

Opportunities and challenges in scaling quantum error detection on hardware

Yanis Le Fur,^{1,2} Ethan Egger,^{3,4,5} Hong-Ye Hu,⁶ Vincent Russo,⁷ William J. Zeng,⁷ Ryan LaRose^{8,9,4,5,*}

¹*Institute of Physics, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland*

²*Institute of Fundamental Physics IFF-CSIC, Calle Serrano 113b, 28006 Madrid, Spain*

³*Department of Computer Science and Engineering,*

Michigan State University, East Lansing, MI 48823, USA

⁴*Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48823, USA*

⁵*Center for Quantum Computing, Science, and Engineering,*

Michigan State University, East Lansing, MI 48823, USA

⁶*Department of Physics, Harvard University, Cambridge, MA 02138, USA*

⁷*Unitary Foundation*

⁸*Department of Computational Mathematics, Science, and Engineering,*

Michigan State University, East Lansing, MI 48823, USA

⁹*Department of Electrical and Computer Engineering,*

Michigan State University, East Lansing, MI 48823, USA

Quantum error detection can produce unbiased expectation values that exponentially converge to noiseless results as the code distance is increased. Despite this, its performance as an error mitigation technique is relatively understudied on quantum hardware because of its two main drawbacks: (i) the number of samples increases exponentially in the circuit depth/noise level, and (ii) the classical pre/post-processing generally grows exponentially in the code distance, though exceptions exist. Additionally, the constant (but often large) overhead of embedding the code and logical operations on hardware can make accuracy worse instead of better. In this work, we seek to provide a clear picture of these opportunities and challenges for scaling quantum error detection on hardware. We do so by performing a detailed benchmarking study on real and simulated noisy quantum computers, using the repetition code and triangular color code for memory experiments and logical computations with up to 74 physical qubits. In addition to these benchmarks, we estimate the pseudo-threshold of codes to map the frontier of error detection on current and future quantum computers. Our results provide a clear view of the challenges that must be overcome to use error detection in end-to-end experiments, but also indicate the opportunity to scale error detection to Megaquop computers.

I. INTRODUCTION

Due to the experimental fragility of quantum states, it is generally accepted that quantum computers must employ active and/or passive methods to deal with errors. Recent landmark experiments in quantum error correction have demonstrated logical error rates that decrease exponentially in the number of physical qubits used [1–3]. These error rates are primarily in “memory experiments” in which a logical state is prepared and then stored on the device for some time. In recent years, there has also been an explosion in the development of so-called quantum error mitigation methods [4, 5]. These error mitigation methods are generally designed to improve the accuracy or precision of computing expectation values on (current) quantum computers. While some error mitigation methods use additional space resources (i.e., qubits) — for example virtual distillation [6] and recent proposed improvements to probabilistic error cancellation [7] — most use additional time, guided by the limitations of current quantum computers. However, the recent findings from quantum error correction confirming that more noisy qubits working together can lead to better results suggest the utility of using additional space resources for quantum error mitigation.

Indeed, the error mitigation technique of quantum error detection shares many desirable features with quantum error correction. In quantum error detection, additional qubits are used to encode logical states, which are then evolved by logical operations. However, instead of attempting to actively correct errors, the logical information is measured and then post-processed to improve expectation values as in error mitigation. Intuitively, the method works by discarding final states that are not in the logical subspace of the chosen error correcting code. Error detection and variants have been proposed under the names of subspace expansion [8–11], error detection [12–14], symmetry verification [15–17], and logical shadow tomography [18]. Experimentally, the method has been used in [13, 19–26]. The main challenges are that the method requires a number of samples (shots) exponential in the size of the quantum circuit under most reasonable noise models, and that the classical pre/post-processing generally scales exponentially in the code size (a problem which several authors have attempted to alleviate [11, 18]), though exceptions (e.g. CSS codes) exist.

Despite these challenges, quantum error detection has very desirable characteristics. Most notably, while virtually all error mitigation techniques are heuristic and produce biased estimators, quantum error detection produces unbiased estimates of expectation values, assuming all errors which occur are correctable. To our knowledge, the only other unbiased quantum error mitigation tech-

* Corresponding author: rmlarose@msu.edu

nique is probabilistic error cancellation (PEC) [27–30], assuming noise can be perfectly characterized. In spite of this, quantum error detection has been relatively understudied on hardware in general, and in particular with respect to its performance as the code size scales.

In this work, we seek to fill this gap by presenting a detailed benchmarking study of quantum error detection on real and simulated quantum computers, highlighting both the opportunities and challenges of the method. Our results demonstrate that the accuracy of expectation values can (exponentially) converge as the size code distance/number of physical qubits increases, analogous to recent experiments in quantum error correction. To demonstrate this, we study the performance of the repetition code and the triangular color code on IBM quantum computers, analyzing the accuracy of expectation values as the code size increases. Our experiments use up to 74 physical qubits and 3146 physical two-qubit gates to prepare logical Bell states, providing better-than-physical performance when evaluating expectation values. We also provide a clear picture of the sampling overhead required and the amount of classical pre/post processing required to achieve these results. Although these remain challenges for the method that we do not seek to address in this work, our experimental results indicate the opportunities for error detection at the scale of Megaqop computers [31] and beyond. To these ends, we provide background on quantum error detection in Sec. II, we present our experimental results in Sec. III, and discuss our methods in Sec. IV.

II. PRELIMINARIES

A. Quantum error detection

Quantum error detection encodes a physical state into a logical state using an error correction code, implements logical operations, then projects the state into the codespace. Specifically, given a physical state ρ , an observable O , and an $[[n, k, d]]$ error correction code defined by stabilizer code $\mathcal{S} = \langle S_1, \dots, S_r \rangle$, the error-mitigated expectation value is given by

$$\langle O \rangle := \frac{\text{Tr} [\Pi \bar{\rho} \Pi^\dagger \bar{O}]}{\text{Tr} [\Pi \bar{\rho} \Pi^\dagger]} \quad (1)$$

where the projection onto the logical codespace Π is defined from the code via

$$\Pi := \prod_{i=1}^r (I + S_i) / 2 = \frac{1}{2^r} \sum_{S \in \mathcal{S}} S. \quad (2)$$

This projector Π has the following properties to mitigate errors. Given a correctable error E acting on a logical state $|\bar{\psi}\rangle$, there exists some stabilizer generator S such that $ES = -SE$. Therefore we have $\frac{I+S}{2}E|\bar{\psi}\rangle = 0$. On the other hand, if there is no error, then $\frac{I+S}{2}|\bar{\psi}\rangle = |\bar{\psi}\rangle$.

Thus, the operator Π projects or filters out correctable errors, while keeping codewords.

