Approximate Autonomous Quantum Error Correction with Reinforcement Learning

Yexiong Zeng,[1,2] Zheng-Yang Zhou,[1] Enrico Rinaldi,[3,1,2,4,5] Clemens Gneiting,[1,2,*] and Franco Nori[1,2,4,†]
[1]Theoretical Quantum Physics Laboratory, Cluster for Pioneering Research, RIKEN, Wakoshi, Saitama 351-0198, Japan
[2]Quantum Computing Center, RIKEN, Wakoshi, Saitama 351-0198, Japan
[3]Quantinuum K.K., Otemachi Financial City Grand Cube 3F, 1-9-2 Otemachi, Chiyoda-ku, Tokyo, Japan
[4]Department of Physics, University of Michigan, Ann Arbor, Michigan 48109-1040, USA
[5]Interdisciplinary Theoretical and Mathematical Sciences Program (iTHEMS), RIKEN, Wakoshi, Saitama 351-0198, Japan

(Received 22 December 2022; accepted 22 June 2023; published 31 July 2023)

Autonomous quantum error correction (AQEC) protects logical qubits by engineered dissipation and thus circumvents the necessity of frequent, error-prone measurement-feedback loops. Bosonic code spaces, where single-photon loss represents the dominant source of error, are promising candidates for AQEC due to their flexibility and controllability. While existing proposals have demonstrated the in-principle feasibility of AQEC with bosonic code spaces, these schemes are typically based on the exact implementation of the Knill-Laflamme conditions and thus require the realization of Hamiltonian distances $d \geq 2$. Implementing such Hamiltonian distances requires multiple nonlinear interactions and control fields, rendering these schemes experimentally challenging. Here, we propose a bosonic code for approximate AQEC by relaxing the Knill-Laflamme conditions. Using reinforcement learning (RL), we identify the optimal bosonic set of code words (denoted here by RL code), which, surprisingly, is composed of the Fock states $|2\rangle$ and $|4\rangle$. As we show, the RL code, despite its approximate nature, successfully suppresses single-photon loss, reducing it to an effective dephasing process that well surpasses the break-even threshold. It may thus provide a valuable building block toward full error protection. The error-correcting Hamiltonian, which includes ancilla systems that emulate the engineered dissipation, is entirely based on the Hamiltonian distance $d = 1$, significantly reducing model complexity. Single-qubit gates are implemented in the RL code with a maximum distance $d_g = 2$.

DOI: 10.1103/PhysRevLett.131.050601

Introduction.—Implementing efficient quantum error correction (QEC) is a prerequisite and the main obstacle toward building a general-purpose quantum computer [1–3]. The purpose of QEC is to restore encoded quantum information that has been corrupted by environmental noise: physical qubits inevitably interact with the surrounding environment [4,5]. In conventional QEC this is achieved by error syndrome measurements and adaptive recovery operations, which involve imperfect measurements and classical feedback loops that themselves raise and propagate errors [6–14].

Autonomous quantum error correction (AQEC) has been developed as an alternative that avoids these additional error sources by realizing QEC through quantum reservoir engineering, where the cleverly designed and continuously acting dissipation processes replace the repeated measurement-feedback cycles [15–19]. More specifically, the interaction between the to-be-protected system and an ancilla system is engineered to transport the decoherence-induced, cumulative entropy from the system to the ancilla, from where it decays into the environment [20–28].

The realization of AQEC in bosonic systems is particularly attractive, as an infinitely large Hilbert space offers ample design opportunities while the noise channels remain fixed. Indeed, several pioneering works have demonstrated the great potential of AQEC to counteract single-photon loss, the dominant error source in bosonic systems; for instance, based on the T4C code [15], the binomial code [29], or the $\sqrt{3}$ code [30]. While these code words fully satisfy the Knill-Laflamme (KL) conditions, and hence in principle allow for exact AQEC [31,32], experimental limitations have so far prohibited their faithful implementation. Therefore, the experimental surpassing of the break-even threshold (i.e., improved performance compared to code words consisting of the Fock states $|0\rangle$ and $|1\rangle$ and no error correction) using AQEC is still lacking.

Rigorous implementation of the KL condition unavoidably implies code words that entail complex and fragile superpositions of Fock states, which are hard to produce. This is reflected by the required Hamiltonian distance, i.e., the number of Fock states that must be bridged by the Hamiltonian. Indeed, an exhaustive search has excluded the existence of a code space that both rigorously satisfies the KL condition and is content with Hamiltonian distance $d = 1$ for AQEC [30]. On the other hand, the larger the Hamiltonian distance, the more high-order interactions are required, accompanied by an increasing number of control fields. The experimental hardness is further aggravated by

the necessity to implement gates, which typically require even larger Hamiltonian distances.

Here, we pursue a different strategy, following the spirit of approximate QEC [33–37]. By relaxing the KL condition, we strive to ease the experimental overhead by lowering the Hamiltonian distance as much as possible. While this excludes the in-principle exact QEC, we consider our strategy successful if the discovered AQEC surpasses the break-even threshold, i.e., if the encoded qubit outperforms the natural qubit and thus lowers the threshold for a concatenated QEC stack.

The search for the optimal approximate AQEC scheme that achieves this involves a complex optimization problem, which we solve with the help of reinforcement learning (RL) [38–48]. We find that the optimal AQEC scheme relies on surprisingly simple code words (denoted here by RL code) composed of the two Fock states $|2\rangle$ and $|4\rangle$. The RL code not only significantly surpasses the break-even with an infidelity reduction of over 80% (surprisingly even outperforming the $\sqrt{3}$ code), but also can be realized with the smallest possible Hamiltonian distance $d = 1$; the respective Hamiltonian distance required to implement single-qubit gates is $d_g = 2$, which again outperforms all previously proposed code spaces.

We demonstrate that the approximate AQEC based on the RL code can be realized by complementing the encoded bosonic mode with an ancillary lossy mode, an ancillary lossy two-level system, and couplings that are readily available on existing platforms.

