Engineering fast bias-preserving gates on stabilized cat qubits

Qian Xu,[1] Joseph K. Iverson,[2] Fernando G. S. L. Brandão,[2,3] and Liang Jiang[1,2,*]
[1]*Pritzker School of Molecular Engineering, The University of Chicago, Chicago, Illinois 60637, USA*
[2]*AWS Center for Quantum Computing, Pasadena, California 91125, USA*
[3]*IQIM, California Institute of Technology, Pasadena, California 91125, USA*

(Received 18 May 2021; revised 29 November 2021; accepted 5 January 2022; published 2 February 2022)

Stabilized cat codes can provide a biased noise channel with a set of bias-preserving (BP) gates, which can significantly reduce the resource overhead for fault-tolerant quantum computing. All existing schemes of BP gates, however, require adiabatic quantum evolution, with performance limited by excitation loss and nonadiabatic errors during the adiabatic gates. In this paper, we apply a derivative-based leakage-suppression technique to overcome nonadiabatic errors, so that we can implement fast BP gates on Kerr-cat qubits with improved gate fidelity while maintaining high noise bias. When applied to concatenated quantum error correction, the fast BP gates not only can improve the logical error rate but also can reduce resource overhead, which enables more efficient implementation of fault-tolerant quantum computing.

DOI: 10.1103/PhysRevResearch.4.013082

I. INTRODUCTION

Quantum error correction (QEC) of generic errors is very challenging, because of the demanding threshold requirements and significant resource overhead. To overcome this challenge, we may adaptively design the QEC codes targeting practically relevant errors in a hardware-efficient way. For example, we can develop various efficient bosonic QEC codes to correct excitation loss errors [1–4], which have been experimentally demonstrated using superconducting circuits [5–9] and trapped ions [10].

With Hamiltonian protection or reservoir engineering, some bosonic codes can continuously suppress practically relevant errors (e.g., excitation loss) and also provide a highly biased noise channel. For example, stabilized cat qubits can exponentially suppress bit-flip errors, because of the large separation of coherent states in the phase space [7,8,11]. Such an encoding with highly biased noise channel can play a unique role in fault-tolerant architecture [12,13], as the higher-level QEC codes can be tailored toward the biased noise to exhibit significantly improved error threshold and resource overhead. To get the maximum benefit from biased noise, it is essential for all gate operations to preserve the noise bias. Recently, a nontrivial set of bias-preserving (BP) gates have been discovered for stabilized cat qubits [8,14–16], which opens up a new direction of fault-tolerant architectural design [17–19].

All existing schemes of BP gates, however, require adiabatic quantum evolution due to the nonadiabatic errors, with performance limited by excitation loss during the slow gates. In this paper, we propose fast BP gates on stabilized cat qubits with unitary control to coherently suppress nonadiabatic errors. We systematically compare the performance of fast BP gates with existing adiabatic BP gates [15,16] and show that the fast BP gates can achieve improved gate fidelity while maintaining high noise bias. We show that by integrating the quantum control and quantum error correction, we not only can improve the logical error rate but also can reduce the resource overhead in the concatenated-code level.

The paper is structured as follows: In Sec. II, we give an introduction to the Kerr cat and the dissipative cat and to a set of bias-preserving gates on them. We discuss the error structure and limitations of the BP gates and compare the Kerr and dissipative BP gates in terms of the different mechanisms of their nonadiabatic errors. In Sec. III, we introduce the systematically designed derivative-based corrections to the Kerr gate Hamiltonians in Ref. [15] and show that the introduced corrections significantly suppress the nonadiabatic bit- and phase-flip errors at the same time. As a result, the fidelity and the noise bias of the gates in the presence of excitation loss can be simultaneously improved. These improvements enable the Kerr gates to outperform the dissipative gates under excitation loss. In Sec. IV, we apply the improved gates to a concatenated QEC scheme, where the stabilized cats are concatenated with a repetition code. Based on a circuit-level noise model, we show that the use of the Kerr gates with our improved control leads to lower logical gate failure rates and/or less resource overhead. In Secs. V and VI, we discuss the experimental realizations and draw conclusions.

II. BIAS-PRESERVING GATES FOR KERR AND DISSIPATIVE CAT

The cat qubit spanned by coherent states $|\alpha\rangle$ and $|-\alpha\rangle$ can be stabilized in a Kerr oscillator with parametric two-photon

*liang.jiang@uchicago.edu

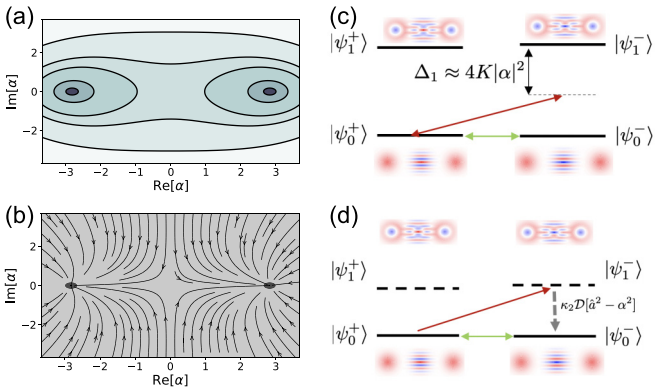

FIG. 1. (a) The semiclassical potential $\langle \hat{H}_{\rm KPO}\rangle$ of a Kerr cat in phase space. (b) The semiclassical dynamics of a dissipative cat in phase space. (c) and (d) Mechanism of the nonadiabatic errors during the z-axis rotation on (c) Kerr cats and (d) dissipative cats. $|\psi_0^{\pm}\rangle$, $|\psi_1^{\pm}\rangle$ denote the first two pairs of eigenstates of $\hat{H}_{\rm KPO}$ with parity $\pm$, and the spacing between states in the vertical axis in (c) and (d) represents the energy gap and dissipative gap of the Kerr cat and the dissipative cat, respectively. For Kerr cats in (c), the nonadiabatic errors induced by the linear drive $[\Omega(t)\hat{a} + \text{H.c.}]$ manifest in the form of off-resonant leakage. For dissipative cats in (d) with a dissipation gap, the linear drive induces decoherence within the cat-state manifold when combined with the engineered two-photon dissipation.

drive [8,14,15,20]. The Hamiltonian of such a Kerr parametric oscillator (KPO) in the frame rotating at the oscillator frequency is

$$\hat{H}_{\rm KPO} = -K(\hat{a}^{2\dagger} - \alpha^2)(\hat{a}^2 - \alpha^2), \tag{1}$$

where K is the strength of Kerr nonlinearity. We may intuitively view the KPO system as a "double-well" potential with two extrema, α and $-\alpha$, in phase space, as shown in Fig. 1(a). In addition to the degenerate ground states $|\pm\alpha\rangle$, $\hat{H}_{\rm KPO}$ supports nearly degenerate pairs of excited states with eigenenergies $\Delta_n \pm \delta_n/2$ for $n = 1, 2, \ldots$, where $\pm$ labels the photon-number parity and δ_n denotes the energy splitting between the members of the nth pair, which is exponentially suppressed by α^2 for $n < \alpha^2/4$ [15]. We denote the subspace spanned by the nth pair of (quasidegenerate) excited states as the nth excited subspace. The excitation gap $\Delta_1 \approx -4K|\alpha|^2$ [21] provides continuous protection of the encoded quantum information.

Alternatively, the cat qubit can also be stabilized by engineered two-photon dissipation and two-photon drive [11,16,17]:

$$\frac{d\rho}{dt} = \kappa_2\mathcal{D}[\hat{a}^2 - \alpha^2]\rho, \tag{2}$$

where κ_2 is the two-photon dissipation rate and $\mathcal{D}[\hat{A}]\hat{\rho} = \hat{A}\hat{\rho}\hat{A}^\dagger - \frac{1}{2}\{\hat{A}^\dagger\hat{A}, \hat{\rho}\}$. We may intuitively understand the stabilization using a semiclassical flow diagram with two stable steady states $|\pm\alpha\rangle$ as illustrated in Fig. 1(b). For quantum evolution, the cat code space is stabilized as an attractive steady-state subspace protected by a dissipative gap.

By choosing the computational basis of the stabilized cat qubit as $|0\rangle_L \equiv \frac{|\psi_0^+\rangle + |\psi_0^-\rangle}{\sqrt{2}} \approx |\alpha\rangle$, $|1\rangle_L \equiv \frac{|\psi_0^+\rangle - |\psi_0^-\rangle}{\sqrt{2}} \approx |-\alpha\rangle$, the noise channel is strongly biased toward phase-flip error $\hat{Z}$, with the noise bias $\eta \equiv P_z/P_x$ increasing exponentially with $|\alpha|^2$ [7,8,11]. Here, and in the rest of this paper, we use P_x to denote all the nondephasing-type errors for simplicity.

A set of bias-preserving gates have been proposed separately for the Kerr cat [15] and dissipative cat [16], which we refer to as Kerr gates and dissipative gates, respectively. Under excitation loss, the total Z and X error probabilities of the BP gates can be written as the following [15,16,18]:

$$\begin{aligned} P_z &= P_z^{\rm NA}(T) + \beta\kappa_1|\alpha|^2 T, \\ P_x &= P_x^{\rm NA}(T) + \gamma\kappa_1|\alpha|^2 T, \end{aligned} \tag{3}$$

where κ_1 denotes the photon loss rate and β is a constant depending on the gate [$\beta = 1$ for Z rotation and $\beta = 2$ for ZZ rotation and the controlled-X (CX) gate]. The first terms, $P_z^{\rm NA}$ and $P_x^{\rm NA}$, are the nonadiabatic errors due to finite gate time and prefer slow gates, while the second terms are photon-loss-induced errors that increase with gate duration. The loss-induced bit flip is negligible since the induced bit-flip rate is suppressed by a factor $\gamma = \gamma[|\alpha|^2, \kappa_1/K(\kappa_2)] \ll 1$, which decreases exponentially for large $|\alpha|^2$ [7,16], and thereby P_x is dominantly given by $P_x^{\rm NA}$. In this paper, we choose the size of the cat as $\alpha^2 = 8$, which is experimentally feasible [8] and also provides us sufficient noise bias [15,18]. Due to the large loss-induced phase-flip rate, fast gates are desirable to obtain high gate fidelity. However, the exponential noise bias might break down in the fast-gate regime due to $P_x^{\rm NA}$. Therefore it is important to design fast BP gates with simultaneously suppressed $P_z^{\rm NA}$ and $P_x^{\rm NA}$.

First, we would like to point out a fundamental difference in the nonadiabatic errors between Kerr gates and dissipative gates. For dissipative gates, the nonadiabatic errors are associated with accumulated dissipation with continuous information leakage into the environment, which is difficult to restore. For example, as shown in Fig. 1(d), the linear drive $\Omega(t)a + \Omega^*(t)a^\dagger$, which implements the Z rotation, can create leakage outside the cat subspace, which becomes continuous phase flips [18] when brought back by the engineered two-photon dissipation. In contrast, for Kerr gates shown in Fig. 1(c), the nonadiabatic errors are coherent off-resonant leakage errors to excited states, which might be reliably eliminated using additional leakage-suppression techniques. Hence the Kerr gates can in principle benefit from quantum coherent controls, while the dissipative gates cannot.

As shown in Fig. 2, the Kerr gates with hard square pulses (blue curves, denoted as "hard") [15] have smaller $P_z^{\rm NA}$ but larger $P_x^{\rm NA}$ than the corresponding dissipative gates (black dashed curves, denoted as "dissipative") [16]. The larger $P_x^{\rm NA}$ error is due to the large leakage induced by the hard pulse, which leads to coherent tunneling between two wells of the KPO through high excited levels. These numerical results manifest the importance of finer control on the Kerr gates to further suppress the leakage, which motivates our work. Although the numerical quantum optimal control method [22] can be applied to optimizing the gates, analytical solutions that can produce smooth and robust pulses and avoid large-scale numerical optimizations for large cats are desirable. We use the derivative-based transition-suppression technique [23], which is a variant of the derivative removal by adiabatic gate (DRAG) technique and closely related to the idea of

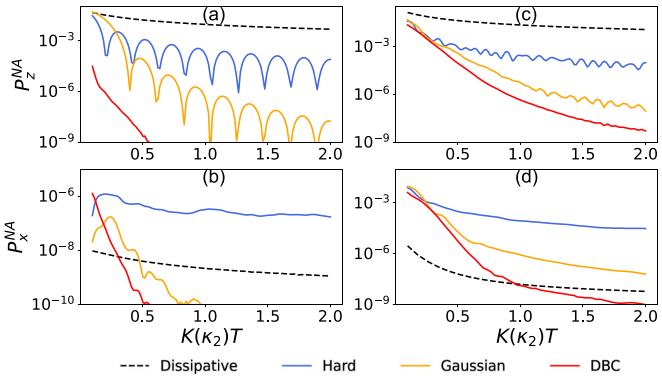

FIG. 2. The nonadiabatic Z (P_z^{NA}) and X (P_x^{NA}) errors of the Z rotation [(a) and (b)] and the CX gate [(c) and (d)] for dissipative gates (black dashed curves), Kerr gates with hard square pulses (blue curves), Kerr gates with truncated Gaussian pulses (orange curves), and Kerr gates with DBC control (red curves). The gate time is in units of K (κ_2) for Kerr (dissipative) gates. The Kerr gates with DBC control can best suppress the nonadiabatic errors.

shortcut to adiabaticity (STA) [23–27], to suppress the leakage of the BP Kerr gates so that they can have simultaneous high gate fidelity and noise bias in the presence of photon loss. The key idea behind this technique is to use quantum interference to suppress the occupation of the excited states by adding corrections to the original Hamiltonian according to the derivatives of the original driving pulse. Please refer to Ref. [23] for details of the technique. We also provide a brief review of the technique in Appendixes B and C. In the rest of the main text, we only present the results derived from the derivative-based technique and refer the reader to Appendixes B and C for more technical details of the derivation.

III. DERIVATIVE-BASED CONTROL OF BP KERR GATES

To simultaneously suppress the nonadiabatic bit- and phase-flip errors of the BP Kerr gates, it is important to identify the associated excited subspace that dominantly contributes to each type of error, respectively. The phase-flip errors come from parity flips, and the occupation of any excited subspace with inconsistent parity compared with the ground subspace leads to phase-flip errors. Therefore, if the drive invoking the diabatic transitions changes the parity, such as the linear drive that is commonly used by the gates that we consider, the dominant error space is the excited subspace with the largest occupation, which is typically the lowest-lying excited subspace with the smallest energy gap. The bit-flip errors, however, come from the tunneling between the two potential wells of $\hat{H}_{\mathrm{KPO}}$ [see Fig. 1(a)]. As a result, the excitation to higher excited subspaces in which the eigenstates are more delocalized brings more bit-flip errors. In fact, the excitation to the the characteristic level $n_c \sim \alpha^2/4$ typically leads to most of the bit-flip errors, since n_c is the lowest-lying level with well-delocalized states just above the potential barrier [15], below which the states are well localized while above which the occupation is weaker. For $\alpha^2 = 8$ considered in this paper, $n_c = 2$. With the above understanding of the error source, we should design the gate controls in a way that can greatly suppress the leakage to the dominant excited subspaces for both bit- and phase-flip errors. The tailored control for each gate is presented in the following section.