From these properties, it follows that the error-mitigated expectation value (1) is unbiased if all errors that occur are correctable. While in practice, bias enters due to uncorrectable errors, including logical errors, this is still a highly desirable property for a quantum error mitigation technique. To our knowledge, the only other unbiased quantum error mitigation technique is probabilistic error cancellation (PEC) [27–30]. Like error detection, this unbiased property in probabilistic error cancellation also depends on an assumption about the noise — namely, that it can be characterized to arbitrary accuracy.

Under most noise models, the probability of measuring codewords becomes exponentially small in the circuit depth. For example, consider a global depolarizing channel $\bar{\rho} \mapsto \mu \bar{\rho} + (1 - \mu)I/d$ after each layer of the circuit. Here, μ is the operation layer fidelity, $0 \leq \mu \leq 1$, and d is the dimension of $\bar{\rho}$. After D layers, the final state is

$$\bar{\rho} \mapsto \mu^D \bar{\rho} + (1 - \mu^D)I/d. \quad (3)$$

Thus, we see the probability of measuring a codeword scales as μ^D ($0 \leq \mu \leq 1$). This reveals a main challenge with quantum error detection: the probability of sampling codewords is exponentially small in the circuit depth under most noise models. (Equivalently, the number of samples required grows exponentially in the circuit depth.) Virtually all quantum error mitigation techniques have similar characteristics, either requiring exponentially many samples or requiring some other resource which grows exponentially with the problem size.

In quantum error detection and correction, a larger code distance is desirable. This is because a larger distance separates codewords more, making it less likely for logical errors to occur. This means that, in principle, using *more* noisy qubits in quantum error detection *reduces* errors in expectation value estimates. In this work, we describe multiple benchmarks performed on IBM quantum computers which show this is true: The main advantage with quantum error detection is that accuracy can be improved by using additional physical qubits (larger code distances).

However, the classical (post-)processing required by quantum error detection is computationally hard. There are two broad techniques to implement error detection (evaluate (1)):

1. Measure all qubits (in the computational basis) and only keep codewords. This is efficient *assuming the codewords are known*. While computing codewords from stabilizer generators is efficient for certain codes (CSS codes), in general this is a hard problem. In Fig. 1, we show runtime of the best-known algorithm to compute codewords from stabilizers. The punchline is that computing codewords from stabilizers scales exponentially and becomes infeasible for relatively small $n \simeq 28$ codes.

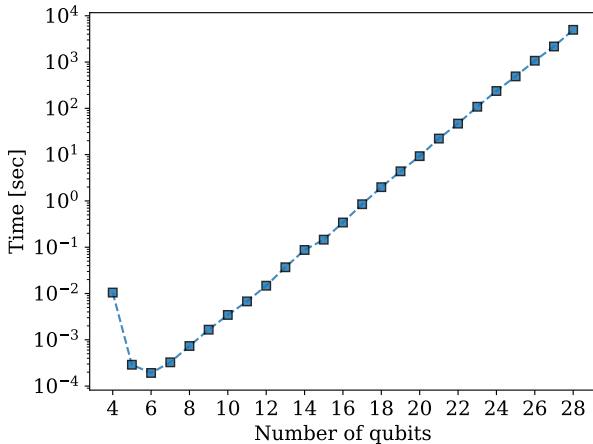


FIG. 1. Time to compute codewords from stabilizer generators for random codes on $4 \leq n \leq 28$ qubits. Using the best algorithm to our knowledge (generating a random state vector and multiplying it by $S + I$ for each generator S) and the best implementation of this algorithm (in Stim [32]), we benchmarked the time to compute codewords from stabilizer generators for random codes with $4 \leq n \leq 28$ physical qubits. The results confirm the exponential scaling and give a sense for the practicality of this strategy: for $n \lesssim 25$ this can be done in minutes, whereas for $25 \lesssim n$ this takes hours or longer, and we ran out of memory for $n > 28$, demonstrating the hard pre/post-processing required for error detection.

- Evaluate the numerator and denominator of (1) as an expectation value problem for the modified observables $\Pi^\dagger \bar{O} \Pi$ (numerator) and Π (denominator). (Note that as long as \bar{O} and Π commute then the modified observable for the numerator can be simplified to $\Pi \bar{O}$.) This means we can implement quantum error detection by estimating the expectation values $\text{Tr}[\rho(\Pi \bar{O})]$ (numerator, non-subsystem stabilizer codes) and $\text{Tr}[\rho \Pi]$ (denominator), where the projector onto the codespace Π is defined in (2). Unfortunately, Π is not Pauli and so does not admit an efficient stabilizer/tableau description. Indeed, Π is in general a summation of 2^r Paulis where $r \equiv n - k$ is the number of stabilizer generators encoding n physical qubits into k logical qubits. This means that the time for classical post-processing generally scales as 2^r .

This reveals a second main challenge with quantum error detection: the classical (post-)processing scales exponentially in the code size.

B. Literature review of quantum error detection experiments

Table I and the papers we highlight in this section provide a current picture of the experimental status of quantum error detection. The theory of quantum error

detection is well established. In an influential 2020 paper, McClean *et al.* [42] proposed it as an “intermediate form of error correction” to enable practical applications, and generalized the projection (2) to a subset of symmetries with general coefficients (i.e., modified (2) to $\Pi = \sum_i c_i S_i$ for general coefficients c_i) that, when measured, leads to a generalized eigenvalue problem. Numerical simulations show improved accuracy for a benchmark with the $[[5, 1, 3]]$ code and for a four-qubit Hydrogen molecule with problem symmetries under a single-qubit depolarizing noise model. Further, recognizing the challenging classical post-processing, the authors propose a stochastic sampling scheme to evaluate the error-mitigated expectation value (1). Last, the paper also proposes modifying the projectors in (2) as $(I + (-1)^{s_i} S_i)/2$ to utilize the i th syndrome bit s_i for “corrections with recovery operations.”

Around the same time, a 2021 paper by Cai explored similar ideas under the names of symmetry verification and symmetry expansion [43], with the primary difference being that the symmetries in this work are problem symmetries, whereas the symmetries in the previous work (McClean *et al.* [42]) are either problem symmetries or stabilizers from error correction codes. Cai numerically investigates their performance with up to $n = 12$ qubit 2D Fermi-Hubbard models with depolarizing noise, considering different sets of symmetries arising from the Hubbard model, and showing improved performance.

