum computation and we assume that we cannot simulate this quantum computation.

Thus, similar as in equation (B.10), the logical error is determined by the probability to get outcome 1 in the final  $q$  measurement. We have numerically estimated the logical errors rates of these three decoders, given in equations (B.8)–(B.10), as a function of  $M$  for  $\Delta = 0.3$  and  $\Delta = 0.4$ , see figure 10 in the main text. The logical error rates for adapted versions of these decoders including corrective displacements after each round are displayed in figure 11.

### Appendix C. Hamiltonian engineering via rotating wave approximations

The goal of this appendix is to discuss the underpinnings and the ‘beyond’ of the commonly-invoked rotating wave approximation of a Hamiltonian of the form  $H = H_0 + V$  with

$$H_0 = \omega_a \left( a^\dagger a + \frac{1}{2} \right) + \omega_b \left( b^\dagger b + \frac{1}{2} \right) + \omega_c \left( c^\dagger c + \frac{1}{2} \right), \quad V = \sum_k \lambda_k \Phi^k, \quad \Phi = \alpha \hat{q}_a + \beta \hat{q}_b + \gamma \hat{q}_c. \quad (\text{C.1})$$

First, in the absence of applying time-dependent drives, the physical basis for an RWA approximation of a  $\Phi^k$ -term can be motivated in at least two ways. One is to move to a rotating frame in which the Hamiltonian remains time-independent, but is a form amenable to Schrieffer–Wolff degenerate perturbation theory so that terms which either (1) contain an unequal number of creation and annihilation operators, and/or (2) contain an equal number of creation and annihilation operators of modes which are sufficiently far detuned, are seen as perturbations to detuned Fock energy levels. For example, in equation (C.1), to argue about the perturbative effect of terms which contain an unequal number of creation and annihilation operators, we observe that they act as off-diagonal elements in the Fock basis of  $H_0$ , changing the number of total excitations. Hence if we view  $H_0$  as block-diagonal with blocks formed by a given total number of excitations in either mode  $a$ ,  $b$  or  $c$ , separated by a gap  $\min(\omega_a, \omega_b, \omega_c)$ , then the effect of such excitation-number changing terms can be examined by Schrieffer–Wolff perturbation theory. In lowest-order perturbation theory, one projects  $V$  onto these blocks labeled by the number of excitations, so that off-diagonal terms have no effect, Kerr and cross-Kerr terms remain, as well as excitation-number preserving terms (e.g.  $a^\dagger b^\dagger c^2$ ). Given the GHz frequencies of the  $\omega_i$ 's versus the relative strength of Josephson-induced coupling, keeping things to this lowest-order is a good approximation and we replace  $V$  by  $V_{\text{eff}}$  which omits these terms which do not preserve the number of excitations. This strictly speaking is the rotating-wave-approximation.

As a next approximation, to handle terms which do not preserve excitation number in any of the *particular* modes, we go to the rotating frame at  $\omega_c$  for all three modes so that  $\tilde{H}_0 = \Delta_{ac}(a^\dagger a + \frac{1}{2}) + \Delta_{bc}(b^\dagger b + \frac{1}{2})$  with  $\Delta_{ic} = \omega_i - \omega_c$  and  $\tilde{V}_{\text{eff}} = V_{\text{eff}}$ . We imagine that  $\tilde{V}_{\text{eff}}$  is expanded in terms which are products of creation and annihilation operators of the modes. The Hamiltonian  $\tilde{H}_0$  has energy eigenspaces  $|x\rangle_a \otimes |y\rangle_b \otimes |\psi\rangle_c$  with Fock states  $x, y = 0, 1, \dots$ , each of which has the degeneracy of the oscillator ( $c$ ) space, and each space is at least  $\min(|\Delta_{ba}|, |\Delta_{ca}|, |\Delta_{bc}|)$  away from another space. The perturbation  $\tilde{V}_{\text{eff}}$  has both diagonal parts with respect to these eigenspaces (e.g. Kerr and cross-Kerr) as well as off-diagonal terms which map between the spaces. In lowest-order perturbation theory, one again projects  $\tilde{V}_{\text{eff}}$  onto these eigenspaces, so that off-diagonal terms have no effect and Kerr and cross-Kerr terms remain. This is then the full (RWA) approximation. To go beyond this, the effect of the off-diagonal terms can be estimated in second or higher-order perturbation theory to obtain an effective Hamiltonian  $H_{\text{eff}}$  which is diagonal in the Fock basis using Schrieffer–Wolff perturbation theory (see e.g. [95–98]). The spectrum of  $H_{\text{eff}}$  approximates that of  $H$  when we are in the perturbative regime with  $\min(\Delta_{ij}) \ll \|V_{\text{eff}}\|/2$ . In second and higher-order perturbation theory, i.e. beyond RWA,  $UH_{\text{eff}}U^\dagger$  provides an approximation for  $\tilde{H}$  where the unitary  $U$  is the perturbatively expanded Schrieffer–Wolff transformation which provides a correction to the Fock eigenbasis. To get an approximation to the original Hamiltonian, one would finally rotate back to the original frame.