Approximate AQEC.—AQEC employs engineered dissipation to recover from the errors that occur due to natural decay processes. In brief, engineered Lindblad operators $\sum_k \mathcal{D}[L_{\mathrm{eng},k}]$ are introduced such that the overall dissipative channel $\mathcal{M}[\cdot]$, which also includes the natural Lindblad operators $\sum_i \mathcal{D}[L_{\mathrm{nat},i}]$, minimizes the growth of the infidelity between an arbitrary qubit state $\rho_{t_0}(\theta,\phi) = |\psi_{\theta\phi}\rangle\langle\psi_{\theta\phi}|$ in the code space and the evolved state $\rho_t(\theta,\phi) = \mathcal{M}[\rho_{t_0}(\theta,\phi)]$. This happens by enlarging the effect of $\sum_k \mathcal{D}[L_{\mathrm{eng},k}]$ [49], where $\mathcal{D}[x] = 2x\rho x^\dagger - x^\dagger x\rho - \rho x^\dagger x$, and the angles θ and ϕ parametrize an arbitrary qubit state in the Bloch-sphere representation.

The mean fidelity $\bar{F}(|0_L\rangle, |1_L\rangle)$, and hence the performance of the AQEC, depends on the choice of the logical code words $|0_L\rangle$ and $|1_L\rangle$. If the logical code words satisfy the KL condition [32],

$$\langle u_L|L_{\mathrm{nat},j}^\dagger L_{\mathrm{nat},i}|v_L\rangle = \alpha_{ji}\delta_{uv}, \quad u,v \in \{0,1\}, \tag{1}$$

where α_{ji} are the elements of a Hermitian matrix, one can, in principle, achieve exact QEC. However, in the context of AQEC for bosonic codes, where $L_{\mathrm{nat},i} \in \{I, a\}$ [50–52], the exact implementation of the KL conditions can be challenging. For instance, the necessity to engineer multiple Lindblad operators, as is the case with the binomial and the T4C code, dramatically increases experimental complexity. A similar challenge occurs if multiple high-order nonlinear interactions, along with multiple control fields, are required, as for the $\sqrt{3}$ code.

It has been shown that approximate QEC (i.e., the code words partially satisfy the KL condition) can combine experimental feasibility with good QEC performance [53–56]. Following this line of research, we relax the KL condition by discarding the constraint that $\langle 1_L|L_{\mathrm{nat},i}^\dagger L_{\mathrm{nat},i}|1_L\rangle = \langle 0_L|L_{\mathrm{nat},i}^\dagger L_{\mathrm{nat},i}|0_L\rangle$; i.e., we allow the logical code words to exhibit different error probabilities. General code words (selecting the even subspace) that satisfy the remaining KL conditions can then be parametrized as

$$|0_L\rangle = \sum_{n=0} c_n^{(0)}|4n\rangle, \quad |1_L\rangle = \sum_{n=0} c_n^{(1)}|4n+2\rangle, \tag{2}$$

where the $c_n^{(0)}$ and $c_n^{(1)}$ denote undetermined real coefficients and satisfy $\sum_n |c_n^{(u)}|^2 = 1$ ($u = 0, 1$). Our goal is now to find coefficients $c_n^{(0)}$ and $c_n^{(1)}$ that optimize the performance of the corresponding AQEC with only a single, fixed QEC jump operator,

$$\begin{aligned} L_{\mathrm{eng}} &= L_{\mathrm{o}}\{\mathrm{Tr}[L_{\mathrm{o}}^\dagger L_{\mathrm{o}}]\}^{-1/2}, \\ L_{\mathrm{o}} &= |0_L\rangle\langle 0_{\mathrm{er}}| + |1_L\rangle\langle 1_{\mathrm{er}}|, \end{aligned} \tag{3}$$

where $|u_{\mathrm{er}}\rangle = a|u_L\rangle/\xi_u$ ($u = 0, 1$) denotes the basis of the error space and $\xi_u = \sqrt{\langle u_L|a^\dagger a|u_L\rangle}$ is the normalization coefficient. The QEC jump operator L_{eng} steers the encoded quantum state from the error space back into the code space and can be expanded as

$$L_{\mathrm{eng}} = \sum_{|d_l|\le d} \sum_n \lambda_{nd_l}|n\rangle\langle n + d_l|, \tag{4}$$

where d is the Hamiltonian distance.

We can model the engineered dissipation through a system-environment interaction by introducing a lossy auxiliary qubit and the coupling $H_{\mathrm{eff}} = g(L_{\mathrm{eng}}\sigma_+ + L_{\mathrm{eng}}^\dagger\sigma_-)$, which results in the engineered Lindblad superoperator $\mathcal{D}[L_{\mathrm{eng}}]$ when the qubit is traced out. The evolution of the whole system is then governed by the master equation ($\hbar = 1$),

$$\frac{d\rho}{dt} = -i[H_{\mathrm{eff}},\rho] + \frac{\gamma_a}{2}\mathcal{D}[a] + \frac{\gamma_b}{2}\mathcal{D}[\sigma_-], \tag{5}$$

where γ_a denotes the single-photon loss rate of mode a and γ_b denotes the decay rate of the auxiliary qubit. We assume that the parameters satisfy $\gamma_a, g \ll \gamma_b$ and $\gamma_a \ll g$, which promotes the unidirectional transition from the error space to the code space effected by $L_{\mathrm{eng}}^\dagger$. As shown in Fig. 1(a), the recovery process is then as follows:

