Correcting coherent errors with surface codes

Sergey Bravyi, 1 Matthias Englbrecht, 2 Robert König, 2 and Nolan Peard 2, 3

1IBM T.J. Watson Research Center
2Institute for Advanced Study & Zentrum Mathematik, Technical University of Munich
3Department of Physics, Massachusetts Institute of Technology
(Dated: October 9, 2017)

We study how well topological quantum codes can tolerate coherent noise caused by systematic unitary errors such as unwanted Z-rotations. Our main result is an efficient algorithm for simulating quantum error correction protocols based on the 2D surface code in the presence of coherent errors. The algorithm has runtime $O(n^2)$, where $n$ is the number of physical qubits. It allows us to simulate systems with more than one thousand qubits and obtain the first error threshold estimates for several toy models of coherent noise. Numerical results are reported for storage of logical states subject to Z-rotation errors and for logical state preparation with general $SU(2)$ errors. We observe that for large code distances the effective logical-level noise is well-approximated by random Pauli errors even though the physical-level noise is coherent. Our algorithm works by mapping the surface code to a system of Majorana fermions.

I. INTRODUCTION

Recent years have witnessed major progress towards the demonstration of quantum error correction and reliable logical qubits [1–5]. Topological quantum codes such as the surface code [6, 7] are among the most attractive candidates for an experimental realization, as they can be implemented on a two-dimensional grid of qubits with local parity check operators.

It is believed that such codes can tolerate a high level of noise [8–10] which is comparable to what can be achieved in the latest experiments [5]. The general confidence in the noise-resilience of topological codes primarily rests on considerations of Pauli noise – a simplified noise model where errors are Pauli operators $X, Y, Z$ drawn at random from some distribution. An example is the case where each qubit $j$ experiences noise described by the channel

$$N_j(\rho) = (1 - \epsilon)\rho + \epsilon_x X\rho X + \epsilon_y Y\rho Y + \epsilon_z Z\rho Z \quad (1)$$

with suitable probabilities $\epsilon_x, \epsilon_y, \epsilon_z$. This kind of noise can be fully described by the stabilizer formalism [11]. In pioneering work, Dennis et al. [8] exploited this algebraic structure to establish the first analytical threshold estimates, see also [12]. The effect of Pauli noise also is efficiently simulable thanks to the Gottesman-Knill theorem, providing numerical evidence for high error thresholds of topological codes [13]. The efficient simulability property has recently been extended beyond Pauli noise to random Cliffords and Pauli-type projectors [14].

While such algebraically defined noise models are attractive from a theoretical viewpoint, they often do not correspond to noise encountered in real-world setups. They are – in a sense – not quantum enough: they model probabilistic processes where errors act randomly on subsets of qubits. Rather than being of such a probabilistic (or incoherent) nature, noise in a realistic device will often be coherent, i.e., unitary, and can involve small rotations acting everywhere. A typical situation where this arises is if e.g., frequencies of oscillator qubits are mis-aligned: this results in systematic unitary over- or under-rotations. On a single-qubit level, this means that (1) should be replaced by noise of the form

$$N_j(\rho) = U_j\rho U_j^\dagger \quad (2)$$

with a suitable unitary operator $U_j \in SU(2)$. Since such errors generally cannot be described within the stabilizer formalism, understanding their effect on a given quantum fault-tolerant scheme is a challenging problem.

Prior theoretical work indicates that the difference between coherent and incoherent errors could be significant. In particular, it was observed [15–19] that coherent errors can lead to large differences between average-case and worst case fidelity measures suggesting that a critical reassessment of commonly used benchmarking measures is necessary. This observation motivates the question of how much coherence is present in the effective logical-level noise [20, 21] experienced by encoded qubits. Depending on whether or not the logical noise is coherent one may choose different metrics for quantifying performance of a given fault-tolerant scheme. Significant progress has been made towards understanding the structure of the logical noise for concatenated codes [20–22]. However, these studies are not directly applicable to large topological codes such as those considered here.

Brute-force simulations of coherent noise in small codes were presented in [23–26] for Steane codes and surface codes with up to 17 qubits. Simulating coherent errors by brute force clearly requires time (and memory) exponential in the number of qubits $n$. For the surface code, Darmawan and Poulin [27] proposed an algorithm with a runtime exponential in $n^{1/2}$ based on tensor networks, and simulated systems with up to 153 qubits. This algorithm can handle arbitrary noise (including e.g., amplitude damping). Unfortunately, its formidable complexity prevents accurate estimation of error thresholds, e.g., for the systematic rotations considered here. In [28], threshold estimates for the 1D repetition code were obtained. To our knowledge, there are no analogous threshold estimates for topological codes subject to coherent noise.
Our setup. Here we show that the effect of coherent errors in surface codes can be studied by means of polynomial-time algorithms. Specifically, we consider coherent errors in the context of two central tasks associated with error correction, namely

(A) fault-tolerant storage of quantum information.

(B) fault-tolerant preparation of a logical basis state.

