Efficient Quantum Error Correction of Dephasing Induced by a Common Fluctuator

David Layden[⬤],[1,2] Mo Chen (陈墨),[1,3] and Paola Cappellaro[1,2]

[1]*Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
[2]*Department of Nuclear Science and Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
[3]*Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*

Quantum error correction is expected to be essential in large-scale quantum technologies. However, the substantial overhead of qubits it requires is thought to greatly limit its utility in smaller, near-term devices. Here we introduce a new family of special-purpose quantum error-correcting codes that offer an exponential reduction in overhead compared to the usual repetition code. They are tailored for a common and important source of decoherence in current experiments, whereby a register of qubits is subject to phase noise through coupling to a common fluctuator, such as a resonator or a spin defect. The smallest instance encodes one logical qubit into two physical qubits, and corrects decoherence to leading-order using a constant number of one- and two-qubit operations. More generally, while the repetition code on n qubits corrects errors to order $t^{O(n)}$, with t the time between recoveries, our codes correct to order $t^{O(2^n)}$. Moreover, they are robust to model imperfections in small- and intermediate-scale devices, where they already provide substantial gains in error suppression. As a result, these hardware-efficient codes open a potential avenue for useful quantum error correction in near-term, pre-fault tolerant devices.

Decoherence, the uncontrolled decay of coherence in open quantum systems, is a central obstacle to developing coherent quantum technologies such as quantum sensors, networks, and computers. This obstacle is compounded by the destructive nature of quantum measurement: straightforward attempts to identify—and ultimately reverse—decoherence destroy the quantum coherence they seek to protect. Quantum error correction (QEC) is a technique for taming decoherence, which sidesteps this issue. It encodes lower-dimensional quantum states into a higher-dimensional quantum system such that decoherence can be detected and approximately reversed without collapsing the encoded state. Specifically, the most common approach encodes k logical qubits into an n-qubit register ($k < n$) whose Hilbert space $\mathcal{H}$ is decomposed into orthogonal subspaces $\mathcal{C}_0, \mathcal{C}_1, \mathcal{C}_2, \ldots$ of dimension 2^k [1]. These subspaces are chosen by specifying operators $E_1, E_2, \ldots$ and demanding that the logical states, which reside in $\mathcal{C}_0$, be mapped to $\mathcal{C}_i$ by E_i without distortion [2]. By performing a partial measurement that reveals *only* which subspace contains the state, and feeding back appropriately, one can reverse the occurrence of any E_i—and, more generally, any error in $\mathcal{E} = \mathrm{span}\{I, E_1, E_2, \ldots\}$. The conventional strategy is to pick E_is so that $\mathcal{E}$ encompasses a broad family of operators on $\mathcal{H}$. Using Pauli operators of weight up to w, for instance, produces a QEC code that corrects arbitrary errors on w qubits. This is a powerful approach, especially in large devices ($n \gg 1$), since it can reverse decoherence with little regard to its physical origins [3,4]. For smaller devices, however, casting such a wide net

requires an overhead of qubits ($n - k$) that is often prohibitive for near-term applications. A more economical strategy for small- and intermediate-scale devices is instead to use a QEC code with $\mathcal{E}$ tailored to include only the dominant, well-characterized decoherence modes. However, while this strategy is well-known (see [3] § 10.6.4), few explicit such codes have been discovered (see, e.g., Refs. [5–7]).

In order to systematically find noise-tailored QEC codes, here we focus on dephasing, since it is the dominant type of decoherence in various experiments. In particular, we consider the common scenario where dephasing in a register of qubits arises primarily due to eigenstate-preserving coupling of each qubit to a common fluctuator, which in turn exchanges energy with an external environment. That is, we consider a Hamiltonian

$$H = H_f^0 + \frac{1}{2}\sum_{j=1}^{n} \omega_j Z_j + H_f^{\mathrm{int}} \otimes \sum_{j=1}^{n} g_j Z_j, \qquad (1)$$

where $[H_f^0, H_f^{\mathrm{int}}] = 0$, and a fluctuator that jumps incoherently between energy eigenstates $\{|\ell\rangle_f\}$ (reflected by a dissipative term in the overall master equation). Moving to the interaction picture, the Hamiltonian (1) becomes

$$\tilde{H} = \sum_{\ell} \lambda_\ell |\ell\rangle\langle\ell|_f \otimes H_E, \qquad (2)$$

where $H_f^{\mathrm{int}} = \sum_\ell \lambda_\ell |\ell\rangle\langle\ell|_f$ and $H_E \coloneqq \sum_{j=1}^{n} g_j Z_j$. When the fluctuator is in state $|\ell\rangle_f$, qubit j has an effective

Hamiltonian $\lambda_\ell g_j Z_j$ in the rotating frame. Jumps of the fluctuator therefore induce spatially correlated random telegraph noise in the register, which causes dephasing [8,9]. This model, which we call *common-fluctuator dephasing* (CFD), often describes the main decoherence mechanism in nuclear spins near spin defects (e.g., nitrogen-vacancy centers in diamond [10]) or quantum dots, and can also be significant in superconducting qubits dispersively coupled to a common resonator with nonzero effective temperature [10–24]. Often the register is read out and/or initialized via the fluctuator, imposing a lower limit on the desirable coupling strengths g_j, and making CFD a significant decoherence mode. Note that CFD does not generally produce a decoherence-free subspace (DFS).

The standard QEC approach to correct dephasing uses E_is comprising Pauli Z operators on at most w qubits (and I on the rest). There are $\sum_{m=0}^{w} \binom{n}{m}$ such matrices; a simple counting argument (the quantum Hamming bound applied to phase noise) therefore suggests that $n \geq 2w + 1$ physical qubits are required to protect $k = 1$ logical qubit from arbitrary phase errors of weight $\leq w$ [3]. Indeed, the repetition code saturates this bound: the smallest instance uses $n = 3$ for $w = 1$, has logical states $|0_L\rangle = |+++\rangle$ and $|1_L\rangle = |---\rangle$ where $|\pm\rangle := (1/\sqrt{2})(|0\rangle \pm |1\rangle)$, and corrects for $\mathcal{E} = \text{span}\{I, Z_1, Z_2, Z_3\}$. It can correct CFD as follows: in any run of the experiment, the register evolves over time t as $U(\theta) = e^{-i\theta H_E}$ for some random variable $\theta \in [t\lambda_{\min}, t\lambda_{\max}]$ that depends on the fluctuator's trajectory. For short t (understood in units of $1/\max_{j\ell}|g_j\lambda_\ell|$, and often reducible through dynamical decoupling [10,25–27]), $U(\theta)$ can be approximated as $U(\theta) = I - i\theta H_E + O(t^2)$. Since $\theta H_E \in \mathcal{E}$ regardless of θ, this three-qubit code corrects dephasing at order $O(t)$. More generally, H_E^q contains Paulis of weight $\leq q$, so correcting to order $O(t^q)$ with the repetition code requires $n = 2q + 1$ qubits (for $k = 1$).

While the value of θ is unknown and varies from one run to the next, the coupling strengths g_j are often fixed and well characterized. This suggests designing a code that corrects expressly for $\mathcal{E} = \text{span}\{I, H_E, H_E^2, \ldots, H_E^q\}$, and depends on the $\{g_j\}$ in a particular device. A similar counting argument as above suggests that such a code would require $q + 1$ subspaces to protect a logical qubit to order $O(t^q)$, and would therefore require

$$n = \lceil 1 + \log_2(q + 1) \rceil \tag{3}$$

qubits—an exponentially smaller overhead. We give a family of such codes here for general q and arbitrary coupling strengths $\{g_j\}$. We focus in particular on the $q = 1$ case, where one logical qubit is encoded in two physical qubits rather than three. We construct recovery and logical operations for this code, which can be implemented using a constant number of one- and two-qubit operations.

The decomposition $\mathcal{H}$ into subspaces $\mathcal{C}_i$ for QEC is equivalent to the Knill-Laflamme conditions [28,29]. For $k = 1$ and $\mathcal{E} = \text{span}\{H_E^j\}_{j=0}^q$, these take the form

$$\langle 0_L | H_E^m | 0_L \rangle = \langle 1_L | H_E^m | 1_L \rangle, \tag{4}$$

$$\langle 0_L | H_E^m | 1_L \rangle = 0, \tag{5}$$

for $0 \leq m \leq 2q$, where we consider values of q that saturate the ceiling in Eq. (3) (that is, $q = 2^{n-1} - 1$). Finding a QEC code that corrects this $\mathcal{E}$ therefore requires finding logical states $|0_L\rangle$ and $|1_L\rangle$ that satisfy Eqs. (4) and (5). We begin with the ansatz

$$|0_L\rangle = \sum_{j=0}^{2^n-1} r_j e^{i\theta_j} |j\rangle \quad |1_L\rangle = \sum_{j=0}^{2^n-1} r_{(2^n-1-j)} e^{i\phi_j} |j\rangle, \tag{6}$$

for $r_j, \theta_j, \phi_j \in \mathbb{R}$, where we use $|j\rangle$ to denote the n-bit binary representation of the integer j. That is, we fix the amplitudes of $|1_L\rangle$ to be those of $|0_L\rangle$ in reverse order. Notice that Eq. (6) always satisfies (4) for even $m \geq 0$, since $X^{\otimes n} H_E^m X^{\otimes n} = (-1)^m H_E^m$. For odd m:

$$\langle 0_L | H_E^m | 0_L \rangle = -\langle 1_L | H_E^m | 1_L \rangle = \vec{z} \cdot \vec{v}_m, \tag{7}$$

where $\vec{z}, \vec{v}_m \in \mathbb{R}^{q+1}$ are defined as $z_i = \langle i | Z_L | i \rangle$, with $Z_L := |0_L\rangle\langle 0_L| - |1_L\rangle\langle 1_L|$, and $(\vec{v}_m)_i = \langle i | H_E^m | i \rangle$ for $i \in [0, q]$ and odd $m \in [0, 2q]$. Therefore, Eq. (4) is satisfied for all relevant m if $\vec{z} \perp \text{span}\{\vec{v}_m\}$. We can always find such a $\vec{z} (\neq \vec{0})$ since the $\vec{v}_m$s have dimension $q + 1$ but there are only q of them, so they cannot form a complete basis. One approach is to construct a matrix V with $\vec{v}_m$s as columns; then, $I - VV^+$ projects onto $\text{span}\{\vec{v}_m\}^\perp$ (where $+$ and $\perp$ denote the pseudoinverse and orthogonal complement, respectively) and therefore has at least one real eigenvector $\vec{u}$ with unit eigenvalue [30]. Taking $\vec{z} = \vec{u}/\|\vec{u}\|_1$ satisfies Eq. (4) since $\vec{u} \cdot \vec{v}_m = 0$ automatically. Finally, building upon a technique developed in Ref. [7] for optimization, we pick r_js as

$$(r_j, r_{(2^n-1-j)}) = \begin{cases} (0, \sqrt{z_j}), & \text{if } z_j \geq 0, \\ (\sqrt{-z_j}, 0), & \text{if } z_j < 0. \end{cases} \tag{8}$$