A. Gate Hamiltonians

To suppress the diabatic transitions, we first replace the hard square pulses with a family of truncated Gaussian pulses because of their favorable frequency selectivity and smoothness [23]:

$$\Omega_{G,m}(t) = A_m \left\{ \exp\left[-\frac{(t - T/2)^2}{2\sigma^2} \right] - \exp\left[-\frac{(T/2)^2}{2\sigma^2} \right] \right\}^m, \tag{4}$$

where m is chosen such that all of the first $m-1$ derivatives of $\Omega_{G,m}$ start and end at 0, A_m is a normalization constant, and σ is chosen to be equal to T in this paper. Then, we introduce the systematically designed derivative-based correction (DBC) Hamiltonian, $\hat{H}_{\mathrm{DBC}}$, to further suppress the leakage for each gate.

We summarize our design of corrections as the following (see Appendix C for details).

1. Z rotation

The Hamiltonian for the Z rotation in Ref. [15] is implemented by applying a linear drive (with hard pulses) to the KPO:

$$\hat{H}_Z = -K(\hat{a}^{2\dagger} - \alpha^2)(\hat{a}^2 - \alpha^2) + \Omega_0(t)\hat{a}^\dagger + \Omega_0^*(t)\hat{a}, \tag{5}$$

where $\Omega_0(t)$ is the hard pulse with amplitude determined by the desired rotation angle. Based on the discussion before, the nonadiabatic phase-flip errors dominantly come from the first excited subspace with the gap energy Δ_1, while the nonadiabatic bit-flip errors mainly come from the second excited subspace with the gap energy Δ_2. To suppress the first-order transitions at these two frequencies, we first replace the base driving pulse $\Omega_0(t)$ with the truncated Gaussian pulse with the second-order smoothness $\Omega_{G,2}(t)$ [see Eq. (4)] and then apply the derivative-based correction, which reshapes the pulse of the linear drive by adding derivatives of the base pulse, $\hat{H}_{\mathrm{DBC}} = u(t)\hat{a}^\dagger + u^*(t)a$, where

$$u(t) = -i\dot{\Omega}_{G,2}\left(\frac{1}{\Delta_1} + \frac{1}{\Delta_2}\right) - \frac{\ddot{\Omega}_{G,2}}{\Delta_1\Delta_2} + 0.07\frac{\Omega_{G,2}^3(t)}{\Delta_1^2}. \tag{6}$$

2. ZZ rotation

The ZZ rotation proposed in Ref. [15] is implemented by applying a beam-splitter interaction (with hard pulses) between the control mode (with subscript c) and the target mode (with subscript t):

$$\begin{aligned} \hat{H}_{ZZ}(t) &= \hat{H}_0 + \hat{V}(t), \\ \hat{H}_0 &= -K\left(\hat{a}_c^{2\dagger} - \alpha^2\right)\left(\hat{a}_c^2 - \alpha^2\right) - K\left(\hat{a}_t^{2\dagger} - \alpha^2\right)\left(\hat{a}_t^2 - \alpha^2\right), \\ \hat{V}(t) &= \Omega_0(t)\hat{a}_c\hat{a}_t^\dagger + \mathrm{H.c.} \end{aligned} \tag{7}$$

We find that the two-mode squeezing, which can also generate the ZZ rotation and has the same order of nonlinearity as the beam-splitter coupling, invokes fewer diabatic transitions and is thereby easier to deal with. So we first replace the beam-splitter term in the Hamiltonian with a two-mode squeezing

term, i.e., $\hat{V}(t) \to \Omega_0(t)\hat{a}_c^\dagger \hat{a}_t^\dagger + \text{H.c.}$ Then we identify the dominant excited subspaces for phase- and bit-flip errors. We use the notation (i, j) to label the subspace given by the tensor product between the ith subspace for the control mode and the jth subspace for the target mode. The excitation to the (1,0) [(0,1)] subspace dominantly leads to phase flips on the control (target) mode, the excitation to the (1,1) subspace dominantly leads to the correlated phase flips, and the excitation to the (2,0) [(0,2)] subspace dominantly produces the bit flips on the control (target) mode. Therefore we aim to suppress three sets of transitions with corresponding gap frequencies $\Delta_a \equiv \Delta_1$, $\Delta_b \equiv 2\Delta_1$, and $\Delta_c \equiv \Delta_1 + \Delta_2$ ($\Delta_2 \approx 2\Delta_1$) in order to suppress the abovementioned bit- and phase-flip errors (and the tensor product between them). To do this, we first replace the base driving pulse $\Omega_0(t)$ with the truncated Gaussian pulse with the third-order smoothness $\Omega_{G,3}(t)$ and then apply the derivative-based correction, which further reshapes the driving pulse of the two-mode squeezing by adding derivatives of $\Omega_{G,3}$, $\hat{H}_{\text{DBC}} = u(t)\hat{a}_c^\dagger \hat{a}_t^\dagger + \text{H.c.}$, where

$$u(t) = -i\dot{\Omega}_{G,3}(t)\left(\frac{1}{\Delta_a} + \frac{1}{\Delta_b} + \frac{1}{\Delta_c}\right) + i\frac{\dddot{\Omega}_{G,3}(t)}{\Delta_a\Delta_b\Delta_c} - \ddot{\Omega}_{G,3}(t)\left(\frac{1}{\Delta_a\Delta_b} + \frac{1}{\Delta_a\Delta_c} + \frac{1}{\Delta_b\Delta_c}\right) + 0.13\frac{\Omega_{G,3}^3(t)}{\Delta_a^2}. \tag{8}$$

3. CX gate

The original Hamiltonian in Ref. [15] for the CX gate between two modes reads

$$\begin{aligned}
\hat{H}_{\text{CX}}(t) &= \hat{H}_{\text{KPO}}^{(c)} + \hat{H}_{\text{KPO}}^{(t)}(t) + \hat{H}_{cp},\\
\hat{H}_{\text{KPO}}^{(c)} &= -K\left(\hat{a}_{\text{c}}^{\dagger 2} - \alpha^2\right)\left(\hat{a}_{\text{c}}^2 - \alpha^2\right),\\
\hat{H}_{\text{KPO}}^{(t)}(t) &= -K\left[\hat{a}_{\text{t}}^{\dagger 2} - \alpha^2 e^{-2i\phi(t)}\left(\frac{\alpha - \hat{a}_{\text{c}}^\dagger}{2\alpha}\right) - \alpha^2\left(\frac{\alpha + \hat{a}_{\text{c}}^\dagger}{2\alpha}\right)\right]\\
&\quad \times \left[\hat{a}_{\text{t}}^2 - \alpha^2 e^{2i\phi(t)}\left(\frac{\alpha - \hat{a}_{\text{c}}}{2\alpha}\right) - \alpha^2\left(\frac{\alpha + \hat{a}_{\text{c}}}{2\alpha}\right)\right],\\
\hat{H}_{cp} &= -\frac{1}{2}\dot{\phi}\frac{(2\alpha - \hat{a}_{\text{c}}^\dagger - \hat{a}_{\text{c}})}{2\alpha}(\hat{a}_{\text{t}}^\dagger \hat{a}_{\text{t}} - \alpha^2),
\end{aligned} \tag{9}$$

where $\hat{H}_{\text{KPO}}^{(t)}(t)$ stabilizes the phase of the target mode conditioned on the control mode and $\hat{H}_{cp}$ serves as a compensation Hamiltonian that partially compensates the nonadiabatic effects coming from the fast controlled phase rotation. The phase rotation is set to be linear, i.e., $\dot{\phi}(t) = \Omega_0(t)$ with $\Omega_0(t)$ being some hard pulse.

In the adiabatic frame (in which the cat states with time-dependent phase span the ground subspace) that block-diagonalizes $\hat{H}_{\text{KPO}}^{(c)} + \hat{H}_{\text{KPO}}^{(t)}(t)$, the term $\hat{H}_{cp}$ invokes diabatic transitions that lead to leakage outside the cat-state manifold. In contrast to the Z and ZZ rotation, the correction to the CX gate is more complicated and requires adding more physical correction terms to $\hat{H}_{\text{DBC}}$, because of the the nontrivial dynamics in the excited subspaces. We design the corrections to suppress the leakage to the (1,0), (0,1), and (1,1) excited levels, which is the dominant process for nonadiabatic phase- and bit-flip errors (the complicated structure of $\hat{H}_{\text{KPO}}^{(t)}$ leads to the fact that the leakage to the abovementioned subspaces also contributes greatly to the bit-flip errors). To realize this, we first replace the linear phase rotation with a more smooth rotation with a first-order Gaussian derivative, i.e., $\dot{\phi}(t) \to \Omega_{G,1}(t)$, and then apply the following derivative-based correction $\hat{H}_{\text{DBC}}$:

$$\begin{aligned}
\hat{H}_{\text{DBC}} &= \hat{H}_{\text{DBC},0} + \hat{H}_{\text{DBC},1} + \hat{H}_{\text{DBC},1} + \hat{H}_{\text{DBC},3},\\
\hat{H}_{\text{DBC},0} &= i\frac{d}{dt}\left[\frac{\dot{\phi}}{\Delta_{11}(t)}\right]\frac{\hat{a}_c - \hat{a}_c^\dagger}{4\alpha} \times (\hat{a}_t^\dagger \hat{a}_t - \alpha^2),\\
\hat{H}_{\text{DBC},1} &= c_1\frac{\dot{\phi}(1 - \cos 2\phi)}{\Delta_{11}(t)}(\hat{a} + \hat{a}^\dagger),\\
\hat{H}_{\text{DBC},2} &= ic_2\frac{\dot{\phi}\sin 2\phi}{\Delta_{11}(t)}(\hat{a}_c^2 - \hat{a}_c^{2\dagger}),\\
\hat{H}_{\text{DBC},3} &= c_3\frac{\dot{\phi}(t)}{\Delta_{11}(t)}[(e^{2i\phi(t)} - 1)\hat{a}_t^{2\dagger} + (e^{-2i\phi(t)} - 1)\hat{a}_t^2],
\end{aligned} \tag{10}$$

where $\Delta_{11}(t)$ is the time-dependent energy gap of the (1,1) subspace in the adiabatic frame that approximately diagonalizes $\hat{H}_{\text{KPO}}^{(t)}(t)$ and c_1, c_2, and c_3 are constants depending on the representation of the $\hat{a}$ operator in the Kerr-cat eigenbasis. See Appendix C for detailed expressions for $\Delta_{11}(t)$, c_1, c_2, and c_3. The correction terms $\hat{H}_{\text{DBC},i}$, $i = 0, 1, 2, 3$, are designed to address different diabatic transitions accordingly. The maximum order of nonlinearity required to implement these corrections is the same as the original Hamiltonian equation (9). We note that due to the high complexity of the CX Hamiltonian, we only derive corrections to suppress the leakage to low-lying excited subspaces, while in principle, similar to the Z and ZZ rotations, further corrections can be added to suppress the leakage to higher-lying subspaces, e.g., (0,2), (2,0), (1,2), or (2,1) subspaces, which may further suppress the nonadiabatic errors, especially the bit-flip errors.

B. Nonadiabatic errors

To illustrate the improvement from the DBC control, we numerically compare the nonadiabatic Z (P_z^{NA}) and X (P_x^{NA}) errors among different BP control schemes for the quantum operations of the Z rotation and the CX gate in Fig. 2. (The nonadiabatic errors of the ZZ rotation are similar to those of the Z rotation; see Fig. 8 in Appendix D.) Both the nonadiabatic Z and X errors of the Kerr gates with DBC control are significantly reduced.

We explicitly provide the scaling of P_z^{NA} with the gate time below. For dissipative gates, $P_z^{\text{NA}} \propto 1/\kappa_2 T$ is given by the integration the induced phase-flip rate over gate time [18]. For Kerr gates with hard or Gaussian control, $P_z^{\text{NA}} \propto |\int_0^T \Omega_0(t)e^{-i\Delta t}dt|^2$ is dominated by the first-order diabatic excitation associated with an energy gap Δ, which is proportional to the Fourier component of the driving pulse at Δ (for the Z rotation, $\Delta = \Delta_1$, while for the CX gate $\Delta = 2\Delta_1$). As such, $P_z^{\text{NA}} \propto 1/(KT)^2$ for hard pulses, while for truncated Gaussian pulses P_z^{NA} first scales exponentially for small KT and then polynomially for large KT due to the sideband excitation [23]. By adding derivative-based corrections to the bare Gaussian pulses we can dramatically suppress the excitation at Δ to the second order and maintain the exponen-

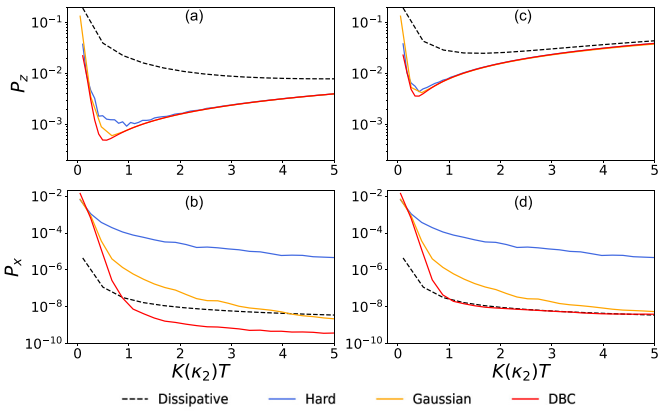

FIG. 3. The total Z and X error probabilities for the CX gate using different control schemes (same color scheme as Fig. 2) in the presence of photon loss. (a) and (b) $\kappa_1/K(\kappa_2) = 5 \times 10^{-5}$. (c) and (d) $\kappa_1/K(\kappa_2) = 5 \times 10^{-4}$. The gate time is in units of K (κ_2) for Kerr (dissipative) gates.

