Fault tolerance with bare ancillary qubits for a [[7,1,3]] code

Muyuan Li, Mauricio Gutiérrez, Stanley E. David, Alonzo Hernandez, and Kenneth R. Brown

Fault Tolerance with Bare Ancillary Qubits for a $[[7, 1, 3]]$ Code

Muyuan Li, Mauricio Gutiérrez, Stanley E. David, Alonzo Hernandez, and Kenneth R. Brown[*]

Schools of Computational Science and Engineering; Chemistry and Biochemistry; and Physics
Georgia Institute of Technology, Atlanta, GA 30332-0400

(Dated: August 9, 2017)

We present a $[[7, 1, 3]]$ quantum error-correcting code that is able to achieve fault-tolerant syndrome measurement using one ancillary qubit per stabilizer for an error model of independent single-qubit Pauli errors.. All single-qubit Pauli errors on the ancillary qubits propagate to form exclusively correctable errors on the data qubits. The situation changes for error models with two-qubit Pauli errors. We compare the level-1 logical error rates under two noise models: the standard Pauli symmetric depolarizing error model and an anisotropic error model. The anisotropic model is motivated by control errors on two-qubit gates commonly applied to trapped ion qubits. We find that one ancillary qubit per syndrome measurement is sufficient for fault-tolerance for the anisotropic error, but is not sufficient for the standard depolarizing errors. We then show how to achieve fault tolerance for the standard depolarizing errors by adding flag qubits to check for errors on select ancilary qubits. Our results on this $[[7, 1, 3]]$ code demonstrates how physically motivated noise models may simplify fault-tolerent protocols.

I. INTRODUCTION

The operation of a universal quantum computer requires protocols to protect it from instrumental and environmental imperfections. Quantum error-correcting (QEC) [1] codes are among the most promising approaches to deliver scalable and reliable fault-tolerant quantum computation. Topological QEC codes have a two-dimensional qubit layout that limits qubit interactions to only nearest neighbors, which is suitable for practical implementation using quantum architectures such as superconducting devices and trapped ions [2–4]. Concatenated codes require more resources to be mapped onto local architectures, but can use less resources overall than topological codes for low error rate systems [5]. In general, the optimal choice of quantum error-correcting code is dependent on the actual noise environment and the quantum circuit being used [6]. Many QEC codes, both topological codes and concatenated codes, belong to the family of Calderbank-Shor-Steane (CSS) codes, where X and Z Pauli errors can be treated separately [7, 8].

For a given QEC code, different methods of syndrome measurement, state preparation and decoding can all affect the resulting logical error rate. For syndrome measurement, using a single (bare) ancillary qubit to measure a stabilizer is in general not fault-tolerant, since errors on the ancillary qubit can propagate to the data and form uncorrectable errors [9, 10]. Although there are various methods to make the syndrome measurement fault-tolerant, these come at the expense of extra resources. Shor's method requires a w-qubit cat state to measure a weight-w stabilizer [11]. An extra qubit is needed to verify the ancillary qubit, but this is not a strict requirement [12, 13]. Steane's method requires the fault-tolerant preparation of a logical state [14], while Knill's method relies on the fault-tolerant preparation of a logical Bell pair [15].

Using bare ancillary qubits for stabilizer measurement can be fault-tolerant if we are guaranteed that single-qubit errors or errors that occur with a probability linearly proportional to the physical error rate do not propagate to form uncorrectable errors. More specifically, if we assume only Pauli errors, then after measuring a stabilizer of weight w a weight $\lfloor w/2 \rfloor$, where $\lfloor x \rfloor$ is the floor of x, error can propagate to the data. This can still be fault-tolerant in several cases. In some codes, the presence of gauge subsystems allow the decomposition of high weight stabilizers into lower weight gauge operators that can be measured in a fault-tolerant fashion using bare ancillary qubits. This is the case of the Bacon-Shor codes [16, 17]. In other cases, codes have a large enough distance such that any error of weight up to $\lfloor w/2 \rfloor$ can be corrected ($d \geq 2\lfloor w/2 \rfloor + 1$). This is the case of large distance surface [3, 18] and color codes [19, 20]. Finally, certain codes with distance $d < 2\lfloor w/2 \rfloor + 1$ allow for stabilizer measurement with bare ancillary qubits because the resulting errors on data qubits are correctable. This is the case for the $[[9, 1, 3]]$ surface code, where the weight-2 errors that propagate when measuring the weight-4 stabilizers are all correctable for specific orderings of the entangling two-qubit gates [21].

In this paper, we present a non-CSS $[[7, 1, 3]]$ QEC code that falls into the latter category: every single-qubit Pauli error on the ancillary qubit propagates to the data qubits to form a correctable error. The code was found by a numerical greedy search for stabilizer codes where single-qubit errors on the ancillary qubits do not lead to a logical error [22]. We refer to this code as the Bare $[[7, 1, 3]]$ code, and we refer to the syndrome measurement method of one ancillary qubit per stabilizer as the bare method.

Given the guarantee that when performing the bare method any single-qubit Pauli error on the ancillary qubits would not lead to a logical error, we want to understand the behavior of the code under higher weight errors. Here we present the results of Monte Carlo stabilizer simulations of the Bare $[[7, 1, 3]]$ code with 3 error correcting steps. We used CHP [23] and a Python-based wrapper to obtain logical error rates for the Bare $[[7, 1, 3]]$ code at the first level of encoding. We

report the logical error rates under two different error models. We find that with the bare method, fault-tolerance is not achieved under the standard depolarizing error model because, although the code is resilient to any single-qubit error on the ancillary qubit, certain two-qubit errors whose probability is linearly proportional to the physical error rate are malignant. To protect our code against these errors, we present how to use two additional flag qubits [24] to achieve fault-tolerant syndrome measurement, which we refer to as the flag method, and report the resulting level-1 pseudothreshold. The bare method is still fault-tolerant under an anisotropic error model applicable to trapped-ion qubits. We report the level-1 pseudothreshold of the code under this error model. In the Appendix we present two alternative Monte Carlo error sampling schemes that we employed in the simulations, and we report the level-1 pseudothreshold for two well studied distance-3 quantum error correcting codes: the Steane code [25–27] and the 5-qubit code [28, 29], for comparison.

The paper is organized as follows. Section II introduces the Bare $[[7,1,3]]$ code, and explains how it handles different errors. Section III describes the two noise models under which we study the performance of the Bare $[[7,1,3]]$ code and demonstrates why the bare method for this error-correcting code is not fault-tolerant under the standard depolarizing noise model. Section IV describes how to use an additional flag qubit to achieve fault-tolerant syndrome measurement for the Bare $[[7,1,3]]$ code. In Section V, we present the simulation scheme and how the simulation results are used to calculate the level-1 pseudothreshold. Section VI presents the results obtained from our simulation for the Bare $[[7,1,3]]$ code under two error models with two different syndrome measurement methods.

II. DETAILS OF THE BARE $[[7,1,3]]$ CODE

The Bare $[[7,1,3]]$ code was found through a numerical greedy search of stabilizer codes with the property that single-qubit errors on the bare ancillary qubit would not lead to a logical error. Table I presents the stabilizer generators and logical X and Z operators of the code. As seen from the stabilizers, the code is non-CSS and degenerate.

