

PAPER • OPEN ACCESS

## Fault-tolerant preparation of approximate GKP states

To cite this article: Yunong Shi *et al* 2019 *New J. Phys.* **21** 093007

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

# New Journal of Physics

The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft 

IOP Institute of Physics

Published in partnership  
with: Deutsche Physikalische  
Gesellschaft and the Institute  
of Physics



## OPEN ACCESS

RECEIVED  
11 May 2019

REVISED  
31 July 2019

ACCEPTED FOR PUBLICATION  
12 August 2019

PUBLISHED  
4 September 2019

## PAPER

# Fault-tolerant preparation of approximate GKP states

Yunong Shi<sup>1,3</sup>, Christopher Chamberland<sup>2,3,4</sup>  and Andrew Cross<sup>2</sup>

<sup>1</sup> Physics department, University of Chicago, Chicago, IL 60637, United States of America

<sup>2</sup> IBM T. J. Watson Research Center, Yorktown Heights, NY 10598, United States of America

<sup>3</sup> YS and CC contributed equally to the work.

<sup>4</sup> Author to whom any correspondence should be addressed.

E-mail: [yunong@uchicago.edu](mailto:yunong@uchicago.edu), [christopher.chamberland@ibm.com](mailto:christopher.chamberland@ibm.com) and [awcross@us.ibm.com](mailto:awcross@us.ibm.com)

**Keywords:** error correction, bosonic codes, GKP states, fault-tolerant state preparation, circuit QED

Original content from this  
work may be used under  
the terms of the [Creative  
Commons Attribution 3.0  
licence](#).

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



## Abstract

Gottesman–Kitaev–Preskill (GKP) states appear to be amongst the leading candidates for correcting errors when encoding qubits into oscillators. However the preparation of GKP states remains a significant theoretical and experimental challenge. Until now, no clear definitions for fault-tolerantly preparing GKP states have been provided. Without careful consideration, a small number of faults can lead to large uncorrectable shift errors. After proposing a metric to compare approximate GKP states, we provide rigorous definitions of fault-tolerance and introduce a fault-tolerant phase estimation protocol for preparing such states. The fault-tolerant protocol uses one flag qubit and accepts only a subset of states in order to prevent measurement readout errors from causing large shift errors. We then show how the protocol can be implemented using circuit QED. In doing so, we derive analytic expressions which describe the leading order effects of the nonlinear dispersive shift and Kerr nonlinearity. Using these expressions, we show that to mitigate the nonlinear dispersive shift and Kerr terms would require the protocol to be implemented on time scales four orders of magnitude longer than the time scales relevant to the protocol for physically motivated parameters. Despite these restrictions, we numerically show that a subset of the accepted states of the fault-tolerant phase estimation protocol maintain good error correcting capabilities even in the presence of noise.

## 1. Introduction

Fault-tolerant quantum computing will be essential for implementing large scale quantum algorithms that offer provable speed-ups over the best known classical algorithms. Currently there are many proposals for encoding qubits into error correcting codes in order to perform universal fault-tolerant quantum computation.

Depending on the underlying physical architecture, some encoding schemes are more suitable than others.

One method proposed by Gottesman, Kitaev and Preskill is to encode a qubit into an oscillator such that small shift errors in both position and momentum can be corrected. Although some bosonic codes have been designed to correct realistic errors arising from noise models encountered in the experiment (e.g. photon loss), recently it has been shown that GKP codes have better error correction capabilities than such codes under the assumption of perfect encoding and decoding [1–3]. In addition, it has been shown how GKP codes can be concatenated with the toric code in order to achieve larger threshold values compared to toric codes with bare physical qubits [4–6]. Lastly, given a supply of GKP-encoded Pauli eigenstates, universal fault-tolerant quantum computation can be achieved using only Gaussian operations [7].

Given the above, it is clear that the fault-tolerant preparation of encoded GKP states is an important problem that needs to be addressed. Various proposals for preparing GKP states have been outlined [3, 8–15]. However to our knowledge, no clear definitions for fault-tolerantly preparing GKP states using qubit-cavity couplings have been proposed. As such, without careful consideration, it is possible that a small number of faults lead to large uncorrectable shift errors. Inspired by [16], in this work we propose new fault-tolerant definitions for preparing GKP states which tolerate small shift errors and a small number of faults occurring on ancilla qubits during the

protocol. We then show how the phase estimation protocol proposed in this work satisfies our fault-tolerance criteria. In particular, the protocol is robust to a single fault occurring on the ancilla qubits in addition to shift errors of magnitude at most 0.06. In order to be fault-tolerant, the protocol uses one flag qubit (and thus requires a total of two ancilla qubits) to prevent damping errors and imperfect implementations of the required gates from causing large shift errors. In addition, we provide an algorithm which only accepts a subset of all output states of phase estimation in order to prevent a single measurement readout error from causing large uncorrectable shift errors. We then proceed to show how our protocol can be implemented using circuit QED. We first analytically derive expressions describing the effects of the nonlinear dispersive shift and Kerr nonlinearity on the evolution of the cavity. We then numerically show that certain states output from the phase estimation protocol are robust to noise processes found in current 2D and 3D cavities since these can still correct small shift errors with high probability.

Our paper is structured as follows. In section 2, we provide new metrics which we use throughout the remainder of the paper to characterize the error correction capabilities of approximate GKP states. The fault-tolerance definitions used throughout this paper are given in section 3. In section 4.1, we briefly review the phase estimation protocol described in [3]. In section 4.2, we obtain a new phase estimation protocol and prove that it is fault-tolerant under our definitions. In section 4.3, we compare the error correction capabilities of various states obtained from the phase estimation protocol in the noise free case. Section 5 is devoted to the implementation and error analysis of our protocol in circuit QED. In section 5.1, we provide analytic expressions for the time evolution of the qubit-cavity coupling when implementing a controlled-displacement gate. The expressions are derived in appendix B. In section 5.2, we numerically solve a master equation which includes all considered noise processes, such qubit damping and dephasing, photon loss in addition to measurement, ancilla state-preparation and gate errors which arise from a depolarizing noise channel. In section 6 we summarize our results and discuss possible future directions.

## 2. Goodness of approximate GKP states

As was explained in [3, 17], preparing perfect GKP states would require an infinite amount of squeezing. In a realistic setting, one can only prepare approximate GKP states with finite squeezing. Perfect GKP states, which are  $+1$  eigenstates of the mutually commuting operators  $S_p = e^{-2i\sqrt{\pi}p}$  and  $S_q = e^{2i\sqrt{\pi}q}$ , can correct shift errors of size at most  $\frac{\sqrt{\pi}}{2}$ . Note that for the displacement operator  $D(\alpha) = e^{\alpha a^\dagger - \alpha^* a}$ , we can write  $S_p = D(\sqrt{2\pi})$  and  $S_q = D(i\sqrt{2\pi})$ . The goal of this paper will be to fault-tolerantly prepare approximate GKP states which can correct small shift errors correctable by perfect GKP states with high probability (in section 3 we will specify what we mean by correctable shift errors). Therefore, it is important to have a metric which allows us to compute the ‘goodness’ of an approximate GKP state.

Recall that for perfect GKP states, the logical  $|\bar{0}\rangle$  and  $|\bar{1}\rangle$  states are given by

$$|\bar{0}\rangle = \sum_{k=-\infty}^{\infty} S_p^k |q=0\rangle$$

$$|\bar{1}\rangle = \sum_{k=-\infty}^{\infty} S_p^k |q=\sqrt{\pi}\rangle, \quad (1)$$

up to normalization. In practice, approximate GKP states analogous to those in equation (1) can be prepared by first preparing finitely squeezed states in  $q$  space and approximately projecting these states onto the  $S_p = 1$  eigenspace. For instance, the  $|q=0\rangle$  state can be written as a squeezed vacuum state  $|sq\rangle$  with squeezing parameter  $\Delta$  which is given by

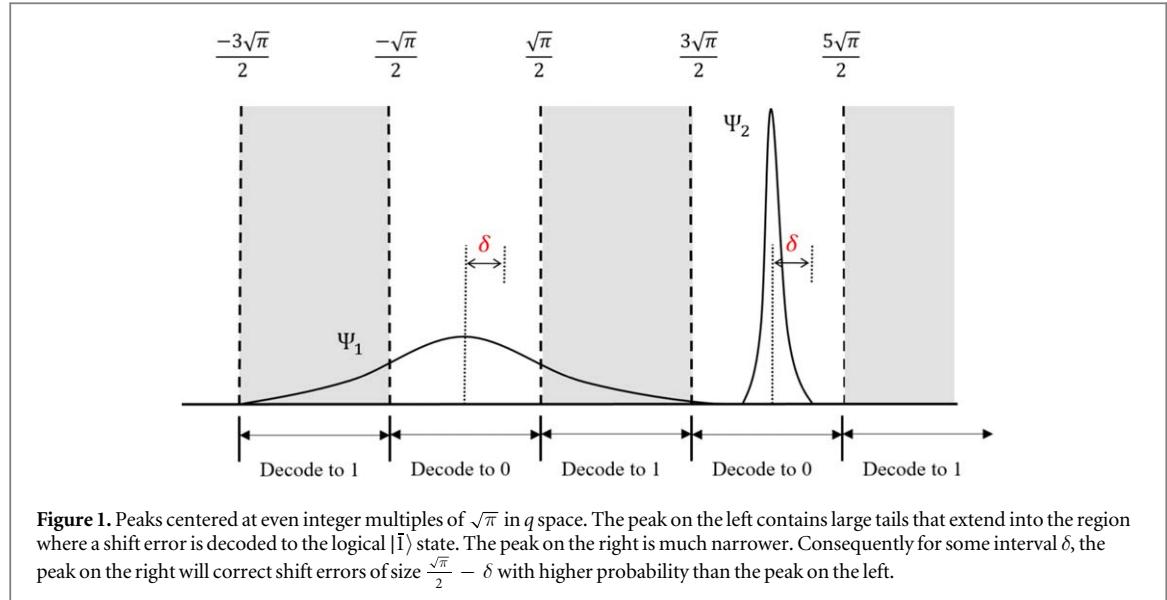
$$|sq\rangle = \int \frac{dq}{(\pi\Delta^2)^{1/4}} e^{-q^2/(2\Delta^2)} |q\rangle. \quad (2)$$

One can then apply a sum of displacements with a Gaussian filter to obtain

$$|\bar{0}\rangle_{\text{approx}} = C \sum_{k=-\infty}^{\infty} e^{-2\pi\tilde{\Delta}^2 k^2} D(k\sqrt{2\pi}) |sq\rangle, \quad (3)$$

where  $C$  is a normalization coefficient. As was shown in [3], it is natural to have  $\tilde{\Delta} = \Delta$ . In section 4, we will present a fault-tolerant version of phase estimation for approximately projecting the state in equation (2) onto the  $+1$  eigenspace of  $S_p$  (see [3] which provides the first description for using phase estimation to prepare approximate GKP states).

Naturally because of the finite width of the peaks of approximate GKP states, it will not be possible to correct a shift error in  $p$  or  $q$  of magnitude at most  $\frac{\sqrt{\pi}}{2}$  with certainty. For example, suppose we have an approximate  $|\bar{0}\rangle$  GKP state with a peak at  $q = 0$  subject to a shift error  $e^{-ivp}$  with  $|v| \leq \frac{\sqrt{\pi}}{2}$ . The finite width of the Gaussian peaks



**Figure 1.** Peaks centered at even integer multiples of  $\sqrt{\pi}$  in  $q$  space. The peak on the left contains large tails that extend into the region where a shift error is decoded to the logical  $|\bar{1}\rangle$  state. The peak on the right is much narrower. Consequently for some interval  $\delta$ , the peak on the right will correct shift errors of size  $\frac{\sqrt{\pi}}{2} - \delta$  with higher probability than the peak on the left.

will have a non-zero overlap in the region  $\frac{\sqrt{\pi}}{2} < q < \frac{3\sqrt{\pi}}{2}$  and  $\frac{-3\sqrt{\pi}}{2} < q < \frac{-\sqrt{\pi}}{2}$ . Thus with non-zero probability the state can be decoded to  $|\bar{1}\rangle$  instead of  $|\bar{0}\rangle$  (see figure 1 for an illustration).

In general, if an approximate GKP state is afflicted by a correctable shift error, we would like the probability of decoding to the incorrect logical state to be as small as possible. A smaller overlap of the approximate GKP state in regions in  $q$  and  $p$  space that lead to decoding the state to the wrong logical state will lead to a higher probability of correcting a correctable shift error by a perfect GKP state. These remarks motivate the following definition

**Definition 1.** Let  $|\tilde{0}\rangle$  be an approximate logical  $|\bar{0}\rangle$  GKP state. We say that  $|\tilde{0}\rangle$  is  $(\frac{\sqrt{\pi}}{2} - \delta, \epsilon)_q$ -GKP correctable if and only if for a given tuple  $(\delta, \epsilon)$  with  $\delta \leq \frac{\sqrt{\pi}}{2}$ , and  $0 \leq \epsilon \leq 1$ , we have

$$\sum_{k=-\infty}^{\infty} \int_{2\sqrt{\pi}k-\delta}^{2\sqrt{\pi}k+\delta} |\langle q|\tilde{0}\rangle|^2 dq > 1 - \epsilon. \quad (4)$$

We say that  $|\tilde{0}\rangle$  is  $(\frac{\sqrt{\pi}}{2} - \delta, \epsilon)_p$ -GKP correctable if and only if for a given tuple  $(\delta, \epsilon)$  with  $\delta \leq \frac{\sqrt{\pi}}{2}$ , and  $0 \leq \epsilon \leq 1$ , we have

$$\sum_{k=-\infty}^{\infty} \int_{\sqrt{\pi}k-\delta}^{\sqrt{\pi}k+\delta} |\langle p|\tilde{0}\rangle|^2 dp > 1 - \epsilon. \quad (5)$$

Note that the bounds in equations (4) and (5) are different since GKP states have peaks defined on a rectangular lattice. Similarly, for an approximate logical  $|\tilde{+}\rangle$  state, we have the following definition

**Definition 2.** Let  $|\tilde{+}\rangle$  be an approximate logical  $|\bar{+}\rangle$  GKP state. We say that  $|\tilde{+}\rangle$  is  $(\frac{\sqrt{\pi}}{2} - \delta, \epsilon)_p$ -GKP correctable if and only if for a given tuple  $(\delta, \epsilon)$  with  $\delta \leq \frac{\sqrt{\pi}}{2}$ , and  $0 \leq \epsilon \leq 1$ , we have

$$\sum_{k=-\infty}^{\infty} \int_{2\sqrt{\pi}k-\delta}^{2\sqrt{\pi}k+\delta} |\langle p|\tilde{+}\rangle|^2 dp > 1 - \epsilon. \quad (6)$$

We say that  $|\tilde{+}\rangle$  is  $(\frac{\sqrt{\pi}}{2} - \delta, \epsilon)_q$ -GKP correctable if and only if for a given tuple  $(\delta, \epsilon)$  with  $\delta \leq \frac{\sqrt{\pi}}{2}$ , and  $0 \leq \epsilon \leq 1$ , we have

$$\sum_{k=-\infty}^{\infty} \int_{\sqrt{\pi}k-\delta}^{\sqrt{\pi}k+\delta} |\langle q|\tilde{+}\rangle|^2 dq > 1 - \epsilon. \quad (7)$$

As an example, suppose we have two approximate GKP states  $|\tilde{0}\rangle_1$  and  $|\tilde{0}\rangle_2$  which are  $(\frac{\sqrt{\pi}}{2} - \delta, \epsilon_1)_q$  and  $(\frac{\sqrt{\pi}}{2} - \delta, \epsilon_2)_q$  correctable for a shift of size  $e^{i(\frac{\sqrt{\pi}}{2} - \delta)p}$  (which is correctable by a perfect  $|\bar{0}\rangle$  GKP state). If  $\epsilon_1 < \epsilon_2$ , then  $|\tilde{0}\rangle_1$  will correct a shift of size  $\frac{\sqrt{\pi}}{2} - \delta$  with greater probability than  $|\tilde{0}\rangle_2$ . This is due to the fact that  $|\tilde{0}\rangle_1$  has a smaller overlap in regions which result in decoding the logical  $|\bar{0}\rangle$  state to the logical  $|\bar{1}\rangle$ . In this sense we say that  $|\tilde{0}\rangle_1$  is better than  $|\tilde{0}\rangle_2$  at correcting shift errors in  $q$  space.

### 3. Fault-tolerant definitions

Given a protocol for preparing an approximate GKP state, errors that occur during the protocol can potentially accumulate resulting in a large shift error in addition to deforming the output state. This can then result in an output state which significantly differs from a good approximate GKP state. By good, we mean a state that for a desired value of  $\delta$ , the state is  $\left(\frac{\sqrt{\pi}}{2} - \delta, \epsilon_1\right)_q$  and  $\left(\frac{\sqrt{\pi}}{2} - \delta, \epsilon_2\right)_q$  correctable with  $\epsilon_1, \epsilon_2 \ll 1$ .