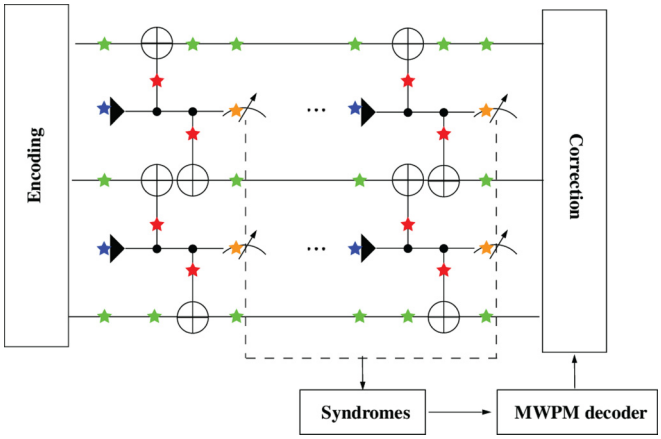

FIG. 4. The QEC circuit of the repetition cat. The potential faulty operations are the state preparation (blue), idling (green), CX gate (red), and measurement (orange). The measured syndromes are fed into the minimum weight perfect matching (MWPM) decoder to determine the errors.

tial error scaling in the regime of gate time of interest (see Appendixes B and C).

We note that since the bit-flip errors mostly come from the nonadiabatic effects of the gates even in the presence of photon loss (which will be discussed in the next section), it is especially important to suppress the nonadiabatic bit flips so that the gates can be implemented fast while preserving high noise bias, which is the prominent feature that makes the biased-noise qubits stand out. For the Kerr CX gate, the hard pulses lead to significant nonadiabatic bit-flip errors due to the large excitation to high excited states, therefore almost destroying the noise bias (for fast gates). Using Gaussian pulses can noticeably reduce the nonadiabatic bit flips since high excitations are suppressed due to the limited bandwidth. By adding the derivative-based corrections the excitations are further suppressed, and as a result, the Kerr CX gate can have large noise bias (even larger than the dissipative gates) with fast implementation ($T \sim 1/K$). To achieve the same level of bit-flip error, the DBC control can be much faster than the Gaussian control and consequently has lower loss-induced dephasing errors—which is crucial to reduce the resource overhead for concatenated QEC (see Sec. IV).

C. Gate performance in the presence of photon loss

In the presence of photon loss, the total Z error probability [Eq. (3)] will have a significant new contribution ($\sim\kappa_1|\alpha|^2T$), associated with the loss-induced parity change. In contrast, the total X error probability is still dominated by the nonadiabatic error $P_x^{\rm NA}$. In Fig. 3, we numerically obtain Z and X error probability of CX gates at different gate times with $\kappa_1/K(\kappa_2) = 5 \times 10^{-5}$ and 5×10^{-4}.

The total Z error probability P_z determines the gate fidelity. According to Eq. (3), the gate time can be optimized to minimize the total Z error probability. The Kerr gate can be implemented faster with higher gate fidelity than the dissipative gate, and compared with using hard or Gaussian control, using DBC control can further speed up the gate and improve the gate fidelity. Plugging in the scaling of $P_z^{\rm NA}$ with T, we can obtain the scaling of the minimal Z error probability P_z^* with $\kappa_1/K(\kappa_2)$. For dissipative gates, $P_{z,\rm Dissi}^* \propto (\frac{\kappa_1}{\kappa_2})^{1/2}$. For Kerr gates with hard control, $P_{z,\rm Hard}^* \propto (\frac{\kappa_1}{K})^{2/3}$. For Kerr gates with Gaussian or DBC control, $P_{z,\rm Gaussian}^*$ and $P_{z,\rm DBC}^*$ can approach the most favorable linear scaling, i.e., $P_{z,\rm Gaussian}^* \propto \frac{\kappa_1}{K}$ and $P_{z,\rm DBC}^* \propto \frac{\kappa_1}{K}$, with $P_{z,\rm DBC}^*$ being smaller. We obtain consistent scaling between P_z^* and $\kappa_1/K(\kappa_2)$ based on numerical fits for the CX gates (see Table III). Given the same small parameter $\kappa_1/K(\kappa_2)$, our fine control scheme can provide a smaller Z error probability and thus a more favorable optimal gate fidelity.

The total X error probability P_x limits the noise bias of the gates. Compared with the dissipative gate, the P_x of the Kerr gate with hard control is too high (at reasonable gate time). In contrast, the Kerr CX gate with Gaussian or DBC control can have P_x comparable to or even below that of the dissipative gate. In terms of gate time, the (counteradiabatic) DBC control significantly outperforms the Gaussian control (limited by the adiabatic requirement). Hence the DBC control can achieve a faster gate with favorable noise bias.

IV. CONCATENATED QUANTUM ERROR CORRECTION

We now compare the performance of different schemes of BP gates in terms of the logical gate failure rates in concatenated QEC. We consider the concatenation of the stabilized cats with a repetition code and simulate the logical gate failure rate using a circuit-level noise model, which includes state preparation errors, idling errors, CX gate errors, and measurement errors. The faulty operations are marked by stars with different colors in the QEC circuit shown in Fig. 4. The physical Z and X error rates of these operations are summarized in Table I. The Z errors of the CX gate comprise three parts: the Z error on the control mode which results from both nonadiabaticity and photon loss, the Z error on the target mode, and correlated Z errors which are induced only by photon loss. $p_0 = \kappa_1|\alpha|^2T_{\rm CX}$ is the characteristic photon-loss-induced Z error probability. For a distance-d repetition code we repeat

TABLE I. The physical error rates of different operations in our error model. $p_0 = \kappa_1|\alpha|^2 T_{\mathrm{CX}}$ is the characteristic Z error rate. The state preparation and measurement in the X basis do not produce bit-flip errors, and the bit-flip error generated during the idling is negligible compared with that generated during the CX gate.

Operation	Idle	$\mathcal{P}_{\vert+\rangle}$	$\mathcal{M}_X$	CX
P_z	p_0	p_0	p_0	$Z_c : p_z^{\mathrm{NA}}(T_{\mathrm{CX}}) + p_0$ $Z_t : \frac{1}{2}p_0$ $Z_cZ_t : \frac{1}{2}p_0$
P_x	≈ 0	0	0	$p_x(\kappa_1, T_{\mathrm{CX}})$

the syndrome extraction d times followed by one round of perfect syndrome extraction (assuming no ancilla errors), decode the full error syndromes using a minimum weight perfect matching (MWPM) decoder for our noise-biased system, and correct the errors by simply updating the Pauli frame.

We consider the total logical error rate of a transversal logical CX gate, which is a function of dimensionless parameters of the photon loss rate $\frac{\kappa_1}{K(\kappa_2)}$, the CX gate time $K(\kappa_2)T_{\mathrm{CX}}$, and the repetition-code distance d. Given κ_1, we can obtain the minimum logical CX gate error rate achievable by the repetition cat by optimizing T_{CX} and d (see Appendix F for details of the optimization):

$$P_L^{**}\left(\frac{\kappa_1}{K(\kappa_2)}\right) = \min_{T_{\mathrm{CX}},d} P_L\left(\frac{\kappa_1}{K(\kappa_2)}, K(\kappa_2)T_{\mathrm{CX}}, d\right). \tag{11}$$

In Fig. 5, we plot P_L^{**} and the corresponding optimal code distance d^{**} and gate time T_{CX}^{**} as a function of photon loss rate when different physical CX gates are used.

Using the Kerr CX gates with DBC control can lead to the lowest logical gate error for $\kappa_1/K < 10^{-3}$, which outperforms all other schemes. The optimized P_L^{**} with dissipative gate is mostly limited by the large P_z and the limited P_x in the low-loss regime [see Figs. 3(a)–3(c)], the P_L^{**} based on hard control is limited by the large nonadiabatic P_x error [see Figs. 3(b) and 3(d)], and the P_L^{**} based on Gaussian control has to follow the adiabatic requirement with the longest gate time [see Fig. 5(c)], which leads to extra overhead in the size of repetition-code d^{**}, especially for the practically relevant intermediate-loss regime ($\kappa_1/K \sim 10^{-4}$) in Fig. 5(b). As expected, the Gaussian control should be comparable for the DBC control for the extremely low loss regime ($\kappa_1/K \sim 10^{-5}$), which is compatible with slower adiabatic gates to maintain high gate performance. The reason is that the slower Gaussian gates mainly result in larger Z errors, while the moderate increase of Z errors (with X errors being similar) *in the far-below-threshold regime* does not affect the logical performance of the repetition code too much in terms of both the optimized logical error rate and the associated resource overhead.

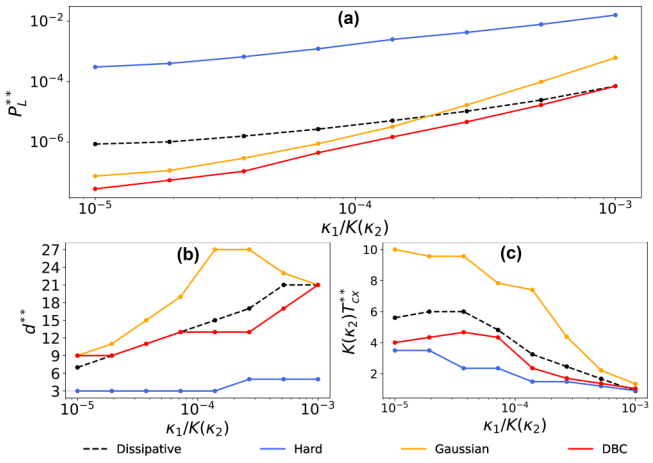

FIG. 5. (a). The minimal gate error P_L^{**} of the logical CX gate using physical BP CX gates with different controls. (b) and (c) The optimal choice of repetition-code distance (b) and the BP CX gate time (c) that minimizes the logical gate error.

In the above we optimize the logical error rates since the repetition cat cannot arbitrarily suppress logical errors in the regime of finite noise bias. However, we should instead optimize the resource overhead required for reaching a target logical error rate when we consider other concatenation QEC schemes that can arbitrarily suppress logical errors, such as the surface cat considered in Refs. [18,19]. In this case, using the fast Kerr gates presented in this paper can reduce the resource overhead towards fault tolerance since the dominant phase-flip errors are suppressed to be far below the threshold.

V. DISCUSSION

So far, the experimental Kerr parametric nonlinear oscillator can achieve $\kappa_1/K \approx 10^{-3}$ (with $\kappa_1/2\pi \approx 10$ kHz, $K/2\pi \approx 6.7$ MHz) [8], which is slightly more favorable than the engineered two-photon dissipation with $\kappa_1/\kappa_2 \approx 10^{-2}$ (with $\kappa_1/2\pi \approx 1.7$ kHz, $\kappa_2/2\pi \approx 170$ kHz) [28], partly because Kerr nonlinearity is less complicated to implement than the two-photon dissipation. Note that the single-photon loss rate is fairly high in all these experiments [7,8,28], to be further reduced in future devices. We expect that $\kappa_1/K(\kappa_2) \leqslant 10^{-4}$ (e.g., $\kappa_1/2\pi \approx 1$ kHz, $K/2\pi \approx 10$ MHz) can be achieved [15,18], which will enable us to achieve high-fidelity logical gates. We note that the presented control technique could potentially be also applied to other codes that require the implementation of bias-preserving gates, e.g., the pair-cat code [29].

VI. CONCLUSION

We use a derivative-based transition-suppression technique to suppress the leakage of the BP gates on the Kerr cat so that both the nonadiabatic phase-flip and nonadiabatic bit-flip errors can be reduced dramatically. In the presence of photon loss, we show that the Kerr gates with our designed DBC control can have higher gate fidelity while maintaining high noise bias. The improved gates, when applied in concatenated QEC, can lead to lower logical error rates and/or lower resource overhead.

ACKNOWLEDGMENTS

We thank Aashish Clerk, Kyungjoo Noh, Shruti Puri, Harry Putterman, and Hugo Ribeiro for helpful discussions. We also thank Christopher Chamberland for useful comments and suggestions on the concatenated quantum error

correction. The authors are also grateful for the support of the University of Chicago Research Computing Center for assistance with the numerical simulations carried out in this paper. We acknowledge support from the ARO (Grants No. W911NF-18-1-0020 and No. W911NF-18-1-0212), the ARO MURI (Grant No. W911NF-16-1-0349), the AFOSR MURI (Grant No. FA9550-19-1-0399), the NSF (Grants No. EFMA-1640959, No. OMA-1936118, and No. EEC-1941583), NTT Research, and the Packard Foundation (Grant No. 2013-39273).

APPENDIX A: THE SHIFTED FOCK BASIS AND THE KERR-CAT EIGENBASIS

Simulating a large cat qubit using the usual Fock basis is inefficient due to the wide photon-number distribution of a large coherent state. In contrast, one can work with the so-called shifted Fock basis, as proposed in Ref. [18], to simplify the analysis since typically only the first few excited states are populated. The shifted basis is defined as

$$|\phi_n, \pm\rangle \equiv \mathcal{N}_{n,\pm}[\hat{D}(\alpha) \pm (-1)^n \hat{D}(-\alpha)]|n\rangle, \tag{A1}$$

where $\hat{D}(\alpha)$ is the displacement operator and $\mathcal{N}_{n,\pm}$ is the normalization constant. In this basis, the Hilbert space is split into two subspaces, labeled by the photon-number parity + (even) and − (odd), respectively. Thus we can effectively represent each shifted Fock state as a tensor product of a "parity qubit" labeling the parity and a Fock state labeling the excitation level:

$$|p\rangle \otimes |n'\rangle \equiv |\phi_n, p\rangle. \tag{A2}$$

Since the logical information of the cat qubit is encoded in the parity, this "parity qubit" can also be viewed as a "logical qubit" carrying one bit of logical information of the cat qubit with Z-basis states $|0\rangle = \frac{1}{\sqrt{2}}(|+\rangle + |-\rangle)$ and $|1\rangle = \frac{1}{\sqrt{2}}(|+\rangle - |-\rangle)$. We note that these shifted Fock states are not exactly mutually orthogonal. However, they are nearly orthogonal for $n < |\alpha|^2/4$. We neglect the nonorthogonality for now when analyzing the low excited states. In this shifted Fock basis, the annihilation operator can be expressed as

$$\hat{a} = \hat{Z} \otimes (\hat{a}' + \alpha), \tag{A3}$$

where $\hat{Z}$ flips the phase of the "logical qubit" (or flips the parity of the "parity qubit") and $\hat{a}' \equiv \sum_n' \sqrt{n'}|n'-1\rangle\langle n'|$ is the bosonic annihilation operator defined on the excitation level of the shifted Fock states. Based on this representation of $\hat{a}$ we can write the Hamiltonian of a Kerr-cat qubit (Kerr parametric oscillator):

$$\begin{aligned}\hat{H}_{\rm KPO} &= -K(\hat{a}^{2\dagger} - \alpha^2)(\hat{a}^2 - \alpha^2) \\ &= -KI \otimes [4\alpha^2 \hat{a}'^{\dagger}\hat{a}' + 2\alpha(\hat{a}'^{2\dagger}\hat{a}' + \hat{a}'^{\dagger}\hat{a}'^2) + \hat{a}'^{2\dagger}\hat{a}'^2].\end{aligned} \tag{A4}$$

In the large-α limit, $H_{\rm KPO}$ is dominant by $-4\alpha^2 KI \otimes a'^{\dagger}a'$, which is diagonal in the shifted Fock basis. However, the off-diagonal elements cannot be neglected in this paper, and they have perturbative effects on the eigenvalues with respect to $1/\alpha$. Since the Kerr Hamiltonian preserves the photon-number parity, we can express its eigenstates as $|\psi_n, p\rangle = |p\rangle \otimes |n''\rangle$. We can perturbatively calculate $|n''\rangle$ in the basis of $\{|n'\rangle\}$ to the first order of $\frac{1}{\alpha}$:

$$\begin{aligned}|n''\rangle = &\frac{\alpha(n-1)\sqrt{n}}{2\alpha^2 + (n-1)}|n-1'\rangle + |n'\rangle \\ &- \frac{\alpha n\sqrt{n+1}}{2\alpha^2 + n}|n+1'\rangle + O((1/\alpha)^2),\end{aligned} \tag{A5}$$

where $n \geqslant 1$ and $|0''\rangle = |0'\rangle$.

