Unbiased Simulation of Near-Clifford Quantum Circuits

Ryan S. Bennink, Erik M. Ferragut, Travis S. Humble,* Jason
A. Laska, James J. Nutaro, Mark G. Pleszkoch, and Raphael C. Pooser

Quantum Computing Institute, Oak Ridge National Laboratory, Oak Ridge, TN, USA
(Dated: 12 May 2017)

Modeling and simulation is essential for predicting and verifying the behavior of fabricated quantum circuits, but existing simulation methods are either impractically costly or require an unrealistic simplification of error processes. We present a method of simulating noisy Clifford circuits that is both accurate and practical in experimentally relevant regimes. In particular, the cost is weakly exponential in the size and the degree of non-Cliffordness of the circuit. Our approach is based on the construction of exact representations of quantum channels as quasiprobability distributions over stabilizer operations, which are then sampled, simulated, and weighted to yield unbiased statistical estimates of circuit outputs and other observables. As a demonstration of these techniques we simulate a Steane [[7,1,3]]-encoded logical operation with non-Clifford errors and compute its fault tolerance error threshold. We expect that the method presented here will enable studies of much larger and more realistic quantum circuits than was previously possible.

I. INTRODUCTION

Modeling and simulation of quantum circuits is indispensable for evaluating the potential of quantum computing devices and guiding their development. State-of-the-art approaches to simulation are limited to circuits involving either a relatively small number of qubits or a highly restricted set of gates (e.g. error-free Clifford gates). As the dynamics of a general quantum physical system are in the complexity class BQP, there is almost certainly no scalable method of simulating universal quantum circuits on conventional computers. This is an important issue for the ongoing development of quantum computing technology, as the utility of modeling and simulation is likely to be severely limited when experimental systems reach sizes significant for quantum computing.

Although the cost of simulating quantum circuits on conventional computers is prohibitive in general, there are important exceptions. In particular, Gottesman and Knill showed that stabilizer circuits—circuits involving only Clifford operations and Pauli measurements—can be simulated efficiently on classical computers via stabilizer propagation [1]. While these circuits are not universal for quantum computing, they are significant for having a prominent role in quantum error correction [2]. Very large instances of these circuits can be simulated due to the efficient scaling of stabilizer propagation in both memory and run time, respectively, $O(n^2)$ and $O(kn^2)$ for n qubits and k gates.

Physical realizations of stabilizer circuits often incur errors as a result of imperfections in gate operation and uncontrolled interactions with the environment. Stabilizer propagation can be extended with Monte Carlo techniques to account for errors modeled as probabilistic mixtures of stabilizer operations. In this approach, the behavior of the circuit is taken to be the average of many realizations of the circuit with random stabilizer errors. However, many common error processes including amplitude damping, over/under-rotation, and misaligned rotation cannot be represented exactly as probabilistic mixtures of stabilizer operations. Several groups have investigated the honesty and accuracy of approximating such error processes as probabilistic mixtures of stabilizer operations [3, 4]. In some of the cases studied the stochastic approximation was found to have little impact on simulation results [4, 5], while in other cases it led to predictions of circuit failure rates that were in error by substantial factors [6].

We address the problem of simulating stabilizer circuits with non-stabilizer errors by representing non-stabilizer errors as *quasiprobability* distributions over stabilizer operations. As we show in the Appendix, relaxing the usual constraint that probabilities be non-negative permits an exact representation of any quantum process (including non-unitary processes, non-unital processes, and general measurements) as a weighted sum of stabilizer operations. Our simulation method begins by representing each gate in the circuit in such a way. The circuit behavior is then estimated by a weighted average over realizations of the circuit with stabilizer errors randomly drawn from each gate's quasiprobability distribution. Thus our method is a direct generalization of existing stochastic stabilizer propagation methods to allow simulation of arbitrary gates and errors. The price to be paid for full generality is that the cost of simulation is generally exponential in the circuit size. But critically, the coefficient of the exponent depends on the degree of "non-Cliffordness" of the circuit as measured by the negativity of the gates' quasiprobability representations. For fault tolerance circuits in experimentally relevant regimes the exponential growth is very weak, so that simulations of circuits with hundreds of qubits are feasible.

Our approach complements the recent work of Pashayan et al., who formalized the role of quasiprobabilities in simulating quantum circuits [7]. Their work identified the significance of negativity in a distribution for

determining the cost of simulation. This can be related to the fact that interference between probabilities (i.e. negativity) is a necessary condition for quantum computational speedup [8]. Others have shown that negativity plays an important role in determining the complexity of simulations involving quasiprobability distributions over circuit input states [9]. While our approach is similar in spirit to Pashayan et al., we have formulated circuit simulation specifically in terms of the qubit stabilizer formalism and stabilizer circuits.

The paper is organized as follows: Sec. II briefly reviews the existing theory of stabilizer circuit simulation before presenting a representation of arbitrary quantum channels as quasiprobability distributions over stabilizer channels. In Sec. III, we discuss the scaling of Monte Carlo simulations using the quasiprobability representation and describe its implementation in Sec. V. We then demonstrate our method by simulating a Steane [[7,1,3]]-encoded logical operation with amplitude damping errors and computing the logical error probability. We offer final remarks in Sec. VII.

II. STABILIZER CHANNEL DECOMPOSITION

Our simulation method is based on the exact representation of quantum channels as quasiprobability distributions over stabilizer channels. Quantum channels represent a very broad class of physical processes including ideal circuit operations and error processes as well as measurements [10]. Mathematically, a quantum channel is a linear mapping on the space of quantum states (the set of positive-semidefinite trace-1 operators). We denote the action of a quantum channel $\mathbf{A}$ on a state ρ as either $\mathbf{A}(\rho)$ or equivalently $\mathbf{A}\vec{\rho}$, where $\vec{\rho}$ denotes the vectorization of ρ. Channels should be contrasted with operators, which may act on states from the left or right. The channel corresponding to conjugation by an operator A ($\rho \to A\rho A^\dagger$) will be denoted by the same symbol in bold ($\mathbf{A}$).

The stabilizer formalism concerns a particularly useful set of quantum states and operations that originally arose in studies of error correction in circuits with Pauli errors [1]. An n-qubit stabilizer state is a fixed point of n independent n-qubit Pauli operators. Stabilizer states include the eigenstates of Pauli operators, as well as many highly entangled states useful for quantum computing and error correction. We call any quantum channel that maps stabilizer states to stabilizer states a *stabilizer channel*. The set of stabilizer channels includes Clifford channels (channels corresponding to the unitary operations known as Clifford operations) as well as non-unitary channels, i.e. channels that are many-to-one on the space of stabilizer states.

The great utility of the stabilizer formalism in the context of quantum error correction has prompted a number of authors to consider how it might be leveraged in broader contexts. Aaronson and Gottesman showed

how arbitrary pure states and arbitrary unitary operators could be expressed as complex superpositions of stabilizer states and Pauli operators, respectively, thereby admitting simulation of arbitrary circuits [11]. They provided a worst-case analysis in which the cost of simulation increases by a factor of 16^m for each non-Clifford m-qubit gate in the circuit. García and Markov introduced the concept of stabilizer frames to represent arbitrary pure states more compactly and demonstrated their effectiveness compared to existing simulation methods [12]. Yoder developed a similar representation for arbitrary states (including mixed states), consisting of a density operator in a stabilizer-defined basis [13]. Most recently, Brayvi and Gosset developed techniques involving linear combinations of stabilizer states to simulate circuits involving Clifford gates and limited numbers of T gates [14].

Others have taken the approach of simulating circuits with gates expressed as probabilistic mixtures (as opposed to complex superpositions) of stabilizer operations [3–5, 15]. Monte Carlo simulation of these circuits has a cost proportional to the cost of stabilizer propagation, which is between linear and quadratic in the number of qubits [11]. However, important quantum operations such as the T gate, as well as realistic errors such as amplitude damping and small unitary errors, cannot be exactly represented as probabilistic mixtures of stabilizer operations, thereby limiting the applicability of that approach.