Several works have performed error detection with the use of ancilla qubits. For example, in 2024 Self *et al.* used ten data qubits and two ancilla qubits to measure stabilizers of the $[[k, k + 2, 2]]$ “Iceberg code” on the Quantinuum H2 computer [36], demonstrating better-than-physical performance on mirror circuit and quantum volume benchmarks. In 2025, Martiel and Javadi-Abhari used up to 18 ancilla qubits and up to 50 data qubits on IBM quantum computers to prepare random stabilizer states, showing a 236x fidelity gain relative to the physical experiment [38]. This paper is also notable for using codes tailored to hardware and introduce techniques to identify and measure low-weight stabilizers. Additionally in 2025, Javadi-Abhari *et al.* used 8 ancilla qubits and up to 120 data qubits to prepare GHZ states [39]. Last, in 2025 Chertkov *et al.* [37] implemented quantum error detection in an experimental study of a dissipative quantum process using the $[[4, 2, 2]]$ code and two ancilla to measure stabilizers.

In 2025, Chertkov *et al.* [37] implemented quantum error detection in an experimental study of a dissipative quantum process. This process contains mid-circuit resets that are applied at random locations/times. In their work, the authors perform a logical version of the computation with each pair of physical qubits encoded with the $[[4, 2, 2]]$ code. Using two ancilla qubits to measure the stabilizer generators of this code, they implement a reset whenever an error is detected via ancilla measurement. In this sense, the authors “avoid the exponential post-selection [sampling] penalty of quantum error de-

Year	Problem	Code	Physical qubits	Computer	Ref
2015	Bell state preparation	[[2, 0, 2]]	4	Four superconducting qubits	[33]
2020	Two-qubit Hydrogen ground state	[[4, 2, 2]]	4 + 2 ancilla	IBM Tokyo	[34]
2022	State preparation	[[5, 1, 3]]	5	12 superconducting qubits	[35]
2024	Mirror circuits/Quantum volume	Iceberg [[k + 2, k, 2]]	10 + 2 ancilla	Quantinuum H2	[36]
2025	Dissipative quantum dynamics	[[4, 2, 2]]	28	Quantinuum H2	[37]
2025	(Random) State preparation	Spacetime	50 + 18 ancilla	IBM Kingston	[38]
2025	GHZ state preparation	GHZ state	120 + 8 ancilla	IBM Aachen (and others)	[39]
2025	Bell state preparation	[[4, 2, 2]]	4	IBM Kyiv	[40]
2026	Bell state preparation	[[2, 0, 2]]	4	Silicon quantum computer	[41]

TABLE I. Quantum error detection experiments on quantum computers.

tection by using an adaptive quantum circuit whose dynamics are controlled by the mid-circuit error detection measurements”. In other words, the authors of this paper similarly identify the exponential sampling overhead of quantum error detection. While they address it, they do so for a particular quantum circuit and application, and not for error detection in general.

III. RESULTS

As a first experiment, we prepare the $|\bar{0}\rangle$ state of the n -qubit repetition code on IBM quantum computers. We then apply an even number d of \bar{X} gates and measure the expectation value $\langle \bar{Z} \rangle$, which would be one on an ideal (noiseless) quantum computer. The results for this experiment on IBM Kyiv are shown in Fig. 2. Here, we see that the error detection results are better than the physical results on average. Further, we see that the error detection results converge exponentially to the ideal expectation value as the code distance (number of physical qubits) increases.

We can gain intuition for the exponential convergence by considering a model of rate p global depolarizing noise. In this case, for any $[[n, 1]]$ code and any logical operator \bar{O} , the error-mitigated expectation value as a function of n is

$$\langle \bar{O} \rangle(n) = \frac{p\text{Tr}[\bar{\rho}\bar{O}]}{p + \frac{1-p}{2^{n-1}}}. \quad (4)$$

Since $\text{Tr}[\bar{\rho}\bar{O}]$ is the ideal (noiseless) expectation value, we see the error-mitigated result converges to the ideal result exponentially in the number of physical qubits n . We remark the limiting case $n = 1$ reproduces the unmitigated (noisy) expectation value $p\text{Tr}[\rho O]$. While Eqn. 4 follows quickly from (1) with the definition of global depolarizing noise, a derivation is included in Appendix A for completeness.

While we see exponential convergence on hardware, the repetition code is classical — it can only correct bit-flip errors. Furthermore, the encoding $|0\rangle \mapsto |\bar{0}\rangle$ is trivial and every stabilizer is diagonal in the Z -basis, so computing (1) reduces to computational basis measurements. In this sense, this presents the best-case scenario for quantum error detection. In the following experiments, we successively relax these conditions to further analyze the

performance of error detection on hardware under more realistic settings.

First, we consider a multi-logical-qubit computation with the repetition code, instead of a single logical qubit memory experiment. In particular, we use the n -qubit repetition code to prepare a logical Bell state $|\bar{\Phi}^+\rangle := (|\bar{0}\bar{0}\rangle + |\bar{1}\bar{1}\rangle)/\sqrt{2}$ and measuring $\langle \bar{Z}\bar{Z} \rangle$. The results are shown in Fig. 3 for $n = 3, 5,$ and 7 , with the physical results shown at $n = 1$ for reference. As can be seen, in this experiment the error detection results are better than the physical results on average. While the results still converge to the ideal noiseless value, the rate of convergence is markedly slower than in the previous memory experiment without any two-qubit logical gates. Although the repetition code has a transversal logical CNOT, the single-qubit logical Hadamard requires a ladder of CNOTs. This makes it more probable for errors to spread. Also, it increases periods in the quantum circuit where qubits are idling. To address errors accumulating during this idle windows, we also perform the experiment with dynamical decoupling (DD). As can be seen in Fig. 3, the encoded + DD results are better than just the encoded results, and the rate of convergence to the ideal value is slightly faster.

While the repetition code simplifies several experimental aspects, it is classical and only able to correct for bit-flip errors. To move beyond this, we turn to the triangular color code. Briefly, the distance d triangular color code encodes $k = 1$ logical qubit using $n = (3d^2 + 1)/4$ physical qubits. For distance $d = 3$, the triangular color code is the seven-qubit Steane code, and higher distances generalize this by a triangular subregion of a hexagonal tiling of a two-dimensional lattice [44]. A full discussion of the triangular color code is provided in Sec. IV. Using this code, we perform the same logical Bell state preparation experiment, this time applying \bar{X}^D for even circuit depth D after the logical Bell state is prepared. We emphasize the difference compared to the repetition code is a non-trivial encoding circuit mapping $|0\rangle^{\otimes n} \mapsto |\bar{0}\rangle^{\otimes k}$. The results are shown in Fig. 4 for the distance $d = 3, 5,$ and 7 triangular color codes on IBM Boston. At distance three ($n = 14$ physical qubits), the encoded/error detection results are better than the physical results for all circuit depths tested. At distance 5 ($n = 38$ physical qubits) however, there is a cross-over region. Namely, the encoded results are better than physical for depths up to 20, then become worse than physical after. This ex-