If we only consider the first three pairs of eigenstates, i.e., $n' \leqslant 2$, we can express $\hat{a}'$ as

$$\hat{a}' = \sigma_{0,1}^- + \sqrt{2}\sigma_{1,2}^- - \lambda_1\Pi_1 - \lambda_2\Pi_2 + \eta\sigma_{0,2}^-, \tag{A6}$$

where $\lambda_1, \lambda_2, \eta = \frac{2\alpha}{2\alpha^2+1}, \frac{3\alpha}{\alpha^2+1}, \frac{\sqrt{2}\alpha}{2\alpha^2+1} + O((\frac{1}{\alpha})^2)$, with reduced Pauli operators and projectors in the Kerr-cat eigenbasis:

$$\begin{aligned}\sigma_{i,j}^- &\equiv |i''\rangle\langle j''|, \\ \Pi_i &\equiv |i''\rangle\langle i''|.\end{aligned} \tag{A7}$$

$\lambda_1, \lambda_2, \eta$ can be calculated more accurately by adding higher-order corrections. For $\alpha = \sqrt{8}$ used in this paper, we can numerically obtain

$$\lambda_1 = 0.4, \quad \lambda_2 = 1.08, \quad \eta = 0.256. \tag{A8}$$

APPENDIX B: ESTIMATION OF OFF-RESONANT EXCITATION IN THE ASYMPTOTIC LIMIT VIA FOURIER ANALYSIS

In this Appendix, following Ref. [23] we provide the estimation of off-resonant excitations in the asymptotic weak-drive limit, based on which we define the order of different transition elements in a Hamiltonian. In the case of constant energy gap, an off-resonant excitation can be estimated via the Fourier spectrum of its driving pulse in the asymptotic weak-drive limit. Only considering a single transition element $\hat{h}_k \equiv |\psi_k^{\rm to}\rangle\langle\psi_k^{\rm from}|$ associated with constant energy gap Δ_k that is driven by $\Omega(t)$, the propagator in the interaction picture is given by

$$\hat{U}_I(t) = \mathcal{T}\exp[-i\int_0^t dt'(\Omega(t')\hat{h}_k e^{i\Delta_k t'} + \text{H.c.})], \tag{B1}$$

where $\mathcal{T}$ is the time-ordering operator. In the weak-drive limit, $U_I(t)$ is dominantly given by the first-order Dyson expansion

$$\hat{U}_I^{(1)}(t) = -i\int_0^t dt'[\Omega(t')\hat{h}_k e^{i\Delta_k t'} + \text{H.c.}]; \tag{B2}$$

then at certain time T the off-resonant transition strength is given by the finite-time Fourier transform of $\Omega(t)$:

$$\langle\psi_k^{\rm to}|\hat{U}_I^{(1)}(T)|\psi_k^{\rm from}\rangle = \mathcal{F}(\Omega, \Delta_k, T) \equiv \int_0^T \Omega(t)e^{-i\Delta_k t}dt, \tag{B3}$$

and the population in the excited state is given by $|\mathcal{F}(\Omega, \Delta_k, t)|^2$. This finite-time Fourier transform can be connected to the standard Fourier transform by assuming that $\Omega(t)$ is truncated outside the $[0, T]$ time window or $\Omega(t)$ smoothly vanishes outside $[0, T]$, which is usually the case for a gate pulse that starts and ends at 0.

For a pulse whose time derivatives also start and end at 0 up to order m, its Fourier spectrum has the property that

$$\mathcal{F}(\Omega,\Delta,T)=(-i)^n\mathcal{F}\left(\frac{\frac{d^n}{dt^n}\Omega(t)}{\Delta^n}\right), \tag{B4}$$

for $n=1,2,\ldots,m+1$. Moreover, the Fourier spectrum of the product of derivatives of Ω can be converted to the Fourier spectrum of polynomials of Ω:

$$\mathcal{F}\left((\Omega(t))^{\sum n_k},\Delta,T\right)=\Theta\left(\mathcal{F}\left(\prod_k\left(\frac{1}{\Delta^k}\frac{d^k\Omega(t)}{dt^k}\right)^{n_k},\Delta,T\right)\right). \tag{B5}$$

To facilitate the perturbative analysis used throughout this paper, we define the order of the off-resonant transition elements with respect to the driving pulse $\Omega(t)$ and the energy gap Δ in the following way. We define the nth-order transition elements as those which will create nth-order excitation $e^{(n)}$ after gate time T:

$$e^{(n)}\propto\left|\mathcal{F}\left(\frac{\Omega^n}{\Delta^{n-1}},\Delta,T\right)\right|^2. \tag{B6}$$

Based on this definition the nth-order elements have a coefficient of $\frac{\Omega^n}{\Delta^{n-1}}$ or a product of time derivatives of Ω according to Eq. (B5). We note that the definition of $e^{(n)}$ coincides with the more conventional definition $e^{(n)}\propto(\frac{\Omega_0}{\Delta})^n$ if $\Omega(t)$ is a hard (square) pulse that has the amplitude Ω_0.

We use Eq. (B6) to estimate the nonadiabatic errors of the gates using different control schemes in the main text. We consider two types of driving pulses in this paper; one is the hard (square) pulse $\Omega_h(t)=\frac{\pi}{T}$, and the other is the truncated Gaussian pulses:

$$\Omega_{G,m}(t)=A_m\left\{\exp\left[-\frac{(t-T/2)^2}{2\sigma^2}\right]-\exp\left[-\frac{(T/2)^2}{2\sigma^2}\right]\right\}^m, \tag{B7}$$

where σ is set to be T.

For gates using hard pulses, the leakage is given by the first-order excitation:

$$e_h^{(1)}(T)\propto|\mathcal{F}(\Omega_h,\Delta,T)|^2=4\frac{\sin^2\Delta T}{(\Delta T)^2}, \tag{B8}$$

which scales quadratically with $1/\Delta T$.

If we simply replace the base driving pulse with truncated Gaussian pulses, the leakage is then $e_{G,m}^{(1)}$; if we add derivative-based corrections to suppress the first-order excitation (which will be discussed in Appendix C), the residual leakage is given by the second-order excitation $e_{G,m}^{(2)}$. $e_{G,m}^{(1)}$ and $e_{G,m}^{(2)}$ are defined as

$$e_{G,m}^{(1)}(T)\propto|\mathcal{F}(\Omega_{G,m},\Delta,T)|^2, \tag{B9}$$

$$e_{G,m}^{(2)}(T)\propto\left|\mathcal{F}\left(\frac{\Omega_{G,m}^2}{\Delta},\Delta,T\right)\right|^2. \tag{B10}$$

Their analytical expressions are lengthy, so we instead numerically plot them in Fig. 6. Compared with the first-order excitation of the hard pulse [Eq. (B8)], shown by the green curve, the first-order excitation of the truncated Gaussian pulse [Eq. (B9)], shown by the red curve, is more narrowband and has a faster-decaying envelope as ΔT. Going to the second order [Eq. (B10), blue curve] gives further dramatic improvement.

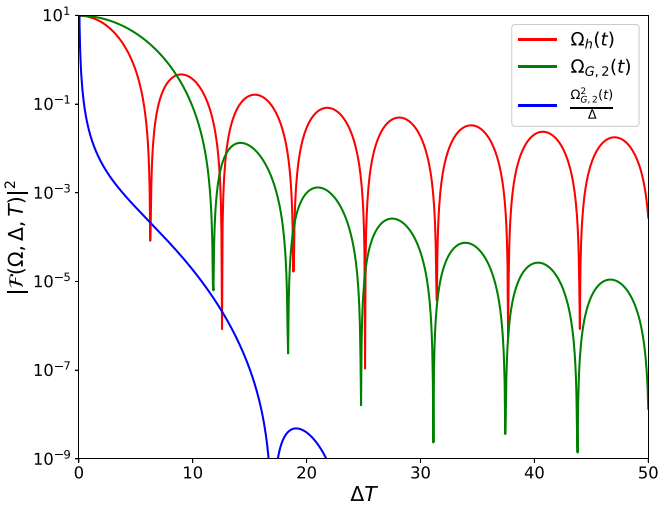

FIG. 6. The off-resonant excitation of different pulses, which is estimated by their finite-time Fourier spectrum. Red curve: the first-order excitation of the hard pulse [Eq. (B8)]. Green curve: the first-order excitation of the truncated Gaussian pulse [Eq. (B9)]. Blue curve: the second-order excitation of the truncated Gaussian pulse [Eq. (B10)].

In addition to facilitating the order counting and scaling analysis, the Fourier analysis also gives us a classical picture of how derivative-based transition suppression works. Using Eq. (B4), for an arbitrarily smooth driving pulse $\Omega(t)$, one can add its higher-order derivatives $\{\frac{d^n}{dt^n}\Omega(t)\}$ to create a set of "spectral holes" at one or more gap frequencies. In general, to create N holes at frequencies $\Delta_1,\Delta_2,\ldots,\Delta_N$, one can modify $\Omega(t)$ as [23]

$$\Omega'=\Omega-i\sum_k\frac{\partial_t\Omega}{\Delta_k}-\sum_k\sum_{j<k}\frac{\partial_t^2\Omega}{\Delta_k\Delta_j}+\cdots+\frac{(-i)^N\partial_t^N\Omega}{\Delta_1\Delta_2\cdots\Delta_N}, \tag{B11}$$

provided that the first $N-1$ derivatives of $\Omega(t)$ start and end at 0.

APPENDIX C: DERIVATIVE-BASED CORRECTIONS TO THE BP GATES

In this Appendix we present the details of our designed derivative-based corrections to each BP gate. The corrections are derived using the derivative-based transition-suppression technique [23,25].

We first derive the corrections to the Z rotation and ZZ rotation as the corrections to these two gates are simpler and illustrate the core idea. The system Hilbert space can be split in the following form: $\mathcal{H}=\mathcal{H}_{\text{logical}}\oplus\mathcal{H}_{\text{leak}}$, where $\mathcal{H}_{\text{logical}}$ is the logical subspace (cat subspace) inside which we would like our system to stay while $\mathcal{H}_{\text{leak}}$ is the subspace which we prevent our system from leaking into. We define $\hat{P}$ as the projector onto $\mathcal{H}_{\text{logical}}$, whereas we define $\hat{Q}$ as the complementary projector onto $\mathcal{H}_{\text{leak}}$. The original Hamiltonian we

apply to generate the desired dynamics is typically in the form

$$\hat{H}_{\text{original}} = \hat{H}_0(t) + \hat{V}(t), \tag{C1}$$

where $\hat{H}_0 = 0 \otimes \Delta \sum_k \frac{\Delta_k}{\Delta} \Pi_k$ is diagonal and only has support on the leakage subspace. $\hat{V}(t) = \Omega_0(t)\hat{V}_0^+ + \text{H.c.}$ is the control Hamiltonian whose projected part $\hat{P}\hat{V}(t)\hat{P}$ generates the desired logical operation while the block off-diagonal part $\hat{P}\hat{V}(t)\hat{Q} + \hat{Q}\hat{V}(t)\hat{P}$ causes leakage. Specifically, $\hat{Q}\hat{V}_0^+\hat{P} = \sum_k \lambda_k \hat{h}_k$ contains diabatic transitions to different leakage levels $|k\rangle$ with transition strength λ_k. These diabatic transitions are off-resonant since $\mathcal{H}_{\text{leak}}$ is gapped from $\mathcal{H}_{\text{logical}}$ by a set of energy gaps $\{\Delta_k\}$ in H_0. The leaked population can be estimated by the finite-time Fourier spectrum of $\Omega_0(t)$ at different gap energies in the asymptotic limit (see Appendix B).

To suppress the leakage, we aim to find a frame transformation, the so-called DRAG frame defined by $\hat{D}(t)$, that can block-diagonalize the original Hamiltonian $\hat{H}_{\text{original}}(t)$ and obtain the effective Hamiltonian in this DRAG frame:

$$\hat{H}_{\text{eff}} = \hat{D}(t)\hat{H}_{\text{original}}\hat{D}^\dagger(t) + i\dot{\hat{D}}(t)\hat{D}^\dagger(t), \tag{C2}$$

where $\hat{D}(t)\hat{H}_{\text{original}}\hat{D}^\dagger(t) = \hat{P}\hat{D}(t)\hat{H}_{\text{original}}\hat{D}^\dagger(t)\hat{P} + \hat{Q}\hat{D}(t)\hat{H}_{\text{original}}\hat{D}^\dagger(t)\hat{Q}$ is block diagonal and $i\dot{\hat{D}}(t)\hat{D}^\dagger(t)$ in general contains the block off-diagonal part. If we are able to remove the diabatic term $i\dot{\hat{D}}(t)\hat{D}^\dagger(t)$ by adding some corrections, we can remove all the leakage in this DRAG frame. If we further ensure that the DRAG frame coincides with the laboratory frame at the beginning and the end of the gate, i.e., $\hat{D}(0) = \hat{D}(T) = 0$, where T is the gate time, we then successfully remove all the leakage by the end of the gate in the laboratory frame. This correction corresponds to modifying the original Hamiltonian by adding a derivative-based correction:

$$\hat{H}_{\text{modified}} = \hat{H}_{\text{original}} + \hat{H}_{\text{DBC}}(t), \tag{C3}$$

where $\hat{H}_{\text{DBC}} = -i\hat{D}^\dagger(t)\dot{\hat{D}}(t)$. However, in practice we can neither perfectly block-diagonalize $\hat{H}_{\text{original}}(t)$ nor perfectly apply the desired corrections. So we have to perturbatively obtain both the frame transformation $\hat{D}(t)$ and the correction $\hat{H}_{\text{DBC}}$.

