Performance and structure of single-mode bosonic codes

Victor V. Albert,[1,2] Kyungjoo Noh,[1] Kasper Duivenvoorden,[3] Dylan J. Young,[1] R. T. Brierley,[1] Philip Reinhold,[1]
Christophe Vuillot,[3] Linshu Li,[1] Chao Shen,[1] S. M. Girvin,[1] Barbara M. Terhal,[3] and Liang Jiang[1]

[1]*Yale Quantum Institute, Departments of Applied Physics and Physics, Yale University, New Haven, Connecticut 06520, USA*

[2]*Walter Burke Institute for Theoretical Physics and Institute for Quantum Information and Matter, California Institute of Technology,
Pasadena, California 91125, USA*

[3]*JARA Institute for Quantum Information, RWTH Aachen University, Aachen 52056, Germany*

The early Gottesman, Kitaev, and Preskill (GKP) proposal for encoding a qubit in an oscillator has recently been followed by cat- and binomial-code proposals. Numerically optimized codes have also been proposed, and we introduce codes of this type here. These codes have yet to be compared using the same error model; we provide such a comparison by determining the entanglement fidelity of all codes with respect to the bosonic pure-loss channel (i.e., photon loss) after the optimal recovery operation. We then compare achievable communication rates of the combined encoding-error-recovery channel by calculating the channel's hashing bound for each code. Cat and binomial codes perform similarly, with binomial codes outperforming cat codes at small loss rates. Despite not being designed to protect against the pure-loss channel, GKP codes significantly outperform all other codes for most values of the loss rate. We show that the performance of GKP and some binomial codes increases monotonically with increasing average photon number of the codes. In order to corroborate our numerical evidence of the cat-binomial-GKP order of performance occurring at small loss rates, we analytically evaluate the quantum error-correction conditions of those codes. For GKP codes, we find an essential singularity in the entanglement fidelity in the limit of vanishing loss rate. In addition to comparing the codes, we draw parallels between binomial codes and discrete-variable systems. First, we characterize one- and two-mode binomial as well as multiqubit permutation-invariant codes in terms of spin-coherent states. Such a characterization allows us to introduce check operators and error-correction procedures for binomial codes. Second, we introduce a generalization of spin-coherent states, extending our characterization to qudit binomial codes and yielding a multiqudit code.

I. INTRODUCTION AND PROBLEM SETUP

Continuous-variable (CV) systems [1–4] continue to gain applications in quantum information processing and communication. The fundamental "moving part" of discrete-variable (DV) systems is one physical qubit, and one has to have a multitude of such qubits to construct a reliable logical qubit. By contrast, one may cleverly utilize the infinite-dimensional space of an oscillator or mode—the fundamental "moving part" of CV systems—in order to realize a comparably reliable logical qubit out of fewer moving parts. While many current linear-optical CV encodings use two modes per qubit in a "dual-rail" scheme [5,6] and CV logical qubits may consist of several modes in the long term, here we focus on a single mode since its theoretical limitations are not yet well understood and since it is useful for communication.

There have been several error-correcting CV encoding schemes proposed to date, formulated in terms of superpositions of either position and momentum eigenstates [7–12], coherent states [13–17], or Fock states [18–24] (see also other hybrid CV-DV schemes [25,26]). Besides the rich variety of quantum codes, there are two prevailing CV noise models: classical (i.e., Gaussian or displacement) noise and pure loss (more generally, thermal noise) [2]. Classical noise is modeled by a distribution of phase-space *displacements* while pure loss contracts phase space to the vacuum and is best understood in terms of *losses*. Because of the differing physical pictures

and mathematical formalisms of these noise models, codes designed to protect against one may or may not protect against the other. However, it is often difficult to rigorously prove protection from noise against which a code was not designed to protect. It is also difficult to study CV codes using the conventional stabilizer formalism because the noise model operators are not as well behaved. This paper closes these gaps by applying tools from qubit-based quantum error correction to CV codes which were not analyzed in this manner before.

A. Codes and error model

The code classes we consider are

$$\text{code} \in \{\text{cat}, \text{bin}, \text{num}, \text{gkps}, \text{gkp}\}. \tag{1.1}$$

The logical states for the first code class—the `cat` codes (5.1)—consist of superpositions of coherent states which are evenly distributed around a circle in phase space [13,15,27]. The second class of codes, the recently developed binomial (`bin`) codes (6.1), are designed to protect exactly against errors consisting of powers of raising or lowering operators up to some maximum order [23]. Here we show that `bin` codes are spin-coherent states embedded in an oscillator. We also include numerically (`num`) optimized codes (some from Ref. [23] and the rest developed here) that were obtained by minimizing the photon number of the code states subject to the constraints of protecting exactly against the first few
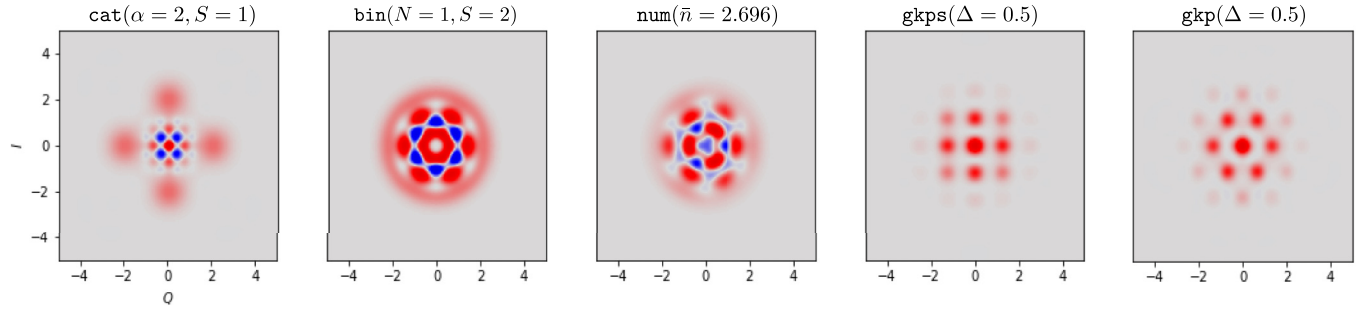
FIG. 1. Wigner function plots for maximally mixed logical states $\frac{1}{2}P_{\text{code}}$ for code being cat (5.1), bin (6.1), num (Appendix B), gkps (7.7b), and gkp (7.8), evaluated for given values of the respective parameters of the codes. On the axes, $Q = \frac{1}{2}\langle \hat{a} + \hat{a}^\dagger \rangle$ and $I = \frac{i}{2}\langle \hat{a}^\dagger - \hat{a} \rangle$; color scales are not the same for all plots.

powers of the lowering operator. The last class consists of GKP codes [9] which are the $+1$ eigenspace of two commuting phase-space displacement operators; since the code space is invariant under both displacements, the code space makes a lattice in phase space. The gkps (7.8) class, with s standing for square, corresponds to a square lattice. The gkp (B4) class consists of GKP codes built out of both the square and other nonrectangular lattices as well as codes whose lattice is shifted by half a lattice spacing from the origin, thus subsuming the gkps class. The gkps codes are presented separately in order to quantify any advantages of other lattices.

A single-mode qubit CV code is a two-dimensional subspace of the bosonic Hilbert space picked to be able to protect quantum information against errors.[1] It is unambiguously represented by the corresponding orthogonal projection onto the subspace,

$$P_{\text{code}} = |0_{\text{code}}\rangle\langle 0_{\text{code}}| + |1_{\text{code}}\rangle\langle 1_{\text{code}}|, \quad (1.2)$$

where code is picked from Eq. (1.1) and $|\mu_{\text{code}}\rangle$ (for $\mu \in \{0,1\}$) is any orthonormal basis for the code subspace. The maximally mixed state $\frac{1}{2}P_{\text{code}}$ thus provides a concise basis-independent fingerprint for each code; we plot the Wigner function of this state in Fig. 1.

We deal exclusively with codes representing a single qubit and are guided by the following question: *Which code best protects against the pure-loss channel?*

We are interested in the pure-loss channel because it is a model for broadband-line and free-space communication [3] and it is the most common *incoherent* error process in optical and microwave cavities [23]. The second most common error is cavity dephasing, which is caused by fluctuations in the cavity frequency. Optical cavities have to be actively stabilized to fix the frequency, but the effects of such fluctuations are small relative to effects of energy loss, particularly in microwave cavities. There are also other *coherent* error processes, such as a Kerr nonlinearity [29], which we briefly consider in Sec. VIII A.

The pure-loss bosonic channel (also known as bosonic amplitude damping or, more simply, as the lossy channel [2]) is Markovian: $\mathcal{N} = \exp(\chi \mathcal{D})$ with superoperator $\mathcal{D}(\cdot) =$

$\hat{a} \cdot \hat{a}^\dagger - \frac{1}{2}\{\hat{n}, \cdot\}$, where $a(a^\dagger)$ is the lowering (raising) operator for the bosonic mode and $\hat{n} \equiv \hat{a}^\dagger\hat{a}$. The dimensionless damping parameter equals $\chi = \kappa t$ for microwave or optical cavities (with excitation loss rate κ and time interval t) or $\chi = l/l_{\text{att}}$ for optical fibers (with fiber length l and attenuation length l_{att}). It is convenient to use the dimensionless *loss rate*

$$\gamma \equiv 1 - e^{-\chi} \quad (1.3)$$

to quantify the severity of the error channel, denoted as $\mathcal{N} \equiv \mathcal{N}_\gamma$ from now on. This channel can be expressed via unraveling [18,30–32] or Lie-algebraic [33] techniques in the Kraus representation, $\mathcal{N}_\gamma(\cdot) = \sum_{\ell=0}^{\infty} E_\ell \cdot E_\ell^\dagger$, with Kraus operators

$$E_\ell \equiv \left(\frac{\gamma}{1-\gamma}\right)^{\ell/2} \frac{\hat{a}^\ell}{\sqrt{\ell!}}(1-\gamma)^{\hat{n}/2}. \quad (1.4)$$

To leading order in γ, expansions of the first two Kraus operators suffice:

$$E_0 = I - \frac{1}{2}\gamma\hat{n} \quad \text{and} \quad E_1 = \sqrt{\gamma}\hat{a}. \quad (1.5)$$

This channel can also be derived by introducing an environment mode $\hat{b}$, coupling our oscillator with the vacuum state of this mode via a beam-splitter interaction

$$\hat{a} \rightarrow \sqrt{1-\gamma}\,\hat{a} + \sqrt{\gamma}\,\hat{b} \quad (1.6)$$

and tracing out the $\hat{b}$ mode [3,34]. The K-mode channel $\mathcal{N}_\gamma^{\otimes K}$ reduces to the multiqubit amplitude damping channel when restricted to the first two Fock states of each mode and reduces to the erasure channel when restricted to the single-excitation subspace.

Notice that this channel does not contain the identity as a Kraus operator for $\gamma \neq 0$. This is due to the *backaction* or *damping* term $(1-\gamma)^{\hat{n}/2}$ in Eq. (1.4), which reduces the probabilities of being in Fock states $|n > 0\rangle$ such that the only state remaining in the $\gamma \rightarrow 1$ limit is the vacuum Fock state $|n = 0\rangle$. Thus, when no losses are recorded (i.e., if $E_{\ell>0}$ has not yet acted on the state), there is still a redistribution of probability caused by E_0. Colloquially for $\gamma > 0$, if one has not lost any photons, then one likely did not have many photons to begin with.

---

[1]While there is a more general definition for multiqubit subsystem codes [28], we stick with the simpler definition since to our knowledge there are no single-mode CV subsystem codes.

B. Channel fidelity and recovery optimization

The combined quantum channel we consider consists of an encoding step $\mathcal{S}_{\text{code}}$, action of the pure-loss channel $\mathcal{N}_\gamma$, a recovery channel $\mathcal{R}$, and a perfect decoding step $\mathcal{S}_{\text{code}}^{-1}$. The encoding step maps the quantum information from the qubit *source space* A [35] into the code subspace of the bosonic Hilbert space; this step is represented by $\mathcal{S}_{\text{code}}$. More precisely, $\mathcal{S}_{\text{code}}(\rho) = S\rho S^{-1}$, where $S = |0_{\text{code}}\rangle\langle 0_{\text{A}}| + |1_{\text{code}}\rangle\langle 1_{\text{A}}|$, ρ is a qubit density matrix in A, and $|0_{\text{A}}/1_{\text{A}}\rangle$ is a basis for A. Therefore, $SS^{-1} = P_{\text{code}}$ and, if $\{|0_{\text{code}}\rangle, |1_{\text{code}}\rangle\}$ are orthonormal, $S^{-1}S = I_{\text{A}}$, the identity on A. The combined channel

$$\mathcal{E} \equiv \mathcal{S}_{\text{code}}^{-1} \circ \mathcal{R} \circ \mathcal{N}_\gamma \circ \mathcal{S}_{\text{code}} \qquad (1.7)$$

thus maps density matrices in A back to A. In contrast, $\mathcal{R} \circ \mathcal{N}_\gamma$ is a map from the bosonic space to the code subspace. The form of $\mathcal{E}$ depends on the code, the loss rate γ, and the recovery $\mathcal{R}$. The channel can be written in the Kraus representation, $\mathcal{E}(\cdot) = \sum_k A_k \cdot A_k^\dagger$, or in the *matrix or Liouville representation*—as a 4×4 matrix with elements

$$\mathcal{E}_{kl} = \tfrac{1}{2}\text{Tr}\{\sigma_k \mathcal{E}(\sigma_\ell)\}, \qquad (1.8)$$

using the three Pauli matrices $\sigma_{1,2,3}$ and identity $I_{\text{A}} \equiv \sigma_0$ (e.g., Ref. [36], Sec. 2.2). Composition "$\circ$" in Eq. (1.7) is equivalent to matrix multiplication in the matrix representation, so we omit the symbol.

None of the codes we consider protect against all errors in $\mathcal{N}_\gamma$, so we have to consider approximate quantum error correction [37,38]. We compare the codes using the *channel fidelity* $F_\mathcal{E}$ [39]—a specific case of entanglement fidelity [40] that is motivated as follows. Let the source qubit A be in a maximally entangled state with ancillary qubit B, i.e., $|\Psi\rangle = (|0_{\text{A}}0_{\text{B}}\rangle + |1_{\text{A}}1_{\text{B}}\rangle)/\sqrt{2}$. Qubit B is left alone (i.e., acted on by the identity superoperator $\mathcal{I}$) while the source qubit is acted on by the channel $\mathcal{E}$ in Eq. (1.7). The channel fidelity $F_\mathcal{E}$ is simply the overlap between the initial state $|\Psi\rangle$ and the final state

$$\rho_\mathcal{E} \equiv \mathcal{E} \otimes \mathcal{I}(|\Psi\rangle\langle\Psi|) \qquad (1.9)$$

(which we define to be the Choi matrix[2] of $\mathcal{E}$):

$$F_\mathcal{E} \equiv \langle\Psi|\rho_\mathcal{E}|\Psi\rangle. \qquad (1.10)$$

Remembering that $\text{Tr}_{\text{B}}\{|\Psi\rangle\langle\Psi|\}$ is the maximally mixed state $\tfrac{1}{2}I_{\text{A}}$ of qubit A, a few simple manipulations yield

$$F_\mathcal{E} = \frac{1}{4}\sum_{k=1}^{4} |\text{Tr}\{A_k\}|^2 = \frac{1}{4}\text{Tr}\{\mathcal{E}\}, \qquad (1.11)$$

where Tr is the trace in the matrix representation (1.8).

Besides clearly being an intrinsic property of $\mathcal{E}$ that is invariant under unitary rotations, several other properties of $F_\mathcal{E}$ make the quantity both meaningful and practically useful. We first mention the property that is crucial for our task, listing the remaining properties in Appendix A. It turns out that the recovery $\mathcal{R}$ which gives the optimal $F_\mathcal{E}$ is computable

via a semidefinite program [44] (see also Refs. [39,45]). This allows us to quickly obtain the highest possible $F_\mathcal{E}$ using a laptop (given reasonable $\bar{n}_{\text{code}}$) and without having to design a recovery for each code. This procedure was applied to the multiqubit context by Fletcher, Shor, and Win [43] (see also Ref. [35] and references therein), and our benchmarking is in some sense a counterpart to that work in the oscillator context. From now on and unless otherwise noted, we let the recovery piece $\mathcal{R}$ of $\mathcal{E}$ (1.7) be one which gives the highest $F_\mathcal{E}$, given a member of the code family and a loss rate γ.

C. Outline of this paper

In Sec. II, we state our main numerical code comparison results and summarize the supporting analytical calculations. In Sec. III, we numerically analyze communication rates of our codes by calculating the hashing bound of $\mathcal{E}$. In Sec. IV, we review the quantum error-correction conditions. In Secs. V, VI, and VII, we calculate these conditions for the cat, bin, and gkp codes, respectively. In Sec. VI C, we characterize single-qubit bin codes in terms of spin-coherent states and relate them to two-mode binomial codes and multiqubit permutation-invariant codes. In Sec. VIII, we analyze code performance after a Kerr interaction is added to the pure loss channel and briefly study the effect of tracking the photon number parity. We summarize our results and discuss future directions in Sec. IX.

II. TAKE-HOME MESSAGES

Here we summarize the results related to code performance, but start off by mentioning two caveats to our primary numerical comparison. Results relating the structure of bin codes to spin-coherent states and other multiqubit codes are summarized in Sec. IX B.

Caveat 1 is that *the encoding, recovery, and decoding are all to be perfect*, meaning that there are no other errors besides $\mathcal{N}_\gamma$ incurred by the state. Therefore, the results of this section should be interpreted as theoretical bounds on code capabilities and not as practical suggestions on the best experimental design. Moreover, optimal recovery procedures are not created equal in the eyes of current technologies: cat code error correction has already been performed [29] while gkp states have yet to be realized. We briefly investigate one additional imperfection in Sec. VIII by including a nonlinearity. There, we also address the consequences of being able to perfectly track the photon number parity.

Caveat 2 has to do with how we quantify the "size" of the codes. Namely, *we organize the codes by mean occupation number*

$$\bar{n}_{\text{code}} \equiv \text{Tr}\{P_{\text{code}}\hat{n}\}/2. \qquad (2.1)$$

While $\bar{n}_{\text{code}}$ is proportional to the average energy required to construct a code state, it does not describe the spread or variance in Fock space, $\sigma_{\text{code}}^2 \equiv \tfrac{1}{2}\text{Tr}\{P_{\text{code}}\hat{n}^2\} - \bar{n}_{\text{code}}^2$. While cat and bin codes follow approximately Poisson and binomial distributions in Fock space, respectively, we will show that gkp codes are *geometrically (i.e., thermally) distributed* and thus have much larger "tails" in Fock space at higher $\bar{n}$. Therefore, for the same $\bar{n}_{\text{code}}$, gkp codes utilize much more of the Fock

---

[2]Note that $\mathcal{I} \otimes \mathcal{E}(|\Psi\rangle\langle\Psi|)$ is used for the Choi matrix in [41,42]. Our convention [35,43] yields $\rho_\mathcal{E} = \sum_k |A_k\rangle\rangle\langle\langle A_k|$ for vectorized Kraus operators $|A_k\rangle\rangle$ of $\mathcal{E}$, but unfortunately makes Alice and Bob switch places.
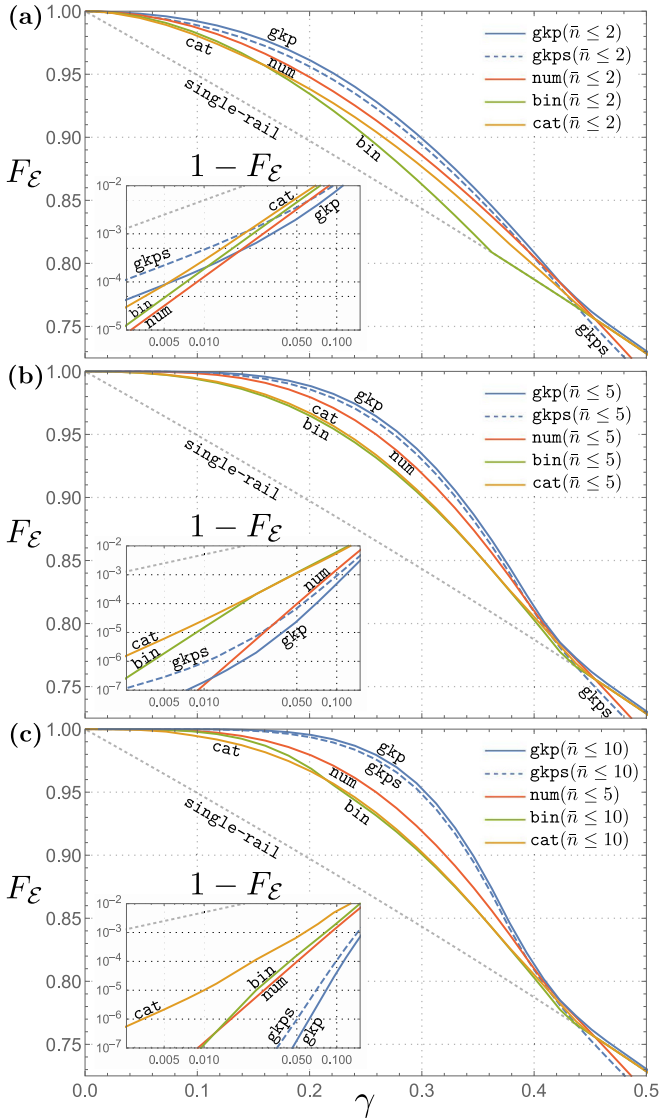
FIG. 2. Channel fidelity $F_\mathcal{E}$ (1.11) given an optimal recovery operation and optimized over all instances of each code given an occupation number constraint $\bar{n}_{\text{code}} \leqslant 2$ (a), 5 (b), or 10 (c). The dotted diagonal line, drawn for reference, is the optimal $F_\mathcal{E}$ for *single-rail encoding* [6] (whose logical states are the Fock states $|0\rangle, |1\rangle$). While gkp codes perform worse than the other codes for sufficiently small γ (see insets), they outperform all other codes as γ is increased despite not being designed to protect against the pure-loss channel. We were not able to obtain significantly better num codes with $\bar{n}_{\text{num}} > 5$ due to a large set of parameters to be optimized over, so red curves in panels (b) and (c) are identical. Parameters for all of the codes used are in Table IV.

space than cat and bin and are therefore "larger" (in the same sense that multiqubit codes constructed out of ten physical qubits are larger than those constructed out of five). A simple energy parameter does not quantify such a notion of size.

A. Numerical comparison

The procedure we use to determine the channel fidelities shown in Fig. 2 is as follows: Recall that each code family (1.1) contains multiple instances of codes. For example, a member

of the bin code family is parameterized by the number of dephasing and loss errors it can correct (N and S, respectively; more details are in Sec. IV). For each loss rate γ in Eq. (1.3), we calculate the optimal $F_\mathcal{E}$ for all instances of each code family subject to the energy constraint that $\bar{n}_{\text{code}} \leqslant 2, 5, 10$ [shown in Figs. 2(a), 2(b) and 2(c), respectively]. Then, we pick the highest $F_\mathcal{E}$ out of all members of the code family and plot it in the figure. We repeat for other values of $\gamma \leqslant 0.5$—the point at which the one-way channel capacity of $\mathcal{N}_\gamma$ becomes zero (see Sec. III for details). As a result of this simultaneous optimization over each code family and over the recovery for a given member of the family, the code which gives the highest $F_\mathcal{E}$ may change with γ and curves in Fig. 2 may display discontinuous derivatives.

Let us first focus on the $\bar{n}_{\text{code}} \leqslant 2$ case in Fig. 2(a) and examine the infidelity ($1 - F_\mathcal{E}$) shown in the log-plot inset. For $\gamma \leqslant 0.025$, specific members of num, bin, and cat perform the best (in that order), showing similar scaling with γ. All three of these codes were designed to deal with errors $\{I, \hat{a}\}$, the dominant terms in $E_{\ell=0,1}$ at small γ (1.5). The num and bin codes show quadratic scaling versus small γ: In a polynomial fit up to order two for $\gamma \leqslant 0.025$, $c_0 + c_1 \gamma + c_2 \gamma^2$, num and bin codes have negligible coefficients $c_0, c_1 \approx 10^{-4}$, and a c_2 of 1.3/1.8, respectively. The cat codes have negligible $c_0, c_1 \approx 10^{-3}$, and $c_2 \approx 2.2$. Following these codes, gkps and gkp perform the worst for small γ, underperforming the other codes as $\gamma \to 0$. This should be expected since these codes were designed to protect against small displacements and not loss events. It is also reasonable that gkp should slightly outperform gkps due to the idea that nonsquare lattices allow for tighter packing than square lattices [9]. The main unexpected behavior for $\bar{n}_{\text{code}} \leqslant 2$ occurs for $\gamma \geqslant 0.025$. There, we see that gkps and gkp actually *outperform* the rest of the codes (this will be discussed later).

We remark here that, for each γ, the amplitude α of the coherent states making up cat codes [see Fig. 1(a)] for the optimal cat code is at a fine-tuned "sweet spot" $\alpha_\star(\gamma)$ which balances the backaction due to the difference in the mean occupation number of the logical states (significant for small α but zero at $\alpha \to \infty$) against the probability that a photon will be lost (zero at $\alpha \to 0$). We discuss this effect more in Sec. V, noting that it has also been studied elsewhere [23,46].

Continuing to $\bar{n}_{\text{code}} \leqslant 5$ in Fig. 2(b), we see substantial increases in performance for all codes. We list notable infidelities for selected γ and $\bar{n}_{\text{code}} \leqslant 5$ in Table I(b). For example, for the relatively lossy channel having $\gamma = 0.0952$, there exist codes in *all five families* which have a channel fidelity higher than 99.4%. Moreover, all such codes have only five photons in them on average, so they could be within reach even with noisy intermediate-scale quantum (NISQ) technologies [47]. At small γ [inset of Fig. 2(b)], we again see polynomial scaling for the cat, bin, and num codes, which are able to deal with loss errors $\{I, \hat{a}, \hat{a}^2\}$. We also once again see gkp outperform the other codes for $\gamma \geqslant 0.05$ and underperform as $\gamma \to 0$. The cat and bin code performances are almost identical, with the exception of small γ, where bin performs slightly better. Expounding on this behavior in Sec. VI A, we show that bin allows for better error suppression than cat in certain ranges of $\bar{n}$. One num code performs the best for all $\gamma \leqslant 0.4$—the code with $\bar{n}_{\text{num}} = 4.15$ (see Appendix B).

TABLE I. Channel infidelity $1 - F_\mathcal{E}$ from Figs. 2(a)–2(c) at selected loss rates γ (with γ rounded to the nearest hundredth).

			(a) $1 - F_\mathcal{E}(\bar{n} \leqslant 2)$		
γ	cat	bin	num	gkps	gkp
0.01	4.2×10^{-4}	2.9×10^{-4}	2.0×10^{-4}	6.0×10^{-4}	2.5×10^{-4}
0.05	5.3×10^{-3}	4.3×10^{-3}	3.1×10^{-3}	3.4×10^{-3}	1.9×10^{-3}
0.10	1.8×10^{-2}	1.6×10^{-2}	1.2×10^{-2}	1.0×10^{-2}	7.1×10^{-3}
0.20	6.2×10^{-2}	6.6×10^{-2}	5.3×10^{-2}	4.5×10^{-2}	3.9×10^{-2}
0.31	1.3×10^{-1}	1.5×10^{-1}	1.2×10^{-1}	1.2×10^{-1}	1.1×10^{-1}
			(b) $1 - F_\mathcal{E}(\bar{n} \leqslant 5)$		
γ	cat	bin	num	gkps	gkp
0.01	4.4×10^{-5}	2.8×10^{-5}	3.7×10^{-7}	1.4×10^{-6}	3.2×10^{-7}
0.05	1.1×10^{-3}	1.1×10^{-3}	8.8×10^{-5}	6.3×10^{-5}	2.2×10^{-5}
0.10	4.9×10^{-3}	5.4×10^{-3}	1.2×10^{-3}	7.6×10^{-4}	3.9×10^{-4}
0.20	3.4×10^{-2}	3.6×10^{-2}	2.1×10^{-2}	1.5×10^{-2}	1.2×10^{-2}
0.31	1.1×10^{-1}	1.1×10^{-1}	9.2×10^{-2}	8.2×10^{-2}	7.7×10^{-2}
			(c) $1 - F_\mathcal{E}(\bar{n} \leqslant 10)$		
γ	cat	bin		gkps	gkp
0.01	1.7×10^{-5}	3.7×10^{-7}		3.0×10^{-10}	1.0×10^{-11}
0.05	6.3×10^{-4}	1.5×10^{-4}		8.2×10^{-7}	1.5×10^{-7}
0.10	4.9×10^{-3}	1.7×10^{-3}		7.9×10^{-5}	2.9×10^{-5}
0.20	3.4×10^{-2}	3.1×10^{-2}		6.5×10^{-3}	4.6×10^{-3}
0.31	1.1×10^{-1}	1.1×10^{-1}		6.3×10^{-2}	5.9×10^{-2}

Let us now consider the case $\bar{n}_{\text{code}} \leqslant 10$ in Fig. 2(c). While these codes may be difficult to create and correct experimentally in the near future, it is nevertheless interesting to see whether doubling the occupation number constraint allows for any improvements of the code. The most noticeable difference between $\bar{n}_{\text{code}} \leqslant 10$ and $\bar{n}_{\text{code}} \leqslant 5$ is that gkp pulls away from the other codes for *all values* of γ sampled. While it is believable that the other codes will still scale more favorably for sufficiently small γ, this occurs only at $\gamma < 0.01$. For larger γ, gkp codes demonstrate $F_\mathcal{E} \geqslant 0.99$ even at $\gamma = 0.2$ (see Table I). Looking at Table IV, the best codes in those families for most γ are those which also have $\bar{n}_{\text{gkp}} \approx 10$. In other words, gkp performs better with increasing $\bar{n}$. A similar monotonic increase in performance occurs for subsets bin$(N, S \approx \xi N)$ of binomial codes (with ξ dependent on γ) when the $\bar{n}_{\text{bin}}$ constraints are relaxed (see Sec. VI B). We explain the bin increase in performance in Secs. VI A–VI B, revealing that they have a larger set of approximately correctable errors than previously thought. This behavior is not seen in cat codes, which do not perform much better than those in Fig. 2(b) and work best at some finite $\bar{n}$. This idea that increasing $\bar{n}$ does not lead to better cat code performance has been observed before in different contexts [46,48,49]. By contrast, the ideal gkp codes have infinite $\bar{n}$, so it seems reasonable that increasing $\bar{n}$ should improve performance. These results support the conjecture that the ordering of the codes with respect to $F_\mathcal{E}$ is gkp > bin > cat when there is no energy constraint.