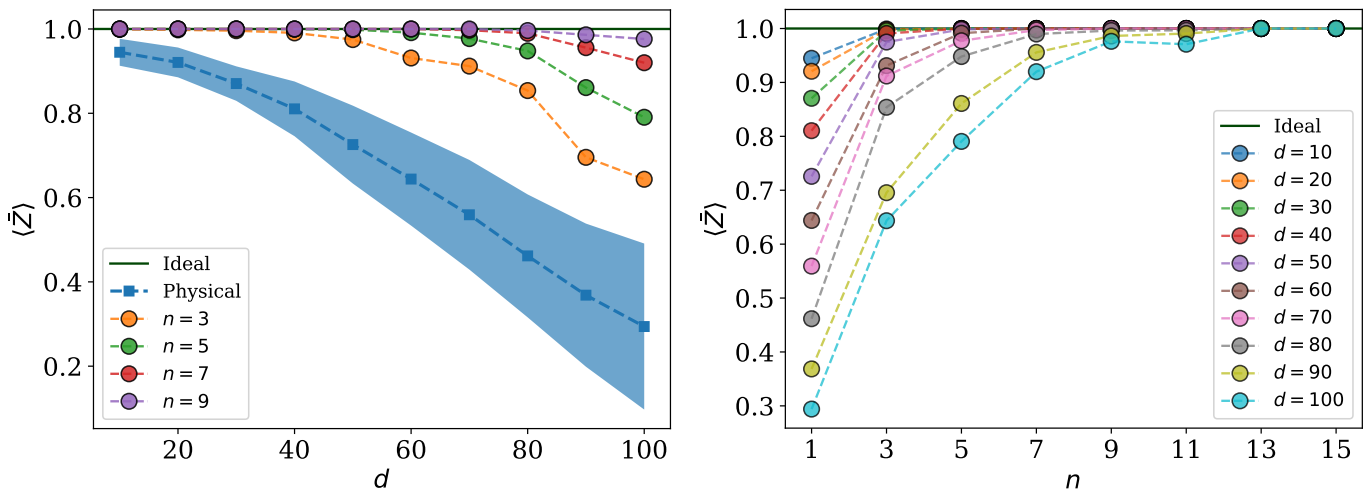


FIG. 2. Repetition code memory experiment results on IBM Kyiv. The left subplot shows $\langle \bar{Z} \rangle$ versus circuit depth d , with the average physical result in blue (colored ± 1 standard deviation) and encoded results for $n = 3, 5, 7, 9$ in different colors. As can be seen, the encoded results are better than the physical results on average, and the accuracy of the encoded results increases as n increases. The right subplot shows the same data but with $\langle \bar{Z} \rangle$ versus n , with different depths d shown in different colors. These curves demonstrate the exponential convergence in n . Note that the average physical results are shown at $n = 1$.

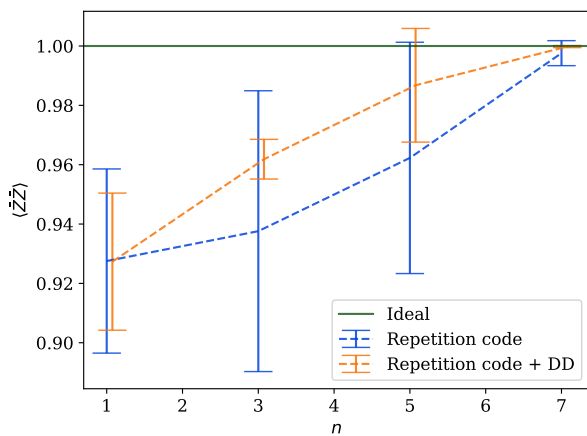


FIG. 3. Logical Bell state preparation with the n -qubit repetition code on IBM Brisbane, measuring $\langle \bar{Z}\bar{Z} \rangle$. Results are shown for $n \in \{3, 5, 7\}$ with the unencoded (physical) result shown at $n = 1$ for reference. The blue curve shows the error detection results, while the orange curve shows the error detection results with XY -4 dynamical decoupling applied to the circuit. Error bars show one standard deviation over ten independent trials.

pected behavior illustrates what we call the *pseudothreshold* — i.e., given a particular circuit, observable, and noise model, the pseudothreshold p^* is the noise level below which error detection provides superior performance and is able to exponentially suppress errors, and above which error detection provides worse performance. This is completely analogous to thresholds of quantum error correction codes, but we use the term pseudothreshold to emphasize the dependence on the particular circuit, ob-

servable, and noise model. Similar to the previous experiment, we also apply dynamical decoupling to the logical circuits. While the performance with dynamical decoupling is better at some depths, it is less emphatic and less clear than the results we obtained for the repetition code. Finally, we also implemented the distance 7 triangular color code corresponding to 74 physical qubits, but did not sample any codewords on hardware after 10^6 shots. To study the code at this distance, we implemented a noise model of single-qubit depolarizing noise after every moment (parallel layer of circuit operations), for a noise rate comparable to hardware. In this case, we consider all-to-all connectivity, instead of the heavy hex connectivity provided by IBM Boston. The noisy simulation results indicate that error detection outperforms the physical experiment at all tested circuit depths. The same noisy simulations were performed for the distance 3 and distance 5 experiments, and are shown for comparison on the respective subplots in Fig. 4.

Although we chose $\bar{Z}\bar{Z}$ for the observable in this experiment, we could have equally well chosen $\bar{X}\bar{X}$ or $\bar{Y}\bar{Y}$ since they are related by a transversal basis rotation in the triangular color code. To illustrate this, Fig. 5 shows the distance three experiment performed with the $\bar{X}\bar{X}$ observable. The only difference in the circuit is a transversal Hadamard applied before the computational basis measurement. A histogram of results over 15 independent experiments on IBM Fez show that the error detection performance is better than physical on average, just as with the $\bar{Z}\bar{Z}$ observable used for the previous experiment.