We define the adiabatic parameter as $\epsilon \equiv |\Omega_0/\Delta|$, upon which we perform the perturbation expansion. Specifically, we define $\hat{D}(t) = \exp[i\hat{S}(t)]$, where $\hat{S}(t) = \sum_{j=1} \hat{S}^{(j)}$ is expanded to different orders in ϵ. Similarly, $\hat{H}_{\text{DBC}} = \sum_j \hat{H}_{\text{DBC}}^{(j)}$ is also expanded to different orders in ϵ. We note the order in which the derivatives of $\Omega_0(t)$ should be counted by the rule $\frac{\partial_t^k \Omega_0}{\Delta^{k+1}} = \Theta(\epsilon)$. The effective Hamiltonian in this DRAG frame can also be expanded to different orders in ϵ using the time-dependent Schrieffer-Wolff (SW) expansion:

$$\begin{aligned}\hat{H}_{\text{eff}} &= \hat{D}(t)\hat{H}_{\text{modified}}\hat{D}^\dagger(t) + i\dot{\hat{D}}(t)\hat{D}^\dagger(t)\\ &= \sum_n \frac{1}{n!}[\hat{H}, -i\hat{S}]_n + (-i)^n \sum_n \frac{1}{(n+1)!}[\dot{\hat{S}}, \hat{S}]_n\\ &= \sum_{j=0} \hat{H}_{\text{eff}}^{(j)}(t),\end{aligned} \tag{C4}$$

where $[\hat{A}, \hat{B}]_n = [[\hat{A}, \hat{B}]_{n-1}, \hat{B}]$ and $[\hat{A}, \hat{B}]_0 = \hat{A}$. We note that $\hat{H}_0$ is of order 0 and $\hat{V}(t)$ is of order 1. We explicitly list $\hat{H}_{\text{eff}}$ up to the second order below:

$$\begin{aligned}\hat{H}_{\text{eff}}^{(0)} &= \hat{H}_0,\\ \hat{H}_{\text{eff}}^{(1)} &= \hat{V} + i[\hat{S}^{(1)}, \hat{H}_0] + \hat{H}_{\text{DBC}}^{(1)} + \dot{\hat{S}}^{(1)},\\ \hat{H}_{\text{eff}}^{(2)} &= i[\hat{S}^{(2)}, \hat{H}_0] + i\big[\hat{S}^{(1)}, \hat{V} + \hat{H}_{\text{DBC}}^{(1)}\big]\\ &\quad - \tfrac{1}{2}[\hat{S}^{(1)}, [\hat{S}^{(1)}, \hat{H}_0]] + \hat{H}_{\text{DBC}}^{(2)} + \dot{\hat{S}}^{(2)} - i[\dot{\hat{S}}^{(1)}, \hat{S}^{(1)}].\end{aligned} \tag{C5}$$

In this paper, we only correct the leakage error to the first few excited states to the first order, i.e., $\text{Tr}[\hat{H}_{\text{eff}}^{(1)}\hat{h}_k] = 0$ for $k = 1, 2, \ldots, N$, which can be satisfied simply by shaping the control pulse:

$$\hat{H}_{\text{DBC}}^{(1)}(t) = u(t)\hat{V}_0^+ + \text{H.c.}, \tag{C6}$$

where $u(t)$ is the classical solution that corresponds to creating N "spectral holes" at N different gap energies $\{\Delta_k\}$[23],

$$u = -i\sum_k \frac{\partial_t \Omega_0}{\Delta_k} - \sum_k \sum_{j<k} \frac{\partial_t^2 \Omega_0}{\Delta_k \Delta_j} + \cdots + \frac{(-i)^N \partial_t^N \Omega_0}{\Delta_1 \Delta_2 \cdots \Delta_N}, \tag{C7}$$

and the corresponding first-order DRAG frame transformation is given by

$$\begin{aligned}\hat{S}^{(1)} &= i\sum_k \frac{\Omega_0}{\Delta_k}\hat{h}_k + \sum_k \sum_{j\neq k} \frac{\partial_t \Omega_0}{\Delta_j \Delta_k}\hat{h}_k\\ &\quad - \sum_k \sum_{j\neq k} \sum_{i\neq j,k} \frac{\partial_t^2 \Omega_0}{\Delta_i \Delta_j \Delta_k}\hat{h}_k + \cdots\\ &\quad + \sum_k \frac{(-1)(-i)^N \partial_t^{N-1}\Omega_0}{\Delta_1 \Delta_2 \cdots \Delta_N}\hat{h}_k + \text{H.c.}\end{aligned} \tag{C8}$$

We note that this frame transformation can be linked to the superadiabatic expansion by applying successive frame transformation defined by $\exp[i\hat{S}_l^{(1)}]$, where l labels the lth superadiabatic transformation and $S_l^{(1)}$ contains the term in Eq. (C8) with the lth-order derivative of $\Omega_0(t)$.

By adding $\hat{H}_{\text{DBC}}^{(1)}$ the leakage errors are then suppressed to the second order, i.e., $O(|\mathcal{F}(\frac{\Omega_0^2}{\Delta}, \Delta, T)|^2)$. However, although the leakage error brought by the higher-order expansions in Eq. (C5) is smaller than the original first-order leakage, there will be phase or rotation errors acting on the logical subspace directly resulting from those higher-order expansions. As these terms are not associated with any energy gap, their contribution can be comparable to or even larger than the residual leakage error. To deal with these errors, we need to calculate $\hat{H}_{\text{eff}}^{(2)}$ (or higher-order terms) given $\hat{S}^{(1)}$ and $\hat{H}_{\text{DBC}}^{(1)}$ and add high-order corrections to $\hat{H}_{\text{DBC}}$.

Now we apply the derived correction strategy to the Z rotation and ZZ rotation, respectively.

1. Z rotation

The original Hamiltonian implementing a Z rotation on a single cat in Ref. [15] is

$$\begin{aligned}\hat{H}_{\text{original}} &= \hat{H}_0 + \hat{V}\\ &= -K_h(\hat{a}^{2\dagger} - \alpha^2)(\hat{a}^2 - \alpha^2) + \Omega_0(t)\hat{a}^\dagger + \Omega_0^*(t)\hat{a},\end{aligned} \tag{C9}$$

TABLE II. The dominant diabatic transitions of the ZZ gate and their associated energy gaps. Here, we use the fact that $\Delta_2 \approx 2\Delta_1$.

Transition	$\hat{Z}_c\hat{Z}_t \otimes (\hat{\sigma}_{0,1}^{+,c}, \hat{\sigma}_{0,1}^{+,t})$	$\hat{Z}_c\hat{Z}_t \otimes (\hat{\sigma}_{0,2}^{+,t}, \hat{\sigma}_{0,2}^{+,c}, \hat{\sigma}_{00,11}^{+})$	$\hat{Z}_c\hat{Z}_t \otimes (\hat{\sigma}_{00,12}^{+}, \hat{\sigma}_{00,21}^{+})$
Energy gap	$\Delta_a \equiv \Delta_1$	$\Delta_b \equiv 2\Delta_1(\Delta_2)$	$\Delta_c \equiv \Delta_1 + \Delta_2$

where $\hat{H}_0 = -K_h(\hat{a}^{2\dagger} - \alpha^2)(\hat{a}^2 - \alpha^2)$ and $\hat{V} = \Omega_0(t)\hat{a} + \Omega_0^*(t)\hat{a}^\dagger$. Ω_0 is a real hard pulse.

Working in the eigenbasis of the Kerr-cat Hamiltonian and only considering the first three pairs of eigenstates, we can express $\hat{H}_0$ and $\hat{V}$ as

$$\begin{aligned}\hat{H}_0 &= \hat{I} \otimes (\Delta_1\hat{\Pi}_1 + \Delta_2\hat{\Pi}_2),\\ \hat{V} &= 2\alpha\Omega_0(t)\hat{Z} \otimes [\hat{\Pi}_0 + (1-\lambda_1)\hat{\Pi}_1 + (1-\lambda_2)\hat{\Pi}_2]\\ &\quad + [\Omega_0(t)\hat{Z} \otimes (\hat{\sigma}_{0,1}^{+} + \sqrt{2}\hat{\sigma}_{1,2}^{+} + \eta\hat{\sigma}_{0,2}^{+}) + \text{H.c.}]. \quad \text{(C10)}\end{aligned}$$

In the second term of $\hat{V}(t)$ there are two diabatic transitions $\hat{Z} \otimes \hat{\sigma}_{0,1}^{+}, \hat{Z} \otimes \hat{\sigma}_{0,2}^{+}$ associated with two gap frequencies Δ_1, Δ_2 controlled by the same driving pulse. Then to suppress the first-order leakage, we first replace the base driving pulse as the Gaussian pulse with second-order smoothness $\Omega_0(t) = \Omega_{G,2}(t)$ and then add H_{DBC} by shaping the base driving pulse, i.e., $\hat{H}_{\text{DBC}} = u(t)\hat{a}^\dagger + u^*(t)\hat{a}$. According to Eq. (C7) the first-order correction pulse is

$$u^{(1)}(t) = -i\dot{\Omega}_0\left(\frac{1}{\Delta_1} + \frac{1}{\Delta_2}\right) - \frac{\ddot{\Omega}_0}{\Delta_1\Delta_2} \quad \text{(C11)}$$

and the corresponding first-order DRAG transformation [according to Eq. (C8)] is

$$\begin{aligned}\hat{S}^{(1)}(t) = i\Omega_0\hat{Z} \otimes &\left(\frac{\hat{\sigma}_{0,1}^{+}}{\Delta_1} + \frac{\hat{\sigma}_{1,2}^{+}}{\Delta_1} + \eta\frac{\hat{\sigma}_{0,2}^{+}}{\Delta_2}\right) + \frac{\dot{\Omega}_0}{\Delta_1\Delta_2}\hat{Z}\\ &\otimes (\hat{\sigma}_{0,1}^{+} + \hat{\sigma}_{1,2}^{+} + \eta\hat{\sigma}_{0,2}^{+}) + \text{H.c.} \quad \text{(C12)}\end{aligned}$$

Then we can block-diagonalize $\hat{H}_{\text{eff}}^{(1)}$, i.e., $\hat{P}\hat{H}_{\text{eff}}^{(1)}\hat{P} = 2\alpha\Omega_x(t)\hat{Z} \otimes \hat{\Pi}_0$, $\hat{Q}\hat{H}_{\text{eff}}^{(1)}\hat{P} = 0$ (only to the first three pairs of eigenstates). The residual error is then of order $O(|\mathcal{F}(\frac{\Omega_0^2}{\Delta}, \Delta, T)|^2)$, which comes from $\hat{H}_{\text{eff}}^{(2)}$ and higher-order expansions. In addition to the leakage error, however, there will be terms from $\hat{H}_{\text{eff}}^{(3)}$ (and higher-order expansions) that cause over-rotation ($\hat{Z} \otimes \hat{\Pi}_0$). The over-rotation angle is given by $\delta\theta \propto \mathcal{F}(\frac{\Omega_0^3}{\Delta_1^2}, 0, T)$, which can be corrected by renormalizing the rotation angle or adding an additional term to u:

$$u(t) = u^{(1)}(t) + c_0\frac{\Omega_0^3(t)}{\Delta_1^2}. \quad \text{(C13)}$$

Instead of doing lengthy calculations we simply numerically optimize c_0 and obtain $c_0 \approx 0.07$.

2. ZZ rotation

The original Hamiltonian implementing a ZZ rotation on two cats in Ref. [15] is

$$\begin{aligned}\hat{H}_{\text{original}}(t) &= \hat{H}_0 + \hat{V}(t),\\ \hat{H}_0 &= -K\left(\hat{a}_c^{2\dagger} - \alpha^2 big\right)\left(\hat{a}_c^2 - \alpha^2\right)\\ &\quad - Kbig(\hat{a}_t^{2\dagger} - \alpha^2)\left(\hat{a}_t^2 - \alpha^2\right),\\ \hat{V}(t) &= \Omega_0(t)\hat{a}_c\hat{a}_t^\dagger + \text{H.c.}, \quad \text{(C14)}\end{aligned}$$

where $\Omega_0(t)$ is again a real hard pulse.

We find that it is hard to apply the derivative-based transition-suppression technique with this original Hamiltonian since there are too many transition elements with distinct gap frequencies to be suppressed. Instead, it will be easier if we replace the beam-splitter interaction with two-mode squeezing $\hat{V}(t) = \Omega_0(t)\hat{a}_c^\dagger\hat{a}_t^\dagger + \text{H.c.}$

Again, working in the Kerr-cat eigenbasis and only considering the first three pairs of eigenstates for each mode, we can express $\hat{H}_0$ and $\hat{V}$ as

$$\begin{aligned}\hat{H}_0 &= \Delta_1(\hat{\Pi}_1^c + \hat{\Pi}_1^t) + \Delta_2(\hat{\Pi}_2^c + \hat{\Pi}_2^t),\\ \hat{P}\hat{V}(t)\hat{P} &= 2\alpha^2\Omega_0\hat{Z}_c\hat{Z}_t \otimes \hat{\Pi}_0,\\ \hat{P}\hat{V}(t)\hat{Q} + \hat{Q}\hat{V}(t)\hat{P} &= \Omega_0(t)\hat{Z}_c\hat{Z}_t \otimes [\alpha(\hat{\sigma}_{00,10}^{+} + \hat{\sigma}_{00,01}^{+})\\ &\quad + \eta\alpha(\hat{\sigma}_{00,20}^{+} + \hat{\sigma}_{00,02}^{+}) + \hat{\sigma}_{00,11}^{+}\\ &\quad + \eta(\hat{\sigma}_{00,12}^{+} + \hat{\sigma}_{00,21}^{+}) + \eta^2\hat{\sigma}_{00,22}^{+}] + \text{H.c.}, \quad \text{(C15)}\end{aligned}$$

where $\hat{\sigma}_{ij,i'j'}^{+} \equiv |i'\rangle_c\langle j'|_t\langle i|_c\langle j|_t$. Here, we only show the projection of $\hat{V}(t)$ in $\mathcal{H}_{\text{logical}}$ and its block off-diagonal part for simplicity.