In this work we generalize the stochastic approach by relaxing the constraint that the weights of the stabilizer operations be non-negative. The following lemma is proved in Appendix A:

Lemma 1. *Any trace-preserving quantum channel on n qubits can be expressed as a real linear combination of Clifford channels and Pauli reset channels on those qubits. If the channel is unital, Clifford channels suffice.*

Here a *Pauli reset channel* refers to the process of measuring a Pauli operator and then conditionally performing a Clifford operation to place the qubit into a prescribed eigenstate of the Pauli operator [15]. An example is the case of initializing a qubit to the $|0\rangle$ state by measuring the Z eigenvalue and performing an X operation if the -1 eigenvalue was obtained. While there are more complicated stabilizer channels that could also be considered— for instance channels corresponding to multiple rounds of alternating Pauli measurements and conditional Clifford operations—they are not considered in this work.

We define a *stabilizer channel decomposition* of a quantum channel $\boldsymbol{\chi}$ as an expression of the form

$$\boldsymbol{\chi} = \sum_i q_i \mathbf{S}_i \tag{1}$$

where each $\mathbf{S}_i$ is a stabilizer channel and each q_i is a real scalar. There are many different ways to decompose a given channel; however, every decomposition of a trace-preserving channel has the property $\sum_i q_i = 1$, which is a consequence of the fact that each constituent

channel $\mathbf{S}_i$ is also trace-preserving. Our representation of arbitrary quantum channels as real linear combinations of stabilizer channels (Clifford plus measurements) should be contrasted with the complex linear combinations of Pauli channels used by Aaronson and Gottesman to represent unitary channels [11]. Notably, our stabilizer channel decomposition can represent both unitary and non-unitary channels.

Each term in a stabilizer channel decomposition individually describes a physical process. Only when all the coefficients are non-negative can they be interpreted as probabilities of a stochastic physical processes. More generally, the coefficients of the stabilizer decomposition in Eq. (1) may be viewed as a quasiprobability distribution, i.e., a distribution in which some of the 'probabilities' are negative. The use of quasiprobability distributions to describe quantum mechanical phenomena traces back to the development of the Wigner function [16], wherein the appearance of a negative value indicates non-classical behavior [17]. More recently, quasiprobabilities have been used to express quantum gates and circuits with the finding that negative coefficients signify quantum behavior and strongly affect the complexity of simulating such circuits [7, 18]. Accordingly, an important metric of a stabilizer decomposition is the total magnitude of its negative coefficients, or *negativity*:

$$\eta \equiv \sum_{i:\ q_i < 0} |q_i|. \tag{2}$$

For a trace-preserving channel the negativity of a decomposition is directly related to its 1-norm: $\sum_i |q_i| = 1 + 2\eta \geq 1$. As will be elaborated in Section IV, the time complexity of our simulation method is strongly dependent on the 1-norms (and hence the negativities) of the decompositions of the circuit operations.

As a demonstration of the stabilizer channel decomposition, we present two examples of quantum channels that do not admit exact representations as probabilistic mixtures of stabilizer operations but may be represented exactly by (quasiprobability) stabilizer channel decompositions. In these examples, $\mathbf{I}$ denotes the identity channel, $\mathbf{Z}$ denotes the channel corresponding to conjugation by the Pauli operator Z, and $\mathbf{S}$ denotes the channel corresponding to conjugation by the phase gate $S = \sqrt{Z}$.

Example 1. *The channel $\mathbf{Z}_\theta$ corresponding to the coherent rotation $\rho \to Z_\theta \rho Z_\theta^\dagger$, where $Z_\theta = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{bmatrix}$, can be decomposed as*

$$\mathbf{Z}_\theta = \frac{1 + \cos\theta - \sin\theta}{2}\mathbf{I} + \frac{1 - \cos\theta - \sin\theta}{2}\mathbf{Z} + \sin\theta\mathbf{S}. \tag{3}$$

For $0 \leq \theta \leq \pi/4$, this decomposition has the least negativity of any exact decomposition. A prominent special case is $\theta = \pi/4$, which corresponds to the T gate:

$$\mathbf{T} = \frac{1}{2}\mathbf{I} + \left(\frac{1}{2} - \frac{1}{\sqrt{2}}\right)\mathbf{Z} + \frac{1}{\sqrt{2}}\mathbf{S}. \tag{4}$$

Example 2. *The amplitude damping channel*

$$\mathbf{A}_\gamma(\rho) = E_0 \rho E_0^\dagger + E_1 \rho E_1^\dagger \tag{5}$$

with $E_0 = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{bmatrix}$, $E_1 = \begin{bmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{bmatrix}$, and damping parameter $0 \leq \gamma \leq 1$ can be decomposed as

$$\mathbf{A}_\gamma = \frac{(1-\gamma) + \sqrt{1-\gamma}}{2}\mathbf{I} + \frac{(1-\gamma) - \sqrt{1-\gamma}}{2}\mathbf{Z} + \gamma\mathbf{R}_Z. \tag{6}$$

where the coefficient of $\mathbf{Z}$ is negative and approximately $-\gamma/4$ for $\gamma \ll 1$. This decomposition has the least negativity of any exact decomposition.

As stated, Lemma 1 applies to channels involving a fixed number of qubits, but the conclusion also applies to channels acting on any finite-dimensional quantum state space. This is because a channel that maps states in a d-dimensional Hilbert space to states in a d'-dimensional Hilbert space can be embedded in a Hilbert space of $n = \lceil \log_2 \max(d, d') \rceil$ qubits, to which Lemma 1 directly applies. Furthermore, the lemma can be extended to processes with observable outcomes (i.e. measurements). A well-known result of Neumark is that any quantum process with m possible outcomes can be cast as a unitary operation involving the system in question and an auxiliary system with m states, followed by projective measurement on the auxiliary system [19, 20]. It then follows from Lemma 1 that such a process can be expressed as a linear combination of Clifford channels acting on the original system plus $\lceil \log_2 m \rceil$ auxiliary qubits, followed by Pauli measurements of the auxiliary qubits. These conclusions are summarized by the following theorem:

Theorem 1. *Any quantum process in a d-dimensional Hilbert space with m possible outcomes can be expressed exactly as a quasiprobability distribution over stabilizer channels on $\lceil \log_2 d \rceil + \lceil \log_2 m \rceil$ qubits, where the stabilizer channels consist of (1) Clifford operations and (2) Pauli measurements followed by Clifford operations conditioned on measurement outcomes.*

For any number of qubits n, the number of Clifford channels substantially outnumbers the dimension of the channel space. Thus there are many different ways to decompose a given channel. Which way is best will depend on the application. Decompositions with the fewest terms may be optimal for analyses that involve a complete enumeration of stabilizer terms. For our simulation method outlined below, decompositions with the smallest 1-norm are optimal as they minimize the time cost. Finding the stabilizer channel decomposition of an arbitrary quantum channel χ that minimizes the 1-norm can be posed as the linear programming problem

$$\min_q \sum_i |q_i| \tag{7a}$$

subject to

$$\chi = \sum_i q_i \boldsymbol{S}_i \qquad (7\text{b})$$

where i indexes the Pauli reset and Clifford stabilizer channels, $q_i \in \mathbb{R}$, and $\boldsymbol{S}_i$ is the i-th stabilizer channel.

For the case of $n = 1$, the linear programming problem in Eq. (7) involves a system of 16 equations in 30 variables (one for each of 24 Cliffords and one for each of 6 Pauli resets). For $n = 2$ there are 256 equations in 11550 variables (11520 Cliffords and 30 Pauli resets), where each equation is sparse with exactly 16 nonzeros when channels are expressed as Pauli transfer matrices [21]. Many relevant quantum logic gates as well as physical noise models may be expressed in terms of 1- or 2-qubit interactions. The 2-qubit channel models may be composed together to generate models for larger channels as needed. However, for $n > 2$, the system of equations quickly becomes impractically large with almost 93 million variables for $n = 3$. This same difficulty arises for approaches that use a positive decomposition to represent the channel.