Stabilizer Generators	Logical Operators
$X_0 X_4$	
$X_1 X_4$	
$X_2 X_5$	$X_L = X_1 X_2 X_3$
$X_3 X_6$	$Z_L = Z_0 Z_1 Z_4$
$Z_2 Z_3 Y_5 Y_6$	
$Z_0 Z_1 Z_2 X_3 Z_4 Z_5$	

TABLE I: List of stabilizers and logical operators X_L, Z_L for the Bare $[[7,1,3]]$ code.

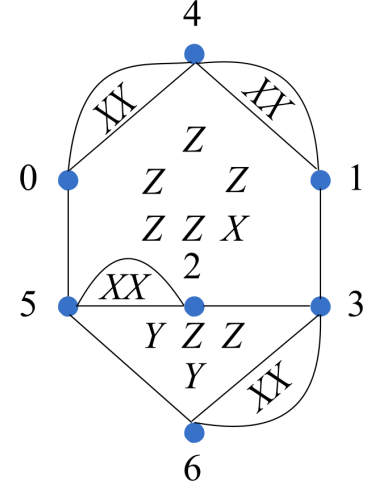FIG 1 shows the configuration of the stabilizers when the Bare $[[7,1,3]]$ code is embedded in a plane.

FIG. 1: The Bare $[[7,1,3]]$ code embeded in a plane. Each vertex represents a data qubit, and each stabilizer generator of the code corresponds to a face in this graph.
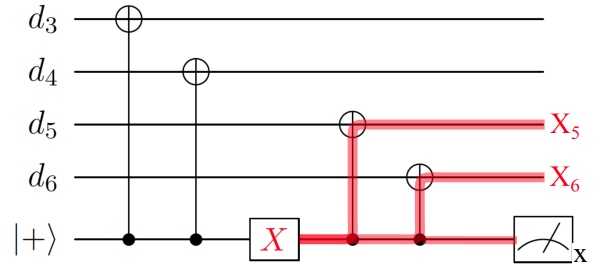

FIG. 2: For the Steane $[[7,1,3]]$ code with bare ancillary qubits, a single-qubit error on the ancillary qubit would lead to a logical error.

A. Single-qubit errors on a bare ancillary qubit

To illustrate how using bare ancillary qubits can lead to uncorrectable errors in some QEC codes, let us consider the $[[7,1,3]]$ Steane code. From Figure 2, we can see that when measuring the stabilizer $X_3 X_4 X_5 X_6$, a single qubit X error on the bare ancillary qubit propagates to form the error $X_5 X_6$ on the data qubits. Although this error is detectable at a later stage, its syndrome is equivalent to an X_0 error. When this correction is applied, the resulting error $X_0 X_5 X_6$ is equivalent to the logical operator $X_L = X_0 X_1 X_2 X_3 X_4 X_5 X_6$ up to a stabilizer.

On the other hand, in the Bare $[[7,1,3]]$ code, all single-qubit errors on the ancillary qubits propagate to become errors with unique syndromes, hence correctable. This is achieved by considering different schedules for coupling the data to the ancillary qubit when measuring the stabilizers. Note that all ancillary qubits are prepared in the $|+\rangle$ state.

$Z_0 \to 100000$	$X_0 \to 000001$	$Y_0 \to 100001$
$Z_1 \to 010000$	$X_1 \to 000001$	$Y_1 \to 010001$
$Z_2 \to 001000$	$X_2 \to 000011$	$Y_2 \to 001011$
$Z_3 \to 000101$	$X_3 \to 000010$	$Y_3 \to 000111$
$Z_4 \to 110000$	$X_4 \to 000001$	$Y_4 \to 110001$
$Z_5 \to 001010$	$X_5 \to 000011$	$Y_5 \to 001001$
$Z_6 \to 000110$	$X_6 \to 000010$	$Y_6 \to 000100$

TABLE II: Single-qubit error syndromes for the Bare $[[7,1,3]]$ code.
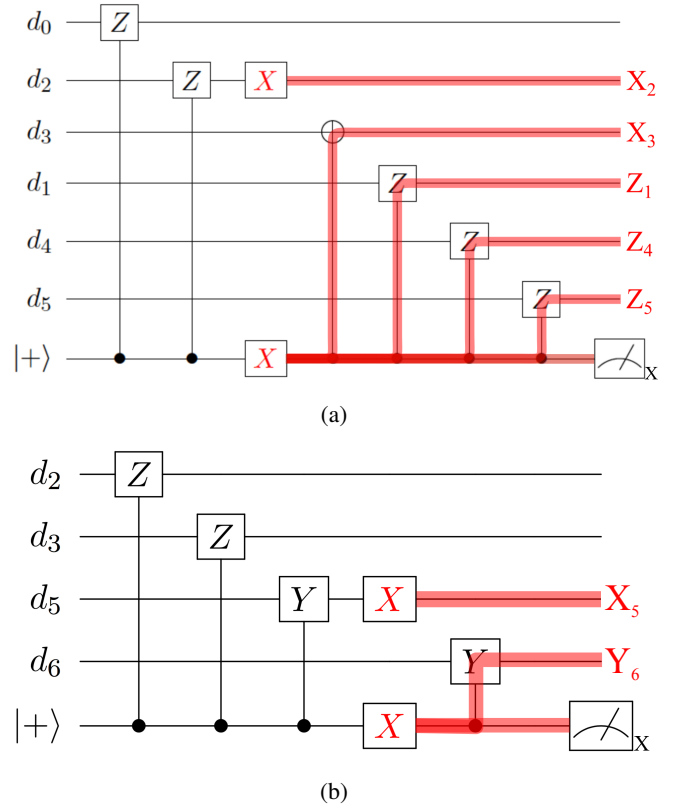

(a)

(b)

FIG. 3: Two examples of 2-qubit XX errors that propagate to become uncorrectable errors. In (a) the error leads to an X_L, while in (b) the error results in a Z_L.

The syndromes for the 21 single-qubit Pauli errors are shown in Table II. Notice that the syndromes of $Z_2 Z_3$ is 001101 which is distinct from all syndromes of single-qubit errors. Similarly, $Z_0 Z_2 \to 101000$, $Z_0 Z_2 X_3 \to 101010$, and $Z_4 Z_5 \to 111010$ all have unique syndromes. Since each gate used in a measurement acts between an ancillary qubit and a data qubit where the ancillary qubit controls a Pauli operator on the data qubit, this observation suggests the following syndrome measurement coupling schedule:

1. For each weight-2 generator, the measurement gates can be coupled to the ancillary qubit in any order.

2. For the stabilizer $Z_2 Z_3 Y_5 Y_6$, couple the measurement gates in left-to-right order.

3. For the stabilizer $Z_0 Z_1 Z_2 X_3 Z_4 Z_5$, couple the measurement gates in order of Z_0, Z_2, X_3, Z_1, Z_4, Z_5.

By using this coupling schedule, for the stabilizer $Z_2 Z_3 Y_5 Y_6$, a single-qubit error on the ancillary qubit can propagate the error $Z_2 Z_3$ onto the data qubits, which has a unique syndrome; for the stabilizer $Z_0 Z_1 Z_2 X_3 Z_4 Z_5$, a single-qubit error on the ancillary qubit can propagate errors $Z_0 Z_2$, $Z_0 Z_2 X_3$, and $Z_4 Z_5$ onto the data qubits, which all have unique syndromes. Therefore, all single-qubit errors on the ancillary qubits propagate to exclusively correctable errors.

