Encoding a qubit into a cavity mode in circuit QED using phase estimation

B. M. Terhal and D. Weigand

JARA Institute for Quantum Information, RWTH Aachen University, 52056 Aachen, Germany

Gottesman, Kitaev, and Preskill have formulated a way of encoding a qubit into an oscillator such that the qubit is protected against small shifts (translations) in phase space. The idea underlying this encoding is that error processes of low rate can be expanded into small shift errors. The qubit space is defined as an eigenspace of two mutually commuting displacement operators S_p and S_q which act as large shifts or translations in phase space. We propose and analyze the approximate creation of these qubit states by coupling the oscillator to a sequence of ancilla qubits. This preparation of the states uses the idea of phase estimation where the phase of the displacement operator, say S_p, is approximately determined. We consider several possible forms of phase estimation. We analyze the performance of repeated and adaptive phase estimation as the simplest and experimentally most viable schemes given a realistic upper limit on the number of photons in the oscillator. We propose a detailed physical implementation of this protocol using the dispersive coupling between a transmon ancilla qubit and a cavity mode in circuit QED. We provide an estimate that in a current experimental setup one can prepare a good code state from a squeezed vacuum state using eight rounds of adaptive phase estimation, lasting in total about 4 μs, with 94% (heralded) chance of success.

I. INTRODUCTION

It is to be expected that quantum error correction will be essential in the implementation of reliable, large-scale, quantum computation. Efforts are underway for superconducting transmon qubits [1] coupled to microwave resonators (circuit-QED) to realize components of a surface code architecture in which robust logical qubits are comprised of many $O(10$–$100)$ elementary transmon qubits and resonators [2,3]. As the qubit and resonator overhead of such encoding is rather daunting and inefficient, one may ask whether there exist more efficient ways of using high-coherence qubits and long lifetime microwave resonators (see some numbers in Table I) to encode robust logical qubits. In this paper, we will explore the proposal by Gottesman, Kitaev, and Preskill [4] to encode a qubit into an oscillator and take a first stab at analyzing how a qubit can be encoded and preserved in a single mode of a microwave cavity. This scheme may not only be interesting as an alternative route towards scalable quantum computation, but also as a way of generating a highly nonclassical state in a (microwave) cavity which one can actively protect against photon loss.

The GKP (Gottesman-Kitaev-Preskill) code is of interest as it offers the possibility to use the cavity mode to store a single qubit, while a transmon qubit coupled to this cavity mode can be used to manipulate the state of the cavity. The goal is to prepare and preserve a qubit in a cavity mode such that it has a coherence time much longer than that of the cavity or the transmon qubit itself. The GKP scheme is attractive in that gates on this qubit are relatively straightforward as Clifford gates can be realized using linear optical components (see [4] and Appendix A). In addition, this qubit-into-oscillator scheme can be concatenated with a surface code or another stabilizer code for scalable protection (see, e.g., [5] for a surface code with oscillators).

However, a realistic scheme for preparing the encoded states and performing quantum error correction has been missing so far (in Sec. I B we list some of the previous work) and this is the focus of this paper. In the remainder of this section, we

review and discuss the code states and their highly nonclassical properties. We argue why, at a heuristic level, using phase estimation can provide a means for approximately generating such quantum states. In Sec. I C, we review and introduce some tools and formalism which are useful in assessing the quality of a preparation protocol, using a model of displacement or shift errors. In Sec. I D, we physically motivate the shift error model by discussing how physical errors can be expanded into such shifts. The shift error model plays a similar role as a Pauli error model for qubits. This section partially reviews some of the insights of [4]; it adds to these by showing in Sec. I D how an expansion into shift errors can lead to a reduction in error rate by several orders of magnitude. Section II is devoted to exploring and understanding how to prepare an approximate code state using the protocol of phase estimation, in this case determining the eigenvalue of a unitary displacement operator. There are many ways to do phase estimation leading to different approximations for the code states and using different numbers of photons. These phase estimation protocols all consist of a coupling sequence of single-qubit ancillas to an oscillator mode sequentially and measuring the ancilla qubit. We argue why we focus on two simplest protocols: one is repeated phase estimation without feedback (Sec. II B) and a second one is a simple form of adaptive phase estimation with feedback (Sec. II C). For those two schemes we explicitly show in Sec. II E how well with a small number of rounds (ancilla qubits) an approximate code state can be prepared. We also consider the average number of photons used and the variance thereof, as these numbers are relevant in any experimental realization of the protocol.

In Sec. III, we discuss a realization of the code state preparation protocol in an experimental circuit-QED setup where the oscillator corresponds to a cavity mode. We review the circuit-QED transmon-cavity coupling and the physical strength of various parameters as reported in recent experiments in circuit-QED. We discuss several aspects of an implementation such as the use of a tunable coupling and the implementation of a (transmon) qubit controlled-displacement

gate by means of a controlled rotation or a direct pulse-driven controlled displacement (details in Appendix D). In Sec. III A, we discuss at a qualitative level the dominant sources of inaccuracy in implementing the protocol; we especially note that the presence of cavity nonlinearities may require a deeper analysis. We end the paper with a discussion and outlook for future work.

In Appendixes A and B, we collect some background technical results which are relevant for implementing gates on the code states and performing quantum error correction. Appendix C gives background on the choice of feedback phases in the phase estimation protocol in Sec. II C.

A. GKP code states

In order to introduce the code states, we assume a harmonic oscillator with annihilation operator (creation operator) a ($a^\dagger$) with which one can define (dimensionless) quadrature operators $p = \frac{i}{\sqrt{2}}(a^\dagger - a)$ (momentum) and $q = \frac{1}{\sqrt{2}}(a + a^\dagger)$ (position) obeying the canonical commutation relations $[q, p] = i$. The ideal GKP code is simply a prescription of a two-dimensional subspace in the infinite-dimensional oscillator space. This subspace, called the code space, is defined as the $+1$ eigenspace of two commuting displacement operators $S_p = e^{-i2\sqrt{\pi}p}$ and $S_q = e^{i2\sqrt{\pi}q}$ (one can verify the commutation of these operators by using the identity $e^A e^B = e^B e^A e^{[A,B]}$ for A, B linear combinations of p and q). The space fixed by S_p and S_q is two dimensional as there are additional operators which we can identify as X and Z which commute with both S_p and S_q, but which mutually anticommute.

States in this two-dimensional subspace will be called code states as they encode a qubit. Of course, there are many possible ways of choosing a two-dimensional subspace of a harmonic oscillator to define a qubit, e.g., choose two energy eigenlevels or pick two orthogonal cat states to define $|0\rangle$ and $|1\rangle$. The subspace of the GKP code is one in which $|0\rangle$ and $|1\rangle$ are highly nonclassical states and is chosen such that small phase space displacements can be undone by quantum error correction. In this sense, the code offers the possibility to realize a long-lived, well-protected qubit.

What is immediately interesting about a state in this subspace is that the condition $S_p = 1$ and $S_q = 1$ fixes *both* momentum $p = 0 \mod \sqrt{\pi}$ and position $q = 0 \mod \sqrt{\pi}$ to be sharply determined. A code state thus escapes Heisenberg's uncertainty principle by being not localized at a single p and q but being a superposition of many equally spaced sharp values for p and q. Due to having low variance in *both* quadratures, we can expect that a code state is useful in metrology for detecting small displacements which shift p or q by less than $\sqrt{\pi}$.

If we write the operators S_p and S_q as displacements, one has $S_p = D(\sqrt{2\pi})$ and $S_q = D(i\sqrt{2\pi})$ [using $D(\alpha) = \exp(\alpha a^\dagger - \alpha^* a)$ with $D(\alpha)|\text{vac}\rangle = |\alpha\rangle$ and $|\alpha\rangle$ a coherent state]. In this code space, one can define the qubit states $|0\rangle$ and $|1\rangle$ as

$$|0\rangle \propto \sum_{t=-\infty}^{\infty} D(t\sqrt{2\pi})|q=0\rangle = \sum_{t=-\infty}^{\infty} S_p^t |q=0\rangle,$$

$$|1\rangle \propto \sum_{t=-\infty}^{\infty} D(t\sqrt{2\pi})|q=\sqrt{\pi}\rangle = \sum_{t=-\infty}^{\infty} S_p^t |q=\sqrt{\pi}\rangle. \quad (1)$$

Thus, $|0\rangle$ is a uniform superposition of eigenstates of position q at *even* integer multiples of $\sqrt{\pi}$, while $|1\rangle$ is a uniform superposition of eigenstates of position q at *odd* integer multiples of $\sqrt{\pi}$. One can also consider these states as superpositions of p eigenstates, i.e., $|0\rangle \propto \int dp \, (\sum_{t=-\infty}^{\infty} S_p^t)|p\rangle \propto \sum_{s=-\infty}^{\infty} |p = s\sqrt{\pi}\rangle$ while $|1\rangle \propto \int dp \, e^{-i\sqrt{\pi}p}(\sum_{t=-\infty}^{\infty} S_p^t)|p\rangle \propto \sum_{s=-\infty}^{\infty} (-1)^s |p = s\sqrt{\pi}\rangle$ [using that $\sum_{t=-\infty}^{\infty} e^{-i2\sqrt{\pi}pt} \propto \sum_{s=-\infty}^{\infty} \delta(p = \sqrt{\pi}s)$]. Thus, both $|0\rangle$ and $|1\rangle$ have nonzero amplitude at integer multiples of $\sqrt{\pi}$ in p space, but due to alternating phases, these amplitudes destructively interfere at odd multiples of $\sqrt{\pi}$ in the state $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ [6]. The operator $X : |0\rangle \leftrightarrow |1\rangle$ is given by $X = e^{-i\sqrt{\pi}p}$, as it shifts each eigenstate $|q\rangle$ by $\sqrt{\pi}$. Note that $X^2 = S_p$ which equals I only on the code space ($+1$ eigenspace of S_p) (see discussion in Appendix A). One can write

$$|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \propto \sum_{t=-\infty}^{\infty} S_q^t |p=0\rangle,$$

$$|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \propto \sum_{t=-\infty}^{\infty} S_q^t |p=\sqrt{\pi}\rangle, \quad (2)$$

with $Z : |+\rangle \leftrightarrow |-\rangle$ given by $Z = e^{i\sqrt{\pi}q}$ and $Z^2 = S_q$.

A simple way to understand the ideal preparation of a code state $|0\rangle$ (or $|1\rangle$) is through Eq. (1): one starts with a $+1$ eigenstate of Z, namely, $|q = 0\rangle$ to which one applies $\Pi_{S_p=1} = \sum_{t=-\infty}^{\infty} S_p^t$ which is the projector onto the $+1$ eigenspace of S_p [7].

Approximate GKP code states

Naturally, Ref. [4] realized that the perfect code states in Eq. (1) are unphysical as their preparation would require infinite squeezing and an infinitely sharp projection onto the $+1$ eigenspace of S_p. Reference [4] suggested using approximate code states: we can understand how this approximation comes about in, say, the definition of $|0\rangle$, as follows.

As it would take an infinite amount of squeezing to prepare $|q = 0\rangle$ we replace this state by a finitely squeezed state (in q) centered around $q = 0$, that is, a squeezed vacuum state [Eq. (3)]. Then, we need to implement an approximate projection onto the $S_p = 1$ eigenspace. Such projection could come about by approximately estimating the eigenvalue of the operator S_p and then *post-selecting* the measurement outcome on this eigenvalue being $+1$ (while discarding other results). Such approximate projection by post-selection, using an ancilla qubit, has been considered in [8], without making reference to any particular technology. However, any post-selection scheme will have a low probability of success and is in fact unnecessary. If one can implement a highly accurate phase estimation measurement of the unitary displacement operator S_p, estimating its eigenvalue as some $e^{i\theta}$, then one can also correct for such an eigenvalue shift and shift, or displace, back to the $+1$ eigenvalue code space.

The formal definition of the approximate states is as follows. One replaces each delta function in position q by a squeezed Gaussian state while the uniform superposition over these localized states is replaced by a Gaussian envelope. For example, for the approximate $|0\rangle$ state, one starts with
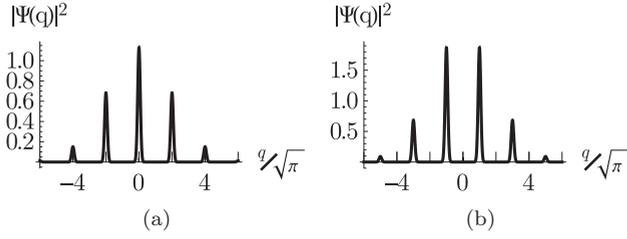
FIG. 1. Absolute value of the wave function of the approximate code state $|0\rangle_{approx}$ (a) and $|1\rangle_{approx}$ (b) for $\Delta = 0.2$.

the squeezed vacuum state $|vac_{sq}\rangle$ with squeezing parameter $\Delta^2 = e^{-2r}$:

$$|vac_{sq}\rangle = \int \frac{dq}{(\pi \Delta^2)^{1/4}} e^{-q^2/(2\Delta^2)} |q\rangle, \tag{3}$$

to which one applies a sum of displacements, as in Eq. (1), with a Gaussian filter, i.e.,

$$|0\rangle_{approx} \propto \sum_{t=-\infty}^{\infty} e^{-2\pi \tilde{\Delta}^2 t^2} D(t\sqrt{2\pi}) |vac_{sq}\rangle$$

$$= \sum_{t=-\infty}^{\infty} \int e^{-2\pi \tilde{\Delta}^2 t^2} e^{-(q-2t\sqrt{\pi})^2/(2\Delta^2)} |q\rangle \, dq,$$

$$|1\rangle_{approx} \propto \sum_{t=-\infty}^{\infty} \int e^{-\pi \tilde{\Delta}^2 (2t+1)^2/2} e^{-[q-(2t+1)\sqrt{\pi}]^2/(2\Delta^2)} |q\rangle \, dq. \tag{4}$$

The code state is thus a Gaussian-weighted sum of displaced squeezed vacuum states. This code state and its Wigner function $W(p,q) = 1/\pi \int_{-\infty}^{\infty} e^{2ipx} \Psi^*(q+x)\Psi(q-x)dx$, where $\Psi(q)$ is the wave function of the state $|0\rangle_{approx}$ in the position basis, are depicted in Figs. 1 and 2 for $\tilde{\Delta} = \Delta = 0.2$ which corresponds to 8.3 dB.

Here and elsewhere in the paper, we calculate dB as $10 \log_{10}(G) = 10 \log_{10}[\cosh^2(r)]$ where G is the amount of gain through the amplifier and $r = \ln(1/\Delta)$ [9]. In the limit of

$\Delta \to 0$ (infinite squeezing) where the width of the Gaussian envelope becomes arbitrarily broad, one obtains the perfect code states. It is worth noting that it is not essential that the parameters $\tilde{\Delta}$ and Δ are identical, nor that the filter is Gaussian: in [10] a very general form of approximate code states depending on two different filter functions was formulated.

One can obtain the form of the code states in p space by taking a Fourier transform. For example, the state $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle_{approx} + |1\rangle_{approx})$ has the following form when $\Delta/\sqrt{\pi}, \tilde{\Delta}\sqrt{\pi} \ll 1$ (see [4]):

$$|+\rangle_{approx} \propto \sum_{t=-\infty}^{\infty} \int e^{-\Delta^2 p^2/2} e^{-(p-2t\sqrt{\pi})^2/(2\tilde{\Delta}^2)} |p\rangle \, dp$$

$$\approx \sum_{t=-\infty}^{\infty} \int e^{-2\pi \Delta^2 t^2} e^{-(p-2t\sqrt{\pi})^2/(2\tilde{\Delta}^2)} |p\rangle \, dp, \tag{5}$$

where one observes that the roles of Δ and $\tilde{\Delta}$ are interchanged. If one encodes in a bosonic mode, it is natural to choose similar approximations in p and q as free evolution of the state evolves these quadratures into each other. This means that the choice $\Delta = \tilde{\Delta}$ is natural. The filter is Gaussian, so that we can use squeezed states which are Gaussian in their quadrature spread. It can be seen in Eqs. (4) and (5) that the wave functions of approximate code states are even in q and p, respectively. As the photon parity operator $P = e^{i\pi a^\dagger a}$ transforms $q \to -q$, $p \to -p$, these code states will always have an even number of photons [this is not true for the approximate shifted code states in Eq. (8)].

For an approximate code state with parameter Δ one can explicitly calculate the mean number of photons $\bar{n} = \langle a^\dagger a \rangle$ and its variance $\sigma^2(n) = \langle (\bar{n} - a^\dagger a)^2 \rangle$. Using that $\bar{n} = \langle \frac{1}{2}(p^2 + q^2) + 1 \rangle = \langle q^2 \rangle + 1$, $\Delta \ll 1/\Delta$ and approximating the infinite sum over t by an integral, one obtains $\bar{n} \approx \frac{1}{2\Delta^2}$. A similar calculation yields

$$\sigma(n) \approx \frac{1}{2\Delta^2} \sim \bar{n}, \tag{6}$$

to leading order in $1/\Delta$. This shows that the approximate code states are highly nonclassical states with large fluctuations in photon number, i.e., scaling with the total number of photons while the approximation parameter Δ controls the number of photons in the state. We will return to the number of photons in the approximately prepared code states in Sec. II E.

B. Previous work

The preparation of the code states has been discussed in several papers (see, e.g., [4,8,11–13]). In [4] it was noted that an interaction $H_{int} = qb^\dagger b$ where q is the position operator of the oscillator (e.g., a cooled micromechanical system) in which the code state is to be encoded and b is the annihilation operator of a different bosonic mode, would be useful in preparing the code states. The protocols discussed in [8,12] use post-selection in preparing the code states making the success probability extremely small, while we find that (almost no) post-selection is required by understanding the preparation step as phase estimation (see details in Sec. II E).

FIG. 2. Surface plot, contour plot, and integral over p,q of the Wigner function $W(p,q)$ for code states $|0\rangle_{approx}$ with $\Delta = 0.5$ (a) and $\Delta = 0.2$ (b). For q being an odd multiple of $\sqrt{\pi}$, the Wigner function rapidly oscillates between positive and negative values when p is an even resp. odd multiple of $\sqrt{\pi}$. Integrating over the p variable thus leaves, due to destructive interference, little amplitude for q at odd multiples of $\sqrt{\pi}$, as can also be seen in Fig. 1.