III. MONTE CARLO SIMULATION OF STABILIZER CHANNEL DECOMPOSITIONS

We now consider the problem of calculating the expectation value for an observable ϕ on the final state of a quantum circuit. Generalization to the case of multiple observables is straightforward. We model an initial state ρ acted upon by a sequence of quantum channels $\boldsymbol{\chi}^{(1)}, \ldots, \boldsymbol{\chi}^{(K)}$ that express the circuit gates and error processes. The purpose of the simulation is to compute the expectation value

$$F = \vec{\phi}^{\dagger} \boldsymbol{\chi}^{(K)} \cdots \boldsymbol{\chi}^{(1)} \vec{\rho}. \qquad (8)$$

We assume that each channel $\boldsymbol{\chi}^{(k)}$ has been decomposed into a weighted sum of stabilizer channels as $\boldsymbol{\chi}^{(k)} = \sum_i q_i^{(k)} \boldsymbol{S}_i$, cf. Eq. (7). Similarly, we assume that the initial state ρ and the observable ϕ have been decomposed into weighted sums of stabilizer states σ_i [12, 13]: $\rho = \sum_i q_i^{(0)} \sigma_i$ and $\phi = \sum_i q_i^{(K+1)} \sigma_i$.

With respect to these stabilizer decompositions, the expectation value F may be expressed as

$$F = \sum_{i_{K+1}} \cdots \sum_{i_0} q_{i_{K+1}}^{(K+1)} \cdots q_{i_0}^{(0)} \left(\vec{\sigma}_{i_{K+1}}^{\dagger} \boldsymbol{S}_{i_K} \cdots \boldsymbol{S}_{i_1} \vec{\sigma}_{i_0} \right) \qquad (9)$$

$$= \sum_i q(i) f(i) \qquad (10)$$

where $i = (i_0, \ldots, i_{K+1})$, $q(i) = q_{i_{K+1}}^{(K+1)} \cdots q_{i_0}^{(0)}$, and $f(i) = \vec{\sigma}_{i_{K+1}}^{\dagger} \boldsymbol{S}_{i_K} \cdots \boldsymbol{S}_{i_1} \vec{\sigma}_{i_0}$. The expression for $f(i)$ describes a stabilizer circuit, hence each $f(i)$ can be computed efficiently. However, explicit computation of Eq. (9) is impractical as the number of terms is generally exponential in K. Instead, F is estimated by sampling terms from the sum. Let p be any probability distribution over i such that $p(i) > 0$ when $q(i) > 0$, and let $i^{(1)}, \ldots, i^{(N)}$ denote values of i drawn independently from p. Then

$$\tilde{F}_N = \frac{1}{N} \sum_{s=1}^{N} \frac{q(i^{(s)}) f(i^{(s)})}{p(i^{(s)})} \qquad (11)$$

is an unbiased estimate of F, i.e. $\langle \tilde{F}_N \rangle = F$. Provided p can be sampled efficiently, $\tilde{F}_N$ can be computed efficiently. For reasons discussed in the next section, we take $p(i) = p_{i_{K+1}}^{(K+1)} \cdots p_{i_0}^{(0)}$ where $p_i^{(k)} \equiv |q_i^{(k)}| / \sum_j |q_j^{(k)}|$. The complete simulation procedure is summarized in Algorithm 1.

Algorithm 1 Observable estimation in terms of stabilizer states and stabilizer channels.

Given: An initial state ρ with stabilizer state decomposition $\rho = \sum_i q_i^{(0)} \sigma_i$
Given: A sequence of channels $\boldsymbol{\chi}^{(1)}, \ldots, \boldsymbol{\chi}^{(K)}$ with stabilizer channel decompositions $\boldsymbol{\chi}^{(k)} = \sum_i q_i^{(k)} \boldsymbol{S}_i$
Given: An observable ϕ with stabilizer state decomposition $\phi = \sum_i q_i^{(K+1)} \sigma_i$
Given: The number N of runs to simulate
    Let $p_i^{(k)} \equiv |q_i^{(k)}| / \sum_j |q_j^{(k)}|$ for $k = 0, 1, \ldots, K+1$.
    $\tilde{F} \leftarrow 0$
    for $r = 1$ to N **do**
        for $k = 0$ to $K+1$ **do**
            Select i_k with probability $p_i^{(k)}$
        end for
        $\rho_* \leftarrow \sigma_{i_0}$
        for $k = 1$ to K **do**
            $\rho_* \leftarrow \boldsymbol{S}_i(\rho_*)$
        end for
        $w \leftarrow q_{i_0}^{(0)} \cdots q_{i_{K+1}}^{(K+1)} / p_{i_0}^{(0)} \cdots p_{i_{K+1}}^{(K+1)}$
        $\tilde{F} \leftarrow \tilde{F} + w \vec{\sigma}_{i_{K+1}}^{\dagger} \vec{\rho}_* / N$
    end for
Return: $\tilde{F} \approx \vec{\phi}^{\dagger} \boldsymbol{\chi}^{(K)} \cdots \boldsymbol{\chi}^{(1)} \vec{\rho}$

IV. PERFORMANCE AND SCALING

The precision of Monte Carlo simulation can be quantified by the variance of the estimator $\tilde{F}_N$:

$$\text{Var}\, \tilde{F}_N \equiv \langle |F^2| \rangle - \langle F \rangle^2 \qquad (12)$$

$$= \frac{1}{N} \left(\langle |qf/p|^2 \rangle - |F|^2 \right) = \frac{1}{N} \text{Var}\, \tilde{F}_1. \qquad (13)$$

(Here the random variable underlying the expectation values is the argument i common to the functions p, q, and f.) The number of samples N needed to achieve a specified variance scales as $\text{Var}\, \tilde{F}_1$, which is determined by the nature of the circuit being simulated as well as the

choice of the sampling distribution p. The minimal value of Var $\tilde{F}_1$ is obtained by sampling from the distribution $p^*(i) = |q(i)f(i)| / \sum_j |q(j)f(j)|$. However, computing p^* is typically as difficult as computing F itself. Instead, we use a distribution that holds to the key principle that p should be large (small) when $|qf|$ is large (small). Our choice for p described above is the product of the optimal sampling distributions for the individual channels, which simultaneously satisfies this key principle and remains easy to sample from.

Let $g_k \equiv \sum_i q_i^{(k)}$ denote the 1-norm of the kth stabilizer decomposition $(k = 0, 1, \ldots, K+1)$. Using the fact that

$$\frac{q(i)}{p(i)} = \text{sign}\left(q_{i_{K+1}}^{(K+1)} \cdots q_{i_0}^{(0)}\right) g_{K+1} \cdots g_0 \qquad (14)$$

the variance can be expressed as

$$\text{Var}\,\tilde{F}_N = \frac{1}{N}\left(g_{K+1}^2 \cdots g_0^2 \left\langle |f|^2 \right\rangle - |F|^2\right). \qquad (15)$$

Now, suppose $f(i)$ were replaced everywhere by $f(i) - \frac{1}{2}$. This would shift the expectation value of $\tilde{F}_N$ by a constant Δ, but would not alter its variance. Thus

$$\text{Var}\,\tilde{F}_N = \frac{1}{N}\left(g_{K+1}^2 \cdots g_0^2 \left\langle |f - \tfrac{1}{2}|^2 \right\rangle - |F - \Delta|^2\right) \quad (16)$$

$$\leq \frac{g_{K+1}^2 \cdots g_0^2}{N} \left\langle |f - \tfrac{1}{2}|^2 \right\rangle. \qquad (17)$$

Recall that $f(i)$ is the projection of one stabilizer state onto another. Since such a projection lies in the interval $[0, 1]$, we have $|f(i) - \frac{1}{2}| \leq \frac{1}{2}$ for all i. It follows that

$$\text{Var}\,\tilde{F}_N \leq \frac{g_{K+1}^2 \cdots g_0^2}{4N} \qquad (18)$$

which is in fact a tight bound. To make the scaling of N with circuit size more explicit, we define $g_{\text{in}} \equiv g_0$, $g_{\text{obs}} \equiv g_{K+1}$, and $g_{\text{ch}} \equiv \max_{k=1,\cdots,K} g_k \geq 1$. Then the number of simulations N needed to estimate F to a given statistical uncertainty ϵ is

$$N \leq \frac{g_{\text{in}}^2 g_{\text{obs}}^2 g_{\text{ch}}^{2K}}{4\epsilon^2}. \qquad (19)$$