B. Two-qubit errors on a bare ancillary qubit

The Bare $[[7,1,3]]$ code is vulnerable to "hook errors". Figure 3a illustrates an instance of a 2-qubit Pauli error that can occur under the standard symmetric depolarizing error model. This XX error propagates to become $Z_1 X_2 X_3 Z_4 Z_5$, which has syndrome 101011, the same syndrome as $Y_1 Z_4 Z_5$. Thus, the correction $Y_1 Z_4 Z_5$ will be applied, and the resulting error is $X_1 X_2 X_3$, the logical X operator.

In Fig.3b the 2-qubit XX error on the control-Y gate propagates to become the error $X_5 Y_6$, which has syndrome 000111, the same as Y_3. After the correction Y_3 is applied, the resulting error $Y_3 X_5 Y_6$ is equivalent to the logical Z operator up to a stabilizer.

III. ERROR MODELS

A. Error Models

To study the properties of the Bare $[[7,1,3]]$ code, we applied two Pauli error models: the standard depolarizing error model and an anisotropic error model.

1. Standard Depolarizing Error Model

The standard depolarizing error model is a common quantum channel where, after each gate, there is a symmetric depolarization with some probability p. After single-qubit gates (including Pauli gates and ancillary qubit preparation gates), a traceless single-qubit Pauli operator, randomly selected from $P = \{X, Y, Z\}$, is applied with probability p_s. After two-qubit gates, a traceless two-qubit Pauli operator, randomly selected from $Q = \{I, X, Y, Z\} \otimes \{I, X, Y, Z\} \setminus \{II\}$, is applied with probablity p_t. Although in general $p_s \neq p_t$, it is common for them to be simulated with the same value [15]. A measurement reports the wrong result with probability p_{meas}.

For the cases studied here, p_{meas} is set to the single qubit error probability p_s.

2. Anisotropic Error Model

For single qubit gates, preparation, and measurement errors remain the same as the standard depolarizing error model. The conceptual motivation for the anisotropic error model is that two-qubit gate errors occur due to errors in the gate coupling. As an example, consider errors in the entangling Mølmer-Sørensen (MS) gate used in trapped ions [30]. The MS gate corresponds to $\exp\left(-i(\pi/4)XX\right)$, and an over or under-rortation could result in the coherent error $\exp\left(-i(\epsilon/2)XX\right)$, where ϵ denotes the over- or under-rotation angle. For ϵ randomly and symmetrically distributed around 0, this error can be described as a random application of the Pauli error XX with probability p_t, which is determined by the distribution of θ.

The key idea is that the two-qubit correlation in the error is not any two-qubit error, but only the two-qubit Pauli error aligned with the gate. The error after a Control-P gate is then ZP with probability p_t. For the specific case of ion traps, this can be derived by including the single qubit rotations necessary to transform the MS gate into a Control-P gate. In this study, we do not include the individual rotations but work at the level of Control-P gates. We expect random errors on the individual qubits in addition to the control error from the two-qubit gate. To account for the random errors, after every two-qubit gate we apply the two-qubit anisotropic error with probability p_t followed by single qubit errors with probability p_s on the two qubits involved in the gate.

B. Fault-tolerance Dependent on Error Model

When performing the bare method under the two error models defined in Section III A, we can see that all two-qubit errors in the anisotropic error model can be detected and corrected, because a Z error on the control qubit does not propagate to form other errors on the data. However, for certain two-qubit errors in the standard depolarizing error model, such as the XX error as seen in Fig. 3a and Fig. 3b, the errors propagate to form uncorrectable hook errors. Therefore, the Bare $[[7, 1, 3]]$ code with the bare method can achieve fault-tolerance under the anisotropic error model because all errors that occur with probability linear in the error rate of physical operations are correctable; however, the same syndrome measurement method cannot achieve fault-tolerance under the standard depolarizing error model since certain two-qubit errors that occur with probability linear in the physical error rate propagate to become uncorrectable logical errors.

IV. FAULT-TOLERANCE WITH A FLAG QUBIT

Since the bare method is not fault-tolerant for the Bare $[[7, 1, 3]]$ code under the standard depolarizing error model, additional resources would be required to ensure fault-tolerance. In this section we propose the method of using two additional flag qubits to perform fault-tolerant syndrome measurement on the Bare $[[7, 1, 3]]$ code. The idea of the flag method was used by Yoder and Kim [4], and presented in detail by Chao and Reichardt [24].

Since the source of logical errors for the Bare $[[7, 1, 3]]$ code are the hook errors propagated from errors on the ancillary qubit, we can use flag qubits to detect errors on the ancillary qubit that can propagate to form these hooks. In order to use the flag method, we need to change the order of the gates to measure the stabilizer $Z_0 Z_1 Z_2 X_3 Z_4 Z_5$ to: $Z_0, X_3, Z_4, Z_2, Z_1, Z_5$. The purpose of this order change is to ensure that all errors that could trigger the flag qubit measurement have distinct syndromes. As we can see in TABLE IV, this change of order makes the Bare [[7,1,3]] code vulnerable to certain single-qubit errors that were correctable with the old order. But since these errors all trigger the flag qubit measurement and all result in distinct syndromes, they are still correctable.

The circuits to fault-tolerantly measure the high-weight stabilizers are shown in Figure 4. The second ancillary qubit initialized in the $|0\rangle$ state in both circuits is the flag qubit. With no errors present, these circuits behave exactly the same as the circuits using bare ancillary qubits for syndrome measurement, and the Z-basis measurements on the flag qubits will always give $|0\rangle$. In Figure 4a errors after gate a and d cannot create uncorrectable hook errors, and in Figure 4b errors after gate a and f cannot create uncorrectable hook errors. Errors after all other gates in these two circuits will be detected by causing a $|1\rangle$ measurement outcome for the flag qubit. These errors and their syndromes are listed in Table III and Table IV. Each of these errors has a distinct syndrome, so it can be corrected. Note that Y errors on the ancilla qubit have the same effect on the data as X errors.

As argued by Chao and Reichardt [24], this method of syndrome measurement is fault-tolerant. Under the standard depolarizing model, all errors that appear with probability linear in the physical error rate can be detected and corrected using the above flag method, hence making the Bare $[[7, 1, 3]]$ code fault-tolerant under the standard depolarizing error model.

V. SIMULATION SCHEME

In this section we present the circuit used in the simulation, the QEC scheme used to perform correction after each round of the Bare $[[7, 1, 3]]$ code with noise, and the calculations we performed to obtain logical error rates. The simulations were

b errors	Data error	syndrome	c errors	Data error	syndrome
IX	Y_5Y_6	001101	IX	Y_6	000100
XX	$X_3Y_5Y_6$	001111	XX	X_5Y_6	000111
YX	$Y_3Y_5Y_6$	001010	YX	Y_5Y_6	001101
ZX	$Z_3Y_5Y_6$	001000	ZX	Z_5Y_6	001110

TABLE III: The nontrivial data errors that can result from a single two-qubit error for the stabilizer $Z_2Z_3Y_5Y_6$ (Figure 4a) and trigger the flag qubit measurement. The errors marked in red can lead to logical errors with the bare method.

b errors	Data error	syndrome	c errors	Data error	syndrome
IX	$Z_1Z_2Z_4Z_5$	100010	IX	$Z_1Z_2Z_5$	010010
XX	$Z_1Z_2X_3Z_4Z_5$	100000	XX	$Z_1Z_2X_4Z_5$	010011
YX	$Z_1Z_2Y_3Z_4Z_5$	100101	YX	$Z_1Z_2Y_4Z_5$	100011
ZX	$Z_1Z_2Z_3Z_4Z_5$	100111	ZX	$Z_1Z_2Z_4Z_5$	100010

d errors	Data error	syndrome	e errors	Data error	syndrome
IX	Z_1Z_5	011010	IX	Z_5	001010
XX	$Z_1X_2Z_5$	011001	XX	X_1Z_5	001011
YX	$Z_1Y_2Z_5$	010001	YX	Y_1Z_5	011011
ZX	$Z_1Z_2Z_5$	010010	ZX	Z_1Z_5	011010

TABLE IV: The nontrivial data errors that can result from a single two-qubit error for the stabilizer $Z_0Z_1Z_2X_3Z_4Z_5$ (Figure 4b) and trigger the flag qubit measurement. The errors marked in red can lead to logical errors with the bare method.