An alternative analysis could be based on the Magnus expansion: to apply this, we move to a rotating frame for each mode at its own frequency for the Hamiltonian in equation (C.1) such that the previously mentioned off-diagonal terms become rapidly time-dependent: as a consequence their effect averages out over sufficiently long times. To observe the dynamics in this rotating frame, we consider the Schrödinger equation for the state  $|\tilde{\psi}(t)\rangle = U_0^\dagger(t)|\psi(t)\rangle$  with  $U_0 = \exp(-iH_0t)$ , while  $|\psi(t)\rangle$  obeys the Schrödinger equation with  $H$ . The state  $|\tilde{\psi}(t)\rangle$  will evolve according to the rotating-frame (or interaction-frame)

Hamiltonian  $\tilde{H} = U_0^\dagger H U_0 + i \frac{dU_0^\dagger}{dt} U_0$ . For example, when  $V = \lambda[a^2 b^\dagger c^\dagger + \text{h.c.}]$ , one has

$$\tilde{H} = \lambda[a^2 b^\dagger c^\dagger e^{i(2\omega_a - \omega_b - \omega_c)t} + \text{h.c.}]. \quad (\text{C.2})$$

For time-dependent Hamiltonians, the Magnus expansion [99] or Magnus–Taylor expansion [100] then forms a convenient representation of the effective dynamics. For the time-evolution (in the rotating frame)

from an initial time  $t = 0$  to a final time  $T$ , the Magnus expansion reads

$$U(T, 0) = \mathcal{T} \exp \left( -i \int_0^T dt \tilde{H}(t) \right) = \exp(-i\bar{H}(T)), \quad (\text{C.3})$$

with  $\bar{H}(T) = \sum_{k=1}^{\infty} \bar{H}^{(k)}(T)$  and the first two terms equal

$$\bar{H}^{(1)}(T) = \int_0^T dt \tilde{H}(t), \quad \bar{H}^{(2)}(T) = -\frac{1}{2} \int_0^T dt' \int_0^{t'} dt [\tilde{H}(t'), \tilde{H}(t)],$$

while for  $k \geq 2$  one gets increasingly higher-order commutators [99]. For the Hamiltonian in equation (C.1), clearly terms diagonal in the Fock basis, are time-independent and will be present in  $\bar{H}^{(1)}(T)$ . As an example of a rotating term, consider a simple time-dependent Hamiltonian  $\tilde{V}(t) = A \exp(i\Delta t) + \text{h.c.}$  where  $A$  is some product of creation and annihilation of some modes and  $\Delta$  is a detuning. For this Hamiltonian, the strength of this lowest-order term decays inversely with  $\Delta$ , i.e.

$$\|\bar{V}^{(1)}(T)\| \leq \frac{2}{\Delta} (\|A\| + \|A^\dagger\|) \sin(\Delta T/2) \leq \frac{4}{\Delta} \|A\|, \quad (\text{C.4})$$

scaling with the perturbative parameter  $\|A\|/\Delta$ .

In general, a sufficient condition for the convergence of the Magnus expansion is that  $\int_0^T dt \|\tilde{H}(t)\| < \pi$ , but, similar as in degenerate perturbation theory, it can still serve as useful asymptotic series expansion.

We thus see that both methods give us perturbative expansions whose validity depends on the strength of the perturbative parameter.