This choice ensures that $\langle j | 0_L \rangle$ or $\langle j | 1_L \rangle$ vanishes for every j, thus satisfying Eq. (5). We now have normalized logical states that form a valid QEC code for all $q \geq 1$. Notice that the components of $|0_L\rangle$ and $|1_L\rangle$ generically have unequal amplitudes r_j by necessity, in marked contrast with classical error-correcting codes and most known QEC codes. The phases θ_j and ϕ_j can be chosen arbitrarily— we demonstrate a convenient choice below. The performance of these codes on $n \leq 5$ qubits is shown in Fig. 1 using an illustrative model of a normally distributed θ. In
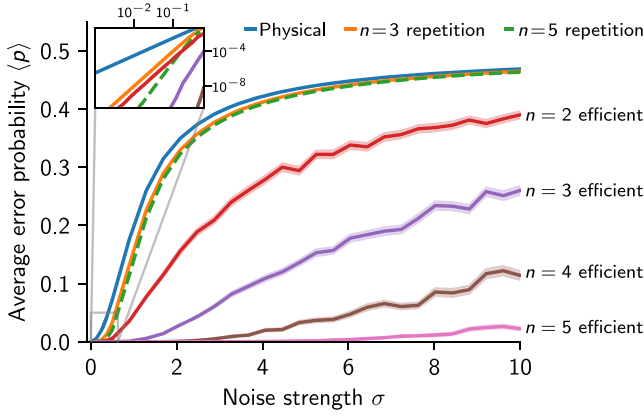
FIG. 1. Comparison of QEC codes performance. We assume that the effect of the quantum fluctuator is to impart a random phase, θ, which follows a Gaussian distribution $\theta \sim \mathcal{N}(0, \sigma)$ with standard deviation σ. By normalizing the g_js to lie in $[0, 1]^n$, σ describes the noise strength. CFD followed by a QEC recovery (if applicable) results in an effective phase- or bit-flip channel $\rho \mapsto (1 - p)\rho + pA\rho A$, where $A = Z$ for the physical qubits, X_L for the repetition codes, and Z_L for hardware-efficient codes. The average infidelity, average trace distance and diamond distance to I are all $\propto p$. As the performance of all strategies shown depends on $\{g_j\}$, we plot the average of p over $\{g_j\} \in [0, 1]^n$. The error bands for the hardware-efficient codes denote the standard error of the mean from Monte Carlo integration. More details on the numerical implementation are given in [31].

addition, we give the pseudothresholds for $n = 2$ and 3 under the same model in the Supplemental Material [31].

To illustrate this QEC code, we consider explicitly the smallest case of $n = 2$ qubits coupled to a two-level fluctuator with $\lambda_{\pm 1} = \pm 1$ [cf. Eq. (2)], at high temperature. We will label the register qubits 1 and 2 such that $|g_1| \geq |g_2|$. Note that here—and in general—$H_E = g_1 Z_1 + g_2 Z_2$ is a combination of weight-1 Pauli operators, not a weight-2 Pauli. This H_E gives $\vec{v}_1 = (g_1 + g_2, g_1 - g_2)^\top$. The matrix $I - VV^+$ has only a one-dimensional eigenspace with unit eigenvalue, spanned by $\vec{u} = (-g_1 + g_2, g_1 + g_2)^\top$, where $\vec{u} \cdot \vec{v}_1 = 0$. If $g_1 > 0$ we find $r_1 = r_3 = 0$ and

$$r_0 = c\sqrt{g_1 - g_2}, \qquad r_2 = c\sqrt{g_1 + g_2}, \qquad (9)$$

where $c = 1/\sqrt{||\vec{u}||_1}$. This gives logical states

$$|0_L\rangle = |\chi_0\rangle|0\rangle, \qquad |1_L\rangle = |\chi_1\rangle|1\rangle, \qquad (10)$$

with

$$|\chi_0\rangle = c(\sqrt{|g_1 - g_2|}e^{i\theta_0}|0\rangle + \sqrt{|g_1 + g_2|}e^{i\theta_2}|1\rangle),$$
$$|\chi_1\rangle = c(\sqrt{|g_1 + g_2|}e^{i\phi_1}|0\rangle + \sqrt{|g_1 - g_2|}e^{i\phi_3}|1\rangle), \qquad (11)$$

where $|0\rangle$ and $|1\rangle$ refer to the states of a qubit. The $g_1 < 0$ case gives the same result up to a relabelling of $|0_L\rangle \leftrightarrow |1_L\rangle$. This code corrects for $\mathcal{E} = \text{span}\{I, H_E\}$; by design, however, it does not correct for $Z_1 Z_2$, nor Z_1 or Z_2 individually, none of which belong to $\mathcal{E}$. Rather, it corrects CFD with fewer qubits than the smallest repetition code precisely because we have chosen not to correct individual Pauli operators.

Observe that Eqs. (10) and (11) reduce to a DFS in the limit where one exists ($|g_1| = |g_2|$), but this is in practice rare. More generally, notice that the choice $\theta_0 = \phi_1 + \pi = -\theta_2 = -\phi_3 = \vartheta$ for arbitrary ϑ proves convenient: first, it gives $\langle\chi_0|\chi_1\rangle = 0$, and a simple action of H_E on logical states:

$$H_E|0_L\rangle \propto |\chi_1\rangle|0\rangle =: |0_E\rangle$$
$$H_E|1_L\rangle \propto |\chi_0\rangle|1\rangle =: |1_E\rangle. \qquad (12)$$

Both lines have the same proportionality constant, and we have defined the error states $|0_E\rangle$ and $|1_E\rangle$. We emphasize that since H_E cannot generically be decomposed as a tensor product, it maps most separable states to entangled states; Eq. (12)—wherein the first qubit is "flipped" by H_E—is due to our choice of $|0_L\rangle$ and $|1_L\rangle$. Second, consider the orthogonal projectors $P_L = |0_L\rangle\langle 0_L| + |1_L\rangle\langle 1_L|$ and $P_E = |0_E\rangle\langle 0_E| + |1_E\rangle\langle 1_E|$ onto $\mathcal{C}_0 = \text{span}\{|0_L\rangle, |1_L\rangle\}$ and $\mathcal{C}_1 = \text{span}\{|0_E\rangle, |1_E\rangle\}$ respectively ($\mathcal{H} = \mathcal{C}_0 \oplus \mathcal{C}_1$). One can detect an error nondestructively by measuring parity in the $|\chi_i\rangle|j\rangle$ basis, which can be done by performing phase estimation (i.e., "phase kickback") on

$$S = P_L - P_E = U_z \otimes Z_2 \qquad (13)$$

with an ancilla [33]. Crucially, the choice of phases in $|0_L\rangle$ and $|1_L\rangle$ makes S separable here, where $U_z := |\chi_0\rangle\langle\chi_0| - |\chi_1\rangle\langle\chi_1|$ is a π rotation about some axis determined by g_1, g_2, and ϑ. This means that the controlled-S (cS) operation used to measure the error syndrome can be implemented through a pair of two-qubit operations (cU_z and cZ), rather than a more challenging three-qubit operation. If an error is detected, it can be corrected by applying $U_x := |\chi_0\rangle\langle\chi_1| + |\chi_1\rangle\langle\chi_0|$ to qubit 1—a π rotation about a different axis. (Both U_x and U_z could be synthesized out of a constant number of Pauli rotations, or implemented directly, e.g., by driving qubit 1 off resonance [34].) The full recovery procedure, which corrects CFD to leading order, is shown in Fig. 2. Note that S behaves like a stabilizer, in the sense of its action on $\mathcal{C}_0$ and $\mathcal{C}_1$. It does not, however, fit in the usual QEC stabilizer formalism since $\{H_E, S\} \neq 0$ generically, because $\{H_E, S\}|\psi\rangle = 0$ for $|\psi\rangle \in \mathcal{C}_0$ but not for $|\psi\rangle \in \mathcal{C}_1$ [35]. This is because H_E maps $\mathcal{C}_0$ to $\mathcal{C}_1$ without distortion, but not vice versa, as H_E is not generically in the Pauli group. (Neither is S.) In spite of these unusual features, the procedure for feeding back on S in Fig. 2 is largely the same as that of the usual stabilizer formalism.
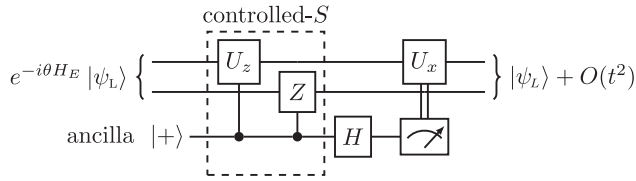
FIG. 2. A recovery procedure for $n = 2$ qubits where $|\psi_L\rangle = \alpha|0_L\rangle + \beta|1_L\rangle$ for arbitrary α and β, H denotes a Hadamard gate, and θ is a random variable. The unitaries U_x and U_z are both π rotations about orthogonal axes on the Bloch sphere which are determined by g_1, g_2, and ϑ.

Finally, (i) the encoding can be realized by applying a $c_2(U_x)_1$ gate to an initial state $|\chi_0\rangle|\psi\rangle$, and (ii) there is a simple way to implement any logical unitary U_L in this code: apply the corresponding physical U to qubit 2 followed by a recovery.

The logical states derived above are also valid for all $q > 1$ (i.e., $n > 2$ qubits), but the corresponding recovery and logical operations are generally more involved. Generically, the analogs of S in (13) are not separable for any choice of θ_j and ϕ_j [36]. One might still synthesize them with one- and two-qubit operations, perform phase kickback through optimal control, or implement a QEC recovery via more general channel-engineering techniques [37–40]. More efficient solutions could even be found by analyzing specific experimental scenarios. One approach could be for example to use devices with $\{g_j\}$ chosen so that the recovery and logical operations can be conveniently implemented. One could also correct to a slightly lower order q [i.e., maintaining $n = O(\log q)$ but not saturating the ceiling in Eq. (3)]; this would yield a continuous family of possible $\vec{z}$ s [cf. Eq. (8)], among which one might find codes with convenient QEC operations. Note finally that for $n > 2$ it is not the bare H_E^ms that map the codespace to the orthogonal subspaces $\{\mathcal{C}_i\}_{i \geq 1}$, but rather linear combinations of them.

These noise-adapted QEC codes involve a trade-off: they correct CFD very efficiently at the cost of leaving most other errors uncorrected. For instance, errors during gates, due to miscalibration of g_js, or from decoherence beyond CFD will generally affect the logical state [31]. Accordingly, these codes are manifestly not fault-tolerant in their current form [41]. Crucially though, they offer such a large error budget under strong CFD—as evidenced by the gaps between QEC codes and physical qubits in Fig. 1— that this trade-off can easily be worthwhile, much like the targeted correction of photon loss in [42]. Indeed, as we show in [31], the gap survives even in the presence of large miscalibration of the g_js. Fault-tolerance could still be achieved using implementation-specific methods as in Ref. [43]. In the long-term, concatenation could potentially reach fault tolerance, using our noise-adapted codes at the lowest level of encoding to protect against the dominant error source, and more conventional codes at higher levels.