We shall consider a particular version of the surface code proposed in Refs. [29, 30]. A distance-\(d\) surface code has one logical qubit and \(n = d^2\) physical qubits located at sites of a square lattice of size \(d \times d\) with open boundary conditions. The code has local stabilizers \(X^\otimes 4\), \(X^\otimes 2\) or \(Z^\otimes 4\), \(Z^\otimes 2\) associated with faces of the lattice as shown in Fig. 1. The stabilizer located on a face \(f\) will be denoted \(B_f\). Logical Pauli operators \(X_L\) and \(Z_L\) acting on the encoded qubit can be chosen as \(X^\otimes d\) and \(Z^\otimes d\) applied to the left and the top boundary of the lattice respectively. The two-dimensional logical subspace is spanned by \(n\)-qubit states \(\psi_L\) satisfying \(B_f|\psi_L\rangle = |\psi_L\rangle\) for all \(f\).

![Surface Codes](image)

FIG. 1. Surface codes with distance \(d = 3\) and \(d = 5\). Qubits and stabilizers are located at sites and faces respectively. Logical Pauli operators \(X_L\) (red) and \(Z_L\) (blue) have support on the left and the top boundary.

To specify the problem (A), consider a logical state \(\psi_L\) initially encoded by the surface code and a coherent error \(U = U_1 \otimes \cdots \otimes U_n\) that applies some (unknown) unitary operator \(U_j\) to each qubit \(j\). To diagnose and correct the error without disturbing the encoded state we adopt the standard protocol based on the syndrome measurement. It works by measuring the eigenvalue (syndrome) \(s_f = \pm 1\) of each stabilizer \(B_f\) on the corrupted state \(U|\psi_L\rangle\) and then applying a Pauli-type correction operator \(C_s\) depending on the measured syndrome \(s = \{s_f\}_f\). The correction \(C_s\) is computed by a classical decoding algorithm (for example, one may choose \(C_s\) as a minimum-weight Pauli error consistent with \(s\)). We note that the syndrome \(s\) is a random variable with some probability distribution \(p(s)\) since the error \(U\) maps the initial logical state to a coherent superposition of states with different syndromes. In this paper we only consider noiseless syndrome measurements. Accordingly, we assume that the correction \(C_s\) always returns the system to the logical subspace resulting in some final logical state \(|\phi_s\rangle\). For this problem we restrict to \(Z\)-rotation errors, that is, we assume that \(U_j = \exp(i\eta_j Z)\) for some (unknown) angles \(\eta_j\). The restriction to \(Z\)-rotations is dictated by the limitations of our simulation algorithm. Thus we shall model a fault-tolerant storage by the following process:

(i) prepare an initial logical state \(|\psi_L\rangle\)

(ii) apply a coherent error \(\otimes_{j=1}^n \exp(i\eta_j Z)\) to \(|\psi_L\rangle\)

(iii) measure the eigenvalues of the stabilizers \(\{B_f\}_f\), resulting in a syndrome \(s = \{s_f\}_f\).

(iv) apply a Pauli correction \(C_s\) returning the system to the logical subspace in some final state \(|\phi_s\rangle\).

To assess how close the final state \(|\phi_s\rangle\) and the initial state \(|\psi_L\rangle\) are, we seek a polynomial-time classical algorithm \(A\) which takes as input \(|\psi_L\rangle\) and the rotation angles \(\eta_1, \ldots, \eta_n\), samples a syndrome \(s\) from the distribution \(p(s)\) specified by the measurement (iii), and outputs \(s\) as well as the associated final state \(|\phi_s\rangle\) (e.g. specified by its Bloch vector). By sampling sufficiently many syndromes, one can learn how frequently and in which ways error correction may fail in the presence of coherent noise.

To specify the problem (B), assume first that we have access to noise-free qubits and operations. In this case, the following standard protocol [10] prepares the encoded stabilizer state \(|+L\rangle\) (the +1 eigenvector of \(X_L\)):

(i) prepare the initial product state \(|+\rangle^\otimes n\).

(ii) measure the eigenvalues of the stabilizers \(\{B_f\}_f\), resulting in a syndrome \(s = \{s_f\}_f\).

(iii) apply a Pauli correction \(C_s\) returning the system to the logical subspace in some final state \(|\phi_s\rangle\).

Using the fact the the initial state is a +1 eigenvector of \(X_L\) one can easily check that \(|\phi_s\rangle = |+L\rangle\) for all \(s\), see [10]. How does this protocol fare in the presence of coherent noise? Let us consider a model where the initial product state cannot be prepared with perfect accuracy, but rather is obtained from \(|+\rangle^\otimes n\) by applying some unwanted unitary operators to every qubit. Thus (i) is replaced by

(i') prepare an initial state \(|\psi_1 \otimes \psi_2 \otimes \cdots \otimes \psi_n\rangle\), where \(|\psi_j\rangle\) are arbitrary single-qubit pure states.

In this case the final state \(|\phi_s\rangle\) may deviate from the target state \(|+L\rangle\) resulting in a logical error. To assess the performance of this protocol, we seek a polynomial time algorithm \(B\) which, on input \(|\psi_1, \ldots, \psi_n\rangle \in \mathbb{C}^2\), outputs a random syndrome \(s\) sampled from the distribution \(p(s)\) specified by the measurement (ii) together with the final logical state \(|\phi_s\rangle\).

Our results. We construct algorithms \(A\) and \(B\) accomplishing the simulation tasks specified above. The runtime of these algorithms scales as \(O(n^2)\), where we measure complexity in terms of the number of additions, multiplications, and divisions on complex numbers that
are required. Using these algorithms, we perform the first numerical study of large topological codes subject to coherent noise, performing simulations for surface codes with up to $n = 2401$ physical qubits, see Table I for a timing analysis. This shows that efficient classical simulation of these fault-tolerance processes is possible, and allows us to extract key characteristics of these codes in the limit of large system size.