In [14], the use of cat states $|0\rangle \propto |\alpha\rangle + |-\alpha\rangle$ and $|1\rangle \propto |i\alpha\rangle + |-i\alpha\rangle$ was explored for encoding a qubit in a cavity mode and dissipative dynamics was considered to preserve the even photon-number code space. Even though dissipative dynamics can drive a state to this code space, it does not imply that such dissipative dynamics realizes quantum error correction as it has to approximately drive an encoded state in this subspace to itself. Unlike the approximate GKP code, the cat-state code by itself (implemented in circuit-QED in [15,16]) is not fully protected against loss of photons. The detection of a photon-number change by 1 constitutes a reversible change (which can be undone), but when one does not detect this change, the cat states decay irreversibly, nor is the code protected against the simultaneous loss of two photons.

In [17], multicomponent cat states were dynamically generated in a cavity mode using the Kerr nonlinearity of the cavity (see, e.g., the description of such a protocol in [18]). A multicomponent cat state is a superposition of coherent states spaced at equal distances on *a circle* in phase space while the approximate code states are created by translations *along a line*. It seems hard to define two approximately orthogonal qubit states using these multicomponent cat states such that (1) the phase space amplitude peaks for $|0\rangle$, $|1\rangle$ are sufficiently separated (so as to protect against small shift errors) *and* (2) $|+\rangle$ and $|-\rangle$ are similarly superpositions of sufficiently separated peaks.

C. Quantum error correction and protection

The primary reason for defining the two-dimensional code space as the $+1$ eigenspace of S_p and S_q is that the code states are protected against small shift errors in phase space, of the form e^{-iup} and e^{-ivq} with $|u|, |v| < \sqrt{\pi}/2$. Such shifts move the peaks of the code state in Fig. 1 as $e^{-iup}|q\rangle = |q+u\rangle$ and $e^{-ivq}|p\rangle = |p-v\rangle$. Thus, when the shift error is less than half the shift represented by the operators X or Z in magnitude, one can undo the error by shifting back the state by the minimal amount.

Imagine that a shift error e^{-ivq} has occurred and the eigenvalue of the check operator S_p is perfectly measured by some means. One will estimate the eigenvalue of S_p as $e^{i2\sqrt{\pi}v}$ (as $S_p e^{-ivq}|\psi\rangle = e^{i2\sqrt{\pi}v}e^{-ivq}|\psi\rangle$ for a state ψ in the code subspace, as follows from the commutation between e^{-ivq} and S_p). For $|v| < \sqrt{\pi}/2$, the phase $\theta = 2\sqrt{\pi}v \in (-\pi, \pi)$ is uniquely given and thus one can correct the shift error by learning the phase θ (and shifting back). For larger $|v|$, θ is consistent with two values of v and choosing the wrong one leads to a logical qubit error. Error correction for shift errors e^{-iup} works similarly by measuring the eigenvalue of S_q (which becomes $e^{i2\sqrt{\pi}u}$). In the phase estimation protocols that we propose, the goal is to approximately measure the eigenvalues of the unitary operator S_p (and S_q). Once one knows the approximate eigenvalue of, say, S_p, one could in principle apply an appropriate corrective displacement such that the eigenvalue of S_p equals 1. In Appendix A, we argue that it is not necessary to do so as any subspace characterized by fixed eigenvalues of S_p and S_q is a good code space (one can compare this phase frame to a Pauli frame used in the description of stabilizer codes).

Perfect code states which are eigenstates of S_p and S_q can be parametrized [11] as basis vectors $|u,v\rangle$ (which form a complete orthonormal basis for the oscillator space) defined as

$$|u,v\rangle = e^{-iup}e^{-ivq}|0\rangle,$$
$$u \in (-\sqrt{\pi}, \sqrt{\pi}], \ v \in (-\sqrt{\pi/2}, \sqrt{\pi/2}]. \quad (7)$$

Note that one has $S_p|u,v\rangle = e^{i2\sqrt{\pi}v}|u,v\rangle$ and $Z|u,v\rangle = e^{i\sqrt{\pi}u}|u,v\rangle$. One can observe that a state $|u,v\rangle$ in q space is simply a sum of peaks where the location of the delta functions is shifted by u, while the wave function at each peak obtains a complex phase given by $\exp(-ivq)$. One can similarly define approximate shifted code states as

$$|u,v,\Delta\rangle \equiv e^{-iup}e^{-ivq}|0\rangle_{\text{approx}}$$

$$\propto \sum_{t=-\infty}^{\infty} \int e^{-2\pi\Delta^2 t^2} e^{-(q-2t\sqrt{\pi})^2/(2\Delta^2)} e^{-ivq} |q+u\rangle \, dq.$$
$$(8)$$

For such an approximate shifted code state, the Gaussian envelope is identical, but each peak within the Gaussian envelope has obtained a phase factor e^{-ivq} (we omit an overall phase factor) and the peaks are located at the shifted positions $q = u \mod 2\sqrt{\pi}$. Of course, the width of the Gaussian envelope with standard deviation $\sigma = 1/\Delta$ should be larger than the maximal shift, i.e., $\Delta \ll 1/\sqrt{\pi} \approx 0.56$, so that all shifts lead to approximate code states within the same overall envelope. It is the goal of the phase estimation protocols in this paper to produce such approximate shifted code states $|u,v,\Delta\rangle$.

It has been shown in [11] that it is possible to do quantum error correction using perfect code states which have undergone some (coherent or stochastic) distribution of shift errors of strength $|u|, |v| < \sqrt{\pi}/6$, assuming otherwise perfect linear optical circuits. In Appendix B, we review this way of doing quantum error correction. We will use the $\sqrt{\pi}/6$ value, corresponding to a phase uncertainty $\delta\theta$ of the stabilizer checks of at most $\pi/3$ (but a phase uncertainty of $\pi/6$ for the measurement of X or Z), as a rough figure of merit to represent how well one can prepare a code state using phase estimation. We will refer to this threshold as the $\sqrt{\pi}/6$ *shift error threshold*. The "Steane quantum error correction" analyzed by Glancy and Knill already assumes the existence of an approximately prepared code state for which we devise a phase estimation method in this paper. Through approximate phase estimation one can perform quantum error correction, hence, the Glancy-Knill method may not be necessary to use at all.

More formally, we define an *effective shift error rate* $P_{\text{error}}^{\sqrt{\pi}/6}$ of a state which makes reference to the $\sqrt{\pi}/6$ threshold as follows. Any density matrix ρ, i.e., an approximately prepared code state, can be written in the basis $|u,v\rangle$ [11] as

$$\rho = \int_{-\sqrt{\pi}}^{\sqrt{\pi}} du \int_{-\sqrt{\pi}}^{\sqrt{\pi}} du'$$

$$\times \int_{-\sqrt{\pi/2}}^{\sqrt{\pi/2}} dv \int_{-\sqrt{\pi/2}}^{\sqrt{\pi/2}} dv' \rho_{uv,u'v'} |u,v\rangle \langle u',v'|. \quad (9)$$

The effective shift error rate $P_{\text{error}}^{\sqrt{\pi}/6}$ is defined as

$$P_{\text{error}}^{\sqrt{\pi}/6} = 1 - \int_{-\sqrt{\pi}/6}^{\sqrt{\pi}/6} du \int_{-\sqrt{\pi}/6}^{\sqrt{\pi}/6} dv \, \rho_{uv,uv}, \qquad (10)$$

so that for a pure state $|\psi\rangle = \int_{-\sqrt{\pi}}^{\sqrt{\pi}} du \int_{-\sqrt{\pi}/2}^{\sqrt{\pi}/2} dv \, c(u,v) |u,v\rangle$, one has $P_{\text{error}}^{\sqrt{\pi}/6} = 1 - \int_{-\sqrt{\pi}/6}^{\sqrt{\pi}/6} du \int_{-\sqrt{\pi}/6}^{\sqrt{\pi}/6} dv |c(u,v)|^2$. For the approximate code state in Eq. (4) [11] has shown that with $\bar{n} \approx$ 22 photons, $\Delta = 0.15$, the probability for shift errors beyond the $\sqrt{\pi}/6$ shift error threshold is at most 1%. If the shift errors are due to independent processes and the probability distribution $\rho_{uv,uv} = \rho_u \tilde{\rho}_v$, then one can use the effective shift error rate in p or q alone, i.e., $P_{\text{error},q}^{\sqrt{\pi}/6} = 1 - \int_{-\sqrt{\pi}/6}^{\sqrt{\pi}/6} du \, \rho_u$ and $P_{\text{error},p}^{\sqrt{\pi}/6} = 1 - \int_{-\sqrt{\pi}/6}^{\sqrt{\pi}/6} dv \, \tilde{\rho}_v$ to estimate the total error $P_{\text{error}}^{\sqrt{\pi}/6}$.

Note that this error rate is only a rough figure of merit as it makes reference to the shift error threshold of $\sqrt{\pi}/6$ which has no intrinsic meaning if quantum error correction is performed through other means. For an approximate code state one can also estimate the X or Z error probability. The X error probability is the probability that the state $|0\rangle_{\text{approx}}$ is incorrectly identified as encoding $|1\rangle$ by means of *perfect* quantum error correction. This is equal to the probability $P_{\text{error},q}^{\sqrt{\pi}/2}$ as shifts with $|u| \geqslant \sqrt{\pi}/2$ lead to the state being incorrectly identified. In [4] this error probability was approximately upperbounded by $\frac{2\Delta}{\pi} \exp(-\pi/4\Delta^2)$ which leads to a 1% error probability for $\Delta = 0.5$. Naturally, this error probability is more optimistic than the $\sqrt{\pi}/6$ threshold value, but since we expect the presence of additional (shift) errors during all protocols, using the threshold value seems a reasonable figure of merit.

In practice, the code states are realized through an approximate noisy implementation of the phase estimation measurement of S_p and S_q. Phase estimation by finite means (see Sec. II) will typically output an estimate $\tilde{\theta}$ for θ such that $\text{Prob}(|\tilde{\theta} - \theta| < \delta) > 1 - \epsilon$ for some ϵ and δ. This means that with probability at least $1 - \epsilon$, one has $|\tilde{\theta} - \theta| = \delta\theta < \delta$ phase uncertainty, which in turn corresponds to $\delta v = \frac{\delta\theta}{2\sqrt{\pi}} < \delta/(2\sqrt{\pi})$ shift uncertainty or error with probability at least $1 - \epsilon$. Note that when one measures the eigenvalue of an operator such as $Z = e^{i\sqrt{\pi}q}$ with phase uncertainty $\delta\theta < \delta$ (with high probability), then this corresponds to having a shift error of strength at most $\delta u < \delta/\sqrt{\pi}$.

D. Physical sources of errors as shift errors

In this section, we discuss how levels of decoherence or inaccuracy *of an arbitrary nature* on the oscillator can be expanded into shift errors in p and q of small strength, at least when this noise is acting on states with a sufficiently low numbers of photons. We will consider this for natural physical sources of noise and inaccuracy such as dephasing, photon loss, and quartic (self-Kerr) interactions of the bosonic mode.

Following [4], any operator E acting on a single bosonic mode can be expanded as $E = \int_{-\infty}^{\infty} d\gamma \, c(\gamma) D(\gamma)$ with complex coefficients $c(\gamma)$, i.e., an expansion into linear combinations of translations $D(\gamma)$ with complex γ in phase space. One can formally obtain the coefficients $c(\gamma)$ as $c(\gamma) = \frac{1}{\pi} \text{Tr}[E D(-\gamma)] = \frac{1}{\pi^2} \int d\alpha \, \langle\alpha| E D(-\gamma) |\alpha\rangle =$

$\frac{1}{\pi} C_W^E(-\gamma)$ where $C_W^E(\lambda)$ is the symmetrically ordered characteristic function of the operator E [18,19], whose Fourier transform is the Wigner function of the operator E. Here, $d\alpha = d \, \text{Re}(\alpha) d \, \text{Im}(\alpha)$. Thus, in principle, one can find an error expansion for any operator by evaluating $c(\gamma)$ and considering whether its support is concentrated on small values of γ. For example, the identity operator I has $c(\gamma) = \delta^2(\gamma)$, a delta function in the real and imaginary parts of γ.

At a more intuitive level, one can understand, however, that the extent to which an expansion into small shifts is warranted should depend on the number of photons in the state. Consider the action of an undesired rotation $\exp(-i\delta a^\dagger a)$ [or photon decay operator $\exp(-\delta a^\dagger a)$] in phase space. For simplicity, we apply it to a coherent state $|\alpha\rangle$: it is clear that such rotation corresponds to a larger state-dependent translation for larger α, i.e., $\exp(-i\delta a^\dagger a) |\alpha\rangle = |\alpha \exp(-i\delta)\rangle \propto D\{\alpha[\exp(-i\delta) - 1]\} |\alpha\rangle \approx D(-i\alpha\delta) |\alpha\rangle$ for $\delta \ll 1$. If $\exp(-i\delta a^\dagger a) = \int d\gamma \, c(\gamma) D(\gamma)$ were only supported on γ with $|\gamma|^2 \leqslant f(\delta)$ with some function $f(\delta)$ which is independent of the number of photons, then this contradicts the fact that there are states $|\alpha\rangle$ with $n = |\alpha|^2$ large enough to be undergoing a large displacement $|\alpha\delta| > |\gamma|$ [20].

This implies that one should consider expansions (of physical noise operators in terms of shifts) which assume an upper bound on the number of photons in the oscillator space: let us call this upper bound on the number of photons $n_{\max}$. In that case, it is clear from the above example that one can only make an expansion in small shifts, of strength at most $\sqrt{\pi}$, of $\exp(-i\delta a^\dagger a)$ when *at least* one satisfies the inequality $|\delta| n_{\max}^{1/2} < \sqrt{\pi}$. A similar argument can be given for a quartic interaction of form $\exp[-i\epsilon(a^\dagger a)^2]$. For such quartic interaction one can write, assuming $\epsilon \ll 1$,

$$\langle\alpha| \exp[i\epsilon(a^\dagger a)^2] a \, \exp[-i\epsilon(a^\dagger a)^2] |\alpha\rangle$$
$$\approx \alpha \, \exp(-i|\alpha|^2\epsilon) \exp(-|\alpha|^2\epsilon^2/2). \qquad (11)$$

Thus, ignoring fluctuations, one has an average phase space rotation (nonlinear phase shift) in addition to coherent state amplitude decay. As before, such average rotation and decay is only expressible in terms of small shifts of strength less than $\sqrt{\pi}$ when at least one obeys the inequality $|\alpha|^3|\epsilon| \leqslant |\epsilon| n_{\max}^{3/2} < \sqrt{\pi}$.

We note that these conditions are necessary, but not here proved to be sufficient as we have only applied these operators to coherent states $|\alpha\rangle$ {while for example a superposition of Fock states $|\text{vac}\rangle + |n_{\max}\rangle$ undergoes a rate of change proportional to $\epsilon n_{\max}^2$ by $\exp[-i\epsilon(a^\dagger a)^2]$}.

Since the approximate code states are superpositions of displaced squeezed vacua, we believe however that these criteria do give a good indication of whether a small-shift expansion is warranted and also capture the strength of shift errors. For the proposed protocol in Sec. III, we will show that the expected process of photon loss during the preparation of the code state has sufficiently small error rate so as to be deeply into this small-shift error regime. We will also discuss the (unwanted) cavity nonlinearities in the effective transmon qubit cavity Hamiltonian given the physical strength of parameters in a possible experiment.

Shift error expansion

Here, we describe a general method to expand an error operator E in terms of shift errors. The error operator E can arise either from open system dynamics or an unwanted unitary transformation. General open system dynamics of the bosonic mode can be modeled by a Lindblad equation of the form

$$\dot{\rho} = -i[H_{\text{ideal}}(t) + V(t), \rho] + \mathcal{D}(\sqrt{\kappa}a)(\rho) + \mathcal{D}(\sqrt{\gamma}a^\dagger a)(\rho),$$
(12)

with the compactly defined superoperator $\mathcal{D}(X)(\rho) = X\rho X^\dagger - \frac{1}{2}(X^\dagger X\rho + \rho X^\dagger X)$. Here, $H_{\text{ideal}}(t)$ is some ideal (time-dependent) dynamics (the execution of the phase estimation protocol of S_p say, or some gate implementation) and $V(t)$ a possible perturbation or correction [e.g., an unwanted self-Kerr term $K(a^\dagger a)^2$]. The Lindblad equation can model photon loss $D(\sqrt{\kappa}a)$ of rate κ or dephasing $D(\sqrt{\gamma}a^\dagger a)$ of rate γ. For simplicity, one can consider the effects of the different sources of noise separately [and ignore $H_{\text{ideal}}(t) + V(t)$] and construct a superoperator for the evolution in a short time interval, as $\mathcal{S}(\rho(t+\tau)) = \sum_i A_i \rho(t) A_i^\dagger$. For the process of photon loss one has $A_0 = I - \frac{\tau\kappa}{2}a^\dagger a + O((\tau\kappa)^2 n_{\text{max}}^2)$ and $A_1 = \sqrt{\kappa\tau}a$. For dephasing one has $A_0 = I - \frac{\tau\gamma}{2}(a^\dagger a)^2 + O((\tau\gamma)^2 n_{\text{max}}^4)$ and $A_1 = \sqrt{\gamma\tau}a^\dagger a$. These Kraus operators are examples of error operators E. It is clear that one can express such operators as low-order polynomials in a and $a^\dagger$ and we will describe how such a low-order polynomial can be expanded in shift errors.

Similarly, a unitary error operator E can be an over-rotation $\exp(-i\delta a^\dagger a)$ or a nonlinearity $\exp[-i\epsilon(a^\dagger a)^2]$ where δ, ϵ are given by an error strength times a time scale. Assuming that $\delta n_{\text{max}} \ll 1$ or $\epsilon n_{\text{max}}^2 \ll 1$, one can Taylor expand this in terms of I and low powers of a and $a^\dagger$.