(a)

(b)

FIG. 4: Flagged syndrome measurement for the Bare $[[7,1,3]]$ code. (a) Circuit to fault-tolerantly measure the syndrome of the stabilizer $Z_2Z_3Y_5Y_6$ using a flag qubit. (b) Circuit to fault-tolerantly measure the syndrome of the stabilizer $Z_0Z_1Z_2X_3Z_4Z_5$ using a flag qubit. Notice the order of gates is different from what we used in the the bare method.

done using the CHP stabilizer simulator [23]. The simulation follows the 1-rectangle (1-Rec) formalism [31] for distance-3 and level-1 encoding, where the 1-Ga we are simulating is the level-1 logical identity gate.

A. Simulation Circuit

Fig. 5 shows the quantum circuit we constructed to simulate one error-correcting step of the bare method for the Bare $[[7,1,3]]$ code. The simulated gate is a logical identity followed by an error-correction gadget. The gates are grouped together for each ancillary qubit measurement, and the order of qubit-coupling between the data qubits and ancillary qubits follows exactly the scheduling as required for fault-tolerant measurements of each stabilizer generators. For each gate shown in this circuit, errors of appropriate type are inserted after it with a probability given by the noise model. The circuit for simulating the flag method requires two additional flag qubits, and has a different order of qubit-coupling between the data and ancillary qubits, as described in Section IV.

The syndrome measurement is repeated up to three times with freshly prepared ancillary qubits. The repetition is necessary due to errors during or between syndrome measurements [32]. If the first two error syndromes agree, correction is performed based on the error syndrome. If the two syndromes
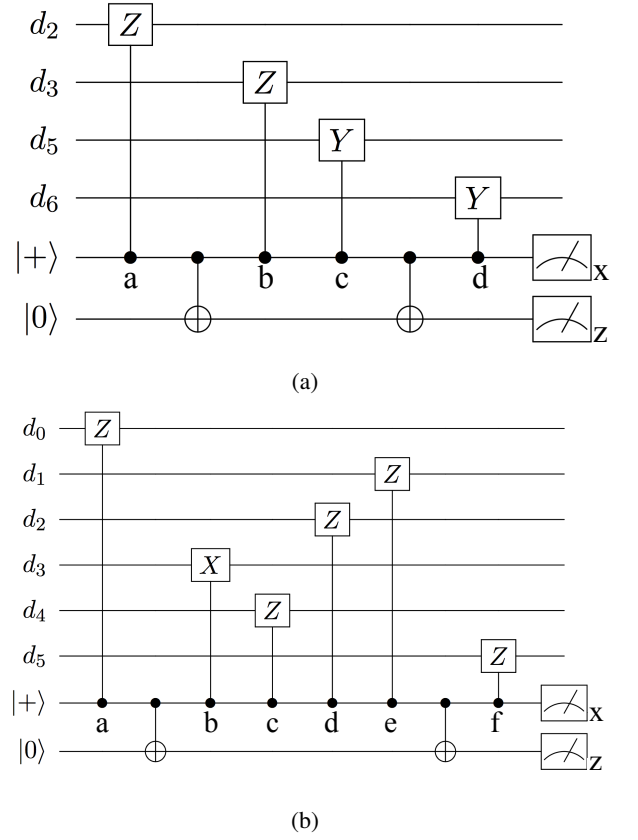
disagree, a third syndrome is measured and correction is performed based on the third syndrome.

B. Logical Error Rate Calculation

We calculate the logical error rate under the two different error models for various physical error strengths. For each run of in the Monte Carlo simulation, we initialize all the data qubits in state $|0\rangle$, 6 of the ancillary qubits in state $|+\rangle$ for syndrome measurements, and 2 of the ancillary qubits in state $|0\rangle$ to be used as flag qubits for the high-weight stabilizers. We then perform one round of noise-free stabilizer measurements to project the state of the data qubits to the logical $|0\rangle$ state. Then the simulation proceeds as follows:

1. Perform 2 or 3 rounds of error correction with random errors inserted using the importance sampling scheme.
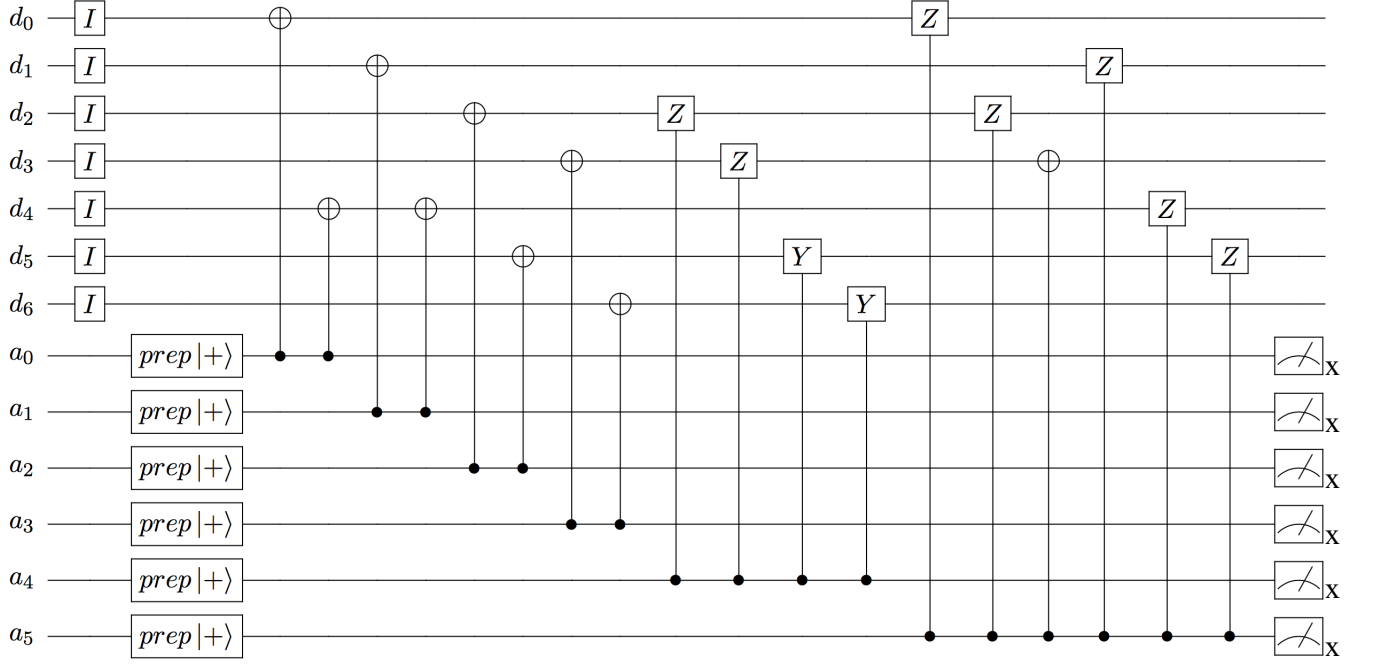
FIG. 5: The circuit simulating the Bare $[[7, 1, 3]]$ code and its measurement steps. The circuit is constructed with only Clifford gates, and errors (single-qubit and two-qubit) are inserted after ancillary qubit preparation, CNOT, Control-Z, Control-Y, and measurements.