We apply algorithm A to study the effect of coherent noise on storage in the surface code. We show that the syndrome probability distribution $p(s)$ is independent of the initial logical state $\psi_L$, whereas the final logical state has the form

$$|\phi_s\rangle = \exp(i\theta_s Z_L)|\psi_L\rangle$$

for some logical rotation angle $\theta_s \in [0, \pi)$ depending on the syndrome $s$. We use the quantity

$$P^L = 2\sum_s p(s)\left|\sin \theta_s\right|$$

as a measure of the logical error rate. We will see that $P^L$ is the average diamond-norm distance between the conditional logical channel $\rho \mapsto e^{i\theta_s Z}\rho e^{-i\theta_s Z}$ and the identity channel. For numerical simulations we consider translation-invariant coherent noise of the form $(e^{i\theta Z})^\otimes n$, where $\theta \in [0, \pi)$ is the only noise parameter. The Pauli correction $C_\theta$ was computed using the standard minimum weight matching decoder [8, 31] with constant weights independent of $\theta$. We are interested in the error threshold, that is, the maximum value $\theta_0$ such that for any $\theta < \theta_0$ the logical error rate $P^L$ goes to zero in the limit $n \to \infty$. We find the numerical estimate

$$0.08\pi \leq \theta_0 \leq 0.1\pi.$$  

(5)

Our numerical experiments confirm that, as expected, the quantity $P^L$ decays exponentially in the code distance for values $\theta < \theta_0$ below the threshold. Surprisingly, the threshold estimate Eq. (5) agrees very well with the so-called Pauli twirl approximation [32, 33] where coherent noise of the form $N(\rho) = e^{i\theta Z}\rho e^{-i\theta Z}$ is replaced by dephasing noise $D(\rho) = (1 - \epsilon)\rho + \epsilon Z\rho Z$, with $\epsilon = \sin^2 \theta$. For the latter the threshold error rate is around $\epsilon_0 \approx 0.11$, see Ref. [8]. Solving the equation $\epsilon_0 = \sin^2 (\theta_0)$ for $\theta_0$ gives $\theta_0 \approx 0.10\pi$, in agreement with Eq. (5). At the same time, we observe that the Pauli twirl approximation significantly underestimates $P^L$ in the sub-threshold regime, confirming that coherence of noise may have a profound effect on a given fault-tolerant scheme, as was previously observed in [22, 27].

Algorithm A allows us to investigate the probability distribution of logical rotation angle $\theta_s$ defined in Eq. (3). We find that for large code sizes, this distribution concentrates around the two points $(0, \pi/2)$ which correspond to the logical Pauli-type errors $\{I, Z_L\}$. To get a deeper insight into this phenomenon, we introduce and numerically study associated measures of “incoherence”. Our findings support the general conjecture that in the limit of large code distances, coherent physical noise gets converted into incoherent logical-level noise.

We apply Algorithm B to study the effect of coherent noise on logical state preparation in the case when the initial product state has the form

$$(\exp (i\varphi X) \exp (i\theta Z) |+) \otimes n$$

Here the angles $\varphi, \theta \in [0, \pi)$ specify the action of a coherent error on the ideal initial state $|+\rangle$. The ideal protocol corresponds to $\theta = 0$. We define the logical error rate $P^L$ as the average trace-norm distance between the final logical state $\phi_s$ and the target state $|+L\rangle$, see Section VI B for details. Our numerical results indicate that the error threshold can be described by a single function $\theta_0(\varphi)$ such that the logical error rate $P^L$ goes to zero in the limit $n \to \infty$ for any $0 \leq \theta < \theta_0(\varphi)$ and $P^L$ is lower bounded by a positive constant for $\theta > \theta_0(\varphi)$. We find the numerical estimate

$$0.1\pi \leq \theta_0(\varphi) \leq 0.15\pi$$

(6)

for all $\varphi$. This indicates that the threshold function $\theta_0(\varphi)$ has a very mild (if any) dependence on $\varphi$. We investigate the behavior of $P^L$ in more detail for $\varphi = 0$ and obtain a more refined estimate $0.13\pi \leq \theta_0(0) \leq 0.14\pi$. The quantity $P^L$ is observed to decay exponentially in the code distance in the sub-threshold regime.

| Code distance | 9 | 19 | 29 | 39 | 49 |
|---------------|---|----|----|----|----|
| Qubits        |   |    |    |    |    |
| Runtime (A)   | 81 | 361 | 841 | 1521 | 2401 |
| Runtime (B)   | 0.001 | 0.04 | 0.2 | 0.7 | 1.7 |

TABLE I. Runtime in seconds for a C++ implementation of algorithms A and B. Timing analysis was performed on a laptop with a 2.6GHz Intel i5 Dual Core CPU.

Outline. The remainder of this paper is structured as follows. Section II provides a high-level overview of our simulation algorithms. In Section III, we describe a representation of the surface code in terms of Majorana fermions. In Sections IV,V, we give the classical simulation algorithms A and B and analyze their complexity. In Section VI we discuss our numerical results. We conclude in Section VII. Appendix A contains a proof of a technical lemma. We provide some background on Majorana fermions and fermionic linear optics in Appendix B.