Let us now discuss the case when we actively drive one of the modes, say mode  $c$ . As before one can apply an RWA, dropping terms which do not preserve total excitation number to obtain an effective  $V_{\text{eff}}$ . If we assume that the only effect of the drive term is to create a coherent state with  $\langle c(t) \rangle = \mathcal{E} e^{-i\omega_p t}$  in oscillator  $c$ , we can remove  $\omega_c c^\dagger c$  from  $H$  and replace  $c$  by  $\langle c(t) \rangle$  everywhere. Given this time-dependence it then seems more convenient to use a Magnus expansion to analyze the effect of the higher-order effects of the non-resonant terms. For this we move to a rotating frame for the oscillators  $a$  and  $b$  at their own frequency, such that depending on the choice of the drive frequency  $\omega_p$  some terms become time-independent. For example, a term like  $a^\dagger bc$  is time-independent for the choice  $\omega_p = \omega_a - \omega_b$ .

A more thorough quantitative analysis of the error induced by the RWA approximation through perturbative or Magnus expansions would be desirable, as it plays into the accuracy of the CZ gate and is influenced by the number of photons in a GKP mode (as the latter influences the strength of the perturbation).

### C.1. Time-dependent displacement frame

Imagine a Hamiltonian (in a rotating frame of the mode  $a$ ) of the form  $H = H_1(a, a^\dagger) + i\mathcal{E}(t)a^\dagger - i\mathcal{E}^*(t)a$  where  $\mathcal{E}(t)$  is some time-dependent envelope of the drive and  $H_1(a, a^\dagger)$  has some functional form on  $a$  and  $a^\dagger$ . We consider the time-evolution of the vector  $|\tilde{\psi}(t)\rangle = U_0^\dagger(0, t)|\psi(t)\rangle$  with  $U_0(0, t) = \mathcal{T} \exp(-i \int_0^t dt' [i\mathcal{E}(t')a^\dagger - i\mathcal{E}^*(t')a])$  which evolves with the Hamiltonian

$\tilde{H}(t) = U_0^\dagger H U_0 + i \frac{\partial U_0^\dagger}{\partial t} U_0$ . In words, we consider the evolution in the time-dependent displacement frame given by the time-dependent drive. When after some final time  $T$ , the total frame evolution is  $U_0(0, T) = I$ , the time-independent Hamiltonian  $\tilde{H}(t)$  will describe the time-evolution of the actual Schrödinger state  $|\psi(t)\rangle$  over the entire period of time  $T$ . We can write  $U_0(0, t) = D(\beta(t))$  where  $\beta(t) = \int_0^t dt' \mathcal{E}(t')$ , so that we evaluate  $\tilde{H}(t) = H_1(a + \beta, a^\dagger + \beta^*) + i\mathcal{E}(t)\beta^* - i\mathcal{E}^*(t)\beta$  and the last term can be omitted as it give rise to an irrelevant phase.

## ORCID iDs

B M Terhal  <https://orcid.org/0000-0003-0218-6614>

J Conrad  <https://orcid.org/0000-0001-6120-9930>

C Vuillot  <https://orcid.org/0000-0002-3445-0179>

## References

- [1] Blais A, Grimsmo A L, Girvin S M and Wallraff A 2020 Circuit quantum electrodynamics (arXiv:2005.12667)
- [2] Blais A, Girvin S M and Oliver W D 2020 Quantum information processing and quantum optics with circuit quantum electrodynamics *Nat. Phys.* **16** 247–56
- [3] Reagar M *et al* 2016 Quantum memory with millisecond coherence in circuit QED *Phys. Rev. B* **94** 014506