The desired values for  $\delta$  depend on the particular fault-tolerant error correction protocol used to correct shift errors on encoded data qubits (see for instance [18–22]). For example, one can use a version of the error correction scheme (which reduces to Steane error correction for qubit CSS stabilizer codes) as proposed by Glancy and Knill in [9]. In this scheme, logical  $|\tilde{0}\rangle$  and  $|\tilde{+}\rangle$  ancillas are prepared and interact with the encoded data qubit via a CNOT gate to correct the shift errors afflicting the data qubit. It is shown that the largest shift error that can be corrected is  $\frac{\sqrt{\pi}}{6}$ . The threshold of  $\frac{\sqrt{\pi}}{6}$  arises from how shift errors that are initially afflicting the encoded data qubit and ancilla states propagate through the CNOT gates and combine prior to the measurement. In practice, if additional shift errors occur during the the error correction scheme (say after applying the CNOT gates), the largest correctable shift error could potentially be smaller than  $\frac{\sqrt{\pi}}{6}$ <sup>5</sup>. In what follows, we will assume that after preparing the desired ancilla states using some state preparation protocol, the ancillas will be used in the error correction scheme of [9] (assumed to be fault-free) to correct shift errors on encoded data qubits. We also point out that due to the propagation of shift errors in the error correction scheme of [9], it is important to prepare approximate  $|\tilde{0}\rangle$  and  $|\tilde{+}\rangle$  states that have small shift errors in *both*  $p$  and  $q$ .

If a small number of errors that occur during the preparation of an approximate GKP state result in large shift errors (or linear combinations of large shift errors) on the output state, then clearly the protocol used would not be practical. Thus it is important that a given state preparation protocol be fault-tolerant. In what follows we will define what we mean by fault-tolerant. We start with the following two definitions:

**Definition 3.** A shift error is said to be correctable if the magnitude of the shift is less than or equal to  $\frac{\sqrt{\pi}}{6}$ . Otherwise, we will say that the shift error is uncorrectable.

**Definition 4.** Suppose we have a protocol for preparing an approximate GKP state. We will say that the output state is an ideal approximate GKP state if no faults occur during the protocol.

Thus by definition 4, any approximate GKP state obtained from a state preparation protocol will be called an ideal approximate GKP state if the protocol is implemented fault free, even if the output state is  $\left(\frac{\sqrt{\pi}}{2} - \delta, \epsilon\right)_{p,q}$  correctable for some desired  $\delta$  with large  $\epsilon$  (so that the probability of correcting a shift  $\frac{\sqrt{\pi}}{2} - \delta$  is small).

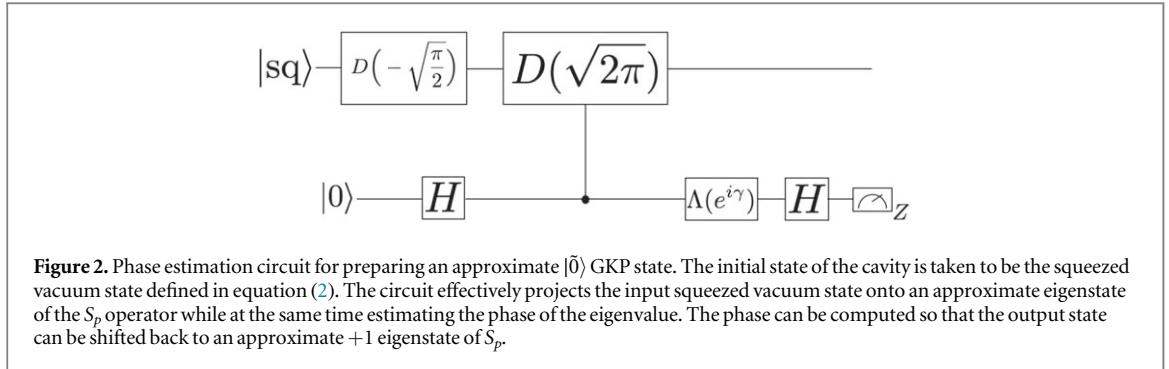
Note that the notion of correctable in definition 3 assumes that only the data and ancilla qubits used in the error correction scheme of [9] have shift errors. If other operations such as the CNOT gates introduce additional errors, the correctable threshold would be smaller than  $\sqrt{\pi}/6$ .

With the above definitions we are now ready to define what it means for a state preparation protocol of an approximate GKP state to be fault-tolerant. We point out that in section 4 we consider a fault-tolerant state preparation protocol based on phase estimation. Hence the definitions given below are specific to the case where an approximate GKP state is obtained by coupling a qubit to an oscillator.

**Definition 5.**  $(m, \tilde{\delta})$ -fault-tolerant state preparation of an approximate GKP state: Suppose we have a protocol for preparing an approximate GKP state which is obtained by coupling qubits to a harmonic oscillator. Suppose also that at most  $m$  faults occur during the protocol on the qubit Hilbert space and in addition, a correctable shift error in either  $p$  or  $q$  of size at most  $\tilde{\delta}$  occurs on the oscillator Hilbert space. We will say that the protocol is an  $(m, \tilde{\delta})$ -fault-tolerant state preparation of an approximate GKP state protocol if the output state differs from an ideal approximate GKP state by a correctable shift error.

A few clarifications are necessary. Firstly, a fault on the qubit Hilbert space corresponds to a location where an error can occur (see for instance [16]). By location we are referring to a particular time step where either a gate is implemented, a qubit is prepared, a qubit is measured or a qubit is idling. On the oscillator Hilbert space, if an error occurs, we can always expand that error into shift errors (see for instance equation (7.12) in [1] and also [3]). By performing the error correction scheme of [9], measuring the  $q$  and  $p$  quadratures to perform error correction will always project the state onto a state with a single shift error in  $q$  and  $p$ . Lastly, we point out that  $\tilde{\delta}$  in definition 5 relates to the largest allowed size of the shift error which occurs during the protocol. This should

<sup>5</sup> Using the optimizations considered in [18], using Knill error correction could potentially increase the threshold of  $\frac{\sqrt{\pi}}{6}$  to a larger value.



not be confused with  $\delta$  in definitions 1 and 2 which relates to the size of a shift error that can be corrected by an approximate GKP state<sup>6</sup>.

Intuitively, a fault-tolerant protocol should have the property that if both the number of qubit errors and the size of shift errors are small, the resulting shift error on the output state should be correctable. Depending on the desired value for  $\tilde{\delta}$ , a particular fault-tolerant error correction protocol might be less tolerant to the size of the shift errors that occurred during the preparation of the approximate GKP state (in practice a different error correction scheme than Steane error correction could be used).

Suppose a state preparation protocol satisfies definition 5. If during the preparation of the state,  $m$  faults occur on the qubit Hilbert space in addition to a shift error of size at most  $\tilde{\delta}$  on the oscillator Hilbert space, the remaining shift error on the output state will be corrected in a perfect version of the error correction protocol in [9]. As we will see in section 4, due to measurement errors, the phase estimation protocol presented in [3] needs to be modified in order to be a  $(1, \tilde{\delta})$  fault-tolerant protocol. There are additional fault locations (apart from measurement errors) which can result in large shift errors that need to be treated with care.

Lastly we describe how we will evaluate the error correction properties of an approximate GKP state obtained from a noisy state preparation protocol. It is important to compare the goodness of a GKP state prepared from a noisy circuit to that of an ideal circuit. If the output state of a noisy state preparation protocol is correctable, the shift error will be removed when performing error correction. Therefore we will compare the  $\left(\frac{\sqrt{\pi}}{2} - \delta, \epsilon\right)_{p,q}$  correctable properties of output states after performing an optimal shift back correction (as long as the shift error is correctable) to the output state. If the shift error is not correctable, the protocol will be deemed too noisy.

The optimal shift back correction is found as follows. In  $q$  space, the optimal shift back  $c_{\max}^q$  is computed as

$$c_{\max}^q = \arg \max_c \sum_{k=-\infty}^{\infty} \int_{2\sqrt{\pi}k+c-\delta}^{2\sqrt{\pi}k+c+\delta} |\langle q|\tilde{0}\rangle|^2 dq. \quad (8)$$

Similarly, in  $p$  space we have

$$c_{\max}^p = \arg \max_c \sum_{k=-\infty}^{\infty} \int_{\sqrt{\pi}k+c-\delta}^{\sqrt{\pi}k+c+\delta} |\langle p|\tilde{0}\rangle|^2 dp. \quad (9)$$

For protocols preparing  $|\tilde{\gamma}\rangle$ , the metric can be defined analogously, but with the integral bounds for  $p$  and  $q$  switched.

## 4. Fault-tolerant phase estimation protocol

In section 4.1 we will give a brief review of the phase estimation protocol presented in [3]. In section 4.2, we will show how the protocol can be modified so that it becomes a  $(1, \tilde{\delta})$ -fault-tolerant state preparation of an approximate GKP state and we will provide the value of  $\tilde{\delta}$ .

### 4.1. Brief review of the phase estimation protocol for preparing approximate GKP states

The phase estimation circuit for preparing an approximate  $|\tilde{0}\rangle$  GKP state is given in figure 2. The  $H$  gate is the Hadamard gate and  $\Lambda(e^{i\gamma}) = \text{diag}(1, e^{i\gamma})$ . After applying several rounds of the circuit in figure 2, the input squeezed vacuum state (given in equation (2)) is projected onto an approximate eigenstate of  $S_p$  with some

<sup>6</sup>The probability  $1 - \epsilon$  of correcting a shift of size  $\frac{\sqrt{\pi}}{2} - \delta$  depends on the location and width of the peaks.

random eigenvalue<sup>7</sup>  $e^{i\theta}$ . Additionally, an estimated value for the phase  $\theta$  is obtained. After computing the phase, the state can be shifted back to an approximate +1 eigenstate of  $S_p$ .

There are a variety of ways to choose the phases  $\gamma$  at each round in the gate  $\Lambda(e^{i\gamma})$ . In a non-adaptive phase estimation protocol, half of the values for  $\gamma$  can be chosen to be 0 and the other half can be chosen to be  $\pi/2$ . In an adaptive phase estimation protocol, if the phase to be estimated is  $\theta$  and assuming no prior knowledge of  $\theta$ , during the first round the phase can be chosen as  $\gamma_1 = 0$ . For later rounds, it was shown in [3] that the next phases can be chosen as

$$\gamma_m = \arg \max_{\gamma} \sum_{x_m=0,1} \left| \int d\theta e^{i\theta} P_{\gamma}(x[m]|\theta) \right|. \quad (10)$$

In equation (10), the probability  $P_{\gamma}(x[m]|\theta)$  is the probability of obtaining measurement outcomes  $x_1, \dots, x_m$  when the produced output eigenstate  $|\psi_{\theta}\rangle$  has eigenvalue  $e^{i\theta}$ . Since the measurement results for each round are independent, the estimated probability is given by

$$P_{\gamma=\gamma_m}(x[m]|\theta) = \prod_{i=1}^m \cos^2 \left( \frac{\theta + \gamma_i}{2} + x_i \frac{\pi}{2} \right). \quad (11)$$

By analytically computing the evolution of the state, the exact probability of obtaining the outcomes  $x[M]$  is derived in appendix D.

Given the final measurement record  $x[M]$ , the estimated phase  $\tilde{\theta}$  is chosen as

$$\tilde{\theta} = \arg \int_{-\pi}^{\pi} d\theta e^{i\theta} P(x[M]|\theta). \quad (12)$$

The shift back correction based on the estimated phase  $\tilde{\theta}$  is given by  $e^{i\frac{\tilde{\theta}}{2\sqrt{\pi}}q}$ . Note that for either the adaptive or non-adaptive protocol, the estimated phase and probabilities are computed using equations (11) and (12).

#### 4.2. (1, $\tilde{\delta}$ ) fault-tolerant state preparation using phase estimation

From the circuit of figure 2 and from equation (12), a measurement readout error can result in the wrong estimated phase which in turn results in an incorrect shift-back correction (application of  $e^{ivq}$  where  $v$  is computed using equations (10) and (11)). Now suppose that the output state is afflicted by a shift error of the form  $e^{-iwq}$ . In this case, the final output state will have a total shift error of the form  $e^{-i(v-w)q}$ . If  $|v - w| \bmod \sqrt{\pi} > \sqrt{\pi}/6$ , then the shift error will be uncorrectable and thus the phase estimation protocol will not be fault-tolerant<sup>8</sup>. In what follows we will identify an output state of the phase estimation protocol as  $x[m]$  if it arises from the measurement outcomes  $x_1, \dots, x_m$  in the fault-free case.

Consider the case where the phase estimation protocol is implemented in  $m$  rounds and let us assume that a single measurement readout error occurs and that all other operations are fault free. In this case the output state  $x_{\text{out}}[m]$  will have one bit which differs from the bit string  $x_{\text{corr}}[m]$  that would have been obtained in the fault free case (so that the Hamming distance  $d_H(x_{\text{out}}[m], x_{\text{corr}}[m]) = 1$ ). The shift-back operation of equation (10) applied to the output state will be the shift correction of an output state that is one Hamming distance away from the actual state that was produced. Thus to ensure that the protocol is (1,  $\tilde{\delta}$ ) fault-tolerant for some  $\tilde{\delta}$ , we should only accept output states  $x_{\text{out}}[m]$  which have the property that applying the shift back correction corresponding to *any* other state  $x'[m]$  with  $d_H(x'[m], x[m]) = 1$  results in a correctable left-over shift error. These remarks motivate the following protocol to prepare an approximate  $|\tilde{0}\rangle$  state which is fault-tolerant to a single measurement readout error.

**Calculation of acceptance set  $A_m$  and  $\tilde{\delta}$ .** Consider all output states obtained from an  $m$  round *fault-free* phase estimation protocol of section 4.1. Let  $A_m = \emptyset$  (which we call the acceptance set) and  $\Gamma_m = \emptyset$ . For each output state  $x[m]$ , do the following

1. Compute the shift correction  $v_m = \frac{1}{2\sqrt{\pi}} \arg \int_{-\pi}^{\pi} d\theta e^{i\theta} P(x[m]|\theta)$ .
2. Let  $j[m]$  be a bit string of size  $m$ . For all  $j \in \{1, \dots, m\}$  such that  $d_H(x[m], j[m]) = 1$ , compute the shift correction  $v_m^{(j)} = \frac{1}{2\sqrt{\pi}} \arg \int_{-\pi}^{\pi} d\theta e^{i\theta} P(j[m]|\theta)$ .
3. If  $\#j$  such that  $|v_m - v_m^{(j)}| \bmod \sqrt{\pi} > \frac{\sqrt{\pi}}{6}$ , append  $x[m]$  to  $A_m$  and  $\forall j \in \{1, \dots, m\}$ , append  $|v_m - v_m^{(j)}|$  to  $\Gamma_m$ .  
If  $A_m \neq \emptyset$ ,  $\tilde{\delta} = \max_m \Gamma_m$ .

<sup>7</sup> The phase  $\theta$  that is obtained depends on the measurement outcome of the ancilla qubit in each round.

<sup>8</sup> Note that  $|v - w|$  should be taken modulo  $\sqrt{\pi}$  since a perfect  $|\tilde{0}\rangle$  GKP state in  $p$  space has peaks at any integer multiple of  $\sqrt{\pi}$ .

**Table 1.** The first row displays the accepted measurement strings (set  $A_4$ ) for the four round non-adaptive phase estimation protocol. If the protocol does not output an element of  $A_4$ , the output is rejected and the protocol begins anew. The second row displays the largest shift difference that can occur when applying the phase estimated shift in the presence of a single measurement readout error. As can be seen, these are all less than  $\sqrt{\pi}/6 \approx 0.295$ . The third row gives the probabilities of obtaining each state at the output of the phase estimation protocol.

Four round phase estimation protocol acceptance set $A_4$	1111	1010	0101	1000	0010
Largest shift difference	0.235	0.225	0.225	0.225	0.225
Probability of obtaining output state	0.1258	0.1287	0.1279	0.0505	0.0499

**(1,  $\delta$ ) fault-tolerant  $m$  round phase estimation protocol for measurement readout errors.** Consider the acceptance set  $A_m \neq \emptyset$  and  $\tilde{\delta}$  computed from the procedure described above. During an implementation of the phase estimation protocol subject to measurement readout errors, if the obtained output state  $x[m] \notin A_m$ , abort the protocol and start anew.