\footnote{Strictly speaking, the simulation time scales as $O(n^2) + t(n)$, where $t(n)$ is the runtime of the decoding algorithm that computes the correction $C_\theta$. In our simulations the decoding time was negligible compared with the time required to sample the syndrome and compute the final logical state.}
II. METHODS

Our main tool is a fermionic representation of the surface code proposed by Kitaev [34] and Wen [29]. It works by encoding each qubit of the surface code into four Majorana fermions in a way that simplifies the structure of the surface code stabilizers. The Kitaev-Wen representation has previously been used by Terhal et al. [35] to design fermionic Hamiltonians with topologically ordered ground states. Here we show that this representation is also well-suited for the design of efficient simulation algorithms. The fermionic version of the surface code will be described in terms of Majorana operators $c_1, \ldots, c_{4n}$ that obey the standard commutation rules $c_p^\dagger = c_p$, $c_p^2 = I$, and $c_p c_q = -c_q c_p$ for $p \neq q$. We will show that the error correction protocols considered in this paper can be decomposed into a sequence of $O(n)$ elementary gates from a gate set known as a fermionic linear optics (FLO), see Refs. [36, 37]. It includes the following operations:

1. Initialize a pair of Majorana modes $p,q$ in a basis state $|0\rangle$ satisfying $ic_p c_q |0\rangle = |0\rangle$.
2. Apply the unitary operator $U = \exp(\gamma c_p c_q)$. Here $\gamma \in [0, \pi]$ is a rotation angle.
3. Apply the projector $\Lambda = (I + ic_p c_q)/2$. Compute the norm of the resulting state.

It is well-known that quantum circuits composed of FLO gates can be efficiently simulated classically [36–39]. The simulation runtime scales as $O(n)$ for gates of type (1,2) and as $O(n^2)$ for gates of type (3). For completeness, we describe the requisite simulation algorithms in Appendix B. By exploiting the geometrically local structure of the surface code we shall be able to reduce the number of modes such that at any given time step the simulator only needs to keep track of $O(n^{1/2})$ modes. Accordingly, each FLO gate can be simulated in time at most $O(n)$. Since the total number of gates is $O(n)$, the total simulation time scales as $O(n^2)$.

III. FROM QUBITS TO MAJORANA FERMIONS

A single Majorana mode $p$ is described by a hermitian operator $c_p$ satisfying $c_p^2 = I$. Operators $c_p$ associated with different modes anti-commute, see Appendix B for formal definitions. A system of four Majorana modes $c_1, c_2, c_3, c_4$ can be used to encode a qubit using a stabilizer code with a single stabilizer

$$S = -c_1 c_2 c_3 c_4$$ (7)

and logical Pauli operators

$$\bar{X} = ic_1 c_2 = ic_3 c_4 S \quad \text{and} \quad \bar{Z} = ic_2 c_3 = ic_1 c_4 S.$$ (8)

We shall refer to this encoding as a $C4$-code.

Consider a surface code with $n$ qubits on a square $d \times d$ lattice, where $n = d^2$. It can described by a planar graph $G = (V,E,F)$ with a set of $n$ vertices $V$, a set of $2n - 2$ edges $E$, and a set of $n - 1$ faces $F$. Qubits are located at vertices $u \in V$ and stabilizers $B_f$ are located at faces $f \in F$ of $G$. Consider a system of 4n Majorana modes $c_1, \ldots, c_{4n}$ distributed over edges and vertices of $G$ as shown on Fig. 2. There are exactly two paired modes located near the endpoints of every edge $e \in E$ and four unpaired modes $c_1, c_2, c_3, c_4$ located near the corners of the lattice as shown on Fig. 2. The paired modes are labeled as $c_5, c_6, \ldots, c_{4n}$ in an arbitrary order.

Let us orient the edges of $G$ as shown on Fig. 2. Suppose $e \in E$ is an edge connecting some pair of modes $c_p, c_q$ such that $c_p$ is the tail of $e$ and $c_q$ is the head of $e$, see Fig. 2. Define the link operator

$$L_e = ic_p c_q.$$ (9)

Note that $L_e$ is hermitian and all link operators pairwise commute. Furthermore, each $L_e$ commutes with the unpaired modes $c_1, c_2, c_3, c_4$.

IV. FROM FLO TO QUBITS

Each $L_e$ can be written as a linear combination of unpaired Majorana modes $c_5, c_6, \ldots, c_{4n}$:

$$L_e = \sum_{p \in \Gamma_u} c_p$$ (10)

where $\Gamma_u$ is a small neighborhood of each vertex $u \in V$ that contains a cluster of four modes, see Fig. 3. We shall denote this cluster $\Gamma_u$. Define a vertex stabilizer

$$S_u = -\prod_{p \in \Gamma_u} c_p.$$ (10)