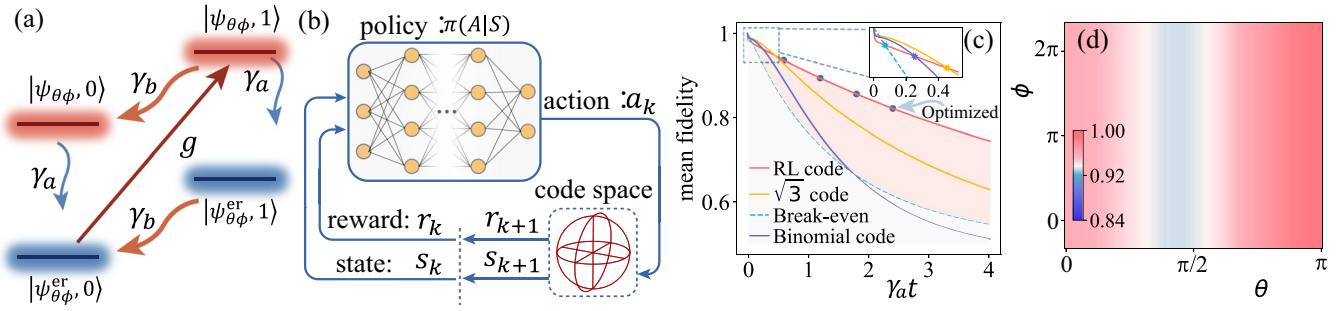

FIG. 1. (a) Energy level diagram illustrating the approximate AQEC process. Induced by a single-photon loss, the encoded state $|\psi_{\theta\phi}\rangle$ undergoes one out of two possible error correction cycles: $|\psi_{\theta\phi},0\rangle \xrightarrow{\gamma_a} |\psi^{\mathrm{er}}_{\theta\phi},0\rangle \xrightarrow{g} |\psi_{\theta\phi},1\rangle \xrightarrow{\gamma_b} |\psi_{\theta\phi},0\rangle$ or $|\psi_{\theta\phi},0\rangle \xrightarrow{\gamma_a} |\psi^{\mathrm{er}}_{\theta\phi},0\rangle \xrightarrow{g} |\psi_{\theta\phi},1\rangle \xrightarrow{\gamma_a} |\psi^{\mathrm{er}}_{\theta\phi},1\rangle \xrightarrow{\gamma_b} |\psi^{\mathrm{er}}_{\theta\phi},0\rangle \xrightarrow{g} |\psi_{\theta\phi},1\rangle \xrightarrow{\gamma_b} |\psi_{\theta\phi},0\rangle$, where the parameters satisfy $\gamma_a, g \ll \gamma_b$ and $\gamma_a \ll g$. (b) Diagram: the cyclic learning process between the RL agent and the encoded system. The policy function $\pi(A|S)$ selects an action based on the current state and reward, and the action acts on the encoded system to generate the new reward r_{k+1} and state s_{k+1}. The states and rewards collected during the sampling process are fed back to the policy network for updating the networks' parameters. (c) Comparison of the AQEC performance of RL code, the lowest-order binomial code, $\sqrt{3}$ code, and break-even. The blue round dots show the optimal mean fidelities obtained from the RL algorithm for different fixed optimization times. The inset highlights an initial transition period where the dynamics is dominated by the single-photon loss. (d) State-dependent fidelity $F(\theta,\phi,t)$ for the RL code versus the angles θ and ϕ at $t = 0.6/\gamma_a$. Other parameters are $\gamma_b/\gamma_a = 1750$ and $g/\gamma_a = 400$.

$|\psi_{\theta\phi},0\rangle \xrightarrow{\gamma_a} |\psi^{\mathrm{er}}_{\theta\phi},0\rangle \xrightarrow{g} |\psi_{\theta\phi},1\rangle \xrightarrow{\gamma_b} |\psi_{\theta\phi},0\rangle$ and $|\psi_{\theta\phi},0\rangle \xrightarrow{\gamma_a} |\psi^{\mathrm{er}}_{\theta\phi},0\rangle \xrightarrow{g} |\psi_{\theta\phi},1\rangle \xrightarrow{\gamma_a} |\psi^{\mathrm{er}}_{\theta\phi},1\rangle \xrightarrow{\gamma_b} |\psi^{\mathrm{er}}_{\theta\phi},0\rangle \xrightarrow{g} |\psi_{\theta\phi},1\rangle \xrightarrow{\gamma_b} |\psi_{\theta\phi},0\rangle$. That is, when an error occurs, the system returns to the logical state $|\psi_{\theta\phi},0\rangle$ by transitioning through one out of two possible error correction cycles.

Optimal code space.—Finding the optimal coefficients $c_n^{(0)}$ and $c_n^{(1)}$, such that the code words Eq. (2) maximize the mean fidelity $\bar{F}(|0_L\rangle,|1_L\rangle)$, at some fixed reference time, represents a complex optimization problem that we solve using reinforcement learning (RL). In brief, each episode is divided into a finite number of steps $k = 1,2,\ldots,K$. As shown in Fig. 1(b), at each step k, the agent observes the current state $s_k \in S$, and chooses an action $a_k \in A$ according to the policy $\pi(A|S)$. The action corresponds to a coefficient vector $[c_n^{(0)}, c_n^{(1)}]$, and the state is described by the fidelity $F(\theta,\phi,t) = \mathrm{Tr}[\rho_{t_0}(\theta,\phi)\rho_t(\theta,\phi)]$, with $\theta \in \{0,\pi/2,\pi\}$ and $\phi \in \{0,\pi/2,\pi,3\pi/2\}$. After each action, the agent receives a reward r_{k+1} and arrives at a new state s_{k+1}, where the reward r_k is set to maximize the difference between the mean fidelity of the logical space and the break-even point. State and reward are obtained by simulating the master equation (5) until $\gamma_a t = 0.6$ with QUTIP [57,58], where the parameters $g/\gamma_a = 400$, $\gamma_b/\gamma_a = 1750$ are chosen to maximize the cooperativity $C = g^2/\gamma_a\gamma_b \approx 91.4$, while remaining experimentally realistic. The states and rewards collected during the sampling process are fed back to the proximal policy optimization algorithm to update the policy $\pi(A|S)$ [59], which is achieved by Ray [60] (see Supplemental Material for the more information [61]).

The optimal code basis found by the RL algorithm, delivering a mean fidelity ≈ 0.95 at $\gamma_a t = 0.6$ and thus significantly exceeding the break-even point at ≈ 0.84, consists of the Fock states: $|0_L\rangle \approx |4\rangle$ and $|1_L\rangle \approx |2\rangle$, i.e., $c_1^{(0)} \approx 1$, $c_0^{(1)} \approx 1$, $c_0^{(0)} \approx c_1^{(1)} \approx 0$. [Note that we truncated the code space at 6 photons. Moreover, we stress that different optimization times $\gamma_a t$ yield consistent outcomes, cf. Fig. 1(c)]. These simple code words can be conveniently prepared in existing experimental setups, e.g., superconducting quantum systems [63–67]. Moreover, the respective QEC jump operator $L_{\mathrm{eng}} \propto |2\rangle\langle 1| + |4\rangle\langle 3|$ has the shortest possible Hamiltonian distance $d = 1$, implying that the QEC Hamiltonian can be efficiently implemented without the need for multiple nonlinear interactions and control fields [15,30]. The Hamiltonian distance of the code words is 2, resulting in single-qubit gates with $d_g = 2$. For instance, the logical Pauli operators are implemented as $\sigma_x = |2\rangle\langle 4| + |4\rangle\langle 2|$, $\sigma_y = i(|4\rangle\langle 2| - |2\rangle\langle 4|)$, and $\sigma_z = |2\rangle\langle 2| - |4\rangle\langle 4|$. For comparison, the lowest-order binomial code requires $d_g = 4$ and the $\sqrt{3}$ code $d_g = 6$. Consequently, single-qubit gates in the RL code space only require second-order nonlinearity, instead of the fourth-order or sixth-order nonlinearities of previous proposals.