We now write a simple Taylor expansion in terms of correctable and incorrectable shift errors for the operator $\sqrt{\delta}a \equiv x_1 + ix_2$ with $x_1 = p\sqrt{\delta/2}$ and $x_2 = q\sqrt{\delta/2}$ assuming $\delta n_{\text{max}} \ll 1$. This expansion and its Hermitian conjugate can then be used to expand operators such as $a^2, a^\dagger a$, etc. With the Taylor expansion of $\arcsin(x) \approx x + \frac{x^3}{6} + \frac{3x^5}{40} + \frac{5x^7}{112} + O(x^9)$, we can expand $x_1 + ix_2$ in terms of $\sin(x_1)$, $\sin^3(x_1)$ and $\sin(x_2)$, $\sin^3(x_2)$, etc. For example, in lowest order one has

$$\sqrt{\delta}a = \sin(p\sqrt{\delta/2}) + \sin^3(p\sqrt{\delta/2})/6$$
$$+ i[\sin(q\sqrt{\delta/2}) + \sin^3(q\sqrt{\delta/2})/6] + O((\delta n_{\text{max}})^{5/2}),$$
(13)

where the $\sin(p\sqrt{\delta/2})$ and $\sin^3(p\sqrt{\delta/2})$ function [and similarly $\sin(q\sqrt{\delta/2})$] can be expanded in terms of $e^{\pm ip\sqrt{\delta/2}}$, $e^{\pm 2ip\sqrt{\delta/2}}$, $e^{\pm 3ip\sqrt{\delta/2}}$, etc. In general, we can thus write $E = \sqrt{\delta}a = E_{\text{correctable}} + E_{\text{incorrectable}}$ where $E_{\text{correctable}}$ contains all terms in the Taylor expansion up to odd kth order such that $k\sqrt{\delta/2} \approx \sqrt{\pi/2}$ and $||E_{\text{incorrectable}}|| \lesssim O((\delta n_{\text{max}})^{1+k/2})$. We can use this to expand an operator $(\sqrt{\delta}a)^p$ [or $(a^\dagger a)^{p/2}$ or $(a^\dagger)^p$] as $(\sqrt{\delta}a)^p = E_{\text{correctable}} + E_{\text{incorrectable}}$ where $||E_{\text{incorrectable}}|| = O((\delta n_{\text{max}})^{1+k/2})$ with $k\sqrt{\delta} \approx \sqrt{\pi}/p$. This shows that on an encoded state, the amplitude of an incorrectable error can be reduced *by several orders of magnitude*, depending on the strength of the errors.

One thing to note is that in order to systematically expand noise processes in terms of shift errors, one should develop the

full Kraus error operator or undesired rotation to a certain order in δn_{max}. In other words, depending on which order in δn_{max} one chooses, one includes higher-order terms in the expansion of the error Kraus operators A_i or the unitary.

One can interpret the protection that the code offers with the following example. Assume that photon loss from a cavity occurs for some time t at rate κ such that $P \equiv \kappa t n_{\text{max}} < 1$. Without encoding, the error operator $E = \sqrt{\kappa t}a$ produces an incorrectable error with probability $\sim P$. For a cat-state encoding (see Sec. I B) this process is correctable, but for the cat encoding, a two-photon loss error operator $E = \kappa t a^2$ is incorrectable (as it is proportional to Z). Hence, incorrectable errors happen with probability $\sim P^2$.

If we expand the two-photon loss error operator in terms of correctable and incorrectable shift errors, the probability for the incorrectable term is of order $(\kappa\tau n_{\text{max}})^{1+k/2} = P^{1+k/2}$ where $k \approx \frac{1}{2}\sqrt{\frac{\pi n_{\text{max}}}{P}}$. Thus, for small $P \ll 1$ the probability for an incorrectable error resulting from two-photon loss can be much reduced if one uses a code that can correct shift errors as compared to a code which can only correct single-photon loss events.

II. PHASE ESTIMATION

The measurement of the eigenvalue $e^{i\theta}$ of a unitary operator U, and the simultaneous projection of the input state onto the corresponding eigenstate $U|\psi_\theta\rangle = e^{i\theta}|\psi_\theta\rangle$, is called phase estimation for U. In our case, we have $U = S_p = D(\sqrt{2\pi})$, say. Phase estimation can be executed by repeatedly running a circuit of a general form depicted in Fig. 3 for varying k and phases φ. Many variants of phase estimation exist (see, e.g., a recent analysis in [21]) depending how or whether one varies k and/or whether one allows φ to depend on earlier qubit measurements, so using feedback, and how one infers the phase from the information obtained from the sequence of qubit measurement outcomes.

We will first consider standard phase estimation as it is used in Shor's factoring algorithm and argue that this method does not give the kind of approximate code state that we are looking for, at least not when we wish to use low numbers of photons. The standard phase estimation has the same performance as Kitaev's phase estimation described in [22]; both require realizations of U^{2^k} for increasing k.

The best form of approximate phase estimation optimizes the accuracy on the phase given a mean number of photons $\bar{n}$ in the approximately prepared code state as we always work with a bounded number of photons. In this respect, it can be noted that for the standard phase estimation protocol described in the following, the phase uncertainty scales as $\delta\theta \sim 2^{-M} \sim \frac{1}{\sqrt{\bar{n}}}$, that is, shot-noise limited. Protocols which can achieve Heisenberg-limited scaling, that is $\delta\theta \sim \frac{1}{\bar{n}}$, have been considered and analyzed in detail in [23–25] (with an improved accuracy analysis in [26]). One scheme that can reach the Heisenberg limit is a modified form of the Kitaev's phase estimation in which the circuit for each controlled-$D(2^k\sqrt{2\pi})$ in Fig. 3 is repeated a number of times depending on k and M [25]. One can thus expect that this Heisenberg-limited scheme performs optimally, but we do not consider it here as the experimental realization of the simpler schemes will
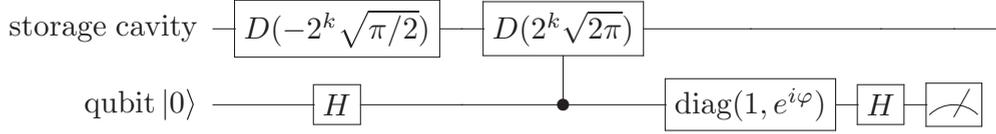
FIG. 3. Phase estimation for the unitary operator $S_p = e^{-i2\sqrt{\pi}p} = D(\sqrt{2\pi})$ where $D(\alpha)$ is the displacement operator. Assume that phase estimation uses M ancillas prepared in $|00\ldots0\rangle$. The circuit in this figure is repeatedly executed for $k = M - 1, \ldots, 0$ starting at $k = M - 1$. The phase φ in the single-qubit rotation around the z axis, $\text{diag}(1, e^{i\varphi})$, will depend on the outcomes of *all* the previously measured ancillas. This sequential realization of phase estimation is identical to normal phase estimation as it merely uses a semiclassical realization of the quantum Fourier transform. Note that the circuit is identical to one in which $\text{diag}(1, e^{i\varphi})$ is moved before the controlled-displacement gate, which is the form of the quantum circuit in [23]. Prior to the controlled-displacement gate, the cavity is (unconditionally) displaced so that the code states are symmetrically centered around the vacuum state and we minimize the total number of photons.

already be demanding. When one allows for a larger number of ancilla qubit rounds, one can expect that switching to such scheme which uses U^2 or U^4, etc., is better.

In Secs. II B and II C, we analyze two simple phase estimation protocols, repeated or nonadaptive phase estimation and phase estimation by feedback or adaptive phase estimation. These schemes only use U and thus no increasingly large displacements (microwave power).

A. Standard phase estimation

Any method for estimating the phase by sequentially acquiring bits of information is information-theoretically bounded: it is clearly optimal if each acquired bit gives us, with certainty, one additional bit of the binary expansion of $\frac{\theta}{2\pi} = 0.\theta_1\theta_2\ldots$, starting with the most significant bit θ_1, etc. This is close to what the standard phase estimation protocol, as it is invoked in Shor's algorithm, achieves [27]. It is known that this phase estimation protocol can be realized sequentially (see e.g. [28–30]) by coupling a sequence of M single qubits to the input state and applying controlled-U^{2^k} gates ($k = 0, \ldots$) with the ancilla qubit as control. This can be understood by replacing the Fourier transform in the standard phase estimation by the semiclassical Fourier transform whose quantum circuit is one in which the first qubit is measured and its outcome is used to apply single-qubit gates on the remaining $M - 1$ qubits. The next qubit is measured and again its outcome determines single-qubit gates on the remaining $M - 2$ qubits, etc. (see Fig. 5.1 in [27]). In this way, the whole procedure can be implemented sequentially, in rounds with one active control qubit per round. We depict one round of such a phase estimation protocol for $U = S_p = D(\sqrt{2\pi})$ in Fig. 3. We can consider the measurement operator M_x that is applied to the input state, upon getting an M-bit outcome $x \in \{0,1\}^M$, with $U = D(\sqrt{2\pi})$, including the compensating displacements in each circuit in Fig. 3 which center the code state around the vacuum state. One obtains $M_x |\psi_{\text{input}}\rangle$, with

$$M_x = \frac{1}{2^M} \sum_{t \in \{0,1\}^M} e^{-2\pi i x t / 2^M} D(\sqrt{2\pi} t) D\left(-\sqrt{2\pi}\frac{2^M - 1}{2}\right)$$

$$\underset{\sim}{\propto} \sum_{t=-2^{M-1}}^{2^{M-1}} e^{-2\pi i x t / 2^M} D(\sqrt{2\pi} t). \qquad (14)$$

[Here, $\underset{\sim}{\propto}$ relates to the fact that we simply approximate $(2^M - 1)/2 \approx 2^{M-1}$.] For phase estimation one can prove that, using $M = \tilde{M} + \log_2(\frac{1}{2} + \frac{1}{2\epsilon})$ ancilla qubits, we obtain the best possible $\tilde{M}$-bit estimate of the binary expansion of $\frac{\theta}{2\pi}$ with probability of success at least $1 - \epsilon$ (when one chooses $\tilde{M} = M$, $P_{\text{success}} = \frac{4}{\pi^2}$). Let $\tilde{\theta} \in (-\pi, \pi]$ represent this $\tilde{M}$-bit estimate. Let x_{round} be the bit-string outcome x rounded off to its $\tilde{M}$-most significant bits. For $x_{\text{round}}/2^{\tilde{M}} \leqslant \frac{1}{2}$, we choose $\tilde{\theta} = \frac{2\pi x_{\text{round}}}{2^{\tilde{M}}}$, while for $x_{\text{round}}/2^{\tilde{M}} > \frac{1}{2}$, one takes $\tilde{\theta} = \frac{2\pi x_{\text{round}}}{2^{\tilde{M}}} - 2\pi$.

Knowing the first $\tilde{M}$ bits of $\theta/(2\pi)$ leads to an error $\delta\theta \leqslant 2\pi 2^{-\tilde{M}}$. This means that one can prepare an approximate S_p eigenstate with $P_{\text{success}} \geqslant \frac{9}{16}$ (for which $M = \tilde{M} + 2$) such that the shift error is less than the $\sqrt{\pi}/6$ shift error threshold (correspond to $\pi/3$ phase error, see Sec. I C) if we take $M \geqslant 4$.

There are a few ways to see that this method of measuring the eigenvalue of S_p is not well suited for the approximation that we are seeking. The protocol requires the implementation of controlled-$D(2^k\sqrt{2\pi})$ for $k = 0, \ldots M - 1$ (Fig. 3). This means that the number of photons in the cavity mode grows exponentially with the number of ancilla qubits used. More precisely, one can calculate the number of photons in an approximate code state produced by applying phase estimation of S_p onto the squeezed vacuum state in Eq. (3), $|\psi_x\rangle \propto M_x |\text{vac}_{\text{sq}}\rangle$. One has

$$\bar{n}_x = \frac{\langle \text{vac}_{\text{sq}}| M_x^\dagger a^\dagger a M_x |\text{vac}_{\text{sq}}\rangle}{\langle \text{vac}_{\text{sq}}| M_x^\dagger M_x |\text{vac}_{\text{sq}}\rangle} \approx 2\pi \frac{1}{2^M} \sum_{k=-2^{M-1}}^{2^{M-1}} k^2 + \bar{n}_{\text{sq}}$$

$$= 2\pi \left(\frac{2^{2M-2}}{3} + \frac{2^{M-1}}{2} + \frac{1}{6}\right) + \bar{n}_{\text{sq}},$$

where $\bar{n}_{\text{sq}}$ is the expected number of photons of the squeezed vacuum state $|\text{vac}_{\text{sq}}\rangle$. Here, we have used that $D(-\alpha)aD(\alpha) = a + \alpha$ and $\langle \text{vac}_{\text{sq}}| D[\sqrt{2\pi}(t - t')] |\text{vac}_{\text{sq}}\rangle \approx 0$ for $|t - t'| \geqslant 1$ [the numerical value is $O(10^{-13})$ for $\Delta = 0.2$]. In principle, $\bar{n}_x$ depends on the outcome x, but the fluctuations with x vanish when we approximate overlaps between different displaced squeezed states by 0. To leading order in M we have $\bar{n}_x \approx 2\pi \frac{2^{2M}}{12} + \bar{n}_{\text{sq}}$, scaling exponentially with M as expected. For $M = 4$, one has $\bar{n} \approx 134 + \bar{n}_{\text{sq}}$. This number is very high compared to the 22 photons in a Gaussian-approximate code state with $\Delta = 0.15$ which has high probability to be below the shift error threshold.
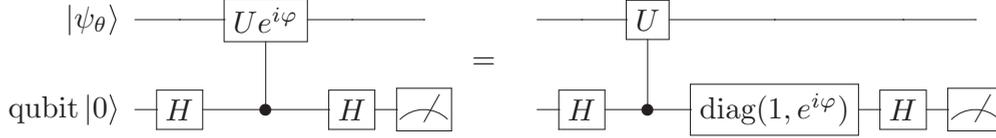
FIG. 4. Phase estimation for a unitary operator U, with varying (feedback) phase φ, and $U|\psi_\theta\rangle = e^{i\theta}|\psi_\theta\rangle$. For repeated phase estimation one chooses either $\varphi = 0$ and $\pi/2$ for half the number of rounds. In a simple form of phase estimation with feedback the phase φ is changed after each outcome so as to optimize the sensitivity of the probability distribution with respect to the currently estimated value for θ.

Another way to understand this is to consider the form of the measurement operator in Eq. (14). Equating $x \approx x_{\rm round}$, we have $M_{x \to \tilde{\theta}} \propto \sum_{t=-2^{M-1}}^{2^{M-1}} e^{-i\tilde{\theta}t} D(\sqrt{2\pi}t)$. One can compare this measurement operator to the projector onto the space with $S_p = e^{i\theta}$ which equals

$$\Pi_{S_p=\exp(i\theta)} = \sum_{t=-\infty}^{\infty} e^{-i\theta t} D(\sqrt{2\pi}t), \qquad (15)$$

and note the similarity between the two operators. This shows that the filter used in standard phase estimation is a wide box-car filter with a hard cutoff, while we are seeking a smooth Gaussian filter. If we were to choose controlled-displacement operators with exponentially growing displacements, one has to show that the duration of these gates is not necessarily exponentially increasing as such growth in time would lead to more loss and decoherence during the QEC cycle and code state preparation. In principle, large controlled displacements can be obtained by large microwave power (see Appendix D), but one may expect that the inaccuracy and undesired side effects of exponentially large displacements are also exponentially increasing.

B. Phase estimation by repetition (nonadaptive)

The simplest way to estimate $\theta \in (-\pi, \pi]$ of a unitary operator $U|\psi_\theta\rangle = e^{i\theta}|\psi_\theta\rangle$ is to repeat the quantum circuit in Fig. 4 with $\varphi = 0$ and $\pi/2$. For the $\varphi = 0$ measurement, each ancilla qubit then has a probability for outcome 0 equal to $P_{\varphi=0}(0|\theta) = \frac{1}{2}[1 + \cos(\theta)]$ while for the $\varphi = \pi/2$ measurement one has $P_{\varphi=\pi/2}(0|\theta) = \frac{1}{2}[1 - \sin(\theta)]$. Note that a simple repetition of the $\varphi = 0$ measurement is insufficient since $P_{\varphi=0}(0|\theta)$ is the same for θ and $-\theta$. Thus, one chooses $\varphi = 0$ for half of the number of rounds and ancilla qubits and $\varphi = \pi/2$ for the other half [28].

In an adaptive phase estimation (see Sec. II C), one takes into account that the sensitivity of the probability distribution $P_\varphi(0|\theta) = \frac{1}{2}[1 + \cos(\theta + \varphi)]$ to θ, i.e., $\frac{dP_\varphi(0|\theta)}{d\theta}$, is a function of θ. One would like to optimize this sensitivity by choosing values for *the feedback* phase φ which depend on previous measurement outcomes. In this optimization one chooses a next phase φ such that the measurement with that φ maximizes the sharpness of the resulting inferred probability distribution over θ. This method has been analyzed and described in detail in [25,31].

One round of phase estimation for $U = S_p$ is depicted in Fig. 5. Instead of starting with an eigenstate, one would like to understand how one approximately projects onto an eigenstate of S_p using this repeated measurement. We can first consider the action of this circuit on a coherent state $|\alpha\rangle$ when the qubit

is measured as $x = 0,1$:

$$|\alpha\rangle \to |\alpha - \sqrt{\pi/2}\rangle + (-1)^x e^{i\varphi}|\alpha + \sqrt{\pi/2}\rangle, \qquad (16)$$

where the resulting state has not been normalized. Repetition of such a circuit on the resulting output state will thus produce a sum of coherent states on a line, each with a phase which depends on the feedback phase and the measurement outcome. If all phases add constructively, this distribution of amplitudes and weights of coherent states will be as in a Pascal triangle, hence binomial, as each coherent state gets split in two equidistant coherent states at every round [see, e.g., Fig. 14(a) in Appendix D].