- [4] Reagor M *et al* 2013 Reaching 10 ms single photon lifetimes for superconducting aluminium cavities *Appl. Phys. Lett.* **102** 192604
- [5] Terhal B M 2015 Quantum error correction for quantum memories *Rev. Mod. Phys.* **87** 307
- [6] Koch J, Yu T M, Gambetta J, Houck A A, Schuster D I, Majer J, Blais A, Devoret M H, Girvin S M and Schoelkopf R J 2007 Charge-insensitive qubit design derived from the Cooper pair box *Phys. Rev. A* **76** 042319
- [7] Haroche S and Raimond J M 2006 *Exploring the Quantum: Atoms, Cavities, and Photons* (Oxford: Oxford University Press)
- [8] Gerry C and Knight P 2005 *Introductory Quantum Optics* (Cambridge: Cambridge University Press)
- [9] Carmichael H 1999 *Statistical Methods in Quantum Optics 1. Master Equations and Fokker–Planck Equations* (Heidelberg: Springer)
- [10] Chuang I L, Leung D W and Yamamoto Y 1997 Bosonic quantum codes for amplitude damping *Phys. Rev. A* **56** 1114–25
- [11] Nielsen M and Chuang I 2000 *Quantum Computation and Quantum Information* (Cambridge Series on Information and the Natural Sciences) (Cambridge: Cambridge University Press)
- [12] Michael M H, Silveri M, Brierley R T, Albert V V, Salmilehto J, Jiang L and Girvin S M 2016 New class of quantum error-correcting codes for a bosonic mode *Phys. Rev. X* **6** 031006
- [13] Hu L *et al* 2019 Quantum error correction and universal gate set operation on a binomial bosonic logical qubit *Nat. Phys.* **15** 503–8
- [14] Cochrane P T, Milburn G J and Munro W J 1999 Macroscopically distinct quantum-superposition states as a bosonic code for amplitude damping *Phys. Rev. A* **59** 2631–4
- [15] Mirrahimi M, Leghtas Z, Albert V V, Touzard S, Schoelkopf R J, Jiang L and Devoret M H 2014 Dynamically protected cat-qubits: a new paradigm for universal quantum computation *New J. Phys.* **16** 045014
- [16] Leghtas Z, Kirchmair G, Vlastakis B, Schoelkopf R J, Devoret M H and Mirrahimi M 2013 Hardware-efficient autonomous quantum memory protection *Phys. Rev. Lett.* **111** 120501
- [17] Ofek N *et al* 2016 Extending the lifetime of a quantum bit with error correction in superconducting circuits *Nature* **536** 441–5
- [18] Preskill J 1998 Fault-tolerant quantum computation *Introduction to Quantum Computation* (Singapore: World Scientific) pp 213–69
- [19] Rosenblum S, Reinhold P, Mirrahimi M, Jiang L, Frunzio L and Schoelkopf R J 2018 Fault-tolerant detection of a quantum error *Science* **361** 266–70
- [20] Puri S, Grimm A, Campagne-Ibarcq P, Eickbusch A, Noh K, Roberts G, Jiang L, Mirrahimi M, Devoret M H and Girvin S M 2019 Stabilized cat in a driven nonlinear cavity: a fault-tolerant error syndrome detector *Phys. Rev. X* **9** 041009
- [21] Bergmann M and van Loock P 2016 Quantum error correction against photon loss using multicomponent cat states *Phys. Rev. A* **94** 042332
- [22] Albert V V *et al* 2018 Performance and structure of single-mode bosonic codes *Phys. Rev. A* **97** 032346
- [23] Grimsmo A L, Combes J and Baragiola B Q 2020 Quantum computing with rotation-symmetric bosonic codes *Phys. Rev. X* **10** 011058
- [24] Niu M Y, Chuang I L and Shapiro J H 2018 Hardware-efficient bosonic quantum error-correcting codes based on symmetry operators *Phys. Rev. A* **97** 032323
- [25] Bergmann M and van Loock P 2016 Quantum error correction against photon loss using NOON states *Phys. Rev. A* **94** 012311
- [26] Ouyang Y and Chao R 2019 Permutation-invariant constant-excitation quantum codes for amplitude damping *IEEE Trans. Inf. Theory* **66** 2921–233
- [27] Lau H K and Plenio M B 2016 Universal quantum computing with arbitrary continuous-variable encoding *Phys. Rev. Lett.* **117** 100501
- [28] Gao Y Y, Lester B J, Chou K S, Frunzio L, Devoret M H, Jiang L, Girvin S M and Schoelkopf R J 2019 Entanglement of bosonic modes through an engineered exchange interaction *Nature* **566** 509–12
- [29] Touzard S *et al* 2018 Coherent oscillations inside a quantum manifold stabilized by dissipation *Phys. Rev. X* **8** 021005
- [30] Grimm A, Frattini N E, Puri S, Mundhada S O, Touzard S, Mirrahimi M, Girvin S M, Shankar S and Devoret M H 2019 The Kerr-cat qubit: stabilization, readout, and gates (arXiv:1907.12131)
- [31] Leghtas Z *et al* 2015 Confining the state of light to a quantum manifold by engineered two-photon loss *Science* **347** 853–7
- [32] Lescanne R, Villiers M, Peronnin T, Sarlette A, Delbecq M, Huard B, Kontos T, Mirrahimi M and Leghtas Z 2020 Exponential suppression of bit-flips in a qubit encoded in an oscillator *Nat. Phys.* **16** 509–13
- [33] Puri S, Boutin S and Blais A 2017 Engineering the quantum states of light in a Kerr-nonlinear resonator by two-photon driving *npj Quantum Inf.* **3** 18
- [34] Wielinga B and Milburn G J 1993 Quantum tunneling in a Kerr medium with parametric pumping *Phys. Rev. A* **48** 2494–6
- [35] Meaney C H, Nha H, Duty T and Milburn G J 2014 Quantum and classical nonlinear dynamics in a microwave cavity *EPJ Quantum Technol.* **1** 7
- [36] Wustmann W and Shumeiko V 2013 Parametric resonance in tunable superconducting cavities *Phys. Rev. B* **87** 184501
- [37] Puri S *et al* 2019 Bias-preserving gates with stabilized cat qubits (arXiv:1905.00450)
- [38] Guillaud J and Mirrahimi M 2019 Repetition cat qubits for fault-tolerant quantum computation *Phys. Rev. X* **9** 041053
- [39] Tuckett D K, Darmawan A S, Chubb C T, Bravyi S, Bartlett S D and Flammia S T 2019 Tailoring surface codes for highly biased noise *Phys. Rev. X* **9** 041031
- [40] Gottesman D, Kitaev A and Preskill J 2001 Encoding a qubit in an oscillator *Phys. Rev. A* **64** 012310
- [41] Terhal B M and Weigand D 2016 Encoding a qubit into a cavity mode in circuit QED using phase estimation *Phys. Rev. A* **93** 012315
- [42] Campagne-Ibarcq P *et al* 2019 A stabilized logical quantum bit encoded in grid states of a superconducting cavity (arXiv:1907.12487)
- [43] Hastrup J, Park K, Bohr Brask J, Filip R and Andersen U L 2019 Measurement-free preparation of grid states (arXiv:1912.12645)
- [44] Duivenvoorden K, Terhal B M and Weigand D 2017 Single-mode displacement sensor *Phys. Rev. A* **95** 012305
- [45] Tzitrin I, Bourassa J E, Menicucci N C and Sabapathy K K 2020 Progress towards practical qubit computation using approximate Gottesman-Kitaev-Preskill codes *Phys. Rev. A* **101** 032315
- [46] Matsuura T, Yamasaki H and Koashi M 2019 On the equivalence of approximate Gottesman-Kitaev-Preskill codes (arXiv:1910.08301)
- [47] Weigand D J and Terhal B M 2018 Generating grid states from Schrödinger-cat states without postselection *Phys. Rev. A* **97** 022341