The numerical results show that codes designed to work well at small γ do not perform well for large γ, and vice versa. More specifically, extensions of the ideas used to correct dominant errors at small γ do not necessarily lead to good codes at larger γ. For instance, the cat and bin codes protect exactly from the first few loss errors by making sure there is adequate *Fock state spacing* S between the states. As an example, an $S = 2$ cat/bin code uses superpositions of Fock states $|0\rangle, |6\rangle, \ldots$

for $|0_{\text{code}}\rangle$ and $|3\rangle, |9\rangle, \ldots$ for $|1_{\text{code}}\rangle$. This guarantees that loss events $E_{\ell=1,2}$, which lower each Fock state by either 1 or 2, do not cause the logical states to overlap with each other. Both cat and bin allow one to increase S arbitrarily, while gkp codes have $S \in \{0, 1\}$, depending on whether their lattice is shifted from the origin or not. Figure 5 shows that, for sufficiently large γ and $\bar{n}_{\text{code}}$, correcting a few errors exactly with spacing (done by cat and bin) is not as helpful as suppressing all errors approximately (done by gkp).

B. Analytical results

To summarize, the family of gkp codes outperforms all of the other codes for most γ, with the exception of small γ (which gets even smaller as the energy constraint is loosened). We see similar behavior analyzing the optimal codes from an information-theoretic perspective in Sec. III. Since all of the other codes were specifically designed to protect against loss errors and gkp codes were designed to protect against displacement errors, gkp have apparently outperformed all of the other codes "at their own game" (albeit a game whose rules were set by caveats 1 and 2). To understand this effect, we have undertaken extensive analytical calculations to determine the quantum error-correction conditions for the pure-loss channel for gkp codes [see Eq. (7.18)] as well as how $1 - F_\mathcal{E}$ scales with γ. As noted in the previous subsection, while other codes protect against errors (to some order in γ) exactly, gkp protect against all errors *approximately*. In other words, other codes protect against the first few errors exactly, but have low fidelity when there is a large probability of an unprotected error occurring. By contrast, gkp codes do not protect against most errors exactly, but the contributions from *all errors* to the infidelity is small.

In order to bound the scaling of $F_\mathcal{E}(\text{gkp})$ and since there is no analytic expression for the optimal $\mathcal{R}$, we calculate a

lower bound on the channel fidelity $F_{\mathcal{E}}^{\text{AGKP}}$ using a recovery $\mathcal{R}^{\text{AGKP}}$ (7.22) which consists of phase-insensitive amplification, followed by conventional gkp recovery consisting of displacement measurements and corrections. The recovery is based on the idea that, after fine-tuning some of the channel parameters,

pure loss + amplification = Gaussian noise,

where Gaussian noise corresponds to uniform diffusion in phase space and its channel has displacements for its Kraus operators ([2], Sec. 2.3; see also Ref. [50]). In other words, amplification (with gain $\frac{1}{1-\gamma}$) exactly compensates the contractive effect of pure loss (with loss rate γ) while at the same time adding noise that, in this context, reduces to Gaussian noise (with variance $\frac{\gamma}{1-\gamma}$). This idea has been considered in the context of communication schemes [51,52]; we apply it to bosonic error correction by noting that Gaussian noise is *exactly* the type of noise that gkp was designed for. That way, we can use earlier tools [9,53] developed to quantify gkp performance against such noise. We consider single-mode behavior here, noting that, in the multimode case, gkp codes can be used to communicate efficiently across more general Gaussian channels [54]. Note that the related idea of "amplication + pure loss = Gaussian noise" has been recently applied to determine bounds on quantum capacities of more general Gaussian channels [54–56].

The probability of failure of $\mathcal{R}^{\text{AGKP}}$ gives a lower bound on $F_{\mathcal{E}}^{\text{AGKP}}$, which in turn bounds $F_{\mathcal{E}}(\text{gkp})$ (see Sec. VII B). The bound contains an essential singularity at $\gamma = 0$,

$$F_{\mathcal{E}}(\text{gkp}) > 1 - \exp\left(-\frac{\pi}{4c}\frac{1-\gamma}{\gamma}\right), \qquad (2.2)$$

where c is a constant determined by the lattice used to construct the code. This exponential suppression of infidelity explains the nontrivial scaling of gkp codes at small γ and accounts for their breakaway performance at higher γ. Because of the nonanalyticity, having different lattices becomes important for $\gamma \ll 1$, where gkp outperforms gkps by an order of magnitude [see Table I(c)].

III. THE HASHING BOUND OF $\mathcal{E}$

Since channel fidelity provides a measure for entanglement preservation, it is also of interest to examine these results from an information-theoretic perspective. The question we aim to answer in this section is the following: *What is the achievable communication rate of $\mathcal{E}$?*

As with the $F_{\mathcal{E}}$, this question also presupposes caveats 1 and 2 from Sec. II. In other words, we assume recovery operations $\mathcal{R}$ can be done perfectly and organize the codes by mean occupation number $\bar{n}_{\text{code}}$, ignoring their "size" in terms of the number of Fock states necessary to express the logical states.

The quantum communication rate of $\mathcal{E} = \mathcal{S}_{\text{code}}^{-1}\mathcal{R}\mathcal{N}_{\gamma}\mathcal{S}_{\text{code}}$ (1.7) is ultimately limited by its most destructive link—the pure-loss channel $\mathcal{N}_{\gamma}$ (1.4). Given an energy constraint of maximum $\bar{n}$, the quantum capacity of $\mathcal{N}_{\gamma}$ is given by ([57], Theorem 8; see also Ref. [58], Eq. (12), and Ref. [59])

$$Q_{\bar{n}} = \max\{0, g((1-\gamma)\bar{n}) - g(\gamma\bar{n})\}, \qquad (3.1)$$

where $g(\bar{n}) = (\bar{n}+1)\log_2(\bar{n}+1) - \bar{n}\log_2\bar{n}$ is the von Neumann entropy of a thermal state with $\bar{n}$. Note that in the limit of $\bar{n} \to \infty$, this capacity approaches the unconstrained quantum capacity[3]

$$Q_{\infty} = \max\left\{0, \log_2\left(\frac{1-\gamma}{\gamma}\right)\right\}. \qquad (3.2)$$

We see how close the codes giving the best $F_{\mathcal{E}}$ (see Fig. 2) come to $Q_{\bar{n}}$ for $\bar{n} \leqslant 2, 5, 10$ by calculating a lower bound on the capacity of $\mathcal{E}$ for each γ and each code listed in Table IV. The lower bound we use is known as the *hashing bound* $D_{\mathcal{E}}$ of $\mathcal{E}$ [62]—the (reverse[2] [63]) coherent quantum information of $\mathcal{E}$'s Choi matrix $\rho_{\mathcal{E}}$ (1.9), where $|\Psi\rangle = (|0_{\text{A}}0_{\text{B}}\rangle + |1_{\text{A}}1_{\text{B}}\rangle)/\sqrt{2}$ is a maximally entangled state of A and B in an arbitrary basis:

$$D_{\mathcal{E}} \equiv H(\text{Tr}_{\text{B}}\{\rho_{\mathcal{E}}\}) - H(\rho_{\mathcal{E}}), \qquad (3.3)$$

where $H(\rho) = -\text{Tr}\{\rho\log_2\rho\}$. This one-shot (i.e., with one application of the channel) coherent information of $\rho_{\mathcal{E}}$ provides an achievable rate of quantum communication and entanglement distillation, assuming many copies of $\mathcal{E}$ are available ([63–65]; see also Corrolary 21.2.1 and Theorem 23.9.1 in Ref. [62]). Therefore, $D_{\mathcal{E}}$ *does not* supply an achievable rate of any one oscillator code, but instead gives an achievable rate of concatenation schemes of the oscillator code with other (outer) codes with the restriction that $\mathcal{R}$ is used as the recovery for the (inner) oscillator code.

The first term in $D_{\mathcal{E}}$ can be simplified to yield an expression only in terms of $\mathcal{E}$,

$$D_{\mathcal{E}} = H\left(\left\{\tfrac{1}{2} \pm \tfrac{1}{2}\sqrt{\tfrac{1}{2}\|\mathcal{E}(I_{\text{A}})\|^2 - 1}\right\}\right) - H(\rho_{\mathcal{E}}), \qquad (3.4)$$

where $H(\{x\}) = -\sum_x x\log_2 x$ for a set of variables $\{x\}$, $\|O\|^2 \equiv \text{Tr}\{O^{\dagger}O\}$ is the Frobenius norm of an operator O, and I_{A} is the qubit identity. Derivation of the first term was done by determining the reduced qubit density matrix $\text{Tr}_{\text{B}}\{\rho_{\mathcal{E}}\}$ in terms $\mathcal{E}$'s matrix representation (1.8) and then diagonalizing to yield the two eigenvalues in the term's argument. Note that this term is maximized when $\mathcal{E}$ is unital [$\mathcal{E}(I_{\text{A}}) = I_{\text{A}}$]. The second term—the von Neumann entropy of the Choi matrix $\rho_{\mathcal{E}}$—increases with the minimal number of Kraus operators needed to express $\mathcal{E}$ and is zero when $\mathcal{E}$ is a unitary channel.

Incidentally, the analytical formula for $D_{\mathcal{E}}$ provides an easily calculable lower bound on the quantum capacity $Q_{\mathcal{E}}$ of any qubit channel $\mathcal{E}$. It turns out that $D_{\mathcal{E}}$ is quite close to $Q_{\mathcal{E}}$ for a family of two Kraus operator qubit channels ([66], Eq. (5)) which includes the dephasing and amplitude damping channel: We have checked numerically that the difference $Q_{\mathcal{E}} - D_{\mathcal{E}} < 0.005$.

A. Hashing bound for codes giving optimal $F_{\mathcal{E}}$

In Fig. 3, we plot the hashing bound $D_{\mathcal{E}}$ for all of the codes which produce the optimal channel fidelity $F_{\mathcal{E}}$ in Fig. 2 for the three energy constraints $\bar{n}_{\text{code}} \leqslant 2, 5, 10$. In other words, this plot is not an optimization of $D_{\mathcal{E}}$ over all codes, but merely

---

[3]Equation (3.2) was proven in Ref. [60], which built on Ref. [61] (see also Remark 17 in Ref. [58]). A standalone derivation is, e.g., in Sec. III of Ref. [54].
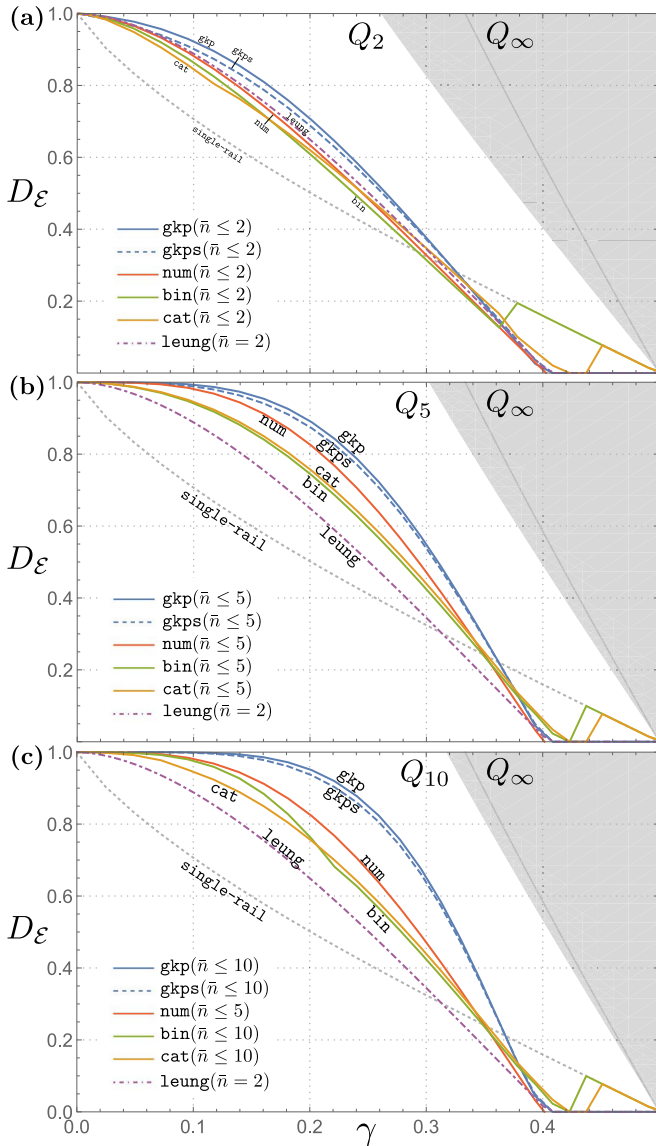
FIG. 3. Hashing bound $D_{\mathcal{E}}$ (3.3) of the codes which optimize $F_{\mathcal{E}}$ for a given γ and given constraints $\bar{n}_{code} \leqslant 2$ (a), 5 (b), or 10 (c). The boundary of the gray region is $Q_{\bar{n}=2,5,10}$ (3.1), the capacity of $\mathcal{N}_{\gamma}$ given the energy constraint of $\bar{n}$. The gray line is the unconstrained capacity $Q_{\bar{n}\to\infty}$ (3.2). The thin dotted diagonal line is $D_{\mathcal{E}}$ for the single-rail encoding, which has the highest $D_{\mathcal{E}}$ at large γ. Recall that cat and bin codes include the single-rail encoding [i.e., this encoding is also cat($\alpha = S = 0$) and bin($N = S = 0$)], so bin and cat eventually jump to match the dotted line. We also show $D_{\mathcal{E}}$ for the four-qubit leung code (3.5) versus qubit amplitude damping (i.e., the pure-loss channel restricted to the first two Fock states of four modes). There are no num codes with $\bar{n}_{num} \geqslant 5$, so red curves in panels (b) and (c) are identical.

a plot of $D_{\mathcal{E}}$ for the codes which give optimal $F_{\mathcal{E}}$. Recall that $D_{\mathcal{E}}$ is a lower bound on the entanglement that is theoretically distillable from using unlimited instances of $\mathcal{E}$ and one-way classical communication from B to A. By contrast, $F_{\mathcal{E}}$ is an overlap which gauges how well entanglement was transmitted over just one instance of $\mathcal{E}$. While $F_{\mathcal{E}}$ bounds one of the terms in $D_{\mathcal{E}}$ [see Eq. (A3)] and the two yield a similar order of

performance of the codes, there is no guarantee that codes giving the optimal value of $F_{\mathcal{E}}$ should also give *optimal* values for $D_{\mathcal{E}}$. In other words, even if $F_{\mathcal{E}}(\text{one}) > F_{\mathcal{E}}(\text{two})$ given two codes one and two, there could still exist an entanglement distillation scheme which extracts more entanglement from the unlimited instances of $\mathcal{E}$ from code two. This is true for our codes at large γ. Let us for example consider $\bar{n} \leqslant 10$. At around $\gamma \approx 0.37$, all of the codes begin to have a lower $D_{\mathcal{E}}$ than the single-rail $\{|0\rangle, |1\rangle\}$ Fock state encoding [thin black line in Fig. 3(c)]. By contrast, the point at which the codes begin to have a lower $F_{\mathcal{E}}$ than the single-rail encoding is $\gamma \approx 0.42$ [see Fig. 2(c)]. The cat and bin codes include this encoding [i.e., single rail is also cat($\alpha = S = 0$) and bin($N = S = 0$)], so we can say for certain that the cat and bin codes which optimize $F_{\mathcal{E}}$ are not those which optimize $D_{\mathcal{E}}$ at $0.37 \leqslant \gamma \leqslant 0.42$.

For $\gamma \lesssim 0.3$, we once again see similar behavior of the codes (relative to each other) as with $F_{\mathcal{E}}$. For $\bar{n}_{code} \leqslant 10$, gkp codes break from the pack and bridge the gap with $Q_{\bar{n}}$ most rapidly. For example, at the high loss rate $\gamma = 0.3$, $D_{\mathcal{E}}[\text{gkp}(\bar{n} \leqslant 10)] \approx 0.63$ bits is about twice that of the naive Fock state code. Moreover, $D_{\mathcal{E}}[\text{gkp}(\bar{n})]$ approaches roughly $\frac{1}{2}Q_{\bar{n}}$ for large γ. In addition, bin codes exhibit better performance with increasing $\bar{n}$ in the $\gamma \lesssim 0.2$ range. This begs the question of how close $D_{\mathcal{E}}$ for gkp and bin comes to $Q_{\bar{n}}$ when one encodes more than a qubit's worth of information and when one utilizes two- or higher mode generalizations of the codes. Such a question is outside the scope of this work, but is being investigated for a subsequent publication.

Since the pure-loss bosonic channel reduces to qubit amplitude damping when restricted to the Fock states $|0\rangle$ and $|1\rangle$, one interesting question to ask is whether the $\bar{n}_{code}$ photons, which so far are concentrated in one mode, will produce a better rate when distributed over the first two Fock states of multiple modes. While comparing single-mode codes to the various discrete-variable codes specialized to protect against qubit amplitude damping (e.g., Refs. [37,67–76]) is outside the scope of this work, we do provide a reference $D_{\mathcal{E}}$ for one specialized code—the four-qubit leung code [37]—that is the smallest known discrete-variable code to protect against one amplitude damping error.[4] Each of the four physical qubits in the leung code,

$$|\pm_{\text{leung}}\rangle = \frac{1}{2}(|00\rangle \pm |11\rangle)^{\otimes 2}, \qquad (3.5)$$

correspond to the first two Fock states of four oscillators. We can then apply $\mathcal{N}_{\gamma}^{\otimes 4}$ (which reduces to amplitude damping within the $|0\rangle, |1\rangle$ Fock state subspace), optimize $F_{\mathcal{E}}$ to yield $\mathcal{E}$, and calculate $D_{\mathcal{E}}$ via the same procedure as with the rest of the codes. A simple calculation yields a total occupation number of $\bar{n}_{\text{leung}} = 2$ photons, which in this case are distributed over the first two Fock states of four modes. The leung code performs similar to the cat, bin, and num ($\bar{n} \leqslant 2$) codes from Fig. 3(a), but is outperformed by the gkp($\bar{n} \leqslant 2$) codes for $\gamma \leqslant 0.35$. This suggests that, at least for intermediate γ and all else being equal, *it is better to encode two photons in a single-mode* gkp

---

[4]Interestingly, this code can also protect against one erasure [77] and be used as an error-detecting code for other errors [78].

state than to spread them out over four modes. The `leung` code is outperformed at almost all γ by all codes considered with $\bar{n}_{\text{code}} \leqslant 5$, but this is not a fair comparison since those codes use more photons.

IV. PRIMER: THE QEC MATRIX

In order to analyze errors for the codes, we consider the quantum error-correction (QEC) conditions [79,80] (see also Ref. [81], Theorem 10.1). The errors we consider are the Kraus operators E_ℓ (1.4), where ℓ denotes the number of photons lost after application of the error. Calculating the effect of the error $E_\ell^\dagger E_{\ell'}$ on the code space yields a 2×2 matrix $\epsilon_{\ell\ell'}$,

$$P_{\text{code}} E_\ell^\dagger E_{\ell'} P_{\text{code}} = \epsilon_{\ell\ell'}^{\text{code}} \in \text{Mat}_{2\times2}. \tag{4.1}$$

We write $\epsilon_{\ell\ell'}^{\text{code}}$ as a superposition of P_{code} and matrices

$$Z_{\text{code}} = |0_{\text{code}}\rangle\langle 0_{\text{code}}| - |1_{\text{code}}\rangle\langle 1_{\text{code}}|, \tag{4.2a}$$

$$X_{\text{code}} = |0_{\text{code}}\rangle\langle 1_{\text{code}}| + |1_{\text{code}}\rangle\langle 0_{\text{code}}|, \tag{4.2b}$$

$$Y_{\text{code}} = |1_{\text{code}}\rangle\langle 0_{\text{code}}| - |0_{\text{code}}\rangle\langle 1_{\text{code}}|. \tag{4.2c}$$

We define our matrix basis as such because both P_{code} and E_ℓ are real for our codes, so the *QEC matrix* ϵ^{code} is real, symmetric, and $2N$ dimensional (with $N \to \infty$ being the dimension of the oscillator). Expanding each 2×2 subblock yields

$$\epsilon_{\ell\ell'}^{\text{code}} = c_{\ell\ell'}^{\text{code}} P_{\text{code}} + x_{\ell\ell'}^{\text{code}} X_{\text{code}} + y_{\ell\ell'}^{\text{code}} Y_{\text{code}} + z_{\ell\ell'}^{\text{code}} Z_{\text{code}}, \tag{4.3}$$

with coefficients denoted by

$$[c,x,y,z]_{\ell\ell'}^{\text{code}} = \tfrac{1}{2}\text{Tr}\{[P,X,Y,Z]_{\text{code}} E_\ell^\dagger E_{\ell'}\}. \tag{4.4}$$

For E_ℓ to be perfectly correctable, one must satisfy the QEC condition

$$\epsilon_{\ell\ell'}^{\text{code}} = c_{\ell\ell'}^{\text{code}} P_{\text{code}} \tag{4.5}$$

(equivalently, $\langle \mu_{\text{code}}|E_\ell^\dagger E_{\ell'}|\nu_{\text{code}}\rangle = c_{\ell\ell'}^{\text{code}}\delta_{\mu\nu}$ for $\mu,\nu \in \{0,1\}$). In words, a correctable error must act as the identity within the code subspace (equivalently, the effect of the error must be the same on both code states). Therefore, the coefficient $c_{\ell\ell'}^{\text{code}}$ represents the correctable part of $\epsilon_{\ell\ell'}^{\text{code}}$ while $\{x_{\ell\ell'}^{\text{code}}, y_{\ell\ell'}^{\text{code}}, z_{\ell\ell'}^{\text{code}}\}$ represent various uncorrectable parts corresponding to bit, phase, and joint bit-phase flips, respectively. Since not all errors E_ℓ can be corrected, we proceed to analyze the magnitude of the uncorrectable parts—the $2N$-dimensional matrix $\epsilon^{\text{code}} - c^{\text{code}}$—with $\epsilon_{\ell\ell'}^{\text{code}}$, $c_{\ell\ell'}^{\text{code}}$ being 2×2 submatrices of ϵ^{code}, c^{code}, respectively.

The QEC matrix block $\epsilon_{\ell\ell'}^{\text{code}}$ can also be interpreted ([81], Fig. 10.5) as a matrix of overlaps between the two error subspaces spanned by $\{E_\ell|\mu_{\text{code}}\rangle\}_{\mu=0}^1$ and $\{E_{\ell'}|\mu_{\text{code}}\rangle\}_{\mu=0}^1$, i.e., the range of $E_\ell P_{\text{code}}$ and $E_{\ell'} P_{\text{code}}$. We call these subspaces $E_\ell P_{\text{code}}$ and $E_{\ell'} P_{\text{code}}$ for short. When no loss events are occurring, the code state undergoes the backaction-induced evolution corresponding to the subspace $E_0 P_{\text{code}}$. As ℓ loss events occur, one's ability to detect them hinges on the orthogonality between $E_0 P_{\text{code}}$ and $E_\ell P_{\text{code}}$, the latter being the space to which a state has gone after losing ℓ photons.

The $\epsilon_{0\ell}^{\text{code}}$ and $\epsilon_{\ell0}^{\text{code}}$ parts of the QEC matrix thus correspond to the ability to distinguish between ℓ losses and no losses, making their satisfaction similar to the satisfaction of the *error-detection conditions* $\delta_\ell^{\text{code}} = P_{\text{code}} E_\ell P_{\text{code}} \propto P_{\text{code}}$ [82]. While the backaction in E_0 makes $\epsilon_{0\ell}^{\text{code}} \neq \delta_\ell^{\text{code}}$, the two converge to each other as $\gamma \to 0$. Since `bin` and `cat` codes satisfy both $\epsilon_{0\ell}^{\text{code}}, \delta_\ell^{\text{code}} \propto P_{\text{code}}$ exactly up to some $\ell \leqslant S$, uncorrectable parts in the QEC matrix blocks $\epsilon_{0\ell}^{\text{code}}$ quantify how well ℓ-photon losses can be detected for those codes.

Uncorrectable parts $\{x_{\ell\ell}^{\text{code}}, y_{\ell\ell}^{\text{code}}, z_{\ell\ell}^{\text{code}}\}$ for "diagonal" errors $E_\ell^\dagger E_\ell$ represent *distortion* of the quantum information within the subspace $E_\ell P_{\text{code}}$ and limit how well one is able to correct the error E_ℓ after detection. Since our code space can become distorted even when there are no loss events, we have to also consider backaction-induced distortion captured by $\epsilon_{00}^{\text{code}}$. The loss event probability distribution is governed by $c_{\ell\ell}^{\text{code}}$ and depends on both γ and $\bar{n}_{\text{code}}$. For a fixed γ and sufficiently large $\bar{n}_{\text{code}}$, we will see that $c_{\ell\ell}^{\text{code}}$ for `cat` (`gkps`) is a Poisson (geometric) distribution having mean $\gamma\bar{n}_{\text{code}}$. In such cases, we can interpret $\gamma\bar{n}_{\text{code}}$ as the average number of photons lost, and only when $\gamma\bar{n}_{\text{code}} \ll 1$ can we say that E_0 is the most likely error for a code.

V. CAT CODES

`cat` code logical states are coherent states projected onto subspaces of occupation number modulo $2(S + 1)$:

$$|\mu_{\text{cat}}\rangle = \frac{\Pi_{(S+1)\mu}|\alpha\rangle}{\sqrt{N_\alpha^{(S+1)\mu}}}, \tag{5.1}$$

with α real (for simplicity), $\mu \in \{0,1\}$, and normalization

$$N_\alpha^{(S+1)\mu} = \langle\alpha|\Pi_{(S+1)\mu}|\alpha\rangle. \tag{5.2}$$

The projections $\{\Pi_0, \Pi_{S+1}\}$ used to define the code states belong to the family (for $r \in \{0,1,\cdots,2S+1\}$[5]

$$\Pi_r = \sum_{n=0}^{\infty} |2n(S+1)+r\rangle\langle 2n(S+1)+r|. \tag{5.3}$$

In the large α limit, i.e., when

$$2\alpha\sin\left(\frac{\pi}{S+1}\right) \gg 1, \tag{5.4}$$

`cat`-code states become equal superpositions of coherent states $\{|\alpha e^{i\frac{\pi}{S+1}k}\rangle\}_{k=0}^{2(S+1)-1}$ distributed equidistantly on a circle of radius α in phase space. In that limit, the seemingly bothersome normalization factors approach the same constant, while when $\alpha \lesssim S$, they become distinct in order to account for the various overlaps between the coherent states. Expressing the normalization factor in terms of such overlaps ([83], Eq. (3.22)), we

---

[5]One can also consider shifted `cat` and `bin` codes by picking subspaces of Fock states $n = s$ and $n = S + 1 + s$ modulo $2(S + 1)$ for shift parameter $s \in \{0,1,\ldots,S\}$. Sampling some of the "sweet spots" for shifted `cat` codes [46] did not alter the qualitative behavior of `cat` relative to the other codes.

have

$$N_\alpha^{(S+1)\mu} = \frac{1}{2(S+1)} \sum_{s=0}^{2S+1} (-1)^{\mu s} \langle \alpha | \alpha e^{i\frac{\pi}{S+1}s} \rangle. \quad (5.5)$$

Since the cat codes which produce the optimal $F_\mathcal{E}$ have $\alpha \lesssim S$, we have to consider the factors $N_\alpha^{(S+1)\mu}$ at intermediate α in order to explain Fig. 2.

Because of properties of coherent states and for $\alpha \neq 0$, cat logical states satisfy

$$(\hat{a}/\alpha)^{2(S+1)} |\mu_\text{cat}\rangle = |\mu_\text{cat}\rangle, \quad (5.6)$$

and cat(S) [and bin(S), as we shall see] can detect exactly S photon loss events using the check operator

$$C_\text{cat} = C_\text{bin} = \exp\left(i\frac{2\pi\hat{n}}{S+1}\right). \quad (5.7)$$

Its square root, $\exp(i\frac{\pi\hat{n}}{S+1})$, makes for a logical Z-operator within cat and bin code subspaces. We call this a check operator (and not a stabilizer) since it has other eigenvalues which are not modulus one and is not part of a set of commuting such operators which is used to construct the code states.