A similar bound can be obtained using Hoeffding's inequality, which for the present problem is

$$\Pr(|\tilde{F}_N - F| > \epsilon) \leq \exp\left(-\frac{2N\epsilon^2}{g_{\text{in}}^2 g_{\text{obs}}^2 g_{\text{ch}}^{2K}}\right) \qquad (20)$$

Rearranging this expression yields a bound on the number of simulations N needed to ensure that the probability of obtaining an estimate with error $\leq \epsilon$ is at least $1 - \delta$:

$$N \leq \frac{g_{\text{in}}^2 g_{\text{obs}}^2 g_{\text{ch}}^{2K}}{2\epsilon^2} \ln \frac{2}{\delta} \qquad (21)$$

Equations (19) and (21) indicate that simulation cost as measured by the number of runs required is generally exponential in the circuit size K. However the base g_{ch} is often close to 1. In particular, if each channel happens to be exactly representable as a positive sum of stabilizer channels, then $g_{\text{ch}} = 1$ and the variance does not increase with K. In that case our simulation method reduces to existing methods that simulate circuits as probabilistic mixtures of stabilizer circuits. For comparison, the 1-norm for optimal decomposition of the T gate given in Eq. (4) is $\sqrt{2}$. Each T gate doubles the variance of the estimator and hence doubles the cost of obtaining an estimate of given precision. The reason the variance grows exponentially when the decompositions involve negative coefficients is that the expected value is obtained as an interference between much larger terms of opposite sign. Any imbalance between the number of sampled positive terms versus the number of sampled negative terms results in a relatively large error in the estimated value. This situation may be seen as a version of the well-known "sign problem" encountered in calculations of the properties of fermionic systems [22].

The targeted application is simulation of a circuit of stabilizer gates with weak (i.e., rare and/or close to identity) non-stabilizer errors. For the error channels considered in the examples, the negativity of the stabilizer decompositions are within small factors of commonly used error measures, such as infidelity and trace distance. We thus have $g_{\text{ch}}^2 \sim 1 + p_{\text{NC}}$ where $p_{\text{NC}} \ll 1$ may be loosely interpreted as the per-gate probability of a non-Clifford error. The number of runs needed to obtain a fixed accuracy estimate scales as $g_{\text{ch}}^{2K} \sim e^{Kp_{\text{NC}}}$, which does not increase rapidly until $Kp_{\text{NC}} \gtrsim 1$. Thus as a rule-of-thumb, our simulation method is efficient for $Kp_{\text{NC}} \lesssim 1$. For example, if the per-gate probability of a non-Clifford error is $\sim 10^{-3}$, using our method one can accurately and efficiently simulate circuits of at least ~ 1000 gates. In contrast, simulations using positive-mixture approximations can be guaranteed accurate only for $Kp_{\text{NC}} \ll 1$. Our method therefore makes it feasible to accurately simulate quantum circuits for which efficient stochastic methods may be significantly inaccurate. Additionally, our method supports a continuous trade-off between cost and model accuracy through the choice of how much negativity is used to model non-Clifford channels.

We have so far discussed the simulation cost in terms of the number of runs needed to obtain an estimate of specified precision The time cost of each run is just that of stabilizer propagation: proportional to the number of gates, and between linear and quadratic in the number of qubits [11]. There is additionally the cost of decomposing the initial state, the channels representing the noisy gates of the circuit, and the final observable. If there is a fixed 1- or 2-qubit error model for each type of gate and there are only a few types of gates, these decompositions can be performed once up front with negligible amortized cost. Finally, we observe that if there are multiple observables to be computed, only the last step of each cir-

cuit simulation needs to be repeated for each observable. Thus the total time cost for simulating a circuit consisting of n qubits, K quantum channels, and M observables scales as $O(n^2(K + M)(1 + p_{NC})^K)$. The space cost is also essentially that of stabilizer propagation, $O(n^2)$ [1]. In practice this simulation method is time-limited rather than memory-limited. As with other Monte Carlo methods, it is embarrassingly parallel and therefore the time cost can be greatly reduced with parallel computation.

V. IMPLEMENTATION

We demonstrate the accuracy of this stabilizer decomposition using a numerical implementation of Algorithm 1. Our implementation relies on a heavily modified version of the CHP stabilizer circuit simulator [11], which propagates quantum states represented as a tableau of signed Pauli operators. Our modifications offer support for non-unital channels by implementing the Pauli Z reset gate as well as methods to efficiently compute the projection of a stabilizer state onto any given stabilizer subspace. As an added benefit, support for measurement of multi-qubit Pauli operators has a time cost $O(n^2)$ with the resulting projection function scaling as $O(n^3)$. This is comparable with the state of the art methods from Cross et al. [23] but with the advantage that an intermediate circuit does not need to be synthesized to reduce the tableau to a canonical form.

The modified stabilizer propagation kernel is embedded in our simulator software responsible for implementing circuit scheduling, fault injection, and error correction logic. The CHP simulator is repeatedly executed to collect the statistics needed to estimate system observables by Monte Carlo sampling. Within the simulator, a parser decomposes quantum circuits into a discrete schedule of gates acting on addressable qubits. We use the well-known QASM pseudo-code notation to specify the input circuit and the internal representations of the simulator [24]. A gate scheduler modifies this circuit description to add subcircuit boundaries, i.e., timestamps, to identify the order in which gates act on qubits. Along each boundary, we insert noise operators that model the errors caused by each subcircuit. The stabilizer decomposition for each gate and noise operator is generated and stored by the simulator.

When an instance of a circuit is simulated, each gate is represented by an element from its stabilizer decomposition. These elements are chosen randomly with respect to the probability distribution determined by the stabilizer decomposition coefficients. An instance of a stabilizer circuit is used as a Monte Carlo sample for the non-stabilizer channel. We translate each stabilizer element in the circuit instance into a sequence of C, H, P, measure, and reset operations. The resulting sequence of CHP operations is simulated using the stabilizer propagation kernel. In the simulations of the Steane code described below, the simulator frequently encounters gates
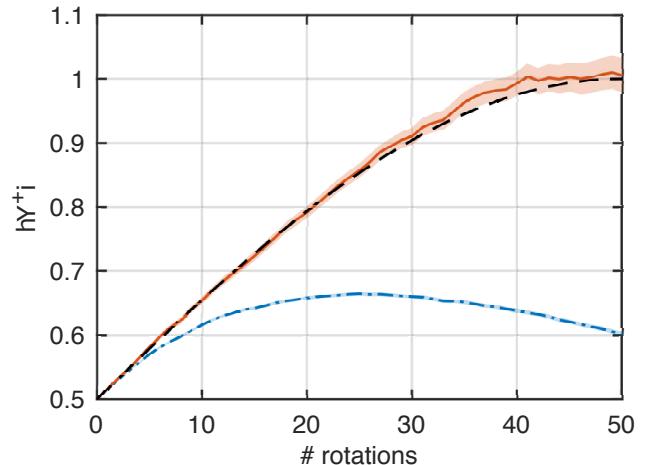

FIG. 1: Monte Carlo simulation of a qubit undergoing repeated coherent rotations, as revealed by its projection onto the +1 eigenstate of the Pauli Y operator. (dashed black line) The analytical result $\langle Y^+ \rangle = (1 + \sin\theta)/2$. (solid red line) Estimate using an exact stabilizer decomposition of the rotation channel. (dash-dot blue line) Estimate using an approximate positive decomposition. Shaded bands cover the area one standard deviation above and below the estimates.

conditioned upon previous measurement outcomes, i.e., syndromes. In our implementation, syndromes are returned by the propagation kernel to the circuit simulator and decoded using a pre-computed lookup table. The gate scheduler prepares the subcircuit that implements the corresponding error correction operation and pushes this new gate sequence to the propagation kernel.