- [48] Vuillot C, Asasi H, Wang Y, Pryadko L P and Terhal B M 2019 Quantum error correction with the toric Gottesman-Kitaev-Preskill code *Phys. Rev. A* **99** 032344
- [49] Weigand D J and Terhal B M 2020 Realizing modular quadrature measurements via a tunable photon-pressure coupling in circuit qed *Phys. Rev. A* **101** 053840
- [50] Flühmann C, Nguyen T L, Marinelli M, Negnevitsky V, Mehta K and Home J P 2019 Encoding a qubit in a trapped-ion mechanical oscillator *Nature* **566** 513–7
- [51] Baragiola B Q, Pantaleoni G, Alexander R N, Karanjai A and Menicucci N C 2019 All-Gaussian universality and fault tolerance with the Gottesman-Kitaev-Preskill code *Phys. Rev. Lett.* **123** 200502
- [52] Campbell E T, Terhal B M and Vuillot C 2017 Roads towards fault-tolerant universal quantum computation *Nature* **549** 172–9
- [53] Braunstein S L 2005 Squeezing as an irreducible resource *Phys. Rev. A* **71** 055801
- [54] Wang Y 2019 Quantum error correction with the GKP code and concatenation with stabilizer codes *MSc Thesis* RWTH Aachen (arXiv:1908.00147)
- [55] Carter S J 1995 Quantum theory of nonlinear fiber optics: phase-space representations *Phys. Rev. A* **51** 3274–301
- [56] Stobińska M, Milburn G J and Wódkiewicz K 2008 Wigner function evolution of quantum states in the presence of self-Kerr interaction *Phys. Rev. A* **78** 013810
- [57] Shi Y, Chamberland C and Cross A 2019 Fault-tolerant preparation of approximate GKP states *New J. Phys.* **21** 093007
- [58] Glancy S and Knill E 2006 Error analysis for encoding a qubit in an oscillator *Phys. Rev. A* **73** 012325
- [59] Wan K H, Neville A and Kolthammer W S 2019 A memory-assisted decoder for approximate Gottesman-Kitaev-Preskill codes (arXiv:1912.00829)
- [60] Brecht T, Pfaff W, Wang C, Chu Y, Frunzio L, Devoret M H and Schoelkopf R J 2016 Multilayer microwave integrated quantum circuits for scalable quantum computing *npj Quantum Inf.* **2** 16002
- [61] Nigg S E, Paik H, Vlastakis B, Kirchmair G, Shankar S, Frunzio L, Devoret M H, Schoelkopf R J and Girvin S M 2012 Black-box superconducting circuit quantization *Phys. Rev. Lett.* **108** 240502
- [62] Bultink C C *et al* 2020 Protecting quantum entanglement from leakage and qubit errors via repetitive parity measurements *Sci. Adv.* **6** 12
- [63] Heonsoo J *et al* 2018 Rapid high-fidelity multiplexed readout of superconducting qubits *Phys. Rev. Appl.* **10** 034040
- [64] Manucharyan V E, Koch J, Glazman L I and Devoret M H 2009 Fluxonium: single Cooper-pair circuit free of charge offsets *Science* **326** 113–6
- [65] Nguyen L B, Lin Y H, Somoroff A, Mencia R, Grabon N and Manucharyan V E 2019 High-coherence fluxonium qubit *Phys. Rev. X* **9** 041041
- [66] Grünhaupt L *et al* 2019 Granular aluminium as a superconducting material for high-impedance quantum circuits *Nat. Mater.* **18** 816–9
- [67] Winkel P, Borisov K, Grünhaupt L, Rieger D, Spiecker M, Valenti F, Ustinov A V, Wernsdorfer W and Pop I M 2019 Implementation of a transmon qubit using superconducting granular aluminium (arXiv:1911.02333)
- [68] Gebauer R *et al* 2019 State preparation of a fluxonium qubit with feedback from a custom FPGA-based platform (arXiv:1912.06814)
- [69] Zhu G, Ferguson D G, Manucharyan V E and Koch J 2013 Circuit QED with fluxonium qubits: theory of the dispersive regime *Phys. Rev. B* **87** 024510