Among all the diabatic transitions in the block off-diagonal part of $\hat{V}$ we consider suppressing three groups of them associated with three energy gaps, Δ_a, Δ_b, and Δ_c, listed in Table II. Similar to the Z rotation, to suppress the first-order leakage, we first replace the base driving pulse as the Gaussian pulse with third-order smoothness $\Omega_0(t) = \Omega_{G,3}(t)$ and then add H_{DBC} by shaping the base driving pulse, i.e., $\hat{H}_{\text{DBC}} = u(t)\hat{a}_c^\dagger\hat{a}_t^\dagger + \text{H.c.}$ According to Eq. (C7) the first-order correction pulse is

$$\begin{aligned}u^{(1)}(t) = &-i\dot{\Omega}_0(t)\left(\frac{1}{\Delta_a} + \frac{1}{\Delta_b} + \frac{1}{\Delta_c}\right)\\ &- \ddot{\Omega}_0(t)\left(\frac{1}{\Delta_a\Delta_b} + \frac{1}{\Delta_a\Delta_c} + \frac{1}{\Delta_b\Delta_c}\right) + i\frac{\dddot{\Omega}_0(t)}{\Delta_a\Delta_b\Delta_c}, \quad \text{(C16)}\end{aligned}$$

and the corresponding first-order DRAG transformation is

$$\begin{aligned}\hat{S}^{(1)}(t) = i\Omega_0&\left(\frac{\hat{h}_a}{\Delta_a} + \frac{\hat{h}_b}{\Delta_b} + \frac{\hat{h}_c}{\Delta_c}\right) + \dot{\Omega}_0\left[\frac{1}{\Delta_a}\left(\frac{1}{\Delta_b} + \frac{1}{\Delta_c}\right)\hat{h}_a\right.\\ &+ \frac{1}{\Delta_b}\left(\frac{1}{\Delta_a} + \frac{1}{\Delta_c}\right)\hat{h}_b + \left.\frac{1}{\Delta_c}\left(\frac{1}{\Delta_a} + \frac{1}{\Delta_b}\right)\hat{h}_c\right]\\ &- \frac{\ddot{\Omega}_x}{\Delta_a\Delta_b\Delta_c}(\hat{h}_a + \hat{h}_b + \hat{h}_c) + \text{H.c.}, \quad \text{(C17)}\end{aligned}$$

where

$$\begin{aligned}\hat{h}_a &= \hat{Z}_c\hat{Z}_t \otimes [\alpha(\hat{\sigma}_{00,10}^{+} + \hat{\sigma}_{00,10}^{+}) + \sqrt{2}\alpha(\hat{\sigma}_{01,02}^{+} + \hat{\sigma}_{10,20}^{+})],\\ \hat{h}_b &= \eta\alpha\hat{Z}_c\hat{Z}_t \otimes [\hat{\sigma}_{00,11}^{+} + \eta\alpha(\hat{\sigma}_{00,02}^{+} + \hat{\sigma}_{00,20}^{+}],\\ \hat{h}_c &= \eta\hat{Z}_c\hat{Z}_t \otimes (\hat{\sigma}_{00,12}^{+} + \hat{\sigma}_{00,21}^{+}). \quad \text{(C18)}\end{aligned}$$

Then the diabatic transitions are suppressed to the second-order, $O(|\mathcal{F}(\frac{\Omega_x^2}{\Delta}, \Delta, T)|^2)$. Similar to the Z rotation, there will be an over-rotation $\hat{Z}_c\hat{Z}_t \otimes \hat{\Pi}_0^c\hat{\Pi}_0^t$ coming from the higher-order DRAG expansion, which can be compensated by rescaling the rotation angle or adding an additional term to u:

$$u(t) = u^{(1)} + c_0\frac{\Omega_0^3(t)}{\Delta_a^2}, \tag{C19}$$

where we also numerically optimize c_0 and obtain $c_0 \approx 0.13$.

3. CX gate

The correction to the CX gate is more complicated and requires adding more physical correction terms to $\hat{H}_{\rm DBC}$ instead of merely pulse shaping.

The original Hamiltonian for the CX gate in Ref. [15] is

$$\begin{aligned}
\hat{H}_{\rm original}(t) &= \hat{H}^{(c)}_{\rm KPO} + \hat{H}^{(t)}_{\rm KPO}(t) + \hat{H}_{cp}, \quad \hat{H}^{(c)}_{\rm KPO} = -K\big(\hat{a}_{\rm c}^{\dagger 2} - \alpha^2\big)\big(\hat{a}_{\rm c}^2 - \alpha^2\big),\\
\hat{H}^{(t)}_{\rm KPO}(t) &= -K\left[\hat{a}_{\rm t}^{\dagger 2} - \alpha^2 e^{-2i\phi(t)}\left(\frac{\alpha - \hat{a}_{\rm c}^\dagger}{2\alpha}\right) - \alpha^2\left(\frac{\alpha + \hat{a}_{\rm c}^\dagger}{2\alpha}\right)\right]\left[\hat{a}_{\rm t}^2 - \alpha^2 e^{2i\phi(t)}\left(\frac{\alpha - \hat{a}_{\rm c}}{2\alpha}\right) - \alpha^2\left(\frac{\alpha + \hat{a}_{\rm c}}{2\alpha}\right)\right],\\
\hat{H}_{cp} &= -\frac{1}{2}\dot{\phi}\frac{(2\alpha - \hat{a}_{\rm c}^\dagger - \hat{a}_{\rm c})}{2\alpha}(\hat{a}_{\rm t}^\dagger\hat{a}_{\rm t} - \alpha^2).
\end{aligned} \tag{C20}$$

We first represent the operators on the control mode in the shifted Fock basis and express $\hat{H}^{(t)}_{\rm KPO}$ and $\hat{H}_{cp}$ as

$$\begin{aligned}
\hat{H}^{(t)}_{\rm KPO} &= -K[\hat{a}_t^{2\dagger} - \alpha^2 e^{-2i\phi}\hat{P}_c^- - \alpha^2\hat{P}_c^+ + \frac{1}{2}\alpha(e^{-2i\phi} - 1)\hat{Z}_c \otimes \hat{a}_c'^\dagger]\left[\hat{a}_t^2 - \alpha^2 e^{2i\phi}\hat{P}_c^- - \alpha^2\hat{P}_c^+ + \frac{1}{2}\alpha(e^{2i\phi} - 1)\hat{Z}_c \otimes \hat{a}_c'\right],\\
\hat{H}_{\rm cp} &= -\dot{\phi}\left[\hat{P}_c^- - \frac{1}{4\alpha}(\hat{a}_c' + \hat{a}_c'^\dagger)\right](\hat{a}_t^\dagger\hat{a}_t - \alpha^2),
\end{aligned} \tag{C21}$$

where $\hat{P}_c^+ \equiv \frac{\hat{I}_c + \hat{Z}_c}{2}$ and $\hat{P}_c^- \equiv \frac{\hat{I}_c - \hat{Z}_c}{2}$.

We define the following adiabatic frame transformation:

$$\hat{U}(t) = \exp\left[-i\int_0^t dt'\dot{\phi}(t')\hat{P}_c^- \otimes (\hat{a}_t^\dagger\hat{a}_t - \alpha^2)\right] = \hat{P}_c^- \otimes \exp[-i\phi(t)(\hat{a}_t^\dagger\hat{a}_t - \alpha^2)] + \hat{P}_c^+ \tag{C22}$$

and obtain the Hamiltonian in the adiabatic frame as

$$\begin{aligned}
\hat{\tilde{H}}_{\rm original}(t) &= \hat{U}\hat{H}_{\rm original}\hat{U}^\dagger + i\dot{\hat{U}}\hat{U}^\dagger = \hat{\tilde{H}}_0 + \hat{\tilde{H}}_v + \hat{\tilde{V}},\\
\hat{\tilde{H}}_0 &= -K\big(\hat{a}_t^{2\dagger} - \alpha^2\big)\big(\hat{a}_t^2 - \alpha^2\big) - K\big(\hat{a}_c^{2\dagger} - \alpha^2\big)\big(\hat{a}_c^2 - \alpha^2\big),\\
\hat{\tilde{H}}_v &= -\frac{1}{2}K\alpha[i\sin 2\phi\hat{Z}_c - (1 - \cos 2\phi)\hat{I}_c] \otimes \hat{a}_c'\big(\hat{a}_t^{2\dagger} - \alpha^2\big) + \text{H.c.} - \frac{1}{2}K\alpha^2(1 - \cos 2\phi)\hat{I}_c\hat{a}_c'^\dagger\hat{a}_c',\\
\hat{\tilde{V}} &= \frac{1}{4\alpha}\dot{\phi}\hat{Z}_c \otimes (\hat{a}_c' + \hat{a}_c'^\dagger) \times (\hat{a}_t^\dagger\hat{a}_t - \alpha^2).
\end{aligned} \tag{C23}$$

Only considering the first two pairs of eigenstates and expressing the annihilation operator in the Kerr-cat eigenbasis as $\hat{a}_{c,t} = \hat{Z}_{c,t} \otimes (\alpha + \hat{\sigma}_{0,1}^{-,c,t} - \lambda_1\hat{\Pi}_1^{c,t})$, we can write the Hamiltonian equation (C23) as

$$\begin{aligned}
\hat{\tilde{H}}_0 &= \Delta_1\hat{I}_t \otimes \hat{\Pi}_1^t + \big[\Delta_1 - \tfrac{1}{2}K\alpha^2(1 - \cos 2\phi)\big]\hat{I}_c \otimes \hat{\Pi}_1^c + 2K\alpha^2\lambda_1^2(1 - \cos 2\phi)\hat{I}_c\hat{I}_t \otimes \hat{\Pi}_1^c\hat{\Pi}_1^t,\\
\hat{\tilde{H}}_v &\approx K\alpha^2(1 - \cos 2\phi)\hat{I}_c\hat{I}_t \otimes \big[\hat{\sigma}_{01,10}^x - \lambda_1\big(\hat{\Pi}_1^c\hat{\sigma}_{0,1}^{x,t} + \hat{\sigma}_{0,1}^{x,c}\hat{\Pi}_1^t\big)\big] - K\alpha^2\sin 2\phi\hat{Z}_c\hat{I}_t \otimes \big[\hat{\sigma}_{01,10}^y - \lambda_1\big(\hat{\sigma}_{0,1}^{y,c}\hat{\Pi}_1^t - \hat{\Pi}_1^c\hat{\sigma}_{0,1}^{y,t}\big)\big],\\
\hat{\tilde{V}} &\approx \frac{1}{4}\dot{\phi}\hat{Z}_c\hat{I}_t \otimes \hat{\sigma}_{0.1}^{x,c}\hat{\sigma}_{0.1}^{x,t}.
\end{aligned} \tag{C24}$$

In this laboratory frame, the dominant leakage transition is $\hat{Z}_c\hat{I}_t \otimes \hat{\sigma}_{00,11}^+$ via $\hat{\tilde{V}}$, and its associated time-dependent energy gap is $\Delta_{11}(t) = 2\Delta_1 - \frac{1}{2}K\alpha^2(1 - 3\lambda_1^2)[1 - \cos 2\phi(t)]$. We can define the first-order DRAG transformation with respect to this transition,

$$\hat{S}^{(1)} = -\frac{1}{4}\frac{\dot{\phi}}{\Delta_{11}(t)}\hat{Z}_c\hat{I}_t \otimes \hat{\sigma}_{00,11}^y, \tag{C25}$$

and obtain the first-order effective Hamiltonian in this DRAG frame:

$$\begin{aligned}
\hat{H}_{\rm eff}^{(1)} &= \hat{\tilde{V}} + i[\hat{S}^{(1)}, \hat{\tilde{H}}_0] + i[\hat{S}^{(1)}, \hat{\tilde{H}}_v] = \frac{1}{4}K\alpha\left(\alpha - \frac{1}{2}\lambda_1\right)\lambda_1\frac{\dot{\phi}(1 - \cos 2\phi)}{\Delta_{11}(t)}\hat{Z}_c\hat{I}_t \otimes \big(\hat{\sigma}_{0,1}^{x,c}\hat{\Pi}_0^t + \hat{\Pi}_0^c\hat{\sigma}_{0,1}^{x,t}\big)\\
&\quad - \frac{1}{4}K\alpha\left(\alpha - \frac{1}{2}\lambda_1\right)\lambda_1\frac{\dot{\phi}\sin 2\phi}{\Delta_{11}(t)}\hat{I}_c\hat{I}_t \otimes \big(\hat{\Pi}_0^c\hat{\sigma}_{0,1}^{y,t} - \hat{\sigma}_{0,1}^{y,c}\hat{\Pi}_0^t\big) - \frac{1}{4}\frac{d}{dt}\left[\frac{\dot{\phi}}{\Delta_{11}(t)}\right]\hat{Z}_c\hat{I}_t \otimes \hat{\sigma}_{00,11}^y.
\end{aligned} \tag{C26}$$

To eliminate the $\hat{Z}_c\hat{I}_t \otimes \hat{\sigma}^{+}_{00,11}$ transition, we can add the standard derivative correction $\hat{H}^{(1)}_{\mathrm{DBC},0} = iu_0(t)\frac{\hat{a}_c-\hat{a}_c^\dagger}{4\alpha} \times (\hat{a}_t^\dagger\hat{a}_t - \alpha^2)$, where $u_0(t) = \frac{d}{dt}[\frac{\dot{\phi}}{\Delta_{11}(t)}]$.

Since $\hat{\tilde{H}}_v$ is not diagonal in the Kerr-cat eigenbasis and is of order ϵ^0, $i[\hat{S}^{(1)}, \hat{\tilde{H}}_v]$ adds new diabatic transitions to $H^{(1)}_{\mathrm{eff}}$, which do not exist in the laboratory frame. So we need to add additional terms to H_{DBC} to suppress these transitions.

To suppress $\hat{Z}_c\hat{I}_t \otimes \hat{\sigma}^{x,c}_{0,1}\hat{\Pi}^t_0$, we can add

$$\hat{H}^{(1)}_{\mathrm{DBC},1} = c_1\frac{\dot{\phi}(1-\cos 2\phi)}{\Delta_{11}(t)}(\hat{a}_c + \hat{a}_c^\dagger),$$
$$c_1 = \frac{1}{4}K\alpha\left(\alpha - \frac{1}{2}\lambda_1\right)\lambda_1. \tag{C27}$$

To suppress $\hat{I}_c\hat{I}_t \otimes \hat{\sigma}^{y,c}_{0,1}\hat{\Pi}^t_0$, we can add

$$\hat{H}^{(1)}_{\mathrm{DBC},2} = ic_2\frac{\dot{\phi}\sin 2\phi}{\Delta_{11}(t)}\left(\hat{a}_c^2 - \hat{a}_c^{2\dagger}\right), \quad c_2 = \frac{1}{8}K\alpha\lambda_1. \tag{C28}$$

Finally, we can simultaneously suppress $\hat{Z}_c\hat{I}_t \otimes \hat{\Pi}^c_0\hat{\sigma}^{x,t}_{0,1}$ and $\hat{I}_c\hat{I}_t \otimes \hat{\Pi}^c_0\hat{\sigma}^{y,t}_{0,1}$ by adding

$$\hat{H}^{(1)}_{\mathrm{DBC},3} = c_3\frac{\dot{\phi}(t)}{\Delta_{11}(t)}\left[(e^{2i\phi(t)}-1)\hat{a}_t^{2\dagger} + (e^{-2i\phi(t)}-1)\hat{a}_t^2\right],$$
$$c_3 = \frac{1}{8}K\alpha\lambda_1. \tag{C29}$$

So, in total, we apply four derivative-based corrections $\hat{H}^{(1)}_{\mathrm{DBC}} = \hat{H}^{(1)}_{\mathrm{DBC},0} + \hat{H}^{(1)}_{\mathrm{DBC},1} + \hat{H}^{(1)}_{\mathrm{DBC},2} + \hat{H}^{(1)}_{\mathrm{DBC},3}$ to suppress the first-order leakage to the first pair of excited states of each mode.