The $S = 0$ cat code was proposed in Ref. [13] for coherent-state quantum computation [84–87]. This code cannot detect loss events since losing a photon causes a logical bit flip, but it of course can be concatenated with, e.g., a bit-flip code [85]. The cat($S = 1,2$) codes were studied first in Refs. [15,27], followed by investigations into cat($S \geqslant 2$) and qudit extensions [46,88,89]. There are several theoretical and experimental schemes for cat state preparation [90]. Schemes designed to protect against loss errors [29] (for $S = 1$) and backaction errors [91] (for $S = 0$), respectively, were realized using microwave cavities coupled to superconducting qubits.

A. A simple example

As an example of the utility of cat codes, consider the simplest nontrivial code family cat($S = 1$) whose logical states are

$$|\mu_\text{cat}^{S=1}\rangle = \frac{|\alpha\rangle + |-\alpha\rangle + (-1)^\mu(|i\alpha\rangle + |-i\alpha\rangle)}{4\sqrt{2[\cosh\alpha^2 + (-1)^\mu\cos\alpha^2]}}. \quad (5.8)$$

In the Fock basis, the above $|0_\text{cat}\rangle$ is a superposition of Fock states $|0\rangle, |4\rangle, |8\rangle, \ldots$ while $|1_\text{cat}\rangle$ is supported by $|2\rangle, |6\rangle, |10\rangle, \ldots$. Therefore, there are exactly $S = 1$ Fock states separating the Fock states supporting $|0_\text{cat}\rangle$ from those supporting $|1_\text{cat}\rangle$, so we call S the *spacing* between logical states. Because of this spacing, $\epsilon_{01}^\text{cat} = 0$—one loss event is always detectable. However, such an event is not always *correctable* since ϵ_{11}^cat contains uncorrectable parts at generic values of α.

In general, $\epsilon_{\ell\ell}^\text{cat} \neq c_{\ell\ell}P_\text{cat}$ because of the backaction term $(1 - \gamma)^{\hat{n}/2}$ in E_ℓ. We will see that uncorrectable parts in $\epsilon_{\ell\ell}^\text{cat}$ are well suppressed as $\alpha \to \infty$, but in that limit the code consists of large-amplitude coherent states and there is a high chance of losing more than one photon (i.e., uncorrectable parts in ϵ_{02}^cat become very large). Therefore, for more general codes with a given spacing S and loss rate γ, there is an optimal or "sweet spot" value $\alpha = \alpha_\star(\gamma)$ that balances the backaction with the loss errors. This is exactly what we see
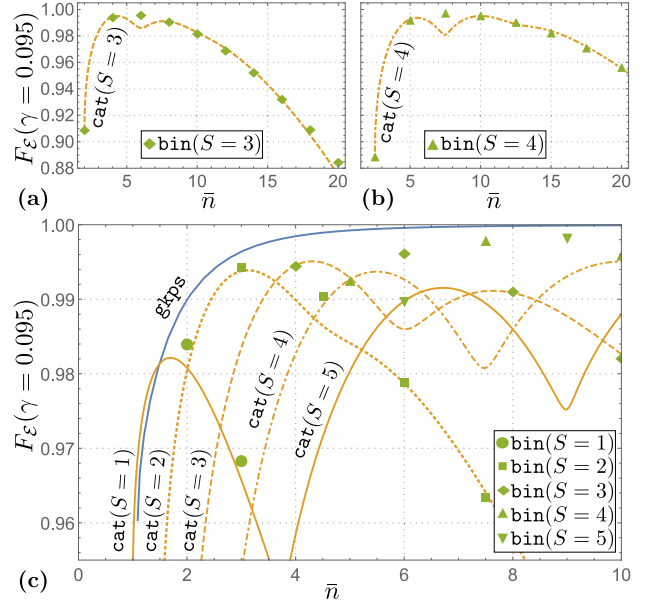

FIG. 4. Channel fidelity $F_\mathcal{E}(\gamma = 0.095)$ (1.11) vs mean occupation number $\bar{n}$ (2.1) for cat and bin at spacing (a) $S = 3$ and (b) $S = 4$. Note that cat(S) depends continuously on $\bar{n}$ and bin exists only at discrete values. For a given S, cat codes perform best at specific $\bar{n}$, corresponding to "sweet-spot" values of $\alpha = \alpha_\star$. However, due to the oscillatory nature of the errors in cat($S \gtrsim 3$), the codes also develop troughs in $F_\mathcal{E}$. On the other hand, bin(S) performs similar to cat(S) at small and large $\bar{n}$ but does not suffer from troughs at intermediate $\bar{n}$. (c) Plots of cat and bin for $S \in \{1,2,3,4,5\}$ along with gkps, which turns out to outperform both cat and bin and increase monotonically with $\bar{n}$. Similar trends are observed for smaller γ.

in Figs. 4(a) and 4(c), where $F_\mathcal{E}$ is plotted versus $\bar{n}_\text{cat}$ for various cat(S) and at a fixed γ. We can see that $F_\mathcal{E}[\text{cat}(S)]$ is maximized at certain $\bar{n}_\text{cat}$ [which in turn corresponds to a certain $\alpha_\star(\gamma)$] and decays with sufficiently large α. Note that there can be multiple $\alpha_\star$ for a given S. By contrast, gkps code fidelities increase monotonically with $\bar{n}_\text{gkps}$.

We now show how the sweet spot can be analytically determined for cat($S = 1$) and in the limit $\gamma \to 0$, discussing general γ in the next subsection. Recalling the form of the Kraus operators (1.5) for small γ, to suppress distortion due to backaction, one has to make sure that both code states have the same occupation number:

$$\delta\bar{n}_\text{cat} = \tfrac{1}{2}\text{Tr}\{Z_\text{cat}\hat{n}\} = 0. \quad (5.9)$$

While $\delta\hat{n}_\text{cat} \to 0$ as $\alpha \to \infty$, there are certain fine-tuned α at which the occupation numbers of the two logical states also coincide due to the oscillatory nature of the normalization factors N_α (5.5). Solving the above equation using $|\mu_\text{cat}^{S=1}\rangle$ (5.8) yields the transcendental equation

$$\tan\alpha^2 = -\tanh\alpha^2, \quad (5.10)$$

whose solution is $\alpha_\star(\gamma \to 0) \approx 1.538$ (corresponding to $\bar{n}_\text{code} \approx 2.324$). Thus, in the small γ limit and at this fine-tuned $\alpha_\star$, a single loss event is both detectable and correctable. This was about the value used in a recent cat-code error-correction experiment [29].

B. QEC matrix for cat codes

Let us briefly elaborate on the discussion regarding sweet spots using the QEC matrix for cat codes for finite (but still small) γ. Because of the spacing of the codes, QEC matrix subblocks $\epsilon_{\ell\ell'}$ (4.1) contain uncorrectable parts only for certain values of ℓ, ℓ'. Let us first study the distortion due to errors $\epsilon_{\ell\ell}^{\text{cat}}$, which we can easily calculate [46,92] by using the representation (5.1) of the cat states, observing that $\Pi_r \hat{a} = \hat{a} \Pi_{r+1}$, and using Eq. (D1a):

$$\epsilon_{\ell\ell}^{\text{cat}} = \frac{(\gamma\alpha^2)^\ell e^{-\gamma\alpha^2}}{\ell!} \begin{pmatrix} \frac{N_{\alpha\sqrt{1-\gamma}}^{-\ell}}{N_\alpha^0} & 0 \\ 0 & \frac{N_{\alpha\sqrt{1-\gamma}}^{S+1-\ell}}{N_\alpha^{S+1}} \end{pmatrix}, \quad (5.11)$$

where $N_{\alpha\sqrt{1-\gamma}}$ are damped versions of the normalization factors in Eq. (5.5). Expansion yields the correctable part (4.4),

$$c_{\ell\ell}^{\text{cat}} \equiv \frac{1}{2} \frac{(\gamma\alpha^2)^\ell e^{-\gamma\alpha^2}}{\ell!} \left(\frac{N_{\alpha\sqrt{1-\gamma}}^{-\ell}}{N_\alpha^0} + \frac{N_{\alpha\sqrt{1-\gamma}}^{S+1-\ell}}{N_\alpha^{S+1}} \right), \quad (5.12)$$

and only one uncorrectable part (4.4),

$$z_{\ell\ell}^{\text{cat}} \equiv \frac{1}{2} \frac{(\gamma\alpha^2)^\ell e^{-\gamma\alpha^2}}{\ell!} \left(\frac{N_{\alpha\sqrt{1-\gamma}}^{-\ell}}{N_\alpha^0} - \frac{N_{\alpha\sqrt{1-\gamma}}^{S+1-\ell}}{N_\alpha^{S+1}} \right). \quad (5.13)$$

The correctable part represents the probability of losing ℓ photons. This distribution is Poisson in the large-α limit, in which the ratios of normalization factors inside the parentheses both go to one exponentially with $-(1-\gamma)\alpha^2$. The uncorrectable part $z_{\ell\ell}^{\text{cat}}$ represents the inability to correct against ℓ loss events. It is suppressed as $\alpha \to \infty$, but is also zero at certain "sweet spots" $\alpha_\star$, which we discuss using another example.

Consider cat($S = 2$) for $\gamma \leqslant 0.0124$ and $\bar{n}_{\text{cat}} \leqslant 5$. This is a case when $\gamma\bar{n} \ll 1$ and so we only have to consider distortion due to backaction $\epsilon_{00}^{\text{cat}}$. We see from Table IV that the cat code which achieves the highest $F_{\mathcal{E}}$ out of all cat codes with $\bar{n} \leqslant 5$ is cat($\alpha = 1.739, S = 2$). This is exactly the sweet spot at which the distortion z_{00}^{cat} is approximately zero: $\alpha_\star(\gamma = 0.005) \approx 1.739$. More generally, $\alpha_\star(\gamma)$ at which $z_{00}^{\text{cat}} \approx 0$ decreases as $\gamma \to 0$ to the value $\alpha_\star(\gamma = 0) \approx 1.737$. [The reason that cat($\alpha = 1.739$) and not cat($\alpha = 1.737$) is the optimal code at $\gamma < 0.005$ is because the resolution in our sampling of α is not sufficient to resolve the difference.] As mentioned in our discussion of cat($S = 1$) in the previous subsection, in the limit $\gamma \to 0$, z_{00}^{cat} becomes proportional to

$$\delta\bar{n}_{\text{cat}} = \frac{1}{2} \text{Tr}\{Z_{\text{cat}} \hat{n}\} = \frac{\alpha^2}{2} \left(\frac{N_\alpha^{-1}}{N_\alpha^0} - \frac{N_\alpha^S}{N_\alpha^{S+1}} \right). \quad (5.14)$$

This $\delta\bar{n}_{\text{cat}}$ tells us how well we are able to correct against distortion due to no loss events, the dominant error when $\gamma\bar{n}_{\text{cat}} \ll 1$. It is zero at exactly the "sweet spot" $\alpha_\star(\gamma = 0)$, and we can similarly relate $\alpha_\star$ to $\delta\bar{n}_{\text{cat}}$ for cat at other values of S. Having only covered cat($S = 2$), we refer the reader to Ref. [46] for such calculations.

We have seen that at $\alpha < S$, backaction-induced errors are not suppressed because α is not sufficiently large [see Eq. (5.4)]. So why are cat codes with high values of α not optimal? This is because for $\alpha \approx S$, the fraction $\gamma\bar{n}_{\text{cat}} \approx \gamma\alpha^2$ of photons lost yields a large probability of losing $S + 1$ photons, an uncorrectable error. More technically, recall that

due to spacing S, the effect of $\ell \leqslant S$ loss events is zero, $\epsilon_{0\ell}^{\text{cat}} = 0$. However, the first uncorrectable loss at $\ell = S + 1$ produces an error that scales unfavorably with α, prohibiting α from getting too large. A calculation yields

$$\epsilon_{0,S+1}^{\text{cat}} = \frac{(\gamma\alpha^2)^{\frac{S+1}{2}} e^{-\gamma\alpha^2}}{\sqrt{(S+1)! N_\alpha^0 N_\alpha^{S+1}}} \begin{pmatrix} 0 & N_{\alpha\sqrt{1-\gamma}}^0 \\ N_{\alpha\sqrt{1-\gamma}}^{S+1} & 0 \end{pmatrix}. \quad (5.15)$$

Once again, the ratios of normalization factors go to one as $\alpha \to \infty$. In that limit, this error becomes proportional to a pure bit flip, $\epsilon_{0,S+1}^{\text{cat}} \sim x_{0,S+1}^{\text{cat}} X_{\text{cat}}$ (where we use the mathematician's definition of "$\sim$"), with

$$x_{0,S+1}^{\text{cat}} \sim \sqrt{c_{00}^{\text{cat}} c_{S+1,S+1}^{\text{cat}}} \quad (5.16)$$

independently of γ. In other words, given even a very small γ, the above equation is still satisfied for a sufficiently large α. We claim that this is the worst possible scaling, making such errors completely undetectable (and therefore uncorrectable). Recall that since the QEC matrix (4.1) is positive semidefinite ($\epsilon \geqslant 0$), in the absence of other errors for a given ℓ, ℓ', all bit-flip errors are bounded by $x_{\ell\ell'} \leqslant \sqrt{c_{\ell\ell} c_{\ell'\ell'}}$. Ideally, we would like to have no error ($x_{\ell\ell'} = 0$), and in the worst case the inequality is saturated ($x_{\ell\ell'} = \sqrt{c_{\ell\ell} c_{\ell'\ell'}}$). In Eq. (5.16), we see that the inequality is saturated as $\alpha \to \infty$. Because the code states are eigenstates of $\hat{a}^{2(S+1)}$, the behavior upon 0 to $2(S+1) - 1$ losses repeats itself after $2(S + 1)$ losses. More generally, a bit flip error of similar intensity occurs at any $\ell - \ell' = S + 1$ modulo $2(S + 1)$. Thus, as $\alpha \to \infty$, the QEC matrix ϵ develops sparse but large bit-flip error entries at each such ℓ, ℓ', implying a large probability of undetectable errors. Therefore, cat codes should work best for values of α at which the probability of losing $\ell = S + 1$ photons is small.

The reader should by now see that, for a given γ and S, a cat code performs optimally at specific $\alpha_\star(\gamma)$. But how does one determine the optimal S? We do not claim to answer this question fully, only noting that (A) the optimal S depends on γ and (B) our energy constraints limit the selection of $\alpha_\star$ to choose from, which in turn limit the selection of S. For an example of (A), notice in Fig. 4(c) that $\bar{n}_{\text{cat}}(\alpha_\star, S = 1) < 2$ while $\bar{n}_{\text{cat}}(\alpha_\star, S > 1) > 2$, so the highest spacing achievable with a sweet spot is $S = 1$. Similarly for (B), for $\bar{n}_{\text{cat}} \leqslant 5$, there are three available $\alpha_\star$, one for each $S = 1, 2, 3$. At $\gamma = 0.095$, the $F_{\mathcal{E}}$ is highest for $S = 3$ [as shown in Fig. 4(c)] while at smaller γ, $S = 2$ is optimal (see Table IV). In summary, cat codes cannot perform well at sufficiently large α and instead are optimal for specific values of $\alpha_\star \lesssim S$.

VI. BINOMIAL CODES

In terms of Fock states, the logical states (defined here in the X_{bin} basis [23]) are

$$|\mu_{\text{bin}}\rangle = \frac{1}{\sqrt{2^{N+1}}} \sum_{m=0}^{N+1} (-1)^{\mu m} \sqrt{\binom{N+1}{m}} |(S+1)m\rangle. \quad (6.1)$$

Their mean occupation number (2.1) is

$$\bar{n}_{\text{bin}} = \frac{1}{2}(N+1)(S+1). \quad (6.2)$$

The non-negative integer N governs the order in γ to which dephasing errors $\{\hat{n}^n\}_{n=0}^N$ can be corrected exactly and is similar

to α in cat codes. The spacing S is the same as that of cat codes since bin(S) and cat(S) are spanned by the same Fock states.[5] Therefore, the ability for bin(S) to perfectly detect $\ell \leqslant S$ loss events using the check operator (5.7) is identical to that of cat(S): $\epsilon_{0\ell}^{\text{bin}} = P_{\text{bin}} E_0^\dagger E_\ell P_{\text{bin}} = 0$ in the language of the QEC matrix (4.1). The difference lies in the other parameter N, the discrete analog of α for cat. Recall that for cat(S), the limit $\alpha \to \infty$ exponentially suppresses all uncorrectable parts in $\epsilon_{\ell\ell}^{\text{cat}}$ for all ℓ. Similarly, in bin(S), tuning N allows one to suppress distortion due to loss events up to a given order $O(\gamma^N)$. Therefore, it should be no surprise that for a fixed S,

$$\text{bin}(N \to \infty, S) \sim \text{cat}(\alpha \to \infty, S). \tag{6.3}$$

While the two codes coincide in this limit, the advantage of bin codes is that, unlike cat, the suppression of distortion in $\epsilon_{\ell\ell}^{\text{bin}}$ occurs *without oscillations*. While the oscillatory nature of the normalization factors in $\epsilon_{\ell\ell}^{\text{cat}(S)}$ (5.11) allows for peaks as well as troughs in $F_\mathcal{E}$ versus $\bar{n}$ [shown in Figs. 4(a)–4(c)], there are no such oscillations in $\epsilon_{\ell\ell}^{\text{bin}(S)}$. Note that this difference only shows up at $S \gtrsim 3$ since only then are there significant oscillations in $F_\mathcal{E}(\text{cat})$. While cat and bin perform about the same at smaller S, there is an intermediate $\bar{n}$ regime for larger S at which cat underperforms (due to being at a trough between two sweet spots $\alpha_\star$) while bin continues to improve. For example, observe the difference between cat and bin at $S = 3$ in Fig. 4(a): $F_\mathcal{E}(\text{bin}) \approx F_\mathcal{E}(\text{cat})$ for all $\bar{n} \leqslant 15$ except at $\bar{n} \approx 6$.

In Sec. VI A, we delve into the performance of bin codes from Figs. 2 and 4. In Sec. VI B, we comment on their performance with no energy constraints. Switching gears in Sec. VI C, we study their structure. We show that qubit bin codes are spin-coherent states embedded into the $\{|(S+1)m\rangle\}_{m=0}^{N+1}$ subspace of the oscillator. In an alternative characterization in Appendix C, we link bin codes to discrete-variable bit-flip codes. These formulations extend to other sets of codes, summarized in Fig. 7(a), and allow one to construct bin check operators for dephasing errors. In Sec. VI D, we introduce a scheme to detect and correct errors in bin using the check operators from before.

A. QEC matrix for binomial codes

Let us fix S and compare bin(S) to cat(S). A bin(N, S) code protects against loss errors $\hat{a}^{\ell \leqslant S}$ (due to spacing S) and dephasing errors $\hat{n}^{n \leqslant N}$ (due to the nature of the binomial distribution). Since loss operators E_ℓ (1.4) consist of superpositions of the two errors, we can readily read off the leading order in γ for which there are uncorrectable parts in $\epsilon_{0\ell}^{\text{bin}}$—$O(\gamma^{S+1})$—and distortion matrices $\epsilon_{\ell\ell}^{\text{bin}}$—$O(\gamma^{N+1})$. However, that is not the whole story.

We know that both bin and cat suppress all distortion errors $z_{\ell\ell}$ with increasing $\bar{n}$. The backaction distortion $z_{\ell\ell}^{\text{bin}}$ does not oscillate with $\bar{n}$ (as opposed to $z_{\ell\ell}^{\text{cat}}$ oscillation with α) and yields a quicker suppression of errors than $z_{\ell\ell}^{\text{cat}}$. We use the basis of positive and negative superpositions $|\pm_{\text{bin}}\rangle$ of the bin states (6.1) to calculate it,

$$P_{\text{bin}} = |+_{\text{bin}}\rangle\langle+_{\text{bin}}| + |-_{\text{bin}}\rangle\langle-_{\text{bin}}|, \tag{6.4}$$

in order to have the backaction-induced errors be along the z axis and match $z_{\ell\ell}^{\text{cat}}$. Note that this amounts to the Z_{bin} basis

of the original paper [23]. The respective probabilities of no loss and backaction distortion can be concisely expressed,

$$c_{\ell\ell}^{\text{bin}} = \frac{\gamma^\ell}{\ell!} \frac{d^\ell}{dx^\ell} \left(\frac{1 + x^{S+1}}{2} \right)^{N+1} \Bigg|_{x=1-\gamma}, \tag{6.5a}$$

$$z_{\ell\ell}^{\text{bin}} = \frac{\gamma^\ell}{\ell!} \frac{d^\ell}{dx^\ell} \left(\frac{1 - x^{S+1}}{2} \right)^{N+1} \Bigg|_{x=1-\gamma}. \tag{6.5b}$$

For $\ell = 0$, the above should be compared to c_{00}^{cat} (5.12) and z_{00}^{cat} (5.13). Clearly, z_{00}^{bin} does not oscillate versus N. While the above is still difficult to analyze analytically for $\ell > 0$, we see numerically that there are no oscillations in $\bar{n}$ of the fidelity of bin(S), leading to certain regimes of $\bar{n}$ at which bin outperforms cat for a given γ. Heuristically, as $\bar{n} \to \infty$, $z_{\ell\ell}^{\text{bin}(S)} \to 0$ order by order in γ while $z_{\ell\ell}^{\text{cat}(S)} \to 0$ as a damped oscillating function with damping coefficient $(1 - \gamma)\alpha^2$. The latter limit turns out to be less controlled, leading to detrimental oscillations in $F_\mathcal{E}^{\text{cat}(S)}$. For example, we plot the uncorrectable parts of $\epsilon_{\ell\ell}^{\text{cat}}$ and $\epsilon_{\ell\ell}^{\text{bin}}$ for $\gamma = 0.095$, $\bar{n}_{\text{code}} \approx 6$, and $S = 3$ in Figs. 5(a) and 5(b), respectively. While the uncorrectable parts $x_{0\ell}^{\text{bin}}, x_{0\ell}^{\text{cat}}$ are comparable in magnitude, one can see that $z_{\ell\ell}^{\text{bin}}$ is visibly less than $z_{\ell\ell}^{\text{cat}}$. However, this effect is most prominent only when E_0 is the only dominant error ($\gamma\bar{n}_{\text{code}} \ll 1$) and when cat oscillations begin to have a detrimental effect ($S \gtrsim 3$ and $\bar{n}_{\text{code}} \geqslant 5$). Inside those regimes, we can see that bin breaks away from cat [see Figs. 4(a)–4(c) and insets in Figs. 2(b) and 2(c)], while outside of those regions, the two codes perform quite similarly.

Another advantage of bin codes manifests itself at large $\bar{n}_{\text{bin}}$. Studying $x_{0\ell}^{\text{bin}}$ and $y_{0\ell}^{\text{bin}}$ is quite difficult, so we explain the advantage by studying subspaces that are mapped to under errors. For cat codes, the undetectable error $\hat{a}^{S+1}$ maps the code *exactly* to the code subspace, $\hat{a}^{S+1} P_{\text{cat}} \propto P_{\text{cat}}$. For bin codes, the mapping is to a subspace that has a component orthogonal to the code space. Quantum information in this component (which may only be one dimensional) can then be mapped back to the code space, yielding an extra layer of approximate error correction. The same is true for $\ell > S + 1$ as long as N is sufficiently high. We consider two examples of this effect, one known and one new.

Let us consider the action of the undetectable error $\hat{a}^2$ on bin($N = 1, S = 1$) and cat($\alpha \gg 1, S = 1$) (α is large only for simplicity; its value is irrelevant to the key point). The logical states $|+_{\text{bin}}\rangle \propto |0\rangle + |4\rangle$ and $|-_{\text{bin}}\rangle = |2\rangle$ are mapped to states $|2\rangle$ and $|0\rangle$, respectively. We see that the latter error state, $\hat{a}^2|-_{\text{bin}}\rangle \propto |0\rangle$, overlaps with $|0\rangle - |4\rangle$, a state orthogonal to the code space. One can thus add an extra Kraus operator to the recovery that maps any quantum information in this extra error subspace back to the code space. This cannot be done for cat, where the logical states $|+_{\text{cat}}\rangle \propto |\alpha\rangle + |-\alpha\rangle$ and $|-_{\text{cat}}\rangle \propto |i\alpha\rangle + |-i\alpha\rangle$ are mapped to $\pm|\pm_{\text{cat}}\rangle$ under $\hat{a}^2$ (recall that $\hat{a}|\alpha\rangle = \alpha|\alpha\rangle$), yielding a completely uncorrectable phase flip. For bin($N = 1, S = 1$), this extra one-dimensional subspace is used to correct from the leading-order backaction error [23]. However, there are enough such extra subspaces for sufficiently high N to correct for both backaction (up to γ^N) *and* some loss errors $\hat{a}^{\ell \geqslant S+1}$.
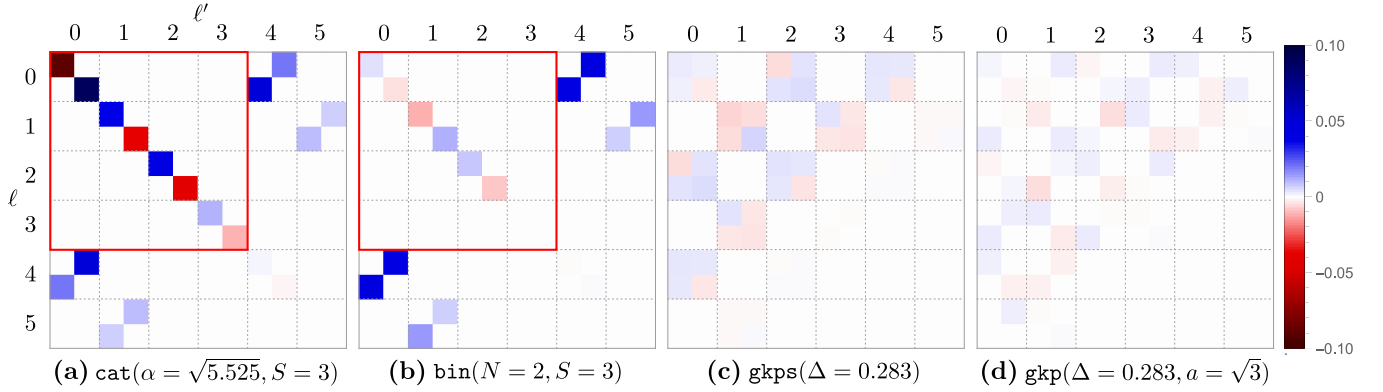
(a) $\mathtt{cat}(\alpha = \sqrt{5.525}, S = 3)$ **(b)** $\mathtt{bin}(N = 2, S = 3)$ **(c)** $\mathtt{gkps}(\Delta = 0.283)$ **(d)** $\mathtt{gkp}(\Delta = 0.283, a = \sqrt{3})$

FIG. 5. Uncorrectable parts of the QEC matrices, $\epsilon^{\mathrm{code}} - c^{\mathrm{code}}$ (see Sec. IV), for $\mathtt{code}$ being (a) $\mathtt{cat}$ (5.1), (b) $\mathtt{bin}$ (6.1), (c) $\mathtt{gkps}$ (7.8), and (d) $\mathtt{gkp}$ (B4) with parameters such that $\bar{n}_{\mathrm{code}} \approx 6$ for all codes and given $\gamma = 0.095$. Recall that ϵ^{code} is a block matrix consisting of 2×2 matrices $\epsilon_{\ell\ell'}^{\mathrm{code}}$ that quantify the overlap between error subspaces $E_\ell P_{\mathrm{code}}$ and $E_{\ell'} P_{\mathrm{code}}$. These 2×2 matrices are delineated by dotted lines. The four entries in $\epsilon_{\ell\ell'}^{\mathrm{code}}$ are presented as colored squares; note that $\epsilon_{\ell\ell'}^{\mathrm{code}}$ has no imaginary part (all $E_\ell P_{\mathrm{code}}$ are real). From Fig. 4(c), we see that $F_{\mathcal{E}}(\mathtt{cat}) < F_{\mathcal{E}}(\mathtt{bin}) < F_{\mathcal{E}}(\mathtt{gkps})$ at $\bar{n}_{\mathrm{code}} \approx 6$, and the above QEC plots nicely corroborate that order of performance. Since $\mathtt{cat}$ and $\mathtt{bin}$ have spacing $S = 3$, there are no off-diagonal errors for $\ell \leqslant 3$ (inside the red square). However, both codes suffer from distortion on the diagonal portions of $\epsilon_{\ell\ell'}^{\mathrm{code}}$, with $\mathtt{bin}$ suffering noticeably less at this particular γ. On the other hand, $\mathtt{gkps}$ (square lattice) and $\mathtt{gkp}$ (shifted hexagonal lattice) barely suffer from *any* noticeable errors. Since $\mathtt{gkps}$ codes have spacing $S = 1$ due to a symmetry of their underlying lattice, $\epsilon_{\ell\ell'}$ does not contain off-diagonal errors for $\ell \leqslant 1$ (more generally, for $\ell - \ell'$ odd). Shifted $\mathtt{gkp}$ codes do not have that symmetry, but perform comparably.