For the numerical demonstrations reported here, we focus on estimating the infidelity of a qubit logically encoded using a quantum-error correction code and stored in the presence of physical noise. Therefore, we conclude each simulation instance by computing the projection of the final state onto a specified (error-free) code state. For calculating this fidelity, we project the simulated noisy state onto the ideal state prepared by the noiseless circuit. We accumulate a running average of the fidelity with the i-th result weighted by the factor q_i, as determined by the coefficients of the i-th randomly selected noisy, circuit instance. A complete simulation repeats the process of sampling the stabilizer channels, propagating the circuit instance, and calculating the desired observable N times, where the number of samples N is specified as input.

We have verified the correctness of our numerical simulator by comparing its results for several test circuits with exact results obtained from an independent density matrix simulator. The results from Monte Carlo simulation were observed to converge to the results of the corresponding density matrix simulation for all test cases provided sufficient samples were collected. As a proof-of-principle example that also illustrates the advantages of this approach, we simulated a simple circuit in which a

qubit initially in the $+1$ eigenstate of the Pauli X operator is rotated in 50 equal steps to the $+1$ eigenstate of the Pauli Y operator. Each rotation operator $Z_{\pi/100}$ is a non-stabilizer operation with an exact decomposition given by Eq. (3). The expected value of Y^+ was computed at each time step (i.e. after each rotation) and added to a running average, yielding a Monte Carlo estimate of the time dependence of $\langle Y^+ \rangle$. A typical outcome for 10,000 Monte Carlo runs is shown in Fig. 1). In this case the final value of $\langle Y^+ \rangle$ was estimated as 1.006 ± 0.027, which is in statistical agreement with correct value (1). While it might seem conspicuous that the estimated probability exceeds 1, the fact that the estimator is unbiased means that if the estimate has any chance of being less than the mean it must also have a chance of exceeding the mean.

For comparison, we also simulated the circuit using the approximate decomposition $\mathbf{Z}_{\pi/100} \approx (1-\sin\theta)\mathbf{I} + \sin\theta\mathbf{S}$, which is the positive decomposition that best reproduces the initial evolution of $\langle Y^+ \rangle$. In a typical run of 10,000 samples we obtained an estimate with mean 0.602 and standard deviation 0.003 (Fig. 1. Clearly, the use of an approximate positive decomposition in this case yields a strongly biased estimator. In addition, the estimator's precision is a very misleading indicator of its accuracy. We further observed that the estimated state using the exact channel decomposition remains approximately pure, whereas the estimated state using the approximate positive decomposition becomes more mixed with each step.

VI. SIMULATIONS OF STEANE [[7,1,3]] CIRCUITS

We used the numerical implementation of Algorithm 1 to investigate the logical error rate for a single-level encoding of the Steane [[7,1,3]] quantum error correction code in the presence of either depolarizing or amplitude damping noise. We used the former noise model as a baseline test case since the depolarizing noise may be represented as a non-negative mixture of stabilizer operators. It is possible to simulate depolarizing noise using conventional stabilizer propagation and we used this as a check for our numerical implementation. We used the latter case of amplitude damping as a test of the applicability of our new method for simulating weakly non-Clifford noise model as given in Eq. (6). To confirm the correct implementation of these noise models in our algorithm, we simulated each noise channel on various 1-qubit input states and found that results matched analytical predictions.

The input to the Steane circuit simulations was a single-qubit state to be prepared as a logical qubit using a noiseless encoding circuit. The encoding was followed by a noisy logical identity operation during which noise may occur on any qubit. Error correction was then performed consisting of syndrome measurement, decoding, and applying the corresponding correction operation.

These correction operations were assumed to be noisy as well. To calculate the logical fidelity of this operation, an additional round of noiseless error correction was performed and the fidelity of the resulting output state was taken with respect to the noiseless, logically encoded input. The specific encoding, syndrome measurement, and decoding circuits used in this numerical study are presented in Appendix B.

We calculated the logical error of the Steane encoded state as the probability that a noisy logical operation followed by noisy QEC operations would yield an unrecoverable error in the logical state. Therefore, we did not include noise within the encoding circuit in order to ensure the input logical state is error-free. In addition, following the noisy error correction operation in Fig. 2, we performed a second round of noiseless error correction. The purpose of the second round of QEC was to remove any correctable errors in the logical state. The logical error was then computed as the infidelity between the two logical states. The infidelity was calculated for six different input states, specifically the eigenstates of the Pauli operators X, Y, Z, and averaged. In the discussion of our results, "infidelity" always refers to the output state infidelity averaged over the six Pauli input states.

Our first example considers the depolarizing noise model, in which the channel

$$\mathbf{D}(\rho) = (1-p)\mathbf{I} + \frac{p}{3}(\mathbf{X} + \mathbf{Y} + \mathbf{Z}).\qquad(22)$$

is applied to each qubit involved in a noisy operation. The infidelity $\bar{I} = 1 - \bar{F}$ of a qubit under this channel is $2p/3$. Fig. 3 shows the estimated infidelity of the logical (Steane-encoded) identity operation as a function of the physical (1-qubit) infidelity. The solid line is the estimate $\tilde{\Phi}$; the shaded band marks the region $\tilde{\Phi} \pm \epsilon$ where ϵ is the standard deviation of $\tilde{\Phi}$ estimated from the sample variance. Our simulations indicate that the logical infidelity is less than the physical infidelity (shown by a dashed black line) when the physical infidelity is less than approximately 2.7×10^{-4} ($p \lesssim 4 \times 10^{-4}$). This result agrees very well with the average level-1 pseudo-threshold computed in [4] (row "DC" of Table VII therein).