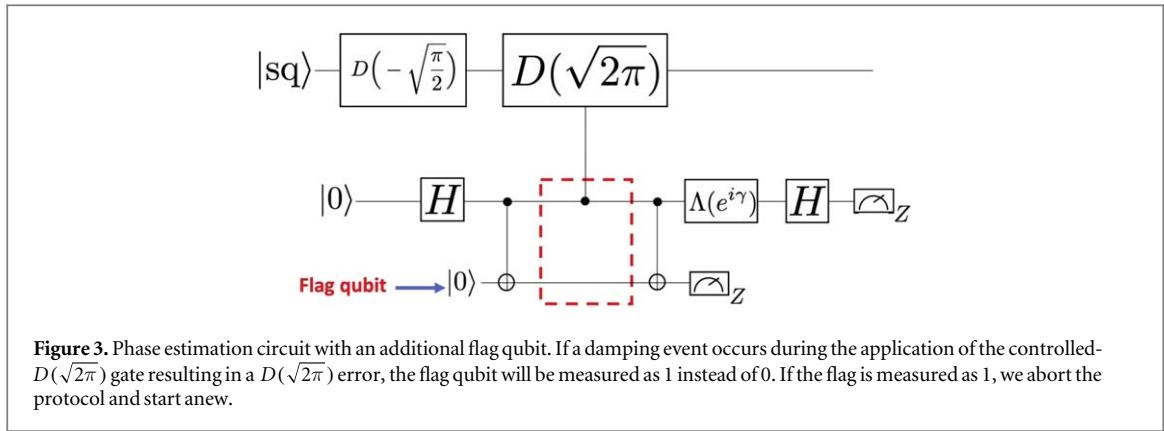
Note that during the protocol, there can be more than one measurement readout error. However if more than one readout error occurs, it is possible that the output state is afflicted by a shift error of magnitude greater than  $\sqrt{\pi}/6$ . Our protocol only guarantees protection against a single readout error.

As an example, consider the four round phase estimation protocol. Applying the procedure described above for the adaptive phase estimation protocol, we find that  $A_4$  is empty when choosing the initial phase to be  $\gamma_0 = 0$  or  $\gamma_0 = \frac{\pi}{2}$ . This indicates that the adaptive phase estimation protocol of four rounds described in section 4.1 is *not* fault-tolerant to single measurement readout errors<sup>9</sup>. If the phases are computed using the non-adaptive phase estimation protocol and applying the protocol described above, we find that  $\tilde{\delta} = \frac{\sqrt{\pi}}{6} - 0.235 \approx 0.0604$ . The set  $A_4$  of accepted measurement strings is given in table 1 and has  $|A_4| = 5$ . The total probability of obtaining a state in  $A_4$  is roughly 48.3%. For the phase estimation protocol with more than four rounds, the acceptance set and the acceptance probability is very sensitive to the initial phase  $\gamma_0$  and the domain of the optimized  $\gamma$ . For example, if we choose the initial phase  $\gamma_0$  to be  $\frac{\pi}{2}$  and the range of  $\gamma$  to be  $[0, 2\pi]$ , we get 18 accepted states (for the adaptive protocol). In this case the total probability of obtaining a state in  $A_8$  is 6.25%. If we choose the initial phase  $\gamma_0$  to be 0, the acceptance set has four states and the probability of obtaining a state in  $A_8$  is 1.3%. For the non-adaptive phase estimation protocol,  $|A_8| = 66$  but the total probability of obtaining a state in  $A_8$  is roughly 79.2%. These results indicate that although the adaptive phase estimation protocol outperforms the non-adaptive protocol in the fault-free case [3, 23] the adaptive phase estimation protocol cannot be used in the presence of measurement readout errors for four rounds, and has significantly fewer states in  $A_8$  for eight rounds. Therefore in a noisy implementation of phase estimation, the non-adaptive protocol is preferable. A list of the states belonging to the acceptance set  $A_8$  for both the adaptive and non-adaptive protocols can be found at [https://github.com/godott/GKP\\_phase\\_estimation.git](https://github.com/godott/GKP_phase_estimation.git).

Suppose now that for  $m$  rounds the set  $A_m$  is not empty. We then know there exists a  $\tilde{\delta}$  such that the protocol is a  $(1, \tilde{\delta})$  fault-tolerant  $m$  round phase estimation protocol for *measurement* readout errors. However there are other fault locations where a single fault resulting in a qubit error could potentially lead to an uncorrectable shift error on output states. Note that an  $X$  error prior to applying the controlled- $D(\sqrt{2\pi})$  gate will do nothing since the qubit is in the  $|+\rangle$  state. A  $Z$  error will just change the sign of one of the peaks of the output state in  $p$  space. However a damping event occurring on the ancilla qubit before or *during* the application of the controlled- $D(\sqrt{2\pi})$  gate can result in incorrectly applying the  $D(\sqrt{2\pi})$  displacement on the cavity. If several damping events occur during the protocol, the output state in  $p$  space could potentially be badly deformed. In [3] it was proposed to replace the ancilla qubit by a  $k$ -qubit cat state so that if a single qubit undergoes damping, a single shift error of size  $D\left(\frac{\sqrt{2\pi}}{k}\right)$  would occur which can be made small for large  $k$ . However this approach would require the fault-tolerant preparation of a large  $k$ -qubit cat state which would significantly increase the required resources in addition to adding substantially more locations where errors can occur. Similar issues were observed in [24] for preparing cat codes. However large cavity displacements were mitigated by interacting the cavity with a three level system  $|f\rangle, |e\rangle$  and  $|g\rangle$  instead of a qubit. A transition from  $|f\rangle$  to  $|e\rangle$  would not cause the incorrect gate from being applied. Thus two damping events would be required to cause the incorrect gate from being applied.

Here we propose an alternative solution for mitigating damping errors during the application of the controlled- $D(\sqrt{2\pi})$  gate which uses a single extra flag qubit [25–30]. Consider the circuit in figure 3. If a bit flip error occurs before or during the application of the controlled- $D(\sqrt{2\pi})$  gate, the flag qubit will be measured as 1 instead of 0. In such a case the strategy will be simply to abort the protocol and start anew. The analysis for a

<sup>9</sup> However it is possible that the adaptive phase estimation protocol described in section 4.1 could be modified to be fault-tolerant to measurement readout errors. We leave such analysis to future work.



**Figure 3.** Phase estimation circuit with an additional flag qubit. If a damping event occurs during the application of the controlled- $D(\sqrt{2\pi})$  gate resulting in a  $D(\sqrt{2\pi})$  error, the flag qubit will be measured as 1 instead of 0. If the flag is measured as 1, we abort the protocol and start anew.

general damping event using the flag qubit is given in appendix A. It is shown that if ever a damping event on the  $|+\rangle$  ancilla qubit causes the wrong  $D(\sqrt{2\pi})$  gate to be applied, the flag qubit will be measured as 1 instead of 0 in which case we abort the protocol and start anew. Thus with the flag qubit, the phase estimation protocol is robust to a damping error during the application of the controlled displacement gate. Note however that if a fault occurs in addition to a damping event, the shift error resulting from the damping event can potentially go undetected. For instance, if in addition to damping, the flag qubit is subject to a measurement error, the measurement outcome can be 0 instead of 1 resulting in acceptance when the protocol should have been aborted.

Suppose that only a single damping event occurs during the four round phase estimation protocol, and that all other operations are fault-free. Using the analytic expressions derived in appendix A and for a damping rate of  $p = 0.5$ , we performed simulations to numerically characterize the effects of qubit damping on the prepared GKP state. We found that the shift error resulting from a single damping event was negligible compared to the largest shift error resulting from a single measurement readout error.

Lastly, an  $X$  error prior to the  $\Lambda(e^{i\gamma})$  gate will change the sign of  $\gamma$ . After propagating through the Hadamard gate, the  $X$  error will become a  $Z$  error which will not affect the measurement outcome of the ancilla qubit. For the four round non-adaptive phase estimation protocol, we performed a simulation which showed that the shift error resulting from a single  $X$  error prior to the  $\Lambda(e^{i\gamma})$  gate is much smaller than the largest shift error arising from a measurement readout error on the ancilla qubit.

Let us assume that the noise model afflicting the oscillator can cause a shift error of size at most  $\tilde{\delta}$  in  $p$  space (i.e. a shift of the form  $e^{-i\tilde{\delta}q}$ ). From the above, the largest shift error that can arise from a single fault afflicting the qubit space during the four round fault-tolerant phase estimation protocol is  $0.235 < \sqrt{\pi}/6$ . Following definition 5, we conclude that the protocol described in this section is a  $(1, \tilde{\delta})$ -fault-tolerant state preparation of an approximate logical  $|\tilde{0}\rangle$  GKP state with  $\tilde{\delta} = \sqrt{\pi}/6 - 0.235 \approx 0.06$ .

In section 5 we will discuss a physical implementation of the phase estimation protocol using circuit QED. A much more detailed analysis of the noise afflicting the cavity and the resulting shift errors will be provided.

#### 4.3. Goodness of approximate $|\tilde{0}\rangle$ states obtained from the noise free phase estimation protocol

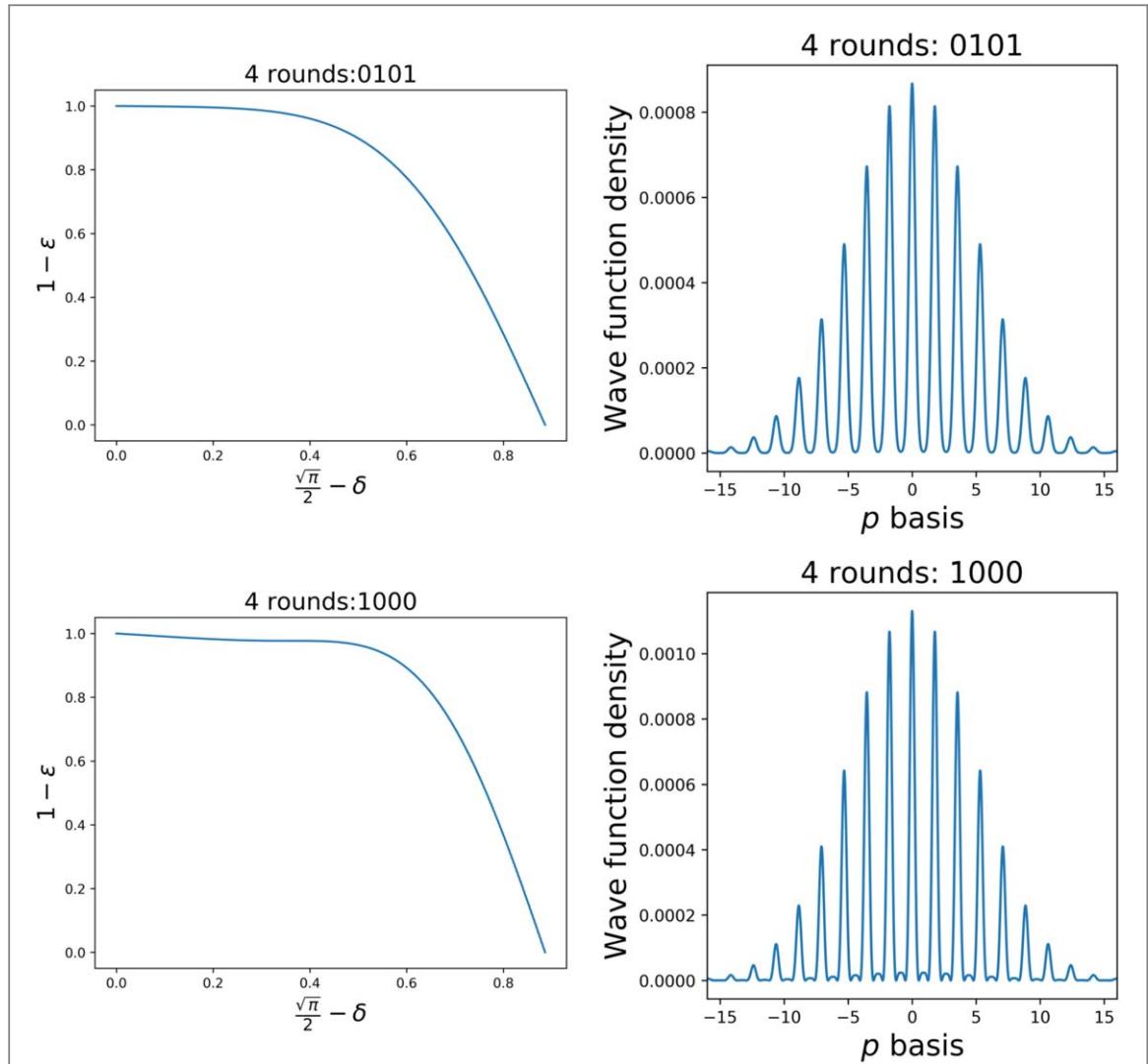
In section 5.2, the fault tolerant implementation of phase estimation described in this section is analyzed for a noise model which introduces gate noise, measurement readout errors and ancilla state preparation errors in addition to photon loss, amplitude damping and dephasing. Here we consider a noiseless implementation of the four round non-adaptive phase estimation protocol described in this section which will be used to benchmark the noisy implementation.

Since the shift correction obtained from non-adaptive phase estimation is in  $p$  space, for a range of values for  $\delta$ , we computed  $\epsilon$  using the integral in definition 1 after performing a shift back correction using equation (9). This allowed us to compute the probability of correcting shift errors in  $p$  of size  $\frac{\sqrt{\pi}}{2} - \delta$ . Plots for the states 0101 and 1000 which belong to  $A_4$  (see table 1) are given in figure 4. The  $\left(\frac{\sqrt{\pi}}{2} - \delta, \epsilon\right)_p$  correctable properties of the other states in  $A_4$  are similar to the ones shown in figure 4. It can be seen that for small values of size  $\frac{\sqrt{\pi}}{2} - \delta$ , both states can correct the shift error with probability close to 1.

## 5. Circuit QED implementation

### 5.1. State evolution in the dispersive regime

In this section we will describe a direct implementation of the controlled- $D(\sqrt{2\pi})$  gate. From [31, 32], the Hamiltonian describing the coupling between a qubit and a cavity in the dispersive regime is given by



**Figure 4.** (Left) Plot illustrating the probability of correcting a shift error of size  $\frac{\sqrt{\pi}}{2} - \delta$  for the states 0101 and 1000 obtained via a non-adaptive noise free simulation of the phase estimation protocol presented in section 4. (Right) Wave function density  $|\psi(p)|^2$  of the states 0101 and 1000 illustrated in  $p$  space. The horizontal axis corresponds to values of  $p$ .

$$H_s = \tilde{\omega}_r a^\dagger a + \tilde{\omega}_a Z + \chi a^\dagger a Z - \phi (a^\dagger a)^2 Z, \quad (13)$$

where  $\phi = \frac{g^4}{\Delta^3}$ ,  $\chi = \frac{g^2}{\Delta} - \phi$ ,  $\tilde{\omega}_r = \omega_r + \phi$  and  $\tilde{\omega}_a = \frac{\omega_r + \chi}{2}$ . Here  $g$  is the coupling strength between the qubit and the cavity and  $\Delta = |\omega_r - \omega_a|$ . The term  $\phi (a^\dagger a)^2 Z$  corresponds to the nonlinear dispersive shift. The Hamiltonian in equation (13) can be derived by performing an exact diagonalization of the Jaynes–Cummings Hamiltonian and keeping only leading order terms in  $\phi$  [32]. Note that a more systematic treatment of the qubit as an anharmonic oscillator leads to an additional term in equation (13) given by  $-\frac{K}{2} (a^\dagger a)^2$  which is referred to as the Kerr nonlinearity [33]. Hence, in our analysis, we choose the system Hamiltonian to be given by

$$H_s = \tilde{\omega}_r a^\dagger a + \tilde{\omega}_a Z + \chi a^\dagger a Z - \phi (a^\dagger a)^2 Z - \frac{K}{2} (a^\dagger a)^2. \quad (14)$$

The direct implementation of the controlled- $D(\sqrt{2\pi})$  gate can be achieved using the drive Hamiltonian

$$H_d(t) = \mathcal{E}(t) a^\dagger e^{-i\omega_d t} + \mathcal{E}^*(t) a e^{i\omega_d t}, \quad (15)$$

where  $\mathcal{E}(t)$  describes the pulse shape of the drive and  $\omega_d$  is the drive frequency. Thus the total Hamiltonian describing the evolution of the qubit-cavity system during the implementation of the controlled- $D(\sqrt{2\pi})$  gate is given by

$$H(t) = H_s + H_d(t). \quad (16)$$

Going into the rotating frame of the qubit and the cavity and defining  $\phi_{\pm} = \phi \pm K$  and  $\omega_{\pm} = \pm \chi$ , in appendix B we show that

**Table 2.** Parameter values for the two simulations of the non-adaptive noisy phase estimation protocol. The values for  $\kappa$  are chosen based on state of the art 3D cavities. The parameters in the second column results in a  $T_1$  time given by  $T_1 = \frac{2\pi}{\kappa} = 10$  ms. For a resonator frequency  $f_r = 7$  GHz, the quality factor is given by  $Q = \frac{f_r}{\kappa / (2\pi)} = 7 \times 10^7$ . For all simulations, the squeezing level of the input squeezed vacuum state is chosen to be 0.2.