A more formal way of showing that the filter of this protocol is binomial and thus approximately Gaussian is as follows. Consider the measurement operator for M rounds with $\varphi = 0$. An outcome bit string $x \in \{0,1\}^M$ will correspond to a measurement operator which can be labeled by the Hamming weight $k = w_H(x)$ as the order of the outcomes of $1s$ and $0s$ is irrelevant. The measurement operator equals

$$M_k \propto [I + D(\sqrt{2\pi})]^{M-k}[I - D(\sqrt{2\pi})]^k D(-M\sqrt{\pi/2})$$

$$= \sum_{p_1=0}^{k} \sum_{p_2=0}^{M-k} \binom{k}{p_1}\binom{M-k}{p_2}(-1)^{p_1}[\sqrt{2\pi}(p_1+p_2-M/2)].$$

When all measurement outcomes $x_i = 0$ (from which one would conclude that $\theta = 0$) the measurement operator has the simple form of a binomial sum of displacements:

$$M_{k=0} \propto \sum_{m=0}^{M} \binom{M}{m} D[\sqrt{2\pi}(m - M/2)]$$

$$\underset{\sim}{\propto} \sum_{t=-M/2}^{M/2} e^{-2t^2/M} D(\sqrt{2\pi}t).$$
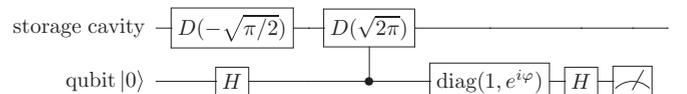
FIG. 5. One round of phase estimation for the unitary operator $S_p = e^{-i2p\sqrt{\pi}} = D(\sqrt{2\pi})$ where $D(\alpha)$ is the displacement operator. The phase $e^{i\varphi}$ of the R_z-qubit rotation can be either fixed to $\varphi = 0$ and $\pi/2$ or adaptively changed per round depending on previous outcomes as in phase estimation by feedback (Sec. II C). Prior to the controlled-displacement gate, the cavity is (unconditionally) displaced so that the code states are symmetrically centered around the vacuum state and we minimize the total number of photons.

One can roughly identify this measurement operator with the Gaussian-filtered projection operator onto approximate code states in Eq. (4) with $\tilde{\Delta}^2 \approx \frac{1}{\pi M}$ (note that the measurement operator inevitably has a hard cutoff while the Gaussian filter does not). This shows that for $\tilde{\Delta} = 0.2$ one can choose $M = 25/\pi \approx 8$ with a much lower number of photons than in regular phase estimation.

Instead of post-selecting on this outcome, we want to use the measurement data to estimate the value of θ, and include the data from the $\varphi = \pi/2$ measurement. We can analyze the efficiency of this method using the Chernoff bound; this argument is explicitly developed in [24]. Section II E gives numerical details on the phase uncertainty and the number of photons. Assume that we use $M/2$ ancilla qubits for $\varphi = 0$ and $M/2$ qubits for $\varphi = \pi/2$ such that $P_{\varphi=0}(0)$ is estimated as $\tilde{P}_{\varphi=0}$ and $P_{\varphi=\pi/2}(0)$ is estimated as $\tilde{P}_{\varphi=\pi/2}$. The estimate $\tilde{\theta}$ is chosen as

$$\tilde{\theta} = \arg[2\tilde{P}_{\varphi=0} - 1 - i(2\tilde{P}_{\varphi=\pi/2} - 1)],$$
$$\tilde{\theta} \in (-\pi, \pi]. \tag{17}$$

Using a Chernoff bound in both cases, $\text{Prob}(|\tilde{P} - P| \geqslant \delta) \leqslant 2e^{-2\delta^2 M}$ and some further bounding arguments [24], one can show that $\text{Prob}[\delta\theta \geqslant \frac{\pi}{3}\sqrt{f(M)/M}] \leqslant 4\exp[-3f(M)/16]$ for any function $f(M) \leqslant M$ for $M \geqslant 1$. This implies that the probability to prepare a code state with phase uncertainty below $\pi/3$ is at least $1 - 4\exp(-3M/16)$. This argument shows that the number of ancillas should be at least $M \geqslant 8$ and the probability of failure will then rapidly vanish. We can note that the phase uncertainty $\delta\theta \sim \frac{1}{\sqrt{M}}$ for M rounds, each of which adds $O(1)$ photons to the state, hence scaling in the expected (shot-noise) way.

C. Phase estimation with feedback (adaptive)

The uncertainty of a probability distribution $P(\theta)$ over phases θ can be measured by the Holevo phase variance defined as $V[P(\theta)] = S[P(\theta)]^{-2} - 1$ with the *sharpness* $S[P(\theta)] \in [0,1]$ given by

$$S[P(\theta)] \equiv |\langle e^{i\theta} \rangle| = \left| \int_{-\pi}^{\pi} d\theta\, e^{i\theta} P(\theta) \right|. \tag{18}$$

For a δ-function distribution in θ, the variance is 0, while for a flat distribution $V[P(\theta)] \to \infty$. One can show (see, e.g., [31]) that for small phase variance $V[P(\theta)] \approx \Delta^2(\theta)$ where $\Delta^2(\theta) = \langle (\theta - \langle\theta\rangle)^2 \rangle$ is the usual variance.

Assume that we execute the circuit in Fig. 4 for M rounds so that one gets an M-bit estimate $\tilde{\theta}$ of the phase θ and let $x[m]$ be the m-bit string of 0 or 1 measurement outcomes after m rounds. The adaptive phases will be set to the values $\varphi_1, \ldots \varphi_m$, $m = 1, \ldots, M$, and the question is how these will be chosen to maximize the sharpness and thus minimize the Holevo phase variance. We assume no prior knowledge on the phase θ, i.e., the initial probability distribution over θ, $P(\theta)$, is assumed to be a flat distribution. This is reasonable as the squeezed (in q) vacuum state to which the S_p measurement is applied is a superposition over eigenstates with rather uniformly distributed phases θ of S_p. If we assume no prior knowledge about θ, one may as well choose the first phase $\varphi_1 = 0$, which is what we do.

It can then be argued (see the self-contained analysis in Appendix C) that one should choose the next phases $\varphi_2, \ldots, \varphi_M$ as follows, depending on the previous measurement outcomes:

$$\varphi_m = \arg\max_{\varphi} \sum_{x_m=0,1} \left| \int d\theta\, e^{i\theta} P_{\varphi}(x[m]|\theta) \right|, \tag{19}$$

where the probability $P_{\varphi}(x[m]|\theta)$ is the probability of obtaining measurement outcomes $x_1, \ldots, x_m$ given an eigenstate $|\psi_\theta\rangle$. This probability has a simple expression as the measurement results of each round are independent, i.e.,

$$P_{\varphi=\varphi_m}(x[m]|\theta) = \prod_{i=1}^{m} P_{\varphi_i}(x_i|\theta) = \prod_{i=1}^{m} \cos^2\left(\frac{\theta+\varphi_i}{2} + x_i\frac{\pi}{2}\right). \tag{20}$$

For a (small) number of measurements, say, $M = 1, \ldots, 20$, one can simply solve this expression for the optimal values for $\varphi_2, \varphi_3, \ldots$ numerically given all previous possible measurement outcomes and store these optimal values as a lookup table, which is what we have done.

Given that one has obtained a final measurement record $x[M]$, how does one choose an estimate for the phase $\tilde{\theta}$? In [31] it is argued that the estimated value $\tilde{\theta}$ should be chosen as

$$\tilde{\theta} = \arg \int_{-\pi}^{\pi} d\theta\, e^{i\theta} P(\theta|x[M]) = \arg \int_{-\pi}^{\pi} d\theta\, e^{i\theta} P(x[M]|\theta). \tag{21}$$

Here, we have used that $P(\theta|x[M]) = P(x[M]|\theta)P(\theta)/P(x[M])$. Note that $P(\theta)$ is a flat distribution and $P(x[M]) = 2^{-M}$ which do not influence the arg function. The probability $P(x[M]|\theta)$, which implicitly depends on the feedback phases $\varphi_1, \ldots, \varphi_m$, was given in Eq. (20). If all feedback phases are set to 0, then $\tilde{\theta}$ will be estimated as 0 or π as the function of which we take the argument is real: this will, however, not occur since $\varphi_2 \neq 0$ after the first bit has been generated. If we do not use any feedback, the value of this estimate $\tilde{\theta}$ coincides with the estimate in Eq. (17).

In Fig. 6, we plot the probability for obtaining a phase error $\delta\theta = |\tilde{\theta} - \theta|$ given a fixed number of rounds M, averaged over random input phases θ and runs through the protocol for both the nonadaptive phase estimation protocol and the adaptive
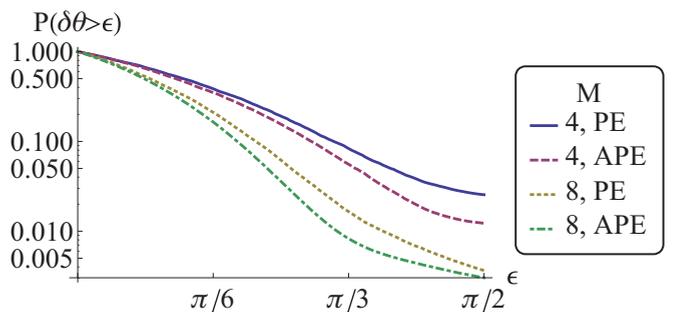
FIG. 6. Total probability $P(\delta\theta > \epsilon)$ with $\delta\theta = |\tilde{\theta} - \theta|$ versus ϵ for $M = 4, 8$ (averaged over $\theta \in (-\pi, \pi]$). APE is the adaptive phase estimation protocol described in this section. PE is the nonadaptive protocol where one sets the feedback phase $\varphi = 0$ for $M/2$ rounds and $\varphi = \pi/2$ for $M/2$ rounds. For $\delta\theta < \pi/3$, one is below the shift error threshold.

phase estimation protocol. One important difference between the two schemes is that in the nonadaptive scheme with $\varphi = 0$ and $\pi/2$ one has a total of $(\frac{M}{2} + 1)^2$ possible outcomes, as the order of the outcomes for a fixed value of φ does not affect the measurement operator which is applied. This is clearly not true for adaptive phase estimation. Setting $\varphi_1 = 0$, we have two possible values for $\varphi_2 = \varphi_2^{x_1}, x_1 = 0, 1$. Then, there are four possible values for $\varphi_3^{x_1 x_2}$, etc., as the optimization in Eq. (19) depends on all previous outcomes. The final estimate for θ in Eq. (21) will depend on all these phases $\varphi_i^{x_1 \ldots x_{i-1}}$ and $\tilde{\theta}$ can thus take on 2^M possible values. These arguments suggest that the adaptive protocol can give a more accurate estimate of θ. Figure 6 shows indeed how adaptive phase estimation (APE) outperforms such simple phase estimation by repetition (PE) for a small number of rounds M. This difference will become more pronounced for larger M (see, e.g., [31]), but the improvement is relatively small here.

In order to understand better how well these phase estimation methods project the input state onto an approximate eigenstate, we explicitly numerically generate the states that are created through these protocols in Sec. II E.

D. Preparing the input state: Two methods

One has two alternatives for preparing a code state such as $|0\rangle$ (or $|+\rangle$) approximately. As we have argued, one can start the protocol in a sufficiently squeezed vacuum state $|\mathrm{vac_{sq}}\rangle$ and approximately measure S_p (or squeeze in the other quadrature and measure S_q to prepare $|+\rangle$). Since the squeezed vacuum state is an approximate eigenstate of Z, one produces an approximate $|0\rangle$ state. How many dB of squeezing does one need in order to get a good eigenstate of Z? Let the squeezed vacuum state be $|\mathrm{vac_{sq}}\rangle = \int dq\, \alpha_q |q\rangle$ [Eq. (3)], with $\Delta = e^{-r}$ and $\alpha_q = \frac{1}{\sqrt[4]{\pi}} e^{r/2} e^{-q^2 e^{2r}/2}$. The probability for shift errors above the shift error threshold is given by $P_{\mathrm{error},q}^{\sqrt{\pi}/6} \leqslant \int_{-\sqrt{\pi}/6}^{\sqrt{\pi}/6} dq\, |\alpha_q|^2 = \mathrm{Erf}(e^r \sqrt{\pi}/6)$. One sees in Fig. 8 that the approach to 100% success probability happens around 8 dB of squeezing. Squeezing of itinerant microwave fields, instead of a confined cavity field, can be achieved using a Josephson parametric amplifier (JPA). For example, in [32] one obtains $\Delta^2 = (\Delta q_{\mathrm{sq}})^2/(\Delta q_{\mathrm{vac}})^2 \approx 12\%$ corresponding to $\Delta \approx 0.35$ and squeezing around 4.1 dB. More recent work has achieved squeezing of at least 10 dB [33]. A more recent proposal was considered in [34]: in this paper it is analyzed how a squeezed microwave drive can be used to produce a squeezed vacuum state as the stationary state of the cavity field (under dissipative photon loss dynamics).

An alternative is to start the entire protocol in the vacuum state of the cavity mode and execute the phase estimation protocol both for Z as well as S_p. The approximate measurement of Z is thus also an effective means of producing a squeezed state from the vacuum state in the cavity mode. The phase uncertainty $\delta\theta$ in the estimate of the eigenvalue of Z needs to be at most $\pi/6$ in order to be below the shift error threshold (instead of $\pi/3$), *twice* as small as compared to the S_p measurement. This suggests that if the phase of S_p is measured in M rounds, then the phase of Z should be measured in $4M$ rounds in order to give an overall similar shift error contribution. As $Z = D(i\sqrt{\pi/2})$ as compared to

$S_p = D(\sqrt{2\pi})$ the average number of photons added per Z round is $\frac{1}{4}$ of a S_p round, hence, the total contribution to the number of photons would be about the same. These arguments show that it may be most advantageous to start with a squeezed vacuum state instead of selecting the code state by phase estimation of Z as the Z measurement would be rather lengthly (at least 32 rounds for $M \geqslant 8$) during which the code state is also decohering. At the same time, the Z-phase estimation protocol by itself may be an interesting way to prepare a squeezed state.

E. Numerical analysis of two phase estimation schemes for S_p

We apply the nonadaptive and adaptive phase estimation schemes for S_p with M rounds to a squeezed vacuum state with parameter Δ in order to create an approximate code state $|0\rangle_{\mathrm{approx}}$. The state obtained through the sequence of measurements is a superposition of displaced squeezed vacuum states, each with a phase which depends on the feedback phase and the measurement outcome $x[M]$:

$$|\Psi(x[M])\rangle = \frac{1}{\sqrt{N}} \sum_{j=0}^{M} c_j(x[M]) D[\sqrt{2\pi}(j - M/2)] |\mathrm{vac_{sq}}\rangle,$$

(22)

with normalization $N \approx \sum_{j=0}^{M} |c_j(x[M])|^2$ {the approximation in the normalization comes from $\langle \mathrm{vac_{sq}}| D[\sqrt{2\pi}(t - t')] |\mathrm{vac_{sq}}\rangle \approx 0$ for $|t - t'| \geqslant 1$}. Here, the coefficients $c_j(x[M]) = \sum_{\{S_j\}} \prod_{k \in S_j} e^{i(\varphi_k + x_k \pi)}$ where the subsets S_j are subsets of $j = |S_j|$ indices chosen from $1, \ldots, M$ (without repetition). There are $\binom{M}{j}$ such subsets S_j. Also, φ_k is the feedback phase for round k and x_k is the outcome bit of round k. From the phase estimation one infers a value $\tilde{\theta} \in (-\pi, \pi]$, upon which one corrects $|\Psi(x[M])\rangle \rightarrow |0(x[M])\rangle_{\mathrm{approx}} = e^{ivq} |\Psi(x[M])\rangle$ with $v = \frac{\tilde{\theta}}{2\sqrt{\pi}}$. Note that the additional displacement to center the code state around the vacuum does not affect the eigenvalue of S_p as it commutes with S_p.

A few examples of the wave functions of states that one can obtain are shown in Fig. 7. These examples show that one is not always guaranteed to get a good wave function and the effective probability of error can vary depending on the measurement outcomes. Because the relation between the effective error
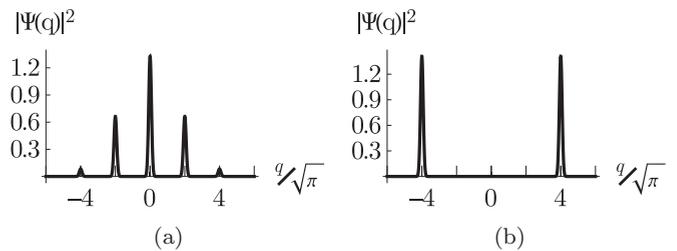
FIG. 7. Absolute values of two examples of wave functions of code states which can be generated by nonadaptive phase estimation for $M = 4$ and $\Delta = 0.2$. After the eigenvalue measurement, the states are shifted onto an approximate code state using the phase estimate. The left state (a) is obtained if the first and the second measurements of both U and iU yield the same result, the right state (b) is obtained if both results differ.
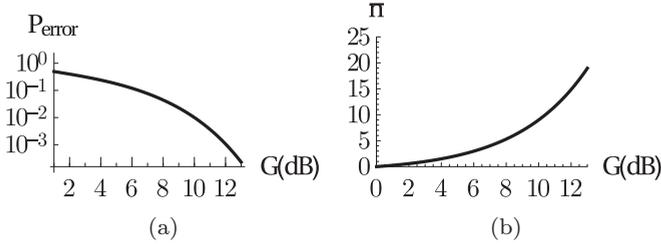
FIG. 8. (a) Upper bound on the effective error probability $P_{\mathrm{error}} = P_{\mathrm{error},q}^{\sqrt{\pi}/6}$ for a squeezed vacuum state in the *squeezed* quadrature. (b) The average number of photons in the squeezed vacuum state depending on the gain G in dB.

FIG. 9. The average number of photons $\langle \bar{n} \rangle_\theta$ after (a) adaptive phase estimation (APE) of S_p and (b) nonadaptive phase estimation (PE) of S_p using M ancilla qubits (rounds) with parameter $\Delta = 0.2$. The fluctuations $\sqrt{\langle (\bar{n} - \langle \bar{n} \rangle_\theta)^2 \rangle_\theta}$ around the average number of photons reflect the different measurement outcomes in the protocol. It is noteworthy that these fluctuations are much smaller for the APE than for the PE protocol.

rate and the measurement result is known, it is possible to use heralding or effectively use some post-selection. If we know (by simulation) that the obtained code state is bad, one would repeat the protocol.