Now consider $\mathtt{bin}(N = 4, S = 1)$. Recall that a $\mathtt{bin}(N = 4)$ protects from backaction up to γ^4, so our calculations are only to that order. While the code exactly detects only one loss E_1, it turns out there is an extra subspace allowing for approximate correction of E_2 and even E_3. The two-dimensional code subspace $\mathrm{Span}\{|\pm_{\mathtt{bin}}\rangle\}$ is supported on the six-dimensional Fock subspace

$$\mathcal{F}_0 = \mathrm{Span}\{|0\rangle, |2\rangle, |4\rangle, |6\rangle, |8\rangle, |10\rangle\}. \tag{6.6}$$

Another two-dimensional subspace is reserved for correcting backaction E_0. This leaves an extra two dimensions $\mathrm{Span}\{Q_0 E_2 |\pm_{\mathtt{bin}}\rangle\} \subset \mathcal{H}_{\mathrm{no\text{-}loss}}$ for approximately correcting the loss error E_2, where the projection Q_0 removes any overlap with the code space and the subspace used for correcting backaction. Indeed, one can add an extra isometry mapping such error states

$$Q_0 E_2 |+_{\mathtt{bin}}\rangle \propto \sqrt{2}\eta^2|2\rangle - (1+\eta^2)|6\rangle + \sqrt{10}|10\rangle, \tag{6.7a}$$

$$Q_0 E_2 |-_{\mathtt{bin}}\rangle \propto \sqrt{10}\eta^2|0\rangle - (1+\eta^2)|4\rangle + \sqrt{2}|8\rangle, \tag{6.7b}$$

back into the code space ($\eta \equiv 1 - \gamma$).

Similarly, this code can also approximately correct the next error E_3. The relevant Fock subspace is now

$$\mathcal{F}_1 = \mathrm{Span}\{|1\rangle, |3\rangle, |5\rangle, |7\rangle, |9\rangle\}. \tag{6.8}$$

The two-dimensional subspace $\mathrm{Span}\{E_1|\pm_{\mathtt{bin}}\rangle\}$ is devoted to correcting E_1, leaving three extra dimensions. Two of those dimensions are then used to correct against E_3. Letting Q_1 be the projection on the remaining three-dimensional subspace $\mathcal{F}_1/\mathrm{Span}\{E_1|\pm_{\mathtt{bin}}\rangle\}$, one can construct a mapping from the extra error subspace $\mathrm{Span}\{Q_1 E_3 |\pm_{\mathtt{bin}}\rangle\}$ back to the code space.

B. Removing energy constraints

We have numerically investigated $\bar{n}_{\mathtt{bin}} \geqslant 10$ in order to see whether a certain direction in the N, S parameter space produces increasing $F_{\mathcal{E}}$ with increasing $\bar{n}_{\mathtt{bin}}$ (6.2). While it is unlikely that $F_{\mathcal{E}} \to 1$ as $\bar{n} \to \infty$ for *any* single-mode code, we have numerical evidence showing that $F_{\mathcal{E}}^{\mathtt{bin}(N)}$ for certain $S = \xi N$ monotonically increases to some value $F_{\mathcal{E}}^{\mathtt{bin}(\infty)} < 1$, with $\xi \geqslant 1$ dependent on γ. For example, as shown in Fig. 6, the limit $S \approx 2N \to \infty$ does seem to be giving us monotonically increasing $F_{\mathcal{E}}$; we have verified this monotonic increase for $N \leqslant 14$ and $S \leqslant 36$ at $\gamma = 0.1$ but it could very well be that the curves eventually decrease for sufficiently high N, S.
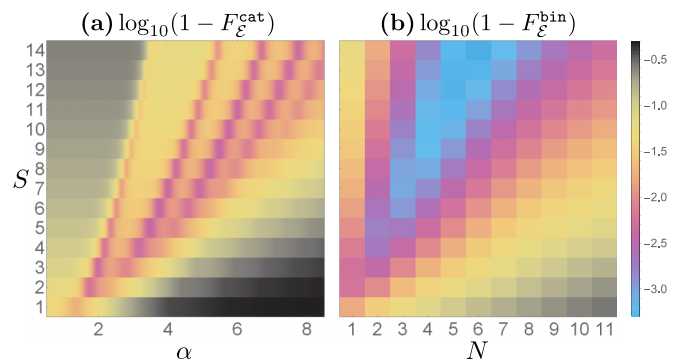

(a) $\log_{10}(1 - F_{\mathcal{E}}^{\mathtt{cat}})$ **(b)** $\log_{10}(1 - F_{\mathcal{E}}^{\mathtt{bin}})$

FIG. 6. Log_{10} plot of $1 - F_{\mathcal{E}}(\gamma = 0.095)$ (1.11) vs (a) $\mathtt{cat}$ code parameters S and α and (b) $\mathtt{bin}$ code parameters S and N (cf. Fig. 1 in Ref. [23]). For a given S, $\mathtt{cat}$ achieves the best performance (purple) at multiple S-dependent sweet spots $\alpha_\star$. While $F_{\mathcal{E}}(\alpha_\star)$ for $\mathtt{cat}(S)$ does not increase with increasing S for the values we have sampled, performance of $\mathtt{bin}(N, S \approx 2N)$ clearly does (cyan). There is thus reason to believe that $\mathtt{bin}$ outperforms $\mathtt{cat}$ when no energy constraints are imposed.

The performance improvement of `bin` codes at large $\bar{n}_{\text{bin}}$ can be attributed to the advantage of having extra error subspaces, discussed in the latter portion of Sec. VI A. To quantify this advantage, one can cook up a recovery procedure consisting of two sets of isometries. The *first-level set* of recovery isometries maps the *correctable* ($\ell \in \{0, 1 \ldots, S\}$) error-subspace code states $E_\ell |\pm_{\text{bin}}\rangle$ back to the code space. This part is similar to the `cat` code recovery scheme from Ref. [46] and to the `bin` scheme in Sec. VI D. The *second-level set* consists of isometries mapping extra error subspaces $Q_\ell E_{\ell+S+1} |\pm_{\text{bin}}\rangle$ back into the code space, where Q_ℓ project out the first-level error subspaces $\text{Span}\{E_\ell |\pm_{\text{bin}}\rangle\}$ (and, in the case of $\ell = 0$, the code space as well). Of course, such a multilevel recovery can be extended to three and more levels. When implemented, the two-level recovery yields a similar scaling with N, S as the optimal recovery in Fig. 6(b) for the N, S we were able to sample. Nevertheless, it does not explain why S increases faster than N for the best codes in that figure. While we have shown that increasing both S and N allows `bin` to correct (at least) approximately for more loss events, the intricate choice of which parameter to increase faster to give the optimal fidelity remains an interesting open question.

C. Relation to spin-coherent states

We characterize all single-mode binomial codes, two-mode binomial codes (closely related to `noon` codes [22]), and multiqubit permutation invariant codes [73] using irreducible representations (irreps) of the Lie algebra $\mathfrak{su}(2)$. Interestingly, `bin`$(S = 0)$ were of interest to the quantum optical community due to their sub-Poissonian distribution [93] (see also Ref. [94], Ch. 5), and the connection to $\mathfrak{su}(2)$ was noticed first back then [95,96].

Consider a spin-J consisting of $2J + 1$ levels and define the standard spin operators

$$J_z \equiv \sum_{m=0}^{2J} (m - J)|J, m - J\rangle\langle J, m - J|, \qquad (6.9)$$

and similarly J_x and J_y ([97], Ch. 7). We can then define its rotated version $J_r(\theta, \phi) = R_{\theta,\phi} J_z R_{\theta,\phi}^\dagger$, where $R_{\theta,\phi}$ is a rotation by azimuthal angle $\theta \in [0, \pi]$ and polar angle $\phi \in [0, 2\pi)$ in the spherical coordinate system parameterizing the spin's Bloch sphere. For each $\{\theta, \phi\}$, $J_r(\theta, \phi)$ has eigenstates $|\theta, \phi\rangle_J$ with eigenvalue J. These are called the $\mathfrak{su}(2)$ or spin-coherent states [98] (see also Ref. [99]):

$$|\theta, \phi\rangle_J = \sum_{m=0}^{2J} \frac{\left(e^{i\phi} \tan\frac{\theta}{2}\right)^m}{\left(1 + \tan^2\frac{\theta}{2}\right)^J} \sqrt{\binom{2J}{m}} |J, m - J\rangle. \quad (6.10)$$

For each $J \in \{0, \frac{1}{2}, 1, \ldots\}$, $\{J_x, J_y, J_z\}$ form an irrep of the $\mathfrak{su}(2)$ algebra, satisfying the well-known angular momentum commutation relations. The labeling by J *exhaustively characterizes* the irreps of $\mathfrak{su}(2)$, so every spin-coherent state corresponds to some irrep J. We go through several codes and show how they all correspond to the spin-coherent states

$$\left|\frac{\pi}{2}, \pi\mu\right\rangle_J = \frac{1}{2^J} \sum_{m=0}^{2J} (-1)^{\mu m} \sqrt{\binom{2J}{m}} |J, m - J\rangle, \quad (6.11)$$

(a)

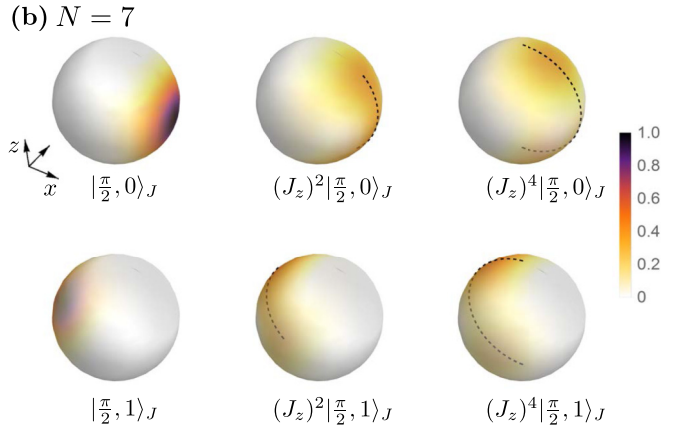| code	basis
bin	Fock states $|(S+1)m\rangle$
perm	M-qubit Dicke states $|D_{(S+1)m}^{M \geq 2J(S+1)}\rangle$
bin2	Fock states $|(S+1)(2J-m), (S+1)m\rangle$

(b) $N = 7$

FIG. 7. (a) Table listing the basis elements used to express binomial (`bin`) (6.1), permutation-invariant (`perm`) (6.14), and two-mode binomial codes (`bin2`) (6.16) (with $m \in \{0, 1, \ldots, 2J\}$). The coefficients next to these basis states are those of spin-coherent states $|\frac{\pi}{2}, \pi\mu\rangle_J$ (6.10) of a spin J. (b) Plots of overlaps $|_J\langle\theta, \phi|\psi\rangle|^2$ vs θ, ϕ given a spin state $|\psi\rangle$. The first (second, third) column shows normalized states $\{(J_z)^p|\frac{\pi}{2}, \pi\mu\rangle_{J=4}\}_{\mu=0}^1$ for $p \in \{0, 2, 4\}$. These plots show that powers of the "error" J_z cause $|\frac{\pi}{2}, 0\rangle_J$ and $|\frac{\pi}{2}, 1\rangle_J$ to approach each other in the Bloch sphere and eventually overlap at the two poles. The dashed arcs connecting $\theta = \pm\frac{\pi}{2J}p$ serve as a guide to the eye. Since J_z is mapped to $\hat{n} - J$ under the Holstein-Primakoff transformation, this provides an interpretation of dephasing errors for `bin` codes, which are correctable as long as $p < J = \frac{1}{2}(N + 1)$. Here, $N = 7$ and one can see from the third column that $(J_z)^4|\frac{\pi}{2}, 0\rangle_J$ and $(J_z)^4|\frac{\pi}{2}, 1\rangle_J$ overlap.

whose basis elements $|J, m - J\rangle$ are mapped either to Fock states of an oscillator(s) or a multiqubit system, summarized in Fig. 7(a). Moreover,

$$\frac{J_x}{J}\left|\frac{\pi}{2}, \pi\mu\right\rangle_J = (-1)^\mu \left|\frac{\pi}{2}, \pi\mu\right\rangle_J, \qquad (6.12)$$

providing us with a logical Z-operator J_x/J and a check operator $(J_x/J)^2$ for all of the codes. In addition, since spin-coherent states resolve the identity operator of the spin, they offer a way to visualize the location of various states before and after certain errors on a generalized Bloch sphere of the spin. Shown in Fig. 7(b), $|\frac{\pi}{2}, \pi\mu\rangle_J$ correspond to spin-coherent states at the antipodal points $(\frac{\pi}{2}, 0)$ and $(\frac{\pi}{2}, \pi)$.

1. Binomial codes

Setting $2J = N + 1$, the coefficients of $|\frac{\pi}{2}, \pi\mu\rangle_J$ (6.11) match those of the `bin` states (6.1). If we additionally map the spin states $|J, m - J\rangle$ into Fock states $|m\rangle$, then

$$\left|\frac{\pi}{2}, \pi\mu\right\rangle_{J = \frac{N+1}{2}} \rightarrow |\mu_{\text{bin}(N, S=0)}\rangle. \qquad (6.13)$$

The operator J_z (6.9) is then mapped to $\hat{n} - J$, revealing the well-known Holstein-Primakoff mapping of a spin into

a boson. For larger spacing $S > 0$, we map $|J, m - J\rangle$ into Fock states $|(S + 1)m\rangle$. Therefore, we have shown that `bin` codes correspond to single-mode irreps of $\mathfrak{su}(2)$ produced by the Holstein-Primakoff mapping.

For `bin` codes, the check operator $(J_x/J)^2$ is related to the protection from dephasing errors (characterized by N), and its nondestructive measurement can be used to detect such errors (see Sec. VID). Moreover, the Bloch sphere picture offers a nice interpretation of why a $\text{bin}(N, S)$ code protects against $k \leqslant N$ dephasing errors. Since $J_z = \hat{n} - J$ in this irrep, the action of a dephasing error $\hat{n}^k$ is directly related to application of J_z and the code states are eigenstates of J_x at antipodal parts of the Bloch sphere. Thus, one action of J_z raises (lowers) the expectation value of J_x for the logical zero (one) state, moving them closer together from their antipodal positions at the equator [Fig. 7(b)]. The states only begin to overlap when a high enough power $(J_z)^{k>N}$ has been applied, which corresponds to an unprotected dephasing error $\hat{n}^{k>N}$ in the bosonic representation.

2. Permutation-invariant codes

These codes (denoted here as `perm`) were introduced by Ouyang [73] (see also Ref. [100]) to tackle single-qubit amplitude damping. Given M qubits and parameters $\{J, S\}$, the logical states are

$$|\mu_{\text{perm}}\rangle = \frac{1}{2^J} \sum_{m=0}^{2J} (-1)^{\mu m} \sqrt{\binom{2J}{m}} |D_{(S+1)m}^M\rangle, \qquad (6.14)$$

where the Dicke state $|D_{(S+1)m}^N\rangle$ is the fully symmetrized M-qubit state with $(S + 1)m$ qubits in state $|1\rangle$ and the remaining qubits in $|0\rangle$. Therefore, we need to have $M \geqslant 2J(S + 1)$ in order to accommodate all of the required Dicke states. If $2J = 3S + 1$ and $M = 3S^2 + 5S$, these codes can protect against qubit amplitude damping errors of weight S [73]. In this context, the spacing S (between excitations of the Dicke states) quantifies a distance of the codes. We can already see the resemblance to spin-coherent states, but here the irrep is more complicated. For simplicity, let us set $S = 0$ and $M = 2J$; the $S > 0$ case is a straightforward extension whose basis elements are shown in Fig. 7(a). For such cases, $|\mu_{\text{perm}}\rangle$ are spin-coherent states of the largest irrep of $\mathfrak{su}(2)$ arising from tensoring M spin-$1/2$ particles, with $J = M/2$ playing the role of a collective spin.

3. Two-mode binomial codes

There is another well-known $\mathfrak{su}(2)$-related construct—the Jordan-Schwinger mapping of a spin into two bosons. Letting $\hat{a}_1, \hat{a}_2$ be the lowering operators of the two bosons and X, Y, Z be the Pauli matrices, we have

$$J_{[x,y,z]} = \frac{1}{2} \sum_{j,k=0}^{1} \hat{a}_j^\dagger [X, Y, Z]_{jk} \hat{a}_k. \qquad (6.15)$$

For example, $J_z = \frac{1}{2}(\hat{a}_1^\dagger \hat{a}_1 - \hat{a}_2^\dagger \hat{a}_2)$. The state space associated with each irrep of this type corresponds to the subspace of fixed total occupation number, i.e., all two-mode Fock states $|n_1, n_2\rangle$ such that $n_1 + n_2 = 2J$ for an irrep of spin J. One can thus see that the total spin is proportional to the "identity" component on the subspace, $J = \frac{1}{2}(\hat{a}_1^\dagger \hat{a}_1 + \hat{a}_2^\dagger \hat{a}_2)$. Any code

within a subspace of fixed J thus consists of eigenstates of the joint backaction $(1 - \gamma)^{(\hat{a}_1^\dagger \hat{a}_1 + \hat{a}_2^\dagger \hat{a}_2)/2}$ in the two-mode pure-loss error operators (1.4) (assuming identical γ's for both modes). Codes having this structure were first considered in Ref. [18].

Spin-coherent states of these irreps correspond to a class of two-mode binomial codes (`bin2`). Mapping the basis $|J, m - J\rangle$ from Eq. (6.11) to Fock states $|(S + 1)(2J - m), (S + 1)m\rangle$ yields the `bin2` code states

$$|\mu_{\text{bin2}}\rangle = \frac{1}{2^J} \sum_{m=0}^{2J} (-1)^{\mu m} \sqrt{\binom{2J}{m}} |2J - (S+1)m, (S+1)m\rangle. \qquad (6.16)$$

As with `bin`, we can parametrize `bin2` in terms of spacing S and a dephasing error parameter N. Then, $J = \frac{1}{2}(N + 1)(S + 1)$ in order to fit all of the required two-mode Fock states. Note that $\text{bin2}(S = 0)$ can be obtained by acting on the Fock state $|S + 1, 0\rangle$ with a 50:50 beam splitter [22] and were considered before in the context of three-mode squeezing [101].

4. Further generalizations

We have covered four bases in which to embed a spin—the spin's own basis $\{|J, m - J\rangle\}_{m=0}^{2J}$, Fock states $\{|m\rangle\}_{m=0}^{2J}$, Dicke states $\{|D_m^M\rangle\}_{m=0}^{2J}$, and two-mode Fock states $\{|2J - m, m\rangle\}_{m=0}^{2J}$. There are other relations between these bases and further code extensions. First, we can go in reverse of what was discussed above and embed any bosonic code into a multiqubit Hilbert space by mapping Fock states to Dicke states. While this produces `perm` codes when the bosonic code is `bin`, it produces previously unexplored codes when the bosonic code is, e.g., `cat` or `gkp` (although such states require $J \to \infty$ due their infinite-dimensional support). Second, Dicke states $\{|D_m^M\rangle\}_{m=0}^{2J}$ converge to Fock states in the limit of fixed J but large $M \gg 2J$ [98]. This famous limit is equivalent to the south pole of the Bloch sphere flattening out into ordinary bosonic phase space in the limit that the Bloch sphere is infinitely large. In this limit,

$$\text{perm}(M \to \infty, J, S) \to \text{bin}(N = 2J - 1, S). \qquad (6.17)$$

Third, the `bin2` states can be tensored $2J$ times to construct logical states for the $4J$-mode `noon` code [22], $|\mu_{\text{noon}}\rangle = |\mu_{\text{bin2}}\rangle^{\otimes 2J}$. The same procedure can of course be applied to `bin` codes. Offering an interesting alternative to spacing, `noon` codes instead concatenate `bin2` with a $2J$-block bit-flip code to correct for up to $2J - 1$ loss errors. Fourth, qubit (qudit) `bin` codes can *themselves* be thought of as bit-flip codes when expressed in a basis of multiqubit (multiqudit) states (see Appendix C).

D. Error-correction procedure for binomial codes

The existence of approximate error recovery maps for the various codes does not explicitly suggest by what means these recovery maps can be implemented nor whether fault-tolerant error recovery is possible for these codes. For qubit stabilizer codes, the theory of fault-tolerant error correction has been developed. For `gkp`, methods of fault-tolerant quantum error correction [9] are possible which simply generalize the techniques of qudit (d-dimensional) stabilizer codes to $d \to \infty$.

In this section, we investigate for the binomial qubit codes what measurements of commuting check operators could give sufficient error information to undo a set of errors. The recovery procedure we give is not necessarily the optimal one obtained by optimization in Sec. II. The binomial code family $\mathtt{bin}(N, S)$ can correct against errors in the error set $\mathscr{E} = \{I, \hat{a}, \ldots, \hat{a}^L, \hat{a}^\dagger, \ldots, (\hat{a}^\dagger)^G, \hat{n}, \ldots \hat{n}^D\}$ with $S = L + G$ and $N = \max(L, G, 2D)$. We know from Eq. (6.11) that the code words correspond to antipodal spin-coherent states of spin $J = \frac{1}{2}(N + 1)$. We will refer to the $N + 1$-dimensional subspace as the spin space.

Imagine that one error from the set $\mathscr{E}$ has taken place on an encoded state. The following procedure describes how to undo this error. First, one nondestructively measures the eigenvalues of the check operator J_x^2 which has eigenvalue $+J^2$ on all states in the code space. Here we assume that the operator J_x only has support on the Fock states $|m(S + 1)\rangle$ and thus has zero eigenvalues elsewhere. Of course, the check operator J_x is not unique and any form of nondestructively learning the value $|m_x|$ is permitted. For odd N (integer spin J), such a measurement has outcomes $|m_x|^2$ with $|m_x| = 0, 1, \ldots, J$. The outcome $m_x = 0$ cannot have come about from one of the dephasing errors of the form $\hat{n}^k$ since this error operator maps an initial state with $|m_x| = J$ to a superposition of states with $|m_x| \geqslant J - k$ so that for $k \leqslant D$, one cannot reach $|m_x| = 0$. For even N (half-integer spin J), m_x will never be zero by the application of a dephasing error. Hence, if one finds the eigenvalue $m_x = 0$, one concludes that photon loss or photon gain errors of the form $\hat{a}^k$, $k \leqslant L$ and $(\hat{a}^\dagger)^l$, $l \leqslant G$ must have occurred. In order to learn more about these photon loss and gain errors, one then measures the photon parity check operator (5.7). If one finds any another value of $|m_x| = k$ for the first measurement, one rotates the two-dimensional $m_x = \pm k$ subspace back to the two-dimensional $m_x = \pm J$ subspace by some unitary transformation. For stabilizer (resp., gkp), this correction can be a Pauli operator (resp., small displacement). Note that, unlike for stabilizer, $\mathtt{cat}$, or $\mathtt{gkp}$ codes, it is necessary to physically apply the correction. In other words, unlike the use of Pauli frames [102], we cannot just record the value of $|m_x|$ and keep the quantum information in this error space with a lower value for $|m_x|$: Subsequent dephasing errors would lead to more laddering up and down in the spin space so the QEC conditions would no longer be met [see Fig. 7(b)]. In case $m_x = 0$, one nondestructively measures the eigenvalues of $C_{\mathtt{bin}}$ (5.7) (via phase estimation, say), allowing one to learn the photon parity k modulo $S + 1$. When $k \leqslant G$ (at most G photons are gained), one applies

$$U_k^+ = \sum_{\mu = 0}^{1} \frac{|\mu_{\mathtt{bin}}\rangle \langle \mu_{\mathtt{bin}}| \hat{a}^k}{\sqrt{\langle \mu_{\mathtt{bin}}| \hat{a}^k (\hat{a}^\dagger)^k |\mu_{\mathtt{bin}}\rangle}} + V_{\mathrm{else}}^+, \qquad (6.18)$$

where V_{else}^+ is chosen to make U_k^+ unitary. When $k > G$ (at most L photons lost), one applies

$$U_l^- = \sum_{\mu = 0}^{1} \frac{|\mu_{\mathtt{bin}}\rangle \langle \mu_{\mathtt{bin}}| (\hat{a}^\dagger)^l}{\sqrt{\langle \mu_{\mathtt{bin}}| (\hat{a}^\dagger)^l \hat{a}^l |\mu_{\mathtt{bin}}\rangle}} + V_{\mathrm{else}}^- \qquad (6.19)$$

with $l = S + 1 - k$. These rotations are not a simple adding or subtracting of photons since some Fock states in $|\mu_{\mathtt{bin}}\rangle$ have been annihilated.

This form of error correction unfortunately does not correct products of dephasing and photon loss or gain errors, which are in principle errors against which the code can correct [23]. Note also that Ref. [23] has shown that specifically for the photon loss channel with errors as in Eq. (1.4), only the measurement of the rotation operator $C_{\mathtt{bin}}$ (5.7) is required (since there is one particular dephasing error associated with a particular number of photon losses). The procedure above thus falls short of giving a general prescription for error correction for the binomial codes. However, for the binomial qubit code $\mathtt{bin}(N = 2, S = 1)$, one can give a scheme which corrects all the errors which meet the quantum error conditions. The code $\mathtt{bin}(N = 2, S = 1)$ can correct against the errors in $\mathscr{E} = \{I, \hat{a}, \hat{n}\}$. For these parameters, the code space is inside the Fock space with a maximum of 6 photons, $\mathcal{F}_6 = \mathrm{Span}\{|0\rangle, \ldots, |6\rangle\}$. We split this space into the direct sum of the even subspace and the odd subspace, that is $\mathcal{F}_6 = \mathcal{F}_6^{\mathrm{even}} \oplus \mathcal{F}_6^{\mathrm{odd}} = \mathrm{Span}\{|0\rangle, |2\rangle, |4\rangle, |6\rangle\} \oplus \mathrm{Span}\{|1\rangle, |3\rangle, |5\rangle\}$.

The code space is inside $\mathcal{F}_6^{\mathrm{even}}$. We identify $\mathcal{F}_6^{\mathrm{even}}$ with a spin $J = 3/2$ via the mapping

$$|6\rangle \equiv |m_z = \tfrac{3}{2}\rangle, \quad |4\rangle \equiv |m_z = \tfrac{1}{2}\rangle,$$
$$|2\rangle \equiv |m_z = -\tfrac{1}{2}\rangle, \quad |0\rangle \equiv |m_z = -\tfrac{3}{2}\rangle,$$

so that $|0_{\mathtt{bin}}\rangle$ and $|1_{\mathtt{bin}}\rangle$ code states are the highest and lowest eigenstates of $J_x(J = 3/2)$,

$$|0_{\mathtt{bin}}\rangle = |m_x = \tfrac{3}{2}\rangle, \qquad |1_{\mathtt{bin}}\rangle = |m_x = -\tfrac{3}{2}\rangle.$$

One dephasing error leaves the code states inside $\mathcal{F}_6^{\mathrm{even}}$. More precisely, it is a linear combination of the identity and $J_z(J = 3/2)$ such that

$$\hat{n}|0_{\mathtt{bin}}\rangle = 2|m_x = \tfrac{1}{2}\rangle + 3|0_{\mathtt{bin}}\rangle, \qquad (6.20\mathrm{a})$$

$$\hat{n}|1_{\mathtt{bin}}\rangle = 2|m_x = -\tfrac{1}{2}\rangle + 3|1_{\mathtt{bin}}\rangle. \qquad (6.20\mathrm{b})$$

For convenience, we relabel the error states $|m_x = \pm\tfrac{1}{2}\rangle \equiv |n\pm\rangle$.

Remarkably, one photon loss maps $\mathcal{F}_6^{\mathrm{even}}$ onto $\mathcal{F}_6^{\mathrm{odd}}$ in such a way that the code states are mapped onto shifted[5] code states for $N = 1$. This is in fact true for general N and S: One photon loss maps the code space (N, S) to the code space $(N - 1, S)$ but shifted by $+S$ in the Fock basis [23]. With this in mind, we can identify $\mathcal{F}_6^{\mathrm{odd}}$ with a spin $J = 1$ with the mapping

$$|5\rangle \equiv |m_z = 1\rangle, \quad |3\rangle \equiv |m_z = 0\rangle, \quad |1\rangle \equiv |m_z = -1\rangle,$$

and then the error states are the highest and lowest eigenstates of $J_x(J = 1)$,

$$\hat{a}|0_{\mathtt{bin}}\rangle \propto |m_x = 1\rangle, \quad \hat{a}|1_{\mathtt{bin}}\rangle \propto |m_x = -1\rangle,$$

which we rename for convenience $|m_x = \pm 1\rangle \equiv |a\pm\rangle$. The third state of this spin-1 subspace is called the unknown state as ending up in this state means the loss of logical information, $|m_x = 0\rangle \equiv |?\rangle$. To complete the description, one can note that the action of $\hat{n}$ on $\mathcal{F}_6^{\mathrm{odd}}$ is also a linear combination of $J_z(J = 1)$ and the identity I, i.e., mapping $|a\pm\rangle$ to a linear combination of itself and $|?\rangle$, and using the fact that $\hat{a}\hat{n} = \hat{n}\hat{a} - \hat{a}$, one can note that $|n\pm\rangle$ is mapped by $\hat{a}$ onto a linear combination of $|a\pm\rangle$ and $|?\rangle$. These relations are summarized in Table II.

TABLE II. Relations between code states and error states for `bin`$(N=2, S=1)$.