Parameter values	Simulation 1	Simulation 2
$p$	$5 \times 10^{-3}$	$10^{-3}$
$\frac{\kappa}{2\pi}$	$1.59 \times 10^{-5}$ MHz	$1.59 \times 10^{-6}$ MHz
$\frac{\gamma_1}{2\pi}$	$1.06 \times 10^{-3}$ MHz	$1.06 \times 10^{-4}$ MHz
$\frac{\gamma_\phi}{2\pi}$	$7.96 \times 10^{-4}$ MHz	$7.96 \times 10^{-5}$ MHz
$g$	8.92 MHz	5.09 MHz
$\Delta$	0.32 GHz	0.32 GHz
$K$ (Kerr)	$10^{-4}$ MHz	$10^{-5}$ MHz

$$V_R(0, T) = \mathcal{T} e^{-i \int_0^T H(t') dt'} = R_+(T) D(A_+) e^{i B_+} |0\rangle \langle 0| + R_-(T) D(A_-) e^{i B_-} |1\rangle \langle 1|, \quad (17)$$

where

$$R_\pm(T) = e^{-i T \omega_\pm a^\dagger a} (1 \pm i T \phi_\pm (a^\dagger a)^2), \quad (18)$$

$$A_\pm = -i \int_0^T \mathcal{E}(t) e^{i \omega_\pm t} dt \mp \phi_\pm (2a^\dagger a - 1) \int_0^T \mathcal{E}(t) t e^{i \omega_\pm t} dt + \mathcal{O}(\phi_\pm^2), \quad (19)$$

$$B_\pm = -\frac{i}{2} \int_0^T dt \int_t^T dt' (\mathcal{E}^*(t') \mathcal{E}(t) e^{i \omega_\pm (t' - t)} - \mathcal{E}(t') \mathcal{E}^*(t) e^{-i \omega_\pm (t' - t)}) \pm \frac{\phi_\pm (2a^\dagger a + 1)}{2} \int_0^T dt \int_t^T dt' (\mathcal{E}^*(t') \mathcal{E}(t) e^{i \omega_\pm (t' - t)} + \mathcal{E}(t') \mathcal{E}^*(t) e^{i \omega_\pm (t' - t)}) (t' - t) + \mathcal{O}(\phi_\pm^2). \quad (20)$$

In deriving equations (18)–(20), we kept only leading order terms in  $\phi_\pm$  since the Hamiltonian in equation (14) neglects higher order terms in  $\phi$  and  $K$ . We keep only leading order terms since with current experimental parameter values,  $\phi_\pm$  is roughly three orders of magnitude smaller than  $\chi$ .

Firstly, notice that the terms  $R_\pm(T)$  introduce a relative rotation between the  $|0\rangle$  and  $|1\rangle$  state of the qubit. Neglecting the terms proportional to  $\phi_\pm$ , it was shown in [3] that by choosing the total interaction time to be  $T = \pi/\chi$ , the relative rotation between  $|0\rangle$  and  $|1\rangle$  can be set to one. However due to the presence of the nonlinear dispersive shift and Kerr terms in equation (18), it is clear that the relative rotation cannot be completely removed. Further,  $R_\pm(T)$  does not depend on the pulse shape  $\mathcal{E}(t)$  of the drive term. Therefore even with an optimized pulse shape (see below), the effects from the nonlinear dispersive shift and the Kerr will cause an undesired relative rotation between the qubit states. Regardless, we will still choose the interaction time  $T$  to be  $\pi/\chi$  to ensure that we eliminate the relative rotation from the leading order terms in equation (18). In appendix B.4, we provide further details using the analytic expressions to show that to mitigate the effects due to the nonlinear dispersive shift and Kerr terms would require the protocol to take place on time scales of order  $\frac{1}{\phi_\pm}$  which (for the parameter values in table 2) is four orders of magnitude larger than  $\frac{\pi}{\chi}$ . For such long time scales, noise terms such as photon loss, dephasing and damping would render the protocol impractical.

The terms in equation (17) that perform the desired controlled displacement gate are given by  $A_\pm$  in equation (19). The goal is to choose a pulse shape that implements the desired gate while at the same time minimizes the contributions from the nonlinear dispersive shift and the Kerr term (terms proportional to  $\phi_\pm$  in equation (19)). In order to be experimentally relevant, it is important to choose a pulse shape that is accessible to near term experiments using 2D and 3D cavities. We chose a Gaussian pulse with the following parameters

$$\mathcal{E}(t) \approx -2.09562 \left( \frac{\sqrt{2\pi}}{T} \right) e^{-\frac{(t-\mu)^2}{2\sigma^2}}, \quad (21)$$

where  $\mu = \frac{\pi}{2\chi}$  and  $\sigma = \frac{\pi}{10\chi}$ . With this Gaussian pulse we obtain

$$A_\pm \approx \mp \frac{\sqrt{2\pi}}{2} \mp \phi_\pm (2a^\dagger a - 1) \int_0^T \mathcal{E}(t) t e^{i \omega_\pm t} dt. \quad (22)$$

The amplitude of the Gaussian, given by  $-2.09562 \left( \frac{\sqrt{2\pi}}{T} \right)$ , is chosen numerically to ensure that the appropriate gate is being performed.

Since both the  $A_+$  and  $A_-$  terms are symmetric with opposite signs, the gate  $D\left(-\sqrt{\frac{\pi}{2}}\right)$  in figure 3 is *not* required. Furthermore, the first term in equation (22) is independent of  $\chi$ . However the contribution from the second term in equation (22) will depend on the coupling strength and relative frequencies between the qubit and the cavity. Thus reducing the value of  $\chi$  will minimize the undesired effects arising from the nonlinear dispersive shift and the Kerr term while still producing the desired gate.

The  $B_\pm$  terms of equation (20) result in a phase difference between the  $|0\rangle$  and  $|1\rangle$  qubit states. However, computing the first integrals in equation (20) (i.e. terms before  $\phi_\pm$ ), we find that for our chosen pulse, the phase difference remains *fixed* during every round of the phase estimation protocol. Therefore after applying the controlled displacement gate, we apply an additional phase gate which removes the phases introduced by the left most integrals of  $B_\pm$ . Note however that the  $\phi_\pm$  terms in  $B_\pm$  will not be canceled and will thus introduce additional shift errors.

## 5.2. Full noise analysis and master equation results

In this section we perform a numerical analysis for the noisy implementation of the phase estimation protocol described in section 4.2. The simulation is performed in several steps that we now describe. First, the controlled displacement gate is modeled using the following master equation

$$\dot{\rho} = -i[H(t), \rho] + \kappa\mathcal{D}[a]\rho + \gamma_1\mathcal{D}[\sigma_-]\rho + \gamma_\phi\mathcal{D}[\sigma_z]\rho, \quad (23)$$

where  $H(t)$  is given by equation (16) after going into the rotating frame and  $\mathcal{D}[L]\rho = (2L\rho L^\dagger - L^\dagger L\rho - \rho L^\dagger L)/2$ . The density matrix corresponds to the joint state of the cavity, ancilla qubit and flag qubit. The parameters  $\kappa$ ,  $\gamma_1$  and  $\gamma_\phi$  correspond to the photon loss rate, the qubit decay rate and qubit dephasing. The pulse shape of the drive is given by equation (21). The total evolution time of the controlled displacement is given by  $\frac{\pi}{\chi}$  with  $\chi = \frac{g^2}{\Delta} - \phi$ .

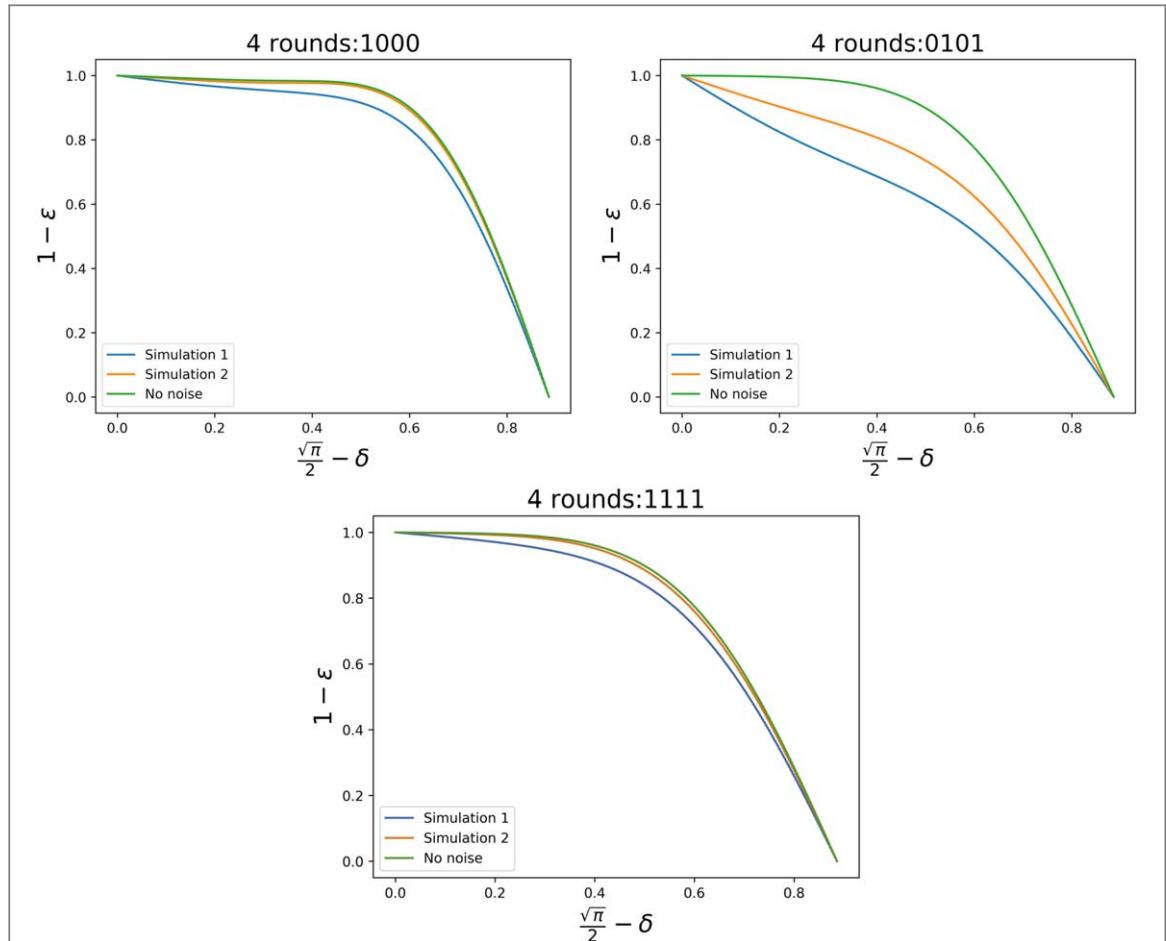
Both before and after the controlled displacement gate, the state of the cavity is subject to photon loss and its evolution is computed by solving a master equation. Based on current gate, state preparation and measurement times, we chose the evolution time from the preparation of the ancilla to the first CNOT gate (after which the controlled-displacement gate is performed) to be  $0.14\ \mu\text{s}$ . Similarly, the evolution time after the controlled-displacement gate to the time the ancilla is measured is also chosen to be  $0.14\ \mu\text{s}$ . Thus during one round, the cavity freely evolves with photon loss for a total of  $0.28\ \mu\text{s}$ . In addition, before and after the controlled displacement gate, we allow all qubit locations to fail with the following depolarizing noise model

1. With probability  $p$ , each two-qubit gate is followed by a two-qubit Pauli error drawn uniformly and independently from  $\{I, X, Y, Z\}^{\otimes 2} \setminus \{I \otimes I\}$ .
2. With probability  $\frac{2p}{3}$ , the preparation of the  $|0\rangle$  state is replaced by  $|1\rangle = X|0\rangle$ .
3. With probability  $\frac{2p}{3}$ , any single qubit measurement has its outcome flipped.
4. With probability  $p/10$ , each Hadamard gate is followed by a Pauli error drawn uniformly and independently from  $\{X, Y, Z\}$ .

We chose  $p/10$  for Hadamard gate failures since for current superconducting architectures, single qubit gate fidelities are about an order of magnitude higher than two-qubit gate fidelities. In practice, the gate errors applied to the qubit Hilbert space will depend strongly on the circuit QED architecture and should also be modeled using a master equation as in equation (23). However, we chose a depolarizing model in order to reduce the computation time and simplify the analysis. We also mention that in our analysis we assumed that  $g$  is tunable. Thus both before and after the controlled-displacement gate, the cavity and qubit system is decoupled and can be treated separately.

The master equation was numerically solved using Qutip [34], our code can be accessed at [https://github.com/godott/GKP\\_phase\\_estimation.git](https://github.com/godott/GKP_phase_estimation.git). Due to the long computation time during the controlled displacement gate, performing a full Monte Carlo simulation to take into account gate errors with the depolarizing model was unfeasible (i.e. gate locations both before and after the controlled displacement gate). Instead, we analytically computed the error probabilities for each Pauli operator (using the depolarizing noise model described above) immediately after the first CNOT gate, before both measurement locations and before the phase gate<sup>10</sup>. At each location, all possible Pauli operators based on their associated probabilities were added. For a given Pauli error, the probabilities (which are expressed as functions of  $p$ ) were then used (in addition to the state evolution

<sup>10</sup> The error probabilities were computed by considering all possible single-fault locations resulting in a given error at the considered location. For instance, a  $Z \otimes I$  error after the first CNOT gate can arise from a  $Z \otimes I$  error from the faulty CNOT gate, but also from a  $Z$  error after the application of the Hadamard gate prior to the CNOT gate.

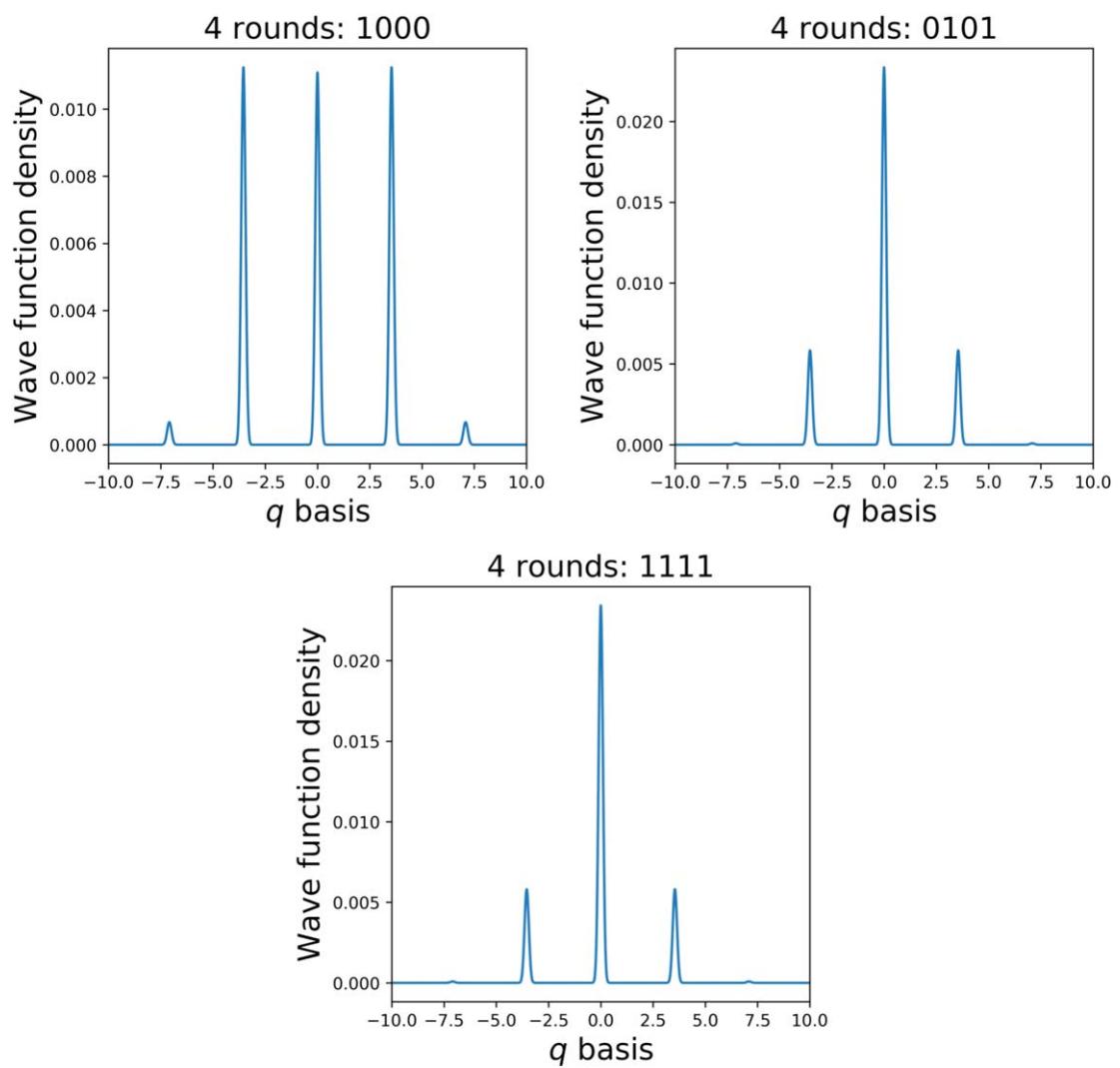


**Figure 5.** Probability of correcting a shift errors of size  $\frac{\sqrt{\pi}}{2} - \delta$  for parameters chosen from the second (labeled ‘Simulation 1’) and third (labeled ‘Simulation 2’) column of table 2 in addition to the case where no noise is present. The plots are for the states 1000, 0101 and 1111 obtained from the four round fault-tolerant phase estimation protocol described in section 4.

obtained from the master equation) to compute the final probability of obtaining a given output state. Instances with two or more faults occurring on the qubit Hilbert space before and after the controlled-displacement gate were neglected. However our analysis is still more complete than previous implementations which considered only measurement errors and errors during the controlled-displacement gate. More details of the Pauli simulation can be found in appendix D.