Consider a surface code with $n$ qubits on a square $d \times d$ lattice, where $n = d^2$. It can described by a planar graph $G = (V,E,F)$ with a set of $n$ vertices $V$, a set of $2n - 2$ edges $E$, and a set of $n - 1$ faces $F$. Qubits are located at vertices $u \in V$ and stabilizers $B_f$ are located at faces $f \in F$ of $G$. Consider a system of 4n Majorana modes $c_1, \ldots, c_{4n}$ distributed over edges and vertices of $G$ as shown on Fig. 2. There are exactly two paired modes located near the endpoints of every edge $e \in E$ and four unpaired modes $c_1, c_2, c_3, c_4$ located near the corners of the lattice as shown on Fig. 2. The paired modes are labeled as $c_5, c_6, \ldots, c_{4n}$ in an arbitrary order.
where a particular product order is chosen for each vertex as shown on Fig. 3. Since $|\Gamma_u| = 4$ for all $u$ and the subsets $\Gamma_u$ are pairwise disjoint, vertex stabilizers are hermitian and pairwise commuting. We shall consider $S_1, \ldots, S_n$ as stabilizers for $n$ independent copies of the C4-code defined in Eqs. (7,8) such that each qubit of the surface code is encoded into its own C4-code. Let $X_u, Z_u,$ and $Y_u \equiv iX_u Z_u$ be the logical Pauli operators for the qubit located at a vertex $u$, see Eq. (8). By definition, each of these logical operators has the form $ic_pc_q$ for some pair of Majorana modes $p, q \in \Gamma_u$, see Fig. 3. The logical operators $X_u, Z_u$ are indicated by small arrows on Fig. 5.

Let $P$ be a Pauli operator acting on the surface code qubits. We shall say that a Majorana operator $\mathcal{P}$ is a C4-encoding of $P$ if $\mathcal{P}$ can be obtained from $P$ by replacing each single-qubit Pauli operator $X_u, Y_u, Z_u$ by its logical counterpart $X_u, Y_u, Z_u$ and, possibly, multiplying by the stabilizer $S_u$. Given a single-qubit state $\psi$, one can define several encoded versions of $\psi$ using the surface code, the C4-code, and the surface code concatenated with $n$-copies of the C4-code. We shall denote these encoded states $\psi_L, \overline{\psi}$, and $\overline{\psi}_L$ respectively. These notations are summarized in Table II.

| Hilbert space          | Encoding                        |
|------------------------|---------------------------------|
| $\psi_L$               | $n$ qubits                      |
| $\overline{\psi}$     | surface code                    |
| $\overline{\psi}_L$   | $4n$ Majorana modes             |
|                        | encode each qubit of $\psi_L$ into the C4-code |

TABLE II. Encoded versions of a single-qubit state $\psi$.

The desired fermionic representation of the surface code is established in Lemmas 1,2,3 below. Consider a face $f \in F$ and let $\partial f \subseteq E$ be the boundary of $f$.

**Lemma 1.** Let $B_f$ be the surface code stabilizer located on a face $f$. Then a C4-encoding of $B_f$ can be chosen as

$$B_f = \prod_{e \in \partial f} L_e. \tag{11}$$

We illustrate Eq. (11) on Fig. 4. The lemma shows that measuring the surface code syndrome can be reduced (after the C4-encoding) to measuring eigenvalues of pairwise commuting link operators $L_e$. We shall see that under certain circumstances such measurements can be efficiently simulated classically.

**Proof.** Consider a face $f$ such that $B_f$ is a Z-stabilizer. Then $B_f = \prod_{u \in f} Z_u$. Consider a vertex $u \in f$ and the C4-code located at $u$. The corresponding logical-Z operator $Z_u = ic_pc_q$ can be chosen such that both modes $c_p, c_q$ are located on the boundary of $f$, see Fig. 3. Thus $B_f$ is proportional to the product of all modes located on the boundary of $f$. The same is true about the operator $\prod_{e \in \partial f} L_e$. Thus $\prod_{e \in \partial f} L_e = \pm B_f$.

By construction, the boundary $\partial f$ alternates between link operators $L_e$ and logical-Z operators $Z_u$, see Fig. 4 for an example. For each link operator $L_e$ with $e \in \partial f$ define a quantity $\omega_f(L_e) = \pm 1$ such that $\omega_f(L_e) = -1$ iff $e$ is oriented clockwise with respect to $f$. Likewise, for each logical-Z operator $Z_u = ic pc_q$ lying on the boundary of $f$ define a quantity $\omega_f(Z_u) = \pm 1$ such that $\omega_f(Z_u) = -1$ iff an arrow $c_p \to c_q$ is oriented clockwise with respect to $f$. See Fig. 5 for examples of such arrows. A simple computation shows that $\prod_{e \in \partial f} L_e = -\omega_f B_f$, where

$$\omega_f = \prod_{e \in \partial f} \omega_f(L_e) \prod_{u \in f} \omega_f(Z_u).$$

Thus we need to check that $\omega_f = -1$ for each face $f \in F$. In other words, the boundary of each face must have an odd number of arrows oriented clockwise. Direct inspection shows that this is indeed the case for the distance-3 code, see Fig. 5. By translation invariance, this also holds for all code distances. The same arguments apply to X-type stabilizers.

We shall need an analogue of Lemma 1 for logical operators of the surface code.

**Lemma 2.** Let $X_L$ and $Z_L$ be the logical operators of the surface code located on the left and the top boundary.
Then $C4$-encodings of $X_L$ and $Z_L$ can be chosen as

$$X_L = ic_1c_2 \prod_{e \in \text{LEFT}} L_e \quad \text{and} \quad Z_L = ic_2c_3 \prod_{e \in \text{TOP}} L_e,$$

where LEFT and TOP are the subsets of edges lying on the left and the top boundaries of the lattice, see Fig. 6.