Even more importantly, our codes could have a near-term impact in applications such as quantum sensing and communication, where long-lived quantum memories are useful even when they are not fault tolerant. We emphasize, however, that these codes are designed expressly for small- and medium-scale qubit registers, and that the exponential reduction in overhead should be understood to apply only in such devices. For one, there is typically a maximum n above which CFD no longer dominates. Also, while the error budget always increases with n in principle, so too do the effects of gate errors, miscalibration of g_js and decoherence beyond CFD, as more qubits introduce more error channels. Conversely, this growing sensitivity suggests an unconventional quantum sensing scheme to measure $\{g_j\}$ for large n, by variationally adjusting one's estimates to maximize code performance. In the nearer term, however, these imperfections will likely set a maximum n in any particular device beyond which one achieves no further gains, depending on their relative importance compared to CFD [31].

The QEC codes presented could be generalized in several ways. First, they can readily be made to correct dephasing due to multiple common fluctuators given enough qubits, at the cost of correcting to lower order in t. Similarly, they can correct spatially correlated phase noise beyond that arising from common fluctuators. For instance, classical white noise in the energy gaps of register qubits leads to Lindblad error operators $L_j = \sqrt{\lambda_j}\vec{c}_j \cdot (Z_1, \ldots, Z_n)$, where $\{\sqrt{\lambda_j}\vec{c}_j\}$ describes the noise's normal modes [44]. In the limit of spatially uncorrelated noise the L_js become Pauli Z operators; however, correlated noise produces L_js with unequal amplitudes $\sqrt{\lambda_j}$. When the noise correlations are appreciable, it could be advantageous to use a QEC code that corrects the stronger noise modes (those with large λ_js) to higher order in t than the weaker ones (smaller λ_js) through an appropriate choice of V. It may also be possible to extend the codes presented here for the setting where a fluctuator's state affects not only the energy gap of each qubit, but also the direction of its Hamiltonian (i.e., its quantization axis) [45]. Eigenstate-preserving coupling arises frequently in practice because a large detuning between a weakly-coupled qubit and fluctuator suppresses noncommuting parts of their interaction Hamiltonian. However, when the coupling to the fluctuator is comparable to the internal Hamiltonian, such as for nuclear spins near defects in diamond, there can remain significant noncommuting terms leading to $H_E \sim \sum_j \vec{g}_j \cdot \vec{\sigma}_j$ in Eq. (2). We analyze this effect's impact on code performance in [31]. Extending the codes introduced here to this more general setting would make them even more widely applicable to near-term experiments, but at the cost of larger overheads, since they would need to contend with a substantially larger space of possible errors. It may be more practical instead to suppress noncommuting

interaction terms at the hardware level by increasing the energy gaps ω_j of the register qubits, or at the "software" level through concatenation [31]. Another interesting generalization would be to efficiently encode $k > 1$ logical qubits, which seems plausible based on the counting argument used throughout involving the dimension of $\mathcal{H}$ vs $\mathcal{E}$. Finally, it would be interesting to use the tools presented here to design codes for other common error sources, such as other types of decoherence, or control and/ or measurement errors.

Our results demonstrate that it is possible to find noise-adapted QEC codes with a well-defined advantage (here exponential) over known, general codes. It is commonly argued that QEC will be of little use in noisy intermediate-scale quantum (NISQ) devices due to its prohibitive overhead [46]. Noise-adapted QEC codes are a promising way to reduce this overhead, although to date they have mostly relied on numerical and variational techniques that lack transparency in terms of what advantage the codes can offer, and when [47–51] (see also [4] Ch. 13 and [52]). In contrast, the codes introduced here exhibit a clear reduction in overhead under a well-characterized and common type of noise. New QEC codes of this type could provide a middle ground between small-scale uncorrected devices and large-scale fault-tolerant ones, where the dominant decoherence mechanisms are tamed through specialized codes with only modest overheads. This view of near-term QEC as quantum "firmware" rather than "software" suggests a possible interplay between theory and experiment, whereby NISQ hardware and efficient QEC codes both guide each other's development.

[1] In general, there could also be a "remainder" subspace $\mathcal{C}_R$ of arbitrary dimension so that $\mathcal{H} = (\bigoplus_i \mathcal{C}_i) \oplus \mathcal{C}_R$.

[2] While it is possible for multiple E_is to have the same effect on the logical states, thus reducing the number of subspaces required for QEC, we will not deal with such degenerate codes here.

[3] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, Cambridge, UK, 2010).

[4] D. Lidar and T. Brun, *Quantum Error Correction* (Cambridge University Press, New York, NY, 2013).

[5] D. W. Leung, M. A. Nielsen, I. L. Chuang, and Y. Yamamoto, Phys. Rev. A **56**, 2567 (1997).

[6] A. Robertson, C. Granade, S. D. Bartlett, and S. T. Flammia, Phys. Rev. Applied **8**, 064004 (2017).

[7] D. Layden, S. Zhou, P. Cappellaro, and L. Jiang, Phys. Rev. Lett. **122**, 040502 (2019).

[8] S. Machlup, J. Appl. Phys. **25**, 341 (1954).

[9] C. Neuenhahn, B. Kubala, B. Abel, and F. Marquardt, Phys. Status Solidi B **246**, 1018 (2009).

[10] M. Chen, W. K. C. Sun, K. Saha, J.-C. Jaskula, and P. Cappellaro, New J. Phys. **20**, 063011 (2018).

[11] Such coupling induces an effective $\sigma_+\sigma_- + \sigma_-\sigma_+$ interaction between qubits, which could be suppressed through dynamical decoupling or through large detunings between qubits.

[12] P. C. Maurer, G. Kucsko, C. Latta, L. Jiang, N. Y. Yao, S. D. Bennett, F. Pastawski, D. Hunger, N. Chisholm, M. Markham, D. J. Twitchen, J. I. Cirac, and M. D. Lukin, Science **336**, 1283 (2012).

[13] J. H. Shim, I. Niemeyer, J. Zhang, and D. Suter, Phys. Rev. A **87**, 012301 (2013).

[14] S. Zaiser, T. Rendler, I. Jakobi, T. Wolf, S.-Y. Lee, S. Wagner, V. Bergholm, T. Schulte-Herbrüggen, P. Neumann, and J. Wrachtrup, Nat. Commun. **7**, 12279 (2016).

[15] P. Bertet, I. Chiorescu, G. Burkard, K. Semba, C. J. P. M. Harmans, D. P. DiVincenzo, and J. E. Mooij, Phys. Rev. Lett. **95**, 257002 (2005).

[16] P. Bertet, I. Chiorescu, C. Harmans, and J. Mooij, https://arxiv.org/abs/cond-mat/0507290.

[17] J. Gambetta, A. Blais, D. I. Schuster, A. Wallraff, L. Frunzio, J. Majer, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Phys. Rev. A **74**, 042318 (2006).

[18] J. Majer, J. Chow, J. Gambetta, J. Koch, B. Johnson, J. Schreier, L. Frunzio, D. Schuster, A. Houck, A. Wallraff *et al.*, Nature (London) **449**, 443 (2007).

[19] A. A. Clerk and D. W. Utami, Phys. Rev. A **75**, 042302 (2007).

[20] A. P. Sears, A. Petrenko, G. Catelani, L. Sun, H. Paik, G. Kirchmair, L. Frunzio, L. I. Glazman, S. M. Girvin, and R. J. Schoelkopf, Phys. Rev. B **86**, 180504(R) (2012).

[21] F. Yan, S. Gustavsson, A. Kamal, J. Birenbaum, A. P. Sears, D. Hover, T. J. Gudmundsen, D. Rosenberg, G. Samach, S. Weber *et al.*, Nat. Commun. **7**, 12964 (2016).

[22] J.-H. Yeh, J. LeFebvre, S. Premaratne, F. C. Wellstood, and B. S. Palmer, J. Appl. Phys. **121**, 224501 (2017).

[23] F. Yan, D. Campbell, P. Krantz, M. Kjaergaard, D. Kim, J. L. Yoder, D. Hover, A. Sears, A. J. Kerman, T. P. Orlando, S. Gustavsson, and W. D. Oliver, Phys. Rev. Lett. **120**, 260504 (2018).

[24] Z. Wang, S. Shankar, Z. K. Minev, P. Campagne-Ibarcq, A. Narla, and M. H. Devoret, Phys. Rev. Applied **11**, 014031 (2019).

[25] L. Viola and S. Lloyd, Phys. Rev. A **58**, 2733 (1998).

[26] M. Ban, J. Mod. Opt. **45**, 2315 (1998).

[27] M. J. Biercuk, H. Uys, A. P. VanDevender, N. Shiga, W. M. Itano, and J. J. Bollinger, Nature (London) **458**, 996 (2009).

[28] E. Knill and R. Laflamme, Phys. Rev. A **55**, 900 (1997).

[29] C. Bény, Phys. Rev. Lett. **107**, 080501 (2011).

[30] Alternatively, the modified Gram-Schmidt procedure provides a less intuitive but more numerically stable method.

[31] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.124.020504, which includes Ref. [32], for stability analyses of various model imperfections, and for further details regarding numerics.

[32] A. Dréau, J.-R. Maze, M. Lesik, J.-F. Roch, and V. Jacques, Phys. Rev. B **85,** 134107 (2012).

[33] R. Cleve, A. Ekert, C. Macchiavello, and M. Mosca, Proc. R. Soc. A **454,** 339 (1998).

[34] D. C. McKay, C. J. Wood, S. Sheldon, J. M. Chow, and J. M. Gambetta, Phys. Rev. A **96,** 022330 (2017).

[35] D. Gottesman, https://arxiv.org/abs/quant-ph/9705052.

[36] E.g., $S_1 = P_L + P_{E1} - P_{E2} - P_{E3}$ and $S_2 = P_L - P_{E1} + P_{E2} - P_{E3}$, which could be measured sequentially to identify an error for $n = 3$.

[37] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, and S. J. Glaser, J. Magn. Reson. **172,** 296 (2005).

[38] P. de Fouquieres, S. Schirmer, S. Glaser, and I. Kuprov, J. Magn. Reson. **212,** 412 (2011).

[39] S. Lloyd and L. Viola, Phys. Rev. A **65,** 010101(R) (2001).

[40] C. Shen, K. Noh, V. V. Albert, S. Krastanov, M. H. Devoret, R. J. Schoelkopf, S. M. Girvin, and L. Jiang, Phys. Rev. B **95,** 134501 (2017).

[41] D. Nigg, M. Müller, E. A. Martinez, P. Schindler, M. Hennrich, T. Monz, M. A. Martin-Delgado, and R. Blatt, Science **345,** 302 (2014).

[42] N. Ofek, A. Petrenko, R. Heeres, P. Reinhold, Z. Leghtas, B. Vlastakis, Y. Liu, L. Frunzio, S. Girvin, L. Jiang *et al.*, Nature (London) **536,** 441 (2016).

[43] S. Rosenblum, P. Reinhold, M. Mirrahimi, L. Jiang, L. Frunzio, and R. J. Schoelkopf, Science **361,** 266 (2018).

[44] D. Layden and P. Cappellaro, npj Quantum Inf. **4,** 30 (2018).