In Fig. 1(c), we compare the mean fidelities of the RL code, $\sqrt{3}$ code, and the lowest-order binomial code, where each code is complemented by a single QEC jump operator Eq. (3). We find that the RL code surpasses break-even and both other codes, with the performance advantage growing in time. The mean fidelity of the RL code exceeds the break-even point by $\approx 36\%$ at $\gamma_a t = 4$.

Note that there is a short transition period where the RL code performs poorer than the other codes. This is because initially, when the system lives entirely in the code space, the jump operator Eq. (3) is not yet effective and the

dynamics is entirely driven by the detrimental single-photon loss [68]. The associated decay rate is proportional to the mean photon number, which is highest for the RL code in our comparison ($\bar{n} = 3$ for the RL code, compared to $\bar{n} = 2$ for the binomial code, $\bar{n} = \sqrt{3}$ for the $\sqrt{3}$ code, and $\bar{n} = 0.5$ for the break-even code). Only after this transition period the jump operator Eq. (3) becomes effective and the RL code can unfold its performance advantage over the other codes. Note that the dip at the onset of the transition period reflects the reversible fraction of the fidelity loss and can be removed by using a quantum trajectory-resolved and temporally coarse-grained fidelity $F^*_\tau(t)$ [61,68].

To demonstrate that the AQEC can efficiently protect all quantum states in the RL code space, we evaluate in Fig. 1(d) the fidelity $F(\theta, \phi, t)$ at $t = 0.6/\gamma_a$, where θ and ϕ are the Bloch-sphere angles. We find that the fidelity remains for all states within the range [0.93, 1], well above the break-even point 0.84.

Protection of the logical qubit.—To better understand the level of protection provided by our AQEC scheme, we solve the master equation in the limit $\gamma_a, g \ll \gamma_b$ and $g^2/(\gamma_a\gamma_b) \gg 1$ (a detailed derivation can be found in the Supplemental Material [61]). The approximate solution, after tracing out the auxiliary qubit and expressed in terms of the first five Fock states, reads

$$\rho_a(t) \approx \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \rho_{22}(0) & 0 & \rho_{24}(0)e^{-u\gamma_a t} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \rho_{42}(0)e^{-u\gamma_a t} & 0 & \rho_{44}(0) \end{pmatrix}, \quad (6)$$

with the initial state elements $\rho_{ij}(0)$. The effective protection factor $u = 3 - 2\sqrt{2} \approx 0.17$ captures the reduction of the residual dephasing rate compared to the $|0\rangle$ and $|1\rangle$ encoding. The resulting mean fidelity is upper bounded by

$$\bar{F}(t) = \frac{2}{3} + \frac{1}{3}\exp(-u\gamma_a t), \quad u = 3 - 2\sqrt{2} \approx 0.17, \quad (7)$$

and thus clearly outperforms the mean fidelity of the break-even:

$$\bar{F}_{\mathrm{be}}(t) = \frac{1}{6}\left[\exp(-\gamma_a t) + 2\exp\left(-\frac{\gamma_a t}{2}\right) + 3\right]. \quad (8)$$

We infer from Eq. (7) that, if the cooperativity C is large enough, the performance of AQEC depends only on the evaluation time $\gamma_a t$. If we choose $\gamma_a t = 0.17$ (e.g., $\gamma_a = 0.1$ kHz, $t = 1.7$ ms), the mean fidelity is 99.05%, which significantly surpasses the break-even point (94.68%). The infidelity reduction of the RL code compared to the break-even is $[1 - \bar{F}(t)]/[1 - \bar{F}_{\mathrm{be}}(t)] \approx 0.17$, corresponding to the gain $G = 1/u \approx 5.83$ (near to 6 times the T4C code [15]). Therefore, the approximate AQEC with the RL code significantly surpasses the break-even point with an infidelity reduction of over 80%. We clarify that, similar to other AQEC schemes, leakage out of the error space ultimately limits the protection time $\gamma_a t$, and eventually the system will end up in the ground state.

Coupling engineering.—In contrast to previous AQEC proposals [15,30], where complex code words necessitate multiple nonlinear interactions and complicated control fields, AQEC based on the RL code only requires a Hamiltonian distance $d = 1$, which can be realized with a comparatively simple setting. Figure 2(a) depicts a possible scheme, where the encoding mode a is intermediately coupled to a dissipative mode c through a qubit. The corresponding system Hamiltonian reads

$$H = \omega_a a^\dagger a + \frac{\omega_b}{2}\sigma_z + \omega_c c^\dagger c + f(t)(a + a^\dagger)\sigma_x + g_c(t)(c^\dagger + c)\sigma_x + \frac{\chi}{2}a^\dagger a\sigma_z, \quad (9)$$

and the control fields are

$$f(t) = \frac{2\alpha_0}{\sqrt{2}}\cos\left[\left(\omega_s + \frac{3\chi}{2}\right)t\right] + \frac{2\alpha_0}{\sqrt{4}}\cos\left[\left(\omega_s + \frac{7\chi}{2}\right)t\right],$$
$$g_c(t) = 2\alpha_1\cos(2\chi t) + 2\alpha_1\cos(4\chi t), \quad (10)$$