In Fig. 15 at the end of the paper we plot the total probability P to prepare a code state with effective error rate $P_{\mathrm{error},p}^{\sqrt{\pi}/6}$ for the (adaptive) phase estimation. As the effective error rate due to finite squeezing and the effective error rate due to the phase estimation scheme are unrelated (one corresponds to shifts in q, while the other corresponds to shifts in p), only the part $P_{\mathrm{error},p}^{\sqrt{\pi}/6}$ due to the phase estimation scheme is shown. The effective error rate including both effects can be computed using $P_{\mathrm{error}}^{\sqrt{\pi}/6} = P_{\mathrm{error},p}^{\sqrt{\pi}/6} + P_{\mathrm{error},q}^{\sqrt{\pi}/6} - P_{\mathrm{error},p}^{\sqrt{\pi}/6} P_{\mathrm{error},q}^{\sqrt{\pi}/6}$. The error rate $P_{\mathrm{error},q}^{\sqrt{\pi}/6}$ is shown in Fig. 8.

Both protocols show a moderate probability to obtain code states with low error rates. What is striking is that the adaptive version of the protocol is much more reliable than the nonadaptive version and produces code states with an effective error rate below 1% in up to 94% of the cases.

It is noteworthy that in most cases, both protocols require some heralding or post-selection, as there is a finite probability to prepare a code state with a large effective error rate. It should also be clear that using phase estimation in this way is much superior to simply post-selecting on obtaining the outcome $x_1 = x_2 = \cdots = x_M = 0$ which occurs with an exponentially small probability in M. If phase estimation is used for quantum error correction, then post-selection is not an option, as we do not wish to throw away the encoded state.

Number of photons

We also consider the mean number of photons and fluctuations thereof for the adaptive and nonadaptive phase estimations of S_p, applied to a squeezed vacuum state with parameter Δ. The expected number of photons in $|\Psi(x[M])\rangle$ (which is only slightly different from the number of photons in $|0(x[M])\rangle$) approximately equals

$$\bar{n}(x[M]) \approx \sinh^2[\ln(\Delta)] + \frac{1}{N}\sum_{j=0}^{M} |(j-M/2)c_j(x[M])\sqrt{2\pi}|^2,$$

(23)

where we again used that $\langle \mathrm{vac_{sq}}| D[\sqrt{2\pi}(t-t')] |\mathrm{vac_{sq}}\rangle \approx 0$ for $|t-t'| \geqslant 1$ and the normalization N has been given above. In Fig. 8, we plot the contribution from squeezing

on the mean photon number. In Fig. 9, we plot the *total photon number* after the measurement of S_p. There, we see $\langle \bar{n} \rangle_\theta = \frac{1}{2\pi} \int d\theta \, P(x[M]|\theta) \bar{n}(x[M])$ with $\bar{n}(x[M])$ as in Eq. (23) as a function of M. We also plot the standard deviation $\sqrt{\langle (\bar{n} - \langle \bar{n} \rangle_\theta)^2 \rangle_\theta}$. Note that this is not the standard deviation $\sigma(n)$ of the state itself which, as we discussed in Sec. I, scales as $\bar{n}$. As the contributions from squeezing and the phase estimation cannot be separated for finite squeezing, we show the photon numbers for $\Delta = 0.2$. Note that both mean value and deviation depend on Δ.

One can observe that $\bar{n} \approx M\pi/2 + \bar{n}_{\mathrm{squeeze}}$: this is what we expect as in each round we symmetrically displace the input state further out to the left and right by an amount $\alpha = \sqrt{\pi/2}$, increasing the expected photon number thus by $\pi/2$. The total expected number of photons in the cavity for $M = 8$ and $\Delta = 0.2$ (8.3 dB) is below 25 such that $\bar{n} \pm \sigma(n)$ is below 50.

F. Usage of adaptive phase estimation in quantum error correction

Adaptive phase estimation of S_p and S_q can in principle be used for quantum error correction. One cycle of quantum error correction will then consist of phase estimation of S_p followed by phase estimation of S_q, each taking a certain number of M rounds. Each time one starts measuring S_p one has prior information about its phase which should be used in choosing the next feedback phases. If the entire QEC protocol is noiseless, then the choice for the feedback phase of the first round of the S_p measurement is simply estimated using the previous feedback phases and outcomes of the previous measurement of S_p, as determined by Eq. (19). It is clear that for such a noiseless protocol, the number of photons thus keeps increasing round by round while the state gets closer and closer to a true eigenstate of S_p (and S_q). In the implementation (see Sec. III), the number of photons is limited and at high number of photons the interactions that are invoked to execute the protocol are no longer accurate. In addition, photon loss is occurring continuously from the cavity during the protocol. This means that for quantum error correction, the choice of feedback phases should ideally take into account an error model of the dynamics during the S_q measurement and how

TABLE I. Ranges of some relevant parameters. The quality factor values (and lifetimes) represent the internal losses inside the cavity while in our protocol the total quality factor of the cavity Q_{tot}, including its intended but flexible coupling to the outside world, is the relevant parameter.

Transmon qubit $\frac{\omega_q}{2\pi}$ and bare cavity frequency $\frac{\omega_r}{2\pi}$	3–11 GHz
Qubit T_1/T_2 time	10–100 μs
3D (storage) cavity lifetime T_{cav}	55 μs [15], 1 ms ($Q > 10^7$) [36]
2D CPW (storage) cavity lifetime T_{cav}	200 μs ($Q > 10^6$) [37]
Controlled-displacement pulse time T_{pulse}	25–100 ns [38]
Dispersive shift $\chi/2\pi$	1–20 MHz
Qubit measurement time t_{meas}	200–300 ns [15,39]
Single-qubit gate	5–10 ns

this dynamics changes the current estimate phase for S_p. We leave such more complete analysis to future work.

III. PROPOSAL FOR REALIZATION IN DISPERSIVE QUBIT-CAVITY SETUP

We consider the following physical setup. A superconducting transmon qubit is capacitively coupled to a two-dimensional (2D) or three-dimensional (3D) microwave high-Q storage cavity, as well as a low-Q readout cavity. The storage cavity will be used to prepare a code state and the readout cavity will be used to measure the state of the qubit. This is the same setup as the cat-state code experiments in, e.g., [16] and we will use values for physical parameters which are similar as in this setup (see Table I). We assume that a particular cavity mode a with resonance frequency ω_r couples most strongly to the transmon qubit and neglect the interactions of the qubit with other cavity modes, as well as the coupling to all higher-energy levels beyond the states $|0\rangle$ and $|1\rangle$.

We assume that the interaction between qubit and storage cavity mode is approximately described by a simple Jaynes-Cummings Hamiltonian $H_{\text{JC}} = -\frac{\omega_q^{\text{bare}}}{2}Z + \omega_r a^\dagger a + g(\sigma_- a^\dagger + \sigma_+ a)$. In the dispersive regime when $\frac{g}{\Delta} \ll 1$ ($\Delta = \omega_q^{\text{bare}} - \omega_r$ is the detuning), one can make a perturbative expansion in $\frac{g}{\Delta}$ and derive an effective Hamiltonian which equals

$$H_{\text{eff}} = (\omega_r - \chi Z)a^\dagger a - \frac{1}{2}\omega_q Z + O\left(\frac{g^4}{\Delta^3}\right), \quad (24)$$

with dispersive shift $\chi \approx \frac{g^2}{\Delta} + O(g^4/\Delta^3)$. Here, $\omega_q = \omega_q^{\text{bare}} + \chi$ where ω_q^{bare} is the bare qubit frequency when the qubit is uncoupled to the cavity. The perturbative expansion is warranted when $\bar{n} < n_{\text{crit}} = \Delta^2/(4g^2)$ where $\bar{n}$ is the mean number of photons in the cavity. The effective Hamiltonian shows that the resonant frequency of the cavity is shifted depending on the state $|0\rangle$ (+) or $|1\rangle$ (−) of the qubit, i.e., its frequency $\omega_r \rightarrow \omega_r^\pm = \omega_r \mp \chi$. In this approximation, we neglect a nonlinear term in the effective Hamiltonian of the form $\frac{5g^4}{3\Delta^3}Z(a^\dagger a)^2 \equiv \chi' Z(a^\dagger a)^2$, a nonlinear dispersive shift.

In a more systematic approach, the qubit-cavity Hamiltonian can be obtained through first determining the normal modes of the coupled LC circuits, after which the nonlinearity due to the Josephson junction is treated as a perturbation [35].

This gives rise to an effective Hamiltonian of the form

$$H_{\text{eff}} = \omega_r a^\dagger a + \tilde{\omega}_q b^\dagger b - \frac{\chi_{rr}}{2}(a^\dagger a)^2 - \frac{\chi_{qq}}{2}(b^\dagger b)^2$$
$$- \chi_{qr} a^\dagger a b^\dagger b, \quad (25)$$

with $\chi_{rr} \equiv 2K = \frac{\chi_{qr}^2}{4\chi_{qq}}$. Thus, both qubit and cavity are represented as coupled nonlinear oscillators such that the anharmonicity for the qubit χ_{qq} is relatively large. When one restricts the b oscillator to the lowest two levels, one can identify $\omega_q \approx \tilde{\omega}_q - \chi_{qq}/2$ and $\chi_{qr} \approx 2\chi$. Higher-order terms in Eq. (25) of the form $(a^\dagger a)^2 b^\dagger b$ would describe the nonlinear dispersive shift. Note that in the model in Eq. (24) the cavity self-Kerr nonlinearity is not present, while its nonzero value has been determined in the experiments. The coupling of the qubit to the readout cavity is described by a similar effective Hamiltonian as in Eq. (24) with a weaker dispersive coupling.

In Table I, we give the ranges of the relevant physical parameters and we choose several specific values in these ranges to demonstrate how well the protocol can be executed. We will take $\chi/2\pi = 2.5$ MHz and let the detuning be $\Delta/2\pi = 1$ GHz in which case $n_{\text{crit}} \approx \frac{\Delta}{4\chi} = 100$. This upper limit on the number of photons is sufficient for creating good code states: for an eight-round protocol $M = 8$, $\bar{n} \approx 25$ (see Fig. 9), so that $\bar{n} \pm \sigma(n) \lesssim 50$ [with $\sigma(n) \sim \bar{n}$], well below this limit. For this choice of parameters $g/2\pi \approx 50$ MHz. The strength of the neglected nonlinear term in Eq. (24) then equals $\chi'/(2\pi) \approx \frac{5g^4}{6\pi\Delta^3} = 10$ kHz. In the experiment described in [16] in which cat states are created with about 55 photons, a value for $\chi'/(2\pi) \approx 4.2$ kHz is estimated for a detuning $\Delta/2\pi = 1.18$ GHz and $\chi/2\pi = 2.4$ MHz which is considerably lower. The strength of the self-Kerr nonlinearity $K(a^\dagger a)^2$ in this experiment is estimated as $K/(2\pi) \approx 3.61$ kHz. The authors estimate $n_{\text{crit}} \lesssim 300$ which is a bit more optimistic than the estimate above.

One round of phase estimation will last a total time T_{round}. During this time, a quantum circuit depicted in Fig. 5 is to be executed. The protocol of a round consists of a short [$O(10)$ ns] interval T_{prep} in which qubit (and storage cavity) are prepared. For the qubit this means it is put in the state $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ [by a Hadamard of $R_x(\pi/2)$ gate]. After T_{prep}, qubit and storage cavity should be coupled by a controlled-displacement transformation or a $D(Z\sqrt{\pi}/2)$ gate (in phase estimation for S_p). For this gate there are two options, as has been discussed in [16,38].

In some physical setups, the dispersive cavity-qubit coupling (both storage and readout cavity) is not tunable and is thus always "on." Such a setup is nonideal in various ways. When the dispersive coupling is always on, it means that one should prepare the code states in a rotating frame (not the laboratory frame) which depends on the qubit state. For example, one chooses the frame in which the qubit is in the state $|0\rangle$ as the frame in which the cavity state should be unchanging, stationary (and we also look at the qubit in its rotating frame as gates on the qubit are done relative to that). In this frame, the effective Hamiltonian in Eq. (24) equals $\tilde{H}_{\text{eff}} = 2\chi |1\rangle \langle 1| a^\dagger a$. It is important that during the qubit measurement, which takes up a considerable amount of time, this effective Hamiltonian induces no further rotations on the partially prepared code state (or if it induces rotations, one should know what they are). This means that during qubit measurement the qubit has to quickly be reset to $|0\rangle$, in order to induce no further rotational dynamics on the cavity state.

Another disadvantage of using a nontunable χ is that the accuracy of single-qubit rotations depends on the number of photons in the cavity. The qubit frequency given a storage cavity with n photons is, by Eq. (24), given by $\omega_q + 2\chi n$. A microwave pulse which should rotate the qubit independent of the number of photons in the cavity should thus qualitatively take at least time $T_{\text{pulse}} = 2/W$ [40] with frequency width $W \geqslant 2\chi\sigma(n)$. This assumes that one sets the center frequency of the pulse at $\omega_q + 2\chi\bar{n}$. Here, $\sigma(n)$ is the standard deviation in a code state, given in Eq. (6), which scales with $\bar{n}$. In [16] it was argued that the unwanted entangling of qubit and cavity due to single-qubit rotations is a leading source of inaccuracies when one goes to higher photon numbers. Even though the number of photons in our proposed protocol for $M = 8$ will never be larger than 50, the protocol consists of many more single-qubit gates than the experiments in [16] so that these errors will accumulate. With a nontunable χ, the dispersive coupling to the readout cavity is of course continuously on during the entire round, while in the figures it is suggested that measurement only occurs at the end of the protocol. It is understood that the coupling between transmon qubit and readout cavity mode is smaller than the coupling between transmon qubit and storage cavity, for example, in [15] the dispersive coupling $\chi/2\pi$ to the readout cavity was estimated as 0.930 MHz. At the same time, this coupling needs to be sufficiently strong to provide a relatively short measurement time for the qubit as T_{round} will be largely determined by the length of the controlled-displacement transformation and the qubit measurement time.

A third disadvantage of a nontunable χ is the the cavity Kerr nonlinearity which is present in Eq. (25) due to the linear coupling between the LC oscillators: the "cavity mode" is in fact a "dressed" cavity mode which sees the Josephson nonlinearity. If χ is turned to a small value, then this Kerr nonlinearity will be correspondingly small.

For these reasons, we imagine that the dispersive coupling χ is tunable and can be turned "on and off" during the execution of a round. Thus, during each round, the dispersive storage cavity-qubit coupling $\chi(t)$ is "off" (i.e., to a low value) during qubit preparation and single-qubit rotation (lasting time T_{prep}) and possibly unconditional cavity displacements [of the form $D(-\sqrt{\pi/2})$]. Then, χ is rapidly turned on after T_{prep} and stays on for a time interval T during which the controlled-displacement gate in Fig. 5 acts. The coupling χ is again rapidly turned off during the last interval T_{post} during which the qubit state undergoes further single-qubit rotations and is being measured. One can similarly imagine that only during the measurement time the coupling to the readout cavity is turned on. The turning off and on of χ could be achieved in two possible ways. One can increase the cavity-qubit detuning Δ (and hence reduce χ) by altering the qubit resonant frequency ω_q^{bare} using a flux-tunable transmon qubit; such switching can take place in $O(1)$ ns. An alternative is to have a variable qubit-cavity capacitive coupling [41,42] which is turned on and off in a $O(1)$ ns time window.

Let us now discuss the realization of the controlled-displacement gate. Note that if χ is in principle turned "off," unless the cavity state needs to be manipulated, it means that we are preparing the code states in the rotating frame of the cavity at frequency ω_r (while single-qubit gates are performed in the frame rotating at ω_q^{bare}). When the coupling is turned on, the effective Hamiltonian in these rotating frames is then $\tilde{H}_{\text{eff}} = -\chi Z a^\dagger a - \frac{1}{2}\chi Z$ (neglecting nonlinearities). The additional Z rotation in this Hamiltonian leads to a phase accumulation on the prepared qubit state $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ which should be taken into account when considering what single-qubit Z rotation of the form $\text{diag}[1, \exp(i\varphi)]$ is done during T_{post}.

In principle, one can enact a controlled-displacement gate by supplementing this dynamics by driving the cavity with a microwave pulse. During the time interval T, a microwave drive of duration $T_{\text{pulse}} \approx T$ is applied. The drive tone of this pulse $\omega_d = \omega_r + \chi$ is resonant with the cavity mode when the qubit is in the state $|1\rangle$ but off-resonant when the qubit is in state $|0\rangle$. Hence, one expects a large cavity displacement in the resonant case and a small negligible displacement when the drive tone is off resonant, thus enacting a controlled-displacement gate. Again, at an intuitive level, the frequency width W of the pulse should be sufficiently narrow as compared to 2χ, which is the difference in resonance frequencies of the oscillator given the qubit state, so that the pulse has few photons at frequency $\omega_r - \chi$. If we center the pulse at $\omega_r + \chi$, then for a Gaussian pulse with $T_{\text{pulse}} = 2/W$, one should at least require that $2\chi > \frac{W}{2} = 1/T_{\text{pulse}}$ or $\chi T_{\text{pulse}} > \frac{1}{2}$. This heuristic argument assumes that the cavity decay rate κ which sets the width of the resonance obeys $\kappa \ll W$ which is warranted for high-Q storage cavities.

During this pulse, the qubit-cavity coupling $-\chi Z a^\dagger a$ also induces a qubit-state-dependent phase space rotation on the cavity state, i.e., of the form $\exp(i\chi Z a^\dagger a t)$, which is in fact *undesired* (see Fig. 14 in Appendix D). It is thus important to choose the interaction time T such that this difference in rotation angle is 2π which implies that the newly created superposition of displaced squeezed states stays on a line. This warrants a precise choice for T, namely, $\chi T = \pi$. In the analysis in Appendix D we model the drive and the corresponding displacements and rotations assuming a simple square pulse (with zero rise time). The conclusion of this analysis is much less negative than what has been stated in [38] where it is written that a direct controlled-displacement gate requires a time $T_{\text{pulse}} \approx 30/\chi$ which would be quite long. The alternative is to do a controlled-displacement gate
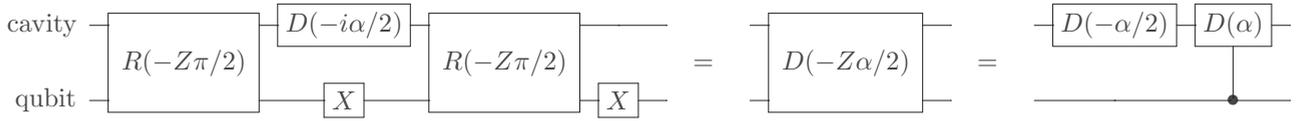
FIG. 10. Realization of a controlled-displacement gate via two controlled rotations for one round of phase estimation of S_p with $\alpha = \sqrt{2\pi}$ or S_q with $\alpha = -i\sqrt{2\pi}$ (see Fig. 5). Here, $R(-Z\pi/2) = \exp(ia^\dagger a Z\pi/2)$.

through a sequence of two controlled rotations interspersed with an unconditional displacement [43], explicitly shown in Fig. 10.