[45] C. D. Aiello and P. Cappellaro, Phys. Rev. A **91,** 042340 (2015).

[46] J. Preskill, Quantum **2,** 79 (2018).

[47] M. Reimpell and R. F. Werner, Phys. Rev. Lett. **94,** 080501 (2005).

[48] A. S. Fletcher, P. W. Shor, and M. Z. Win, Phys. Rev. A **75,** 012338 (2007).

[49] R. L. Kosut, A. Shabani, and D. A. Lidar, Phys. Rev. Lett. **100,** 020502 (2008).

[50] S. Taghavi, R. L. Kosut, and D. A. Lidar, IEEE Trans. Inf. Theory **56,** 1461 (2010).

[51] P. D. Johnson, J. Romero, J. Olson, Y. Cao, and A. Aspuru-Guzik, arXiv:1711.02249.

[52] K. Noh, V. V. Albert, and L. Jiang, IEEE Trans. Inf. Theor. **65,** 2563 (2018).

Supplemental Material
Efficient quantum error correction of dephasing induced by a common fluctuator

David Layden,[1,2] Mo Chen (陈墨),[1,3] and Paola Cappellaro[1,2]

[1]*Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
[2]*Department of Nuclear Science and Engineering,*
Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA
[3]*Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*

Appendix A: Probability distribution of θ and effective channel form

We consider in the main text a register of n qubits coupled to a common fluctuator. In the interaction picture, the register evolves by $U(\theta) = e^{-i\theta H_{\mathrm{E}}}$ in any run of the experiment, where θ is a random variable that depends on the fluctuator's dynamics, its coupling to the register, and the elapsed time. The fluctuator can behave differently in each run of the experiment, producing a different value of θ in each realization, and thus causing decoherence in the register.

As far as the register is concerned, the fluctuator's behavior can be fully captured by specifying a (classical) probability distribution for θ. Generically, this distribution will depend on the underlying physics of the fluctuator. For instance, a fluctuator consisting of a Nitrogen-vacancy center electronic spin at room temperature under dynamical decoupling can produce a somewhat different distribution for θ than a fluctuator consisting of a microwave resonator at cryogenic temperature. This variation between different systems precludes a fully general performance analysis not just of the QEC codes introduced here, but also of existing codes under CFD. Instead we consider a simple probability distribution for θ for the sake of illustration, both here and in Fig. 1 of the main text.

In Fig. 1 we choose $\theta \sim \mathcal{N}(0, \sigma)$ to be normally distributed with zero mean and standard deviation σ. (A non-zero mean of θ could always be absorbed into the rotating frame to give the aforementioned distribution.) This choice has two main motivations:

1. It depends only on a single parameter σ that can straightforwardly be interpreted as the noise strength, since we normalize the coupling strengths to lie in $[0, 1]$.

2. It induces a generally monotonic loss of coherence in the register with the noise strength σ. In contrast, more complicated distributions, such as those arising from random telegraph noise in certain regimes, can induce an oscillating loss of coherence. Such oscillations would becloud some of the analyses below. (For instance, there is some ambiguity in defining pseudothresholds when the physical and logical infidelities, or the like, display small oscillations and intersect several times.)

For the sake of comparison, however, we plot in Fig. A1 quantities analogous to those in Fig. 1 of the main text, but with a different underlying distribution of θ. Notice that the curves are deformed slightly from those in Fig. 1, but are qualitatively similar.

Returning to $\theta \sim \mathcal{N}(0, \sigma)$, we now show that the effective logical channel describing CFD followed by a recovery has the form $\rho_{\mathrm{L}} \mapsto (1-p)\rho_{\mathrm{L}} + pZ_{\mathrm{L}}\rho_{\mathrm{L}}Z_{\mathrm{L}}$, where p depends on σ, n and $\{g_j\}$. The average effect of $U(\theta)$ can be readily calculated for $\theta \sim \mathcal{N}(0, \sigma)$. First, we define the superoperators $\mathcal{H}_{\mathrm{E}}$ and $\mathcal{U}_\theta$ by their action on a generic matrix X as $\mathcal{H}_{\mathrm{E}}(X) := [H_{\mathrm{E}}, X]$ and $\mathcal{U}_\theta(X) := U(\theta)XU(\theta)^\dagger = e^{-i\theta\mathcal{H}_{\mathrm{E}}}(X)$. To find $\langle \mathcal{U} \rangle := \int_{\mathbb{R}} \mathcal{U}_\theta \, p(\theta) \, d\theta$, the superoperator describing the register's average evolution, we begin by diagonalizing $\mathcal{H}_{\mathrm{E}}$. Notice that if $H_{\mathrm{E}}|j\rangle = E_j|j\rangle$, then $\{|j\rangle\langle k|\}_{j,k=0}^{2^n-1}$ is a complete eigenbasis for $\mathcal{H}_{\mathrm{E}}$. In particular:

$$\mathcal{H}_{\mathrm{E}}\big(|j\rangle\langle k|\big) = (E_j - E_k)|j\rangle\langle k|, \tag{A1}$$

and so the spectral decomposition of $\mathcal{H}_{\mathrm{E}}$ reads

$$\mathcal{H}_{\mathrm{E}}(X) = \sum_{j,k=0}^{2^n-1} (E_j - E_k)\langle j|X|k\rangle \, |j\rangle\langle k|. \tag{A2}$$
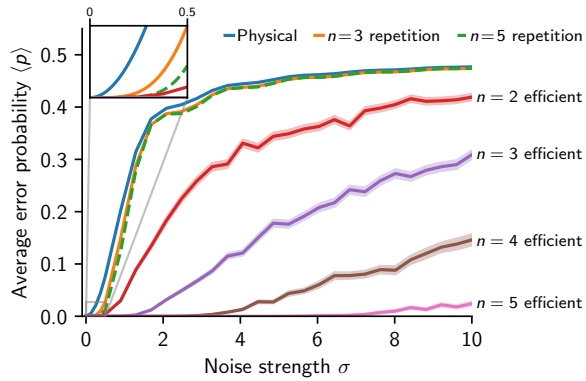
FIG. A1: We assume that the effect of the quantum fluctuator is to impart a random phase, θ. In contrast with Fig. 1 of the main text, we consider here a uniformly distributed $\theta \sim \text{unif}(-\theta_{\max}, \theta_{\max})$ for some maximum rotation angle $\theta_{\max} \geq 0$. To simplify the comparison with Fig. 1, we plot $\langle p \rangle$ versus the standard deviation $\sigma = \theta_{\max}/\sqrt{3}$, rather than versus $\theta_{\max}$ directly. By normalizing the g_j's to lie in $[0,1]^n$, σ describes the noise strength. CFD followed by a QEC recovery (if applicable) results in an effective phase- or bit-flip channel $\rho \mapsto (1-p)\rho + p\,A\rho A$, where $A = Z$ for the physical qubits, X_{L} for the repetition codes, and Z_{L} for hardware-efficient codes. The average infidelity, average trace distance and diamond distance to I are all $\propto p$. As the performance of all strategies shown depends on $\{g_j\}$, we plot the average of p over $\{g_j\} \in [0,1]^n$. The error bands for the hardware-efficient codes denote the standard error of the mean from Monte Carlo integration.

Therefore,

$$\langle \mathcal{U} \rangle(X) = \sum_{j,k=0}^{2^n-1} \left\langle e^{-i\theta(E_j - E_k)} \right\rangle \langle j|X|k\rangle \, |j\rangle\langle k|. \tag{A3}$$

Finally, we use that the characteristic function of $\theta \sim \mathcal{N}(0,\theta)$ is $\langle e^{i\theta x} \rangle = e^{-\frac{\sigma^2 x^2}{2}}$ for any $x \in \mathbb{R}$ to arrive at

$$\langle \mathcal{U} \rangle(X) = \sum_{j,k=0}^{2^n-1} e^{-\sigma^2(E_j - E_k)^2/2} \, \langle j|X|k\rangle \, |j\rangle\langle k|, \tag{A4}$$

or more compactly: $\langle \mathcal{U} \rangle = \exp(-\sigma^2 \mathcal{H}_{\text{E}}^2/2)$. This last expression is convenient for numerical analyses, as it removes the need to average over many realizations of θ numerically. For $\theta \sim \text{unif}(-\theta_{\max}, \theta_{\max})$, a similar calculation gives $\langle \mathcal{U} \rangle = \text{sinc}(\mathcal{H}_{\text{E}} \theta_{\max})$.

We consider throughout the usual (i.e., transpose) recovery channel $\mathcal{R}(\rho) = \sum_j = U_j^\dagger P_j \rho P_j U_j$, where U_j and P_j are defined as in §10.3 of [1]. A straightforward calculation allows one to express $\mathcal{R}$ directly in terms of P_{L} and H_{E} as

$$\mathcal{R}(\rho) = \sum_{j,k=0}^{p} (M^+)_{jk} P_{\text{L}} H_{\text{E}}^j \rho H_{\text{E}}^k P_{\text{L}}, \tag{A5}$$

where $M = (m_{jk})_{j,k=0}^p$ is the matrix of Knill-Laflamme coefficients $P_{\text{L}} H_{\text{E}}^j H_{\text{E}}^k P_{\text{L}} =: m_{jk} P_{\text{L}}$ and M^+ is the Moore-Penrose pseudoinverse of M (which reduces to M^{-1} when M is invertible). By construction, $m_{jk} = 0$ when $j + k$ is odd, i.e., when j and k have different parities. Furthermore, the m_{jk}'s are equal along anti-diagonal bands of M with $j + k = \text{const}$. It follows that every eigenvector $\vec{u}$ of M can be chosen such that $u_i = 0$ either for all even i or for all odd i. This implies that $(M^+)_{jk} = 0$ when j and k have different parities, so the only non-vanishing terms in Eq. (A5) will have j and k both even or both odd[1].

―――――

[1] One can also use this observation to separate M into two submatrices with $(j,k) = (\text{even, even})$ or (odd, odd), for which the pseudoinverses can be computed separately and then combined to give M^+. This is not only faster, tends to be more numerically stable for larger n.