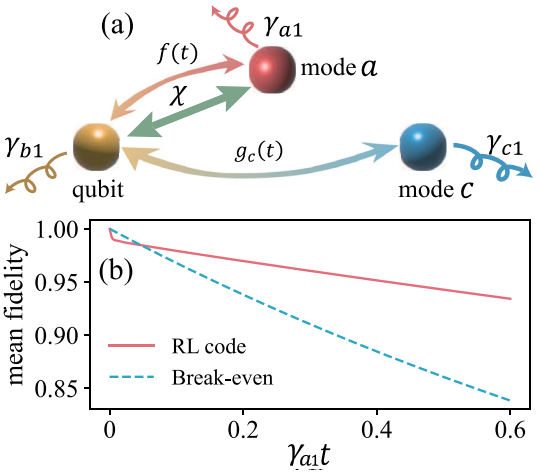

FIG. 2. (a) Proposed system-environment coupling. Blue sideband transitions between a qubit and the encoding mode a are selected by the control field $f(t)$. Red sideband transitions, mediated by the control field $g_c(t)$, transfer the excitation of the qubit to the high-decay mode c, where it dissipates into the environment. (b) Mean fidelity of the RL code versus time for this coupling model under parameters $\omega_a/2\pi = 3.5$ GHz, $\omega_b/2\pi = \omega_c/2\pi = 5$ GHz, $\chi/2\pi = 3$ MHz, $\alpha_0/2\pi = 0.05$ MHz, $\alpha_1/2\pi = 0.07$ MHz, $\gamma_{a1}/2\pi = 0.2$ kHz, $\gamma_{b1}/2\pi = 2$ kHz, and $\gamma_{c1}/2\pi = 0.12$ MHz.

where ω_i $(i = a, b, c)$ are the resonant frequencies of the modes and the qubit $(\omega_b = \omega_c)$, $\omega_s = \omega_a + \omega_b$ is the resonant frequency for the blue sideband transitions, χ is the nonlinear coefficient, and α_j $(j = 0, 1)$ are the strengths of the control fields. (For detailed information about the control fields, see the Supplemental Material [61]). We can recover the master equation (5) by adiabatically eliminating the high-decay mode c under the conditions $\omega_a, \omega_b, \omega_c \gg \chi$, and $\gamma_{c1} \gg \alpha_1 \geq \alpha_0 \gg \gamma_{a1}, \gamma_{b1}$, where γ_{i1} $(i = a, b, c)$ are the respective decay rates. The controllable coupling between qubits and bosonic modes has been thoroughly investigated both in theory and in experiment [69–76]. The nonlinear term $a^\dagger a\sigma_z$ can be effectively obtained by a linear coupling between mode and qubit, $\beta(a^\dagger + a)\sigma_x$, with a large detuning $\Delta = |\omega_a - \omega| \gg \beta$, where the dispersive coefficient $\chi = g^2/\Delta$ can be larger than $7 \times 2\pi$ MHz [77–85]. It is thus possible to realize the scheme with linearly coupled systems. A low single-photon loss of $0.118 \times 2\pi$ kHz has been achieved in experiments with 3D coaxial cavities, and the decay of the transmon qubits can be suppressed to $1.8 \times 2\pi$kHz [86–88]. We simulate the mean fidelity of the full quantum system in Fig. 2(b) with typical experimental parameters in superconducting circuits, $\omega_a/2\pi = 3.5$ GHz, $\omega_b/2\pi = \omega_{c1}/2\pi =$ 5 GHz, $\gamma_{c1}/2\pi = 0.12$ MHz, $\chi/2\pi = 3$ MHz, $\alpha_0/2\pi =$ 0.05 MHz, $\alpha_1/2\pi = 0.07$ MHz, $\gamma_{a1}/2\pi = 0.2$ kHz, and $\gamma_b/2\pi = 2$ kHz. We find that the mean fidelity of the RL code by far surpasses break-even threshold, and the model may be realized by engineering the coupling between a 3D coaxial cavity, a transmon qubit, and a dissipative cavity.

Discussion and conclusion.—We propose a bosonic code space for approximate AQEC with the shortest possible Hamiltonian distance $d = 1$, allowing for significantly reduced model complexity. Nevertheless, our scheme comfortably surpasses break-even threshold, outperforming other AQEC schemes that rely on larger Hamiltonian distances and disproving a previous claim that surpassing break-even threshold requires $d \geq 2$. The code words consist of Fock states rather than complex superposition states, and single-qubit logic gates in the RL code space are implemented with a maximum distance of $d_g = 2$. The mean fidelity of the RL code can exceed break-even point by more than 36% at $\gamma_a t = 4$, and the expected gain $G = 1/(3 - 2\sqrt{2}) \approx 5.83$ is more than twice the currently best experimental value $G = 2.27 \pm 0.07$ [89]. This demonstrates that, despite our AQEC scheme being approximate, which upper bounds the level of protection, it delivers logical qubits with significantly improved quality and thus may greatly facilitate subsequent QEC steps toward fully fault-tolerant qubits.

We would like to acknowledge the help of Professor Jie-Qiao Liao and the valuable suggestions of Dr. Wei Qin, Dr. Ye-Hong Chen, Dr. Fabrizio Minganti, and Dr. Ran Huang. F. N. is supported in part by Nippon Telegraph and Telephone Corporation (NTT) Research, the Japan Science and Technology Agency (JST) [via the Quantum Leap Flagship Program (Q-LEAP) and the Moonshot R&D Grant No. JPMJMS2061], the Asian Office of Aerospace Research and Development (AOARD) (via Grant No. FA2386-20-1-4069), and the Foundational Questions Institute Fund (FQXi) via Grant No. FQXi-IAF19-06.