Let $R(\theta) = \exp(-i\theta a^\dagger a)$ be a rotation in phase space. The conditional rotation $R(-Z\pi/2) = \exp(ia^\dagger a Z\pi/2)$ is obtained by $\tilde{H}_{\text{eff}}$, i.e., coupling the transmon qubit for time t with $t\chi = \pi/2$ to the cavity mode. The inverse rotation equals $R(Z\pi/2) = X R(-Z\pi/2) X$ where X is a Pauli X (π rotation) on the qubit. The advantage of this way of doing a controlled displacement is that its implementation is always fast, as it requires $O(10)$ ns unconditional displacements and single-qubit rotations as well as a total qubit-cavity interaction time of $T = \pi/\chi$ during which the cavity state is rotated depending on the qubit state. For our chosen χ, one has $T = \pi/\chi = 200$ ns. Note that the last X gate on the qubit can be absorbed in the post-processing stage during which one thus applies a single-qubit rotation which is composed of Pauli X in Fig. 10, the possible feedback rotation diag$[1, \exp(i\varphi)]$, including corrections for accumulated phases, and then a Hadamard gate, shown in Fig. 5. However, in order for this implementation to work, one needs to turn off the dispersive coupling χ in the middle period during which an unconditional displacement and a single-qubit rotation act. Thus, this scheme would require one to turn χ on (and off) twice during a round, while a direct conditional displacement would allow one to turn χ on and off once.

In either implementation of the controlled-displacement gate, T_{round} will be largely determined by the duration of the controlled-displacement gate and the qubit measurement time. If we assume that the qubit measurement time is approximately 300 ns, we have $T_{\text{round}} \approx 500$ ns. During this time qubit decoherence (see Table I) is negligible as $T_{\text{round}}/T_1(= 50 \, \mu\text{s}) = 1 \times 10^{-2}$. If we assume a 3D storage cavity lifetime of 1 ms (see Table I), then during the execution of $M = 8$ rounds for the full phase estimation of S_p, in total lasting $T_{\text{PE}} = 4 \, \mu\text{s}$, one has $T_{\text{PE}}/T_{\text{cav}} = 4 \times 10^{-3}$. One cycle of quantum error correction using the measurement of S_p and S_q will thus last $T_{\text{QEC-cycle}} = 8 \, \mu\text{s}$, still considerably faster than the cavity decay time.

For 3D cavities, no flux-tunable transmon qubits have been used so far. For a 2D coplanar waveguide cavity, to which transmon qubits have been coupled with a tunable χ, the best lifetime can be 200 μs (see Table I) which would give estimates which are factor of $\frac{1}{5}$ worse as compared to a 3D cavity.

The adaptive phase estimation protocol uses feedback which implies that the phase in a single-qubit rotation in the next 500 ns round depends on the outcome of the qubit measurement in the previous round. In current technology, it is possible for a qubit measurement outcome to determine the execution of a single-qubit gate 200 ns later, so this does not pose a problem in the implementation.

A. Shift error and noise during the protocol

There are two sources of errors during a round of the protocol, namely, errors on the transmon qubit and direct errors on the cavity mode. We first consider errors on the transmon qubit which can propagate to the cavity mode in two ways.

We can call a protocol strictly fault tolerant if an error with low probability (or amplitude) on the qubit can induce a shift error u or v with low probability and amplitude *and* low strength $|u|, |v| \ll \sqrt{\pi}/2$ on the cavity mode.

An error on the qubit during a round can alter the measurement outcome of this qubit (for example, an error which flips the measurement outcomes). This will lead to an incorrect estimate for $\tilde{\theta}$. However, if the rate at which these errors take place is sufficiently low, so most rounds are error free, then it should be clear that the error on the estimate is also small, hence the induced shift error is small, implying fault tolerance. One can see this as fault tolerance which arises from the repetition of the rounds; how well this works will depend on whether the rate of qubit errors is sufficiently low.

Another type of error on the qubit can directly result in a shift error on the cavity mode. For example, during the execution of the direct controlled displacement or rotation, the qubit state decays to $|0\rangle$, resulting in a displacement error which can be as large as $D(\sqrt{2\pi})$. Such error shows that the phase estimation protocol is not strictly fault tolerant since, even though this error is very unlikely, it does lead to a large incorrectable shift error. Using several, say k, ancilla qubits per round and putting them into a cat state $|00\ldots0\rangle + |11\ldots1\rangle$ so that each one of the k qubits is only used to perform a much smaller controlled displacement $D(\sqrt{2\pi}/k)$ can mitigate this problem and make the protocol more fault tolerant. Effectively, we are trading a low probability for a large shift error to a larger probability (as one uses more qubits) to have only small correctable shift errors.

Given the fact that the accuracy of single-qubit gates and controlled rotations is quite high and qubit decoherence on the time scale of a round very low, we do not anticipate that qubit errors are the dominant source of errors.

Let us next consider direct errors on the cavity mode which can result from photon loss, a self-Kerr nonlinearity, or a nonlinear dispersive shift. In principle, for short enough time intervals all these processes can be expanded as linear combinations of shift errors, and these shift errors can be propagated through the ideal circuit in a round, or in-between rounds, such that they remain shift errors which gradually add up in strength. This is true as the ideal circuit only contains (conditional) displacements and (conditional) rotations.

Given the points made in Sec. I D, one can consider the effect of cavity decay during preparation and the presence of nonlinearities which we expect to be the dominant source of shift errors. For cavity decay with rate κ with

$\kappa = 1/T_{\text{cav}}$ of a 3D cavity, one can easily meet the condition $\delta \equiv \kappa T_{\text{PE}} \approx 4 \times 10^{-3} \ll \frac{1}{n_{\text{max}}}$ when n_{max} created during the eight-round protocol is no more than 50 photons. The arguments in Sec. I D then imply that cavity decay modeled by some Lindblad equation induces dimensionless shift errors u,v of strength $|u|,|v| \leqslant O(\sqrt{\kappa T_{\text{PE}}}) \approx 0.06$ [using only the first term in the shift expansion in Eq. (13)] during the phase estimation measurement of S_p.

We can also consider the effect of the self-Kerr nonlinearity in Eq. (25) or the nonlinear dispersive shift $\chi' Z(a^\dagger a)^2$ where the strength of $\chi'/(2\pi)$ and $K/(2\pi)$ can be taken to be, say, 4 kHz. If we represent the overall nonlinearity during the protocol as $\exp[-i\epsilon(a^\dagger a)^2]$ one has $|\epsilon| = K T_{\text{PE}} = 2\pi \times 16 \times 10^{-3} \approx 0.1$ which is two orders of magnitude *larger* than $\frac{1}{n_{\text{max}}^2}$ for $n_{\text{max}} = 25$. This means that the shift errors due to the nonlinearity will add up throughout the protocol and do not necessarily remain small correctable shift errors.

Even though the physical values of these numbers are not clad in stone, there is a simple argument which demonstrates that higher-order terms in a perturbatively derived dispersive coupling Hamiltonian will lead to errors which require a more thorough analysis. This argument comes about from the fact that the running time of the protocol, let us say T_{round}, is determined by $\chi \sim g^2/\Delta$, i.e., $T_{\text{round}} \sim 1/\chi$. The next order in the perturbative coupling between cavity and qubit scales as $\frac{g^4}{\Delta^3}$ and we want that its overall effect during T_{round} to be sufficiently small, but still have a sizable number of photons in the cavity. We consider

$$|\epsilon| n_{\text{max}}^2 \sim \frac{g^4}{\Delta^3} T_{\text{round}} n_{\text{max}}^2 \sim \frac{g^2}{\Delta^2} n_{\text{max}}^2 = \frac{n_{\text{max}}^2}{4 n_{\text{crit}}} \ll 1, \quad (26)$$

while n_{max} should be 10–100 photons in order to obtain a good code state. One has $n_{\text{max}} < n_{\text{crit}}$ but the stringent condition in Eq. (26) would require making n_{crit} much larger. Since n_{crit} should be enlarged without making the protocol last much longer (as this would enhance the strength of other sources of errors), it would be best to work at a larger detuning Δ rather than a smaller capacitive coupling between transmon qubit and cavity (coupling g).

One should note that this problem rapidly gets better at higher-order terms in perturbation theory. Assume that the first unwanted term comes in at kth order in perturbation theory, i.e., with strength $g(\frac{g}{\Delta})^{k-1}$, so that $T_{\text{round}} g(\frac{g}{\Delta})^{k-1} n_{\text{max}}^{k/2} \ll 1$. One then obtains the condition

$$4 n_{\text{crit}} \left(\frac{n_{\text{max}}}{4 n_{\text{crit}}} \right)^{k/2} \ll 1. \quad (27)$$

For a quartic term $k = 4$ this condition is quite demanding, but for the next higher-order term $k = 6$ and $n_{\text{max}} \ll 4 n_{\text{crit}}$ the condition is much more mild.

These arguments suggest that it is better to not treat the nonlinear dynamics as a source of errors, but rather treat it as a source of known systematic errors. One approach is to try to actively cancel their effect during the evolution so as to obtain the same code states [44]. Another approach is to take as a given that these interactions exist and seek a code formalism that captures their effect. Assuming that χ is tunable, one may at least hope to reduce the nonlinear dispersive shift and

the self-Kerr nonlinearity in strength during the time that the dispersive interaction is off.

These arguments show that further numerical analysis including the nonlinearity and open system dynamics modeled by a Lindblad equation are warranted to assess their overall effect during the phase estimation protocol.

IV. DISCUSSION

In order to implement phase estimation in a shorter amount of time, it is possible to couple several transmon qubits simultaneously to a 2D (or possibly 3D) cavity. It may be interesting to consider whether a form of Shor quantum error correction (using cat-state ancillas) instead of Steane quantum error correction is possible for the GKP code states. It will be worthwhile to numerically study the code state preparation protocol including the effect of nonlinearities and open system dynamics.

A somewhat different scheme for preparing code states, eigenstates of S_p and S_q, involves two oscillators as follows, borrowing an idea in [45]. The basic idea is that by (sequentially) coupling a qubit to the oscillator to gather phase information, we are getting little information, at most one bit, per qubit used. At the same time, this transmon qubit is effectively realized as a nonlinear oscillator as in Eq. (25). So, can one not get information faster by coupling the storage cavity to another oscillator?

Assume that we can prepare one ancilla oscillator in the state $|q_a = 0\rangle$ where q_a is its position quadrature (or some squeezed version thereof) and the other "storage" oscillator is in some initial state $|\psi_s\rangle = \int dp\, \alpha_p |p\rangle$. Assume an interaction between the oscillators of the form $H_{\text{int}} = p_s \otimes p_a$ for some time $t = 2\sqrt{\pi}$ so that $e^{-i H_{\text{int}} t} |\psi_s\rangle \otimes |q_a = 0\rangle = \int dp\, \alpha_p |p\rangle |q_a = 2p\sqrt{\pi}\rangle$. If we can measure the q quadrature of the ancilla oscillator modulo 2π, i.e., we determine $q_{\text{meas}} = q_a \mod 2\pi$, we project the system oscillator into a superposition of states with $p = \frac{q_{\text{meas}}}{2\sqrt{\pi}} \mod \sqrt{\pi}$. Shifting back this projected state depending on q_{meas} gives the output state $\sum_{k=-\infty}^{\infty} \alpha_p |p = k\sqrt{\pi}\rangle$, thus preparing an eigenstate of S_p. A similar procedure for S_q on this state would then prepare the full code state. The linear coupling between the two oscillators is very natural; superconducting LC circuits with capacitive coupling realize such interaction. However, measuring a quadrature modulo 2π is not simple: one has to make sure not to get too much information. If the ancilla oscillator were modified to include a Josephson junction, it would make one of the quadratures, namely, the phase variable, 2π periodic and thus suitable. A rapid turning on of the Josephson junction, and thus a rapid change of the quadratic potential in q or ϕ to a periodic cosine potential $U(\phi) = -E_J \cos(\phi)$ could freeze this state (taking $E_J \gg E_C$, potential energy much larger than kinetic energy) and information could be read out. It is an open question whether a rapid turning-on of a Josephson junction (see, e.g., [10] for a similar circuit switch idea) is experimentally feasible and what the details of such a scheme would look like.

APPENDIX A: USE OF PHASE OR DISPLACEMENT FRAME AND LOGICAL GATES

We prepare a code state or perform error correction through phase estimation of the displacement operators S_p and S_q. The phase estimation protocol outputs an estimate for the eigenvalue $e^{i\theta_p}$ ($e^{i\theta_q}$) of S_p (S_q), while simultaneously projecting the input state onto an approximate eigenstate $|\psi_{\theta_p,\theta_q}\rangle$ of S_p and S_q. Error correction would then correspond to displacing this oscillator state by the corrective displacement $D_{\theta_p,\theta_q} = e^{iq\theta_p/(2\sqrt{\pi})}e^{ip\theta_q/(2\sqrt{\pi})}$ such that $D_{\theta_p,\theta_q}|\psi_{\theta_p,\theta_q}\rangle \approx |\psi_{\theta_p=0,\theta_q=0}\rangle$, an approximate code state.

Here, we show that it is not necessary to do these additional displacement operations, but one can work with a phase frame similar as the Pauli frame for stabilizer codes [5,46]. We will use the notation $\overline{X} = X$ and $\overline{U}$ for operators which have the action of Pauli X and a unitary gate U on the states in the code space. We call such operators logical operators or logical gates. In addition, we write $\overline{X}_{\theta_p}$ for the Pauli X operator on a code space labeled by a phase θ_p for S_p.

Let us first assume that we work with perfect phase estimation and perfect code states. The operator $e^{-i\sqrt{\pi}p}$ has been called $\overline{X}$ because it maps the perfect state $|0\rangle$ onto $|1\rangle$ and vice versa. We can note that $\overline{X}$ is not Hermitian, we have in fact $\overline{X} = \overline{X}^\dagger S_p$. The action of $\overline{X}$ and $\overline{X}^\dagger$ on the perfect code space is, however, identical, thus on the code space $\overline{X} \equiv \overline{X}^\dagger$ so that $\overline{X}^2 = I$. If we use a code space labeled by the eigenvalue $e^{i\theta_p}$ of S_p, we have $\overline{X}^2 = e^{i\theta_p}I$. This means that we need to redefine the logical X operator on the code space characterized by eigenvalue $e^{i\theta_p}$ as $\overline{X}_{\theta_p} = D_{\theta_p,\theta_q}^\dagger \overline{X} D_{\theta_p,\theta_q} = \overline{X}e^{-i\theta_p/2}$ for which $\overline{X}_\theta^2 = I$ on the θ_p code space. Similarly, $\overline{Z}_{\theta_q} = D_{\theta_p,\theta_q}^\dagger \overline{Z} D_{\theta_p,\theta_q}$. Clearly, $\overline{X}_{\theta_p}$ and $\overline{Z}_{\theta_q}$ only differ from $\overline{X}$ and $\overline{Z}$ by phases, so transform similarly.

It is known what operations are necessary to perform logical gates on the code states in the $\theta_p = 0$, $\theta_q = 0$ code space, such as the Clifford gates (the CNOT gate, the Hadamard gate, the S gate) and the T gate (see [4]), as one can verify their proper action on the logical operators $\overline{X}$ and $\overline{Z}$.

The logical Clifford group gates can be realized by linear optical transformations, i.e., linear transformations on the set of positions and momenta of n oscillators $(q_1,p_1,\ldots,q_n,p_n)$, which preserve their commutation relations. A circuit for the CNOT gate is shown in Fig. 12. The Hadamard gate represents a $\pi/2$ phase delay enacting $q \to p$ and $p \to -q$. The $S = \mathrm{diag}(1,i)$ gate enacts the transformation $q \to q$ and $p \to p - q$.

Under the action of these gates, shift errors (displacements) remain shift errors, i.e., a general displacement operator on these n oscillators of the form $D = \exp[i\sum_{i=1}^n(\alpha_i p_i + \beta_i q_i)]$ (with real α_i,β_i) transforms to $D' = UDU^\dagger$ where D' has coefficients $\{\alpha_i',\beta_i'\}$. One can propagate the shift errors through a Clifford circuit: small shift errors in several modes can add up to large shifts in one mode, similar as Pauli errors can propagate to become high-weight incorrectable Pauli errors. Furthermore, even though the Clifford gates do not amplify shift errors, shift errors inside, say, the realization of a CNOT gate can get somewhat deamplified or amplified in strength as

the circuit uses squeezing and squeezing acts, e.g., as $p \to e^r p, q \to e^{-r} q$.

For logical Clifford gates $\overline{U}$ one can verify their action on the θ_p,θ_q code space. One has

$$\overline{U}|\psi_{\theta_p,\theta_q}\rangle = \overline{U}D_{\theta_p,\theta_q}^\dagger \overline{U}^\dagger \overline{U}|\psi_{0,0}\rangle = \tilde{D}_{\theta_p,\theta_q}\overline{U}|\psi_{0,0}\rangle, \quad \text{(A1)}$$

where $\tilde{D}_{\theta_p,\theta_p}$ is the new shift correction (displacement). This means that one never needs to do the corrective displacement, but can just keep track of the phase frame in software, just as for the Pauli frame in a computation with only Clifford gates.

When the encoding is only approximate (and we know the eigenvalues of S_p and S_q approximately), every code state is in principle given by a perfect code state and a distribution of shift errors, i.e., a density matrix expanded in the shift error basis as in Eq. (9). The gates are still defined by their action on the perfect code states. The distribution of shift errors or the approximate encoding may be slightly different for every code state when these are generated, say, by the protocol in this paper. It is important to note that this does affect the proper functioning of the Clifford group gates as these gates only propagate the shift errors between logical qubits.