We performed two different simulations where for each simulation, the chosen parameter values are given in table 2. The parameters chosen for the first simulation (middle column in table 2) are based on current experimental values for 2D and 3D cavities [24, 35–39]. The parameters chosen for the second simulation are based on values that might be obtained with improved future technologies. Plots showing the probability of correcting shift errors  $\frac{\sqrt{\pi}}{2} - \delta$  after performing a shift correction of the output states of the noisy phase estimation protocol using equation (9) are given in figure 5. In what follows we will refer to these plots as  $\epsilon - \delta$  plots. Note that noise during the phase estimation protocol introduces more shift errors in  $p$  space. Therefore in figure 5, only  $\epsilon - \delta$  plots in  $p$  space are shown.

For the four round phase estimation protocol, we find that the state 1010 has a similar  $\epsilon - \delta$  plot to the one for the state 0101 whereas the state 0010 has a similar  $\epsilon - \delta$  plot to the one for the state 1000. The probability of obtaining each state for the two noisy simulations (parameters in table 2) are given in table 3. It can be seen that the probability of the state 0101 to correct shift errors is significantly more affected by the noise than the state 1000 or 1111. To understand why this is the case, it is useful to look at the wave function densities for these three states in the  $q$  basis when no noise is present (see figure 6). Comparing the wave function densities, it can be observed that the state 1000 has three peaks with similar amplitudes, whereas the states 0101 and 1111 have two smaller peaks compared to the peak at the center. Performing a numerical analysis, it is found that when only damping is present (with all other noise terms including Kerr and nonlinear dispersive shift set to zero), damping has a negligible effect on the height of the peaks for all considered states. However performing a numerical simulation with only the Kerr and nonlinear dispersive shift terms present, these terms significantly reduce the height of the peaks for the state 0101 and 1010. Therefore, these states are left with one large peak in  $q$



**Figure 6.** Wave function density  $|\psi(q)|^2$  of the states 1000, 0101 and 1111 obtained from the noise free non-adaptive phase estimation protocol.

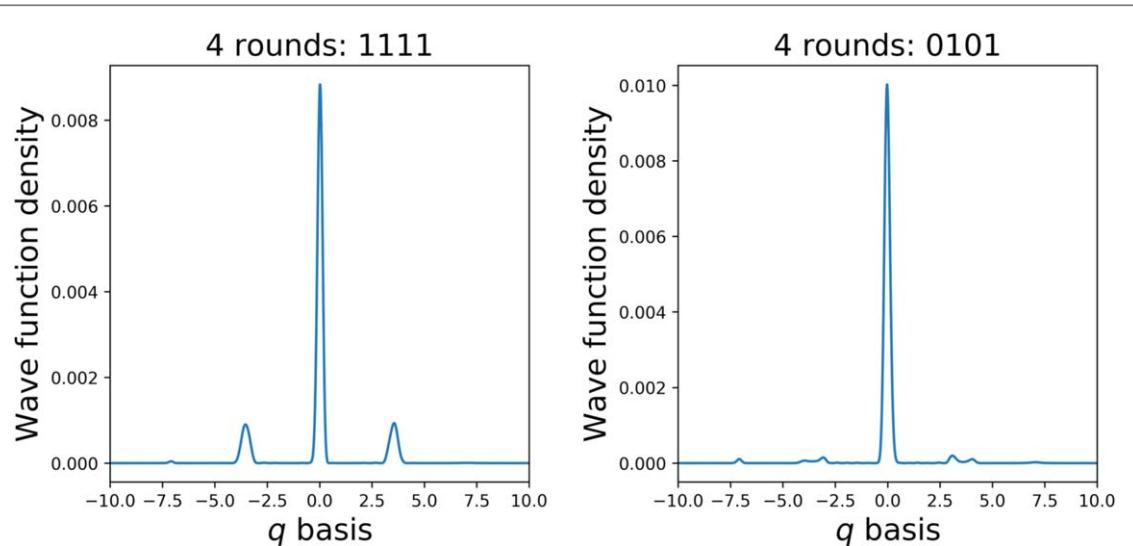
**Table 3.** The second row corresponds to the probabilities of obtaining the output states of  $A_4$  for the parameters chosen from the second column of table 2 conditioned on all flag measurements being 0. The third row is identical but for the parameters chosen from the third column of table 2. The probabilities are smaller for noisier circuits since the flag qubit has a higher chance of being measured as 1 causing the protocol to be aborted.

Four round phase estimation protocol acceptance set $A_4$	1111	1010	0101	1000	0010
Probability of obtaining output state noisy simulation 1	0.1297	0.0727	0.0737	0.0474	0.0429
Probability of obtaining output state noisy simulation 2	0.1468	0.0982	0.0978	0.0514	0.0448

with all other peaks close to zero which results in a wave function density with very low resolution in  $p$  space (which thus affects the ability for these states to correct shift errors in  $p$ ). Interestingly, the state 1111 is more robust to the Kerr and nonlinear dispersive shift contributions since there is a much smaller reduction in the amplitudes of the smaller peaks compared to those for the states 0101 and 1010 (see figure 7). Furthermore, the states 1000 and 0010 have three large peaks in  $q$  and thus even with a reduced amplitude, these states have a higher wave function density resolution in  $p$  space.

## 6. Conclusion

In this work we presented a fault-tolerant state preparation protocol for preparing GKP states using phase estimation. In section 2 we provided metrics for comparing how good approximate GKP states are at correcting shift errors in both  $q$  and  $p$  space. In section 3, we provided a definition for the  $(m, \tilde{\delta})$ -fault-tolerant state



**Figure 7.** Wave function density of the states 1111 and 0101 where all noise terms are excluded apart from the Kerr and nonlinear dispersive shift terms. Comparing to figure 6, the Kerr and nonlinear dispersive shift terms reduce the amplitudes of the two small peaks of the 0101 state significantly more than those for the 1111 state.

preparation of an approximate GKP state where  $m$  is the maximum number of allowable faults which can occur on the qubit Hilbert space and  $\delta$  is the maximum allowed shift error that can affect the oscillator. Using a non-adaptive phase estimation protocol with one ancilla qubit and one flag qubit, in section 4 we showed how the protocol can be made into a (1, 0.06)-fault-tolerant state preparation of an approximate GKP state. The flag qubit is used to detect damping errors. In addition, it was shown how the adaptive phase estimation protocol of section 4.1 can not be made fault-tolerant in the presence of measurement readout errors. For the four round non-adaptive phase estimation protocol, 5 of the 16 output states can be used in the presence of measurement readout errors. The total probability of obtaining the accepted states is approximately 0.48. In section 5, we considered how the protocol can be implemented in circuit QED. We first provided (to leading order) analytic expressions of the nonlinear dispersive shift and Kerr terms during the evolution of the qubit and cavity in the fault-tolerant phase estimation protocol. We used these expressions to find a Gaussian pulse shape that allows one to implement the desired gates. However to mitigate effects due to the nonlinear dispersive shift and Kerr terms would require the protocol to be implemented on time scales four orders of magnitude larger than those that were considered in this work. Due to the noise processes afflicting the system, such long time scales would render the protocol impractical. Performing two different simulations for both current and futuristic parameter values found in 2D and 3D cavities, we numerically solved a master equation to study the affects of qubit damping, dephasing and photon loss on the cavity in addition to gate and measurement errors during the protocol. It is shown numerically that 3 of the 5 accepted states (for the four round non-adaptive phase estimation protocol) are much more robust to noise arising from the nonlinear dispersive shift and Kerr terms and maintain good error correction capabilities even in the presence of noise.

The pulse shape used in this work to implement the controlled-displacement gate of the protocol was obtained from the analytical expressions describing the time evolution of the qubit-cavity system. An important direction for future work is to find a pulse shape using methods such as optimal control [33, 40] in order to obtain a pulse which can further mitigate the effects from the nonlinear dispersive shift and Kerr terms. Since the nonlinear dispersive shift and Kerr terms are the dominant source of noise that reduce the error correcting capabilities of approximate GKP states, using optimal control could potentially allow the protocol to be implemented using a larger number of rounds in order to obtain better approximate GKP states.

The fault-tolerant state preparation of approximate GKP states presented in this work is tailored to protocols that use phase estimation. An interesting direction for future work would be to find fault-tolerant implementations for preparing approximate GKP states that apply to broader schemes such as those found in [10, 13]. In addition, fault-tolerant state preparation protocols for hexagonal GKP codes could be analyzed since these offer better error correction capabilities than GKP codes on a square lattice. It would also be interesting to extend the ideas presented in this work beyond state preparation. For instance, fault-tolerant protocols for the implementation of logical gates would be of great interest.

## Acknowledgments

We thank Barbara Terhal, Daniel Weigand and Daniel Gottesman for useful feedback and discussions. We give special thanks to Zlatko Minev for the many useful discussions regarding circuit QED. YS performed the majority of the numerical simulations. CC performed the majority of the analytical calculations, developed the fault-tolerant schemes and wrote the majority of the manuscript. YS and AC acknowledges scientific discussions with Fred Chong that took place during the completion of this project. YS is funded in part by EPiQC, an NSF Expedition in Computing, under grant CCF-1730449 and NSF award STAQ, PFCQC-1818914. YS is also funded in part by the NSF QISE-NET fellowship under grant number 1747426. This work was completed in part with resources provided by the University of Chicago Research Computing Center.

## Appendix A. Shift error arising from amplitude damping before the controlled-displacement gate

Consider amplitude damping on the  $|+\rangle$  ancilla state before the controlled displacement gate as shown in figure A1. Let  $|\varphi\rangle$  be the input state part of the cavity Hilbert space. Defining  $p$  as the damping rate, after the controlled-displacement gate, the state of the system can be written as

$$|\psi^{(1)}\rangle = \frac{1}{\sqrt{2}} \{ |00\rangle |\varphi\rangle + \sqrt{1-p} |11\rangle D(\sqrt{2\pi}) |\varphi\rangle + \sqrt{p} |01\rangle |\varphi\rangle \}. \quad (\text{A1})$$

Here we have omitted writing the state of the environment interacting with the first ancilla qubit. However, this does not affect the result of the calculations that follow.

After the second CNOT gate, the state becomes

$$|\psi^{(2)}\rangle = \frac{1}{\sqrt{2}} \{ |00\rangle |\varphi\rangle + \sqrt{1-p} |10\rangle D(\sqrt{2\pi}) |\varphi\rangle + \sqrt{p} |01\rangle |\varphi\rangle \}. \quad (\text{A2})$$

If the flag qubit is measured as 1, the protocol is aborted. So assume that the flag is measured as 0. In this case the state becomes (tracing out the flag qubit)

$$|\psi^{(3)}\rangle = \frac{1}{\sqrt{2-p}} \{ |0\rangle |\varphi\rangle + \sqrt{1-p} |1\rangle D(\sqrt{2\pi}) |\varphi\rangle \}. \quad (\text{A3})$$

Applying the remaining gates in figure A1, the final state prior to the measurement of the ancilla qubit is

$$|\psi^{(4)}\rangle = \frac{1}{\sqrt{2(2-p)}} \{ |0\rangle (1 + e^{i\gamma} \sqrt{1-p} D(\sqrt{2\pi})) |\varphi\rangle + |1\rangle (1 - e^{i\gamma} \sqrt{1-p} D(\sqrt{2\pi})) |\varphi\rangle \}. \quad (\text{A4})$$

If the measurement of the ancilla is 0, the output state will be

$$|\psi^{\text{out},0}\rangle = \frac{1}{\sqrt{2-p}} \{ (1 + e^{i\gamma} \sqrt{1-p} D(\sqrt{2\pi})) |\varphi\rangle \}. \quad (\text{A5})$$

If the measurement of the ancilla is 1, the output state will be

$$|\psi^{\text{out},1}\rangle = \frac{1}{\sqrt{2-p}} \{ (1 - e^{i\gamma} \sqrt{1-p} D(\sqrt{2\pi})) |\varphi\rangle \}. \quad (\text{A6})$$

Both outcomes occur with probability 1/2.

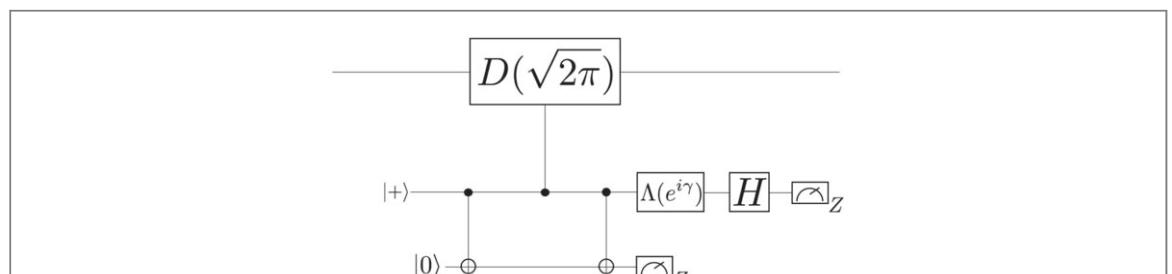


Figure A1. Controlled-displacement circuit with the flag qubit. The protocol is aborted if the flag qubit measurement is non-trivial.

## Appendix B. Analytic derivation of the unitary operator describing the evolution of the qubit-cavity system during the implementation of the controlled- $D(\sqrt{2\pi})$ gate

### B.1. Implementation of the controlled- $D(\sqrt{2\pi})$ gate in the lab frame

In this section, we will derive the unitary operator describing the evolution of the qubit-cavity state during the application of the control-displacement gate. In the dispersive regime, the qubit-cavity interaction can be described as:

$$H(t) = H_s + H_d(t), \quad (\text{B1})$$

where

$$H_s = \tilde{\omega}_r a^\dagger a + \tilde{\omega}_a Z + \chi a^\dagger a Z - \phi(a^\dagger a)^2 Z. \quad (\text{B2})$$

and

$$H_d(t) = \mathcal{E}(t)(a + a^\dagger). \quad (\text{B3})$$

In equation (B2),  $\phi = \frac{g^4}{\Delta^2}$ ,  $\tilde{\omega}_r = \omega_r + \phi$ ,  $\chi = \frac{g^2}{\Delta} - \phi$  and  $\tilde{\omega}_a = \frac{\omega_a + \chi}{2}$ . The term  $\phi(a^\dagger a)^2 Z$  corresponds to the nonlinear dispersive shift. Note that we have not included the Kerr term  $-\frac{K}{2}(a^\dagger a)^2$ . At the end of this section we will modify our results to take into account its effect. We also note that in writing the drive Hamiltonian in equation (B3), we neglected a term of the form  $\lambda X$  (where  $\lambda = (g/\Delta)$ ) and all higher powers in  $\lambda$ . For parameter values considered in this paper, we found the effect of this term to be negligible. Additionally, we chose a pulse shape which is real valued.

In equation (B3), we represent the drive pulse  $\mathcal{E}(t)$  as

$$\mathcal{E}(t) = \Omega_x(t) \cos(\omega_d t) + \Omega_y(t) \sin(\omega_d t), \quad (\text{B4})$$

where  $\omega_d = \tilde{\omega}_r - \chi$  is the drive frequency.

Since the Hamiltonian does not commute at different times, the unitary operator describing the time evolution under the Hamiltonian of equation (B1) is given by

$$V(0, T) = \mathcal{T}e^{-i \int_0^T H(t') dt'}. \quad (\text{B5})$$

The right-hand side of equation (B5) can be computed using the Suzuki-Trotter decomposition, so that

$$\begin{aligned} V(0, T) &= \left( \lim_{n \rightarrow \infty} \prod_{j=1}^n e^{-i\tilde{\omega}_r t_j} (a^\dagger a) e^{-i\tilde{\omega}_a t_j} (a + a^\dagger) \right) e^{-i\tilde{\omega}_a T} \\ &= \left( \lim_{n \rightarrow \infty} \prod_{j=1}^n e^{-i\tilde{\omega}_r t_j} (a^\dagger a) e^{-i\tilde{\omega}_a t_j} (a + a^\dagger) \right) e^{-i\tilde{\omega}_a T} |0\rangle \langle 0| \\ &\quad + \left( \lim_{n \rightarrow \infty} \prod_{j=1}^n e^{-i\tilde{\omega}_r t_j} (a^\dagger a) e^{-i\tilde{\omega}_a t_j} (a + a^\dagger) \right) e^{i\tilde{\omega}_a T} |1\rangle \langle 1|, \end{aligned} \quad (\text{B6})$$

where  $\tilde{t} \equiv T/n$ ,  $\omega_{\pm} \equiv \tilde{\omega}_r \pm \chi$ .