Notice that $(\mathcal{R}\mathcal{U}_\theta)(\rho_{\mathrm{L}})$ will comprise a weighted sum of terms with the form $\mathcal{R}(H_{\mathrm{E}}^\ell \rho_{\mathrm{L}} H_{\mathrm{E}}^m)$. Using that

$$P_{\mathrm{L}} H_{\mathrm{E}}^m P_{\mathrm{L}} = \frac{1}{2} \operatorname{tr}(Z_{\mathrm{L}}^m H_{\mathrm{E}}^m) Z_{\mathrm{L}}^m \tag{A6}$$

from Eqs. (6)–(8) of the main text, and that $\rho_{\mathrm{L}} = P_{\mathrm{L}} \rho_{\mathrm{L}} P_{\mathrm{L}}$ for encoded states, we find

$$\mathcal{R}(H_{\mathrm{E}}^\ell \rho_{\mathrm{L}} H_{\mathrm{E}}^m) = \frac{1}{4} \sum_{j,k=0}^{p} (M^+)_{jk} \operatorname{tr}(Z_{\mathrm{L}}^{j+\ell} H_{\mathrm{E}}^{j+\ell}) \operatorname{tr}(Z_{\mathrm{L}}^{k+m} H_{\mathrm{E}}^{k+m}) Z_{\mathrm{L}}^{j+\ell} \rho_{\mathrm{L}} Z_{\mathrm{L}}^{k+m}. \tag{A7}$$

It follows immediately that $\mathcal{R}\langle\mathcal{U}\rangle = \langle\mathcal{R}\mathcal{U}_\theta\rangle$ must have the form

$$\langle\mathcal{R}\mathcal{U}_\theta\rangle(\rho_{\mathrm{L}}) = a_0 \rho_{\mathrm{L}} + a_1 Z_{\mathrm{L}} \rho_{\mathrm{L}} + a_2 \rho_{\mathrm{L}} Z_{\mathrm{L}} + a_3 Z_{\mathrm{L}} \rho_{\mathrm{L}} Z_{\mathrm{L}} \tag{A8}$$

for appropriate a_j's. In particular,

$$\langle\mathcal{R}\mathcal{U}_\theta\rangle(\rho_{\mathrm{L}}) = \sum_{\ell,m=0}^{\infty} \frac{(-1)^\ell i^{\ell+m} \langle\theta^{\ell+m}\rangle}{\ell!\, m!} \mathcal{R}(H_{\mathrm{E}}^\ell \rho_{\mathrm{L}} H_{\mathrm{E}}^m). \tag{A9}$$

For $\theta \sim \mathcal{N}(0,\sigma)$, $\langle\theta^{\ell+m}\rangle = 0$ if $\ell+m$ is odd, i.e., if ℓ and m have different parities. Therefore, the only non-vanishing terms in Eq. (A9) have ℓ and m both even or both odd. Comparing with Eq. (A7), one immediately sees that the a_1 and a_2 cross-terms vanish. Finally, since $\langle\mathcal{R}\mathcal{U}_\theta\rangle$ is completely positive and trace-preserving (CPTP), we can write $a_0 = 1-p$ and $a_3 = p$ for some $p \in [0,1]$, as claimed[2]. For $n=2$ the resulting function p is

$$p = \frac{1}{4g_1^2} e^{-2(g_1+g_2)^2\sigma^2} \left[2g_1^2 e^{2(g_1+g_2)^2\sigma^2} + (g_1-g_2)g_2 - 2(g_1^2-g_2^2)e^{2g_1(g_1+2g_2)\sigma^2} - e^{8g_1 g_2 \sigma^2} g_2(g_1+g_2) \right], \tag{A10}$$

where we have assumed $|g_1| \geq |g_2| \geq 0$ without loss of generality. The expressions for p quickly become complicated for larger n.

More straightforwardly, the effect of CFD on the physical qubit j can be shown to have the same form, except with $p = [1 - \langle\cos(2g_j\theta)\rangle]/2$, where $\langle\cos(2g_j\theta)\rangle = e^{-2g_j^2\sigma^2}$ for $\theta \sim \mathcal{N}(0,\sigma)$. Similarly, the logical channels for the $n=3$ and $n=5$ qubit repetition codes (i.e., phase-flip codes) can be shown to have the form of a bit-flip channel $\rho_{\mathrm{L}} \mapsto (1-p)\rho_{\mathrm{L}} + p X_{\mathrm{L}} \rho_{\mathrm{L}} X_{\mathrm{L}}$, for different functions p. Alternatively, they could be expressed as phase-flip channels with the same p's by using the labeling convention $|\pm_{\mathrm{L}}\rangle = |\pm\rangle^{\otimes n}$. For $n=3$,

$$p = \frac{1}{16} \Big(8 - 4e^{-2g_1^2\sigma^2} - 4e^{-2g_2^2\sigma^2} - 4e^{-2g_3^2\sigma^2} + e^{-2(g_1+g_2+g_3)^2\sigma^2} + e^{8(g_1+g_2)g_3\sigma^2 - 2(g_1+g_2+g_3)^2\sigma^2}$$
$$+ e^{8g_2(g_1+g_3)\sigma^2 - 2(g_1+g_2+g_3)^2\sigma^2} + e^{8g_1(g_2+g_3)\sigma^2 - 2(g_1+g_2+g_3)^2\sigma^2} \Big). \tag{A11}$$

The corresponding expression for $n=5$ is more complicated. Notice that while the codewords for the repetition code do not depend on $\{g_j\}$, Eq. (A11) does, via $\langle\mathcal{U}\rangle$.

Many common measures of performance, e.g., average infidelity, average trace distance or diamond distance from the identity channel, have the form $p \times \text{const.}$ for bit- or phase-flip channels. It is therefore highly convenient that the channels above have this form, as one can capture all of these performance measures at once (up to constant pre-factors) simply by considering p. This avoids any ambiguity in choosing one performance measure over others.

Appendix B: Pseudothresholds and non-commuting interaction terms

We now examine the pseudothresholds for our hardware-efficient codes with $\theta \sim \mathcal{N}(0,\sigma)$. That is, for a given n, we find the largest σ for which p of our codes is smaller than the p's of all physical qubits. Naturally, this pseudothreshold

---

[2] The same argument holds for symmetric distributions about $\theta = 0$ more generally, such as $\theta \sim \text{unif}(-\theta_{\max}, \theta_{\max})$.

value, which we denote σ_{th}, depends on $\{g_j\}$. As discussed above, the pseudothresholds derived from p are the same as those from average infidelity, and average trace and diamond distance from the identity.

The left panel of Fig. B1 shows the pseudothresholds for the efficient code with $n = 2$. Notice that σ_{th} diverges around the region where $g_1 = g_2$, as the codespace becomes a decoherence-free subspace (DFS) here. In order to put these values of σ_{th} in perspective, we repeat this analysis for the 3-qubit repetition code. Its pseudothreshold under CFD will depend on $\{g_1, g_2, g_3\}$, making it difficult to visualize. To get around this difficulty, we fix $\max\{g_j\} = 1$ and leave the other coupling strengths free. One could interpret this as expressing σ_{th} and $\{g_j\} \setminus \{g_{\max}\}$ in units of $g_{\max}$. To simplify comparison between two- and three-qubit codes, the right panel of Fig. B1 shows σ_{th} in this way for the $n = 2$ efficient code. Fig. B2 then shows the pseudothresholds for the $n = 3$ repetition code (left) and efficient code (right) under CFD. We used different color maps to emphasize that these plots depict very different ranges of σ_{th}, since their features would be obscured if they were shown on a single shared color map with the same scale. The pseudothresholds in Figs. B1 and B2 assume perfect operations, exact knowledge of $\{g_j\}$, and no other decoherence besides the CFD.
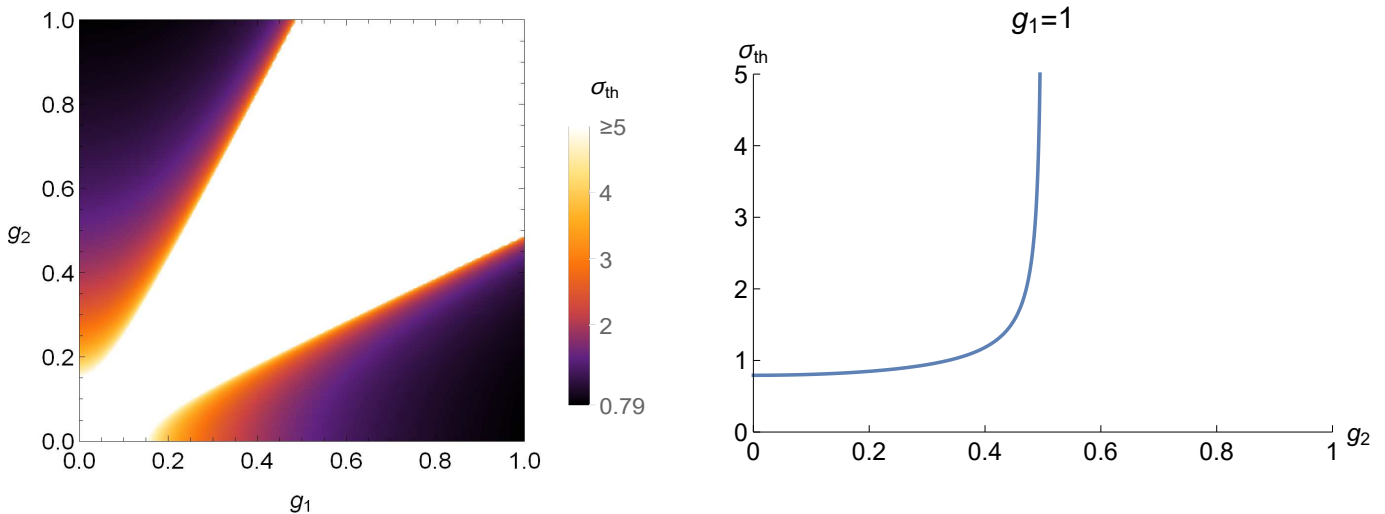

FIG. B1: The pseudothresholds for the $n = 2$ efficient code. Without loss of generality, we consider $g_1, g_2 \geq 0$. Left: the pseudothreshold vs. the coupling strengths (g_1, g_2). Right: a slice of the left panel with g_1 fixed to 1, for comparison with Fig. B2.
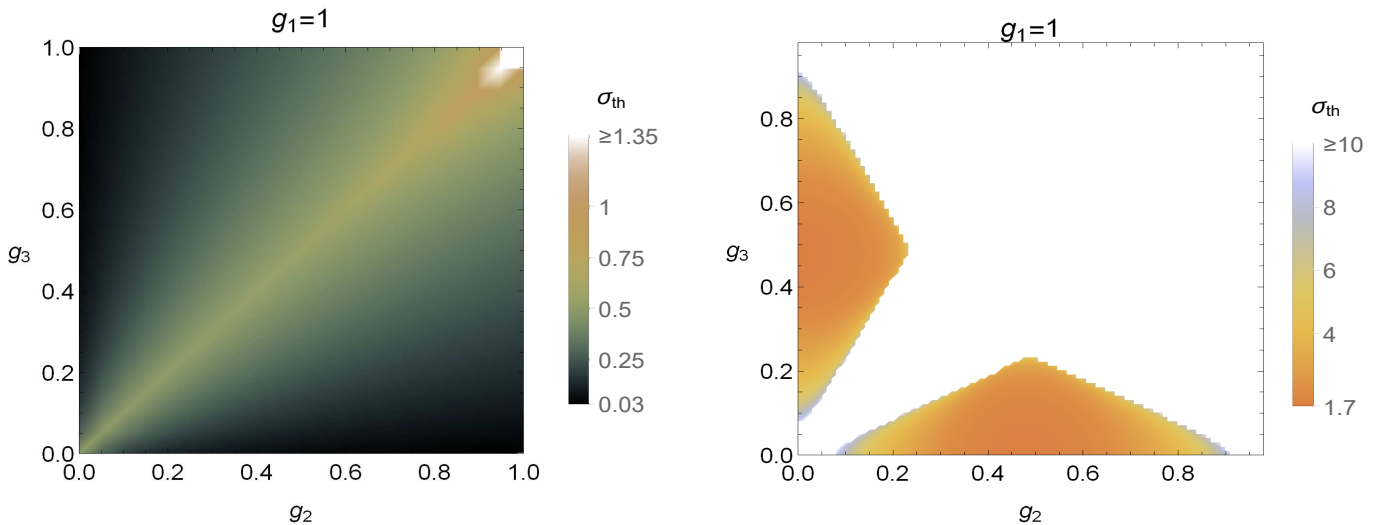

FIG. B2: Left: the pseudothresholds for the $n = 3$ repetition code. Right: the pseudothresholds for the $n = 3$ efficient code. In both panels $g_{\max} = g_1$ is fixed to 1. Without loss of generality, we consider $g_1, g_2 \geq 0$.