Finally, we explore the pseudothreshold behavior exhibited in Fig. 4 more deeply. Again, we define the pseudothreshold as the physical noise rate p^* below which error detection provides better performance, provided a given circuit, observable, and noise model. Just as in

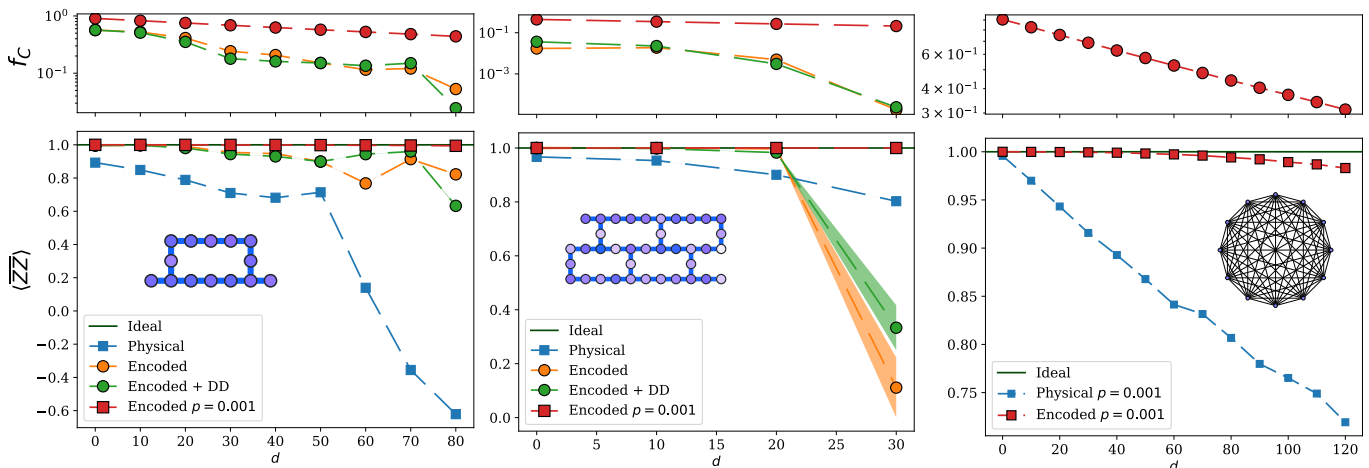


FIG. 4. Results of preparing logical Bell states with the triangular color code on IBM Boston and measuring the expectation value $\langle ZZ \rangle$. From left to right are distances three, five, and seven, corresponding to $n = 14, 38,$ and 74 physical qubits. For distance three, the encoded results are better than the physical results for all executed depths d . For distance five, the encoded results become worse than the physical results at depth $d = 30$. Colors show one standard deviation calculated from the fraction of codewords f_C sampled. For the distance seven code, no codewords were sampled at depth $d = 0$ with 10^6 shots, and results are shown for depolarizing noise with all-to-all connectivity, also shown in the distance three and five plots for reference. As in previous experiments, circuit depths are increased by inserting an even number d of \bar{X} gates after the logical Bell state is prepared. Inset graphs show the device subgraphs used in the experiment (for distance seven, the graph is all-to-all on 74 qubits, shown for 12 nodes for visualization purposes).

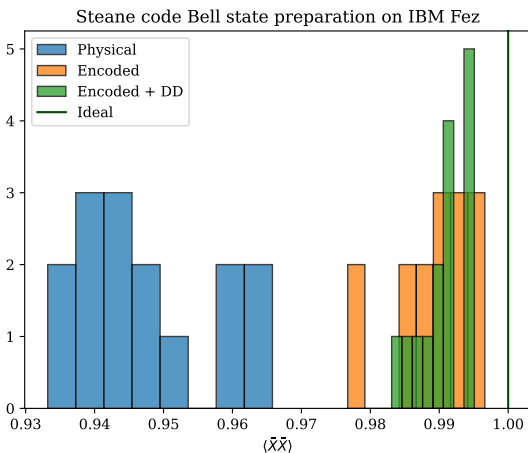


FIG. 5. Histogram of results for preparing the logical Bell state $|\bar{\Phi}^+\rangle$ on IBM Fez and measuring $\langle \bar{X}\bar{X} \rangle$ with the distance $d = 3$ triangular color code (Steane code).

error correction, (pseudo)thresholds are difficult to evaluate analytically, and this can only be done in special cases. To explore the pseudothreshold here, we consider the same logical Bell state preparation experiment with the triangular color code and a model of rate p single-qubit depolarizing noise applied after every parallel layer of operations. To estimate the pseudothreshold, we simulate the physical experiment and the encoded experiment at multiple distances, then estimate the crossover point. The results are shown in Fig. 6. As expected, we see that

the pseudothreshold p^* drops exponentially in the depth of the circuit. This sub-threshold region characterizes the circuit volumes and noise rates for which error detection provides better-than-physical performance.

IV. METHODS

The n -qubit repetition code is an $[[n, 1, n]]$ code (for odd n) with stabilizer generators $\mathcal{S} = \langle Z_1 Z_2, Z_2 Z_3, \dots, Z_{n-1} Z_n \rangle$ and logical operators $\bar{Z} = Z_1$ and $\bar{X} = X_1 X_2 \dots X_n$ [45, 46]. As a classical code, it can only detect and correct bit-flip (X) errors. The logical CNOT gate is transversal: $\overline{\text{CNOT}} = \prod_i \text{CNOT}_{i, i+n}$. The logical Hadamard-like gate is not transversal and requires a ladder of CNOT gates.

The distance- d triangular color code is an $[[n, 1, d]]$ code with $n = (3d^2 + 1)/4$ physical qubits arranged on a triangular lattice with three-colorable faces [44, 47, 48]. Unlike the repetition code, it is a quantum code capable of correcting both bit-flip and phase-flip errors on any physical qubit. For distance $d = 3$ (the Steane code, $n = 7$), the stabilizer group is generated by six operators (three X -type and three Z -type) corresponding to the faces of the lattice, and this generalizes to higher distances with larger lattices. In the triangular color code, both the logical Hadamard and logical CNOT are transversal.

Unlike the repetition code in which logical state preparation $|0\rangle^{\otimes n} \mapsto |\bar{0}\rangle^{\otimes k}$ is trivial, the triangular color code requires an encoding method. Generally this can always