Now, the two terms on the right-hand side of equation (B6) can be decomposed as

$$\lim_{n \rightarrow \infty} \prod_{j=1}^n e^{-i\tilde{\omega}_r t_j} (a^\dagger a) e^{-i\tilde{\omega}_a t_j} (a + a^\dagger) = \lim_{n \rightarrow \infty} R_n^n (R_n^{-n} D_{t_n} R_n^n) \cdots (R_n^{-2} D_{t_2} R_n^2) (R_n^{-1} D_{t_1} R_n^1), \quad (\text{B7})$$

where

$$R_n = e^{-i\tilde{\omega}_r t_n} (a^\dagger a), \quad (\text{B8})$$

and

$$D_{t_j} = e^{-i\tilde{\omega}_a t_j} (a + a^\dagger). \quad (\text{B9})$$

Hence from the above, we see that we need to compute terms of the form

$$R_n^{-k} D_{t_k} R_n^k = e^{i\tilde{\omega}_r k t_n} (a^\dagger a) e^{-i\tilde{\omega}_a k t_n} (a + a^\dagger) e^{-i\tilde{\omega}_r k t_n} (a^\dagger a) e^{i\tilde{\omega}_a k t_n} (a + a^\dagger). \quad (\text{B10})$$

In order to compute the products appearing in equation (B10), we will use the identity  $Ae^B A^{-1} = e^{ABA^{-1}}$ . Defining

$$H' = k(\omega_r - \phi a^\dagger a) a^\dagger a, \quad (\text{B11})$$

and

$$A_{\tilde{t}} = e^{-i\tilde{t}\mathcal{E}(t_k)(a+a^\dagger)}, \quad (B12)$$

with

$$A(\tilde{t}) = e^{iH'\tilde{t}}A_{\tilde{t}}e^{-iH'\tilde{t}} \quad (B13)$$

we have that

$$A(\tilde{t}) = \exp(-i\tilde{t}\mathcal{E}(t_k)e^{iH'\tilde{t}}(a + a^\dagger)e^{-iH'\tilde{t}}), \quad (B14)$$

The term  $e^{iH'\tilde{t}}(a + a^\dagger)e^{-iH'\tilde{t}}$  in equation (B14) can be computed using Heisenberg's equation of motion. We obtain

$$\frac{da(\tilde{t})}{d\tilde{t}} = i[H', a(\tilde{t})], \quad (B15)$$

with

$$[H', a] = k(\phi' + 2\phi a^\dagger a)a, \quad (B16)$$

where we defined

$$\phi' \equiv \phi - \omega_+. \quad (B17)$$

We thus have the following set of coupled differential equations:

$$\frac{da(\tilde{t})}{d\tilde{t}} = ik(\phi' + 2\phi a^\dagger(\tilde{t})a(\tilde{t}))a(\tilde{t}), \quad (B18)$$

and

$$\frac{da^\dagger(\tilde{t})}{d\tilde{t}} = -ika^\dagger(\tilde{t})(\phi' + 2\phi a^\dagger(\tilde{t})a(\tilde{t})), \quad (B19)$$

The solutions to equations (B18) and (B19) are given by

$$a(\tilde{t}) = ae^{ik\tilde{t}(\phi' + 2\phi(a^\dagger a - 1))}, \quad (B20)$$

and

$$a^\dagger(\tilde{t}) = e^{-ik\tilde{t}(\phi' + 2\phi(a^\dagger a - 1))}a^\dagger. \quad (B21)$$

To see this, we take a derivative of equation (B20) to obtain

$$\frac{da(\tilde{t})}{d\tilde{t}} = ik\phi' a(\tilde{t}) + 2ik\phi a(\tilde{t})(a^\dagger a - 1). \quad (B22)$$

Comparing equations (B18) and (B22), we must have that

$$a(\tilde{t})(a^\dagger a - 1) = a^\dagger(\tilde{t})a^2(\tilde{t}). \quad (B23)$$

Using equations (B20) and (B21) and defining  $H_{te} \equiv -kt(\phi' + 2\phi(a^\dagger a - 1))$ , we have that

$$\begin{aligned} a^\dagger(\tilde{t})a^2(\tilde{t}) &= e^{iH_{te}\tilde{t}}a^\dagger e^{-iH_{te}\tilde{t}}e^{iH_{te}\tilde{t}}ae^{-iH_{te}\tilde{t}}ae^{-iH_{te}\tilde{t}} \\ &= a^\dagger e^{-2i\phi k\tilde{t}}ae^{2i\phi k\tilde{t}}ae^{-iH_{te}\tilde{t}} \\ &= a^\dagger aae^{-iH_{te}\tilde{t}} \\ &= a^\dagger aa(\tilde{t}), \end{aligned} \quad (B24)$$

as desired. A similar calculation can be done for  $a^\dagger(\tilde{t})$ . Hence we have that

$$A(\tilde{t}) = \exp(-i\tilde{t}\mathcal{E}(t_k)(ae^{i\Phi_k\tilde{t}} + e^{-i\Phi_k\tilde{t}}a^\dagger)), \quad (B25)$$

for

$$\Phi_k = k(\phi' + 2\phi(a^\dagger a - 1)). \quad (B26)$$

Writing  $V(0, T)$  (equation (B6)) as

$$V(0, T) = V_+(0, T)|0\rangle\langle 0| + V_-(0, T)|1\rangle\langle 1|, \quad (B27)$$

using equation (B25) and the fact that  $R_n^n = e^{-iT(\omega_+ - \phi a^\dagger a)a^\dagger a} \equiv R_+(T)$ , we can write

$$V_+(0, T) = R_+(T) \lim_{n \rightarrow \infty} \prod_{k=1}^n \exp\{A_k a^\dagger - a A_k^\dagger\}, \quad (B28)$$

where we defined

$$A_k \equiv -\frac{iT}{n}\mathcal{E}(t_k)e^{-\frac{iT\Phi_k}{n}}. \quad (B29)$$

Notice that we can write  $V_+(0, T)$  as products of displacements  $D(A_k)$ . However, now  $A_k$  is an operator instead of a complex number. In order to compute the products in (B28), we will need to obtain a relation for terms of the form  $D(A_k)D(A_j)$  (where  $A_k$  is given in equation (B29)). Since  $A_k$  is proportional to  $T/n$  and remembering that we will take the limit where  $n \rightarrow \infty$ , using Baker–Campbell–Hausdorff, we have the exact relation

$$\exp \{A_k a^\dagger - aA_k^\dagger + A_j a^\dagger - aA_j\} = D(A_k)D(A_j)\exp \left\{ -\frac{1}{2}[A_k a^\dagger - aA_k^\dagger, A_j a^\dagger - aA_j] \right\}. \quad (\text{B30})$$

The commutator on the right-hand side of equation (B30) has four terms

$$[A_k a^\dagger - aA_k^\dagger, A_j a^\dagger - aA_j] = [A_k a^\dagger, A_j a^\dagger] - [A_k a^\dagger, aA_j^\dagger] - [aA_k^\dagger, A_j a^\dagger] + [aA_k^\dagger, aA_j^\dagger]. \quad (\text{B31})$$

As we will show, two of these terms vanish when taking the limit  $n \rightarrow \infty$ .

First, we compute

$$[A_k a^\dagger, A_j a^\dagger] = \left( \frac{-iT}{n} \right)^2 \mathcal{E}(t_k) \mathcal{E}(t_j) [e^{-\frac{iT}{n}\Phi_k} a^\dagger, e^{-\frac{iT}{n}\Phi_j} a^\dagger], \quad (\text{B32})$$

with

$$\begin{aligned} [e^{-\frac{iT}{n}\Phi_k} a^\dagger, e^{-\frac{iT}{n}\Phi_j} a^\dagger] &= e^{-\frac{iT}{n}\Phi_k} a^\dagger e^{-\frac{iT}{n}\Phi_j} a^\dagger - e^{-\frac{iT}{n}\Phi_j} a^\dagger e^{-\frac{iT}{n}\Phi_k} a^\dagger \\ &= e^{-\frac{iT}{n}\Phi_k} a^\dagger e^{\frac{iT}{n}\Phi_k} e^{-\frac{iT}{n}\Phi_{(k+j)}} a^\dagger e^{\frac{iT}{n}\Phi_{(k+j)}} e^{-\frac{iT}{n}\Phi_{k+j}} - e^{-\frac{iT}{n}\Phi_j} a^\dagger e^{\frac{iT}{n}\Phi_j} e^{-\frac{iT}{n}\Phi_{(k+j)}} a^\dagger e^{\frac{iT}{n}\Phi_{(k+j)}} e^{-\frac{iT}{n}\Phi_{(k+j)}}. \end{aligned} \quad (\text{B33})$$

Now, terms such as  $e^{-\frac{iT}{n}\Phi_j} a^\dagger e^{\frac{iT}{n}\Phi_j}$  can be computed by defining

$$H_{\text{temp}} \equiv -j\phi'(1 + \tilde{\delta}(a^\dagger a - 1)), \quad (\text{B34})$$

and invoking Heisenberg's equation of motion. Using  $[H_{\text{temp}}, a^\dagger] = -j\phi'\tilde{\delta}a^\dagger$  and solving the differential equation, we obtain

$$e^{-\frac{iT}{n}\Phi_j} a^\dagger e^{\frac{iT}{n}\Phi_j} = e^{-\frac{iT}{n}j\phi'\tilde{\delta}} a^\dagger. \quad (\text{B35})$$

Using the result of equation (B35) into equation (B33), we have

$$\begin{aligned} [e^{-\frac{iT}{n}\Phi_k} a^\dagger, e^{-\frac{iT}{n}\Phi_j} a^\dagger] &= e^{-\frac{iT}{n}k\phi'\tilde{\delta}} a^\dagger e^{-\frac{iT}{n}(k+j)\phi'\tilde{\delta}} a^\dagger e^{-\frac{iT}{n}\Phi_{(k+j)}} - e^{-\frac{iT}{n}j\phi'\tilde{\delta}} a^\dagger e^{-\frac{iT}{n}(j+k)\phi'\tilde{\delta}} e^{-\frac{iT}{n}\Phi_{(k+j)}} \\ &= e^{-\frac{iT}{n}(j+k)\phi'\tilde{\delta}} (e^{-\frac{iT}{n}k\phi'\tilde{\delta}} - e^{-\frac{iT}{n}j\phi'\tilde{\delta}}) (a^\dagger)^2 e^{-\frac{iT}{n}\Phi_{(k+j)}}. \end{aligned} \quad (\text{B36})$$

Using  $\phi'\tilde{\delta} = 2\phi$  and expanding equation (B36) to leading order in  $\phi$ , we obtain

$$\begin{aligned} [e^{-\frac{iT}{n}\Phi_k} a^\dagger, e^{-\frac{iT}{n}\Phi_j} a^\dagger] &\approx \left( 1 - \frac{2iT}{n}(j+k)\phi \right) \left( \frac{-2iT}{n}\phi \right) (a^\dagger)^2 e^{-\frac{iT}{n}\Phi_{(k+j)}} \\ &\approx \frac{-2iT}{n}(k-j)\phi (a^\dagger)^2 e^{-\frac{iT}{n}\Phi_{(k+j)}}. \end{aligned} \quad (\text{B37})$$

Inserting equation (B37) into equation (B32), we obtain

$$[A_k a^\dagger, A_j a^\dagger] = 2 \left( \frac{-iT}{n} \right)^3 (k-j)\phi \mathcal{E}(t_k) \mathcal{E}(t_j) (a^\dagger)^2 e^{-\frac{iT}{n}\Phi_{(k+j)}} + \mathcal{O}(\phi^2). \quad (\text{B38})$$

A similar calculation shows that

$$[aA_k^\dagger, aA_j^\dagger] = 2 \left( \frac{iT}{n} \right)^3 (k-j)\phi \mathcal{E}(t_k) \mathcal{E}(t_j) e^{\frac{iT}{n}\Phi_{(k+j)}} a^2 + \mathcal{O}(\phi^2). \quad (\text{B39})$$

Next we compute the cross terms. We first have

$$[A_k a^\dagger, aA_j^\dagger] = \left( \frac{iT}{n} \right)^2 \mathcal{E}(t_k) \mathcal{E}(t_j) [e^{-\frac{iT}{n}\Phi_k} a^\dagger, a e^{\frac{iT}{n}\Phi_j}], \quad (\text{B40})$$

with

$$\begin{aligned} [e^{-\frac{iT}{n}\Phi_k} a^\dagger, a e^{\frac{iT}{n}\Phi_j}] &= e^{-\frac{iT}{n}\Phi_k} a^\dagger a e^{\frac{iT}{n}\Phi_j} - a e^{\frac{iT}{n}\Phi_j} e^{-\frac{iT}{n}\Phi_k} a^\dagger \\ &= e^{-\frac{iT}{n}\Phi_k} a^\dagger e^{\frac{iT}{n}\Phi_k} e^{-\frac{iT}{n}\Phi_k} a e^{\frac{iT}{n}\Phi_k} e^{-\frac{iT}{n}\Phi_k} e^{\frac{iT}{n}\Phi_j} - a e^{\frac{iT}{n}\Phi_{(k-j)}} a^\dagger \\ &= a^\dagger a e^{\frac{-iT}{n}\Phi_{(k-j)}} - a e^{\frac{-iT}{n}\Phi_{(k-j)}} a^\dagger e^{\frac{iT}{n}\Phi_{(k-j)}} e^{-\frac{iT}{n}\Phi_{(k-j)}} \\ &= (a^\dagger a - a a^\dagger e^{\frac{-2iT}{n}\phi(k-j)}) e^{\frac{-iT}{n}\Phi_{(k-j)}} \\ &= (a^\dagger a (1 - e^{\frac{-2iT}{n}\phi(k-j)}) - e^{\frac{-2iT}{n}\phi(k-j)} e^{\frac{-iT}{n}\Phi_{(k-j)}}) e^{\frac{-iT}{n}\Phi_{(k-j)}}. \end{aligned} \quad (\text{B41})$$

Expanding the term proportional to  $a^\dagger a$  to leading order in  $\phi$ , equation (B41) becomes

$$[e^{\frac{-iT}{n}\Phi_k}a^\dagger, a e^{\frac{iT}{n}\Phi_j}] \approx \left( \frac{2iT}{n} \phi a^\dagger a(k-j) - e^{\frac{-2iT}{n}\phi(k-j)} \right) e^{\frac{-iT}{n}\Phi_{(k-j)}}. \quad (\text{B42})$$

Inserting equation (B42) into equation (B40), we obtain

$$[A_k a^\dagger, A_j a^\dagger] \approx 2 \left( \frac{iT}{n} \right)^3 \phi(k-j) \mathcal{E}(t_k) \mathcal{E}(t_j) a^\dagger a e^{\frac{-iT}{n}\Phi_{(k-j)}} - \left( \frac{iT}{n} \right)^2 \mathcal{E}(t_k) \mathcal{E}(t_j) e^{\frac{-2iT}{n}\phi(k-j)} e^{\frac{-iT}{n}\Phi_{(k-j)}}. \quad (\text{B43})$$

The last commutator to compute is

$$[a A_k^\dagger, A_j a^\dagger] = \left( \frac{-iT}{n} \right)^2 \mathcal{E}(t_k) \mathcal{E}(t_j) [a e^{\frac{iT}{n}\Phi_k}, e^{\frac{-iT}{n}\Phi_j} a^\dagger], \quad (\text{B44})$$

with

$$\begin{aligned} [a e^{\frac{iT}{n}\Phi_k}, e^{\frac{-iT}{n}\Phi_j} a^\dagger] &= a e^{\frac{iT}{n}\Phi_{(k-j)}} a^\dagger - e^{\frac{-iT}{n}\Phi_j} a^\dagger a e^{\frac{iT}{n}\Phi_k} \\ &= e^{\frac{iT}{n}\Phi_{(k-j)}} e^{\frac{-iT}{n}\Phi_{(k-j)}} a e^{\frac{iT}{n}\Phi_{(k-j)}} a^\dagger - e^{\frac{iT}{n}\Phi_{(k-j)}} e^{\frac{-iT}{n}\Phi_k} a^\dagger e^{\frac{iT}{n}\Phi_k} e^{\frac{-iT}{n}\Phi_k} a e^{\frac{iT}{n}\Phi_k} \\ &= e^{\frac{iT}{n}\Phi_{(k-j)}} (a^\dagger a (e^{\frac{2iT}{n}\phi(k-j)} - 1) + e^{\frac{2iT}{n}\phi(k-j)}) \\ &\approx e^{\frac{iT}{n}\Phi_{(k-j)}} \left( \frac{2iT}{n} \phi(k-j) a^\dagger a + e^{\frac{2iT}{n}\phi(k-j)} \right) \end{aligned} \quad (\text{B45})$$

so that

$$[a A_k^\dagger, A_j a^\dagger] \approx 2 \left( \frac{iT}{n} \right)^3 \phi(k-j) \mathcal{E}(t_k) \mathcal{E}(t_j) e^{\frac{iT}{n}\Phi_{(k-j)}} a^\dagger a + \left( \frac{-iT}{n} \right)^2 \mathcal{E}(t_k) \mathcal{E}(t_j) e^{\frac{iT}{n}\Phi_{(k-j)}} e^{\frac{2iT}{n}\phi(k-j)} \quad (\text{B46})$$