- [70] Pfaff W, Axline C J, Burkhardt L D, Vool U, Reinhold P, Frunzio L, Jiang L, Devoret M H and Schoelkopf R J 2017 Controlled release of multiphoton quantum states from a microwave cavity memory *Nat. Phys.* **13** 882
- [71] Roy A and Devoret M 2018 Quantum-limited parametric amplification with Josephson circuits in the regime of pump depletion *Phys. Rev. B* **98** 045405
- [72] Abdo B, Kamal A and Devoret M 2013 Nondegenerate three-wave mixing with the Josephson ring modulator *Phys. Rev. B* **87** 014508
- [73] Bergeal N, Vijay R, Manucharyan V E, Siddiqi I, Schoelkopf R J, Girvin S M and Devoret M H 2010 Analog information processing at the quantum limit with a Josephson ring modulator *Nat. Phys.* **6** 296–302
- [74] Vrajitoarea A, Huang Z, Groszkowski P, Koch J and Houck A 2019 Quantum control of an oscillator using a stimulated Josephson nonlinearity *Nat. Phys.* **16** 211–7
- [75] Sameti M, Potočnik A, Browne D E, Wallraff A and Hartmann M J 2017 Superconducting quantum simulator for topological order and the toric code *Phys. Rev. A* **95** 042330
- [76] Rymarz M, Bosco S, Ciani A and DiVincenzo D P 2020 Hardware-encoding grid states in a non-reciprocal superconducting circuit (arXiv:2002.07718)
- [77] Frattini N E, Vool U, Shankar S, Narla A, Sliwa K M and Devoret M H 2017 Three-wave mixing Josephson dipole element *Appl. Phys. Lett.* **110** 222603
- [78] Hillmann T, Quijandria F, Johansson G, Ferraro A, Gasparinetti S and Ferrini G 2020 Universal gate set for continuous-variable quantum computation with microwave circuits (arXiv:2002.01402)
- [79] Versluis R, Poletto S, Khammassi N, Tarasinski B, Haider N, Michalak D J, Bruno A, Bertels K and DiCarlo L 2017 Scalable quantum circuit and control for a superconducting surface code *Phys. Rev. Appl.* **8** 034021
- [80] Varbanov B M, Battistel F, Tarasinski B M, Ostroukh V P, O'Brien T E, DiCarlo L and Terhal B M 2020 Leakage detection for a transmon-based surface code (arXiv:2002.07119)
- [81] Noh K and Chamberland C 2020 Fault-tolerant bosonic quantum error correction with the surface-Gottesman-Kitaev-Preskill code *Phys. Rev. A* **101** 012316
- [82] Andrist R S, Katzgraber H G, Bombin H and Martin-Delgado M A 2016 Error tolerance of topological codes with independent bit-flip and measurement errors *Phys. Rev. A* **94** 012318
- [83] Ohno T, Arakawa G, Ichinose I and Matsui T 2004 Phase structure of the random-plaquette  $z_2$  gauge model: accuracy threshold for a toric quantum memory *Nucl. Phys. B* **697** 462–80
- [84] Takeda K, Sasamoto T and Nishimori H 2005 Exact location of the multicritical point for finite-dimensional spin glasses: a conjecture *J. Phys. A: Math. Gen.* **38** 3751–74
- [85] Brink M, Chow J M, Hertzberg J, Magesan E and Rosenblatt S 2018 Device challenges for near term superconducting quantum processors: frequency collisions *2018 IEEE Int. Electron Devices Meeting (IEDM)* pp 6.1.1–3
- [86] Blumoff J Z *et al* 2016 Implementing and characterizing precise multiqubit measurements *Phys. Rev. X* **6** 031041
- [87] Heeres R W, Vlastakis B, Holland E, Krastanov S, Albert V V, Frunzio L, Jiang L and Schoelkopf R J 2015 Cavity state manipulation using photon-number selective phase gates *Phys. Rev. Lett.* **115** 137002