2. Apply the decoder to determine the corresponding error configuration and correct accordingly.

3. Perform noise-free correction to the final state to project the state back to the codespace.

4. If the final state is different from the initial one, count as one logical error.

In the importance sampling scheme (see Appendix A.2), we classify the error configurations into subsets according to the number of single- and two-qubit errors present in the configuration. The subset consisting of configurations of s single-qubit errors and t two-qubit errors is labelled by (s, t). For example, if during the execution of the QEC circuit, two single-qubit preparations were faulty and an error occurred after one of the CNOTs, this error configuration belongs to the subset $(2, 1)$.

We perform Monte Carlo simulations on selected error subsets and compute the logical error rate per subset, $p_L(s, t)$, by calculating the ratio of the successful runs over the total number of runs. The total logical error rate, p_L, is calculated as a weighted average over the selected error subsets:

$$p_L(p_s, p_t) = \sum_{s,t} A_{s,t}(p_s, p_t) p_L(s, t),$$

where $A_{s,t}$ is the probability of occurrence of subset (s, t), that is, the total probability of occurrence of all error configurations with s single-qubit errors and t two-qubit errors.

We use the calculated logical error rates at different physical error strengths to estimate the pseudothreshold of the code under a particular error model [33]. The pseudothreshold is the intersection between the physical error rate line $y = 2p/3$ and the logical error rate $p_L(p_s, p_t)$. The reason for using this line instead of $y = p$ is that, if we assume a symmetric depolarizing noise model on a single qubit, the infidelity is given by $2p/3$. More intuitively, if we focus on a single-qubit Pauli state, like $|0\rangle$, only two Pauli errors (X and Y in this case) will cause an error. Furthermore, this will result in a slightly more pessimistic value for the pseudothreshold.

VI. RESULTS

In this section we present and analyze the results from numerical Monte Carlo simulations of the Bare $[[7, 1, 3]]$ code under different noise models using the two different syndrome measurement methods. We first focus on the standard depolarizing error model and then on the anisotropic error model. We show that using the bare method, the Bare $[[7, 1, 3]]$ code does not have a pseudothreshold under the standard depolarizing error model, due to some uncorrectable two-qubit errors that have probability linear in the physical error strength.

A. Standard Depolarizing Error Model

Fig. 6 presents the logical error rates for the standard depolarizing model with $p_s = p_t = p$ for several values of p. For the bare method, the best fit corresponds to a function with linear term as leading term, while for the flag method the best fit corresponds to a function with quadratic term as leading term.

For the bare method, although there is an intersection between the logical error rate and the physical error rate, the curves still remain parallel for physical error rates below the intersection. The fact that there is still a linear term in the logical error rate implies that we would not observe an exponential suppression on the logical error rate with subsequent levels of concatenation. Thus, this intersection is not a pseudothreshold for the Bare $[[7, 1, 3]]$ code, and the code is not fault-tolerant under the standard depolarizing error model. Examples of why this would happen are shown in Fig. 3.

For the flag method, all errors that occur with probability linearly proportional to the physical error rate can be corrected. By performing a quadratic fit for the logical error rates and computing the intersection of the fitting curve with the error rate of an unencoded qubit, we obtain a level-1 pseudothreshold of 1.08×10^{-3}.
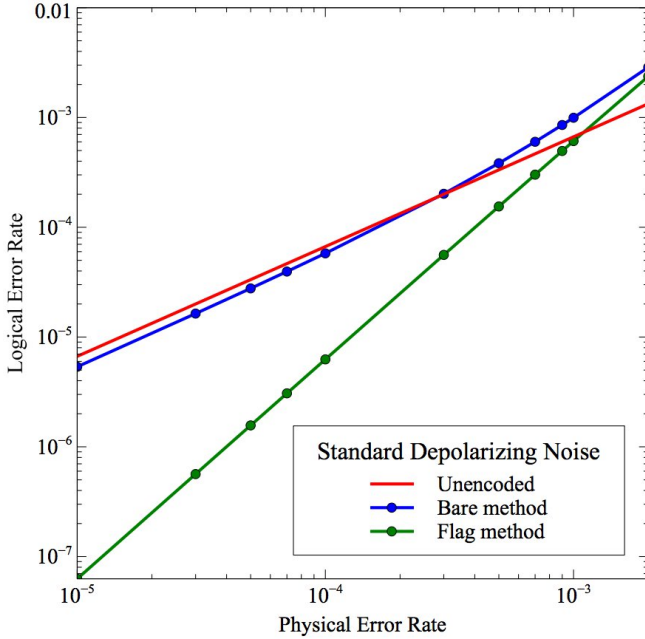

FIG. 6: Logical error rate for the Bare $[[7, 1, 3]]$ code under the standard error model. For the bare method, the logical error rate of the code remains parallel to the physical line for physical error rates below the intersection. This implies that the logical error rate is linear in p. For the flag method, we observe a level-1 pseudothreshold of 1.08×10^{-3}.

B. Anisotropic Error Model

We now calculate the pseudothreshold of the Bare $[[7, 1, 3]]$ code under the anisotropic error model. In this model, the physical error strength after each gate is p, and $p_s = p_t = p$.
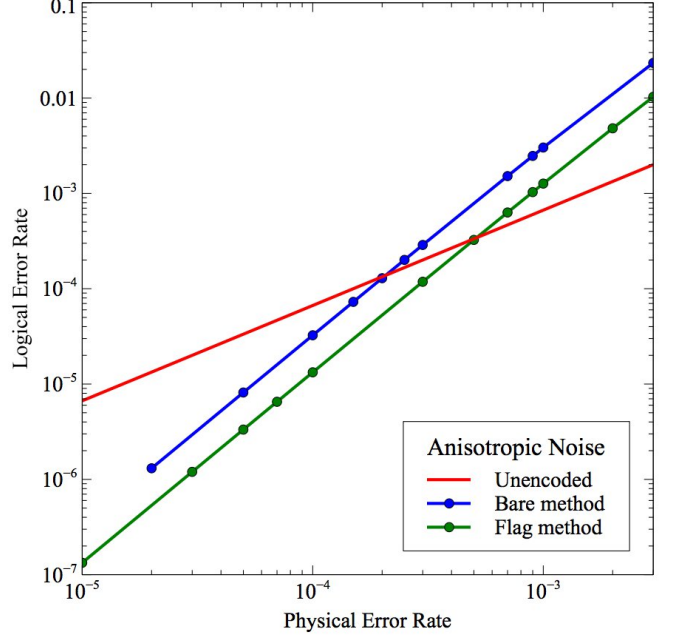

FIG. 7: Logical error rate for the Bare $[[7, 1, 3]]$ code under the anisotropic error model, $p_s = p_t$. For the bare method, the level-1 pseudothreshold is at $p = 2.0 \times 10^{-4}$. For the flag method, the level-1 pseudothreshold is at $p = 4.0 \times 10^{-4}$. The Bare $[[7, 1, 3]]$ code is fault-tolerant under the anisotropic error model with both syndrome measurement methods.