Inserting equations (B38), (B39), (B43) and (B46) into equation (B30), we have

$$\begin{aligned} D(A_k + A_j) &= D(A_k) D(A_j) \exp \left[ -\frac{1}{2} \left\{ 2 \left( \frac{-iT}{n} \right)^3 (k-j) \phi \mathcal{E}(t_k) \mathcal{E}(t_j) (a^\dagger)^2 e^{\frac{-iT}{n}\Phi_{(k+j)}} \right. \right. \\ &\quad + 2 \left( \frac{iT}{n} \right)^3 (k-j) \phi \mathcal{E}(t_k) \mathcal{E}(t_j) e^{\frac{iT}{n}\Phi_{(k+j)}} a^2 - 2 \left( \frac{iT}{n} \right)^3 \phi(k-j) \mathcal{E}(t_k) \mathcal{E}(t_j) a^\dagger a e^{\frac{-iT}{n}\Phi_{(k-j)}} \\ &\quad + \left( \frac{iT}{n} \right)^2 \mathcal{E}(t_k) \mathcal{E}(t_j) e^{\frac{-2iT}{n}\phi(k-j)} e^{\frac{-iT}{n}\Phi_{(k-j)}} - 2 \left( \frac{iT}{n} \right)^3 \phi(k-j) \mathcal{E}(t_k) \mathcal{E}(t_j) e^{\frac{iT}{n}\Phi_{(k-j)}} a^\dagger a \\ &\quad \left. \left. - \left( \frac{-iT}{n} \right)^2 \mathcal{E}(t_k) \mathcal{E}(t_j) e^{\frac{iT}{n}\Phi_{(k-j)}} e^{\frac{2iT}{n}\phi(k-j)} \right\} \right]. \end{aligned} \quad (\text{B47})$$

Since the product  $D(A_n) D(A_{n-1}) \cdots D(A_2) D(A_1)$  will produce sums in the exponents proportional to  $n^2$ , all terms in equation (B47) proportional to  $\left( \frac{T}{n} \right)^3$  will vanish in the limit where  $n \rightarrow \infty$ . Hence equation (B47) simplifies to

$$\begin{aligned} D(A_k + A_j) &= D(A_k) D(A_j) \exp \left[ i \mathcal{E}(t_k) \mathcal{E}(t_j) \left( \frac{T}{n} \right)^2 \left\{ \frac{e^{\frac{-2iT}{n}\phi(k-j)} e^{\frac{-iT}{n}\Phi_{(k-j)}} - e^{\frac{2iT}{n}\phi(k-j)} e^{\frac{iT}{n}\Phi_{(k-j)}}}{2i} \right\} \right] \\ &= D(A_k) D(A_j) \exp \left[ -i \mathcal{E}(t_k) \mathcal{E}(t_j) \left( \frac{T}{n} \right)^2 \sin \left( \frac{T}{n} \bar{\Phi}_{(k-j)} \right) \right], \end{aligned} \quad (\text{B48})$$

where  $\bar{\Phi}_{(k-j)}$  is defined as

$$\bar{\Phi}_{(k-j)} \equiv (k-j)(\phi - \omega_+ + 2\phi a^\dagger a). \quad (\text{B49})$$

We thus have that the product of the modified displacement operators is given by

$$D(A_k) D(A_j) = D(A_k + A_j) \exp \left[ i \mathcal{E}(t_k) \mathcal{E}(t_j) \left( \frac{T}{n} \right)^2 \sin \left( \frac{T}{n} \bar{\Phi}_{(k-j)} \right) \right]. \quad (\text{B50})$$

We now wish to compute

$$P_n \equiv D(A_n) D(A_{n-1}) D(A_{n-2}) \cdots D(A_2) D(A_1). \quad (\text{B51})$$

First notice that for any operators  $A$  and  $B$ , to leading order  $e^A e^B = e^B e^A e^{[A, B]}$ . Now for terms of the form  $\exp \left[ i \mathcal{E}(t_k) \mathcal{E}(t_j) \left( \frac{T}{n} \right)^2 \sin \left( \frac{T}{n} \bar{\Phi}_{(k-j)} \right) \right] D(\sum_m A_m)$ , the commutator of the two exponents will be proportional to  $\left( \frac{T}{n} \right)^3$  which will vanish in the limit where  $n \rightarrow \infty$ . Hence, we can commute all terms  $\exp \left[ i \mathcal{E}(t_k) \mathcal{E}(t_j) \left( \frac{T}{n} \right)^2 \sin \left( \frac{T}{n} \bar{\Phi}_{(k-j)} \right) \right]$  to the right-hand side of  $P_n$ . Hence we have

$$P_n = D \left( \sum_{k=1}^n A_k \right) \exp \left[ i \sum_{k < j} \frac{T^2}{n^2} \mathcal{E}(t_k) \mathcal{E}(t_j) \sin \left( \frac{T}{n} \bar{\Phi}_{(k-j)} \right) \right]. \quad (\text{B52})$$

Now, taking the limit where  $n \rightarrow \infty$  of  $P_n$ , we obtain

$$\begin{aligned} \lim_{n \rightarrow \infty} P_n &= D \left( \lim_{n \rightarrow \infty} \sum_{k=1}^n \left( \frac{-iT}{n} \right) \mathcal{E}(t_k) e^{-\frac{iT}{n} \Phi_k} \right) \exp \left[ i \lim_{n \rightarrow \infty} \sum_{k < j} \frac{T^2}{n^2} \mathcal{E}(t_k) \mathcal{E}(t_j) \sin \left( \frac{T}{n} \bar{\Phi}_{(k-j)} \right) \right] \\ &= D \left( -i \int_0^T \mathcal{E}(t) e^{-i\Phi t} dt \right) \exp \left[ -i \int_0^T dt \int_t^T dt' \mathcal{E}(t) \mathcal{E}(t') \sin(\bar{\Phi}(t' - t)) \right] \end{aligned} \quad (\text{B53})$$

We conclude that

$$V_+(0, T) = R_+(T) D(A_+) e^{iB_+}, \quad (\text{B54})$$

where

$$R_+(T) = e^{-iT(\omega_+ - \phi a^\dagger a) a^\dagger a}, \quad (\text{B55})$$

$$A_+ = -i \int_0^T \mathcal{E}(t) e^{-i\Phi t} dt, \quad (\text{B56})$$

and

$$B_+ = - \int_0^T dt \int_t^T dt' \mathcal{E}(t) \mathcal{E}(t') \sin(\bar{\Phi}(t' - t)). \quad (\text{B57})$$

Hence to conclude, the unitary evolution of the Hamiltonian described in equation (B1) is given by

$$V(0, T) = R_+(T) D(A_+) e^{iB_+} e^{-i\tilde{\omega}_a T} |0\rangle \langle 0| + R_-(T) D(A_-) e^{iB_-} e^{i\tilde{\omega}_a T} |1\rangle \langle 1|. \quad (\text{B58})$$

Recall that in deriving the expression for  $V_+(0, T)$ , in several steps of the calculation we expanded to leading order in  $\phi$ . Hence the expressions obtained in equations (B55)–(B57) are only valid to leading order in  $\phi$ . Hence to leading order in  $\phi$ , we have

$$R_\pm(T) = e^{-iT\omega_\pm a^\dagger a} (1 \pm iT\phi(a^\dagger a)^2) + \mathcal{O}(\phi^2), \quad (\text{B59})$$

$$A_\pm = -i \int_0^T \mathcal{E}(t) e^{i\omega_\pm t} dt \mp \phi(2a^\dagger a - 1) \int_0^T \mathcal{E}(t) t e^{i\omega_\pm t} dt + \mathcal{O}(\phi^2), \quad (\text{B60})$$

and

$$\begin{aligned} B_\pm &= \int_0^T dt \int_t^T dt' \mathcal{E}(t) \mathcal{E}(t') \sin(\omega_\pm(t' - t)) \pm \phi(2a^\dagger a + 1) \int_0^T dt \\ &\quad \times \int_t^T dt' \mathcal{E}(t) \mathcal{E}(t') \cos(\omega_\pm(t' - t))(t' - t) + \mathcal{O}(\phi^2). \end{aligned} \quad (\text{B61})$$

## B.2. Including the Kerr nonlinearity

When including the Kerr nonlinearity, the Hamiltonian during the control-displacement gate is now given by

$$H(t) = \tilde{\omega}_r a^\dagger a + \tilde{\omega}_a Z + \chi a^\dagger a Z - \phi(a^\dagger a)^2 Z - \frac{K}{2} (a^\dagger a)^2. \quad (\text{B62})$$

In this case, we have that

$$\begin{aligned} V(0, T) &= \left( \lim_{n \rightarrow \infty} \prod_{j=1}^n e^{-i\tilde{t}(\omega_+ - (\phi + \frac{K}{2}) a^\dagger a) a^\dagger a} e^{-i\tilde{t} \mathcal{E}(t_j) (a + a^\dagger)} \right) e^{-i\tilde{\omega}_a T} |0\rangle \langle 0| \\ &\quad + \left( \lim_{n \rightarrow \infty} \prod_{j=1}^n e^{-i\tilde{t}(\omega_- + (\phi - \frac{K}{2}) a^\dagger a) a^\dagger a} e^{-i\tilde{t} \mathcal{E}(t_j) (a + a^\dagger)} \right) e^{i\tilde{\omega}_a T} |1\rangle \langle 1|, \end{aligned} \quad (\text{B63})$$

Hence, we see that the analysis of section B.1 is exactly the same, with  $\phi \rightarrow \phi \pm \frac{K}{2}$  in  $R_\pm(T)$ ,  $A_\pm$  and  $B_\pm$ .

## B.3. Controlled-displacement gate in the rotating frame

In this section, we consider the same Hamiltonian as in equation (B1) but with a drive term of the form

$$H_d(t) = \mathcal{E}(t) a^\dagger e^{-i\omega_d t} + \mathcal{E}^*(t) a e^{i\omega_d t}. \quad (\text{B64})$$

We will go into the rotating frame of both the qubit and the cavity. We define

$$H'_s = \tilde{\omega}_r a^\dagger a + \tilde{\omega}_a Z. \quad (\text{B65})$$

With  $U(t) = e^{iH_s' t}$  and applying the transformation  $U(t)H(t)U^{-1}(t) - iU(t)\frac{\partial}{\partial t}U^{-1}(t) \equiv H_R(t)$  to the Hamiltonian in equation (B1) with  $H_d(t)$  given in equation (B64), we obtain

$$H_R(t) = \chi Za^\dagger a - \phi(a^\dagger a)^2 Z + \mathcal{E}(t)a^\dagger e^{i(\tilde{\omega}_r - \omega_d)t} + \mathcal{E}^*(t)ae^{-i(\tilde{\omega}_r - \omega_d)t} \quad (B66)$$

The frequency dependence in the drive term of equation (B66) can be eliminated by choosing  $\omega_d = \tilde{\omega}_r$ . Again, we wish to compute the unitary evolution

$$V_R(0, T) = \mathcal{T}e^{-i\int_0^T H_R(t')dt'}. \quad (B67)$$

Using the Suzuki–Trotter decomposition

$$\begin{aligned} V_R(0, T) &= \left( \lim_{n \rightarrow \infty} \prod_{j=1}^n e^{-i\tilde{\tau}(\omega_+ - \phi a^\dagger a)a^\dagger a} e^{-i\tilde{\tau}(\mathcal{E}(t_j)a^\dagger + \mathcal{E}^*(t_j)a)} \right) |0\rangle \langle 0| \\ &\quad + \left( \lim_{n \rightarrow \infty} \prod_{j=1}^n e^{-i\tilde{\tau}(\omega_- + \phi a^\dagger a)a^\dagger a} e^{-i\tilde{\tau}(\mathcal{E}(t_j)a^\dagger + \mathcal{E}^*(t_j)a)} \right) |1\rangle \langle 1|, \end{aligned} \quad (B68)$$

where now we have that  $\omega_{\pm} = \pm\chi$ .

Comparing equation (B68) to equation (B6), we see that we can follow the same steps as in appendix B.1 by using the new values for  $\omega_{\pm}$  and replacing  $\mathcal{E}$  by  $\mathcal{E}^*$  in the conjugate expressions. Doing so, we obtain

$$V_R(0, T) = R_+(T)D(A_+)e^{iB_+}|0\rangle \langle 0| + R_-(T)D(A_-)e^{iB_-}|1\rangle \langle 1|, \quad (B69)$$

where

$$R_{\pm}(T) = e^{-iT\omega_{\pm}a^\dagger a}(1 \pm iT\phi(a^\dagger a)^2), \quad (B70)$$

$$A_{\pm} = -i \int_0^T \mathcal{E}(t)e^{i\omega_{\pm}t}dt \mp \phi(2a^\dagger a - 1) \int_0^T \mathcal{E}(t)te^{i\omega_{\pm}t}dt + \mathcal{O}(\phi^2), \quad (B71)$$

$$\begin{aligned} B_{\pm} &= -\frac{i}{2} \int_0^T dt \int_t^T dt' (\mathcal{E}^*(t')\mathcal{E}(t)e^{i\omega_{\pm}(t'-t)} - \mathcal{E}(t')\mathcal{E}^*(t)e^{-i\omega_{\pm}(t'-t)}) \\ &\pm \frac{\phi(2a^\dagger a + 1)}{2} \int_0^T dt \int_t^T dt' (\mathcal{E}^*(t')\mathcal{E}(t)e^{i\omega_{\pm}(t'-t)} + \mathcal{E}(t')\mathcal{E}^*(t)e^{i\omega_{\pm}(t'-t)})(t' - t) + \mathcal{O}(\phi^2). \end{aligned} \quad (B72)$$

#### B.4. Effects of the nonlinear dispersive shift and Kerr term on the unitary evolution of the qubit-cavity system

In this section we will show that regardless of the chosen pulse shape (even with numerical tools such as optimal control), to leading order in  $\phi_{\pm}$ , the changes to the unitary evolution of the qubit-cavity system due to the nonlinear dispersive shift and Kerr terms cannot be completely removed for time scales  $T \ll \frac{1}{\phi_{\pm}}$ . Using equations (B55), (B60) and (B61), we can write (for instance choosing the terms affecting the  $|0\rangle \langle 0|$ )

$$R_+(T)D(A_+) = e^{-i\omega_+ T a^\dagger a} e^{i\omega_+ T \phi_+ (a^\dagger a)^2} e^{I_1 a^\dagger - I_1^* a - I_2 \phi_+ (2a^\dagger a - 1)a^\dagger + I_2^* \phi_+ a (2a^\dagger a - 1)}, \quad (B73)$$

where we define

$$I_1 \equiv -i \int_0^T \mathcal{E}(t)e^{i\omega_+ t}dt, \quad (B74)$$

and

$$I_2 \equiv \int_0^T \mathcal{E}(t)te^{i\omega_+ t}dt. \quad (B75)$$

Let  $A = i\omega_+ T \phi_+ (a^\dagger a)^2$  and  $B = I_1 a^\dagger - I_1^* a - I_2 \phi_+ (2a^\dagger a - 1)a^\dagger + I_2^* \phi_+ a (2a^\dagger a - 1)$ . Using the Baker–Campbell–Hausdorff lemma and keeping only leading order terms in  $\phi_+$ , we have

$$e^A e^B = e^{A+B+\frac{1}{2}[A,B]-\frac{1}{12}[B[A,B]]-\frac{1}{220}[B,[B,[B,A]]]}. \quad (B76)$$

Computing the commutators, we find

$$\begin{aligned} e^A e^B &= e^{\frac{1}{6}i\omega_+ T \phi_+ \left(-\frac{1}{5}|I_1|^4 + |I_1|^2\right)} e^{\left(I_2 \phi_+ + \left(1 + \frac{1}{2}i\omega_+ T \phi_+\right) I_1\right) a^\dagger - \left(I_2^* \phi_+ + \left(1 - \frac{1}{2}i\omega_+ T \phi_+\right) I_1^*\right) a + \frac{1}{6}i\omega_+ T \phi_+ (I_1^2 (a^\dagger)^2 + 4|I_1|^2 a^\dagger a + (I_1^*)^2 a^2)} \\ &\quad + i\omega_+ T \phi_+ (I_1 a^\dagger a^\dagger a + I_1^* a^\dagger a a) - 2I_2 \phi_+ a^\dagger a a^\dagger + 2I_2^* \phi_+ a a^\dagger a + \omega_+ T \phi_+ (a^\dagger a)^2 \end{aligned} \quad (B77)$$

Notice that the last term in equation (B77) is proportional to  $(a^\dagger a)^2$  and independent of the pulse shape (performing the same calculation as above including the  $B_+$  term will not change the conclusion). The terms dependent on the pulse shape are expressed as lower powers of  $a$  and  $a^\dagger$ . One cannot eliminate all terms

proportional to  $\phi_+$  unless one chooses a time scale comparable to  $\frac{1}{\phi_+}$ . For time scales on the order of  $\frac{1}{\phi_+}$ , higher order terms in  $\phi_+$  will be relevant and therefore it might possible to choose pulse (with numerical techniques such as optimal control) which can eliminate effects from the nonlinear dispersive shift and Kerr terms. However using the parameters of table 2,  $\frac{1}{\phi_+}$  is roughly four orders of magnitude longer than the chosen time scale ( $T = \frac{\pi}{\chi}$ ) of our protocol. For such long time scales, effects due to photon loss, damping and dephasing would render the protocol impractical.