**Proof.** Let us add a “logical edge” connecting the modes $c_2$ and $c_3$ to the graph $G$, see Fig. 2. This creates an extra “logical face” $f$ attached to the top boundary of the lattice. The new edge carries a link operator $L_e = ic_2c_3$. The same arguments as in the proof of Lemma 1 show that $Z_L = -\omega_f \prod_{e \in \partial f} L_e$, where $\omega_f = \pm 1$ is the parity of the number of arrows lying on the boundary of $f$ and oriented clockwise with respect to $f$. From Fig. 5 one gets $\omega_f = -1$. The same argument applies to $X_L$. \qed

![Fig. 6. The sets of edges LEFT (red) and TOP (blue).](image)

**Lemma 3.** Let $\phi_{\text{link}}$ be the state of $4n - 4$ Majorana modes $c_5, c_6, \ldots, c_{4n}$ stabilized by all link operators,

$$|\phi_{\text{link}}\rangle\langle \phi_{\text{link}}| = \prod_{e \in E} \frac{1}{2}(I + L_e). \quad (13)$$

Let $\psi$ be any single-qubit state. Then

$$|\psi_L\rangle \sim \prod_{u \in V} \frac{1}{2}(I + S_u)|\psi\rangle \otimes |\phi_{\text{link}}\rangle. \quad (14)$$

Here we used the notations from Table II.

The lemma will allow us to replace the initial logical state $\psi_L$ in the error correction protocol by a simpler state $\phi_{\text{link}}$ at the cost of measuring certain additional stabilizers. We shall see that the state $\phi_{\text{link}}$ is a fermionic Gaussian state, see Section B for details. Furthermore, a state obtained from $\phi_{\text{link}}$ by applying a coherent error (encoded by the $C4$-code) is also Gaussian. These features will be instrumental for our simulation algorithm.

**Proof.** Suppose first that $|\psi\rangle = |0\rangle$. Let us add a pair of “logical edges” connecting modes $c_2, c_3$ and $c_1, c_4$ to the graph $G$. This creates an extra pair of “logical faces” attached to the top and the bottom boundaries. The new edges carry link operators $ic_2c_3$ and $ic_1c_4$. Lemmas 1,2 imply that the state $|\psi\rangle \otimes |\phi_{\text{link}}\rangle$ is stabilized by operators $B_f$ and $Z_f$. Furthermore, since these operators commute with all vertex stabilizers, the state on the right-hand side of Eq. (14) is stabilized by $B_f$ and $Z_f$. Since it is also stabilized by all $S_u$, this state has the same set of stabilizers as $|0_L\rangle$, which proves Eq. (14) for $|\psi\rangle = |0\rangle$. Note that$$X_L|0_L\rangle \otimes |\phi_{\text{link}}\rangle = |I_L\rangle \otimes |\phi_{\text{link}}\rangle,$$see Lemma 2. Furthermore, $X_L$ commutes with all vertex stabilizers. Thus applying $X_L$ to both sides of Eq. (14) with $|\psi\rangle = |0\rangle$ proves Eq. (14) for $|\psi\rangle = |1\rangle$. By linearity, it holds for all $\psi$. \qed

**IV. LOGICAL STATE PREPARATION**

We shall first describe the algorithm for simulating the logical state preparation because it is much simpler than the storage simulation. For each syndrome $s = \{s_f\}_{f \in F}$ define a syndrome projector

$$\Pi_s = \prod_{f \in F} \frac{1}{2}(I + s_f B_f). \quad (15)$$

It projects onto the subspace spanned by $n$-qubit states with the syndrome $s$. Note that $\sum_s \Pi_s = I$. Let $\Pi_0$ be the projector onto the logical subspace of the surface code. Since the Pauli correction $C_s$ maps any state with a syndrome $s$ to the logical subspace, it must satisfy

$$\Pi_s = C_s \Pi_0 C_s. \quad (16)$$

Here and below we assume that $C_s^4 = C_s$.

Suppose that our initial state has the product form

$$|\psi\rangle = |\psi_1 \otimes \psi_2 \otimes \cdots \otimes \psi_n\rangle.$$

Here $\psi_j$ are arbitrary single-qubit states. Our goal is to sample a syndrome $s$ from the probability distribution

$$p(s) = \langle \psi | \Pi_s | \psi \rangle \quad (17)$$

and compute the final logical state conditioned on the syndrome. The latter has the form

$$|\phi_s\rangle = \frac{1}{\sqrt{p(s)}} C_s \Pi_s |\psi\rangle \quad (18)$$

Let us first discuss how to simulate the syndrome measurement. We shall encode each qubit into the $C4$-code as discussed in Section III. Let $|\psi_a\rangle$ be the encoded version of $|\psi_a\rangle$. Using the Euler angle decomposition one
can write \(|\psi_a\rangle = e^{-i\alpha X}e^{-i\beta Z}e^{-i\gamma X}|0\rangle\). Replacing Pauli operators by their C4-encodings defined in Eq. (8) gives
\[
|\bar{\psi}_a\rangle = \exp(\alpha c_1 c_2) \exp(\beta c_2 c_3) \exp(\gamma c_1 c_2) |0\rangle.
\]
The state \(|0\rangle\) is stabilized by \(Z = ic_2 c_3\) and \(\bar{Z} = ic_1 c_4\), see Eq. (8). Thus the states \(|\bar{\psi}_a\rangle\) can be prepared using only FLO gates. Applying this independently to each qubit gives a sequence of \(O(n)\) FLO gates that prepares the state
\[
|\bar{\psi}\rangle = |\bar{\psi}_1 \otimes \cdots \otimes \bar{\psi}_n\rangle.
\]
Clearly, measuring syndromes of stabilizers \(B_f\) on the state \(|\psi\rangle\) is equivalent to measuring syndromes of the encoded stabilizers \(\bar{B}_f\) on the state \(|\bar{\psi}\rangle\). By Lemma 1, the latter can be reduced to measuring syndromes \(m_e = \pm 1\) of the link operators \(L_e\) and then classically computing face syndromes \(s_f = \prod_{e \in \partial f} m_e\). Since each link operator has a form \(L_e = ic_p c_q\), measuring the eigenvalue of \(L_e\) is a FLO gate. Thus we have realized the full syndrome measurement using \(O(n)\) FLO gates.