For quantum universality, one needs a $T = \mathrm{diag}[1, \exp(i\pi/4)]$ gate. This gate can in principle be realized using an ancilla $T|+\rangle$ or alternatively the ancilla $|H = 1\rangle \propto SHT|+\rangle$ which is a $+1$ eigenstate of the Hadamard gate, in addition to Clifford group gates (see, e.g., [5]). The logical Hadamard gate $\overline{H} : \overline{X} \leftrightarrow \overline{Z}$ equals $\exp(i\frac{\pi}{2}a^\dagger a)$ which means that a $+1$ eigenstate of this operator has a photon number which is 0 mod 4. One can create the state $|\overline{H = 1}\rangle$ by doing phase estimation for S_p and S_q on the vacuum state $|\mathrm{vac}\rangle$ for which $e^{i\frac{\pi}{2}a^\dagger a}|\mathrm{vac}\rangle = |\mathrm{vac}\rangle$ and post-selecting on the outcomes $S_p \approx 1$ and $S_q \approx 1$. The resulting state is encoded and a $+1$ eigenstate of $\overline{H}$ as $\overline{H}$ commutes with S_p and S_q. One can imagine that noisy ancillas $|\overline{H = 1}\rangle$ are further distilled into fewer higher-quality ancillas before being used. It is important to note that a $+1$ eigenstate of $\overline{H}$ for the phase frame $\theta_q = \theta_p = 0$ is not related by a displacement to a $+1$ eigenstate of $\overline{H}$ in an arbitrary phase frame θ_q,θ_p. So, before one uses such ancillas one has to at least know or physically fix the phase frame that is in use. We leave a realistic method to implement T (and Clifford) gates on the approximate code states in circuit-QED hardware to future work.

APPENDIX B: QUANTUM ERROR CORRECTION USING ENCODED ANCILLAS

In this section, we briefly review a form of quantum error correction suggested in [4] and a simplified implementation in [11], which, for stabilizer codes, is called Steane error correction. In Steane error correction, the errors are corrected by coupling the encoded data to an encoded ancilla using a CNOT gate and measuring the encoded ancilla. The advantage of using this procedure for quantum error correction is that the quantum circuit is fault tolerant, made from linear optical elements, phase shifters, and squeezers and is fully deterministic. Fault tolerance means that shift errors are not amplified in strength by the circuit (but they do propagate) as it involves only linear optical components. A disadvantage of
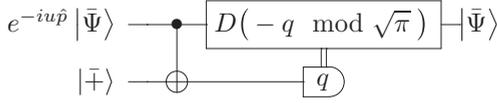
FIG. 11. Steane error correction for shifts in q generated by the error shift operator e^{-iup} which are detected by the stabilizer check S_q. The measurement of the ancilla in the encoded $|\bar{+}\rangle$ state is a homodyne measurement of q. The displacement $D(-q \mod \sqrt{\pi})$ is a correction which is like a Pauli frame for a stabilizer code and does not physically need to be realized. Note that the modulo function should be taken in the interval $(-\sqrt{\pi}/2, \sqrt{\pi}/2]$. In the main text we have referred to $|\bar{+}\rangle$ as $|+\rangle$.

the method is that it requires an ancilla code state which has to be first prepared with sufficient accuracy.

The basic circuit for correcting shifts in q (detected by S_q) is shown in Fig. 11. A similar circuit holds for S_p where the encoded ancilla is prepared in $|\bar{0}\rangle$ (denoted as $|0\rangle$ in the main text). The CNOT gate (a linear optical gate) has the following effect on the two quadratures of control (c) and target (t) mode: $q_c \rightarrow q_c, q_t \rightarrow q_c + q_t, p_c \rightarrow p_c - p_t, p_t \rightarrow p_t$. The CNOT gate cannot be implemented with only beam splitters and phase shifters as the symplectic transformation matrix of the CNOT, i.e., the matrix applied to the vector (q_c, p_c, q_t, p_t) is not orthogonal, while it is always orthogonal for any passive linear optics transformation (see, e.g., [47]). The CNOT gate can be realized by beam splitters and squeezing [48]. The optimal CNOT circuit expressed in terms of these elementary components is depicted in Fig. 12.

In the Glancy and Knill method in [11] the CNOT gate which uses two beam-splitters and two squeezers is replaced by a single beam splitter and a single (but stronger) squeezer. In addition, the simple corrective displacement is replaced by a displacement by a more complicated function $f_{GK}(q)$ of the measurement result, which is given explicitly in [11]. The $\sqrt{\pi}/6$ shift error bound is arrived at by arguing how shift input errors on data and ancilla propagate through the quantum error correction circuit leading to shift errors which have to be corrected in the next round.

Basically, a shift error $\exp(-iup)\exp(-ivq)$ that needs to be properly corrected in steady state (which requires $|u|, |v| < \sqrt{\pi}/2$) results from the propagation of three separate shift errors $u = u_I + u_p + u_q, v = v_I + v_p + v_q$, which are the initial errors u_I, v_I and the errors that occur somewhere
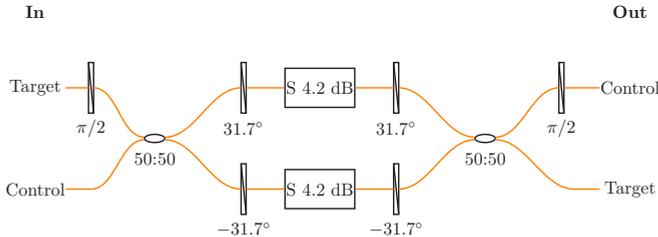
FIG. 12. A CNOT gate on two bosonic modes, target and control. S denotes squeezing and its strength in dB [see squeezing convention below Eq. (4)]. The phase shifters shown explicitly can of course be absorbed in the action of the squeezers or the 50:50 beam splitters.
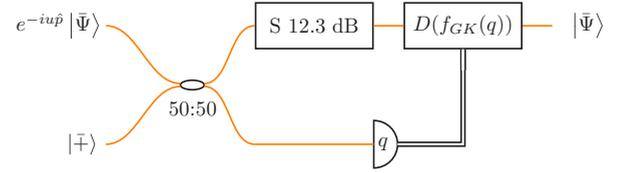
FIG. 13. Steane quantum error correction using the method of Glancy and Knill for shifts e^{-iup} which are detected by the stabilizer check S_q. The circuit only uses a single beam splitter, a homodyne measurement of q, and much stronger squeezing followed by a different corrective displacement $D(f_{GK}(q))$ which is explicitly given in [11].

in the quantum error correction circuits for S_p (u_p, v_p) and S_q (u_q, v_q), hence, all $|u_i|, |v_i| < \sqrt{\pi}/6$. Thus, if all initial shift errors on encoded data and ancilla state are of strength at most $|u|, |v| < \sqrt{\pi}/6$, errors remain correctable through the repeated application of these circuits, assuming that the linear optical circuits themselves are faultless. If the linear optical circuits are faulty, one has to propagate the shift errors that occur during the circuit forward, which will result in a somewhat lower threshold value.

APPENDIX C: CHOICE OF FEEDBACK PHASES

Let us assume that the feedback phases $\varphi_1, \ldots, \varphi_{m-1}$ have been fixed and the bit string $x[m-1] \equiv \{x_1 \ldots x_{m-1}\}$ has been generated. How do we choose the next phase φ_m in an optimal way? Using conditional probability distributions [i.e., the identities $P(A|BC) = P(C|AB)P(A|B)/P(C|B)$ and $P(C|AB) = P(CB|A)/P(B|A)$], the probability distribution for θ after obtaining the measurement result $x[m] = \{x_m, x[m-1]\}$ equals

$$P_\varphi(\theta|x[m]) = \frac{P(\theta|x[m-1])P_{\varphi_m}(x_m|\theta, x[m-1])}{P_\varphi(x_m|x[m-1])}$$

$$= \frac{P(\theta|x[m-1])P_\varphi(x[m]|\theta)}{P_\varphi(x_m|x[m-1])P(x[m-1]|\theta)}. \quad \text{(C1)}$$

Here, we have explicitly kept the dependence of probability distributions on the phase φ which is chosen for the mth qubit measurement as this is the quantity that one wants to optimize over. This phase φ is chosen as the one which maximizes the sharpness of the *a posteriori* probability distribution $P_\varphi(\theta|x[m])$ averaged over possible outcomes x_m (given the previous outcomes and choices for phases), or using Eq. (18),

$$\varphi_m = \arg\max_\varphi \sum_{x_m=0,1} P_\varphi(x_m|x[m-1])S[P_\varphi(\theta|x[m])]$$

$$= \arg\max_\varphi \sum_{x_m=0,1} \left| \int d\theta \, e^{i\theta} \frac{P(\theta|x[m-1])P_\varphi(x[m]|\theta)}{P(x[m-1]|\theta)} \right|. \quad \text{(C2)}$$

The expression for $P_\varphi(x[m]|\theta)$ is simple as the measurement results of each round are independent, i.e.,

$$P_{\varphi=\varphi_m}(x[m]|\theta) = \prod_{i=1}^{m} P_{\varphi_i}(x_i|\theta) = \prod_{i=1}^{m} \cos^2\left(\frac{\theta + \varphi_i}{2} + x_i\frac{\pi}{2}\right). \quad \text{(C3)}$$

If we assume no prior knowledge on θ one can show that $P(\theta|x[m-1]) \propto P(x[m-1]|\theta)$ where the proportionality constant c is independent of θ and φ_m (but may depend on previous phases φ_{m-1}, etc.). The argument is as follows. For the first measurement we write $P_{\varphi_1}(\theta|x_1) = P_{\varphi_1}(x_1|\theta)P(\theta)/P(x_1) \propto P_{\varphi_1}(x_1|\theta)$ if we assume that $P(\theta)$ is a flat distribution (we have no prior knowledge about θ) and $P(x_1) = \frac{1}{2}$. Using Eq. (C1) we see that this then holds for for all measurement outcomes by induction, i.e., $P(\theta|x[m-1]) \propto P_{\varphi_m}(x[m-1]|\theta)/P_{\varphi_{m-1}}(x_{m-1}|x[m-2])$.

This implies that the optimization in Eq. (C2) is equivalent to

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

which can also be rewritten as

$$\varphi_m = \arg\max_\varphi \sum_{x_m=0,1} \left| \int d\theta \, e^{i\theta} \cos^2\left(\frac{\theta+\varphi}{2} + x_m \frac{\pi}{2}\right) \prod_{j=1}^{m-1} \right.$$
$$\left. \times \cos^2\left(\frac{\theta+\varphi_i}{2} + x_i \frac{\pi}{2}\right) \right|. \tag{C4}$$

APPENDIX D: DIRECT CONTROLLED DISPLACEMENT

We model the effect of applying a microwave drive to the cavity mode without including cavity decay. The reason for not including cavity decay is that the parameters of the protocol should be chosen to work without cavity decay such that the loss of photons from the cavity acts as a source of (shift) errors during this ideal protocol.

The drive field is represented as a term $H_{\text{drive}}(t) = \lambda(t)(a + a^\dagger)$ where $\lambda(t) = \Omega_x(t)\cos(\omega_d t) + \Omega_y(t)\sin(\omega_d t)$ with ω_d the drive frequency and $\Omega_x(t), \Omega_y(t)$ the envelope of the pulse. The drive frequency ω_d is chosen as $\omega_d = \omega_r + \chi$. The total Hamiltonian $H_{\text{tot}}(t) = H_{\text{eff}} + H_{\text{drive}}(t)$ (in the laboratory frame) with H_{eff} in Eq. (24) equals

$$H_{\text{tot}}(t) = [\omega_r I - \chi(t)Z]a^\dagger a + \lambda(t)(a + a^\dagger) - \frac{\omega_q Z}{2}. \tag{D1}$$

In order to analyze the effect of the dynamics due to $H_{\text{tot}}(t)$, we use the following useful tool (fact). For a Hamiltonian of the form $H(t) = \omega a^\dagger a + \lambda(t)(a + a^\dagger)$, the unitary time evolution can be written as a product of an overall rotation, an overall displacement, and a phase shift, i.e.,

$$\mathcal{T}e^{-i\int_0^T dt' H(t')} = R(\omega T)D(\gamma)\exp(i\Psi), \tag{D2}$$

with $\gamma = -i\int_0^T dt'\lambda(t')e^{-i\omega t'}$ and $\Psi = \int_0^T dt \int_t^T dt'\lambda(t)\lambda(t')\sin[\omega(t'-t)]$. We have $R(\omega T) = \exp(-i\omega T a^\dagger a)$. This equality can be derived using the
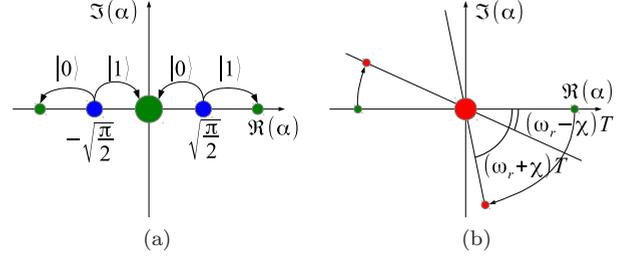
FIG. 14. Phase space sketches: the size of the blobs represents amplitude not quadrature uncertainty. (a) A cat state $|\alpha\rangle + |-\alpha\rangle$ with $\alpha = \sqrt{\pi/2}$ (blue dots) gets displaced by $\pm\sqrt{\pi/2}$ depending on a qubit state $|0\rangle$ or $|1\rangle$ along the $\text{Re}(\alpha)$ line producing the green dots. Assuming constructive interference of phases, the amplitude for $|\text{vac}\rangle$ is double that of the outlying states. Repetition of this scheme with constructive interference produces a binomial distribution of coherent states according to a Pascal triangle. In the phase estimation protocol, the input state is not a cat state but the squeezed vacuum [along the $\text{Re}(\alpha)$ direction] and the protocol generates superpositions of displaced squeezed vacua on a line. In (b) it is shown how an additional phase space rotation which is different for $|0\rangle$ as for $|1\rangle$ produces a superposition of three states (red dots) which are no longer on a line, hence will not represent a code state. When the condition $2\chi T = 2\pi$ is met, this relative unwanted rotation vanishes.

Suzuki-Trotter decomposition:

$$\mathcal{T}e^{-i\int_0^T dt' H(t')}$$

$$= \lim_{n\to\infty} \prod_{j=1}^n \left(\underbrace{e^{\frac{-iT}{n}\lambda(t_n)(a+a^\dagger)}}_{D_{t_j}} \underbrace{e^{\frac{-i\omega T}{n}a^\dagger a}}_{R_n} \right)$$

$$= \lim_{n\to\infty} R_n^n\left(R_n^{-n}D_{t_n}R_n^n\right)\ldots\left(R_n^{-2}D_{t_2}R_n^2\right)\left(R_n^{-1}D_{t_1}R_n\right)$$

$$= R(\omega T)D\left[\lim_{n\to\infty}\sum_{j=1}^n \frac{-i\lambda(t_j)T}{n}e^{-i\omega T\frac{j}{n}}\right]e^{i\Psi},$$

where Ψ is determined using $D(\alpha)D(\beta) = \exp[i\text{Im}(\alpha\beta^*)]D(\alpha+\beta)$ and taking the limit $n\to\infty$.

We can apply this equality to the cavity-transmon dynamics $V(0,T) \equiv \mathcal{T}\exp[-i\int_0^T dt' H_{\text{tot}}(t')]$ as it is diagonal in the $\{|0\rangle, |1\rangle\}$ basis:

$$V(0,T) = R[(\omega_r - \chi)T]D(\gamma_-)e^{i(\Psi_- + \omega_q T/2)}|0\rangle\langle 0|$$
$$+ R[(\omega_r + \chi)T]D(\gamma_+)e^{i(\Psi_+ - \omega_q T/2)}|1\rangle\langle 1|,$$

where $\Psi_\pm = \int_0^T dt \int_t^T dt'\lambda(t)\lambda(t')\sin[(\omega_r \pm \chi)(t'-t)]$ and $\gamma_\pm = -i\int_0^T dt'\lambda(t')e^{-i(\omega_r \pm \chi)t'}$, the Fourier transform of $\lambda(t)$ at frequencies $\omega_r \pm \chi$. We note that the relative rotation when qubit is in state $|1\rangle$ versus $|0\rangle$ is $R(2\chi T)$. Thus, without constraining χT, the interaction between qubit and cavity field realizes a controlled rotation in addition to the desired controlled-displacement transformation $D(\gamma_\pm)$. We can understand the undesired effect of the relative rotation in the sketch in Fig. 14 in which a controlled-displacement
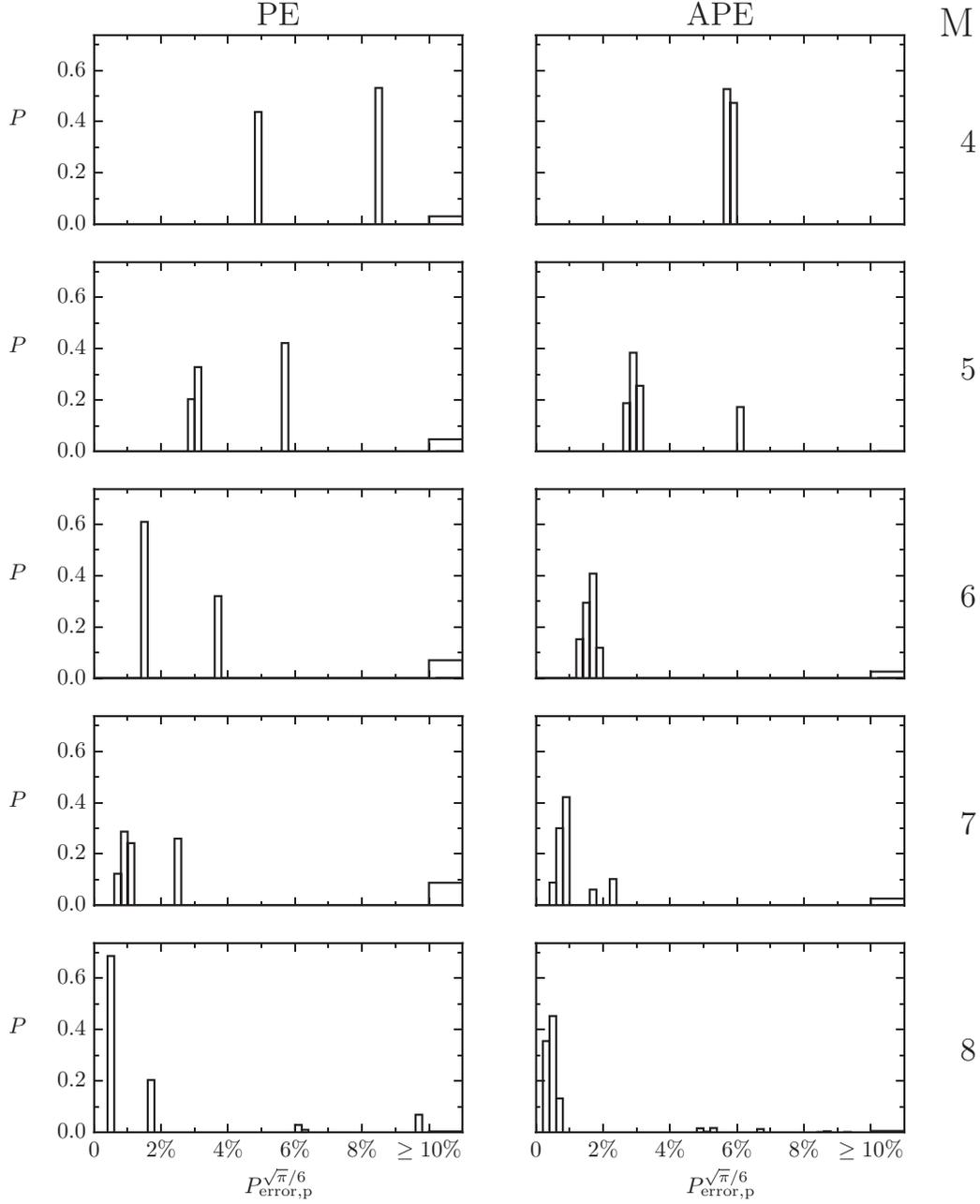
FIG. 15. Probability P to prepare a code state with an effective error rate $P_{\mathrm{error},p}^{\sqrt{\pi}/6}$ using M ancilla qubits (rounds) and an infinitely squeezed vacuum state as initial state, binned to 0.2%. Errors due to finite squeezing are uncorrelated to the effective error rate shown here and can be taken into account using Fig. 8. The wide bin $P_{\mathrm{error},p}^{\sqrt{\pi}/6} \geqslant 10\%$ contains all measurement outcomes with effective error rate $\geqslant 10\%$. Left: nonadaptive phase estimation of S_p. Right: adaptive phase estimation of S_p.