$J_x(J=3/2)$	$\mathcal{F}_6^{\text{even}}$		$\mathcal{F}_6^{\text{odd}}$	$J_x(J=1)$
$\frac{3}{2}$	$\lvert\mu_{\text{bin}}=0\rangle$	$\overset{\hat{a}}{\rightarrow}$	$\lvert a+\rangle$	1
	$\hat{n}\downarrow$		$\downarrow\hat{n}$	
$\frac{1}{2}$	$\lvert n+\rangle$	$\overset{\hat{a}}{\rightarrow}$	$\lvert ?\rangle$	0
$-\frac{1}{2}$	$\lvert n-\rangle$			
	$\hat{n}\uparrow$		$\uparrow\hat{n}$	
$-\frac{3}{2}$	$\lvert\mu_{\text{bin}}=1\rangle$	$\overset{\hat{a}}{\rightarrow}$	$\lvert a-\rangle$	-1

One possible way to extract error information is then to measure (via phase estimation) the eigenvalues of the following unitary

$$U = \exp\left\{\frac{2\pi i}{b}[a(J_x(3/2))^2 \oplus (J_x(1))^2]\right\}, \qquad (6.21)$$

where the two parameters a and b can be chosen to obtain good spacing between different eigenvalues. For example, one can choose $a=8$ and $b=5$, leading to Table III.

To obtain the four eigenvalues of U (via phase estimation), one needs at least two qubit ancillas. A more direct method would be to first measure photon parity. If odd, then correct for photon loss. If even, then one measures the eigenvalues of J_x^2 by measuring $U = \exp(i32\pi J_x(3/2)^2/9)$, which has eigenvalue $+1$ for the no-error case and eigenvalue $\exp(i32\pi/36) \approx -1$ in the dephasing error case.

VII. GKP CODES

While their error-correcting properties were first revealed in Ref. [9], `gkp` states have connections to quantum foundations [103], solid-state physics [104], and signal processing (where their analogs are frequency combs). The ideal (i.e., infinite $\bar{n}$) square lattice `gkps` code space, denoted by its projection $P_{\text{gkps}}^{\text{ideal}}$, is the simultaneous $+1$ eigenspace of the two commuting stabilizers

$$S_{\mathbf{x}} = D_{\sqrt{2\pi}} \quad \text{and} \quad S_{\mathbf{p}} = D_{i\sqrt{2\pi}}, \qquad (7.1)$$

where $D_\alpha \equiv e^{\alpha\hat{a}^\dagger - \alpha^\star\hat{a}}$ is the displacement operator (note that $D_{\sqrt{2\pi}} = e^{-i2\sqrt{\pi}\hat{p}}$). The projection onto the code can be constructed out of all of their powers,

$$P_{\text{gkps}}^{\text{ideal}} \equiv \left(\frac{1}{\sqrt{\pi}}\sum_{n\in\mathbb{Z}}S_{\mathbf{x}}^n\right)\left(\frac{1}{\sqrt{\pi}}\sum_{n\in\mathbb{Z}}S_{\mathbf{p}}^n\right) \equiv P_{\mathbf{x}}P_{\mathbf{p}}. \qquad (7.2)$$

Applying the Poisson summation formula allows us to express $P_{\mathbf{x}}$ ($P_{\mathbf{p}}$) as a sum of projections onto eigenstates $\lvert n\sqrt{\pi}\rangle_{\hat{x}}$

TABLE III. Eigenstructure of the proposed unitary to be measured for error correction of `bin`$(N=2, S=1)$.

Eigenstates	$\lvert\mu=0 \text{ or } 1\rangle$	$\lvert n+\rangle, \lvert n-\rangle$	$\lvert a+\rangle, \lvert a-\rangle$	$\lvert ?\rangle$
Eigenvalues	$e^{i6\pi/5}$	$e^{i4\pi/5}$	$e^{i2\pi/5}$	1
Decoding	No error	Dephasing	Photon loss	Failure

$(\lvert n\sqrt{\pi}\rangle_{\hat{p}})$ of $\hat{x}$ ($\hat{p}$). We demonstrate this for $P_{\mathbf{p}}$:

$$P_{\mathbf{p}} = \frac{1}{\sqrt{\pi}}\sum_{n\in\mathbb{Z}}e^{i2\sqrt{\pi}n\hat{x}} \qquad (7.3a)$$

$$= \sum_{n\in\mathbb{Z}}\delta(\hat{x}-\sqrt{\pi}n) \qquad (7.3b)$$

$$= \sum_{n\in\mathbb{Z}}\lvert\sqrt{\pi}n\rangle_{\hat{x}}\langle\sqrt{\pi}n\rvert. \qquad (7.3c)$$

These sets of positions and momenta make up the *code lattice*, the lattice dual to the stabilizer lattice (in the language of Ref. [9]) and generated by the logical operators

$$X_{\text{gkps}} = D_{\sqrt{\pi/2}} = S_{\mathbf{x}}^{1/2}, \qquad (7.4a)$$

$$Z_{\text{gkps}} = D_{i\sqrt{\pi/2}} = S_{\mathbf{p}}^{1/2}. \qquad (7.4b)$$

The maximally mixed state $\frac{1}{2}P_{\text{gkps}}$ reproduces this lattice, shown in the fourth panel in Fig. 1.

Conventionally, `gkps` logical states are expressed in terms of squeezed states,

$$\lvert\mu_{\text{gkps}}^{\text{ideal}}\rangle \propto \sum_{n\in\mathbb{Z}}\lvert\sqrt{\pi}(2n+\mu)\rangle_{\hat{x}}. \qquad (7.5)$$

One can obtain an equivalent (see Appendix D 2) representation in terms of coherent states by projecting the vacuum state $\lvert 0\rangle$ onto the code and the ± 1 eigenstates of Z_{gkps}:

$$\lvert\mu_{\text{gkps}}^{\text{ideal}}\rangle \propto [I+(-1)^\mu Z_{\text{gkps}}]P_{\text{gkps}}^{\text{ideal}}\lvert 0\rangle$$

$$= \sum_{\vec{n}\in\mathbb{Z}^2}D_{\sqrt{\frac{\pi}{2}}(2n_1+\mu)}D_{i\sqrt{\frac{\pi}{2}}n_2}\lvert 0\rangle, \qquad (7.6)$$

where $\vec{n}\equiv(n_1,n_2)$. The above displacements generate the two *state lattices* for Z_{gkps}-logical states, whose horizontal spacing is twice that of the code lattice due to the $I\pm Z_{\text{gkps}}$ term. In general, the state lattice depends on the logical basis used (see Fig. 8) while the code lattice is basis independent.

The usual way to make the states (7.5) have finite $\bar{n}$ (and therefore be normalizable) is to assume finite squeezing for each position eigenstate and add a Δ^2-dependent Gaussian envelope, producing the `gkps` states in Eq. (7.7a) below. Alternatively, one can add a Gaussian envelope to Eq. (7.6), yielding a representation in terms of coherent states, Eq. (7.7b). A third finite-$\bar{n}$ representation can be written in terms of $\lvert\mu_{\text{gkps}}^{\text{ideal}}\rangle$ smeared by a Gaussian distribution of displacements [9], making contact with the errors that the codes are designed to correct. This is the third equation below:

$$\lvert\mu_{\text{gkps}}^\Delta\rangle \propto \sum_{n\in\mathbb{Z}}e^{-\frac{\pi}{2}\Delta^2(2n+\mu)^2}D_{\sqrt{\frac{\pi}{2}}(2n+\mu)}S_{-\ln\Delta}\lvert 0\rangle \qquad (7.7a)$$

$$\sim \sum_{\vec{n}\in\mathbb{Z}^2}e^{-\frac{\pi}{2}\Delta^2[(2n_1+\mu)^2+n_2^2]}D_{\sqrt{\frac{\pi}{2}}(2n_1+\mu)}D_{i\sqrt{\frac{\pi}{2}}n_2}\lvert 0\rangle$$

$$(7.7b)$$

$$\sim \int d^2\alpha\,\frac{e^{-\lvert\alpha\rvert^2/\Delta^2}}{\sqrt{\pi\Delta^2/2}}D_\alpha\lvert\mu_{\text{gkps}}^{\text{ideal}}\rangle, \qquad (7.7c)$$

where $\mu\in\{0,1\}$ and $S_r = e^{+\frac{1}{2}r(\hat{a}^2-\hat{a}^{\dagger 2})}$ is the squeezing operator. We use $\Delta\in[0,1]$ for both the envelope and squeezing
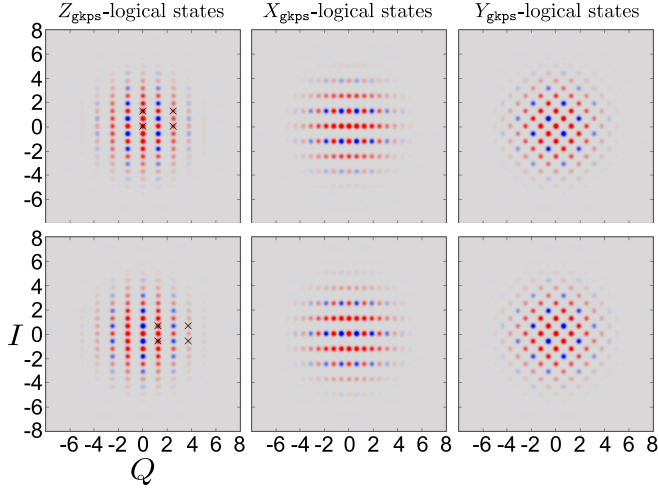
FIG. 8. Wigner function sketches of the two Z_{gkps}-, X_{gkps}-, Y_{gkps}-logical states. Comparing to the fourth panel in Fig. 1, which shows that the code lattice is square, here we see that the lattices formed by the logical states may be square or rectangular, depending on which logical operator is considered. The unit cell of the state lattices (7.6) is marked by "×" in the two leftmost panels; the remaining dots appear as a result of the coherences between different coherent states.

parameters for simplicity. These representations numerically converge to each other very quickly in the $\Delta \to 0$ limit, but there are visual differences between them for small envelopes. A fourth representation in terms of Fock states is possible using Eq. (86) from Ref. [94]. Note that $|0_{\text{gkps}}^{\Delta}\rangle$ and $|1_{\text{gkps}}^{\Delta}\rangle$ are nonorthogonal for nonzero Δ, and this source of error manifests itself in the QEC matrix.

Recall that gkps($\Delta \to 0$) states can protect against displacement errors $D_{\alpha_1 + i\alpha_2}$ in phase space, where $|\alpha_1|, |\alpha_2| < \sqrt{\pi/8}$. The representations can easily generalize to more tightly packed lattices, yielding a slightly larger volume of correctable displacements. To construct the coherent-state representation of the Z_{gkp}-logical states, one first constructs commuting stabilizers (following Ref. [9]) and repeats Eq. (7.6). Adding an envelope, this representation (7.7b) is particularly simple to express:

$$\left|\mu_{\text{gkp}}^{\Delta}\right\rangle \propto \sum_{\alpha \in L(\mu)} e^{-\Delta^2 |\alpha|^2} e^{-i\alpha_1\alpha_2} |\alpha\rangle, \qquad (7.8)$$

where $|\alpha = \alpha_1 + i\alpha_2\rangle$ is a coherent state and $L(\mu)$ is the state lattice for each code state μ. We considered both these lattices and their shifted versions (B4) for the gkp numerics (see Appendix B). For all analytics below, we use the finite-$\bar{n}_{\text{gkps}}$ unshifted square-lattice states $|\mu_{\text{gkps}}^{\Delta}\rangle$ (7.7c), noting any generalizations to other lattices.

We have calculated moments of the occupation number, yielding a geometric (i.e., thermal) distribution:

$$\overline{n_{\text{gkps}}^{\ell}} \equiv \tfrac{1}{2}\text{Tr}\{P_{\text{gkps}}\hat{n}^{\ell}\} \sim \ell!\bar{n}_{\text{gkps}}^{\ell}, \qquad (7.9)$$

where the average occupation number is

$$\bar{n}_{\text{gkps}} \sim \frac{1}{2\Delta^2} - \frac{1}{2}. \qquad (7.10)$$

As expected, the moments diverge as the states become unnormalizable in the small Δ limit.

A. QEC matrix for GKP codes

Recall that any trace class bosonic operator A (i.e., satisfying $\text{Tr}\{A^{\dagger}A\} < \infty$) can be expanded in terms of displacement operators using the orthogonality condition of D_{α} at the superoperator level [105],

$$\text{Tr}\{D_{\alpha}^{\dagger}D_{\beta}\} = \pi\delta^2(\alpha - \beta). \qquad (7.11)$$

The expansion is then $A = \int \frac{d^2\alpha}{\pi}\text{Tr}\{D_{\alpha}^{\dagger}A\}D_{\alpha}$, where the integral is over all of phase space and $\text{Tr}\{D_{\alpha}^{\dagger}A\}$ is the *characteristic function* of A. Protection of gkp against pure loss was previously discussed using an approximation of $\hat{a}$ in terms of a *sum* of displacements instead of an integral, at first very briefly [9] and subsequently taking into account the maximum number of photons in the oscillator [106]. Here we calculate the QEC matrix $\epsilon_{\ell\ell'}^{\text{gkps}}$ (4.1) by expressing Kraus operators in terms of the full integral expansion.

Unlike $\hat{a}$, the error operator E_{ℓ} and its variants are trace class due to the damping term, yielding

$$E_{\ell}^{\dagger}E_{\ell'} = \int \frac{d^2\alpha}{\pi}e^{-\frac{1}{2}(1-\gamma)|\alpha|^2}\langle\ell|D_{\alpha^\star}|\ell'\rangle D_{\alpha\sqrt{\gamma}}, \qquad (7.12)$$

where $\langle\ell|D_{\alpha^\star}|\ell'\rangle$ are matrix elements of the displacement operator D_{α} in the Fock state basis (D2). To obtain this, one can express the trace in $\text{Tr}\{D_{\alpha}^{\dagger}E_{\ell}^{\dagger}E_{\ell'}\}$ as a sum over Fock states, plug in Eq. (D2), and use the generating function of Laguerre polynomials (D3). Complementing the expansion of Gaussian noise $\{D_{\alpha}\}$ in terms of photon creation and annihilation operators (e.g., Ref. [107]), the above equation completes the "Rosetta stone," expressing each of the two primary noise models in the language of the other.

To gain a flavor of the calculations below, let us first examine how we can calculate $c_{\ell\ell}^{\text{gkps}} = \tfrac{1}{2}\text{Tr}\{P_{\text{gkps}}E_{\ell}^{\dagger}E_{\ell}\}$. Using Eq. (7.12), this calculation boils down to determining $\text{Tr}\{P_{\text{gkps}}D_{\alpha}\}$. Consider first the infinite $\bar{n}_{\text{gkps}}$ limit, recalling from Fig. 1 that P_{gkps} is a sum of (unphysical) points of fixed position and momentum arranged in a square code lattice. Then, $\text{Tr}\{P_{\text{gkps}}D_{\alpha}\}$ will be nonzero only for those α which displace the lattice back on top of itself, i.e., α displaces by a multiple of the lattice's unit cell. For those cases, the overlap of each point with itself will be infinite, and so the total result is $\sum_{\Lambda}\delta^2(\alpha - \Lambda)$, where the sum is over all displacements $\Lambda \in \sqrt{\pi/2}(n_1, n_2)$ (with integers $n_{1,2}$) preserving the code lattice. Coming back to finite $\bar{n}_{\text{gkps}}$, a natural guess would be to substitute the Gaussian representation $\frac{1}{\Delta}e^{-\frac{1}{2\Delta^2}|\alpha-\Lambda|^2}$ for the Dirac δ function in the sum. This almost obtains the right result, but there are two more steps. The first is normalization, which cancels the $\frac{1}{\Delta}$ in front of the Gaussian representation of the δ function. The second is addition of the Gaussian envelope, yielding

$$\frac{1}{2}\text{Tr}\{P_{\text{gkps}}D_{\alpha}\} \sim \sum_{\Lambda} e^{-\frac{1}{2\Delta^2}|\alpha-\Lambda|^2}e^{-\frac{\Delta^2}{2}|\Lambda|^2}. \qquad (7.13)$$

Notice that what used to be a Dirac δ is now a Kronecker $\delta_{\alpha,\Lambda} \sim e^{-\frac{1}{2\Delta^2}|\alpha-\Lambda|^2}$ in the small Δ limit. As a sanity check, setting $\alpha = 0$ yields unity in that limit.

Calculating these overlaps is more involved (see Appendix D 3), but we can nevertheless use the above intuition to

understand the more general element (with $\nu \in \{0,1\}$)

$$
\langle \mu_{\text{gkps}}^{\Delta} | D_{\alpha} | \nu_{\text{gkps}}^{\Delta} \rangle
$$

$$
\sim \sum_{\vec{n} \in \mathbb{Z}^2} e^{i\pi(n_1 + \frac{\mu+\nu}{2})n_2} e^{-\frac{1}{2\Delta^2}|\alpha - \Lambda_{\delta\mu}^{\vec{n}}|^2} e^{-\frac{\Delta^2}{2}|\Lambda_{\delta\mu}^{\vec{n}}|^2}, \quad (7.14)
$$

where $\delta\mu \equiv \mu - \nu$ and $\Lambda_{\delta\mu}^{\vec{n}} \equiv \sqrt{\frac{\pi}{2}}[(2n_1 + \delta\mu) + in_2]$. Since there are two different state lattices, there are extra phases in the sum and the sum is over displacements $\alpha \in \Lambda_{\delta\mu}^{\vec{n}}$ which overlap the two lattices. We can now plug this into $E_{\ell}^{\dagger} E_{\ell'}$ (7.12) and proceed to calculate the integral (see Appendix D 4), yielding

$$
\langle \mu_{\text{gkps}}^{\Delta} | E_{\ell}^{\dagger} E_{\ell'} | \nu_{\text{gkps}}^{\Delta} \rangle
$$

$$
\sim \sqrt{c_{\ell\ell}^{\text{gkps}} c_{\ell'\ell'}^{\text{gkps}}} \sum_{\vec{n} \in \mathbb{Z}^2} e^{-\frac{(1-\gamma)}{2\gamma}|\Lambda_{\delta\mu}^{\vec{n}}|^2}
$$

$$
\times e^{i\pi(n_1 + \frac{\mu+\nu}{2})n_2} e^{-\frac{\Delta^2}{2}|\Lambda_{\delta\mu}^{\vec{n}}|^2} \langle \ell | D_{(\Lambda_{\delta\mu}^{\vec{n}})^{\star}/\sqrt{\gamma}} | \ell' \rangle, \quad (7.15)
$$

for $\gamma\bar{n}_{\text{gkps}} \to \infty$, where $c_{\ell\ell}^{\text{gkps}}$ will turn out to be the probability of losing ℓ photons,

$$
c_{\ell\ell}^{\text{gkps}} = \frac{1}{2}\text{Tr}\{P_{\text{gkps}} E_{\ell}^{\dagger} E_{\ell}\} \sim \frac{(\gamma\bar{n}_{\text{gkps}})^{\ell}}{(\gamma\bar{n}_{\text{gkps}} + 1)^{\ell+1}}. \quad (7.16)
$$

Thus, the photon loss distribution for gkps is a asymptotically thermal with mean $\gamma\bar{n}$. Ignoring the Δ^2 envelope term from now on, all Δ dependence of Eq. (7.15) is contained in $c_{\ell\ell}^{\text{gkps}}$.

Notice that $|\langle \ell | D_{(\Lambda_{\delta\mu}^{\vec{n}})^{\star}/\sqrt{\gamma}} | \ell' \rangle| \leqslant 1$ because they are overlaps between two states. Thus, the only quantity regulating the sum (7.15) is $e^{-\frac{(1-\gamma)}{2\gamma}|\Lambda_{\delta\mu}^{\vec{n}}|^2}$. Assuming $\gamma \ll 1$, the "on-site" term ($\vec{n} = \vec{0}$) in Eq. (7.15) is μ independent and thus contributes to

$$
c_{\ell\ell'}^{\text{gkps}} \sim c_{\ell\ell}^{\text{gkps}} \delta_{\ell\ell'}, \quad (7.17)
$$

while the "nearest-neighbor" terms ($|\vec{n}| = 1$) contribute to the leading-order uncorrectable parts

$$
|z_{\ell\ell'}^{\text{gkps}}| \sim \sqrt{c_{\ell\ell}^{\text{gkps}} c_{\ell'\ell'}^{\text{gkps}}} e^{-\frac{\pi}{4}\frac{1-\gamma}{\gamma}} \langle \ell | (D_{\sqrt{\frac{\pi}{2\gamma}}} + D_{\sqrt{\frac{\pi}{2\gamma}}}^{\dagger}) | \ell' \rangle \quad (7.18)
$$

and $|x_{\ell\ell'}^{\text{gkps}}| = |z_{\ell\ell'}^{\text{gkps}}|$. (There is no $|y_{\ell\ell'}^{\text{gkps}}|$ to this order.) The uncorrectable parts are the same (up to sign) due to the identical effect of position and momentum displacements on the code. So, while the nearest neighbors $\vec{n} = (0,\pm 1)$ contributed to $z_{\ell\ell'}^{\text{gkps}}$ and $\vec{n} = (\pm 1, 0)$ contributed to $x_{\ell\ell'}^{\text{gkps}}$, the two quantities have to be equal in magnitude due to this effect. (Considering more general lattices can, of course, break this balance.) We can also see another symmetry manifest itself—the invariance of the lattice under parity $(-1)^{\hat{n}}$. Since the sum of displacements is even under parity, it does not connect even Fock states to odd ones and guarantees that $z_{\ell\ell'}^{\text{gkps}} = 0$ unless $\ell - \ell'$ is even. This means that technically gkps codes have spacing $S = 1$. However, this spacing disappears when the lattice is slightly shifted and the symmetry lost, but the performance of the codes remains. This should not be surprising since shifted gkps code spaces are akin to Pauli frames in the stabilizer formalism [106]. The most striking result is that the reason for this high performance is not due to the spacing, but to the suppression by the γ-dependent exponential factor in Eq. (7.18). Namely, while $\epsilon_{\ell\ell'}^{\text{gkps}}$ contains uncorrectable parts for all ℓ, ℓ' (modulo symmetry constraints), *all of these parts*

are suppressed exponentially by $e^{-\frac{\pi}{4}\frac{1-\gamma}{\gamma}}$ when $\gamma \ll 1$. As an example, we show the comparable strength of the exponential suppression of uncorrectable parts for gkps($\bar{n} \approx 6$) in Fig. 5. Assuming that the infidelity $1 - F_{\mathcal{E}}$ to leading order in γ is polynomial in all uncorrectable parts, one expects $1 - F_{\mathcal{E}}$ to also be exponentially suppressed by $\frac{1-\gamma}{\gamma}$. We proceed to show this by bounding $1 - F_{\mathcal{E}}$ using an explicit recovery.

B. Removing energy constraints

In Fig. 2, we have observed that $F_{\mathcal{E}}^{\text{gkps}}$ is significantly higher than that for all other codes for most γ. While the nontrivial exponential suppression of uncorrectable parts (7.18) of the QEC matrix hints at an analytical explanation, this still does not tell us how $F_{\mathcal{E}}$ scales with γ. For this, we need to consider a specific analytically tractable recovery. Having investigated several recoveries, the simplest one we found is based on the fact that the combination of amplification and pure loss produces Gaussian (i.e., displacement) noise [2]—a channel which most naturally fits the error-correction capabilities of gkps.

Coming back to the formulation of $\mathcal{N}$ in terms of a beam splitter (1.6), consider amplifying the signal

$$
\hat{a} \to \sqrt{G}\hat{a} + \sqrt{G-1}\hat{b}^{\dagger} \quad (7.19)
$$

after application of $\mathcal{N}$. Here, G is the gain of the amplifier, which we set to $G = e^{\chi} = \frac{1}{1-\gamma}$ to compensate the effect of damping. Tracing out the $\hat{b}$ mode, amplification is simply the properly normalized transpose of pure loss,

$$
\mathcal{A}(\cdot) = (1-\gamma)\mathcal{N}^{\ddagger}(\cdot), \quad (7.20)
$$

where "$\ddagger$" is the adjoint in the matrix representation. The Kraus operators of $\mathcal{A}$ are $\sqrt{1-\gamma}E_{\ell}^{\dagger}$ for $\ell \in \{0, 1, \dots\}$ {e.g., Ref. [34], Eq. (5.5)} and it is simple to verify that $\mathcal{A}$ is indeed a channel: $(1-\gamma)\sum_{\ell=0}^{\infty} E_{\ell}E_{\ell}^{\dagger} = I$.

Upon amplification, the pure-loss channel $\mathcal{N}$ is transformed into a Gaussian noise channel with variance $\frac{\gamma}{1-\gamma}$. The noise comes from two parts: the intrinsic noise due to amplification and the amplified noise due to pure loss. More explicitly, we can apply Eq. (7.12) to express $\mathcal{A}\mathcal{N}$ in terms of displacements and use displacement orthogonality (7.11):

$$
\mathcal{A}\mathcal{N}(\rho) = (1-\gamma) \int \frac{d^2\alpha}{\pi} \frac{d^2\beta}{\pi} e^{-\frac{1-\gamma}{2}(|\alpha|^2 + |\beta|^2)} D_{\frac{\alpha}{\sqrt{\gamma}}} \rho D_{\frac{\beta}{\sqrt{\gamma}}}
$$

$$
\times \sum_{\ell,\ell'=0}^{\infty} \langle \ell | D_{\alpha^{\star}} | \ell' \rangle \langle \ell' | D_{\beta^{\star}} | \ell \rangle
$$

$$
= (1-\gamma) \int \frac{d^2\alpha}{\pi} d^2\beta\, e^{-\frac{1-\gamma}{2}(|\alpha|^2 + |\beta|^2)} D_{\frac{\alpha}{\sqrt{\gamma}}} \rho D_{\frac{\beta}{\sqrt{\gamma}}} \delta^2(\alpha + \beta)
$$

$$
= \frac{1-\gamma}{\gamma} \int \frac{d^2\alpha}{\pi} e^{-\frac{1-\gamma}{\gamma}|\alpha|^2} D_{\alpha} \rho D_{\alpha}^{\dagger}. \quad (7.21)
$$

Appending amplification with the conventional gkp recovery $\mathcal{R}^{\text{GKP}}$ which measures and corrects displacements within the correctable unit cell [9], the total recovery we consider is

$$
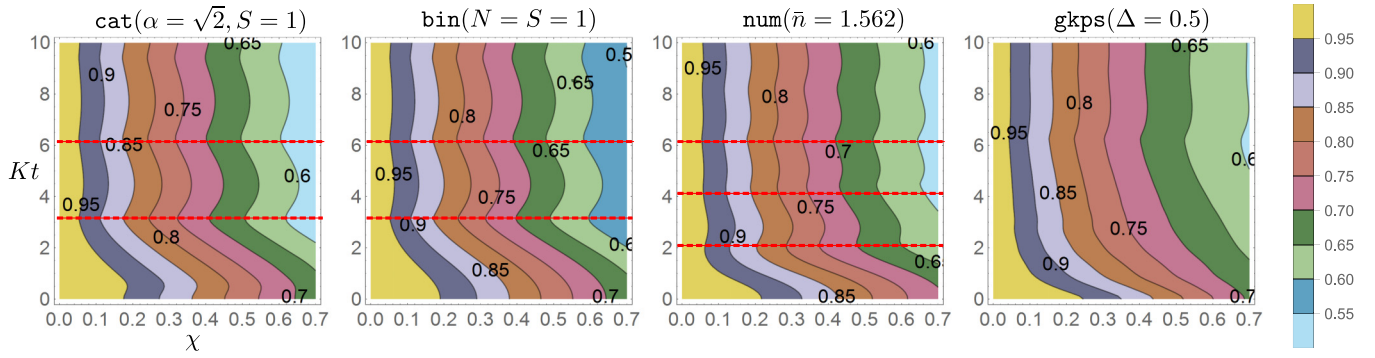\mathcal{R}^{\text{AGKP}} = \mathcal{R}^{\text{GKP}}\mathcal{A}. \quad (7.22)
$$

FIG. 9. Contour plot of $F_{\mathcal{E}}$ vs dimensionless Kerr parameter Kt and damping parameter $\chi \equiv \kappa t$ for cat, bin, num, and gkps picked such that they all have $\bar{n} \approx 2$. Here, K is the strength of the Kerr Hamiltonian (8.1), κ is the cavity decay rate, and t is time. Starting with a fixed χ and tracking increasing Kt, we see that $F_{\mathcal{E}}$ quickly decreases to a constant for all codes considered, implying a potentially universal failure of error correction when Kt is large. However, $F_{\mathcal{E}}$ is minimal at high-symmetry points of the codes (dashed red lines) and maximal in between. For example, $\frac{1}{2}P_{\text{num}}$ for the num code is threefold symmetric (see Fig. 1), and $F_{\mathcal{E}}(\text{num})$ is minimal at Kt, being the first few multiples of $2\pi/3$.