- [88] O'Brien T E, Tarasinski B and DiCarlo L 2017 Density-matrix simulation of small surface codes under current and projected experimental noise *npj Quantum Inf.* **3** 39
- [89] Rol M A *et al* 2019 Fast, high-fidelity conditional-phase gate exploiting leakage interference in weakly anharmonic superconducting qubits *Phys. Rev. Lett.* **123** 120502
- [90] Arute F *et al* 2019 Quantum supremacy using a programmable superconducting processor *Nature* **574** 505–10
- [91] Place A P M *et al* 2020 New material platform for superconducting transmon qubits with coherence times exceeding 0.3 milliseconds (arXiv:2003.00024)
- [92] Mumford D 2007 *Tata Lectures on Theta I Modern Birkhäuser Classics* 2nd edn (Basel: Birkhäuser) <https://springer.com/gp/book/9780817645724>
- [93] Romik D 2018 The Taylor coefficients of the Jacobi theta constant  $\theta_3$  (arXiv: 1807.06130)
- [94] Flajolet P and Sedgewick R 2009 *Analytic Combinatorics* (Cambridge: Cambridge University Press)
- [95] Blais A, Huang R S, Wallraff A, Girvin S M and Schoelkopf R J 2004 Cavity quantum electrodynamics for superconducting electrical circuits: an architecture for quantum computation *Phys. Rev. A* **69** 062320
- [96] Bravyi S, DiVincenzo D P, Loss D and Terhal B M 2008 Quantum simulation of many-body Hamiltonians using perturbation theory with bounded-strength interactions *Phys. Rev. Lett.* **101** 070503
- [97] Bravyi S, DiVincenzo D P and Loss D 2011 Schrieffer–Wolff transformation for quantum many-body systems *Ann. Phys., NY* **326** 2793–826
- [98] Consani G and Warburton P A 2020 Effective Hamiltonians for interacting superconducting qubits: local basis reduction and the Schrieffer–Wolff transformation *New J. Phys.* **22** 053040
- [99] Blanes S, Casas F, Oteo J and Ros J 2009 The Magnus expansion and some of its applications *Phys. Rep.* **470** 151–238
- [100] Zeuch D, Hassler F, Slim J and DiVincenzo D P 2018 Exact rotating wave approximation (arXiv:1807.02858)