*clemens.gneiting@riken.jp
†fnori@riken.jp

[1] J. Chiaverini, D. Leibfried, T. Schaetz, M. D. Barrett, R. B. Blakestad, J. Britton, W. M. Itano, J. D. Jost, E. Knill, C. Langer, R. Ozeri, and D. J. Wineland, Realization of quantum error correction, Nature (London) **432**, 602 (2004).
[2] P. Schindler, J. T. Barreiro, T. Monz, V. Nebendahl, D. Nigg, M. Chwalla, M. Hennrich, and R. Blatt, Experimental repetitive quantum error correction, Science **332**, 1059 (2011).
[3] B. M. Terhal, Quantum error correction for quantum memories, Rev. Mod. Phys. **87**, 307 (2015).
[4] F. Gaitan, *Quantum Error Correction and Fault Tolerant Quantum Computing* (Taylor & Francis, Andover, England, 2017).
[5] D. A. Lidar and T. A. Brun, *Quantum Error Correction* (Cambridge University Press, Cambridge, England, 2013).
[6] C. Ryan-Anderson, J. G. Bohnet, K. Lee, D. Gresh, A. Hankin, J. P. Gaebler, D. Francois, A. Chernoguzov, D. Lucchetti, N. C. Brown, T. M. Gatterman, S. K. Halit, K. Gilmore, J. A. Gerber, B. Neyenhuis, D. Hayes, and R. P. Stutz, Realization of Real-Time Fault-Tolerant Quantum Error Correction, Phys. Rev. X **11**, 041058 (2021).
[7] M. Sarovar and G. J. Milburn, Continuous quantum error correction by cooling, Phys. Rev. A **72**, 012306 (2005).
[8] J. Cohen and M. Mirrahimi, Dissipation-induced continuous quantum error correction for superconducting circuits, Phys. Rev. A **90**, 062344 (2014).
[9] V. V. Albert, S. O. Mundhada, A. Grimm, S. Touzard, M. H. Devoret, and L. Jiang, Pair-cat codes: autonomous error-correction with low-order nonlinearity, Quantum Sci. Technol. **4**, 035007 (2019).
[10] Y. Ma, Y. Xu, X. Mu, W. Cai, L. Hu, W. Wang, X. Pan, H. Wang, Y. P. Song, C.-L. Zou, and L. Sun, Error-transparent operations on a logical qubit protected by quantum error correction, Nat. Phys. **16**, 827 (2020).
[11] K. C. Young, M. Sarovar, and R. Blume-Kohout, Error Suppression and Error Correction in Adiabatic Quantum Computation: Techniques and Challenges, Phys. Rev. X **3**, 041013 (2013).
[12] J. Atalaya, S. Zhang, M. Y. Niu, A. Babakhani, H. C. H. Chan, J. M. Epstein, and K. B. Whaley, Continuous quantum error correction for evolution under time-dependent Hamiltonians, Phys. Rev. A **103**, 042406 (2021).
[13] Y.-H. Chen, R. Stassi, W. Qin, A. Miranowicz, and F. Nori, Fault-Tolerant Multiqubit Geometric Entangling Gates Using Photonic Cat-State Qubits, Phys. Rev. Appl. **18**, 024076 (2022).
[14] Y.-H. Chen, W. Qin, R. Stassi, X. Wang, and F. Nori, Fast binomial-code holonomic quantum computation with

ultrastrong light-matter coupling, Phys. Rev. Res. **3**, 033275 (2021).
[15] J. M. Gertler, B. Baker, J. Li, S. Shirol, J. Koch, and C. Wang, Protecting a bosonic qubit with autonomous quantum error correction, Nature (London) **590**, 243 (2021).
[16] J. P. Barnes and W. S. Warren, Automatic Quantum Error Correction, Phys. Rev. Lett. **85**, 856 (2000).
[17] D. R. Pérez and E. Kapit, Improved autonomous error correction using variable dissipation in small logical qubit architectures, Quantum Sci. Technol. **6**, 015006 (2020).
[18] 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).
[19] A. M. Steane, Error Correcting Codes in Quantum Theory, Phys. Rev. Lett. **77**, 793 (1996).
[20] E. Kapit, Hardware-Efficient and Fully Autonomous Quantum Error Correction in Superconducting Circuits, Phys. Rev. Lett. **116**, 150501 (2016).
[21] J. Kerckhoff, H. I. Nurdin, D. S. Pavlichin, and H. Mabuchi, Designing Quantum Memories with Embedded Control: Photonic Circuits for Autonomous Quantum Error Correction, Phys. Rev. Lett. **105**, 040502 (2010).
[22] F. Reiter, A. S. Sørensen, P. Zoller, and C. A. Muschik, Dissipative quantum error correction and application to quantum sensing with trapped ions, Nat. Commun. **8**, 1822 (2017).
[23] S. Krastanov, M. Heuck, J. H. Shapiro, P. Narang, D. R. Englund, and K. Jacobs, Room-temperature photonic logical qubits via second-order nonlinearities, Nat. Commun. **12**, 191 (2021).
[24] W. Cai, Y. Ma, W. Wang, C.-L. Zou, and L. Sun, Bosonic quantum error correction codes in superconducting quantum circuits, Fundam. Res. **1**, 50 (2021).
[25] Q. Xu, G. Zheng, Y.-X. Wang, P. Zoller, A. A. Clerk, and L. Jiang, Autonomous quantum error correction and fault-tolerant quantum computation with squeezed cat qubits, arXiv:2210.13406.
[26] T. Hillmann and F. Quijandría, Quantum error correction with dissipatively stabilized squeezed cat qubits, Phys. Rev. A **107**, 032423 (2023).
[27] P. Zanardi, J. Marshall, and L. Campos Venuti, Dissipative universal Lindbladian simulation, Phys. Rev. A **93**, 022312 (2016).
[28] J.-M. Lihm, K. Noh, and U. R. Fischer, Implementation-independent sufficient condition of the Knill-Laflamme type for the autonomous protection of logical qudits by strong engineered dissipation, Phys. Rev. A **98**, 012317 (2018).
[29] L. Hu, Y. Ma, W. Cai, X. Mu, Y. Xu, W. Wang, Y. Wu, H. Wang, Y. P. Song, C.-L. Zou, S. M. Girvin, L.-M. Duan, and L. Sun, Quantum error correction and universal gate set operation on a binomial bosonic logical qubit, Nat. Phys. **15**, 503 (2019).
[30] Z. Wang, T. Rajabzadeh, N. Lee, and A. H. Safavi-Naeini, Automated discovery of autonomous quantum error correction schemes, PRX Quantum **3**, 020302 (2022).
[31] E. Knill and R. Laflamme, Theory of quantum error-correcting codes, Phys. Rev. A **55**, 900 (1997).
[32] E. Knill, R. Laflamme, and L. Viola, Theory of Quantum Error Correction for General Noise, Phys. Rev. Lett. **84**, 2525 (2000).
[33] 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).
[34] C. Bény and O. Oreshkov, General Conditions for Approximate Quantum Error Correction and Near-Optimal Recovery Channels, Phys. Rev. Lett. **104**, 120501 (2010).
[35] C. Bény, Perturbative Quantum Error Correction, Phys. Rev. Lett. **107**, 080501 (2011).
[36] L. Kong and Z.-W. Liu, Near-optimal covariant quantum error-correcting codes from random unitaries with symmetries, PRX Quantum **3**, 020314 (2022).
[37] P. Faist, S. Nezami, V. V. Albert, G. Salton, F. Pastawski, P. Hayden, and J. Preskill, Continuous Symmetries and Approximate Quantum Error Correction, Phys. Rev. X **10**, 041018 (2020).
[38] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, 2nd ed. (The MIT Press, Cambridge, MA, 2018).
[39] M. Bukov, A. G. R. Day, D. Sels, P. Weinberg, A. Polkovnikov, and P. Mehta, Reinforcement Learning in Different Phases of Quantum Control, Phys. Rev. X **8**, 031086 (2018).
[40] T. Fösel, P. Tighineanu, T. Weiss, and F. Marquardt, Reinforcement Learning with Neural Networks for Quantum Feedback, Phys. Rev. X **8**, 031084 (2018).
[41] Y. Baum, M. Amico, S. Howell, M. Hush, M. Liuzzi, P. Mundada, T. Merkh, A. R. R. Carvalho, and M. J. Biercuk, Experimental deep reinforcement learning for error-robust gate-set design on a superconducting quantum computer, PRX Quantum **2**, 040324 (2021).
[42] S.-F. Guo, F. Chen, Q. Liu, M. Xue, J.-J. Chen, J.-H. Cao, T.-W. Mao, M. K. Tey, and L. You, Faster State Preparation across Quantum Phase Transition Assisted by Reinforcement Learning, Phys. Rev. Lett. **126**, 060401 (2021).
[43] S. Borah, B. Sarma, M. Kewming, G. J. Milburn, and J. Twamley, Measurement-Based Feedback Quantum Control with Deep Reinforcement Learning for a Double-Well Nonlinear Potential, Phys. Rev. Lett. **127**, 190403 (2021).
[44] K. Bartkiewicz, C. Gneiting, A. Černoch, K. Jiráková, K. Lemr, and F. Nori, Experimental kernel-based quantum machine learning in finite feature space, Sci. Rep. **10**, 12356 (2020).
[45] A. Bolens and M. Heyl, Reinforcement Learning for Digital Quantum Simulation, Phys. Rev. Lett. **127**, 110502 (2021).
[46] J. Yao, L. Lin, and M. Bukov, Reinforcement Learning for Many-Body Ground-State Preparation Inspired by Counterdiabatic Driving, Phys. Rev. X **11**, 031070 (2021).
[47] T. Xiao, J. Fan, and G. Zeng, Parameter estimation in quantum sensing based on deep reinforcement learning, npj Quantum Inf. **8**, 2 (2022).
[48] P. Peng, X. Huang, C. Yin, L. Joseph, C. Ramanathan, and P. Cappellaro, Deep Reinforcement Learning for Quantum Hamiltonian Engineering, Phys. Rev. Appl. **18**, 024033 (2022).