It remains to compute the final logical state \(\psi_s\). Let \(\bar{b}_s = (\bar{b}_x^s, \bar{b}_y^s, \bar{b}_z^s)\) be the Bloch vector of \(\psi_s\) such that
\[
\bar{b}_s = \langle \phi_s | X_L | \phi_s \rangle, \quad b_x^s = \langle \phi_s | Y_L | \phi_s \rangle, \quad b_z^s = \langle \phi_s | Z_L | \phi_s \rangle.
\]
Below we focus on computing \(b_x^s\). Define \(\lambda_s = \pm 1\) such that \(C_s X_L = \lambda_s X_L C_s\). Then
\[
b_x^s = \lambda_s \frac{\langle \psi | \Pi_s X_L | \psi \rangle}{\langle \psi | \Pi_s | \psi \rangle} = \lambda_s \frac{\langle \bar{\psi} | \Pi_s \bar{X}_L | \bar{\psi} \rangle}{\langle \bar{\psi} | \Pi_s | \bar{\psi} \rangle}.
\]
(19)

Here in the first equality we noted that \(X_L\) commutes with \(\Pi_s\). In the second equality we encoded every qubit using the C4-code. Let \(m \in \{+1,-1\}\) be the combined syndrome of link operators measured in the first part of the algorithm such that \(m_e\) is the measured eigenvalue of \(L_e\). Define the corresponding syndrome projector
\[
\Pi^\text{link}_m = \prod_{e \in E} \frac{1}{2}(I + m_e L_e).
\]
We claim that
\[
b_x^s = \lambda_s \prod_{e \in \text{LEFT}} m_e \cdot \frac{\langle \bar{\psi} | \Pi^\text{link}_m i c_1 c_2 | \bar{\psi} \rangle}{\langle \bar{\psi} | \Pi^\text{link}_m | \bar{\psi} \rangle}.
\]
(20)

Here \(\text{LEFT}\) is the subset of edges lying on the left boundary of the lattice, see Fig. 6. The ratio in Eq. (20) can be computed by taking the normalized state \(\Pi^\text{link}_m | \bar{\psi} \rangle\) obtained after measuring the link syndromes (the first part of the algorithm) and measuring the eigenvalue of \(i c_1 c_2\). The latter requires a single FLO gate. This gives the desired value of \(b_x^s\). Likewise, measuring the eigenvalues of \(i c_2 c_3\) and \(-ic_1 c_3\) on the final state \(\Pi^\text{link}_m | \bar{\psi} \rangle\) gives the remaining components of the Bloch vector \(b_x^s\) and \(b_y^s\) respectively.

To conclude, FLO gates enable simulation of the syndrome measurement and computation of the final logical Bloch vector conditioned on the measured syndrome. The resulting FLO circuit can be simulated classically in time \(O(n^3)\) since it includes \(O(n)\) projection gates \((1/2)(I + m_e L_e)\). The simulation runtime can be significantly improved using the following simple observations. First, once a link operator \(L_e = ic_p c_q\) has been measured, the modes \(c_p, c_q\) are completely disentangled from the rest of the system. Such disentangled modes can be removed from the simulator reducing the total number of modes and the computational cost of subsequent steps. Second, we can exploit the fact that the initial state \(\bar{\psi}\) has a product form. In particular, a four-mode cluster \(\Gamma_u\) that supports the state \(\bar{\psi}_u\) only needs to be loaded into the simulator at a time step when some mode \(p \in \Gamma_u\) participates in the measurement of a link operator. Thus at any given time step the simulator only needs to keep track of “active” clusters \(\Gamma_u\) such that at least one mode \(p \in \Gamma_u\) has been measured and at least one mode \(q \in \Gamma_u\) has not been measured. One can easily choose the order of measurements such that the number of active clusters is \(O(n^{1/2})\) at any time step (for example, one can measure link operators column by column). Accordingly, the cost of simulating a single FLO gate is at most \(O(n)\). Since the number of gates is \(O(n)\), the total simulation cost is \(O(n^2)\).