As discussed in the main text, there are experimental settings in which the fluctuator's state affects not only the

energy gaps of the register qubits, but also their quantization axes. This not only dephases the qubits, but more generally depolarizes them. Our hardware-efficient codes do not correct such decoherence at present; instead, the underlying interaction terms can be suppressed at the hardware level by increasing the qubits' free energy gaps. Or, they could be corrected at the "software" level by concatenating these codes with a more conventional QEC code. Fig. B3 shows the impact of such terms on the pseudothresholds of the $n = 2$ efficient code.

We model the effect of these depolarizing terms (here for 2 qubits) by the Hamiltonian,

$$H_{\mathrm{E}} = g_1^\perp X_1 + g_1^\parallel Z_1 + g_2^\perp X_2 + g_2^\parallel Z_2, \tag{B1}$$

choosing the non-commuting ($\perp$) components to lie along X_1 and X_2, without loss of generality. In the interaction picture, the total Hamiltonian (cf. Eq. (2) of the main text) takes the form

$$\tilde{H}_{\mathrm{int}}(t) = \sum_\ell \lambda_\ell |\ell\rangle\langle\ell|_f \otimes \Big(g_1^\perp \big[\cos(\omega_1 t) X_1 - \sin(\omega_1 t) Y_1 \big] + g_1^\parallel Z_1 + g_2^\perp \big[\cos(\omega_2 t) X_2 - \sin(\omega_2 t) Y_2 \big] + g_2^\parallel Z_2 \Big). \tag{B2}$$

In each run of the experiment, the register's dynamics is therefore generated by

$$\tilde{H}_{\mathrm{eff}}(t) = \lambda(t) \Big(g_1^\perp \big[\cos(\omega_1 t) X_1 - \sin(\omega_1 t) Y_1 \big] + g_1^\parallel Z_1 + g_2^\perp \big[\cos(\omega_2 t) X_2 - \sin(\omega_2 t) Y_2 \big] + g_2^\parallel Z_2 \Big), \tag{B3}$$

where the stochastic process $\lambda(t)$ is defined such that $\lambda(t) = \lambda_\ell$ when the fluctuator is in state ℓ at time t. When $g_j^\perp = 0$, this reduces to the case from the main text, with $g_j \leftrightarrow g_j^\parallel$ and $\theta = \int_0^{\Delta t} \lambda(t) dt$, where Δt is the elapsed time between recoveries. In the more general setting here, however, the U from Eq. (??) can instead be expressed as a Magnus series, which we truncate to leading order under the assumption that the couplings are weak compared to Δt^{-1}:

$$U \approx \exp\Big[-i \int_0^{\Delta t} \tilde{H}_{\mathrm{eff}}(t) \, dt \Big]. \tag{B4}$$

For the sake of illustration, we take $\omega_1 = \omega_2 =: \omega$ and $g_1^\perp/g_1^\parallel = g_2^\perp/g_2^\parallel$. Moreover, as per the hardware-level approach described above, we will assume that ω is made large compared to the coupling strengths and Δt^{-1}. We do not, however, make any assumptions about the relative strengths of $g_j^\perp$ and $g_j^\parallel$.

The integral in Eq. (B4) contains two kinds of terms: (i) those considered in the main text, of the form $\theta \, (g_1^\parallel Z_1 + g_2^\parallel Z_2)$, and (ii) rapidly oscillating terms of the form

$$g_j^\perp \int_0^{\Delta t} \lambda(t) \big[\cos(\omega t) X_j - \sin(\omega t) Y_j \big] dt. \tag{B5}$$

As with θ, here we want to describe these integrals, $\xi_x = \int_0^{\Delta t} \lambda(t) \cos(\omega t) dt$ and $\xi_y = \int_0^{\Delta t} \lambda(t) \sin(\omega t) dt$, as stochastic variables with a given distribution. Defining the Hamiltonian and propagator superoperators, $\mathcal{H}$ and $\mathcal{U}$, as done Sec. (A), we would find that the X_j and Y_j components (normalized by Δt) give quadratures of the power spectral density of λ at ω, $S_\lambda(\omega)$. As they represent quadratures of the spectrum at a (shifted) frequency ω, we can thus model $\xi_{x,y}$ as independent of θ and of each other, identically distributed, and as both following a normal distribution with zero mean and variance $S_\lambda(\omega)\Delta t/2$ (so that the total power $S_\lambda(\omega)$ is split evenly between both quadratures). Concretely, we take

$$U = \exp\left\{ -i\Big[\sum_{j=1}^2 g_j^\perp \big(\xi_x X_j + \xi_y Y_j \big) + \theta \sum_{j=1}^2 g_j^\parallel Z_j \Big] \right\}, \tag{B6}$$

and average over $\theta \sim \mathcal{N}(0, \sigma)$ and $\xi_x, \xi_y \sim \mathcal{N}(0, \sqrt{\frac{S_\lambda(\omega)\Delta t}{2}})$ (all independent). Notice that the non-commuting terms' importance depends not only on the relative strengths of $g_j^\perp$ and $g_j^\parallel$, but also on $S_\lambda(\omega)$, which is a property of the fluctuator. In fact, using the scaling property of normal distributions, Eq. (B6) can be re-written as

$$U = \exp\left\{ -i \sum_{j=1}^2 g_j^\parallel \Big[\Xi_x X_j + \Xi_y Y_j + \theta Z_j \Big] \right\}, \tag{B7}$$

where $\Xi_x, \Xi_y \sim \mathcal{N}\left(0, \sqrt{\frac{S_\lambda(\omega)\Delta t}{2}} \frac{g_j^\perp}{g_j^\parallel}\right)$. Therefore, just as σ sets the strength of the commuting noise, $\sqrt{S_\lambda(\omega)\Delta t}\, g_j^\perp/g_j^\parallel$ (which is independent of j) sets the strength of the non-commuting noise[3]. Fig B3 shows the resulting pseudothresholds for different ratios

$$r = \frac{\sqrt{S_\lambda(\omega)\Delta t}\, g_j^\perp/g_j^\parallel}{\sigma} \tag{B8}$$

of these two quantities. The ratio r describes the relative importance of the two noise types: $r \ll 1$ when commuting ($\parallel$) noise dominates, while $r \gg 1$ when non-commuting ($\perp$) noise dominates.
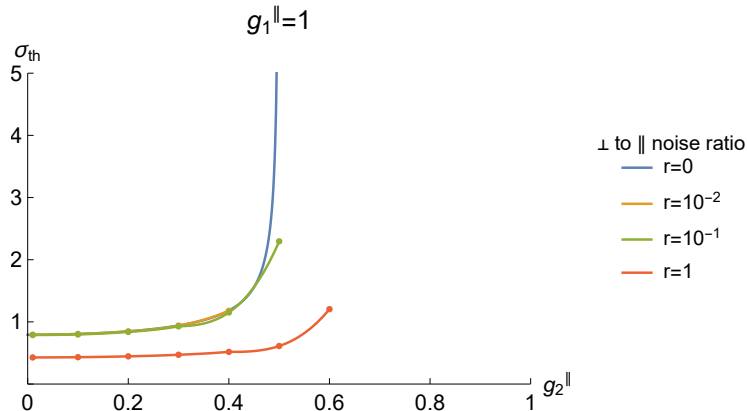
FIG. B3: The analogous pseudothresholds to Fig. B1, but under non-commuting interaction terms of various strengths. The $r = 0$ curve is the same as in Fig. B1. The $r > 0$ curves were calculated using the average infidelity (as opposed to $\langle p \rangle$), as their logical channels no longer have a simple phase-flip form. These curves appear to diverge like the $r = 0$ curve, but at larger values of $g_2^\parallel$. Only when the non-commuting ($\perp$) terms become dominant (i.e., when $r > 1$, not shown) do the pseudothresholds start to behave erratically and become ill-defined. The pseudothresholds for $r > 0$ were computed at discrete intervals; the interpolating curves serve to guide the eye.

Pseudothresholds under CFD become difficult to visualize for $n \geq 4$. Moreover, since those of efficient codes diverge around DFS's, we cannot average σ_{th} over $\{g_j\}$. Therefore, in order to analyze the sensitivity of efficient codes to calibration errors (i.e., uncertainty) in $\{g_j\}$ versus n, we will revert to plotting $\langle p \rangle$ versus σ, as in Fig. 1 of the main text. p is always bounded, so the average $\langle p \rangle$ over $\{g_j\}$ is always well-defined, in contrast with σ_{th}.

Appendix C: Monte Carlo averaging

We use Monte Carlo integration to compute $\langle p \rangle = \int_0^1 dg_1 \cdots \int_0^1 dg_n\, p$ in Figs. 1, A1, D1, D2 and D3. Specifically, we estimate $\langle p \rangle$ through

$$\langle p \rangle = \langle p(\sigma) \rangle \approx \frac{1}{N} \sum_{i=1}^N p(\vec{g}^{(i)}, \sigma) =: \mu, \tag{C1}$$

where $\vec{g} = (g_1, \ldots, g_n)$ is sampled uniformly N times from $[0,1]^n$. The sample variance of p over $\vec{g} \in [0,1]^n$ is

$$\text{Var}(p) = \frac{1}{N-1} \sum_{i=1}^N \left[p(\vec{g}^{(i)}, \sigma) - \mu \right]^2, \tag{C2}$$

so the standard error in approximating $\langle p \rangle$ by μ is $\sqrt{\text{Var}(p)/N}$, which is shown as error bars/bands in these plots. To validate this numerical averaging, we compare in Fig. C1 the Monte Carlo estimates of $\langle p \rangle$'s with analytical

---

[3] The factor of $1/\sqrt{2}$ drops out since there are two independent quadratures that contribute.