Our second example considers the amplitude damping noise model, wherein the channel $\mathbf{A}$ defined in eq. (5) is applied to each qubit involved in a noisy operation. For our simulations we used the stabilizer decomposition given by Eq. (6), whose negativity is approximately $-\gamma/4$, where γ ranges from 0 to 1. For $\gamma \ll 1$ the infidelity of this channel is $\bar{I} \approx \gamma/3$, which was obtained the relation $\bar{F} = (2F_{\mathbf{A},\mathbf{I}} + 1)/3$ [25], where $F_{\mathbf{A},\mathbf{I}} = (1 + \sqrt{(1 - \gamma)})^2/4$ is the fidelity of the process matrices for channels $\mathbf{A}$ and $\mathbf{I}$. Fig. 3 shows the estimated logical infidelity as a function of the physical infidelity under the amplitude damping noise model. Again, the solid line denotes the estimate, the shaded band shows the estimated uncertainty, and the dashed line shows the physical infidelity. The slight bend in the logical infidelity appears to be an artifact of statistical
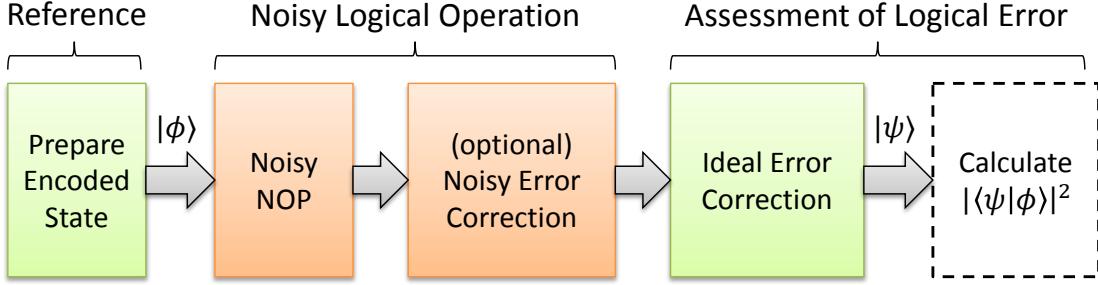
FIG. 2: Block structure of the circuit simulated as demonstration of our simulation method. The circuit prepares a reference logical state in the Steane [[7,1,3]] encoding. The encoded state is then subjected to errors, and optionally, a round of noisy error correction. The logical error of the resulting state is obtained by removing correctable errors with a round of error-free error correction, then projecting the result onto the originally encoded state. Details of the subcircuits, including specific gates and locations of error insertion, are provided in Appendix B.
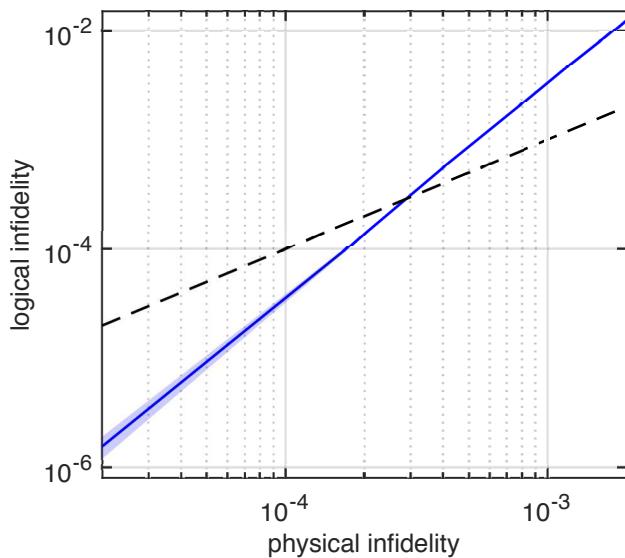
FIG. 3: (sold blue line with shaded uncertainty band) Estimated infidelity of a Steane [[7,1,3]] logical identity operation as a function of physical qubit depolarization. The circuit implementing the logical operation is specified by Fig. 2. For comparison, the physical infidelity is shown by the dashed black line.
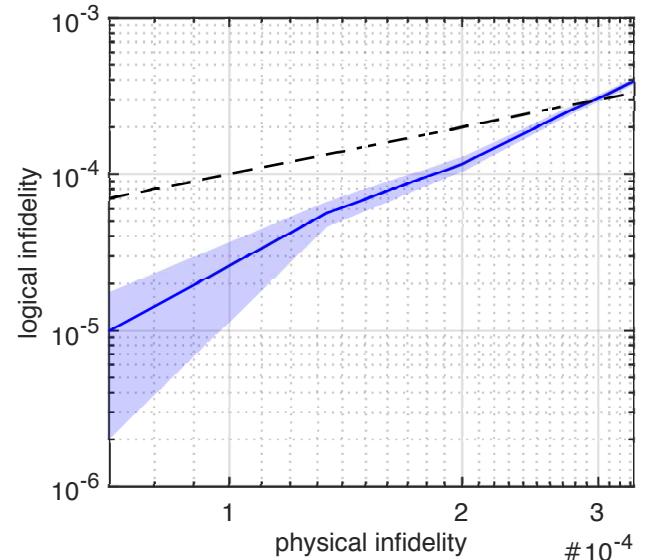
FIG. 4: (sold blue line with shaded uncertainty band) Estimated infidelity of a Steane [[7,1,3]] logical identity operation as a function of physical qubit amplitude damping. The circuit implementing the logical operation is specified by Fig. 2. For comparison, the physical infidelity is shown by the dashed black line.

fluctuations. The logical infidelity is less than the physical infidelity when the latter is less than approximately 2.9×10^{-4} ($\gamma \lesssim 9 \times 10^{-4}$). This is about a factor of two greater than the pseudo-threshold estimated by Gutierrez et al. for a similar circuit [4]. This difference may be attributed to several differences in our approach: Whereas we use majority vote when three rounds of syndrome extraction are performed, Gutierrez et al. rely on whatever results arises from the third round. In addition, they use the trace distance as the basis for threshold calculations and average over a larger set of input states.

VII. CONCLUSION

Modeling and simulation play an important role in verifying and validating quantum information technologies. We have presented a new method for simulating arbitrary quantum circuits that leverages the efficiency of stabilizer propagation and Monte Carlo sampling. A key aspect of this method is that it offers an exact representation of the circuit described by quasiprobability distributions over efficiently simulatable operations. The cost of this simulation method depends smoothly on the negativity of the quasiprobability distributions, which may serve as

a measure of their non-Cliffordness. Consequently, this simulation method is particularly well suited for studying quantum error correction circuits with weak non-Clifford noise models. A promising application is the study of coherent error models which can be expressed exactly and simulated using our methods. Coherent errors in quantum error correction have recently shown interesting results [26–28].

Unbiased Monte Carlo simulation of the stabilizer channel decomposition offers several convenient features over existing simulation methods. This includes a memory scaling that is polynomial in circuit size while maintaining an exact description of the quantum circuit. The Monte Carlo sampling method is highly parallelizable and portable to large-scale distributed computing environments. In addition, the time scaling is controllable via the number of samples requested to achieve a desired estimator variance. Finally, by limiting the amount of negativity allowed in the quasiprobability representations of circuit operations, the trade-off between simulation accuracy and efficiency is adjustable. Quantifying this trade-off is an important direction for future work to better understanding the relationship between the nature of a quantum channel and its optimal quasiprobability representation. Future work will also include an investigation into efficient methods for obtaining optimal or near-optimal stabilizer decompositions of quantum channels involving two or more qubits.

[1] D. Gottesman, arXiv preprint quant-ph/9807006 (1998).
[2] F. Gaitan, *Quantum error correction and fault tolerant quantum computing* (CRC Press, 2008).
[3] E. Magesan, D. Puzzuoli, C. E. Granade, and D. G. Cory, Phys. Rev. A **87**, 012324 (2013), URL http://link.aps.org/doi/10.1103/PhysRevA.87.012324.
[4] M. Gutiérrez and K. R. Brown, Phys. Rev. A **91**, 022335 (2015), URL http://link.aps.org/doi/10.1103/PhysRevA.91.022335.
[5] D. Puzzuoli, C. Granade, H. Haas, B. Criger, E. Magesan, and D. G. Cory, Phys. Rev. A **89**, 022306 (2014), URL http://link.aps.org/doi/10.1103/PhysRevA.89.022306.
[6] M. Gutiérrez, C. Smith, L. Lulushi, S. Janardan, and K. R. Brown, Phys. Rev. A **94**, 042338 (2016), URL https://link.aps.org/doi/10.1103/PhysRevA.94.042338.
[7] H. Pashayan, J. J. Wallman, and S. D. Bartlett, Phys. Rev. A **115**, 070501 (2015).
[8] D. Stahlke, Phys. **90**, 022302 (2014).
[9] V. Veitch, C. Ferrie, D. Gross, and J. Emerson, New Journal of Physics **14**, 113011 (2012), URL http://stacks.iop.org/1367-2630/14/i=11/a=113011.
[10] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information* (Cambridge university press, 2010).
[11] S. Aaronson and D. Gottesman, Phys. Rev. A **70**, 052328 (2004), URL http://link.aps.org/doi/10.1103/PhysRevA.70.052328.
[12] H. J. García and I. L. Markov, Computers, IEEE Transactions on **64**, 2323 (2015).
[13] T. J. Yoder (2012), URL http://www.scottaaronson.com/showcase2/report/ted-yoder.pdf.
[14] S. Bravyi and D. Gosset, arXiv preprint arXiv:1601.07601 (2016).
[15] M. Gutiérrez, L. Svec, A. Vargo, and K. R. Brown, Phys. Rev. A **87**, 030302 (2013), URL http://link.aps.org/doi/10.1103/PhysRevA.87.030302.
[16] E. Wigner, Physical Review **40**, 749 (1932).
[17] A. Kenfack and K. Życzkowski, Journal of Optics B: Quantum and Semiclassical Optics **6**, 396 (2004).
[18] H. F. Hofmann, J. Phys. A: Math. Theor. **42**, 275304 (2009).
[19] A. Peres, Foundations of Physics **20**, 1441 (1990), ISSN 0015-9018.
[20] W. R. M. Rabelo, A. G. Rodrigues, and R. O. Vianna, International Journal of Modern Physics C **17**, 1203 (2006).
[21] J. M. Chow, J. M. Gambetta, A. D. Córcoles, S. T. Merkel, J. A. Smolin, C. Rigetti, S. Poletto, G. A. Keefe, M. B. Rothwell, J. R. Rozen, et al., Phys. Rev. Lett. **109**, 060501 (2012), URL http://link.aps.org/doi/10.1103/PhysRevLett.109.060501.
[22] M. Troyer and U.-J. Wiese, Phys. Rev. Lett. **94**, 170201 (2005).
[23] H. García, I. Markov, and A. Cross, arxiv:1210.6646v3 [cs.ET] (2013).
[24] S. Balensiefer, L. Kreger-Stickles, and M. Oskin (2005), vol. 5815, pp. 103–114, URL http://dx.doi.org/10.1117/12.604073.
[25] M. A. Nielsen, Physics Letters A **303**, 249 (2002), ISSN 0375-9601.
[26] A. S. Darmawan and D. Poulin, arXiv:1607.06460 ((2016)).
[27] Y. Suzuki, K. Fujii, and M. Koashi, arxiv:1703.03671 ((2017)).
[28] J. P. Barnes, C. J. Trout, D. G. Lucarelli, and B. Clader,

arXiv:1704.03961 ((2017)).