However, by adding $\hat{H}^{(1)}_{\mathrm{DBC},1}$ and $\hat{H}^{(1)}_{\mathrm{DBC},3}$ we also induce extra unitary Z rotation on the control mode, which can be compensated by applying an additional Z rotation $\hat{Z}(\delta\theta)$ on the control mode after the CX gate, which has negligible error compared with the CX gate. $\delta\theta$ can be calculated as

$$\delta\theta = \frac{1}{4}K\alpha^2\lambda_1(3\alpha - \lambda_1)\int_0^T dt\frac{\dot{\phi}(t)[1-\cos 2\phi(t)]}{\Delta_{11}(t)}. \tag{C30}$$

APPENDIX D: DETAILS OF THE NUMERICAL SIMULATIONS

All simulations in this paper use the Kerr-cat eigenbasis $\{|\psi_n^\pm\rangle\}$ with $n = 0, 1, \ldots, d-1$, where d is the truncation number and the dimension of the truncated Hilbert space for each mode is $2 \times d$. The simulations of Z and ZZ rotation are in the laboratory frame. However, to simulate the CX gate using the Kerr-cat basis we need to move to the adiabatic frame defined in Eq. (C22). The projector $\hat{P}_c^-$ is obtained approximately by

$$\hat{P}_c^- = \sum_{n=0}^{d-1}\frac{1}{2}(|\psi_n^+\rangle + (-1)^n|\psi_n^-\rangle)(\langle\psi_n^+| + (-1)^n\langle\psi_n^-|)$$
$$= \sum_{n=0}^{d-1}|1\rangle_L \otimes |n''\rangle, \tag{D1}$$

where the second line uses the subsystem decomposition.

To numerically extract the gate errors shown in the main text (Figs. 2 and 3), we initialize the cats to certain initial states, simulate the gate dynamics, apply a strong two-photon dissipation ($\mathcal{D}[a^2 - \alpha^2]$) at the end of each gate, and calculate the final state fidelity with target states. The application of the two-photon dissipation, which is a mathematical completely positive trace-preserving (CPTP) map that preserves the parity and locality in phase space, thus preserving the logical information, pushes all the leakage back into the cat-state manifold

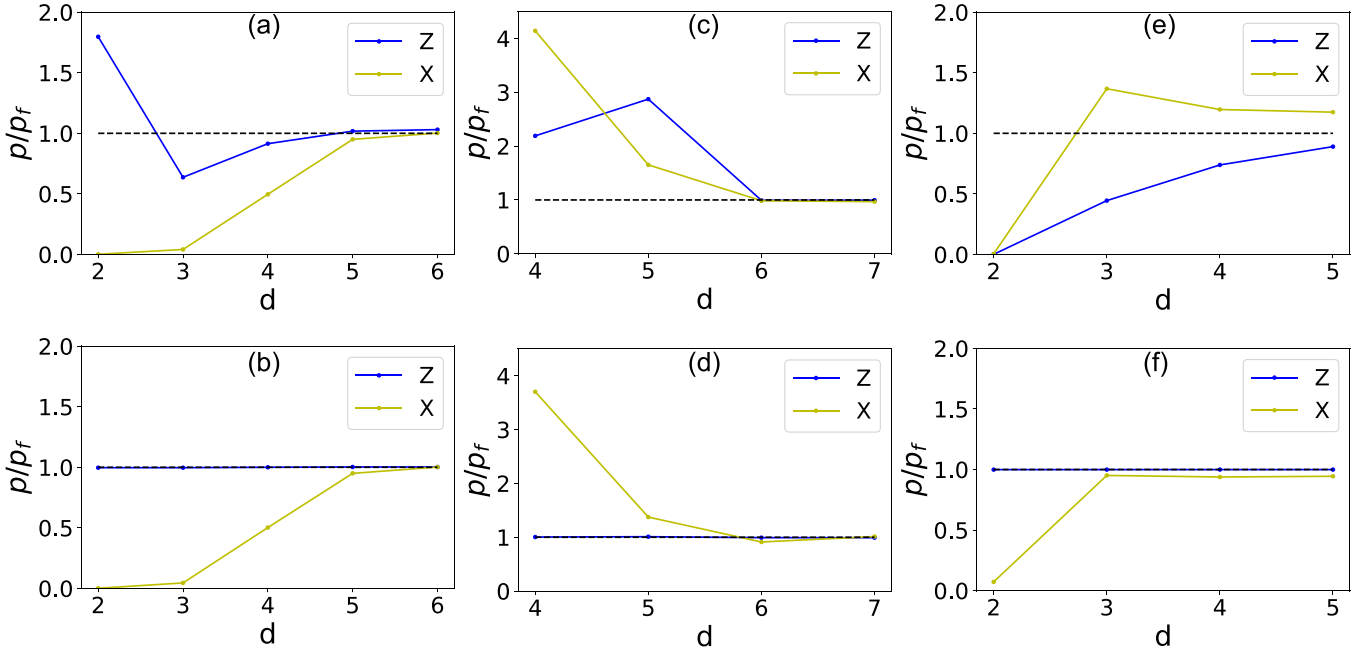

FIG. 7. Convergency of the numerical simulation using the Kerr-cat eigenbasis for $\alpha = \sqrt{8}$. (a), (c), and (e) show the results when there is no photon loss. (b), (d), and (f) show the results when there is photon loss with $\kappa_1/K = 10^{-4}$. The choice of the gate time for these three types of gates is $T_{\mathrm{Z}} = T_{\mathrm{ZZ}} = 0.2/K$, $T_{\mathrm{CX}} = 1/K$.

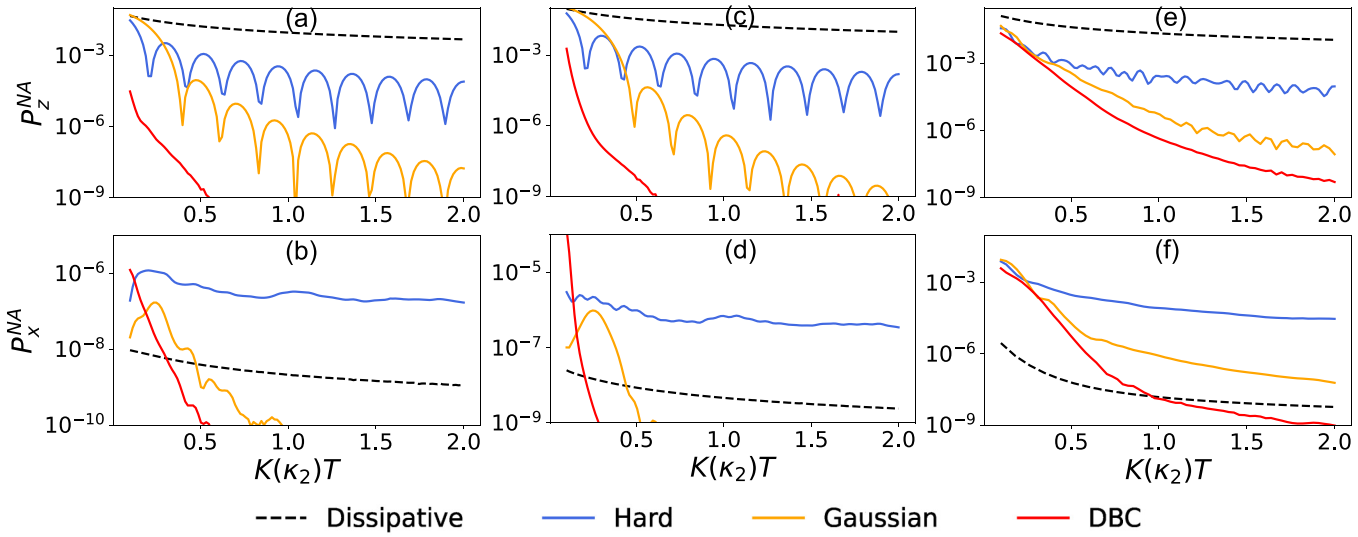

FIG. 8. The nonadiabatic errors of all three gates considered in this paper including the *ZZ* rotation, which is not included in the main text. (a) and (b) *Z* rotation. (c) and (d) *ZZ* rotation. (e) and (f) CX gate.

in which we evaluate the logical information. We note that alternatively, we can directly evaluate the logical information by measuring the logical sector while tracing out the gauge sector in the subsystem decomposition shown in Appendix A. In principle, the later measurement can be done by measuring the photon-number parity for the (logical) z-axis measurement and doing the homodyne measurement (to distinguish between $|\alpha\rangle$ and $|-\alpha\rangle$) for the (logical) x-axis measurement. For simplicity, we choose $|+\rangle_L$ or $|++\rangle_L$ as initial states to extract Z errors, whereas we choose $|0\rangle_L$ or $|00\rangle_L$ to extract X errors. The choice of the above states for error extraction is justified in Appendix E, where we provide the full error channel of the gates.

The extracted gate errors using the Kerr-cat eigenbasis converge to those using the standard Fock basis as d increases. In Fig. 7 we plot the ratio $\frac{P}{P_f}$ for different gates as a function of d, where P is the total Z or X error probability obtained using the Kerr-cat eigenbasis and P_f is the corresponding error probability obtained using the Fock basis. Based on Fig. 7, the values of d we choose to simulate Z rotation, ZZ rotation, and the CX gate are 5, 6, and 4, respectively

In Fig. 8 we show the numerically extracted nonadiabatic errors of all three types of gates considered in this paper, including the *ZZ* rotation, which is not covered in the main text. In Fig. 9 we plot $P_z^{\rm NA}$ versus $1/K(\kappa_2)T$ in a log-log scale for gates with dissipative (black dashed curves) and hard control (blue solid curves), which highlights that $P_z^{\rm NA} \propto 1/\kappa_2 T$ for dissipative gates while $P_z^{\rm NA} \propto 1/(KT)^2$ for Kerr gates with hard control, which supports the scaling analysis in the main text. In Table III we show the numerically fitted scalings of the minimal Z error probability of the CX gates with photon loss rate κ_1, which are in good agreement with the analysis in the main text.

APPENDIX E: ERROR CHANNEL OF THE CX GATE

Here, we present the numerically extracted error channel of the Kerr CX gate in the presence of photon loss. In Figs. 10(a) and 10(b) we plot the real part and imaginary part of the error matrix $\chi^{\rm err}$ obtained from numerical process tomography, which describes the CPTP map:

$$\mathcal{E}(\rho) = \sum_{mn} \chi_{mn}^{\rm err} P_m \tilde{\rho} P_n^\dagger, \tag{E1}$$

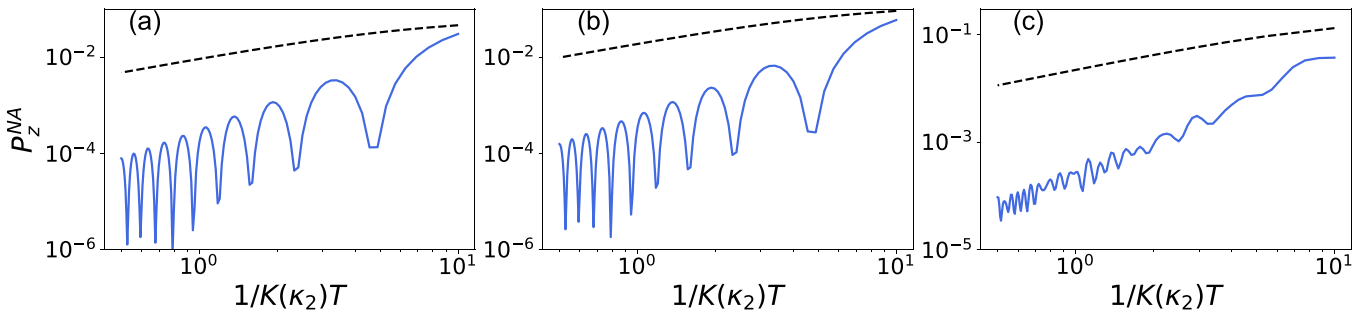

FIG. 9. Plots of $P_z^{\rm NA}$ vs $1/K(\kappa_2)T$ in a log-log scale for gates with dissipative (black dashed curves) and hard control (blue solid curves), which highlight that $P_z^{\rm NA} \propto 1/\kappa_2 T$ for dissipative gates while $P_z^{\rm NA} \propto 1/(KT)^2$ for Kerr gates with hard control. (a) *Z* rotation. (b) *ZZ* rotation. (c) CX gate.

TABLE III. The numerically fitted κ_1 dependence of the minimal Z error probability P_z^* of CX gates with different control schemes.

Control scheme	Dissipative	Hard	Gaussian	DBC
$P_z^* \propto (\frac{\kappa_1}{K(\kappa_2)})^p$	$p = 0.49$	$p = 0.68$	$p = 0.81$	$p = 0.84$

where $\tilde{\rho} = CX\rho CX$ is the image of ρ by an ideal CX gate and P_n ($n = 0, 1, 2, 3$) denotes the Pauli operators. The parameters used for the presented numeric result are $\kappa_1/K = 5 \times 10^{-5}$ and $KT = 1$. The dominant errors are Z-type errors Z_C, Z_t, and Z_cZ_t, while all non-Z-type errors are much smaller. We define the total Z error probability p_Z^{tomo} as the sum of diagonal coefficients of χ^{err} corresponding to Z_c, Z_t, and Z_cZ_t, whereas we define the total X error probability p_X^{tomo} as the sum of the rest of the diagonal error coefficients. We compare p_Z^{tomo} (p_X^{tomo}) with the total error probability p_Z (p_X) extracted by calculating the target state infidelity with initial states $|++\rangle_L$ ($|00\rangle_L$): $p_Z^{\text{tomo}} = 0.0008$, $p_Z = 0.0008$; $p_X^{\text{tomo}} = 1.71 \times 10^{-8}$, $p_X = 1.41 \times 10^{-8}$. So the gate errors extracted using the chosen initial states well capture the error channel of the CX gate, and the numerical results presented in this paper are well justified.