expressions [i.e., from integrating Eqs. (A10), (A11) etc.], when the latter can be calculated in Mathematica. The values of $\langle p \rangle$ for physical qubits and repetition codes shown in Figs. 1, A1, D1, D2 and D3 were found analytically to reduce unnecessary statistical noise in these plots. (That is, they use the yellow lines in Fig. C1 rather than the blue dots.) Those for efficient codes are Monte Carlo averages.
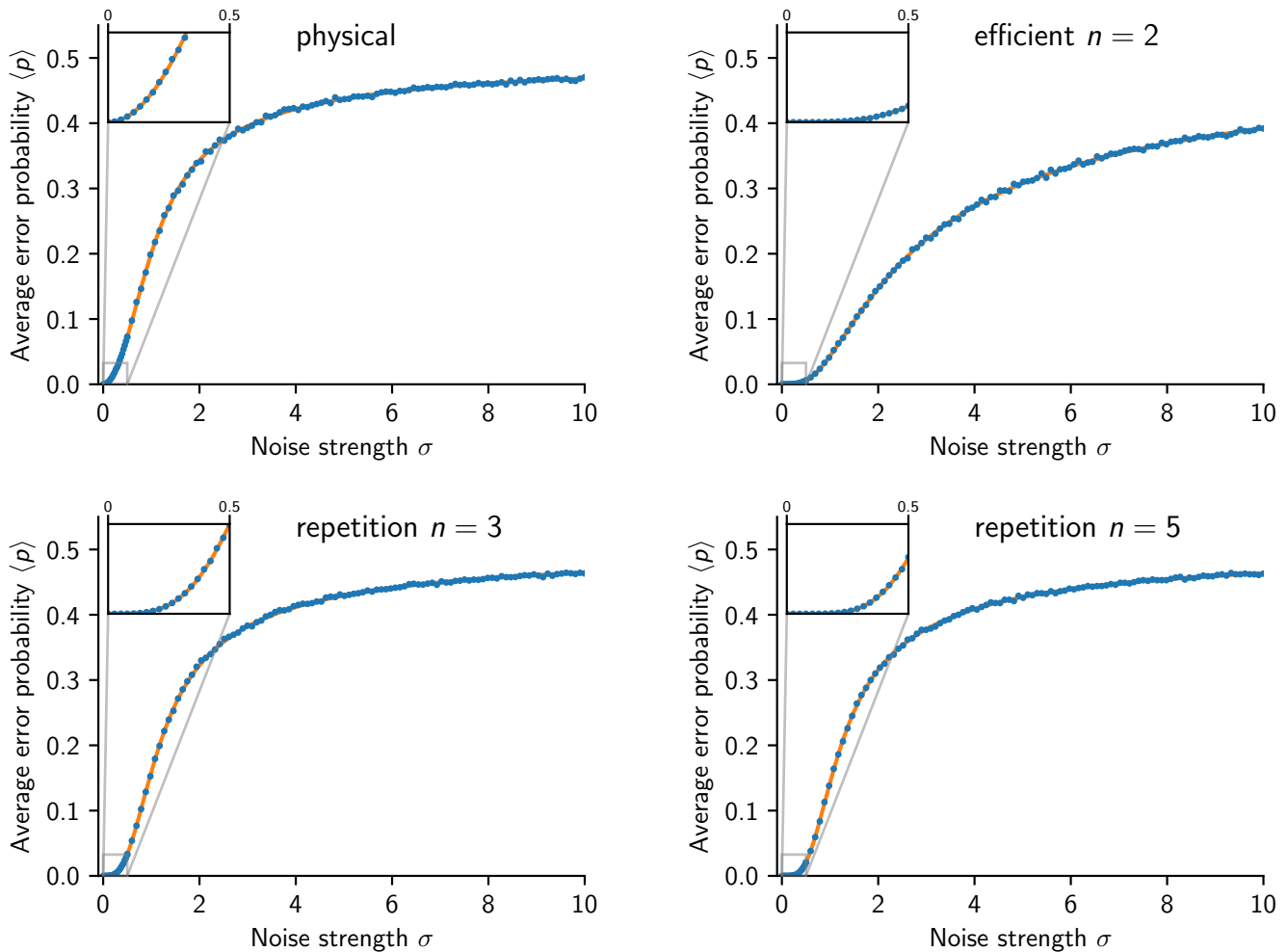

FIG. C1: Validation of Monte Carlo averaging. The yellow line is the exact $\langle p \rangle$ found analytically, and the blue dots are the Monte Carlo estimates of the same quantity. Error bars are too small to be seen due to the large number of samples used.

As further validation, we also plot the average error probability $\langle p \rangle$ in the ultra-low-noise regime in Fig. C2, compared to the corresponding power law in σ.

Appendix D: Sensitivity to calibration errors

We now analyze the effects of calibration error (i.e., uncertainty) in $\{g_j\}$. Since the encoding and recovery operations for our codes depend explicitly on $\{g_j\}$, error in measuring these coupling strengths will generally hurt overall performance. We model such errors through additive Gaussian noise on g_j: that is, if g_j is the true coupling strength, we suppose that one instead estimates $g_j + \delta g_j$, where $\delta g_j \sim \mathcal{N}(0, \Delta g_j)$. For the sake of illustration, we assume furthermore that the errors δg_j are independent across qubits, and that the measurement precision Δg_j is the same for all j. We then estimate $\langle p \rangle$ through Monte Carlo averaging both over $\vec{g} \in [0, 1]^n$ and δg_j.

Our goal is to analyze the sensitivity of efficient codes to measurement precision Δg_j as a function of n. To this end,
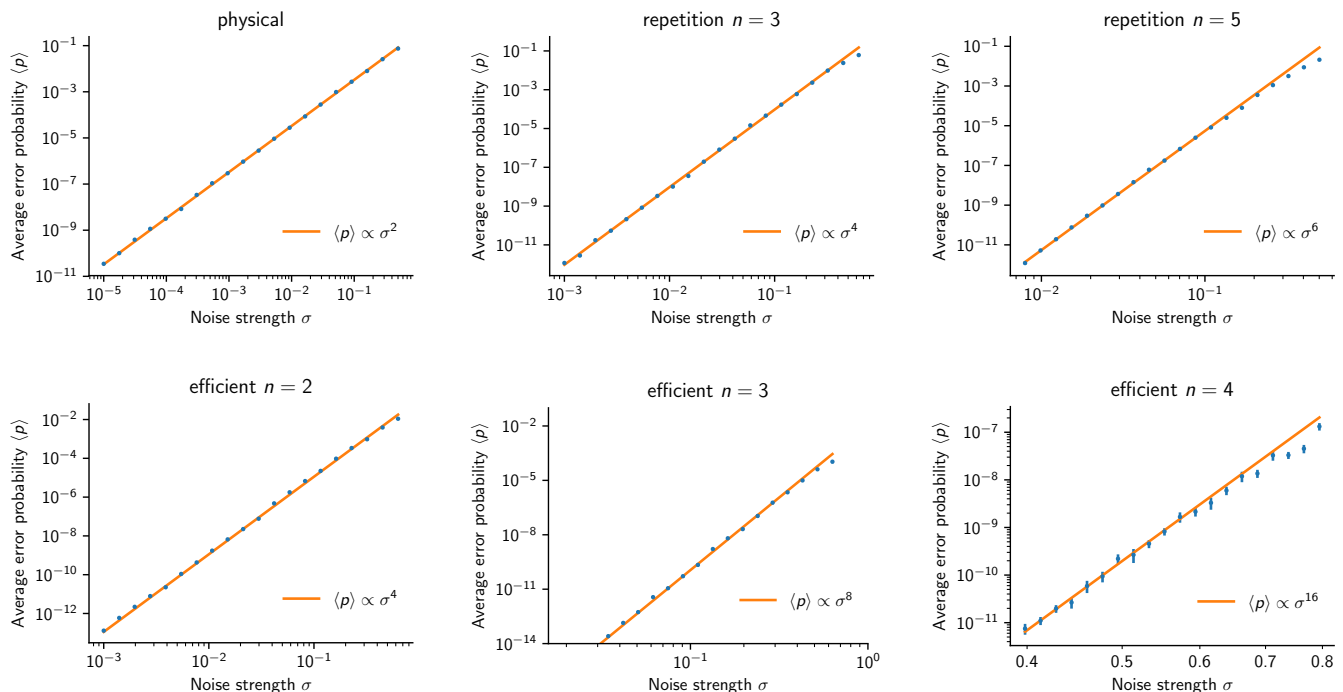
FIG. C2: Further validation of Fig. 1 from the main text, in the ultra-low-noise regime. Each code is shown in a different panel for clarity (along with an unencoded qubit for reference). The blue dots are data points; error bars, when visible, denote standard error of the mean in Monte Carlo averaging. The yellow lines are the corresponding power laws whose slopes serve to guide the eye (their offsets, which were determined through fitting, are not of primary importance here). For clarity, the panels use different ranges of σ (unlike the inset of Fig. 1), chosen to be as wide as possible, such that (i) they represent a perturbative regime ($\sigma \lesssim 1$) where $\langle p \rangle$ is described by a power law, and (ii) they do not produce unmanageably small values of $\langle p \rangle$ prone to significant numerical errors. The hardware-efficient code with $n = 5$ is not shown because it suppressed noise so strongly that it was difficult to find values of σ satisfying both (i) and (ii).

we begin by computing analogues of Fig. 1 from the main text (which assumes $\Delta g_j = 0$) with increasing calibration error. The results are shown in Fig. D1. Notice that at $\Delta g_j = 0.01$ our efficient codes all perform only slightly worse than in Fig. 1. As Δg_j increases further, however, we begin to see saturation: that is, there is an $n_{\max}$ (decreasing with Δg_j) above which there is no further improvement. Finally, when one has almost no knowledge of the coupling strengths at $\Delta g_j = 0.5$, the efficient codes are no longer effective for any n. This behavior is expected: our codes achieve a high level of protection by exploiting knowledge of the noise. In the limit where we lose this knowledge, we necessarily also lose the protection. Note that the measurement precision reported in [2] for nuclear spins near a Nitrogen vacancy center corresponds to $\Delta g_j \sim 0.01$.

Fig. D2 presents similar information in a different way. It shows explicitly the behavior of $\langle p \rangle$ vs n at different representative noise strengths and levels of calibration error. In each panel, we fix n and examine the dependence of $\langle p \rangle$ on Δg_j. In the low- and intermediate-noise regimes, this dependence is well-described by $\langle p \rangle = A_n \Delta g_j^2 + B_n$, whereas at high noise $\langle p \rangle = A_n \Delta g_j + B_n$ gives a better fit. The coefficients A_n (which correspond to $d\langle p \rangle/d(\Delta g_j^2)$ and $d\langle p \rangle/d(\Delta g_j)$ respectively, and were found using smaller steps of Δg_j than shown in Figs. D1 and D2) are plotted versus n in Fig. D3. Notice that the sensitivity to calibration uncertainty always increases with n. This is expected, since adding more qubits with $\Delta g_j > 0$ introduces more uncertainty to the system. The asymptotic behavior of the sensitivity versus n is less clear. In the high-noise regime, for instance, it is consistent with an exponential increase. In the low- and intermediate-noise regimes, however, the scaling appears sub-exponential (and perhaps even sub-linear in the latter regime).