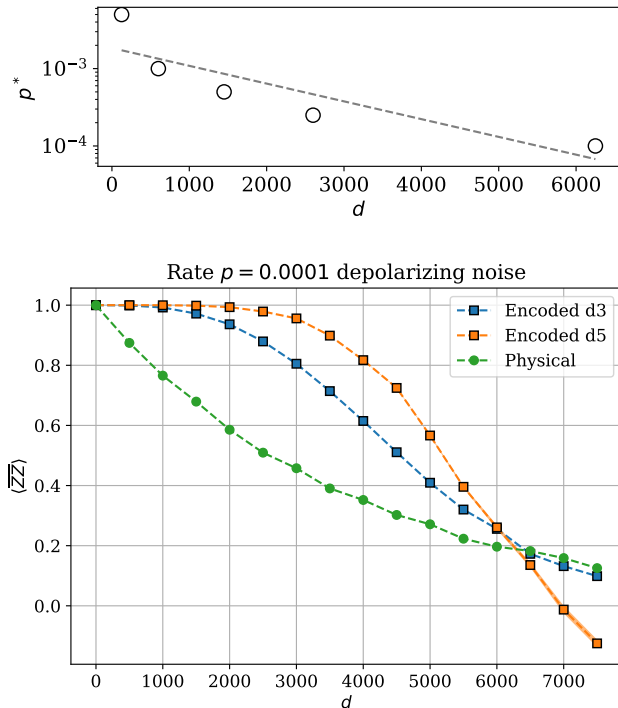


FIG. 6. Estimated pseudothreshold p^* (top panel) for the triangular color code Bell state preparation experiment, applying \bar{X}^d for a given depth d . The pseudothreshold p^* is defined as the physical noise rate p below which error detection outperforms the physical experiment. Markers show calculated thresholds and the dashed line shows an exponential fit. Thresholds are calculated by simulating the experiment with different code distances and estimating the crossover point (bottom plot), the same method for estimating thresholds of quantum error correction codes.

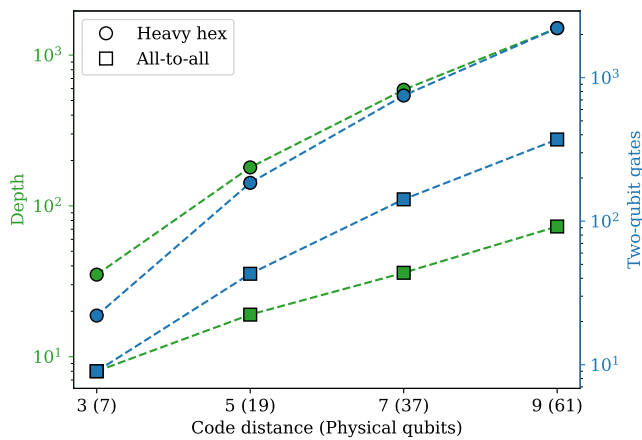


FIG. 7. Number of two-qubit gates and circuit depths for unitarily preparing the logical zero state of the triangular color code from the all zero physical state, used for state preparation in Fig. 4.

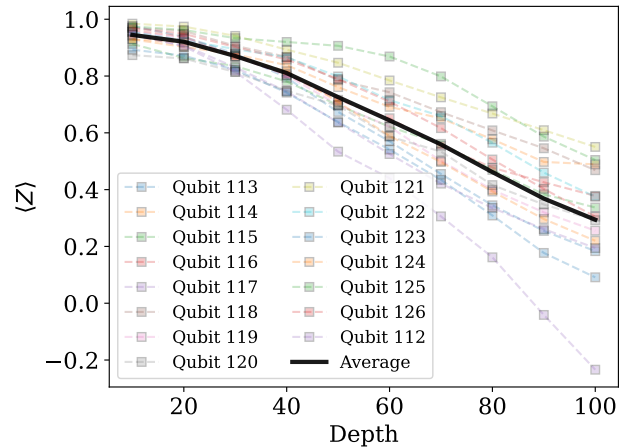


FIG. 8. Individual physical qubit results for the repetition code memory experiment from Fig. 2.

be done for error correcting codes by measuring each stabilizer and applying an appropriate correction to ensure the $+1$ eigenstate of all stabilizers — i.e., the $|\bar{0}\rangle$ state — is prepared. This is equivalent to a round of error correction applied to the $|0\rangle^{\otimes n}$ physical state. On IBM quantum computers we are currently unable to reliably perform this encoding technique, so we instead implement a unitary circuit U such that $U|0\rangle^{\otimes n} = |\bar{0}\rangle^{\otimes k}$. This can always be done by performing Gaussian elimination on the stabilizer tableau (check matrix) of the code and implementing the corresponding circuit operations. Fig. 7 shows the number of two-qubit gates and circuit depths for unitarily encoding the logical zero state of the distance d triangular color code, both on all-to-all connectivity used in noisy simulations and on the heavy hex topology on IBM quantum computers. This plot shows the relative complexity of unitary encoding as well as the overhead incurred by embedding to the heavy hex architecture.

Unless otherwise mentioned, we performed physical experiments with 10^4 shots and logical experiments with 10^5 shots on hardware. All hardware experiments were performed with IBM quantum computers. For each experiment, we select the best subset of qubits to run on from the latest calibration data. Qubits generally exhibit non-trivial variance over space and time. Figure 8 shows the individual qubit results from the average physical results calculated in Fig. 2. For experiments, we choose contiguous qubit subsets to minimize the number of two-qubit gates required in circuits. For the repetition code these subsets are simply lines of qubits for both memory and logical Bell state preparation experiments. For the triangular color code, we choose adjacent hexagons of qubits on the device. Noisy simulations were performed with a Clifford circuit simulator [32] assuming all-to-all connectivity and a model of single-qubit depolarizing noise applied after every moment.

V. CONCLUSION

In this work, we have experimentally demonstrated that quantum error detection can improve the accuracy of expectation values exponentially in the number of physical qubits, provided the hardware noise rate is below a threshold. This mirrors the threshold behavior observed in quantum error correction and provides a key opportunity for performing quantum error detection experiments on hardware. For the repetition code memory experiments, we observe clear exponential convergence as n increases, consistent with the simple theoretical prediction of (4). While the Bell state preparation experiments also show the opportunities for quantum error detection, they exhibit the challenges of scaling as well. In particular, as the code size grows, compiled circuits become deeper and introduce additional noise that can push the effective error rate above the pseudothreshold.

We expect that optimized encoding circuits would further improve our experimental results. Further, we expect that non-unitary encoding circuits would improve our results when experimentally available, although these also involve additional noise arising from noisy ancilla state preparation and syndrome measurements. It is an interesting future direction to develop quantum error correction codes tailored to particular hardware to alleviate these challenges. While we characterize the post-

selection rate (fraction of codewords f_C measured), we did not normalize the improvement of quantum error detection by the additional resources (shots) used.

Our benchmarks show that quantum error detection holds strong promise as an unbiased error mitigation technique. If new techniques are introduced to alleviate the exponential sampling overhead and the exponential classical processing, even larger-scale, application-oriented experiments with quantum error detection will be able to be performed. Such techniques have been proposed in recent literature, e.g. generalized subspace expansion [11] and logical shadow tomography [18], and are of significant interest for future work.

CODE AND DATA AVAILABILITY

Code to reproduce experiments is available at [49].

ACKNOWLEDGMENTS

We thank IBM for providing access to their quantum computers through the Quantum Research Credits Award program. The views expressed in this article are those of the authors and do not necessarily reflect those of IBM.