[29] S. Kimmel, M. P. da Silva, C. A. Ryan, B. R. Johnson, and T. Ohki, Physical Review X **4**, 011050 (2014).

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

APPENDIX A: CLIFFORD DECOMPOSITION

We extend the argument made in Ref. [29] to prove Theorem 1, which states that any trace-preserving quantum channel can be expressed as a real linear combination of stabilizer channels and, specifically, Clifford channels and Pauli measure-reset channels. Let $\mathcal{P}_n$ denote the set of unsigned n-qubit Pauli operators and let $\mathbf{1}^{(P',P)}$ denote the (generally unphysical) channel that maps $P \in \mathcal{P}_n$ to $P' \in \mathcal{P}_n$ and maps every other element of $\mathcal{P}_n$ to 0. Since $\mathcal{P}_n$ is a basis for n-qubit states, $\{\mathbf{1}^{(P',P)}\}$ is a basis for the space of n-qubit channels. We will show that each basis channel can be written as a real linear combination of stabilizer channels, and thus that any quantum channel can be expressed as such a linear combination.

We use the fact that each non-identity operator $P \in \mathcal{P}_n$ commutes with exactly half of the elements of $\mathcal{P}_n$ and anticommutes with the remaining elements. Thus

$$\sum_{Q \in \mathcal{P}_n} [\![Q, P]\!] = \begin{cases} 4^n & P = I^{\otimes n} \\ 0 & \text{else} \end{cases} \qquad (A1)$$

where $[\![Q, P]\!] \equiv 1\,(-1)$ in the case that Q commutes (anticommutes) with P. Furthermore, $[\![Q, P]\!][\![Q, P']\!] = [\![Q, PP']\!]$ which gives

$$\sum_{Q \in \mathcal{P}_n} [\![Q, P]\!][\![Q, P']\!] = \begin{cases} 4^n & P = P' \\ 0 & \text{else} \end{cases}. \qquad (A2)$$

Consider the linear combination of stabilizer channels

$$\frac{1}{4^n} \sum_{Q \in \mathcal{P}_n} [\![P, Q]\!] \mathbf{Q} \qquad (A3)$$

where $\mathbf{Q}$ is the channel corresponding to operator Q. We have

$$\frac{1}{4^n} \sum_{Q \in \mathcal{P}_n} [\![P, Q]\!] \mathbf{Q}(P') = \frac{1}{4^n} \sum_{Q \in \mathcal{P}_n} [\![Q, P]\!] Q P' Q^\dagger$$

$$= P' \frac{1}{4^n} \sum_{Q \in \mathcal{P}_n} [\![Q, P]\!][\![Q, P']\!] \quad (A4)$$

$$= \begin{cases} P' & P' = P \\ 0 & \text{else} \end{cases} \qquad (A5)$$

$$= \mathbf{1}^{(P,P)}. \qquad (A6)$$

The expression on the left is thus a stabilizer decomposition of $\mathbf{1}^{(P,P)}$.

We now obtain a decomposition of $\mathbf{1}^{(P',P)}$ where $P' \neq P$ and $P, P' \neq I^{\otimes n}$. Let $\mathbf{C}$ be any Clifford channel that maps P to P'. Such a channel always exists. Then

$$\mathbf{C}\mathbf{1}^{(P,P)}(P'') = \left\{ \begin{array}{ll} P' & P'' = P \\ 0 & \text{else} \end{array} \right\} = \mathbf{1}^{(P',P)}. \qquad (A7)$$

To obtain a channel of the form $\mathbf{1}^{(P,I^{\otimes n})}$ where $P \neq I^{\otimes n}$, let $\mathbf{R}_P$ denote a channel that sets the qubits to a $+1$ eigenstate of P. This may be implemented as a measurement of P conditionally followed by a Clifford channel $\mathbf{N}_P$ that maps P to $-P$, where $\mathbf{N}_P$ is applied only in the case that measurement yielded the -1 eigenvalue. Using the fact that $(I^{\otimes n} \pm P)/2$ is the projector onto the ± 1 eigenspace of P, we have

$$\mathbf{R}_P(I^{\otimes n}) = \frac{I^{\otimes n} + P}{2} I^{\otimes n} \frac{I^{\otimes n} + P}{2} +$$
$$\mathbf{N}_P \frac{I^{\otimes n} - P}{2} I^{\otimes n} \frac{I^{\otimes n} - P}{2} \qquad (A8)$$

$$= \frac{I^{\otimes n} + P}{2} + \mathbf{N}_P \frac{I^{\otimes n} - P}{2} \qquad (A9)$$

$$= I^{\otimes n} + P. \qquad (A10)$$

Then

$$\mathbf{R}_P \mathbf{1}^{(I^{\otimes n}, I^{\otimes n})}(P') = \begin{cases} I^{\otimes n} + P & P' = I^{\otimes n} \\ 0 & P' \neq I^{\otimes n} \end{cases} \qquad (A11)$$

yields the decomposition $\mathbf{1}^{(P,I^{\otimes n})} = \mathbf{R}_P \mathbf{1}^{(I^{\otimes n}, I^{\otimes n})} - \mathbf{1}^{(I^{\otimes n}, I^{\otimes n})}$.

Summarizing, we have

$$\mathbf{1}^{(P',P)} = \frac{1}{4^n} \sum_{Q \in \mathcal{P}_n} [\![P, Q]\!] \mathbf{C}\mathbf{Q} \qquad (A12)$$

$$\mathbf{1}^{(P',I^{\otimes n})} = \frac{1}{4^n} \sum_{Q \in \mathcal{P}_n} (\mathbf{R}_{P'} - 1) \mathbf{Q} \qquad (A13)$$

where $P, P' \in \mathcal{P}_n$ with $P \neq I^{\otimes n}$, $\mathbf{C}$ is any Clifford channel that maps P to P', and $\mathbf{R}_P$ is any channel that resets P to a $+1$ eigenvalue. Each of the expressions above is clearly a linear combination of stabilizer channels, where each term is a product of at most one Pauli measurement and at most two Clifford operations. All that remains are channels of the form $\mathbf{1}^{(I^{\otimes n}, P)}$, and we do not need to decompose these since a trace-preserving channel cannot have such a channel as a component. (If it did, the trace-1 states $(I^{\otimes n} + P)/2$ and $(I^{\otimes n} - P)/2$ would map to states of unequal trace.) Thus the space of trace-preserving n-qubit channels is spanned by the set of n-qubit stabilizer channels.