APPENDIX F: DETAILS OF THE CONCATENATED QEC

In this Appendix we provide the details of the concatenated QEC considered in this paper, including the detailed error model and how we obtain the logical Z error rate and the final optimal logical gate error rate.

Using our model defined in Table I, the logical Z error probability P_Z^L of the logical CX gate will in general be a function of the photon loss rate κ_1, the repetition-code distance d, and the physical CX gate time T_{CX}. In the following we will show how we estimate P_Z^L in the low-loss regime while avoiding large Monte Carlo (MC) simulations.

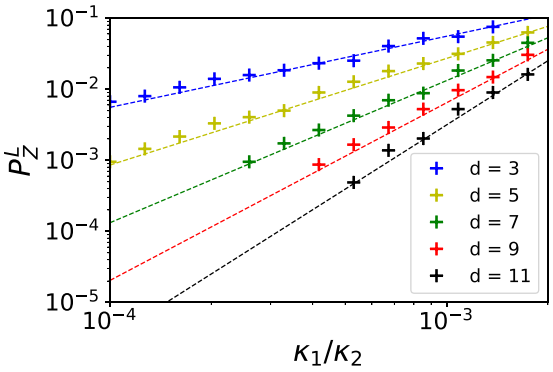

FIG. 11. Logical Z error probability of the logical CX gate encoded in a repetition cat using dissipative CX gates. The pluses are the MC simulation results, while the dashed lines are the fitting results $P_Z^L = A[B\frac{\kappa_1}{\kappa_2}]^{\frac{d+1}{4}}$. For this plot, η_T is set to 1, and the fitted coefficients are $A = 0.16$, $B = 352.1$.

We define a dimensionless parameter $\eta \equiv \frac{T_{\text{CX}}}{T^*}$ as the ratio between the chosen physical CX gate time T_{CX} and the gate time T^* that maximizes the physical CX gate fidelity.

Using the dissipative CX gate, $T^* = \frac{\pi}{8\alpha^2\sqrt{2\kappa_1\kappa_2}}$ [18] and both the nonadiabatic error probability P_z^{NA} and the characteristic Z error probability p_0 in Table III will be proportional to $\sqrt{\frac{\kappa_1}{\kappa_2}}$ for a given η. Therefore we expect an empirical scaling of P_Z^L:

$$P_Z^L(\kappa_1, \eta, d) = A(\eta)\left[B(\eta)\frac{\kappa_1}{\kappa_2}\right]^{\frac{d+1}{4}}. \tag{F1}$$

This is numerically verified using the MC simulation, and the coefficients $A(\eta)$ and $B(\eta)$ are fitted. As an example, the numerically obtained and fitted P_Z^L for $\eta = 1$ is shown in Fig. 11. The fitted coefficients $A(\eta)$ and $B(\eta)$ are shown in Figs. 12(a) and 12(b). By interpolating $A(\eta)$ and $B(\eta)$ using the numerically fitted data in Figs. 12(a) and 12(b) we can

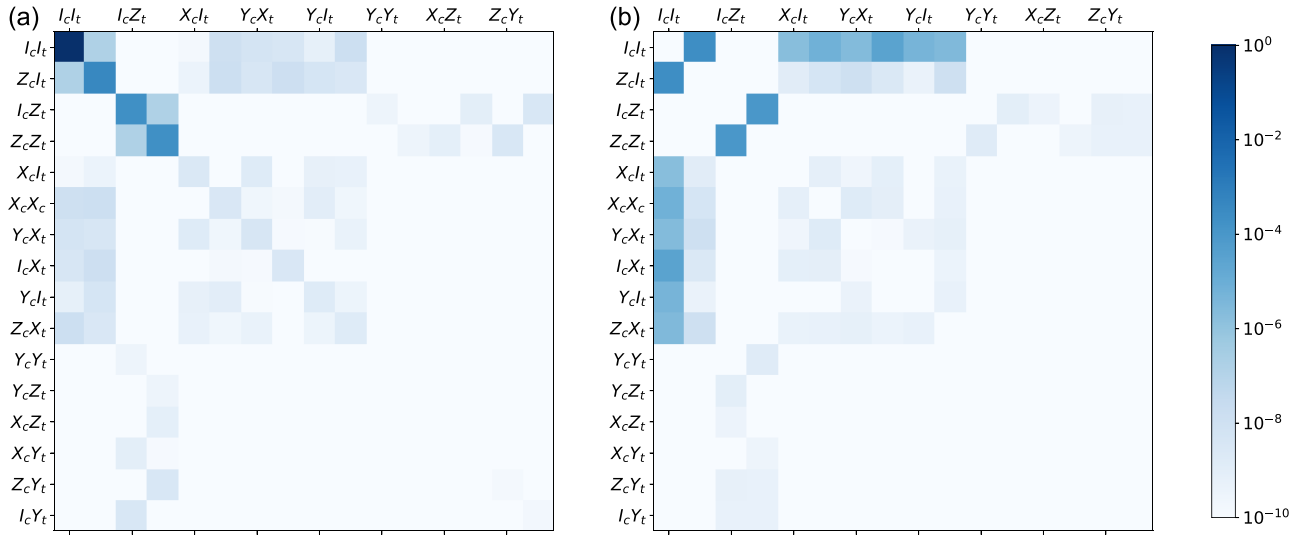

FIG. 10. Numerically extracted error channel of the CX gate. The real part $|\text{Re}(\chi^{\text{err}})|$ (a) and the imaginary part $|\text{Im}(\chi^{\text{err}})|$ (b) of the error process matrix are plotted separately.

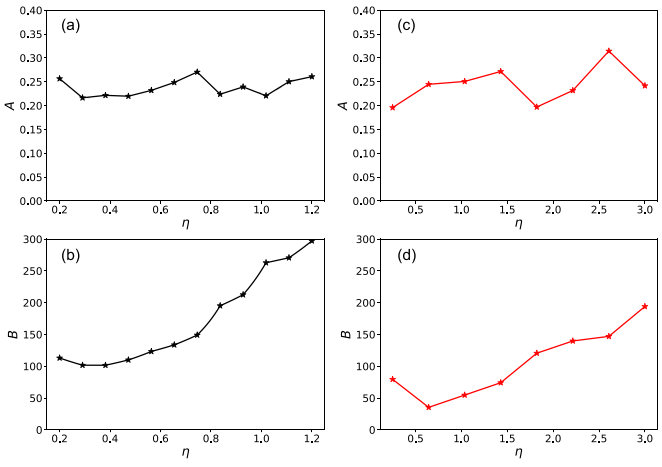

FIG. 12. (a) and (b) The fitted coefficients A and B in Eq. (F1) as functions of η. (c) and (d) The fitted coefficients A and B in Eq. (F2) as functions of η.

then use Eq. (F1) to estimate the logical Z error probability P_Z^L when choosing different dissipative CX gate times.

Using the physical CX gate on the Kerr cat with hard control, $T^* = (\frac{\pi^2}{512\alpha^6\kappa_1 K^2})^{1/3}$ and both the nonadiabatic error probability P_z^{NA} and the characteristic Z error probability p_0 in Table III will be proportional to $(\frac{\kappa_1}{K})^{\frac{2}{3}}$ for a given η. So, similarly, we can fit the logical Z error probability as

$$P_Z^L(\kappa_1, \eta, d) = A(\eta)\left[B(\eta)\frac{\kappa_1}{K}\right]^{\frac{d+1}{3}}. \tag{F2}$$

The fitted coefficients $A(\eta)$ and $B(\eta)$ are shown in Figs. 12(c) and 12(d).

When using the physical CX gate on the Kerr cat with Gaussian or DBC control, we notice that for the range of gate times of interest, the nonadiabatic Z error probability P_z^{NA} is negligible compared with the characteristic Z error probability p_0. So P_Z^L is only a function of p_0 and the code distance d, which can be numerically fitted as

$$P_z^L(\kappa_1, T_{\mathrm{CX}}, d) = P_z^L(p_0(\kappa_1, T_{cx}), d) \approx 0.086 * (454p_0)^{\frac{d+1}{2}}, \tag{F3}$$

where $p_0 = \kappa_1\alpha^2 T_{\mathrm{CX}}$.

In the end, we can obtain the minimal total logical error probability of the logical CX gate for given κ_1 by optimizing over d and T_{CX}:

$$\begin{aligned} P_L^*(\kappa_1) &= \min_{T_{\mathrm{CX}},d} P^L(\kappa_1, T_{\mathrm{CX}}, d) \\ &= \min_{T_{\mathrm{CX}},d} P_Z^L(\kappa_1, T_{\mathrm{CX}}, d) + P_X^L(\kappa_1, T_{\mathrm{CX}}, d), \end{aligned} \tag{F4}$$

where $P_Z^L(\kappa_1, T_{\mathrm{CX}}, d)$ is calculated using Eqs. (F1)–(F3).

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

[2] Z. Leghtas, G. Kirchmair, B. Vlastakis, R. Schoelkopf, M. Devoret, and M. Mirrahimi, Hardware-Efficient Autonomous Quantum Error Correction, Phys. Rev. Lett. **111**, 120501 (2013).

[3] 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).

[4] V. V. Albert, K. Noh, K. Duivenvoorden, D. J. Young, R. T. Brierley, P. Reinhold, C. Vuillot, L. Li, C. Shen, S. M. Girvin, B. M. Terhal, and L. Jiang, Performance and structure of single-mode bosonic codes, Phys. Rev. A **97**, 032346 (2018).

[5] 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).

[6] 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).

[7] 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).

[8] A. Grimm, N. E. Frattini, S. Puri, S. O. Mundhada, S. Touzard, M. Mirrahimi, S. M. Girvin, S. Shankar, and M. H. Devoret, Stabilization and operation of a Kerr-cat qubit, Nature (London) **584**, 205 (2020).

[9] P. Campagne-Ibarcq, A. Eickbusch, S. Touzard, E. Zalys-Geller, N. E. Frattini, V. V. Sivak, P. Reinhold, S. Puri, S. Shankar, R. J. Schoelkopf, L. Frunzio, M. Mirrahimi, and M. H. Devoret, Quantum error correction of a qubit encoded in grid states of an oscillator, Nature (London) **584**, 368 (2020).

[10] C. Flühmann, T. L. Nguyen, M. Marinelli, V. Negnevitsky, K. Mehta, and J. P. Home, Encoding a qubit in a trapped-ion mechanical oscillator, Nature (London) **566**, 513 (2019).

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

[12] D. K. Tuckett, S. D. Bartlett, and S. T. Flammia, Ultrahigh Error Threshold for Surface Codes with Biased Noise, Phys. Rev. Lett. **120**, 050505 (2018).

[13] J. P. Bonilla Ataides, D. K. Tuckett, S. D. Bartlett, S. T. Flammia, and B. J. Brown, The XZZX surface code, Nat. Commun. **12**, 2172 (2021).

[14] S. Puri, S. Boutin, and A. Blais, Engineering the quantum states of light in a Kerr-nonlinear resonator by two-photon driving, npj Quantum Inf. **3**, 18 (2017).

[15] S. Puri, L. St-Jean, J. A. Gross, A. Grimm, N. E. Frattini, P. S. Iyer, A. Krishna, S. Touzard, L. Jiang, A. Blais, S. T. Flammia, and S. M. Girvin, Bias-preserving gates with stabilized cat qubits, Sci. Adv. **6**, eaay5901 (2020).

[16] J. Guillaud and M. Mirrahimi, Repetition Cat Qubits for Fault-Tolerant Quantum Computation, Phys. Rev. X **9**, 041053 (2019).

[17] J. Guillaud and M. Mirrahimi, Error rates and resource overheads of repetition cat qubits, Phys. Rev. A **103**, 042413 (2021).

[18] C. Chamberland, K. Noh, P. Arrangoiz-Arriola, E. T. Campbell, C. T. Hann, J. Iverson, H. Putterman, T. C. Bohdanowicz, S. T. Flammia, A. Keller, G. Refael, J. Preskill, L. Jiang, A. H. Safavi-Naeini, O. Painter, and F. G. S. L. Brandão, Building a fault-tolerant quantum computer using concatenated cat codes, arXiv:2012.04108 [quant-ph].

[19] A. S. Darmawan, B. J. Brown, A. L. Grimsmo, D. K. Tuckett, and S. Puri, Practical quantum error correction with the ZXXZ code and Kerr-Cat qubits, PRX Quantum **2**, 030345 (2021).

[20] H. Goto, Universal quantum computation with a nonlinear oscillator network, Phys. Rev. A **93**, 050301(R) (2016).

[21] Despite the fact that the energy gap $\Delta_1 < 0$ is negative in the rotating frame, the excitation energy $\omega_0 + \Delta_1 > 0$ is still positive in the laboratory frame, with oscillator frequency ω_0.

[22] E. R. Pinch, *Optimal Control and the Calculus of Variations* (Oxford University Press, Oxford, 1995).

[23] F. Motzoi and F. K. Wilhelm, Improving frequency selection of driven pulses using derivative-based transition suppression, Phys. Rev. A **88**, 062318 (2013).

[24] D. Guéry-Odelin, A. Ruschhaupt, A. Kiely, E. Torrontegui, S. Martínez-Garaot, and J. G. Muga, Shortcuts to adiabaticity: Concepts, methods, and applications, Rev. Mod. Phys. **91**, 045001 (2019).

[25] L. S. Theis, F. Motzoi, S. Machnes, and F. K. Wilhelm, Counteracting systems of diabaticities using DRAG controls: The status after 10 years, EPL (Europhys. Lett.) **123**, 60001 (2018).

[26] F. Motzoi, J. M. Gambetta, P. Rebentrost, and F. K. Wilhelm, Simple Pulses for Elimination of Leakage in Weakly Nonlinear Qubits, Phys. Rev. Lett. **103**, 110501 (2009).

[27] H. Ribeiro, A. Baksic, and A. A. Clerk, Systematic Magnus-Based Approach for Suppressing Leakage and Nonadiabatic Errors in Quantum Dynamics, Phys. Rev. X **7**, 011021 (2017).

[28] S. Touzard, A. Grimm, Z. Leghtas, S. O. Mundhada, P. Reinhold, C. Axline, M. Reagor, K. Chou, J. Blumoff, K. M. Sliwa, S. Shankar, L. Frunzio, R. J. Schoelkopf, M. Mirrahimi, and M. H. Devoret, Coherent Oscillations inside a Quantum Manifold Stabilized by Dissipation, Phys. Rev. X **8**, 021005 (2018)..

[29] 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).