Note that the above derivation of Gaussian noise is exact for *all values* of γ. We bound $F_{\mathcal{E}}$ by calculating the success probability that $\mathcal{R}^{\text{AGKP}}$ will succeed in correcting the above Gaussian noise $\mathcal{AN}$ starting with *ideal* (i.e., infinite $\bar{n}_{\text{gkps}}$) code states:

$$P_{\text{succ}}(\gamma) = \frac{1-\gamma}{\gamma} \int_{\blacksquare} \frac{d\alpha_1 d\alpha_2}{\pi} e^{-\frac{1-\gamma}{\gamma}(\alpha_1^2 + \alpha_2^2)}, \quad (7.23)$$

where the integration is over correctable displacements $|\alpha_1|, |\alpha_2| \leqslant \sqrt{\pi/8}$ denoted by $\blacksquare$. The channel infidelity $1 - F_{\mathcal{E}}^{\text{AGKP}}$ can then be estimated using the failure probability, which is the complementary integral outside of the unit cell. Upper bounding that integral by integrating the complement of the circle with radius $\sqrt{\pi/8}$ (instead of the complement of the square with length $\sqrt{\pi/8}$) yields

$$P_{\text{fail}}(\gamma) < \frac{1-\gamma}{\gamma} \int_{|\alpha| \geqslant \sqrt{\frac{\pi}{8}}} \frac{d^2\alpha}{\pi} e^{-\frac{1-\gamma}{\gamma}|\alpha|^2} = e^{-\frac{\pi}{8}\frac{1-\gamma}{\gamma}}. \quad (7.24)$$

We remark that this bound can be improved to $e^{-\frac{\pi}{4\sqrt{3}}\frac{1-\gamma}{\gamma}}$ using ideal hexagonal gkp.

VIII. ADDITIONAL FEATURES

A. Nonlinearity

We have tried to address the effect of pure loss on our codes, but real-world microwave cavities have undesired unitary evolution (i.e., coherent errors). In general, the joint effect of pure loss and a unitary process on the state depends not only on how many losses have occurred but also their specific times. The purpose of this subsection is to answer the following: *Does adding coherent errors reduce code performance?* The answer to this, at least in our case, is a firm "yes."

The coherent (i.e., unitary) error we add is generated by a Kerr nonlinearity with Hamiltonian

$$H_K \equiv \frac{1}{2}K\hat{n}(\hat{n}-1) = \frac{1}{2}K\hat{a}^{\dagger 2}\hat{a}^2, \quad (8.1)$$

with Kerr parameter K. Here, we show what happens when our codes get exposed to the joint evolution of pure loss and Kerr, namely, the channel

$$\mathcal{N}_{\chi,Kt}(\cdot) = e^{-iKt[\frac{1}{2}\hat{n}(\hat{n}-1),\cdot]+\chi\mathcal{D}(\cdot)}, \quad (8.2)$$

where $\mathcal{D}(\cdot)$ is the Lindbladian for the pure loss channel from Sec. IA. Since the Kerr nonlinearity is prominent in cavities coupled with transmons (as opposed to optical fibers), we use the excitation loss rate κ to quantify the strength of pure loss (recall that $\gamma = 1 - e^{-\kappa t}$). Thus, the two unitless scales of the problem are $\chi \equiv \kappa t$ and Kt. An analytic Kraus representation for $\mathcal{N}_{\chi,Kt}$ has yet to be obtained, but various approaches have come close [108,109].

Figure 9 plots $F_{\mathcal{E}}(Kt,\chi)$ for four code families at $\bar{n}_{\text{code}} \approx 2$. For $Kt = 0$ (horizontal axis in each plot), we see the same behavior in $F_{\mathcal{E}}$ vs pure loss strength as we saw before. For the other extreme of $\chi = 0$ (vertical axis), we see unit $F_{\mathcal{E}}$ since Kerr nonlinearity alone is a perfectly correctable unitary process. Starting with a fixed nonzero χ and looking up at the vertical line of increasing Kt, we see that $F_{\mathcal{E}}$ quickly decreases to a (roughly) constant value for all codes. Since the optimal recovery is not able to ascertain exactly when photon loss events occurred, Kerr evolution induces rotations of unknown angle between those events, $e^{-iH_Kt}\hat{a} \propto \hat{a}e^{i\hat{n}Kt}e^{-iH_Kt}$, and thus destroys the quantum information. Metrology protocols are also susceptible to this effect [110]. The value to which $F_{\mathcal{E}}$ decreases at large Kt seems to be (roughly) universal across all codes, so there might be a fundamental limit to correcting large $\bar{n}$-dependent coherent errors in the presence of incoherent errors. However, it is still possible to use error correction to our advantage in, e.g., the $Kt \approx 1$ regime (given $\bar{n} \approx 2$). Incidentally, in that regime, cat shows an *increase* in $F_{\mathcal{E}}$, implying that a slight amount of Kerr is actually helping cat-code performance. We are investigating this effect in a subsequent publication.

Lastly, we want to mention that $F_{\mathcal{E}}$ for a given code is minimal when Kt is at an angle by which rotating $\frac{1}{2}P_{\text{code}}$ leaves the projection invariant. Recall that evolution under e^{-iH_Kt} causes a coherent state to transform into cat states at certain rational t [111–113], but such recurrences are quickly degraded under pure loss [114]. Nevertheless, we see some periodicity in code performance, e.g., in the case of num in the third panel in Fig. 9. From Fig. 1, we know that $\frac{1}{2}P_{\text{num}}$ for this num code is threefold rotationally symmetric in phase space. Coincidentally, $F_{\mathcal{E}}(\text{num})$ is minimal at Kt being the first few multiples of $2\pi/3$ (dashed red lines) and maximal in between.
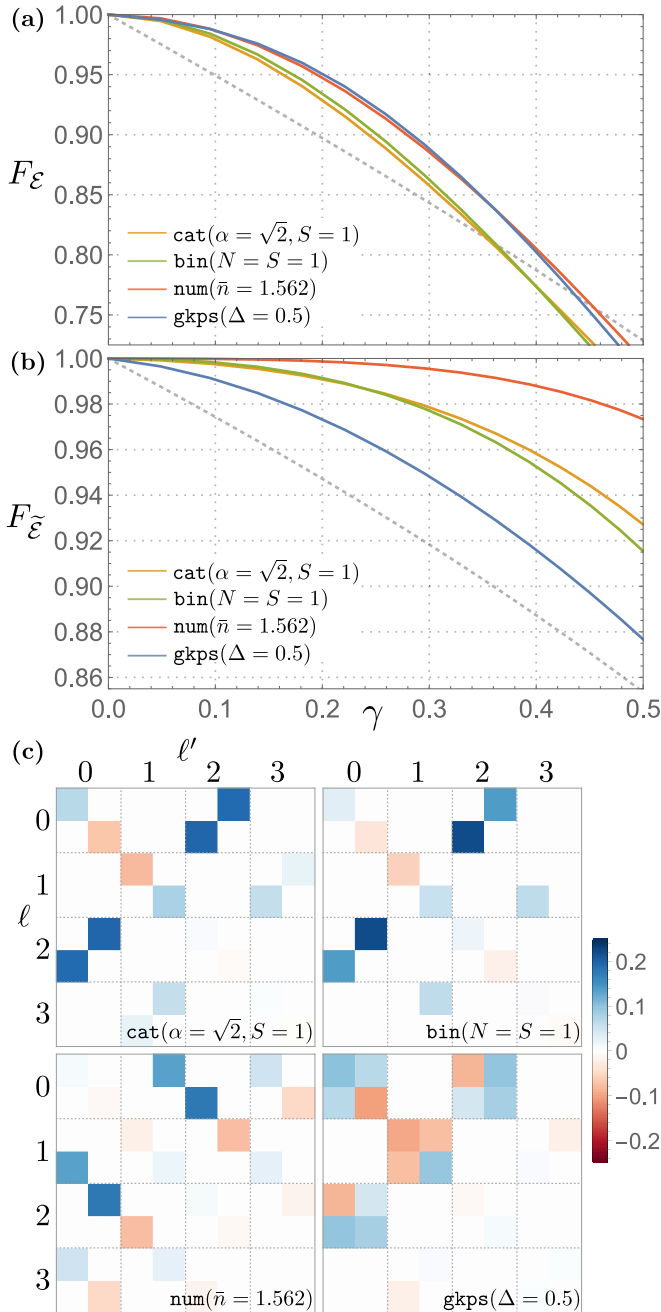
FIG. 10. Channel fidelity for `cat`, `bin`, `num`, and `gkps` picked such that they all have $\bar{n} \approx 2$, given (a) the pure-loss channel $\mathcal{N}$ (1.4) and (b) the pure-loss channel with the capability of knowing the number of photons lost, modeled by the quantum instrument $\widetilde{\mathcal{N}}$ (8.4). (c) Uncorrectable parts of the QEC matrices, $\epsilon^{\text{code}} - c^{\text{code}}$ (see Sec. IV), for the four codes (cf. Fig. 5). Since uncorrectable parts in the diagonal blocks $\epsilon_{\ell\ell}^{\text{code}}$ are all that matter in recovering from $\widetilde{\mathcal{N}}$, we see that `cat`, `bin`, and `num` outperform `gkp` at this $\bar{n}$.

We do not know the reason for this effect, but one can see that it occurs for all codes to various extent.

B. Parity measurements

Here, we briefly consider the following question: What if we know how many photons were lost?

While microwave cavities still do not have the capability to directly count photons, one can perform nondemolition photon parity measurements of microwave cavity modes [29,115] and vibrational modes of trapped ions [116]. If we assume that (1) we have a fixed-parity initial state and (2) we can measure parity $(-1)^{\hat{n}}$ during the loss portion of the channel $\mathcal{E}$ perfectly and faster than any time scale of the system, then we can in principle track every loss event $\hat{a}$ without destroying the state. This results in an *unraveled* [117] system, where part of the knowledge reserved for the environment—the number and times of the loss events—is now learned by the experimenter. For example, a trajectory lasting time t during which jumps occurred at times $\tau_2 \geqslant \tau_1$ would incur the conditional evolution

$$\widetilde{E}_2 |\psi\rangle \equiv e^{-\frac{1}{2}\kappa(t-\tau_2)\hat{n}} \hat{a} e^{-\frac{1}{2}\kappa(\tau_2-\tau_1)\hat{n}} \hat{a} e^{-\frac{1}{2}\kappa\tau_1\hat{n}} |\psi\rangle, \qquad (8.3)$$

where $|\psi\rangle$ is the initial state of the oscillator and we have not yet renormalized the state. Defining $\widetilde{E}_\ell$ in similar fashion and permuting all $\hat{a}$'s to the left leaves us with $\widetilde{E}_\ell = f_\ell E_\ell$, where E_ℓ (1.4) is the Kraus operator for the pure-loss channel and f_ℓ is a function of the jump times. Since f_ℓ is a scalar, knowledge of jump times is irrelevant to error correction against pure loss, and we can ignore it from now on. We model this process using an extended version of pure loss $\mathcal{N}$—the quantum instrument [62]

$$\widetilde{\mathcal{N}}(\rho) = \sum_{\ell=0}^{\infty} E_\ell \rho E_\ell^{\dagger} \otimes |\ell\rangle\langle\ell|, \qquad (8.4)$$

where ρ is a single-mode density matrix. The second tensor factor represents our knowledge of ℓ, making sure that each $E_\ell \rho E_\ell^{\dagger}$ is mapped into an orthogonal subspace of the extended Hilbert space. The corresponding recovery $\mathcal{R}$ has Kraus operators $U_\ell^{\text{code}} \otimes \langle\ell|$, with U_ℓ^{code} being a unitary mapping $E_\ell P_{\text{code}}$ into P_{code}.

We compare $F_{\mathcal{E}}$ (1.11) with $F_{\widetilde{\mathcal{E}}}$, the fidelity given the extended loss channel $\widetilde{\mathcal{N}}$ (8.4), in Figs. 10(a) and 10(b), respectively. Code performance improves for all codes, even the naive 0/1 Fock state encoding (dashed gray line). Note that, given this extra knowledge, only the diagonal blocks $\epsilon_{\ell\ell}^{\text{code}}$ of the QEC matrix (4.1) are relevant. The `gkps` codes have the largest uncorrectable parts in those blocks at this $\bar{n}_{\text{gkps}}$ [see Fig. 10(c)], so their performance increases the least out of all the codes. The `num` code winds up being the optimal encoding in the eyeball norm, but the parity-tracking procedure described above is not applicable to this code since its states (B2a) are not of fixed parity.

IX. CONCLUSION

The results of this paper can be categorized into two parts: one regarding code performance and one regarding code structure. Here, we summarize these results and comment on future directions and open questions.

A. Code performance

We gave a numerical and analytical performance comparison of the four primary single-mode continuous-variable quantum codes—cat [13,15,27], binomial (`bin`) [23], numerically optimized (`num`, Ref. [23] and here), and `gkp` [9] codes—against the pure loss channel with loss rate γ. For the numerical

part, we compared the codes' ability to preserve entanglement using channel-adapted error correction [35] subject to two caveats: (1) the encoding, recovery, and decoding are all assumed perfect and (2) the codes are grouped by their mean occupation number $\bar{n}_{\text{code}} \leqslant 2,5,10$. For the analytical part, we calculated the quantum error-correction (QEC) conditions for `cat`, `bin`, and `gkp`. We briefly discuss our results below, but encourage the reader to peruse Sec. II for more details.

Even though `cat` and `bin` follow the traditional convention of correcting exactly against a subset of errors, their performance is significantly worse than that of `gkp` for many γ. While `gkp` does not exactly correct against any errors, we find that the violation of the QEC conditions for each error is insignificant compared to the violation of the leading-order uncorrectable errors for `cat` and `bin` for most $\bar{n}_{\text{code}}$.

In the limit of vanishing γ, we observe the following order of performance: `gkp` < `cat` < `bin` < `num`. On the one hand, since `cat` and `bin` codes correct exactly against the first few errors, their performance scales polynomially with γ. We further reveal the regions in the $\{\gamma, \bar{n}_{\text{code}}\}$ parameter space in which `bin` codes outperform `cat` codes. On the other hand, we analytically show that `gkp` code entanglement infidelity is $O(e^{-\frac{c}{\gamma}})$ (with c a constant dependent on the type of `gkp` code).

As γ increases, `gkp` quickly overpowers the rest of the codes and their performance persists even for high γ. For example, optimal recovery of a `gkp` state with 10 photons on average yields a fidelity of 99.5% given $\gamma \approx 20\%$, compared to a fidelity of 96.9% for `bin` and 96.6% for `cat`. At $\gamma \approx 30\%$, where about one in every three photons are lost, the fidelity of `gkp` is still 95%, which is 5% higher than that of `cat` and `bin`. At such high γ, we observed the following order of performance: `bin` $\lesssim$ `cat` < `num` < `gkp`.

We extended our analysis of entanglement preservation to determine achievable rates of quantum communication using these codes, where we saw similar orders of performance. We also show that sending a `gkp` state with an average occupation number of two photons produces a higher communication rate than distributing two photons among four modes using the smallest encoding protecting against one loss error.

Relaxing the $\bar{n}_{\text{code}}$ constraint, we have numerical evidence showing that performance of `gkp` codes and a subset of `bin` codes increases with increasing $\bar{n}_{\text{code}}$. Since the ideal `gkp` states indeed have infinite $\bar{n}_{\text{gkp}}$, it is reasonable that their performance increases monotonically as they become more ideal. To back this claim analytically, we cook up a simple recovery procedure involving phase-insensitive amplification that converts the pure-loss channel into Gaussian noise; this procedure can also be used in multimode extensions of `gkp` codes [54]. The `bin` increase in performance can be justified by showing they have a larger set of approximately correctable errors than previously thought; we do so in Secs. VIA–VIB.

We added a unitary error in the form of a Kerr nonlinearity K in order to see how code performance is changed. We observed that `cat` code performance increases slightly at small K and that all code performance oscillates with periods depending on their symmetries; these are subjects of future investigation. We also observed that, at sufficiently large K, the performance of all codes fails at about the same rate, signaling the need to keep such coherent errors low in a real device. We also briefly addressed changes in code performance if one is able to learn how many photons were lost.

There are obvious generalizations of this analysis to other multimode codes mentioned in the introduction, storing multiple qubits worth of information; we are currently pursuing some of them. Another direction has to do with having `gkp` and `bin` codes catch up to `cat` codes in terms of experimental realizability. While `gkp` codes may have been considered by some to be unphysical in the past, recent technological advances in, e.g., microwave cavity [118], atomic ensemble [119,120], and trapped ion [121] control, suggest that making these states may be within reach. In fact, there are recent theoretical proposals related to making and maintaining `gkp` states in two of the aforementioned technologies [106,122] (see Refs. [123–128] for other proposals) and a related trapped-ion experiment [129]. While the comparison offered here is completely free from consideration of experimental imperfections, we hope that our conclusions will motivate the community to pursue quantum information processing and communication schemes with `bin` and `gkp` states.

B. Code structure

We discussed a connection between `bin` codes and spin-coherent states and used it to characterize related two-mode binomial codes as well as multiqubit permutation-invariant codes [73]. This connection yields a check operator for `bin` dephasing errors, and we discussed an error-correction scheme that utilizes this operator. This connection was also extended to qudit versions of the aforementioned codes, yielding a generalization of spin-coherent states and a check operator for qudit codes.

By mapping the coefficients of the qudit `bin` code into a particular subspace of multiple qudits, we introduced a multiqudit (i.e., discrete variable) code that extends the multiqubit permutation-invariant codes. The extension, which we call `perm'` in order to differentiate it from another extension [75], turns out to be nothing but a multiqudit bit-flip code when expressed in the basis of products of the individual qudit states. However, when expressed in terms of a qudit extension of Dicke states, the coefficients next to those states match those of the qudit binomial codes. This relates the protection of the continuous-variable binomial codes to that of a discrete-variable bit-flip code. A similar bit-flip-like trick was used for another code—the `noon` code [22]—where two-mode `noon` states and their multimode generalizations were tensored together to form codes protecting against pure loss. Such intriguing connections between discrete- and continuous-variable codes should be investigated further. In addition, the generalization of spin-coherent states introduced here may be useful in experimental settings such as atomic ensembles (e.g., Refs. [120,130]) and magnetometry (e.g., Refs. [131–133]).

APPENDIX A: THE MANY FACES OF CHANNEL FIDELITY

A well-known property of $F_{\mathcal{E}}$ is the relation to the average input-output fidelity of $\mathcal{E}$ [134] (see also Ref. [135]),

$$\int d\psi \langle\psi|\mathcal{E}(|\psi\rangle\langle\psi|)|\psi\rangle = \frac{dF_{\mathcal{E}} + 1}{d + 1}, \quad \text{(A1)}$$

where d is the dimension of the system. Above, $\langle\psi|\mathcal{E}(|\psi\rangle\langle\psi|)|\psi\rangle$ is the input-output fidelity for some initial state $|\psi\rangle$ of the source qubit and $d\psi$ is a uniform distribution over all pure states. Because of the above equality, one should not be surprised that the capacity of entanglement transmission determined by $F_{\mathcal{E}}$ is equivalent to the capacity of pure state preservation determined by the input-output fidelity [136]. In addition, since $F_{\mathcal{E}}$ is a fidelity between two states, it gives rise to a metric, is stable under addition of ancillary systems, and satisfies the chaining property (meaning that it can be used to provide a bound on the error of a larger quantum computation). These properties can be proven using Ref. [137], where $F_{\mathcal{E}} = F_{\text{pro}}(\mathcal{E}, \mathcal{I})$.

The channel fidelity can be related to the worst-case input-output fidelity $\min_{|\psi\rangle}\langle\psi|\mathcal{E}(|\psi\rangle\langle\psi|)|\psi\rangle$ [138],

$$1 - d\sqrt{1 - F_{\mathcal{E}}^2} \leqslant \min_{|\psi\rangle}\langle\psi|\mathcal{E}(|\psi\rangle\langle\psi|)|\psi\rangle. \quad \text{(A2)}$$

The dependence of the bound on the dimension as well as the square of $F_{\mathcal{E}}$ suggests that $F_{\mathcal{E}}$ is not a good measure of the worst-case scenario. There is indeed a discrepancy between average and worst-case behavior for channels that contain a combination of coherent and incoherent noise [139,140]. Such an example here is pure loss with an additional Kerr nonlinearity, considered in Sec. VIII. However, the pure loss channel alone contains only incoherent noise and so $F_{\mathcal{E}}$ is a reasonable marker of even worst-case behavior; moreover, one can prove that the dimension dependence goes away entirely [141]. Note that Eq. (A2) was derived starting from the worst-case infidelity $1 - \min_{|\psi\rangle}\langle\psi|\mathcal{E}(|\psi\rangle\langle\psi|)|\psi\rangle$, applying the Fuchs–van de Graaf inequalities [142] to convert the infidelity to the maximum trace distance $\max_{|\psi\rangle}\|(\mathcal{E} - \mathcal{I})(|\psi\rangle\langle\psi|)\|_{\text{tr}}$ (with $\|A\|_{\text{tr}} \equiv \text{Tr}\{\sqrt{A^\dagger A}\}$), upper bounding said trace distance by its stabilized version $\max_{|\psi\rangle}\|(\mathcal{E} \otimes \mathcal{I} - \mathcal{I}^{\otimes 2})(|\psi\rangle\langle\psi|)\|_{\text{tr}}$ (i.e., the diamond norm), upper bounding the diamond norm by the trace norm $d\|\rho_{\mathcal{E}} - \rho_{\mathcal{I}}\|_{\text{tr}}$ between the Choi matrices of $\mathcal{E}$ and $\mathcal{I}$ using Lemma 7 from Refs. [143,144] (this is where the d

dependence comes in), and once again applying Fuchs–van de Graaf to turn the trace norm into $d\sqrt{1 - F_{\mathcal{E}}^2}$.

An information-theoretic property of $F_{\mathcal{E}}$ is its presence in the quantum Fano inequality [145] (see also Ref. [81], Theorem 12.9). This is an upper bound on the von Neumann entropy of $\rho_{\mathcal{E}}$,

$$H(\rho_{\mathcal{E}}) \leqslant H(\{F_{\mathcal{E}}, 1 - F_{\mathcal{E}}\}) + (1 - F_{\mathcal{E}})\log_2(d^2 - 1), \quad \text{(A3)}$$

where $H(\rho_{\mathcal{E}}) = -\text{Tr}\{\rho_{\mathcal{E}}\log_2\rho_{\mathcal{E}}\}$. The entropy $H(\rho_{\mathcal{E}})$ is also called the entropy exchange since it quantifies the entropy gained by the environment responsible for the nonunitary nature of $\mathcal{E}$. There is also the anti-Fano inequality [36], a lower bound on $F_{\mathcal{E}}$ in terms of the entropy,

$$F_{\mathcal{E}} \geqslant e^{-2H(\rho_{\mathcal{E}})}. \quad \text{(A4)}$$

A similar relation to information-theoretic quantities can be made regarding a specific error map—the erasure channel. Let us divide B into two regions, B_1 and B_2, and trace over B_2. Then, the $F_{\mathcal{E}}$ given the optimal recovery which reconstructs B_2 using only B_1 satisfies

$$F_{\mathcal{E}} \geqslant e^{-\frac{1}{2}I(A:B_2|B_1)}, \quad \text{(A5)}$$

where $I(A : B_2|B_1)$ is the conditional mutual information quantifying correlations between A and B_2 given information from B_1 [146]. In this context, $F_{\mathcal{E}}$ is also called the fidelity of recovery [147].

From yet another information theory perspective ([41], Theorem 2), the optimal $F_{\mathcal{E}}$ yields the equality

$$dF_{\mathcal{E}} = 2^{-H_{\min}(B|A)_{\mathcal{NS}}}, \quad \text{(A6)}$$

where $H_{\min}(B|A)_{\mathcal{NS}}$ is the conditional min entropy of the Choi matrix $\rho_{\mathcal{NS}}$ (1.9) of the encoding and loss portions of $\mathcal{E}$. This inequality can be adapted from the equation below Eq. (4.5) in Ref. [42] by noting that $\mathcal{R}^{\ddagger}\mathcal{S}$ is a unital map. In this context, A and B (of dimensions ∞ and two, respectively) share a state $\rho_{\mathcal{NS}} = \mathcal{NS} \otimes \mathcal{I}(|\Psi\rangle\langle\Psi|)$, and $H_{\min}(B|A)_{\mathcal{NS}}$ is the most conservative way to quantify the uncertainty about the state of B after the state of A is sampled.

APPENDIX B: NUMERICAL BENCHMARKING DETAILS

The parameters for the specific members of the `cat`, `bin`, `gkps`, and `gkp` code families that optimize $F_{\mathcal{E}}(\gamma)$ in Fig. 2 are given in Table IV.

There are a total of five `num` codes, organized by their approximate mean occupation number

$$\bar{n}_{\text{num}} \in \{1.562, 2.696, 2.770, 4.149, 4.336\}. \quad \text{(B1)}$$

Interestingly, the first code—the $\sqrt{17}$ code [23]—can be expressed as

$$|0_{\text{num}}^{\bar{n}\approx 1.562}\rangle = \frac{1}{\sqrt{6}}(\sqrt{7 - \sqrt{17}}|0\rangle + \sqrt{\sqrt{17} - 1}|3\rangle), \quad \text{(B2a)}$$

$$|1_{\text{num}}^{\bar{n}\approx 1.562}\rangle = \frac{1}{\sqrt{6}}(\sqrt{9 - \sqrt{17}}|1\rangle - \sqrt{\sqrt{17} - 3}|4\rangle). \quad \text{(B2b)}$$

All five code states are listed in the ancillary MATHEMATICA notebook accompanying this paper as Supplemental Material [148]. The $\bar{n}_{\text{num}} \in \{1.562, 2.696, 4.149\}$ codes are from

TABLE IV. Code parameters for the code giving the highest $F_{\mathcal{E}}$ (1.11) out of all codes of a given code family with the constraint $\bar{n}_{code} \leqslant 2,5,10$ and a given loss rate γ (1.3). The first column lists the γ's sampled while the next three sets of seven columns give the code parameters for $\mathtt{cat}(\alpha,S)$ (5.1), $\mathtt{bin}(N,S)$ (6.1), $\mathtt{gkps}(\Delta)$ (7.8), and $\mathtt{gkp}(\Delta,a)$ (B4). Each of the three sets corresponds to one of the three energy constraints. Optimal code values below $\gamma \leqslant 0.0124$ do not change significantly and so are not shown. For $\gamma \geqslant 0.4512$, a small $\bar{n}$ is preferable for all codes and optimal $\mathtt{cat}$ and $\mathtt{bin}$ switch to encoding into the first two Fock states. For $\bar{n}_{gkp} \leqslant 10$ and $\gamma \leqslant 0.05$, denoted with the " symbol in the last column, we bound $F_{\mathcal{E}}$ with the channel fidelity $F_{\mathcal{E}}^{QR}$ that uses the quadratic recovery $\mathcal{R}^{QR}$ [149] (see also Refs. [150,151]) for unshifted hexagonal $\mathtt{gkp}$ (7.8) because $F_{\mathcal{E}}$ decreases significantly around that regime due to numerical precision limitations of the optimization. We also use $F_{\mathcal{E}}^{QR}$ to bound $F_{\mathcal{E}}$ in Sec. VI B.

| | $\bar{n}_{code} \leqslant 2$ | | | | | | | $\bar{n}_{code} \leqslant 5$ | | | | | | | $\bar{n}_{code} \leqslant 10$ | | | | | | |
| | cat | | bin | | gkps | gkp | | cat | | bin | | gkps | gkp | | cat | | bin | | gkps | gkp | |
γ	α	S	N	S	Δ	Δ	a	α	S	N	S	Δ	Δ	a	α	S	N	S	Δ	Δ	a
0.0124	1.440	1	1	1	0.481	0.477	1.550	1.739	2	2	2	0.309	0.309	1.650	2.890	3	3	4	0.221	0.221	"
0.0247	1.440	1	1	1	0.481	0.477	1.618	1.746	2	2	2	0.309	0.309	1.650	2.890	3	3	4	0.221	0.221	"
0.0488	1.396	1	1	1	0.481	0.477	1.618	1.962	3	1	2	0.309	0.309	1.650	3.162	4	2	4	0.221	0.221	"
0.0723	1.369	1	1	1	0.481	0.477	1.618	1.969	3	1	3	0.309	0.309	1.700	3.162	4	2	5	0.221	0.221	"
0.0952	1.351	1	1	1	0.481	0.477	1.618	1.975	3	1	3	0.309	0.309	1.700	1.975	3	2	5	0.221	0.221	"
0.1175	1.332	1	1	1	0.481	0.477	1.618	1.981	3	1	3	0.309	0.309	1.700	1.981	3	2	5	0.221	0.221	1.725
0.1393	1.508	2	1	1	0.481	0.477	1.618	1.987	3	1	3	0.309	0.309	1.700	1.987	3	2	5	0.221	0.221	1.725
0.1605	1.508	2	1	1	0.481	0.477	1.618	1.994	3	1	3	0.309	0.309	1.700	1.994	3	2	5	0.221	0.221	1.725
0.1813	1.508	2	1	1	0.481	0.477	1.618	1.994	3	1	3	0.309	0.309	1.700	1.994	3	2	5	0.221	0.221	1.725
0.2015	1.508	2	1	1	0.481	0.477	1.618	2.000	3	1	3	0.309	0.309	1.700	2.000	3	2	5	0.221	0.221	1.725
0.2212	1.508	2	1	1	0.481	0.477	1.618	2.000	3	1	3	0.309	0.309	1.700	2.000	3	2	5	0.221	0.221	1.725
0.2404	1.508	2	1	1	0.481	0.477	1.618	2.000	3	1	3	0.309	0.309	1.700	2.000	3	1	3	0.221	0.221	1.725
0.2592	1.508	2	1	1	0.481	0.477	1.618	2.000	3	1	3	0.309	0.309	1.700	2.000	3	1	3	0.221	0.221	1.725
0.2775	1.508	2	1	1	0.481	0.477	1.618	1.994	3	1	3	0.309	0.309	1.700	1.994	3	1	3	0.221	0.221	1.725
0.2953	1.508	2	1	1	0.481	0.477	1.618	1.994	3	1	3	0.309	0.309	1.700	1.994	3	1	3	0.221	0.221	1.725
0.3127	1.508	2	1	1	0.481	0.477	1.618	1.987	3	1	3	0.309	0.309	1.700	1.987	3	1	3	0.221	0.221	1.725
0.3297	1.508	2	1	1	0.481	0.477	1.618	1.981	3	1	3	0.309	0.309	1.700	1.981	3	1	3	0.221	0.221	1.725
0.3462	1.508	2	1	1	0.481	0.477	1.618	1.975	3	1	3	0.309	0.309	1.700	1.975	3	1	3	0.221	0.221	1.725
0.3624	1.508	2	1	1	0.481	0.477	1.618	1.969	3	1	3	0.309	0.309	1.700	1.969	3	1	3	0.221	0.221	1.725
0.3781	1.194	1	0	0	0.481	0.477	1.618	1.643	2	1	3	0.309	0.309	1.700	1.643	2	1	3	0.246	0.243	1.700
0.3935	1.183	1	0	0	0.481	0.477	1.525	1.636	2	1	2	0.316	0.312	1.732	1.636	2	1	2	0.316	0.312	1.732
0.4084	1.173	1	0	0	0.481	0.477	1.525	1.628	2	1	2	0.392	0.394	1.650	1.628	2	1	2	0.392	0.394	1.650
0.4231	1.173	1	0	0	0.481	0.510	1.450	1.612	2	1	2	0.471	0.510	1.450	1.612	2	1	2	0.471	0.510	1.450
0.4373	1.162	1	0	0	0.500	0.659	1.350	1.162	1	0	0	0.500	0.659	1.350	1.162	1	0	0	0.500	0.659	1.350
0.4512	0	0	0	0	0.535	0.659	1.350	0	0	0	0	0.535	0.659	1.350	0	0	0	0	0.535	0.659	1.350
0.4647	0	0	0	0	0.577	0.913	1.250	0	0	0	0	0.577	0.913	1.250	0	0	0	0	0.577	0.913	1.250
0.4780	0	0	0	0	0.632	0.933	1.200	0	0	0	0	0.632	0.933	1.200	0	0	0	0	0.632	0.933	1.200
0.4908	0	0	0	0	0.632	0.953	1.150	0	0	0	0	0.632	0.953	1.150	0	0	0	0	0.632	0.953	1.150
0.5034	0	0	0	0	0.632	0.976	1.100	0	0	0	0	0.632	0.976	1.100	0	0	0	0	0.632	0.976	1.100

Ref. [23] while the $\bar{n}_{num} \in \{2.770, 4.336\}$ codes were obtained here using a different optimization routine, described as follows. In order to find logical states $|\mu_{num}\rangle$ for $\mu \in \{0,1\}$ which allow for the correction of error operators $e_\ell \in \{I, \hat{a}, \hat{a}^2\}$, we create a cost function from the QEC matrix $f_{\mu\nu\ell\ell'} = \langle\mu_{num}|e_\ell^\dagger e_{\ell'}|\nu_{num}\rangle$, $c_1 = \sum_{\ell,\ell'} |f_{00\ell\ell'} - f_{11\ell\ell'}|^2 + |f_{01\ell\ell'}|^2$. In order to prefer lower occupation, the penalty $c_2 = \lambda_{\bar{n}} \bar{n}_{num}$ is introduced with $\lambda_{\bar{n}} = 10^{-3}$. Code words are produced by numerically optimizing the total cost over complex unit vectors:

$$\underset{|\psi_0\rangle, |\psi_1\rangle \in \mathbb{C}^{N_{max}}}{\text{minimize}} \quad c_1 + c_2, \tag{B3}$$

where N_{max} is the Fock space cutoff. The $\sqrt{17}$ code is the only code below $\bar{n}_{num} = 2$. For $\bar{n}_{num} \leqslant 5, 10$, the best-performing code for $\gamma \leqslant 0.3935$ is $\mathtt{num}(4.149)$, for $\gamma = 0.4084$ is $\mathtt{num}(2.770)$, and for $\gamma \geqslant 0.4231$ is the $\sqrt{17}$ code.