APPENDIX B: QUANTUM CIRCUITS FOR NUMERICAL SIMULATION

We present the quantum circuits used for the numerical results presented in Sec. VI. This includes the encoding and error correction subcircuits for the Steane

[[7,1,3]] code [30], which corresponds to the functional blocks shown in Fig. 2. The first two blocks in that figure correspond to the subcircuits shown in Figs. 5 and 6, respectively. The block "Error Correction" in Fig. 2 consists of the six syndrome extraction circuits, which correspond to Figs. 7–12 below. Syndrome extraction is followed by a recovery operation (not shown) prescribed by the measurement results. In the case of noisy error correction, the syndrome extraction circuit is repeated three times; a correction is applied only if at least two of syndromes are the same. In these diagrams, MR denotes a measurement in the computational basis followed by a reset of the qubit to the $|0\rangle$ state. Dashed vertical lines denote the boundaries at which simulated errors occur. At such times, each qubit is independently subjected to either amplitude damping or depolarizing error, depending on the noise model chosen for the simulation.
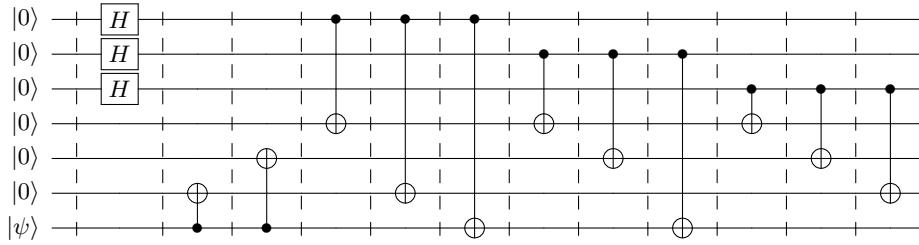
FIG. 5: Quantum circuit for encoding Steane [[7,1,3]] quantum error correction codeword.
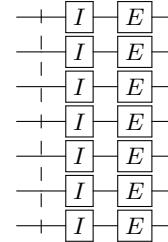
FIG. 6: Quantum circuit for logical identity operation, referred to as a no-op.
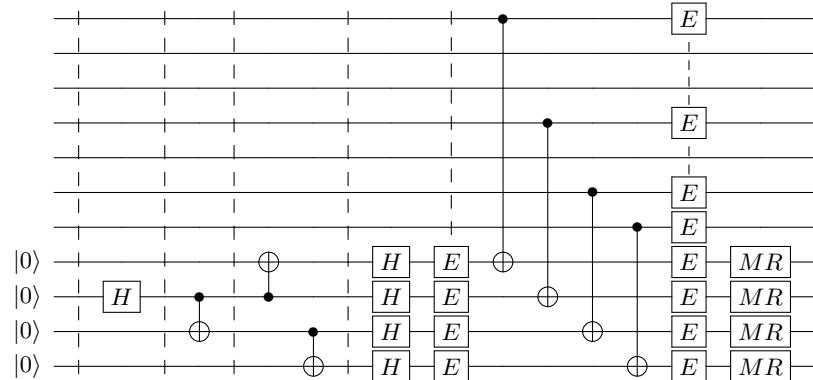


FIG. 7: Quantum circuit for extracting syndrome 1.
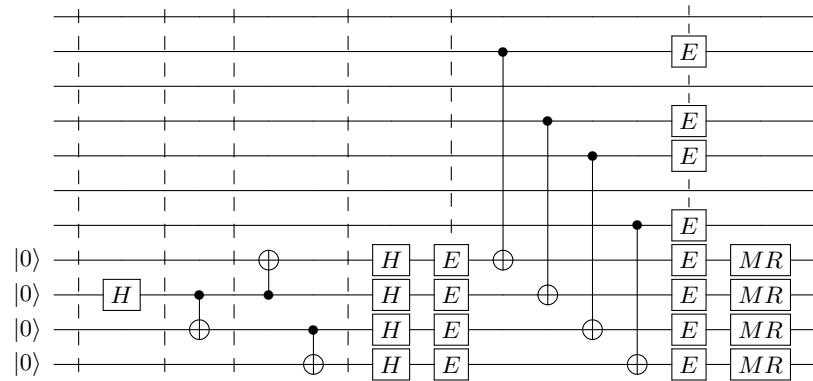
FIG. 8: Quantum circuit for extracting syndrome 2.
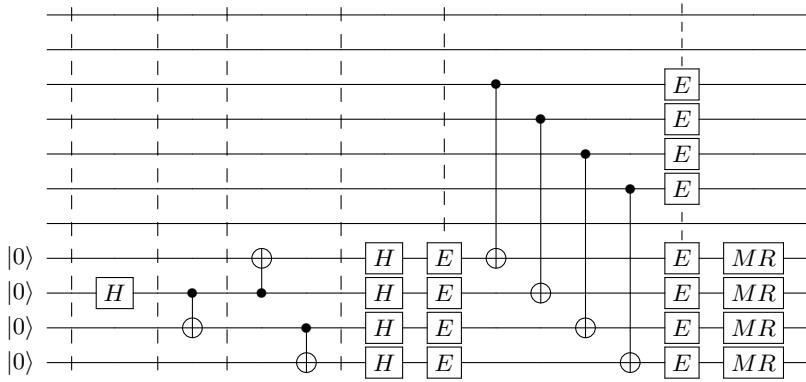
FIG. 9: Quantum circuit for extracting syndrome 3.
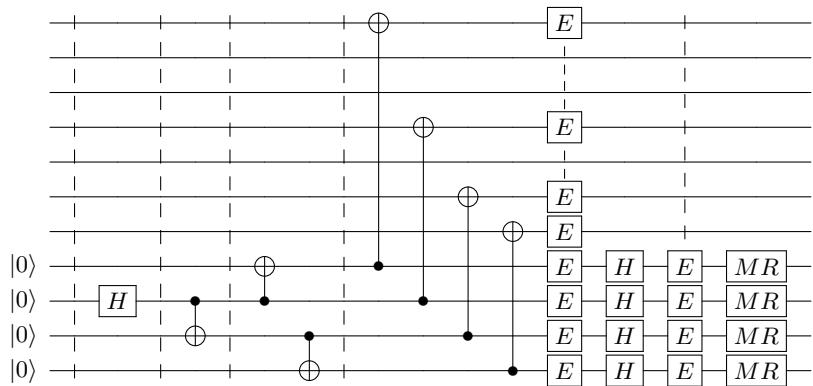
FIG. 10: Quantum circuit for extracting syndrome 4.
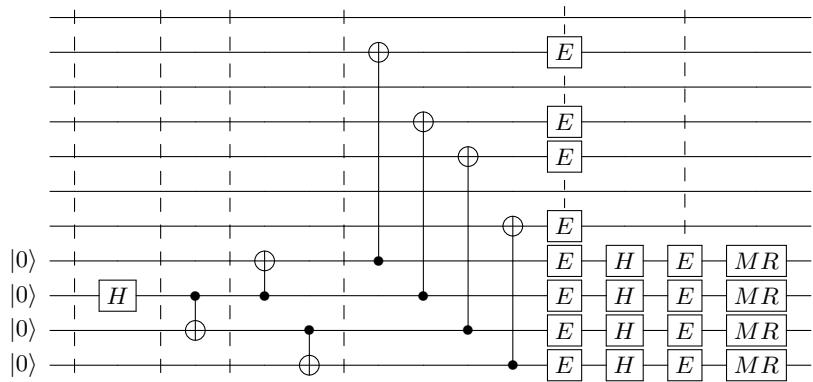

FIG. 11: Quantum circuit for extracting syndrome 5.
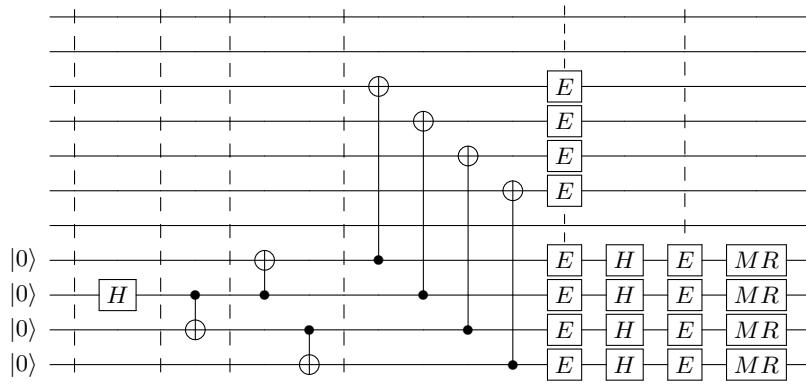
FIG. 12: Quantum circuit for extracting syndrome 6.