The upper limit of $n = 5$ in these plots is due to numerical instabilities during Monte Carlo averaging. When the coupling strengths $\{g_j\}$ admit a DFS, the matrix M pertaining to Eq. (A5) becomes trivial, as does the recovery $\mathcal{R}$. Physically, this is a very fortunate situation, as it gives a good quantum memory without the need for active error correction. The same is true when $\{g_j\}$ give only an approximate DFS. However, while physically convenient, this
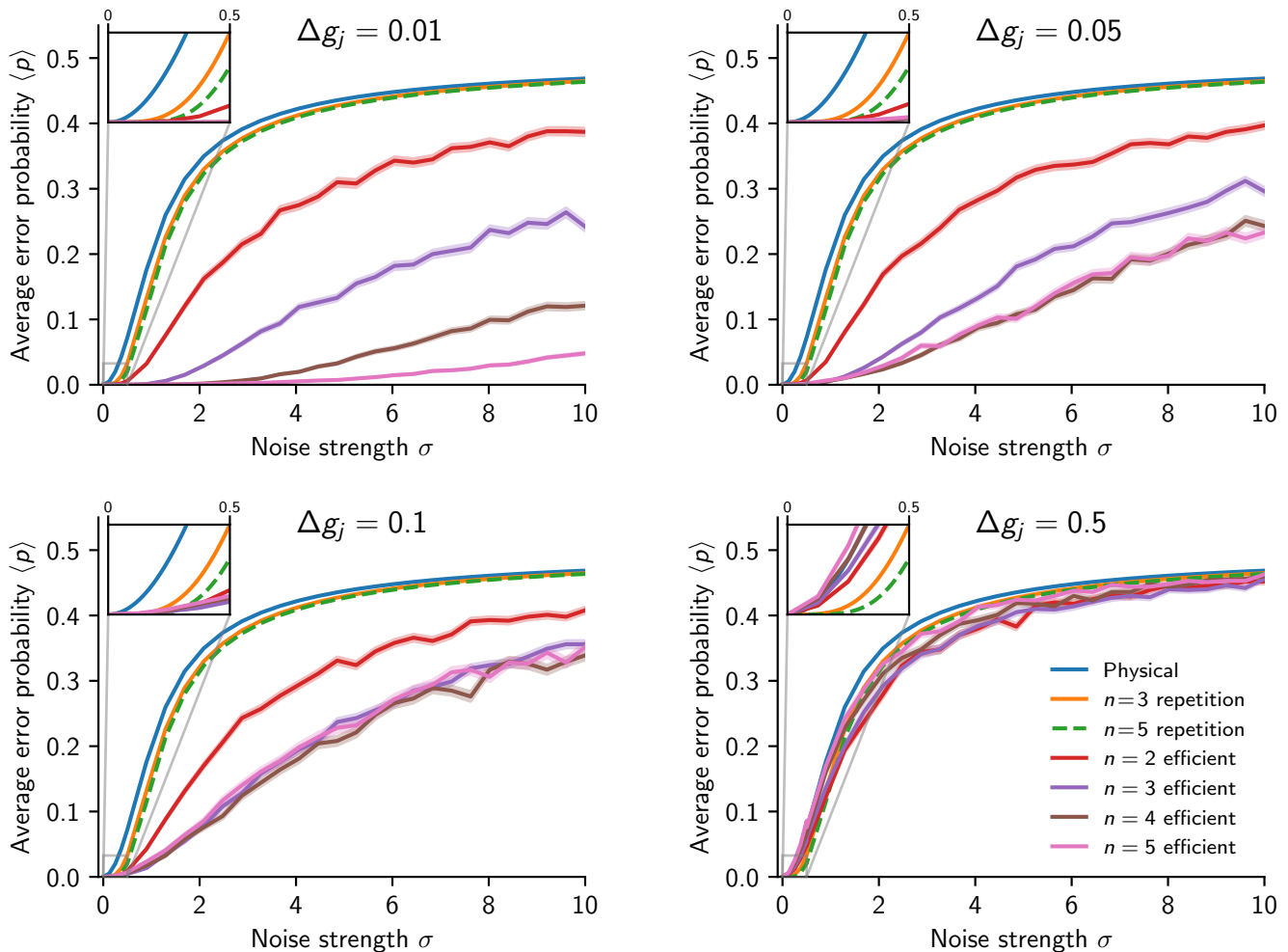
FIG. D1: The quantities analogous to those in Fig. 1 of the main text, but where with increasing measurement uncertainty Δg_j in the coupling strengths between register qubits and the common resonator. Error bands denote standard error of the mean from Monte Carlo integration.

latter case is problematic for the Monte-Carlo averaging we perform here, as it gives an ill-conditioned matrix M. This, in turn, can produce large numerical errors in p when performed automatically as a subroutine of Monte Carlo integration. Such approximate DFS's become more frequent as n increases (since it becomes more likely that two elements of a random $\vec{g} \in [0,1]^n$ be nearly equal), thus making it difficult to compute *average* p's for $n \gtrsim 6$.

[1] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, Cambridge, UK, 2010).
[2] A. Dréau, J.-R. Maze, M. Lesik, J.-F. Roch, and V. Jacques, Phys. Rev. B **85**, 134107 (2012).
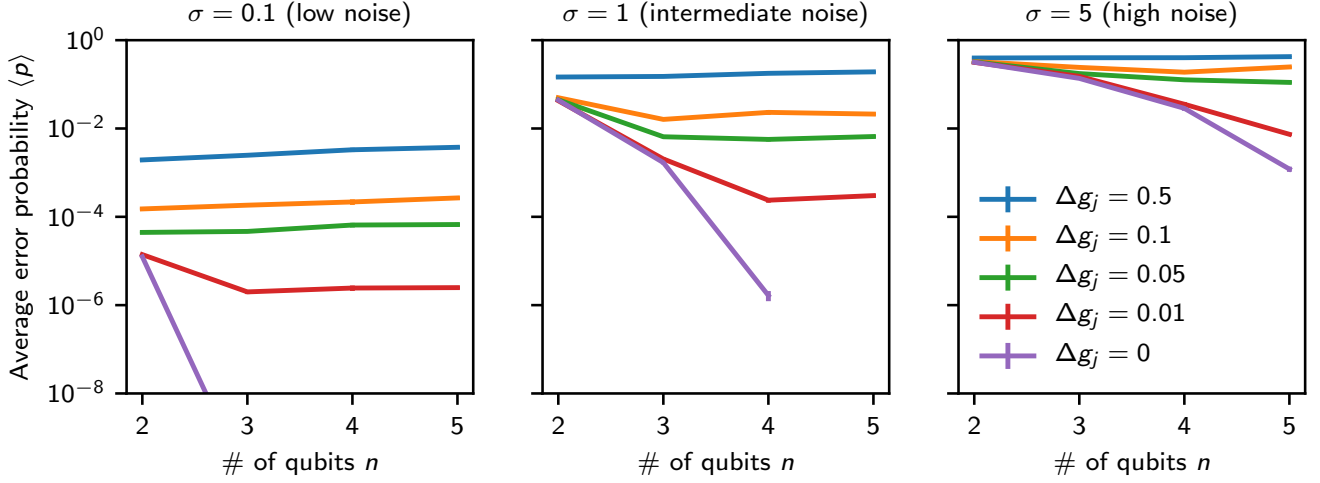
FIG. D2: Average performance of efficient QEC codes by Δg_j in three illustrative noise strength regimes. Error bars denote standard error of the mean from Monte Carlo integration, and lines are to guide the eye.
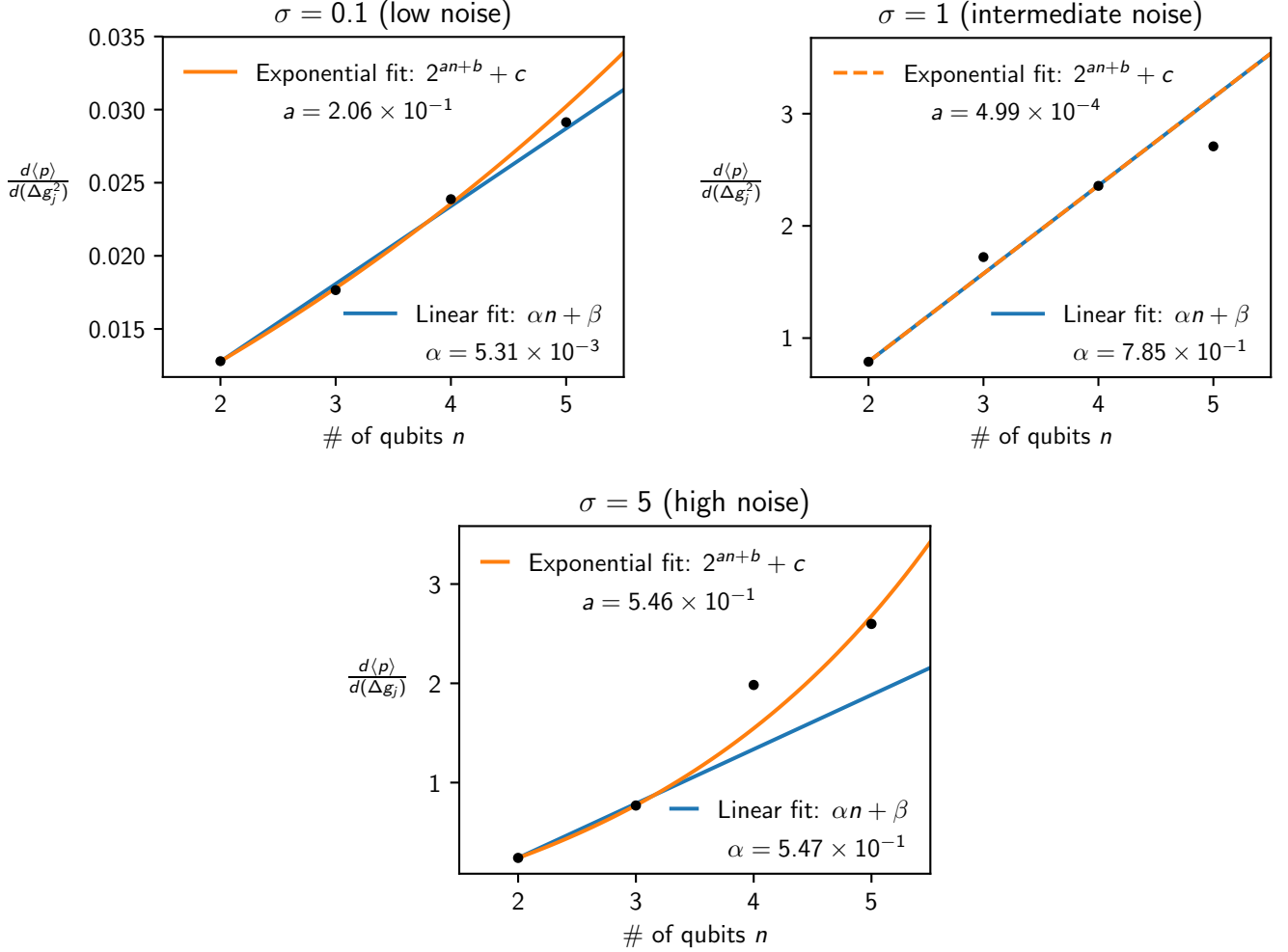

FIG. D3: Sensitivity of efficient codes to calibration errors versus n in three illustrative noise strength regime. The blue and yellow curves are the best linear and exponential fits, respectively.