We were unable to find good codes with $\bar{n}_{num} > 5$ due to the prominence of shallow local minima.

For the numerical comparison, we swept all values of the code parameters for $\mathtt{cat}(\alpha,S)$, $\mathtt{bin}(N,S)$, $\mathtt{num}(\bar{n})$, and $\mathtt{gkps}(\Delta)$ subject to the energy constraints. For $\mathtt{gkp}(\Delta,a)$, we only considered values of Δ which gave $\bar{n}_{gkp} \approx 2$, 5, or 10 (since we knew from the $\mathtt{gkps}$ calculations that increasing $\bar{n}$ generally increased $F_{\mathcal{E}}$ for all but the largest values of γ). We also did not consider all possible nonrectangular lattices, but instead implemented a subset of them by sweeping the lattice parameter $a \in [1,2]$ in the following coherent-state representation for the shifted nonsquare $\mathtt{gkp}$ code states:

$$|\mu_{gkp}^{shift}\rangle \propto \sum_{\vec{n} \in \mathbb{Z}} (-1)^{\mu n_1} e^{-i\frac{\pi}{2} n_2 (2n_1 + \mu)} e^{-\frac{\pi a}{4} \Delta^2 [(2n_1 + \mu)^2 + (\frac{2}{a} n_2)^2]}$$

$$\times \left| \frac{\sqrt{\pi a}}{2} \left(2n_1 + \mu + i\frac{2}{a} n_2 \right) \right\rangle. \tag{B4}$$

The $\Delta \to 0$ states are stabilized by

$$S_{\mathbf{x}} = -D_{\sqrt{\frac{\pi a}{2}}(1+i\frac{2}{a})} \quad \text{and} \quad S_{\mathbf{p}} = D_{4i\sqrt{\frac{\pi}{2a}}}.$$

The resulting code lattice formed by $\frac{1}{2}P_{\text{gkp}}^{\text{shift}}$ is shifted, retaining its error-correcting properties but not having a lattice point at the origin. Both shifted and unshifted lattices are used in the numerics, but only unshifted lattices are used for analytical calculations in Appendix D. Note that $|0_{\text{gkp}}^{\text{shift}}\rangle$ is odd while $|1_{\text{gkp}}^{\text{shift}}\rangle$ is even under parity $(-1)^{\hat{n}}$, so there is no spacing S.

Since cat and gkp code states are formally superpositions of an infinite number of Fock states, we have to truncate them and use a finite Fock state superposition $\sum_{n=0}^{N_{\max}} c_n|n\rangle$ for each logical state. We picked $N_{\max}$ such that $\sum_{n=0}^{N_{\max}} |c_n|^2 \geqslant 0.99999$ for both logical states. For gkps/gkp code states, we used the coherent state representation and picked only the lattice points values $s,t \leqslant \lfloor 4/\Delta \rfloor$. The codes were generated with MATHEMATICA while the semidefinite program was executed using the CVX package [152] in MATLAB, with the MATLINK add-on [153] for MATHEMATICA acting as the bridge. Helpful routines were borrowed from Cubitt [154].

APPENDIX C: QUDIT bin AND bin2 CODES

Here we extend the analogy from Sec. VIC between bin codes and spin-coherent states to qu*d*it bin codes, introducing a multiqudit code perm$'$ in the process and eventually yielding a logical X_{code} operator and a check operator for the qu*d*it bin [23] and bin2 codes. We consider the following generalization of spin-coherent states (6.10),

$$|\theta,\phi\rangle_{J,d} = \sum_{m=0}^{2J} \frac{\sqrt{\binom{N+1}{m}_d} \, e^{i\phi m} \tan^{2m}\frac{d-1}{d}\theta}{\sqrt{\left(\sum_{\mu=0}^{d-1} \tan^{2\mu}\frac{d-1}{d}\theta\right)^{N+1}}} |J,m-J\rangle,$$

(C1)

where $|J,m-J\rangle$ is the spin basis of a spin $J = \frac{1}{2}(d-1)(N+1)$, and $\binom{N+1}{m}_d$ are extended binomial coefficients [155] (also called polynomial coefficients [156]), defined by

$$(1 + x + \cdots + x^{d-1})^{N+1} = \sum_{m=0}^{(d-1)(N+1)} \binom{N+1}{m}_d x^m. \quad \text{(C2)}$$

(Note a similar generalization in the proof of Ref. [75], Theorem 1.2). For $d = 2$, $\{|\frac{\pi}{2},\pi\mu\rangle_{J,d=2}\}_{\mu=0}^{1}$ reduce to the two antipodal spin-coherent states discussed in the main text. In general, the d states $\{|\frac{\pi}{2},\frac{2\pi}{d}\mu\rangle_{J,d}\}_{\mu=0}^{d-1}$ are similar to squeezed spin-coherent states equidistantly distributed along the equator of the Bloch sphere. For a fixed J, the amount of squeezing increases with increasing d, as shown in an example in Fig. 11. This is sensible since increasing d for fixed J means fitting more quantum information in the same amount of space.

Mapping the basis $|J,m-J\rangle$ to Fock states $|(S+1)m\rangle$ or two-mode Fock states

$$|(S+1)[(d-1)(N+1) - m],(S+1)m\rangle$$

yields qudit versions of bin and bin2, respectively, for general parameters N,S. It was shown in Ref. [23] that qudit bin codes can protect to order $O(\gamma^N)$ against dephasing and $O(\gamma^S)$ against loss. We do not prove this for bin2 here, but anticipate this to also be the case for those codes. Observing the protection
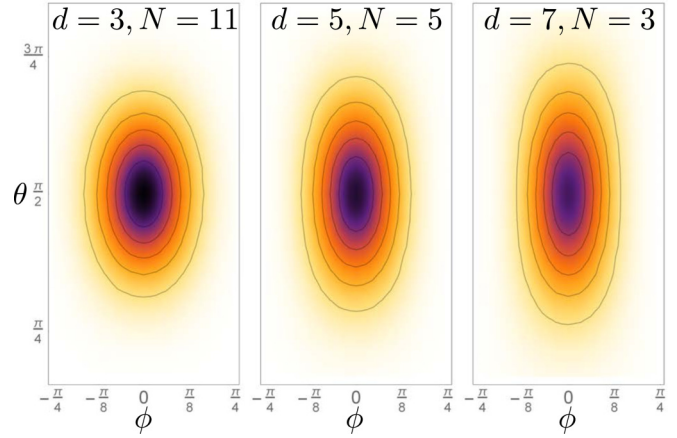

FIG. 11. Overlap $|_J\langle\theta,\phi|\frac{\pi}{2},\frac{2\pi}{d}\mu\rangle_{J,d}|^2$ vs θ,ϕ (in radians) for qudit states $|\frac{\pi}{2},\frac{2\pi}{d}\mu\rangle_{J,d}$ (C1) with $\mu = 0$ and d,N picked such that the spin $J = \frac{1}{2}(d-1)(N+1) = 12$. These states resemble spin-squeezed states and characterize the qu*d*it bin and bin2 codes (see Appendix C). For fixed J, the degree of squeezing increases with d.

offered by noon codes [22], it is also reasonable to believe that tensor products $|\frac{\pi}{2},\frac{2\pi}{d}\mu\rangle_{J,d}^{\otimes M}$ (with $|J,m-J\rangle$ mapped to one- or two-mode Fock states) will yield yet another class of multimode codes. Regarding the multiqubit mapping, we introduce perm$'$—a different extension of qubit perm codes to qudits. These codes can be obtained by mapping $|J,m-J\rangle$ to a qudit generalization of Dicke states $|D_m^{N+1}\rangle$ that we denote as the *extended binomial states* $|E_m^{d-1,N+1}\rangle$, i.e.,

$$|\mu_{\text{perm}'}\rangle = \sum_{m=0}^{(d-1)(N+1)} \frac{e^{i\frac{2\pi}{d}\mu m}}{\sqrt{d^{N+1}}} \sqrt{\binom{N+1}{m}_d} |E_m^{d-1,N+1}\rangle.$$

(C3)

The states $|E_m^{d-1,N+1}\rangle$ are defined as normalized equal superpositions of all multiqudit states having a total of m excitations distributed over $N+1$ qudits,

$$\big|E_m^{d-1,N+1}\big\rangle = \frac{1}{\sqrt{\binom{N+1}{m}_d}} \sum_{\substack{v_1,\cdots,v_{N+1}=0 \\ \sum_i v_i = m}}^{d-1} |v_1,\ldots,v_{N+1}\rangle. \quad \text{(C4)}$$

The normalization of these states happens to be exactly the extended binomial coefficient because $\binom{N+1}{m}_d$ is, by definition, the number of ways of obtaining m as the sum of $N+1$ independent random variables which take values from 0 to $d-1$ [156].

The code perm$'$ is *different* from the qudit perm codes [75] because those utilize a different generalization of qubit Dicke states. For example, extended binomial states for $d = 3$ and $N = 1$ are

$$\big|E_0^{2,2}\big\rangle = |00\rangle, \quad \big|E_1^{2,2}\big\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle), \quad \text{(C5a)}$$

$$\big|E_2^{2,2}\big\rangle = \frac{1}{\sqrt{3}}(|02\rangle + |11\rangle + |20\rangle), \quad \text{(C5b)}$$

$$\big|E_3^{2,2}\big\rangle = \frac{1}{\sqrt{2}}(|21\rangle + |12\rangle), \quad \big|E_4^{2,2}\big\rangle = |22\rangle. \quad \text{(C5c)}$$

In contrast, qudit Dicke states [75] are superpositions of a multiqudit state which has a fixed number of excitations *for each qudit* and all of that state's permutations. For the above case, the qudit Dicke states are $|E_0^{2,2}\rangle, |E_1^{2,2}\rangle, |E_3^{2,2}\rangle, |E_4^{2,2}\rangle$ along with $\frac{1}{\sqrt{2}}(|02\rangle + |20\rangle)$ and $|11\rangle$. In the general case of $N+1$ qu*d*its, there are $\binom{N+d}{d-1}$ qudit Dicke states while only $(d-1)(N+1)+1$ extended binomial states. While the qudit Dicke states span the entire fully symmetric $N+1$-qudit subspace, extended binomial states span only a subspace of that subspace. After introduction of a spacing $S \neq 0$ in similar fashion to `perm` codes (see Sec. VI C 2), it may be that $|\mu_{\text{perm}'}\rangle$ protects against multiqubit amplitude damping, but such properties have yet to be proven.

We conclude this section by relating `perm'` to eigenstates of an N-qu*d*it generalization of an N-qubit collective spin operator. This reveals that such codes are closely related to bit-flip codes and provides a check operator for qudit `bin` and `bin2` codes.

1. Relating `perm'` codes to bit-flip codes

We start with the spin-coherent states from Sec. VI (i.e., $d=2$) written in the irrep for which J_x is a collective operator for a $2J = M$-qubit system with $S=0$. In other words, $J_x = \frac{1}{2}\sum_{k=1}^{M} X_k$, where X_k is the Pauli matrix of the kth qubit. For those parameters, the qubit states $|\mu_{\text{perm}}\rangle$ in this irrep are simply tensor products of eigenstates $|(-1)^\mu\rangle_k$ of X_k,

$$|\mu_{\text{perm}}\rangle = \bigotimes_{k=1}^{M} |(-1)^\mu\rangle_k = \frac{1}{\sqrt{2^M}} \bigotimes_{k=1}^{M} [|0\rangle_k + (-1)^\mu |1\rangle_k] \tag{C6}$$

with $X_k |(-1)^\mu\rangle_k = (-1)^\mu |(-1)^\mu\rangle_k$. Proving this is simple if one writes out $\bigotimes_{k=1}^{M} |(-1)^\mu\rangle_k$ in terms of the Dicke states $\{|D_m^M\rangle\}_{m=1}^{M}$. Observe that, after performing all tensor products, $|\mu_{\text{perm}}\rangle$ will consist of an equal linear superposition of multiqubit states (denoted by binary strings) with coefficients $\pm 1/\sqrt{2^M}$. To change basis to Dicke states, we group multiqubit states by their total number of excitations (i.e., the number of 1's in each binary string). For m excitations out of M qubits, the number of such states is $\binom{M}{m}$. Moreover, since each additional excitation brings about an additional factor of -1, all states with the same number of excitations have matching coefficients. Thus, we can group each superposition of states with fixed excitations into unnormalized Dicke states. Multiplying and dividing each unnormalized Dicke state by $\binom{M}{m}^{-1/2}$ yields the original form of $|\mu_{\text{perm}}\rangle$ in Eq. (6.14).

We can now generalize the above setup to qudits. Consider M qudits of dimension d and let

$$X = \sum_{v=0}^{d-1} |v\rangle\langle v+1 \bmod d| \tag{C7}$$

now be the shift operator for a qudit (i.e., defined such that $X|d-1\rangle = |0\rangle$). This X has d eigenstates

$$|e^{i\frac{2\pi}{d}\mu}\rangle = \frac{1}{\sqrt{d}} \sum_{v=0}^{d-1} e^{i\frac{2\pi}{d}\mu v} |v\rangle \tag{C8}$$

(with $\mu \in \{0,1,\ldots,d-1\}$) whose eigenvalues are $e^{i\frac{2\pi}{d}\mu}$. Using the same procedure as above, one can consider tensor products of $|e^{i\frac{2\pi}{d}\mu}\rangle$,

$$|\mu_{\text{perm}'}\rangle = (|e^{i\frac{2\pi}{d}\mu}\rangle)^{\otimes M}, \tag{C9}$$

and express them in the extended binomial basis. Now, $|\mu_{\text{perm}'}\rangle$ consists of equal superpositions of multiqudit states with coefficients $\{e^{i\frac{2\pi}{d}\mu v}/\sqrt{d^M}\}_{v=0}^{d-1}$, but the coefficients in front of multiqudit states of fixed total excitation match. The normalization of $|E_m^{d-1,M}\rangle$ is the square root of the number of multiqudit states in $|E_m^{d-1,M}\rangle$, which we have already defined to be $\binom{M}{m}_d$. This yields the `perm'` code states from Eq. (C3) with $M = N+1$.

An important consequence of the above description is that now all qudit codes $\in \{\text{bin},\text{bin2}\}$ admit a logical operator

$$X_{\text{code}} = \frac{1}{M} \sum_{k=1}^{M} X_k, \tag{C10}$$

where X_k is X for the kth qudit, and a corresponding check operator $(X_{\text{code}})^M$. It is implied that both of these are projected only onto the subspace spanned by the extended binomial states $\{|E_m^{d-1,M}\rangle\}_{m=0}^{(d-1)M}$ (vs the full permutation-symmetric subspace spanned by the qudit Dicke states discussed above). Thus, the logical operator will be a matrix of dimension $(d-1)M$ satisfying

$$X_{\text{code}} |\mu_{\text{perm}'}\rangle = e^{i\frac{2\pi}{d}\mu} |\mu_{\text{perm}'}\rangle. \tag{C11}$$

Mapping the extended binomial state basis to the corresponding Fock states thus creates analogous check operators for `bin` and `bin2`. One can also consider products of X_k's and form other check operators. Such check operators should prove useful in experimental realizations of the error-correcting procedures of these codes.

APPENDIX D: CALCULATIONS FOR `gkp` CODES

1. Useful identities

Throughout the text, we have used the following standard identities for coherent states $D_\alpha |0\rangle = |\alpha\rangle$ and $|\beta\rangle$:

$$e^{x\hat{n}}|\alpha\rangle = e^{-\frac{1}{2}|\alpha|^2(1-|e^x|^2)}|\alpha e^x\rangle, \tag{D1a}$$

$$D_\alpha D_\beta = e^{\frac{1}{2}(\alpha\beta^\star - \alpha^\star\beta)} D_{\alpha+\beta}, \tag{D1b}$$

$$\langle\alpha|\beta\rangle = e^{-\frac{1}{2}(|\alpha|^2+|\beta|^2)+\alpha^\star\beta}. \tag{D1c}$$

We also use the Fock space matrix elements of D_α [157],

$$\langle\ell|D_\alpha|\ell'\rangle = e^{-\frac{|\alpha|^2}{2}} \sqrt{\frac{\ell'!}{\ell!}} L_{\ell'}^{(\ell-\ell')}(|\alpha|^2) \alpha^{\ell-\ell'}, \tag{D2}$$

for $\ell \geq \ell'$ and $\langle\ell|D_\alpha|\ell'\rangle = (\langle\ell'|D_{-\alpha}|\ell\rangle)^\star$ for $\ell < \ell'$, where $L_n^{(a)}(x)$ is the generalized Laguerre polynomial. The generating function of these polynomials is

$$\sum_{p=0}^{\infty} \frac{(m+p)!}{p!} t^p L_{m+p}^{(\alpha)}(x) = \frac{m! e^{-\frac{tx}{1-t}}}{(1-t)^{m+\alpha+1}} L_m^{(\alpha)}\left(\frac{x}{1-t}\right). \tag{D3}$$

Finally, we use the Poisson summation formula; for a function $f(x)$,

$$\sum_{n\in\mathbb{Z}} f(n) = \sum_{n\in\mathbb{Z}} \int_{-\infty}^{\infty} dx\, e^{2\pi i n x} f(x). \qquad \text{(D4)}$$

2. Equivalence between squeezed and coherent state representations for gkp

We sketch a derivation of Eq. (7.5) from Eq. (7.6). Writing the displacements in Eq. (7.7b) in terms of position and momentum operators $\hat{x}$ and $\hat{p}$, inserting a resolution of the identity in terms of position eigenstates between the displacements, and using $e^{-i x_2 \hat{p}} |x_1\rangle_{\hat{x}} = |x_1 + x_2\rangle_{\hat{x}}$ yields

$$|\mu_{\text{gkps}}^{\text{ideal}}\rangle \propto \sum_{\vec{n}\in\mathbb{Z}^2} \int dx\, |x + \sqrt{\pi}(2n_1+\mu)\rangle_{\hat{x}}\,{}_{\hat{x}}\langle x| e^{i\sqrt{\pi} n_2 \hat{x}} |\text{vac}\rangle, \qquad \text{(D5)}$$

where $\vec{n} = (n_1, n_2)$, $|\text{vac}\rangle$ is the Fock state $|0\rangle$, and we use "$\propto$" to ignore normalization and any constant prefactors that we obtain throughout the calculation. Now, we recall that ${}_{\hat{x}}\langle x|\text{vac}\rangle \propto \exp(-\frac{1}{2}x^2)$ and apply the Poisson summation (D4) to the sum over n_2, yielding a sum over Dirac δ functions. We can then easily evaluate the integral over x, yielding

$$|\mu_{\text{gkps}}^{\text{ideal}}\rangle \propto \sum_{\vec{n}\in\mathbb{Z}_2} e^{-2\pi n_2^2} |\sqrt{\pi}(2n_1 + 2n_2 + \mu)\rangle_{\hat{x}}. \qquad \text{(D6)}$$

Finally, we can redefine indices and evaluate one of the new sums to yield

$$|\mu_{\text{gkps}}^{\text{ideal}}\rangle \propto \sum_{n\in\mathbb{Z}} |\sqrt{\pi}(2n+\mu)\rangle_{\hat{x}}. \qquad \text{(D7)}$$

3. Projecting displacements onto the gkp code space

We have utilized all three representations (7.7a)–(7.7c) to verify the calculations below, initially calculating overlaps $\langle \mu_{\text{gkps}}^{\Delta} | \hat{a}^{\dagger p} \hat{a}^q | \mu_{\text{gkps}}^{\Delta}\rangle$ and summing them up to yield the QEC matrix $\epsilon_{\ell\ell'}^{\text{gkps}}$. We will not report on these calculations, noting that they are cumbersome, but do yield the right answers.

We evaluate matrix elements $\langle \mu_{\text{gkps}}^{\text{ideal}} | D_\alpha | v_{\text{gkps}}^{\text{ideal}}\rangle$ of the displacement operator for ideal gkps states (written in terms of position eigenstates $|\sqrt{\pi}(2n+\mu)\rangle_{\hat{x}}$) from Eq. (7.5). We can split $D_\alpha = D_{\alpha_1 + i\alpha_2}$ into a shift by $\sqrt{2}\alpha_1$ in position and by $\sqrt{2}\alpha_2$ in momentum. The latter translates $|\sqrt{\pi}(2n+\mu)\rangle_{\hat{x}}$ while the former turns into a phase since $|\sqrt{\pi}(2n+\mu)\rangle_{\hat{x}}$ are eigenstates of $\hat{x}$. We can then use the orthogonality of position eigenstates, ${}_{\hat{x}}\langle x_1 | x_2\rangle_{\hat{x}} = \delta(x_1 - x_2)$, and change indices to obtain

$$\langle \mu_{\text{gkps}}^{\text{ideal}} | D_\alpha | v_{\text{gkps}}^{\text{ideal}}\rangle = \sqrt{\frac{2}{\pi}} \sum_{n_1, n_2 \in \mathbb{Z}} e^{-i\alpha_1 \alpha_2} e^{-i\sqrt{2\pi}(2n_2+\mu)\alpha_2}$$
$$\times \delta\left(\alpha_1 - \sqrt{\frac{\pi}{2}}(2n_1 + \delta\mu)\right), \qquad \text{(D8)}$$

where we multiplied each $|\mu_{\text{gkps}}^{\text{ideal}}\rangle$ by $(\frac{2}{\sqrt{\pi}})^{1/2}$ to remove constants in front of the sum (D9) below. We now apply the Poisson summation formula (D4) to turn the sum of n_2-dependent

phases into another sum of Dirac δ functions for α_2, yielding

$$\langle \mu_{\text{gkps}}^{\text{ideal}} | D_\alpha | v_{\text{gkps}}^{\text{ideal}}\rangle = \sum_{\vec{n}\in\mathbb{Z}} e^{i\pi(n_1 + \frac{\mu+\nu}{2})n_2} \delta^2\left(\alpha - \Lambda_{\delta\mu}^{\vec{n}}\right), \qquad \text{(D9)}$$

where $\Lambda_{\delta\mu}^{\vec{n}} = \sqrt{\frac{\pi}{2}}[(2n_1 + \delta\mu) + i n_2]$.

Now let us consider finite gkps states in the smeared representation (7.7c) and calculate

$$\langle \mu_{\text{gkps}}^{\Delta} | D_\alpha | v_{\text{gkps}}^{\Delta}\rangle$$
$$= \int \frac{d^2\beta\, d^2\gamma}{\pi \Delta^2/2} e^{-\frac{1}{\Delta^2}(|\beta|^2 + |\gamma|^2)} \langle \mu_{\text{gkps}}^{\text{ideal}} | D_{-\beta} D_\alpha D_\gamma | v_{\text{gkps}}^{\text{ideal}}\rangle. \qquad \text{(D10)}$$

We add the displacements and use Eq. (D9), whose δ functions allow us to immediately evaluate one of the integrals. The remaining Gaussian integral is also simply evaluated to yield

$$\langle \mu_{\text{gkps}}^{\Delta} | D_\alpha | v_{\text{gkps}}^{\Delta}\rangle$$
$$= \sum_{\vec{n}\in\mathbb{Z}^2} e^{i\pi(n_1 + \frac{\mu+\nu}{2})n_2} e^{-\frac{1}{2\Delta^2}|\alpha - \Lambda_{\delta\mu}^{\vec{n}}|^2} e^{-\frac{\Delta^2}{8}|\alpha + \Lambda_{\delta\mu}^{\vec{n}}|^2}. \qquad \text{(D11)}$$

We can then substitute $\alpha \sim \Lambda_{\delta\mu}^{\vec{n}}$ into the envelope function in the $\Delta \to 0$ limit, yielding Eq. (7.14).