-
- [1] Google Quantum AI, *Exponential suppression of bit or phase errors with cyclic error correction*, *Nature* **595**, 383 (2021).
- [2] Google Quantum AI, *Suppressing quantum errors by scaling a surface code logical qubit*, *Nature* **614**, 676 (2023).
- [3] Google Quantum AI, *Quantum error correction below the surface code threshold*, *Nature* **638**, 920 (2025).
- [4] Z. Cai, R. Babbush, S. C. Benjamin, S. Endo, W. J. Huggins, Y. Li, J. R. McClean, and T. E. O’Brien, *Quantum error mitigation*, *Reviews of Modern Physics* **95**, 045005 (2023).
- [5] V. Russo, [Quantum error mitigation zoo](#), (2026).
- [6] W. J. Huggins, S. McArdle, T. E. O’Brien, J. Lee, N. C. Rubin, S. Boixo, K. B. Whaley, R. Babbush, and J. R. McClean, *Virtual distillation for quantum error mitigation*, *Physical Review X* **11**, 041036 (2021).
- [7] B. DalFavero and R. LaRose, *Error mitigation for partially error-corrected quantum computers*, (2025), [10.48550/arXiv.2510.10905](#), arXiv:2510.10905 [quant-ph].
- [8] J. R. McClean, M. E. Kimchi-Schwartz, J. Carter, and W. A. De Jong, *Hybrid quantum-classical hierarchy for mitigation of decoherence and determination of excited states*, *Physical Review A* **95**, 042308 (2017).
- [9] J. R. McClean, Z. Jiang, N. C. Rubin, R. Babbush, and H. Neven, *Decoding quantum errors with subspace expansions*, *Nature communications* **11**, 636 (2020).
- [10] W. J. Huggins, J. R. McClean, N. C. Rubin, Z. Jiang, N. Wiebe, K. B. Whaley, and R. Babbush, *Efficient and noise resilient measurements for quantum chemistry on near-term quantum computers*, *npj Quantum Information* **7**, 23 (2021).
- [11] N. Yoshioka, H. Hakoshima, Y. Matsuzaki, Y. Tokunaga, Y. Suzuki, and S. Endo, *Generalized quantum subspace expansion*, *Physical Review Letters* **129**, 020502 (2022).
- [12] E. Knill, *Quantum computing with realistically noisy devices*, *Nature* **434**, 39 (2005).
- [13] A. D. Córcoles, E. Magesan, S. J. Srinivasan, A. W. Cross, M. Steffen, J. M. Gambetta, and J. M. Chow, *Demonstration of a quantum error detection code using a square lattice of four superconducting qubits*, *Nature communications* **6**, 6979 (2015).
- [14] N. M. Linke, M. Gutierrez, K. A. Landsman, C. Figgatt, S. Debnath, K. R. Brown, and C. Monroe, *Fault-tolerant quantum error detection*, *Science advances* **3**, e1701074 (2017).
- [15] S. McArdle, X. Yuan, and S. Benjamin, *Error-mitigated digital quantum simulation*, *Physical review letters* **122**, 180501 (2019).
- [16] X. Bonet-Monroig, R. Sagastizabal, M. Singh, and T. O’Brien, *Low-cost error mitigation by symmetry verification*, *Physical Review A* **98**, 062339 (2018).
- [17] Z. Cai, *Quantum error mitigation using symmetry expansion*, *Quantum* **5**, 548 (2021).
- [18] H.-Y. Hu, R. LaRose, Y.-Z. You, E. Rieffel, and Z. Wang, *Logical shadow tomography: Efficient estimation of error-mitigated observables*, arXiv preprint arXiv:2203.07263 (2022).