By performing a quadratic fit for the logical error rates and computing the intersection of the fitting curve with the error rates of an unencoded qubit, we obtain level-1 pseudothreshold for the Bare $[[7, 1, 3]]$ code with both syndrome measurement methods. This is because all single-qubit and two-qubit errors that occur with probability linear in p are guaranteed to be detected and corrected by the Bare $[[7, 1, 3]]$ code. From Fig. 7 we can see that for the bare method, the level-1 pseudothreshold is at 2.0×10^{-4}. For the flag method, the level-1 pseudothreshold is at 4.0×10^{-4}.

The main reason for the difference between the two pseudothresholds is that the number of uncorrectable error configurations with probability quadratically proportional to the physical error rate is higher for the bare method. Consider, for example, that some errors with probability quadratically proportional to the physical error rate can behave the same as the two-qubit errors listed in Table III and Table IV. Those errors could lead to logical errors with the bare method, but can be corrected with the flag method.

C. Comparison of Results across Error Models

From the results presented in Fig. 6 and Fig. 7 we can see that the level-1 pseudothreshold under the anisotropic error model is lower than that under the standard depolarizing error model. This is because in the anisotropic error model, after each two-qubit gate we are simulating a single-qubit random Pauli error on each qubit with probability p_s (Section III A), leading to more possible error locations in the overall circuit. If we assume the single-qubit errors and two-qubit errors are equally probable, as we did in the results shown in Fig. 6 and Fig. 7, errors after the two-qubit gates are 3 times more likely to occur under the anisotropic error model than under the standard depolarizing model, because of the extra single-qubit Pauli errors. Hence, the Bare $[[7, 1, 3]]$ code under the anisotropic error model has a lower pseudothreshold than under the standard depolarizing model.

VII. CONCLUSIONS

In this paper we have presented properties of a new $[[7, 1, 3]]$ stabilizer code that can achieve fault-tolerant syndrome measurements using a single ancillary qubit under the anisotropic error model, and showed why certain two-qubit errors under the standard depolarizing error model would prevent this code from achieving fault-tolerant measurements. In particular, the limit on total number of syndrome outcomes makes it impossible for the lookup table decoder to detect and correct all two-qubit errors in the standard depolarizing error model. Additionally, we showed that using the flag method the Bare $[[7, 1, 3]]$ code can achieve fault-tolerance under both the standard depolarizing model and the anisotropic model. Overall, with the flag method the Bare $[[7, 1, 3]]$ code shows better performance under realistic error models.

We note that the $[[7, 1, 3]]$ triangle code presented by Yoder and Kim [4] also does not require additional ancillary qubit preparations and is able to achieve fault-tolerance under the standard depolarizing error model. However, their code would not be identified by the search criteria that identified our code

[22]: a syndrome measurement with strictly bare ancillary qubits on their code would not be able to correct all hook errors originated from a single qubit error on the ancillary qubit. Instead they proposed an interwoven syndrome measurement method to achieve fault-tolerance. Since the interwoven method is not applicable for the Bare $[[7, 1, 3]]$ code, no direct comparison between the two codes can be easily made. But if the flag method [24] was to be applied on both codes, they can both achieve fault-tolerance under the standard depolarizing error model with the same amount of resources, and the pseudothreshold we obtained for the Bare $[[7, 1, 3]]$ code for the 1-Rec is comparable to the pseudothreshold reported by Yoder and Kim for the exREC [31].

A natural direction of future work would be to study the behavior of the Bare $[[7, 1, 3]]$ code under realistic other error models such as coherent errors and amplitude damping [34], and tailor this code for architecture-specific noise environments where limitations on ancillary qubit resources and measurement times are required.

The fact that the Bare $[[7, 1, 3]]$ code with the bare method does not perform well under the theoretical standard error model but can achieve fault-tolerant measurements under a more realistic error model suggests that more effort should be put into designing and finding error correcting codes that help achieve fault-tolerant quantum computation under realistic scenarios [35]. As realistic error models are becoming more and more relevant, development of environment-specific error correcting codes will become more important.

ACKNOWLEDGMENTS

[1] P. W. Shor, Phys. Rev. A **52**, R2493 (1995).

[2] R. Raussendorf and J. Harrington, Phys. Rev. Lett. **98**, 190504 (2007).

[3] A. G. Fowler, A. M. Stephens, and P. Groszkowski, Phys. Rev. A **80**, 052312 (2009).

[4] T. J. Yoder and I. H. Kim, arXiv preprint arXiv:1612.04795 (2016).

[5] K. M. Svore, D. P. Divincenzo, and B. M. Terhal, Quantum Info. Comput. **7**, 297 (2007), ISSN 1533-7146.

[6] M. Suchara, A. Faruque, C. Y. Lai, G. Paz, F. T. Chong, and J. Kubiatowicz, in *In Proceedings of IEEE International Conference on Computer Design* (IEEE, 2013).

[7] A. Steane, Proceedings: Mathematical, Physical and Engineering Sciences **452**, 2551 (1996), ISSN 13645021.

[8] A. R. Calderbank and P. W. Shor, Phys. Rev. A **54**, 1098 (1996).

[9] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information* (Cambridge university press, 2010).

[10] J. Preskill, in *Introduction to quantum computation and information* (Singapore: World Scientific, 1998), pp. 213–269.

[11] P. W. Shor, in *Foundations of Computer Science, 1996. Proceedings., 37th Annual Symposium on* (IEEE, 1996), p. 56 65.

[12] D. P. DiVincenzo and P. Aliferis, Phys. Rev. Lett. **98**, 020501 (2007).

[13] Y. Tomita, M. Gutiérrez, C. Kabytayev, K. R. Brown, M. R. Hutsel, A. P. Morris, K. E. Stevens, and G. Mohler, Phys. Rev. A **88**, 042336 (2013).

[14] A. M. Steane, Phys. Rev. Lett. **78**, 2252 (1997).

[15] E. Knill, Nature **434**, 39 (2005).

[16] D. Bacon, Phys. Rev. A **73**, 012340 (2006).

[17] P. Aliferis and A. W. Cross, Phys. Rev. Lett. **98**, 220502 (2007).

[18] A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, Phys. Rev. A **86**, 032324 (2012).

[19] H. Bombin and M. A. Martin Delgado, Phys. Rev. Lett. **97**, 180501 (2006).

[20] A. J. Landahl, J. T. Anderson, and P. R. Rice, arXiv:1108.5738 [quantph] (2011).

[21] Y. Tomita and K. M. Svore, Phys. Rev. A **90**, 062320 (2014).

[22] A. Cross, Personal Communication (2016).

[23] S. Aaronson and D. Gottesman, Phys. Rev. A **70**, 052328 (2004).

[24] R. Chao and B. W. Reichardt, arXiv preprint arXiv:1705.02329 (2017).

[25] A. M. Steane, Phys. Rev. Lett. **77**, 793 (1996).