[49] J. Lebreuilly, K. Noh, C.-H. Wang, S. M. Girvin, and L. Jiang, Autonomous quantum error correction and quantum computation, arXiv:2103.05007.
[50] R. Lescanne, M. Villiers, T. Peronnin, A. Sarlette, M. Delbecq, B. Huard, T. Kontos, M. Mirrahimi, and Z. Leghtas, Exponential suppression of bit-flips in a qubit encoded in an oscillator, Nat. Phys. **16**, 509 (2020).
[51] A. Blais, A. L. Grimsmo, S. M. Girvin, and A. Wallraff, Circuit quantum electrodynamics, Rev. Mod. Phys. **93**, 025005 (2021).
[52] X. Gu, A. F. Kockum, A. Miranowicz, Y.-x. Liu, and F. Nori, Microwave photonics with superconducting quantum circuits, Phys. Rep. **718–719**, 1 (2017).
[53] D. A. Herrera-Martí, T. Gefen, D. Aharonov, N. Katz, and A. Retzker, Quantum Error-Correction-Enhanced Magnetometer Overcoming the Limit Imposed by Relaxation, Phys. Rev. Lett. **115**, 200501 (2015).
[54] S. Kwon, S. Watabe, and J.-S. Tsai, Autonomous quantum error correction in a four-photon Kerr parametric oscillator, npj Quantum Inf. **8**, 40 (2022).
[55] F. G. S. L. Brandão, E. Crosson, M. B. Şahinoğlu, and J. Bowen, Quantum Error Correcting Codes in Eigenstates of Translation-Invariant Spin Chains, Phys. Rev. Lett. **123**, 110502 (2019).
[56] I. Tzitrin, J. E. Bourassa, N. C. Menicucci, and K. K. Sabapathy, Progress towards practical qubit computation using approximate Gottesman-Kitaev-Preskill codes, Phys. Rev. A **101**, 032315 (2020).
[57] J. R. Johansson, P. D. Nation, and F. Nori, QUTIP: An open-source Python framework for the dynamics of open quantum systems, Comput. Phys. Commun. **183**, 1760 (2012).
[58] J. R. Johansson, P. D. Nation, and F. Nori, QUTIP 2: A Python framework for the dynamics of open quantum systems, Comput. Phys. Commun. **184**, 1234 (2013).
[59] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, Proximal policy optimization algorithms, arXiv:1707.06347.
[60] E. Liang *et al.*, RLlib:Industry-Grade Reinforcement Learning, 2018, https://github.com/ray-project/ray/tree/master/rllib.
[61] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.131.050601 for more details, which includes Ref. [62].
[62] Y. C. Liu, X. Luan, H. K. Li, Q. Gong, C. W. Wong, and Y. F. Xiao, Coherent Polariton Dynamics in Coupled Highly Dissipative Cavities, Phys. Rev. Lett. **112**, 213602 (2014).
[63] Y. xi Liu, L. F. Wei, and F. Nori, Generation of nonclassical photon states using a superconducting qubit in a microcavity, Europhys. Lett. **67**, 941 (2004).
[64] M. Hofheinz, H. Wang, M. Ansmann, R. C. Bialczak, E. Lucero, M. Neeley, A. D. O'Connell, D. Sank, J. Wenner, J. M. Martinis, and A. N. Cleland, Synthesizing arbitrary quantum states in a superconducting resonator, Nature (London) **459**, 546 (2009).
[65] M. Hofheinz, E. M. Weig, M. Ansmann, R. C. Bialczak, E. Lucero, M. Neeley, A. D. O'Connell, H. Wang, J. M. Martinis, and A. N. Cleland, Generation of Fock states in a superconducting quantum circuit, Nature (London) **454**, 310 (2008).
[66] Note that the imbalance between the code word error probabilities of the RL code can in principle be removed by employing a two-mode qubit structure [67], however at the cost of significantly increased model complexity.
[67] H. Mabuchi and P. Zoller, Inversion of Quantum Jumps in Quantum Optical Systems under Continuous Observation, Phys. Rev. Lett. **76**, 3108 (1996).
[68] G. Sarma and H. Mabuchi, Gauge subsystems, separability and robustness in autonomous quantum memories, New J. Phys. **15**, 035014 (2013).
[69] A. Wallraff, D. I. Schuster, A. Blais, J. M. Gambetta, J. Schreier, L. Frunzio, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Sideband Transitions and Two-Tone Spectroscopy of a Superconducting Qubit Strongly Coupled to an On-Chip Cavity, Phys. Rev. Lett. **99**, 050501 (2007).
[70] Y. Lu, S. Chakram, N. Leung, N. Earnest, R. K. Naik, Z. Huang, P. Groszkowski, E. Kapit, J. Koch, and D. I. Schuster, Universal Stabilization of a Parametrically Coupled Qubit, Phys. Rev. Lett. **119**, 150502 (2017).
[71] P. Roushan *et al.*, Chiral ground-state currents of interacting photons in a synthetic magnetic field, Nat. Phys. **13**, 146 (2017).
[72] L. Bresque, P. A. Camati, S. Rogers, K. Murch, A. N. Jordan, and A. Auffèves, Two-Qubit Engine Fueled by Entanglement and Local Measurements, Phys. Rev. Lett. **126**, 120605 (2021).
[73] P. Campagne-Ibarcq, E. Zalys-Geller, A. Narla, S. Shankar, P. Reinhold, L. Burkhart, C. Axline, W. Pfaff, L. Frunzio, R. J. Schoelkopf, and M. H. Devoret, Deterministic Remote Entanglement of Superconducting Circuits through Microwave Two-Photon Transitions, Phys. Rev. Lett. **120**, 200501 (2018).
[74] T. Niemczyk, F. Deppe, H. Huebl, E. P. Menzel, F. Hocke, M. J. Schwarz, J. J. Garcia-Ripoll, D. Zueco, T. Hümmer, E. Solano, A. Marx, and R. Gross, Circuit quantum electrodynamics in the ultrastrong-coupling regime, Nat. Phys. **6**, 772 (2010).
[75] P. J. Leek, S. Filipp, P. Maurer, M. Baur, R. Bianchetti, J. M. Fink, M. Göppl, L. Steffen, and A. Wallraff, Using sideband transitions for two-qubit operations in superconducting circuits, Phys. Rev. B **79**, 180511(R) (2009).
[76] Y.-x. Liu, L. F. Wei, J. R. Johansson, J. S. Tsai, and F. Nori, Superconducting qubits can be coupled and addressed as trapped ions, Phys. Rev. B **76**, 144518 (2007).
[77] D. I. Schuster, A. A. Houck, J. A. Schreier, A. Wallraff, J. M. Gambetta, A. Blais, L. Frunzio, J. Majer, B. Johnson, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Resolving photon number states in a superconducting circuit, Nature (London) **445**, 515 (2007).
[78] J. Majer, J. M. Chow, J. M. Gambetta, J. Koch, B. R. Johnson, J. A. Schreier, L. Frunzio, D. I. Schuster, A. A. Houck, A. Wallraff, A. Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Coupling superconducting qubits via a cavity bus, Nature (London) **449**, 443 (2007).
[79] J. Q. You and F. Nori, Quantum information processing with superconducting qubits in a microwave field, Phys. Rev. B **68**, 064509 (2003).
[80] M.-L. Cai, Z.-D. Liu, W.-D. Zhao, Y.-K. Wu, Q.-X. Mei, Y. Jiang, L. He, X. Zhang, Z.-C. Zhou, and L.-M. Duan, Observation of a quantum phase transition in the quantum Rabi model with a single trapped ion, Nat. Commun. **12**, 1126 (2021).