- [19] M. Urbanek, B. Nachman, and W. A. de Jong, *Error detection on quantum computers improving the accuracy of chemical calculations*, *Physical Review A* **102**, 022427 (2020).
- [20] M. Gong, X. Yuan, S. Wang, Y. Wu, Y. Zhao, C. Zha, S. Li, Z. Zhang, Q. Zhao, Y. Liu, *et al.*, *Experimental exploration of five-qubit quantum error-correcting code with superconducting qubits*, *National Science Review* **9**, nwab011 (2022).
- [21] C. N. Self, M. Benedetti, and D. Amaro, *Protecting expressive circuits with a quantum error detection code*, *Nature Physics* **20**, 219 (2024).
- [22] E. Chertkov, A. C. Potter, D. Hayes, and M. Foss-Feig, *Error detection without post-selection in adaptive quantum circuits*, arXiv preprint arXiv:2509.25326 (2025).
- [23] S. Martiel and A. Javadi-Abhari, *Low-overhead error detection with spacetime codes*, arXiv preprint arXiv:2504.15725 (2025).
- [24] A. Javadi-Abhari, S. Martiel, A. Seif, M. Takita, and K. X. Wei, *Big cats: entanglement in 120 qubits and beyond*, arXiv preprint arXiv:2510.09520 (2025).
- [25] A. Vezvae, V. Tripathi, M. Morford-Oberst, F. Butt, V. Kasatkin, and D. A. Lidar, *Demonstration of high-fidelity entangled logical qubits using transmons*, *Nature Communications* (2026).
- [26] C. Zhang, C. Li, Z. Tian, Y. Jiang, F. Xu, S. Zhang, H. Wang, Y.-N. Zhang, X. Bai, B. Zhao, *et al.*, *Quantum error detection in a silicon quantum processor*, *Nature Electronics*, 1 (2026).
- [27] K. Temme, S. Bravyi, and J. M. Gambetta, *Error mitigation for short-depth quantum circuits*, *Physical Review Letters* **119**, 180509 (2017).
- [28] S. Endo, S. C. Benjamin, and Y. Li, *Practical quantum error mitigation for near-future applications*, *Physical Review X* **8**, 031027 (2018).
- [29] S. Zhang, Y. Lu, K. Zhang, W. Chen, Y. Li, J.-N. Zhang, and K. Kim, *Error-mitigated quantum gates exceeding physical fidelities in a trapped-ion system*, *Nature Communications* **11**, 587 (2020).
- [30] E. Van Den Berg, Z. K. Mineev, A. Kandala, and K. Temme, *Probabilistic error cancellation with sparse Pauli-Lindblad models on noisy quantum processors*, *Nature Physics*, 1 (2023).
- [31] J. Preskill, *Beyond nisq: The megaquop machine*, *ACM Transactions on Quantum Computing* **6**, 1–7 (2025), arXiv:2502.17368 [quant-ph].
- [32] C. Gidney, *Stim: a fast stabilizer circuit simulator*, *Quantum* **5**, 497 (2021).
- [33] A. D. Córcoles, E. Magesan, S. J. Srinivasan, A. W. Cross, M. Steffen, J. M. Gambetta, and J. M. Chow, *Demonstration of a quantum error detection code using a square lattice of four superconducting qubits*, *Nature Communications* **6**, 6979 (2015).
- [34] M. Urbanek, B. Nachman, and W. A. d. Jong, *Error detection on quantum computers improves accuracy of chemical calculations*, *Physical Review A* **102**, 022427 (2020), arXiv:1910.00129 [quant-ph].
- [35] M. Gong, X. Yuan, S. Wang, Y. Wu, Y. Zhao, C. Zha, S. Li, Z. Zhang, Q. Zhao, Y. Liu, F. Liang, J. Lin, Y. Xu, H. Deng, H. Rong, H. Lu, S. C. Benjamin, C.-Z. Peng, X. Ma, Y.-A. Chen, X. Zhu, and J.-W. Pan, *Experimental exploration of five-qubit quantum error correcting code with superconducting qubits*, *National Science Review* **9**, nwab011 (2022), arXiv:1907.04507 [quant-ph].
- [36] C. N. Self, M. Benedetti, and D. Amaro, *Protecting expressive circuits with a quantum error detection code*, *Nature Physics* **20**, 219–224 (2024), arXiv:2211.06703 [quant-ph].
- [37] E. Chertkov, A. C. Potter, D. Hayes, and M. Foss-Feig, *Error detection without post-selection in adaptive quantum circuits*, (2025), 10.48550/arXiv.2509.25326, arXiv:2509.25326 [quant-ph].
- [38] S. Martiel and A. Javadi-Abhari, *Low-overhead error detection with spacetime codes*, (2025), 10.48550/arXiv.2504.15725, arXiv:2504.15725 [quant-ph].
- [39] A. Javadi-Abhari, S. Martiel, A. Seif, M. Takita, and K. X. Wei, *Big cats: entanglement in 120 qubits and beyond*, (2025), 10.48550/arXiv.2510.09520, arXiv:2510.09520 [quant-ph].
- [40] A. Vezvae, V. Tripathi, M. Morford-Oberst, F. Butt, V. Kasatkin, and D. A. Lidar, *Demonstration of high-fidelity entangled logical qubits using transmons*, (2025), 10.48550/arXiv.2503.14472, arXiv:2503.14472 [quant-ph].
- [41] C. Zhang, C. Li, Z. Tian, Y. Jiang, F. Xu, S. Zhang, H. Wang, Y.-N. Zhang, X. Bai, B. Zhao, Y.-F. Zhang, H. Shu, J. Liu, K. Wu, C. Huang, K. Shi, M. Duan, T. Xin, P. Huang, T. Pan, S. Liu, G. Wang, G. Hu, Y. He, and D. Yu, *Quantum error detection in a silicon quantum processor*, *Nature Electronics*, 1–9 (2026).
- [42] J. R. McClean, Z. Jiang, N. C. Rubin, R. Babbush, and H. Neven, *Decoding quantum errors with subspace expansions*, arXiv:1903.05786 [physics, physics:quant-ph] (2019), arXiv: 1903.05786.
- [43] Z. Cai, *Quantum error mitigation using symmetry expansion*, *Quantum* **5**, 548 (2021).
- [44] C. Chamberland, A. Kubica, T. J. Yoder, and G. Zhu, *Triangular color codes on trivalent graphs with flag qubits*, *New Journal of Physics* **22**, 023019 (2020).
- [45] S. M. Girvin, *Introduction to quantum error correction and fault tolerance*, *SciPost Physics Lecture Notes* (2023), 10.21468/SciPostPhysLectNotes.70.
- [46] V. V. Albert and P. Faist, in *The Error Correction Zoo*, edited by V. V. Albert and P. Faist (errorcorrection-zoo.org, 2023).
- [47] H. Bombin, *An introduction to topological quantum codes*, (2013), arXiv:1311.0277 [quant-ph].
- [48] V. V. Albert and P. Faist, in *The Error Correction Zoo*, edited by V. V. Albert and P. Faist (errorcorrection-zoo.org, 2024).
- [49] R. LaRose, rmlarose/encoded github, <https://github.com/rmlarose/encoded> (2026).

Appendix A: Derivations

To see why (4) is true, the error-mitigated value is defined as

$$\langle \bar{O} \rangle(n) := \frac{\text{Tr}[\Pi \mathcal{E}_p(\bar{\rho}) \Pi^\dagger \bar{O}]}{\text{Tr}[\Pi \mathcal{E}_p(\bar{\rho}) \Pi^\dagger]}, \quad (\text{A1})$$

where \mathcal{S} is a generating set for the stabilizer group, $\Pi := \prod_{S \in \mathcal{S}} (I + S)/2$ is the projector onto the codespace, and $\mathcal{E}_p := pI/d + (1-p)\bar{\rho}$ is the action of rate p depolarizing noise on the encoded state $\bar{\rho} := |0^{\otimes n}\rangle\langle 0^{\otimes n}|$. Using

the facts that $[\Pi, \bar{O}] = 0$, $\Pi = \Pi^\dagger$, and $\Pi^2 = \Pi$, the numerator of (A1) is

$$\text{Tr}[\Pi \mathcal{E}_p(\bar{\rho}) \Pi^\dagger \bar{O}] = \frac{1-p}{d} \text{Tr}[\Pi \bar{O}] + p \text{Tr}[\bar{\rho} \Pi \bar{O}] = p \text{Tr}[\bar{\rho} \bar{O}]. \quad (\text{A2})$$

The second equality follows because every term in $\Pi \bar{O}$ is a Pauli so that $\text{Tr}[\Pi \bar{O}] = 0$, and because $\bar{\rho}$ is a codeword

so that $\bar{\rho} \Pi = \bar{\rho}$. Similarly, the denominator of (A1) is

$$(1-p) \text{Tr}[\Pi]/d + p \text{Tr}[\bar{\rho} \Pi] = (1-p)/2^{n-1} + p. \quad (\text{A3})$$

This follows because $\Pi = \frac{1}{2^{n-1}} \sum_{S \in \mathcal{S}} S$ and every non-identity Pauli is traceless, so $\text{Tr}[\Pi] = \frac{1}{2^{n-1}} \text{Tr}[I] = \frac{2^n}{2^{n-1}} = 2$, giving $\text{Tr}[\Pi]/d = 1/2^{n-1}$. Similarly, $\text{Tr}[\bar{\rho} \Pi] = 1$ since $\bar{\rho}$ is a codeword. Combining the numerator and denominator yields (4).