### Appendix C. Probability of measuring accepted measurement strings

In this section, we derive the analytic expression for output states obtained from phase estimation protocols and their corresponding probabilities.

#### C.1. Phase estimation measurement operator of 1 round

From figure 2, the operator describing the evolution of a single round of phase estimation in terms of the measurement operator set  $\left\{ M_x = \frac{D(-\sqrt{2\pi}) + (-1)^x e^{i\gamma} D(\sqrt{2\pi})}{2} |x = 0, 1\right\}$  can be written as

$$U_{\text{measure}}^\gamma = \{|0\rangle\langle 0| \otimes M_0 + |1\rangle\langle 1| \otimes M_1\} \quad (C1)$$

After the measurement, the state become

$$\begin{aligned} |\psi_{\text{output}}\rangle &= \frac{U_{\text{measure}}^\gamma |\psi_{\text{input}}\rangle}{\sqrt{Pr_x}} \\ &= \frac{1}{2\sqrt{Pr_x}} (D(-\sqrt{2\pi}) |\psi_{\text{input}}\rangle + (-1)^x e^{i\gamma} D(\sqrt{2\pi}) |\psi_{\text{input}}\rangle) \\ &= \frac{1}{2\sqrt{Pr_x}} (D(-\sqrt{2\pi}) |\psi_{\text{input}}\rangle + e^{i(\gamma+x\pi)} D(\sqrt{2\pi}) |\psi_{\text{input}}\rangle) \end{aligned} \quad (C2)$$

where  $Pr_x$  is the probability of measuring  $x \in \{0, 1\}$ .

#### C.2. Phase estimation measurement operator after $M$ rounds

After  $M$  rounds of the phase estimation circuit, the following state will be generated:

$$|\Phi(x[M])\rangle = \frac{1}{2^M \sqrt{N}} \prod_{m=1}^M \left[ D\left(-\sqrt{\frac{\pi}{2}}\right) + e^{i(\gamma_m+x_m\pi)} D\left(\sqrt{\frac{\pi}{2}}\right) \right] |\text{sq.vac}\rangle, \quad (C3)$$

where  $N = \prod_i Pr_{x_i}$  is the normalization factor. From the product of  $M$  sums, if we choose  $j$  of them to be positive shifts  $\left(D\left(\sqrt{\frac{\pi}{2}}\right)\right)$ ,  $M - j$  to be negative shifts  $\left(D\left(-\sqrt{\frac{\pi}{2}}\right)\right)$ , we can produce a peak at  $\sqrt{\pi/2}(j - (M - j)) = \sqrt{2\pi}(j - M/2)$  in  $q$  space, and the produced phase is  $\prod_{l \in S_j} e^{i(\gamma_l+x_l\pi)}$  ( $S_j$  is the set of rounds we choose to be positive shifts). There are  $\binom{M}{j}$  ways to choose such an  $S_j$ , by choosing any of such  $S_j$  we can produce a peak at  $\sqrt{2\pi}(j - M/2)$ . Thus,  $|\Phi(x[M])\rangle$  can be written as:

$$|\Phi(x[M])\rangle = \frac{1}{2^M \sqrt{N}} \sum_{j=0}^M c_j(x[M]) D(\sqrt{2\pi}(j - M/2)) |\text{sq.vac}\rangle, \quad (C4)$$

where  $N \approx \sum_{j=0}^M |c_j(x[M])|^2$  with  $c_j(x[M]) = \sum_{\{S_j\}} \prod_{l \in S_j} e^{i(\gamma_l+x_l\pi)}$ .

#### C.3. Probability

Assume we already performed  $M - 1$  rounds of phase estimation and the measurement results  $x_1, \dots, x_{M-1}$  and angle parameters we chose (denoted as  $\gamma_1, \dots, \gamma_{M-1}$ ) are known. Then from (C4), the input state to the  $M$ th round is (including the qubit):

$$|\Phi_{\text{input}}\rangle_M = \frac{1}{2^{M-1} \sqrt{Pr_{x_1} \dots Pr_{x_{M-1}}}} \sum_{j=0}^{M-1} c_j(x[M-1]) D(\sqrt{2\pi}(j - (M-1)/2)) |0\rangle \otimes |\text{sq.vac}\rangle. \quad (C5)$$

The probability of measuring  $x_M$  is thus given by

$$\begin{aligned} Pr_{x_M} &= \langle \Phi_{\text{input}} | (U_{\text{measure}}^{\gamma})^{\dagger} U_{\text{measure}}^{\gamma} | \Phi_{\text{input}} \rangle_M \\ &= \left| \frac{1}{2^M \sqrt{Pr_{x_1} \dots Pr_{x_{M-1}}}} \sum_{j=0}^M c_j(x[M]) D(\sqrt{2\pi}(j - (M)/2)) | \text{sq.vac} \rangle \right|^2. \end{aligned} \quad (\text{C6})$$

Notice that the transfer probability from one peak to another is almost 0 ( $< 10^{-35}$ ) assuming an initial squeezing level of 0.2. Therefore, we can make the following approximation

$$\approx \frac{1}{2^{2M} Pr_{x_1} \dots Pr_{x_{M-1}}} \sum_{j=0}^M |c_j(x[M]) D(\sqrt{2\pi}(j - (M)/2)) | \text{sq.vac} \rangle|^2 \quad (\text{C7})$$

$$= \frac{1}{2^{2M} Pr_{x_1} \dots Pr_{x_{M-1}}} \sum_{j=0}^M |c_j(x[M])|^2 \underbrace{\langle \text{sq.vac} | D^{\dagger}(\sqrt{2\pi}(j - (M)/2)) | D(\sqrt{2\pi}(j - (M)/2)) | \text{sq.vac} \rangle}_1 \quad (\text{C8})$$

$$= \frac{1}{2^{2M} Pr_{x_1} \dots Pr_{x_{M-1}}} \sum_{j=0}^M |\sum \prod_{l \in S_j} e^{i(\gamma_l + x_l \pi)}|^2 \quad (\text{C9})$$

$$= \frac{1}{2^{2M} Pr_{x_1} \dots Pr_{x_{M-1}}} \sum_{j=0}^M \left[ \left( \sum_{\{S_j^m\}} \prod_{l \in S_j^m} e^{-i(\gamma_l + x_l \pi)} \right) \left( \sum_{\{S_j^n\}} \prod_{k \in S_j^n} e^{i(\gamma_k + x_k \pi)} \right) \right]. \quad (\text{C10})$$

The above expression is recursive, hence substituting  $Pr_{x_{M-1}}$  into it, we find

$$Pr_{x_M} \approx \frac{\sum_{j=0}^M \left[ \left( \sum_{\{S_j^m\}} \prod_{l \in S_j^m} e^{-i(\gamma_l + x_l \pi)} \right) \left( \sum_{\{S_j^n\}} \prod_{k \in S_j^n} e^{i(\gamma_k + x_k \pi)} \right) \right]}{4 \sum_{j=0}^{M-1} \left[ \left( \sum_{\{S_j^m\}} \prod_{l \in S_j^m} e^{-i(\gamma_l + x_l \pi)} \right) \left( \sum_{\{S_j^n\}} \prod_{k \in S_j^n} e^{i(\gamma_k + x_k \pi)} \right) \right]}. \quad (\text{C11})$$

## Appendix D. Details of numerical simulation

From the depolarizing noise model described in section 5.2, we computed error probabilities for all Pauli errors arising from a single fault at the locations indicated by the blue boxes in figure D1. For instance, the probability of an  $I \otimes Z$  error at location  $L_1$  is given by

$$P_{IZ} = \frac{p}{15} (P1_X P1_Z + P1_X (1 - P1_Z) + (1 - P1_X) (1 - P1_Z)) + (1 - p) (1 - P1_X) P1_Z, \quad (\text{D1})$$

where

$$P1_X = \frac{p}{15}, \quad (\text{D2})$$

during the first round, and

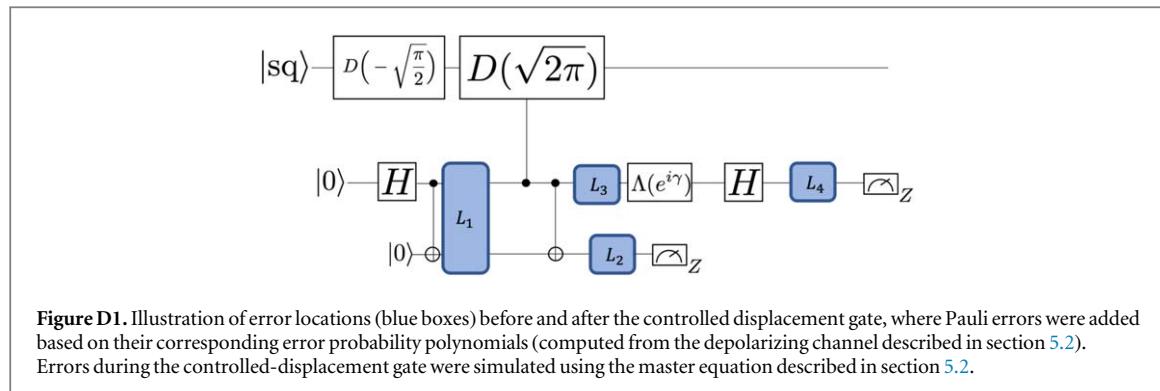
$$P1_X = \frac{2p}{3}, \quad (\text{D3})$$

in later rounds. Similarly

$$P1_Z = \frac{p}{15} \left( 1 - \frac{p}{10} \right) + \frac{p^2}{450} + \left( 1 - \frac{p}{15} \right) \frac{p}{15}, \quad (\text{D4})$$

for the first round and

$$P1_Z = \frac{2p}{3} \left( 1 - \frac{p}{10} \right) + \frac{2p}{3} \frac{p}{30} + \left( 1 - \frac{2p}{3} \right) \frac{p}{15}. \quad (\text{D5})$$



With a Pauli error added at either locations  $L_1, L_2, L_3$  or  $L_4$ , we update the state of the qubit before performing the master equation simulation described in section 5.2. With the analytic error probability for the Pauli error, in addition to the evolution of the state during the master equation, we can compute the total probability of obtaining the output state. As mentioned in section 5.2, such an analysis ignored second order Pauli error events (before and after the controlled displacement gate).

## ORCID iDs

Christopher Chamberland  <https://orcid.org/0000-0003-3239-5783>

## References

- [1] Albert V V *et al* 2018 Performance and structure of single-mode bosonic codes *Phys. Rev. A* **97** 032346
- [2] Noh K, Albert V V and Jiang L 2019 Quantum capacity bounds of gaussian thermal loss channels and achievable rates with Gottesman–Kitaev–Preskill codes *IEEE Trans. Inf. Theory* **65** 2563–82
- [3] 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
- [4] 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
- [5] Fukui K, Tomita A, Okamoto A and Fujii K 2018 High-threshold fault-tolerant quantum computation with analog quantum error correction *Phys. Rev. X* **8** 021054
- [6] Noh K and Chamberland C 2019 Fault-tolerant bosonic quantum error correction with the surface-GKP code arXiv:1908.03579
- [7] 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 arXiv:1903.00012
- [8] Travaglione B C and Milburn G J 2002 Preparing encoded states in an oscillator *Phys. Rev. A* **66** 052322
- [9] Glancy S and Knill E 2006 Error analysis for encoding a qubit in an oscillator *Phys. Rev. A* **73** 012325
- [10] Alonso J, Leupold F M, Soler Z U, Fadel M, Marinelli M, Keitch B C, Negnevitsky V and Home J P 2016 Generation of large coherent states by bang–bang control of a trapped-ion oscillator *Nat. Commun.* **7** 11243
- [11] Fluhmann 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
- [12] Su W-C, Lin C-Y and Wu S-T 2018 Encoding qubits into harmonic-oscillator modes via quantum walks in phase space arXiv:1808.08722
- [13] Eaton M, Nehra R and Pfister O 2019 Gottesman–Kitaev–Preskill state preparation by photon catalysis arXiv:quant-ph/1903.01925
- [14] Douce T, Markham D, Kashefi E, van Loock P and Ferrini G 2019 Probabilistic fault-tolerant universal quantum computation and sampling problems in continuous variables *Phys. Rev. A* **99** 012344
- [15] Puri S, Grimm A, Campagne-Ibarcq P, Eickbusch A, Noh K, Roberts G, Jiang L, Mirrahimi M, Devoret M H and Girvin S M 2018 Stabilized cat in driven nonlinear cavity: a fault-tolerant error syndrome detector arXiv:1807.09334
- [16] Aliferis P, Gottesman D and Preskill J 2006 Quantum accuracy threshold for concatenated distance-3 codes *Quantum Inf. Comput.* **6** 97–165
- [17] Gottesman D, Kitaev A and Preskill J 2001 Encoding a qubit in an oscillator *Phys. Rev. A* **64** 012310
- [18] Cross A W, Divincenzo D P and Terhal B M 2009 A comparative code study for quantum fault tolerance *Quantum Inf. Comput.* **9** 541–72
- [19] Aliferis P and Cross A W 2007 Subsystem fault tolerance with the bacon–shor code *Phys. Rev. Lett.* **98** 220502
- [20] Chamberland C, Jochym-O’Connor T and Laflamme R 2016 Thresholds for universal concatenated quantum codes *Phys. Rev. Lett.* **117** 010501
- [21] Chamberland C, Jochym-O’Connor T and Laflamme R 2017 Overhead analysis of universal concatenated quantum codes *Phys. Rev. A* **95** 022313
- [22] Chamberland C and Ronagh P 2018 Deep neural decoders for near term fault-tolerant experiments *Quantum Sci. Technol.* **3** 044002
- [23] Duivenvoorden K, Terhal B M and Weigand D 2017 Single-mode displacement sensor *Phys. Rev. A* **95** 012305
- [24] 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
- [25] Chao R and Reichardt B W 2018 Quantum error correction with only two extra qubits *Phys. Rev. Lett.* **121** 050502
- [26] Chao R and Reichardt B W 2018 Fault-tolerant quantum computation with few qubits *NPJ Quantum Inf.* **4** 42
- [27] Chamberland C and Beverland M E 2018 Flag fault-tolerant error correction with arbitrary distance codes *Quantum* **2** 53

- [28] Tansuwanont T, Chamberland C and Leung D 2018 Flag fault-tolerant error correction for cyclic css codes arXiv:[1803.09758](https://arxiv.org/abs/1803.09758)
- [29] Reichardt B W 2018 Fault-tolerant quantum error correction for steaneas seven-qubit color code with few or no extra qubits arXiv:[quant-ph/1804.06995](https://arxiv.org/abs/1804.06995)
- [30] Chamberland C and Cross A W 2019 Fault-tolerant magic state preparation with flag qubits *Quantum* **3** 143
- [31] Gambetta J, Blais A, Boissonneault M, Houck A A, Schuster D I and Girvin S M 2008 Quantum trajectory approach to circuit QED: quantum jumps and the Zeno effect *Phys. Rev. A* **77** 012112
- [32] Boissonneault M, Gambetta J M and Blais A 2009 Dispersive regime of circuit QED: photon-dependent qubit dephasing and relaxation rates *Phys. Rev. A* **79** 013819
- [33] Ofek N *et al* 2016 Demonstrating quantum error correction that extends the lifetime of quantum information *Nature* **536** 441–5
- [34] Johansson J, Nation P and Nori F 2013 Qutip 2: a Python framework for the dynamics of open quantum systems *Comput. Phys. Commun.* **184** 1234–1240
- [35] Chou K S, Blumoff J Z, Wang C S, Reinhold P C, Axline C J, Gao Y Y, Frunzio L, Devoret M H, Jiang L and Schoelkopf R J 2018 Deterministic teleportation of a quantum gate between two logical qubits *Nature* **561** 1476–4687
- [36] 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** 1476–4687
- [37] Reagor M *et al* 2013 Reaching 10 ms single photon lifetimes for superconducting aluminum cavities *Appl. Phys. Lett.* **102** 192604
- [38] Minev Z K, Serniak K, Pop I M, Leghtas Z, Sliwa K, Hatridge M, Frunzio L, Schoelkopf R J and Devoret M H 2016 Planar multilayer circuit quantum electrodynamics *Phys. Rev. Appl.* **5** 044021
- [39] 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
- [40] Khaneja N, Reiss T, Kehlet C, Schulte-Herbruggen T and Glaser S J 2005 Optimal control of coupled spin dynamics: design of NMR pulse sequences by gradient ascent algorithms *J. Magn. Reson.* **172** 296–305