gate (and a some relative controlled rotation) is applied on, say, a cat state.

If we demand that T is chosen such that $\chi T = \pi k$, the relative rotation acts trivially $R(2\chi T) = I$. One thus has to choose $T = \pi/\chi$ which for $\chi/2\pi = 2.5$ MHz (which is the same as $\chi/2\pi$ in the experiment in [16]), equals 200 ns.

We can write the expressions for the displacements $D(\gamma_{\pm})$. For a square displacement pulse (with fictitious zero rise time) which is turned on for the entire interval T, the displacements $D(\gamma_{\pm})$ are given by

$$\gamma_+ = \frac{(-i\Omega_x - \Omega_y)T}{2} + \delta_+,$$

$$\gamma_- = \frac{(-i\Omega_x - \Omega_y)T}{2} e^{-i\omega_r T} \mathrm{sinc}(\omega_r T)$$

$$+ \frac{(-i\Omega_x + \Omega_y)T}{2} e^{i\chi T} \mathrm{sinc}(\chi T), \qquad (D3)$$

with $\mathrm{sinc}(x) = \sin(x)/x$ and a displacement error $\delta_+ = e^{-i(\omega_r+\chi)T} \frac{T(-i\Omega_x-\Omega_y)}{2} \mathrm{sinc}[(\omega_r+\chi)T]$. We wish that $\gamma_- \approx 0$ (cavity is off resonance) and note that this depends, in principle, on $\mathrm{sinc}(\chi T)$ and $\mathrm{sinc}(\omega_r T)$. One can always bound $|\mathrm{sinc}(x)| \leqslant 1/|x|$, which shows that the dominant term in the expression for the off-resonant displacement γ_- scales in strength like $\frac{|\Omega_x + i\Omega_y|T}{2\chi T}$. However, for our chosen pulse length $T = \pi/\chi$, $\mathrm{sinc}(\chi T) = 0$ and this analysis is too pessimistic.

For $\omega_r/2\pi = 10$ GHz (for example), one has $\frac{\omega_r T}{2\pi} = \frac{\omega_r}{2\chi} = 2 \times 10^3$ so that $|\mathrm{sinc}(\omega_r T)| < 0.5 \times 10^{-3}$. For the well-justified approximation $\omega_r \gg \chi$, $|\delta_+| \approx |\gamma_-| \approx |\gamma_+|\mathrm{sinc}(\omega_r T) < |\gamma_+| \times 0.5 \times 10^{-3}$. This means that the off-resonant displacement is a factor 10^3 less than the on-resonant displacement.

For the S_p measurement, one can then choose $\Omega_x = 0$ and $\gamma_+ = \sqrt{2\pi} = \frac{-\Omega_y T}{2}$ (assuming $\gamma_- \approx 0$). Similarly, for the S_q measurement one can choose $\Omega_y = 0$, etc. The value of the parameter Ω_y (or Ω_y) depends on the external coupling of the cavity with the drive line as well as the applied microwave power, e.g., $\Omega_y \propto \sqrt{\kappa_{\mathrm{ext}} F_t}$ where F_t is the photon flux per unit time. For the square pulse this photon flux per unit time is assumed to be constant during the interval T. Thus, $P \propto \omega_d T F_t = \omega_d n_{\mathrm{pulse}}$ is the total input power and n_{pulse} the total number of photons in the displacement pulse. We thus find that one can enact the controlled-displacement gate when $2\pi \approx \kappa_{\mathrm{ext}} n_{\mathrm{pulse}} T$. Given that T is fixed, depending on χ, and κ_{ext} is required to be small in order to have a high-Q cavity, this shows that by increasing n_{pulse} or the total input power one can always achieve the desired displacement.

It is important to note that if the condition $T = \pi/\chi$ is not accurately fulfilled, and one uses the pessimistic bound $|\mathrm{sinc}(x)| \leqslant \frac{1}{x}$, the relative difference in displacement equals $|\gamma_-|/|\gamma_+| \sim \frac{1}{\chi T} = \frac{1}{\pi}$. Also, it is not required that $\gamma_- \approx 0$ as long as the value of the *relative* displacement (for, say, S_p) is large enough, i.e., $\gamma_+ - \gamma_- = \sqrt{2\pi}$ and the values for $\gamma_\pm$ are known (calibrated). This means that we do not need to fulfill the condition $T_{\mathrm{pulse}} = T = \pi/\chi$ exactly. What happens when $T_{\mathrm{pulse}} < T$ and we also include the finite rise time of the pulse? The expressions derived above will remain valid, but lead to a different $\gamma_\pm$ (and $\Psi_\pm$) which could be fine tuned using the pulse shape $\Omega_x(t)$, $\Omega_y(t)$. These values of $\gamma_\pm$ can also be calibrated by testing the controlled-displacement gate on the vacuum state.

Let us comment on the phases $\Psi_\pm$ which affect the transmon qubit and thus need to be taken into account if the qubit undergoes further rotations. Obviously, only the relative phase $\Psi_+ - \Psi_-$ will affect the qubit state. If this relative phase remains *fixed* for every round of the S_p phase estimation measurement (and possibly different but fixed for the S_q measurement), it will not affect the accuracy of the prepared code states, even if this relative phase is unknown. It essentially means that whenever one does controlled-S_p in the protocol, one implements controlled-$S_p e^{i(\Psi_+ - \Psi_-)}$ instead. If the phase difference is not fixed, but known (so it represents a systematic error), it can be corrected with a rotation of the qubit. A fluctuating phase difference $\Psi_+ - \Psi_-$ will, however, lead to a noisier estimate in the phase estimation algorithm, hence shift errors in the resulting code state.

---

[1] J. Koch, Terri M. Yu, J. Gambetta, A. A. Houck, D. I. Schuster, J. Majer, A. Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Charge-insensitive qubit design derived from the Cooper pair box, Phys. Rev. A **76**, 042319 (2007).

[2] M. H. Devoret and R. J. Schoelkopf, Superconducting circuits for quantum information: An outlook, Science **339**, 1169 (2013).

[3] A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, Surface codes: Towards practical large-scale quantum computation, Phys. Rev. A **86**, 032324 (2012).

[4] D. Gottesman, A.Yu. Kitaev, and J. Preskill, Encoding a qubit in an oscillator, Phys. Rev. A **64**, 012310 (2001).

[5] Barbara M. Terhal, Quantum error correction for quantum memories, Rev. Mod. Phys. **87**, 307 (2015).

[6] Please note the somewhat confusing $\sqrt{2}$ conversion factor between displacements and quadrature values, i.e., $q = \sqrt{2} \mathrm{Re}(\alpha)$ and $p = \sqrt{2} \mathrm{Im}(\alpha)$: The spacing between peaks in the code states in $\mathrm{Re}(\alpha)$ is $\sqrt{2\pi}$ while it is $2\sqrt{\pi}$ in q space.

[7] One can verify this by noting that $\Pi_{S_p=1}^2 = \Pi_{S_p=1}$ and using the Poisson summation formula $\sum_{t=-\infty}^{\infty} e^{i\theta t} \propto \sum_s \delta(\theta = 2\pi s)$.

[8] B. C. Travaglione and G. J. Milburn, Preparing encoded states in an oscillator, Phys. Rev. A **66**, 052322 (2002).

[9] In some references, squeezing dB is estimated as $10\log_{10}(e^{2r}) = -10\log_{10}[2(\Delta q_{\mathrm{sq}})^2]$ where $(\Delta q_{\mathrm{sq}})^2$ is the variance of the squeezed state. For finite amounts of squeezing the difference between these estimates can be considerable, i.e., for $\Delta = 0.2$ one would obtain $10\log_{10}(1/\Delta^2)$ dB ≈ 13.98 dB of squeezing.

[10] P. Brooks, A. Kitaev, and J. Preskill, Protected gates for superconducting qubits, Phys. Rev. A **87**, 052306 (2013).

[11] S. Glancy and E. Knill, Error analysis for encoding a qubit in an oscillator, Phys. Rev. A **73**, 012325 (2006).

[12] H. Vasconcelos, L. Sanz, and S. Glancy, All-optical generation of states for "Encoding a qubit in an oscillator", Opt. Lett. **35**, 3261 (2010).

[13] S. Pirandola, S. Mancini, D. Vitali, and P. Tombesi, Constructing finite-dimensional codes with optical continuous variables, Europhys. Lett. **68**, 323 (2004).

[14] M. Mirrahimi, Z. Leghtas, V. V. Albert, S. Touzard, R. J. Schoelkopf, L. Jiang, and M. H. Devoret, Dynamically protected cat-qubits: A new paradigm for universal quantum computation, New J. Phys. **16**, 045014 (2014).

[15] L. Sun *et al.*, Tracking photon jumps with repeated quantum non-demolition parity measurements, Nature (London) **511**, 444 (2014).

[16] B. Vlastakis *et al.*, Deterministically encoding quantum information using 100-photon Schroedinger cat states, Science **342**, 607 (2013).

[17] G. Kirchmair *et al.*, Observation of quantum state collapse and revival due to the single-photon Kerr effect, Nature (London) **495**, 205 (2013).

[18] S. Haroche and J.-M. Raimond, *Exploring the Quantum: Atoms, Cavities, and Photons* (Oxford University Press, Oxford, UK, 2006).

[19] Leonard Mandel and Emil Wolf, *Optical Coherence and Quantum Optics* (Cambridge University Press, Cambridge, England, 1995).

[20] In [4] the phase space rotation $\exp(-i\delta a^\dagger a)$ was expanded in shifts γ such that $|\gamma|^2 \sim \delta$ (so small shift errors when δ is small), but the stationary phase approximation that was invoked was not proved in any rigorous fashion.

[21] K. M. Svore, M. B. Hastings, and M. Freedman, Faster Phase Estimation, Quantum Inf. Comput. **14**, 306 (2013).

[22] A. Yu. Kitaev, A. H. Shen, and M. N. Vyalyi, *Classical and Quantum Computation.*, Vol. 47 of Graduate Studies in Mathematics (American Mathematical Society, Providence, RI, 2002).

[23] B. L. Higgins, D. W. Berry, S. D. Bartlett, H. M. Wiseman, and G. J. Pryde, Entanglement-free Heisenberg-limited phase estimation, Nature (London) **450**, 393 (2007).

[24] B. L. Higgins, D. W. Berry, S. D. Bartlett, M. W. Mitchell, H. M. Wiseman, and G. J. Pryde, Demonstrating Heisenberg-limited unambiguous phase estimation without adaptive measurements, New J. Phys. **11**, 073023 (2009).

[25] D. W. Berry, B. L. Higgins, S. D. Bartlett, M. W. Mitchell, G. J. Pryde, and H. M. Wiseman, How to perform the most accurate possible phase measurements, Phys. Rev. A **80**, 052114 (2009).

[26] S. Kimmel, G. Hao Low, and T. J. Yoder, Robust Single-Qubit Process Calibration via Robust Phase Estimation, Phys. Rev. A **92**, 062315 (2015).

[27] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, UK, 2000).

[28] A. Yu. Kitaev, Quantum measurements and the Abelian stabilizer problem, arXiv:quant-ph/9511026.

[29] M. Mosca and A. Ekert, The hidden subgroup problem and eigenvalue estimation on a quantum computer, in *Proceedings of 1st NASA QCQC Conference*, Lecture Notes in Computer Science, Vol. 1509 (Springer, Berlin, 1998).

[30] R. Cleve, A. Ekert, C. Macchiavello, and M. Mosca, Quantum algorithms revisited, in *Proceedings of the Royal Society of London*, Vol. A454 (Royal Society Publishing, London, 1998), pp. 339–354.

[31] D. W. Berry, H. M. Wiseman, and J. K. Breslin, Optimal input states and feedback for interferometric phase estimation, Phys. Rev. A **63**, 053804 (2001).

[32] F. Mallet, M. A. Castellanos-Beltran, H. S. Ku, S. Glancy, E. Knill, K. D. Irwin, G. C. Hilton, L. R. Vale, and K. W. Lehnert, Quantum State Tomography of an Itinerant Squeezed Microwave Field, Phys. Rev. Lett. **106**, 220502 (2011).

[33] M. A. Castellanos-Beltran, K. D. Irwin, G. C. Hilton, L. R. Vale, and K. W. Lehnert, Amplification and squeezing of quantum noise with a tunable Josephson metamaterial, Nat. Phys. **4**, 929 (2008).

[34] M. Elliott and E. Ginossar, Enhancement and state tomography of a squeezed vacuum with circuit quantum electrodynamics, Phys. Rev. A **92**, 013826 (2015).

[35] S. E. Nigg, H. Paik, B. Vlastakis, G. Kirchmair, S. Shankar, L. Frunzio, M. H. Devoret, R. J. Schoelkopf, and S. M. Girvin, Black-Box Superconducting Circuit Quantization, Phys. Rev. Lett. **108**, 240502 (2012).

[36] M. Reagor *et al.*, Reaching 10 ms single photon lifetimes for superconducting aluminum cavities, Appl. Phys. Lett. **102**, 192604 (2013).

[37] A. Bruno, G. de Lange, S. Asaad, K. L. van der Enden, N. K. Langford, and L. DiCarlo, Reducing intrinsic loss in superconducting resonators by surface treatment and deep etching of silicon substrates, Appl. Phys. Lett. **106**, 182601 (2015).

[38] Z. Leghtas, G. Kirchmair, B. Vlastakis, M. H. Devoret, R. J. Schoelkopf, and M. Mirrahimi, Deterministic protocol for mapping a qubit to coherent state superpositions in a cavity, Phys. Rev. A **87**, 042315 (2013).

[39] M. Hatridge, S. Shankar, M. Mirrahimi, F. Schackert, K. Geerlings, T. Brecht, K. M. Sliwa, B. Abdo, L. Frunzio, S. M. Girvin, R. J. Schoelkopf, and M. H. Devoret, Quantum back-action of an individual variable-strength measurement, Science **339**, 178 (2013).

[40] S. Barzanjeh, D. P. DiVincenzo, and B. M. Terhal, Dispersive qubit measurement by interferometry with parametric amplifiers, Phys. Rev. B **90**, 134515 (2014).

[41] S. J. Srinivasan, N. M. Sundaresan, D. Sadri, Y. Liu, J. M. Gambetta, T. Yu, S. M. Girvin, and A. A. Houck, Time-reversal symmetrization of spontaneous emission for quantum state transfer, Phys. Rev. A **89**, 033857 (2014).

[42] A. J. Hoffman, S. J. Srinivasan, J. M. Gambetta, and A. A. Houck, Coherent control of a superconducting qubit with dynamically tunable qubit-cavity coupling, Phys. Rev. B **84**, 184515 (2011).

[43] P. van Loock, W. J. Munro, Kae Nemoto, T. P. Spiller, T. D. Ladd, Samuel L. Braunstein, and G. J. Milburn, Hybrid quantum computation in quantum optics, Phys. Rev. A **78**, 022303 (2008).

[44] R. W. Heeres, B. Vlastakis, E. Holland, S. Krastanov, V. V. Albert, L. Frunzio, L. Jiang, and R. J. Schoelkopf, Cavity State Manipulation Using Photon-Number Selective Phase Gates, Phys. Rev. Lett. **115**, 137002 (2015).

[45] S. Lloyd, Hybrid quantum computing, arXiv:quant-ph/0008057.

[46] E. Knill, Quantum computing with realistically noisy devices, Nature (London) **434**, 39 (2005).

[47] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, Gaussian quantum information, Rev. Mod. Phys. **84**, 621 (2012).

[48] Samuel L. Braunstein, Squeezing as an irreducible resource, Phys. Rev. A **71**, 055801 (2005).