[26] M. Gutiérrez and K. R. Brown, Phys. Rev. A **91**, 022335 (2015).

[27] M. Gutiérrez, C. Smith, L. Lulushi, S. Janardan, and K. R. Brown, Phys. Rev. A **94**, 042338 (2016).

[28] D. P. DiVincenzo and P. W. Shor, Phys. Rev. Lett. **77**, 3260 3263 (1996).

[29] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Phys. Rev. A **54**, 3824 (1996).

[30] A. Sørensen and K. Mølmer, Phys. Rev. Lett. **82**, 1971 (1999).

[31] P. Aliferis, D. Gottesman, and J. Preskill, Quantum Info. Comput. **6**, 97 (2006), ISSN 1533-7146.

[32] F. J. MacWilliams and N. J. A. Sloane, *The theory of error correcting codes* (Elsevier, 1977).

[33] K. M. Svore, A. W. Cross, I. L. Chuang, and A. V. Aho, Quantum Information and Computation **6**, 193 (2006).

[34] M. Gutiérrez, L. Svec, A. Vargo, and K. R. Brown, Phys. Rev. A **87**, 030302 (2013).

[35] A. Robertson, C. Granade, S. D. Bartlett, and S. T. Flammia, arXiv preprint arxiv:1703.08179 (2017).

[36] S. Janardan, Y. Tomita, M. Gutiérrez, and K. R. Brown, Quantum Information Processing **15**, 3065 3079 (2016).

[37] S. Bravyi and A. Vargo, Phys. Rev. A **88**, 062308 (2013).

Appendix A: Efficient Sampling Algorithm for Monte Carlo Simulations

For a given QEC circuit and a physical noise model, obtaining an exact algebraic expression for the logical error rate is in principle possible. In the context of a stabilizer code and a noise model consisting of stochastic discrete errors, this amounts to (1) enumerating every possible error configuration on the circuit, (2) calculating its probability of occurrence, and (3) determining whether or nor it results in a logical error. The logical error rate is then given by:

$$p_L = \sum_{i=0}^{N_c} A_i\, p_i, \qquad (A1)$$

where N_c is the total number of error configurations, A_i is the probability of occurrence of error configuration i, and $p_i = 0(1)$ if the error configuration i is correctable (uncorrectable). Although possible in principle, the exact computation of a logical error rate is infeasible in practice, due to the high cardinality of the error configuration set. For a circuit with n_g gates and s possible different errors after each gate, $N_c = (s+1)^{n_g}$. Even a modest circuit, like a level-1 QEC routine for a distance-3 code consisting of 3 rounds of stabilizer measurements, contains more than 100 gates, making the exact computation of the logical error rate impractical.

It is, therefore, common to employ Monte Carlo methods to estimate p_L [2, 3, 15, 36]. The basic procedure to obtain a logical error rate consists of two steps: (1) the generation of a faulty circuit (an error configuration) based on the physical noise model and (2) the simulation of the circuit to determine if that particular error configuration is correctable. In this section, we describe the two different Monte Carlo methods used in the first step. The first one utilizes traditional direct sampling of the whole error configuration set. The second one relies on importance sampling of the error configuration subsets that are relevant to the logical error rate. Both methods are based on direct sampling, i.e., each error configuration is completely uncorrelated from the previous ones.

1. Traditional Sampler

To generate an error configuration, the traditional sampler traverses the circuit exhaustively and after each gate an error is inserted with a probability given by the noise model. This approach is convenient for high error rates. However, it is problematic if the error rate is low, because most of the times no error is inserted. For illustrative purposes, consider a simple case where the error rate $p = 0.1\%$ is the same for ech gate and the circuit has $N_g = 100$ gates. The traditional sampler will return an error-free circuit $(1-p)^{n_g} \approx 90\%$ of the runs. Furthermore, if the circuit corresponds to a fault-tolerant QEC protocol of a distance-3 code, by construction no error configuration of weight-1 will result in a logical error. This implies that $(1-p)^{n_g} + n_g p(1-p)^{n_g-1} \approx 99.5\%$ of the runs will

generate an error configuration which is known a priori to be correctable. The limitation of the traditional sampler becomes even more dramatic for lower error rates and codes of higher distance [37].

2. Importance Sampler

It is possible to split the error configuration set into subsets based on each configuration's error weight (number of errors). The key advantage of the importance sampler relies on two features of this particular subset splitting: (1) it is straightforward to compute analytically the total probability of occurrence of each subset (the sum of the probabilities of occurrence of the error configurations in the subset) and (2) for low error rates we can safely ignore high weight error subsets, since their probability of occurrence will be vanishingly small.

Here we consider subsets in terms of both errors that occur after single-qubit gates and two-qubit gates. For example, a subset labelled by (s, t) contains all error configurations of the circuit with s errors after single-qubit gates and t errors after two-qubit gates. Notice that an error after a two-qubit gate can still be of weight 1. Let n_s and n_t be the total number of single-qubit gates and two-qubit gates in this circuit. We assume an error model where the error probability is the same for all single-qubit gates (p_s) and the error probability is also the same for all two-qubit gates (p_t). The probability of occurrence of the error subset (s, t) is then:

$$A_{s,t}(p_s, p_t) = \binom{n_s}{s} p_s^s (1 - p_s)^{n_s - s} \binom{n_t}{t} p_t^t (1 - p_t)^{n_t - t}. \tag{A2}$$

Given a quantum circuit and a noise model, the algorithm to estimate the logical error rate using the importance sampler consists of three steps:

1. Select a tolerance value. This corresponds to the total added probability of occurrence of the high weight subsets that will not be sampled. In the worst scenario imaginable, every error configuration in the excluded subsets would result in a logical error. The tolerance value represents the worst-case discrepancy between the real and the obtained logical error rate. In particular, it provides a worst-case upper bound to the logical error rate.

2. For each error subset (s, t), use direct Monte Carlo sampling to approximate its logical error rate $p_L(s, t)$. The sampling is done by randomly selecting s single-qubit gates and t two-qubit gates and adding errors after them.

3. Calculate the total logical error rate for the circuit:

$$p_L(p_s, p_t) = \sum_{s,t} A_{s,t}(p_s, p_t) p_L(s, t), \tag{A3}$$

where $A_{s,t}(p_s, p_t)$ is computed analytically using Equation A2.

Notice that the logical error rate $p_L(s, t)$ for a particular (s, t) subset is not a function of the physical error rates p_s and p_t. This means that, for a given circuit and noise model, we can pre-compute the logical error rate for each relevant subset, and use those values to evaluate the total logical error rate for different physical error rates. In contrast to the traditional sampler, there is no need to re-run the Monte Carlo simulations for different values of p_s and p_t. We simply calculate new values for the probabilities of subset occurrence $A_{s,t}$ and compute p_L using Equation A3. Once the Monte Carlo simulations for each relevant subset are done, the logical error rates for different physical error rates can be computed at no additional cost, making the importance sampler much more efficient than the traditional one. However, for high physical error rates (about 10^{-2} and higher) the importance sampler becomes either inaccurate or very slow. The trade-off arises because as the error rate increases so does the probability of occurrence of higher weight subsets and therefore the number of subsets that need to be included to obtain an accurate result. In this error regime, it is suitable to employ the traditional sampler.