[81] Z.-L. Xiang, S. Ashhab, J. Q. You, and F. Nori, Hybrid quantum circuits: Superconducting circuits interacting with other quantum systems, Rev. Mod. Phys. **85**, 623 (2013).
[82] S. Ashhab and F. Nori, Qubit-oscillator systems in the ultrastrong-coupling regime and their potential for preparing nonclassical states, Phys. Rev. A **81**, 042311 (2010).
[83] D. Zueco, G. M. Reuther, S. Kohler, and P. Hänggi, Qubit-oscillator dynamics in the dispersive regime: Analytical theory beyond the rotating-wave approximation, Phys. Rev. A **80**, 033846 (2009).
[84] M.-J. Hwang, R. Puebla, and M. B. Plenio, Quantum Phase Transition and Universal Dynamics in the Rabi Model, Phys. Rev. Lett. **115**, 180404 (2015).
[85] T. Brown, E. Doucet, D. Ristè, G. Ribeill, K. Cicak, J. Aumentado, R. Simmonds, L. Govia, A. Kamal, and L. Ranzani, Trade off-free entanglement stabilization in a superconducting qutrit-qubit system, Nat. Commun. **13**, 1 (2022).
[86] 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).
[87] S. Rosenblum, P. Reinhold, M. Mirrahimi, L. Jiang, L. Frunzio, and R. J. Schoelkopf, Fault-tolerant detection of a quantum error, Science **361**, 266 (2018).
[88] C. Wang, Y. Y. Gao, P. Reinhold, R. W. Heeres, N. Ofek, K. Chou, C. Axline, M. Reagor, J. Blumoff, K. M. Sliwa, L. Frunzio, S. M. Girvin, L. Jiang, M. Mirrahimi, M. H. Devoret, and R. J. Schoelkopf, A Schrödinger cat living in two boxes, Science **352**, 1087 (2016).
[89] V. V. Sivak, A. Eickbusch, B. Royer, S. Singh, I. Tsioutsios, S. Ganjam, A. Miano, B. L. Brock, A. Z. Ding, L. Frunzio, S. M. Girvin, R. J. Schoelkopf, and M. H. Devoret, Real-time quantum error correction beyond break-even, Nature (London) **616**, 50 (2023).