4. QEC matrix for gkp codes

To compute the QEC matrix for gkps, let us sandwich both sides of Eq. (7.12) by $\langle \mu_{\text{gkps}}^{\Delta}|$ and $|v_{\text{gkps}}^{\Delta}\rangle$:

$$\langle \mu_{\text{gkps}}^{\Delta} | E_\ell^\dagger E_{\ell'} | v_{\text{gkps}}^{\Delta}\rangle$$
$$= \int \frac{d^2\alpha}{\pi} e^{-\frac{(1-\gamma)}{2}|\alpha|^2} \langle \ell | D_{\alpha^\star} | \ell'\rangle \langle \mu_{\text{gkps}}^{\Delta} | D_{\sqrt{\gamma}\alpha} | v_{\text{gkps}}^{\Delta}\rangle. \qquad \text{(D12)}$$

Plugging in Eq. (D11) with $\alpha \sim \Lambda_{\delta\mu}^{\vec{n}}$ in the Δ^2-dependent envelope and switching the sum and integral, one obtains

$$\langle \mu_{\text{gkps}}^{\Delta} | E_\ell^\dagger E_{\ell'} | v_{\text{gkps}}^{\Delta}\rangle$$
$$\sim \sum_{\vec{n}\in\mathbb{Z}^2} e^{i\pi(n_1 + \frac{\mu+\nu}{2})n_2} e^{-\frac{\Delta^2}{2}|\Lambda_{\delta\mu}^{\vec{n}}|^2}$$
$$\times \int \frac{d^2\alpha}{\pi} e^{-\frac{(1-\gamma)}{2}|\alpha|^2} \langle \ell | D_{\alpha^\star} | \ell'\rangle e^{-\frac{\gamma}{2\Delta^2}|\alpha - \Lambda_{\delta\mu}^{\vec{n}}/\sqrt{\gamma}|^2}. \qquad \text{(D13)}$$

Next, we evaluate the integral by changing to polar coordinates $\alpha = |\alpha| e^{i\theta}$ and evaluating the angular integral first. This integral turns out to be integral representation of the modified Bessel function of the first kind, $I_n(z) = \int_0^\pi \frac{d\theta}{\pi} e^{z\cos\theta} \cos(n\theta)$. Recalling that $\langle \ell | D_{\alpha^\star} | \ell'\rangle$ (D2) contain Laguerre polynomials, the remaining integral over $|\alpha|$ contains both $L_{\ell_{\min}}^{(|\ell-\ell'|)}$ and $I_{|\ell-\ell'|}$. Luckily, it can be evaluated using Ref. [158], Sec. 2.19.12, Eq. (6):

$$\int_0^\infty dx\, x^{\frac{\lambda}{2}} e^{-px} I_\lambda(2b\sqrt{x}) L_n^{(\lambda)}(x)$$
$$= b^\lambda \frac{(p-1)^n}{p^{\lambda+n+1}} e^{\frac{b^2}{p}} L_n^{(\lambda)}\left(\frac{b^2}{p(p-1)}\right). \qquad \text{(D14)}$$

The prefactor $\frac{(p-1)^n}{p^{n+1}}$ eventually gives the thermal weights in the QEC coefficients $c_{\ell\ell}^{\text{gkps}}$ (7.16). The resulting Laguerre polynomials can then be re-expressed in terms of displacement matrix elements (D2). During this simplification, we take the limit

$$\gamma\left(\frac{1}{2\Delta^2} - \frac{1}{2}\right) \sim \gamma\bar{n}_{\text{gkps}} \gg 1, \qquad \text{(D15)}$$

relating γ to $\bar{n}_{\text{gkps}}$. This yields the QEC matrix elements

$$\langle \mu_{\text{gkps}}^{\Delta} | E_{\ell}^{\dagger} E_{\ell'} | \nu_{\text{gkps}}^{\Delta} \rangle$$

$$\sim \frac{(\gamma\bar{n}_{\text{gkps}})^{\frac{\ell+\ell'}{2}}}{(\gamma\bar{n}_{\text{gkps}}+1)^{\frac{\ell+\ell'}{2}+1}} \sum_{\vec{n}\in\mathbb{Z}^2} e^{-\frac{(1-\gamma)}{2\gamma}|\Lambda_{\delta\mu}^{\vec{n}}|^2}$$

$$\times e^{i\pi(n_1+\frac{\mu+\nu}{2})n_2} e^{-\frac{\Delta^2}{2}|\Lambda_{\delta\mu}^{\vec{n}}|^2} \langle \ell | D_{(\Lambda_{\delta\mu}^{\vec{n}})^{\star}/\sqrt{\gamma}} | \ell' \rangle, \qquad \text{(D16)}$$

where we can once again let $\Delta \to 0$ to produce Eq. (7.15).

[1] S. L. Braunstein and P. van Loock, Quantum information with continuous variables, Rev. Mod. Phys. **77**, 513 (2005).

[2] N. J. Cerf, G. Leuchs, and E. S. Polzik, *Quantum Information with Continuous Variables of Atoms and Light* (World Scientific, London, 2007).

[3] 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).

[4] A. Serafini, *Quantum Continuous Variables: A Primer of Theoretical Methods* (CRC Press, Boca Raton FL, 2017).

[5] I. L. Chuang and Y. Yamamoto, Simple quantum computer, Phys. Rev. A **52**, 3489 (1995).

[6] P. Kok, W. J. Munro, K. Nemoto, T. C. Ralph, J. P. Dowling, and G. J. Milburn, Linear optical quantum computing with photonic qubits, Rev. Mod. Phys. **79**, 135 (2007).

[7] S. Lloyd and J.-J. E. Slotine, Analog Quantum Error Correction, Phys. Rev. Lett. **80**, 4088 (1998).

[8] S. L. Braunstein, Error Correction for Continuous Quantum Variables, Phys. Rev. Lett. **80**, 4084 (1998).

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

[10] N. C. Menicucci, Fault-Tolerant Measurement-Based Quantum Computing with Continuous-Variable Cluster States, Phys. Rev. Lett. **112**, 120504 (2014).

[11] P. Hayden, S. Nezami, G. Salton, and B. C. Sanders, Spacetime replication of continuous variable quantum information, New J. Phys. **18**, 083043 (2016).

[12] A. Ketterer, A. Keller, S. P. Walborn, T. Coudreau, and P. Milman, Quantum information processing in phase space: A modular variables approach, Phys. Rev. A **94**, 022325 (2016).

[13] P. T. Cochrane, G. J. Milburn, and W. J. Munro, Macroscopically distinct quantum-superposition states as a bosonic code for amplitude damping, Phys. Rev. A **59**, 2631 (1999).

[14] J. Niset, U. L. Andersen, and N. J. Cerf, Experimentally Feasible Quantum Erasure-Correcting Code for Continuous Variables, Phys. Rev. Lett. **101**, 130503 (2008).

[15] Z. Leghtas, G. Kirchmair, B. Vlastakis, R. J. Schoelkopf, M. H. Devoret, and M. Mirrahimi, Hardware-Efficient Autonomous Quantum Memory Protection, Phys. Rev. Lett. **111**, 120501 (2013).

[16] F. Lacerda, J. M. Renes, and V. B. Scholz, in *2016 IEEE International Symposium on Information Theory (ISIT)* (IEEE, 2016), pp. 2499–2503.

[17] V. V. Albert, S. O. Mundhada, A. Grimm, S. Touzard, M. H. Devoret, and L. Jiang, Multimode cat codes, arXiv:1801.05897.

[18] I. L. Chuang, D. W. Leung, and Y. Yamamoto, Bosonic quantum codes for amplitude damping, Phys. Rev. A **56**, 1114 (1997).

[19] E. Knill, R. Laflamme, and G. J. Milburn, A scheme for efficient quantum computation with linear optics, Nature (London) **409**, 46 (2001).

[20] T. C. Ralph, A. J. F. Hayes, and A. Gilchrist, Loss-Tolerant Optical Qubits, Phys. Rev. Lett. **95**, 100501 (2005).

[21] W. Wasilewski and K. Banaszek, Protecting an optical qubit against photon loss, Phys. Rev. A **75**, 042316 (2007).

[22] M. Bergmann and P. van Loock, Quantum error correction against photon loss using NOON states, Phys. Rev. A **94**, 012311 (2016).

[23] M. H. Michael, M. Silveri, R. T. Brierley, V. V. Albert, J. Salmilehto, L. Jiang, and S. M. Girvin, New Class of Quantum Error-Correcting Codes for a Bosonic Mode, Phys. Rev. X **6**, 031006 (2016).

[24] M. Y. Niu, I. L. Chuang, and J. H. Shapiro, Hardware-efficient bosonic quantum error-correcting codes based on symmetry operators, arXiv:1709.05302.

[25] S.-W. Lee and H. Jeong, Near-deterministic quantum teleportation and resource-efficient quantum computation using linear optics and hybrid qubits, Phys. Rev. A **87**, 022326 (2013).

[26] E. Kapit, Hardware-Efficient and Fully Autonomous Quantum Error Correction in Superconducting Circuits, Phys. Rev. Lett. **116**, 150501 (2016).

[27] 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).

[28] D. Kribs and D. Poulin, in *Quantum Error Correct*, edited by D. A. Lidar, T. A. Brun, and T. Brun (Cambridge University Press, Cambridge, UK), pp. 163–180.

[29] N. Ofek, A. Petrenko, R. Heeres, P. Reinhold, Z. Leghtas, B. Vlastakis, Y. Liu, L. Frunzio, S. M. Girvin, L. Jiang, M. Mirrahimi, M. H. Devoret, and R. J. Schoelkopf, Extending the lifetime of a quantum bit with error correction in superconducting circuits, Nature (London) **536**, 441 (2016).

[30] M. Ueda, Probability-density-functional description of quantum photodetection processes, Quantum Opt. **1**, 131 (1989).

[31] C. T. Lee, Superoperators and their implications in the hybrid model for photodetection, Phys. Rev. A **49**, 4888 (1994).

[32] D. Vitali, P. Tombesi, and G. J. Milburn, Quantum-state protection in cavities, Phys. Rev. A **57**, 4930 (1998).

[33] A. B. Klimov and S. M. Chumakov, *A Group-Theoretical Approach to Quantum Optics* (Wiley, Weinheim, 2009).

[34] J. S. Ivan, K. K. Sabapathy, and R. Simon, Operator-sum representation for bosonic Gaussian channels, Phys. Rev. A **84**, 042311 (2011).

[35] A. S. Fletcher, Channel-adapted quantum error correction, Ph.D. thesis, MIT, Cambridge, MA, 2007, arXiv:0706.3400.

[36] C. M. Caves, Quantum error correction and reversible operations, J. Supercond. **12**, 707 (1999).

[37] D. W. Leung, M. A. Nielsen, I. L. Chuang, and Y. Yamamoto, Approximate quantum error correction can lead to better codes, Phys. Rev. A **56**, 2567 (1997).

[38] C. Crépeau, D. Gottesman, and A. Smith, in *Advances in Cryptology EUROCRYPT 2005*, Lecture Notes in Computer Science Vol. 3494, edited by R. Cramer (Springer, Berlin, 2005), pp. 285–301.

[39] M. Reimpell and R. F. Werner, Iterative Optimization of Quantum Error Correcting Codes, Phys. Rev. Lett. **94**, 080501 (2005).

[40] B. W. Schumacher and M. D. Westmoreland, *Quantum Processes, Systems, and Information* (Cambridge University Press, New York, 2010).

[41] R. Konig, R. Renner, and C. Schaffner, The operational meaning of min- and max-entropy, IEEE Trans. Inf. Theory **55**, 4337 (2009).

[42] M. Tomamichel, A framework for non-asymptotic quantum information theory, Ph.D. thesis, ETH Zurich, Zurich, Switzerland, 2012, arXiv:1203.2142.

[43] A. S. Fletcher, P. W. Shor, and M. Z. Win, Optimum quantum error recovery using semidefinite programming, Phys. Rev. A **75**, 012338 (2007).

[44] K. M. R. Audenaert and B. De Moor, Optimizing completely positive maps using semidefinite programming, Phys. Rev. A **65**, 030302 (2002).

[45] R. L. Kosut and D. A. Lidar, Quantum error correction via convex optimization, Quant. Info. Proc. **8**, 443 (2009).

[46] L. Li, C.-l. Zou, V. V. Albert, S. Muralidharan, S. M. Girvin, and L. Jiang, Cat Codes with Optimal Decoherence Suppression for a Lossy Bosonic Channel, Phys. Rev. Lett. **119**, 030502 (2017).

[47] J. Preskill, Quantum computing in the NISQ era and beyond, arXiv:1801.00862.

[48] M. Mirrahimi, Cat-qubits for quantum computation, C. R. Phys. **17**, 778 (2016).

[49] Z. Leghtas and M. Mirrahimi (private communication).

[50] F. Caruso, V. Giovannetti, and A. S. Holevo, One-mode bosonic Gaussian channels: A full weak-degradability classification, New J. Phys. **8**, 310 (2006).

[51] D. Gottesman and J. Preskill, Secure quantum key distribution using squeezed states, Phys. Rev. A **63**, 022309 (2001).

[52] R. Garcia-Patron, C. Navarrete-Benlloch, S. Lloyd, J. H. Shapiro, and N. J. Cerf, Majorization Theory Approach to the Gaussian Channel Minimum Entropy Conjecture, Phys. Rev. Lett. **108**, 110505 (2012).

[53] J. Harrington and J. Preskill, Achievable rates for the Gaussian quantum channel, Phys. Rev. A **64**, 062301 (2001).

[54] K. Noh, V. V. Albert, and L. Jiang, Improved quantum capacity bounds of Gaussian loss channels and achievable rates with Gottesman-Kitaev-Preskill codes, arXiv:1801.07271.

[55] M. Rosati, A. Mari, and V. Giovannetti, Narrow bounds for the quantum capacity of thermal attenuators, arXiv:1801.04731.

[56] K. Sharma, M. M. Wilde, S. Adhikari, and M. Takeoka, Bounding the energy-constrained quantum and private capacities of phase-insensitive Gaussian channels, arXiv:1708.07257.

[57] M. M. Wilde, P. Hayden, and S. Guha, Quantum trade-off coding for bosonic communication, Phys. Rev. A **86**, 062306 (2012).

[58] M. M. Wilde and H. Qi, Energy-constrained private and quantum capacities of quantum channels, arXiv:1609.01997v1.

[59] M. M. Wilde, P. Hayden, and S. Guha, Information Trade-Offs for Optical Quantum Communication, Phys. Rev. Lett. **108**, 140501 (2012).

[60] M. M. Wolf, D. Pérez-García, and G. Giedke, Quantum Capacities of Bosonic Channels, Phys. Rev. Lett. **98**, 130501 (2007).

[61] A. S. Holevo and R. F. Werner, Evaluating capacities of bosonic Gaussian channels, arXiv:quant-ph/9912067 [quant-ph].

[62] M. M. Wilde, *Quantum Information Theory*, 2nd ed. (Cambridge University Press, Cambridge, UK, 2013).

[63] R. García-Patrón, S. Pirandola, S. Lloyd, and J. H. Shapiro, Reverse Coherent Information, Phys. Rev. Lett. **102**, 210501 (2009).

[64] S. Pirandola, R. García-Patrón, S. L. Braunstein, and S. Lloyd, Direct and Reverse Secret-Key Capacities of a Quantum Channel, Phys. Rev. Lett. **102**, 050503 (2009).

[65] S. Pirandola, R. Laurenza, C. Ottaviani, and L. Banchi, Fundamental limits of repeaterless quantum communications, Nat. Commun. **8**, 15043 (2017).

[66] M. M. Wolf and D. Perez-Garcia, Quantum capacities of channels with small environment, Phys. Rev. A **75**, 012303 (2007).

[67] D. Gottesman, Stabilizer codes and quantum error correction, Ph.D. thesis, California Institute of Technology, Pasadena, CA, 1997, arXiv:quant-ph/9705052.

[68] R. Lang and P. W. Shor, Nonadditive quantum error correcting codes adapted to the ampltitude damping channel, arXiv:0712.2586.

[69] A. S. Fletcher, P. W. Shor, and M. Z. Win, Channel-adapted quantum error correction for the amplitude damping channel, IEEE Trans. Inf. Theory **54**, 5705 (2008).

[70] R. Duan, M. Grassl, Z. Ji, and B. Zeng, in *2010 IEEE International Symposium on Information Theory* (IEEE, 2010), pp. 2672–2676.

[71] P. W. Shor, G. Smith, J. A. Smolin, and B. Zeng, High performance single-error-correcting quantum codes for amplitude damping, IEEE Trans. Inf. Theory **57**, 7180 (2011).

[72] C. Han, Quantum error correction of photon loss with quantum encoding, J. Korean Phys. Soc. **62**, 721 (2013).

[73] Y. Ouyang, Permutation-invariant quantum codes, Phys. Rev. A **90**, 062317 (2014).

[74] T. Jackson, M. Grassl, and B. Zeng, in *2016 IEEE International Symposium on Information Theory (ISIT)* (IEEE, 2016), pp. 2269–2273.

[75] Y. Ouyang, Permutation-invariant qudit codes from polynomials, Linear Algebr. Appl. **532**, 43 (2017).

[76] Y. Ouyang and J. Fitzsimons, Permutation-invariant codes encoding more than one qubit, Phys. Rev. A **93**, 042340 (2016).

[77] M. Grassl, T. Beth, and T. Pellizzari, Codes for the quantum erasure channel, Phys. Rev. A **56**, 33 (1997).

[78] L. Vaidman, L. Goldenberg, and S. Wiesner, Error prevention scheme with four particles, Phys. Rev. A **54**, R1745 (1996).

[79] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Mixed-state entanglement and quantum error correction, Phys. Rev. A **54**, 3824 (1996).

[80] E. Knill and R. Laflamme, Theory of quantum error-correcting codes, Phys. Rev. A **55**, 900 (1997).

[81] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, UK, 2011).

[82] J. Preskill, Lecture notes on Quantum Computation (2004), http://www.theory.caltech.edu/people/preskill/ph229/.

[83] V. V. Albert and L. Jiang, Symmetries and conserved quantities in Lindblad master equations, Phys. Rev. A **89**, 022118 (2014).

[84] H. Jeong and M. S. Kim, Efficient quantum computation using coherent states, Phys. Rev. A **65**, 042305 (2002).

[85] T. C. Ralph, A. Gilchrist, G. J. Milburn, W. J. Munro, and S. Glancy, Quantum computation with optical coherent states, Phys. Rev. A **68**, 042319 (2003).

[86] S. Glancy, H. M. Vasconcelos, and T. C. Ralph, Transmission of optical coherent-state qubits, Phys. Rev. A **70**, 022317 (2004).

[87] A. P. Lund, T. C. Ralph, and H. L. Haselgrove, Fault-Tolerant Linear Optical Quantum Computing with Small-Amplitude Coherent States, Phys. Rev. Lett. **100**, 030503 (2008).

[88] V. V. Albert, C. Shu, S. Krastanov, C. Shen, R.-B. Liu, Z.-B. Yang, R. J. Schoelkopf, M. Mirrahimi, M. H. Devoret, and L. Jiang, Holonomic Quantum Control with Continuous Variable Systems, Phys. Rev. Lett. **116**, 140502 (2016).

[89] M. Bergmann and P. van Loock, Quantum error correction against photon loss using multicomponent cat states, Phys. Rev. A **94**, 042332 (2016).

[90] U. L. Andersen, J. S. Neergaard-Nielsen, P. van Loock, and A. Furusawa, Hybrid discrete- and continuous-variable quantum information, Nat. Phys. **11**, 713 (2015).

[91] Z. Leghtas, S. Touzard, I. M. Pop, A. Kou, B. Vlastakis, A. Petrenko, K. M. Sliwa, A. Narla, S. Shankar, M. J. Hatridge, M. Reagor, L. Frunzio, R. J. Schoelkopf, M. Mirrahimi, and M. H. Devoret, Confining the state of light to a quantum manifold by engineered two-photon loss, Science **347**, 853 (2015).

[92] V. V. Albert, Lindbladians with multiple steady states: Theory and applications, Ph.D. thesis, Yale University, New Haven, CT, 2017, arXiv:1802.00010.

[93] D. Stoler, B. Saleh, and M. Teich, Binomial states of the quantized radiation field, Opt. Acta (London) **32**, 345 (1985).

[94] V. Dodonov and V. Man'ko, *Theory of Nonclassical States of Light* (CRC Press, New York, 2003).

[95] G. Dattoli, J. Gallardo, and A. Torre, Binomial states of the quantized radiation field: Comment, J. Opt. Soc. Am. B **4**, 185 (1987).

[96] F. Hong-yi and J. Si-cong, Connection of a type of q-deformed binomial state with q-spin coherent states, Phys. Rev. A **50**, 1909 (1994).

[97] L. I. Schiff, *Quantum Mechanics* (McGraw-Hill, New York, 1968).

[98] F. T. Arecchi, E. Courtens, R. Gilmore, and H. Thomas, Atomic coherent states in quantum optics, Phys. Rev. A **6**, 2211 (1972).

[99] R. R. Puri, *Mathematical Methods of Quantum Optics* (Springer, Berlin, 2001).

[100] Y. Ouyang, Transmitting quantum information reliably across various quantum channels, Ph.D. thesis, University of Waterloo, Waterloo, Canada, 2013.

[101] Z. Shaterzadeh-Yazdi, P. S. Turner, and B. C. Sanders, SU(1,1) symmetry of multimode squeezed states, J. Phys. A: Math. Theor. **41**, 055309 (2008).

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

[103] Y. Aharonov, H. Pendleton, and A. Petersen, Modular variables in quantum theory, Int. J. Theor. Phys. **2**, 213 (1969).

[104] J. Zak, Finite Translations in Solid-State Physics, Phys. Rev. Lett. **19**, 1385 (1967).

[105] K. E. Cahill and R. J. Glauber, Ordered expansions in boson amplitude operators, Phys. Rev. **177**, 1857 (1969).

[106] B. M. Terhal and D. Weigand, Encoding a qubit into a cavity mode in circuit QED using phase estimation, Phys. Rev. A **93**, 012315 (2016).

[107] P. P. Hofer, J.-R. Souquet, and A. A. Clerk, Quantum heat engine based on photon-assisted Cooper pair tunneling, Phys. Rev. B **93**, 041418 (2016).

[108] A. B. Klimov and J. L. Romero, An algebraic solution of Lindblad-type master equations, J. Opt. B: Quantum Semiclassical Opt. **5**, S316 (2003).

[109] L. Kunz, M. G. A. Paris, and K. Banaszek, Noisy propagation of coherent states in a lossy Kerr medium, J. Opt. Soc. Am. B **35**, 214 (2018).

[110] K. Duivenvoorden, B. M. Terhal, and D. Weigand, Single-mode displacement sensor, Phys. Rev. A **95**, 012305 (2017).

[111] B. Yurke and D. Stoler, Generating Quantum Mechanical Superpositions of Macroscopically Distinguishable States Via Amplitude Dispersion, Phys. Rev. Lett. **57**, 13 (1986).

[112] A. Miranowicz, R. Tanas, and S. Kielich, Generation of discrete superpositions of coherent states in the anharmonic oscillator model, Quantum Opt. **2**, 253 (1990).

[113] K. Tara, G. S. Agarwal, and S. Chaturvedi, Production of Schrodinger macroscopic quantum-superposition states in a Kerr medium, Phys. Rev. A **47**, 5024 (1993).

[114] G. J. Milburn and C. A. Holmes, Dissipative Quantum and Classical Liouville Mechanics of the Anharmonic Oscillator, Phys. Rev. Lett. **56**, 2237 (1986).

[115] L. Sun, A. Petrenko, Z. Leghtas, B. Vlastakis, G. Kirchmair, K. M. Sliwa, A. Narla, M. Hatridge, S. Shankar, J. Blumoff, L. Frunzio, M. Mirrahimi, M. H. Devoret, and R. J. Schoelkopf, Tracking photon jumps with repeated quantum non-demolition parity measurements, Nature (London) **511**, 444 (2014).

[116] L. G. Lutterbach and L. Davidovich, Method for Direct Measurement of the Wigner Function in Cavity QED and Ion Traps, Phys. Rev. Lett. **78**, 2547 (1997).

[117] A. Barchielli and M. Gregoratti, *Quantum Trajectories and Measurements in Continuous Time: The Diffusive Case* (Springer, Berlin, 2009).

[118] R. W. Heeres, P. Reinhold, N. Ofek, L. Frunzio, L. Jiang, M. H. Devoret, and R. J. Schoelkopf, Implementing a universal gate set on a logical qubit encoded in an oscillator, Nat. Commun. **8**, 94 (2017).

[119] W. Chen, J. Hu, Y. Duan, B. Braverman, H. Zhang, and V. Vuletić, Carving Complex Many-Atom Entangled States by Single-Photon Detection, Phys. Rev. Lett. **115**, 250502 (2015).

[120] R. McConnell, H. Zhang, J. Hu, S. Cuk, and V. Vuletic, Entanglement with negative Wigner function of almost 3,000 atoms heralded by one photon, Nature (London) **519**, 439 (2016).

[121] D. Kienzler, H.-Y. Lo, V. Negnevitsky, C. Flühmann, M. Marinelli, and J. P. Home, Quantum Harmonic Oscillator State Control in a Squeezed Fock Basis, Phys. Rev. Lett. **119**, 033602 (2017).

[122] K. R. Motes, B. Q. Baragiola, A. Gilchrist, and N. C. Menicucci, Encoding qubits into oscillators with atomic ensembles and squeezed light, Phys. Rev. A **95**, 053819 (2017).

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

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

[125] S. Pirandola, S. Mancini, D. Vitali, and P. Tombesi, Continuous variable encoding by ponderomotive interaction, Eur. Phys. J. D **37**, 283 (2006).

[126] S. Pirandola, S. Mancini, D. Vitali, and P. Tombesi, Generating continuous variable quantum codewords in the near-field atomic lithography, J. Phys. B: At. Mol. Opt. Phys. **39**, 997 (2006).

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

[128] D. J. Weigand and B. M. Terhal, Generating grid states from Schrodinger cat states without post-selection, Phys. Rev. A **97**, 022341 (2018).

[129] C. Flühmann, V. Negnevitsky, M. Marinelli, and J. P. Home, Sequential Modular Position and Momentum Measurements of a Trapped Ion Mechanical Oscillator, Phys. Rev. X (to be published).

[130] J. Hu, W. Chen, Z. Vendeiro, H. Zhang, and V. Vuletic, Entangled collective-spin states of atomic ensembles under nonuniform atom-light interaction, Phys. Rev. A **92**, 063816 (2015).

[131] D. Budker and M. Romalis, Optical magnetometry, Nat. Phys. **3**, 227 (2007).

[132] J. M. Taylor, P. Cappellaro, L. Childress, L. Jiang, D. Budker, P. R. Hemmer, A. Yacoby, R. Walsworth, and M. D. Lukin, High-sensitivity diamond magnetometer with nanoscale resolution, Nat. Phys. **4**, 810 (2008).

[133] V. M. Acosta, E. Bauch, M. P. Ledbetter, C. Santori, K.-M. C. Fu, P. E. Barclay, R. G. Beausoleil, H. Linget, J. F. Roch, F. Treussart, S. Chemerisov, W. Gawlik, and D. Budker, Diamonds with a high density of nitrogen-vacancy centers for magnetometry applications, Phys. Rev. B **80**, 115202 (2009).

[134] M. Horodecki, P. Horodecki, and R. Horodecki, General teleportation channel, singlet fraction, and quasidistillation, Phys. Rev. A **60**, 1888 (1999).

[135] M. A. Nielsen, A simple formula for the average gate fidelity of a quantum dynamical operation, Phys. Lett. A **303**, 249 (2002).

[136] H. Barnum, E. Knill, and M. A. Nielsen, On quantum fidelities and channel capacities, IEEE Trans. Inf. Theory **46**, 1317 (2000).

[137] A. Gilchrist, N. K. Langford, and M. A. Nielsen, Distance measures to compare real and ideal quantum processes, Phys. Rev. A **71**, 062310 (2005).

[138] H. K. Ng and P. Mandayam, Simple approach to approximate quantum error correction based on the transpose channel, Phys. Rev. A **81**, 062342 (2010).

[139] R. Kueng, D. M. Long, A. C. Doherty, and S. T. Flammia, Comparing Experiments to the Fault-Tolerance Threshold, Phys. Rev. Lett. **117**, 170502 (2016).

[140] J. J. Wallman, Bounding experimental quantum error rates relative to fault-tolerant thresholds, arXiv:1511.00727.

[141] S. T. Flammia, Quantifying photon loss errors in quantum gates (private communication).

[142] C. Fuchs and J. van de Graaf, Cryptographic distinguishability measures for quantum-mechanical states, IEEE Trans. Inf. Theory **45**, 1216 (1999).

[143] J. J. Wallman and S. T. Flammia, Randomized benchmarking with confidence, New J. Phys. **16**, 103032 (2014).

[144] J. J. Wallman and S. T. Flammia, Corrigendum: Randomized benchmarking with confidence (2014 New J. Phys. 16 103032), New J. Phys. **18**, 079501 (2016).

[145] B. Schumacher, Sending entanglement through noisy quantum channels, Phys. Rev. A **54**, 2614 (1996).

[146] O. Fawzi and R. Renner, Quantum conditional mutual information and approximate Markov chains, Commun. Math. Phys. **340**, 575 (2015).

[147] K. P. Seshadreesan and M. M. Wilde, Fidelity of recovery, squashed entanglement, and measurement recoverability, Phys. Rev. A **92**, 042321 (2015).

[148] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevA.97.032346 for a MATHEMATICA notebook listing the five num code states discussed in Appendix B.

[149] J. Tyson, Two-sided bounds on minimum-error quantum measurement, on the reversibility of quantum dynamics, and on maximum overlap using directional iterates, J. Math. Phys. **51**, 092204 (2010).

[150] C. Bény and O. Oreshkov, General Conditions for Approximate Quantum Error Correction and Near-Optimal Recovery Channels, Phys. Rev. Lett. **104**, 120501 (2010).

[151] C. Bény and O. Oreshkov, Approximate simulation of quantum channels, Phys. Rev. A **84**, 022333 (2011).

[152] M. Grant and S. Boyd, CVX: MATLAB software for disciplined convex programming (2013), http://cvxr.com/cvx/.

[153] R. S. Menon and S. Horvát, MATLink (2013), http://matlink.org/.

[154] T. S. Cubitt, MATLAB code (2016), http://www.dr-qubit.org.

[155] T. Neuschel, A note on extended binomial coefficients, J. Integer Seq. **17**, 14.10.4 (2014).

[156] C. Caiado and P. Rathie, in *Eighth Annual Conference of the Society of Special Functions and their Applications*, edited by A. M. Mathai, M. A. Pathan, K. Jose, and J. Jacob (Society for Special Functions and Their Applications, 2007).

[157] W. Vogel and D.-G. Welsch, *Quantum Optics*, 3rd ed. (Wiley, Weinheim, 2006).

[158] A. P. Prudnikov, Y. A. Brychkov, and O. I. Marichev, *Special Functions*, Integrals and Series Vol. 2 (Gordon and Breach Science, 1998).