3. Comparing Performance of Sampling Algorithms

In order to test the performance of the importance sampler, we simulated the logical error rate for the Steane $[[7, 1, 3]]$ code, the five-qubit code, and the Bare $[[7, 1, 3]]$ code with both the traditional and the importance sampling algorithm. To make sure that the results of the two sampling algorithms are comparable, we analytically proved that the probability for an arbitrary error to occur at any location in the circuit is the same for both sampling algorithms.

From both Fig. 8 and Fig. 9 we can see that the results of the two sampling algorithms coincide exactly at lower physical error rates, but can start to diverge as physical error rate increases, with the logical error rate obtained from the traditional sampler slightly higher than that from the importance sampler. This is because for a given number of subsets, at lower physical error rates these subsets are enough to calculate the cumulative logical error rate to a high accuracy, but for higher physical error rates the importance sampler must sample a larger number of error subsets in order to achieve the same level of accuracy in the cumulative logical error rate. In this case, the traditional sampler accounts for all error subsets of this circuit, while the importance sampler only accounts for a small number of subsets, resulting in a smaller logical error rate.

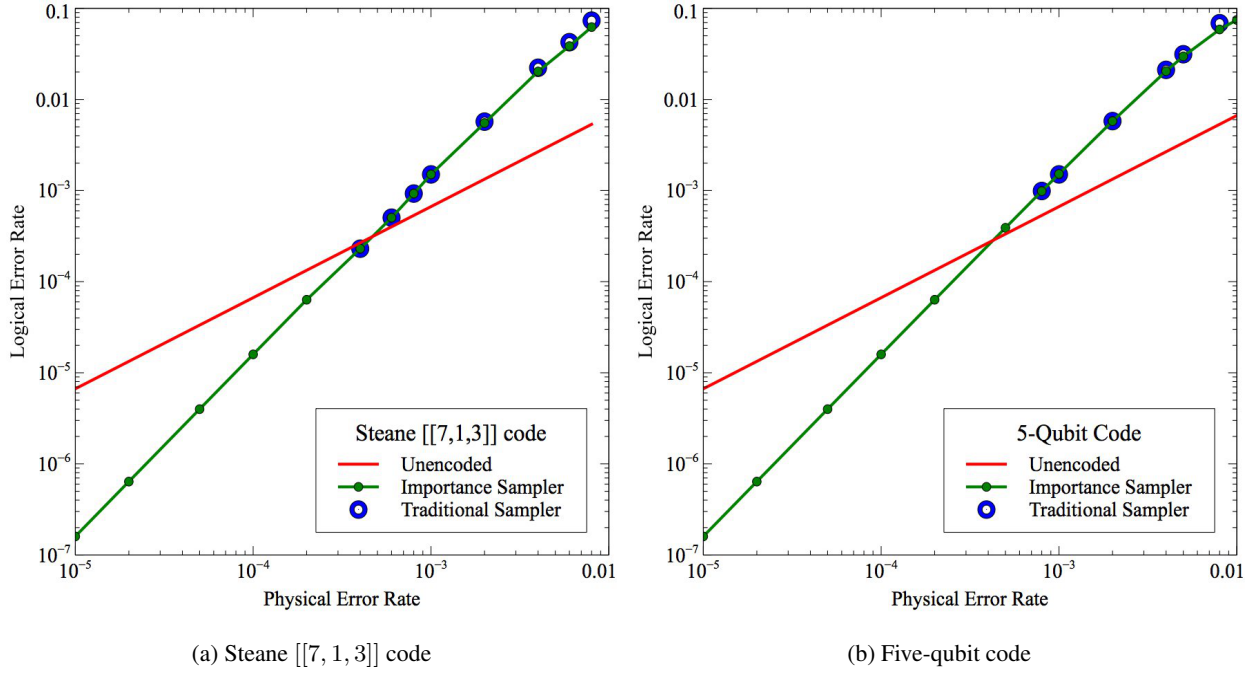(a) Steane $[[7, 1, 3]]$ code

(b) Five-qubit code

FIG. 8: Logical error rate with the importance sampler and the traditional sampler for the Steane $[[7, 1, 3]]$ code and the five-qubit code with Shor-style ancillary qubits under the standard depolarizing error model. Up to about $p = 2.0 \times 10^{-3}$, the two samplers result in essentially the same logical error rate.

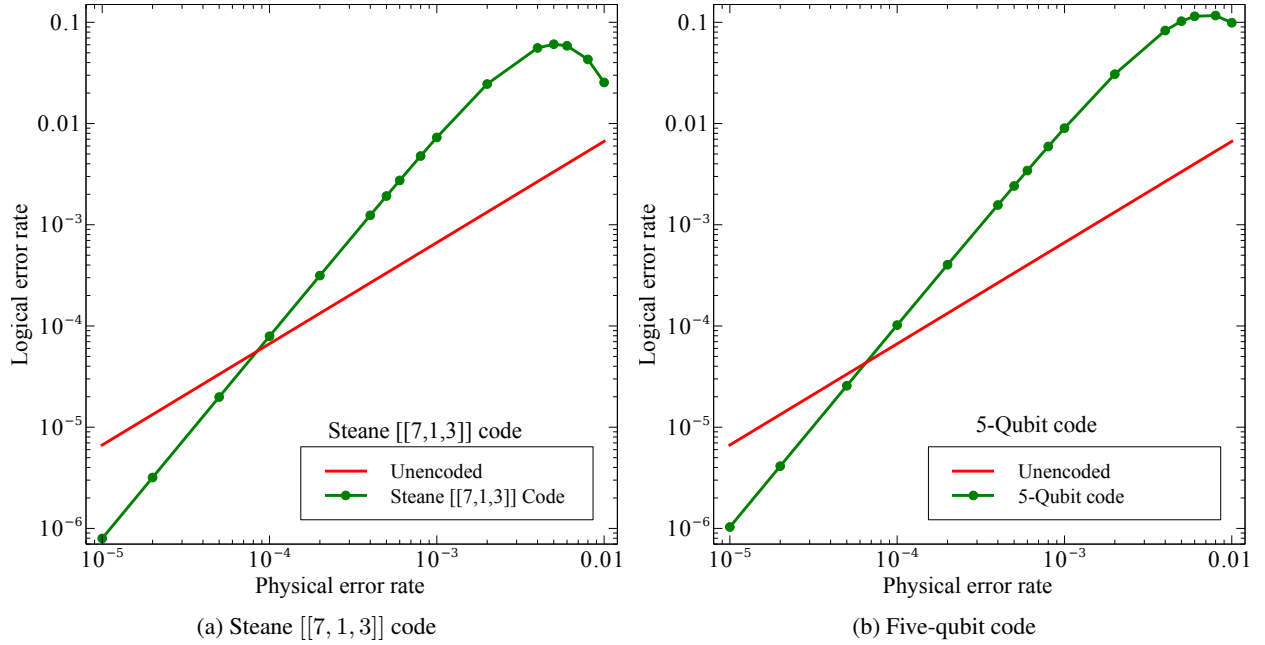(a) Steane $[[7, 1, 3]]$ code

(b) Five-qubit code

FIG. 9: Logical error rate with the importance sampler for the Steane $[[7, 1, 3]]$ code and the five qubit code with Shor-style ancillary qubits under the anisotropic error model. Around $p = 1.0 \times 10^{-2}$, the probability of occurrence of the high-weight errors becomes significant and ignoring them causes a dip on the logical error rate. The dip occurs at a lower p under the anisotropic error model because there are more faulty locations than under the standard model.