It remains to prove Eq. (20). Let \(\delta\) be a map from link syndromes to the corresponding face syndromes, that is, \(s = \delta(m)\) iff \(s_f = \prod_{e \in \partial f} m_e\) for all \(f \in F\). By Lemma 1,
\[
\Pi_s = \sum_{m : \delta(m) = s} \Pi^\text{link}_m.
\]
(21)

Below we prove the following

**Proposition 1.** Suppose \(m\) and \(m'\) are link syndromes such that \(\delta(m) = \delta(m')\). Then there exists a subset of vertices \(W \subseteq V\) such that
\[
\Pi^\text{link}_m = T T^\text{link}_{m'} T, \quad \text{where} \quad T = \prod_{u \in W} S_u.
\]

We postpone the proof until the end of the section. Let \(m\) and \(m'\) be any link syndromes with \(\delta(m) = \delta(m')\). Then the proposition implies that
\[
\langle \bar{\psi} | \Pi^\text{link}_{m'} | \psi \rangle = \langle \bar{\psi} | T T^\text{link}_m T | \psi \rangle = \langle \bar{\psi} | \Pi^\text{link}_m | \psi \rangle
\]
(22)
since \(S_u \bar{\psi} = \bar{\psi}\) for all \(u\). Likewise,
\[
\langle \bar{\psi} | \Pi^\text{link}_{m'} \bar{X}_L | \psi \rangle = \langle \bar{\psi} | \Pi^\text{link}_m \bar{X}_L | \psi \rangle
\]
(23)
since \(\bar{X}_L\) commutes with \(T\). From Eqs. (19,21,22,23) one gets
\[
b_x^s = \lambda_s \frac{\langle \bar{\psi} | \Pi^\text{link}_m \bar{X}_L | \psi \rangle}{\langle \bar{\psi} | \Pi^\text{link}_m | \psi \rangle}
\]
(24)
for any link syndrome \(m\) such that \(\delta(m) = s\). In particular, one can choose \(m\) as the link syndrome measured in the first part of the algorithm. Substituting \(\bar{X}_L\) from Lemma 2 gives the desired result Eq. (20). It remains to prove Proposition 1.
Proof. By linearity, we can assume that $m$ is the trivial syndrome, that is, $m_e = 1$ for all $e \in E$. Then $\delta(m') = \delta(m)$ iff $m'$ is a flat connection on the surface code lattice, that is, $m'$ is a 1-chain with $\mathbb{Z}_2$ coefficients such that the $\mathbb{Z}_2$-valued magnetic flux through every face is $+1$. Since the lattice is topologically trivial, any flat connection is a co-boundary of some 0-chain $W \subseteq V$. Since a vertex stabilizer $S_u$ flips the link syndromes on all edges incident to $u$, the condition that $m'$ is a co-boundary of $W$ is equivalent to $\prod_{m'} = T T_{m'}^T$ for $T = \prod_{u \in W} S_u$. \hfill \Box

V. STORAGE OF A LOGICAL STATE

Assume now that our initial state has the form

$$|\psi\rangle = U|\psi_L\rangle, \quad U = e^{i \eta_1 Z} \otimes e^{i \eta_2 Z} \otimes \cdots \otimes e^{i \eta_n Z}$$

where $\psi_L$ is some (unknown) logical state of the surface code and $\eta_1, \ldots, \eta_n$ are arbitrary rotation angles. The unitary $U$ describes a coherent error applied to each qubit before the syndrome measurement. Our goal is to sample a syndrome $s$ from the probability distribution

$$p(s) = \langle \psi | \Pi_s | \psi \rangle$$

and compute the final logical state conditioned on the syndrome,

$$|\phi_s\rangle = \frac{1}{\sqrt{p(s)}} C_s \Pi_s |\psi\rangle.$$  \hspace{1cm} (26)

Clearly, since $\psi$ contains only $Z$-type errors, the observed syndrome of $Z$-stabilizers is always trivial. Accordingly, we shall assume that the correction $C_s$ is a $Z$-type Pauli. We shall need the following fact.

**Lemma 4.** The probability $p(s)$ does not depend on the initial logical state $\psi_L$. The map $\psi_L \rightarrow \phi_s$ is a logical $Z$-rotation by some angle $\theta_s \in [0, \pi]$, that is,

$$|\phi_s\rangle = \exp [i \theta_s Z_L] |\psi_L\rangle.$$  \hspace{1cm} (27)

We shall refer to $\theta_s$ as a logical rotation angle. Here and below all states are defined modulo an overall phase factor. We defer the proof of the lemma to Appendix A.

A. Simulating the syndrome measurement

Let us first discuss how to sample $s$ from $p(s)$. By Lemma 4, $p(s)$ does not depend on $\psi_L$, so below we set $|\psi_L\rangle = |+L\rangle$. Since only syndromes of $X$-stabilizers may be non-trivial, it suffices to measure eigenvalues $m_u = \pm 1$ of single-qubit Pauli operators $X_u$ and then classically compute the face syndrome $s_f = \prod_{u \in f} m_u$ for each face $f$ that supports an $X$-stabilizer. Let $m \in \{+1, -1\}^V$ be the combined $X$-measurement outcome and $p^x(m)$ be the probability of an outcome $m$. We have

$$p^x(m) = \langle \psi | \Pi^x_m | \psi \rangle, \quad \Pi^x_m = \prod_{u \in V} \frac{1}{2} (I + m_u X_u).$$  \hspace{1cm} (28)

Let us order qubits column by column as shown on Fig. 7. Let $p^x_t(m_1, \ldots, m_t)$ be the marginal distribution of $p^x(m)$ describing the first $t$ qubits and let

$$p^x_t(m_t|m_1, \ldots, m_{t-1}) = \frac{p^x_t(m_1, \ldots, m_t)}{p^x_{t-1}(m_1, \ldots, m_{t-1})}$$  \hspace{1cm} (29)

be the conditional distribution of $m_t$ given $m_1, \ldots, m_{t-1}$.

Let us show how to sample $m_t$ from the distribution Eq. (29) using FLO gates. Partition the set of all qubits as $|n\rangle = AB$ where $A = \{1, \ldots, t\}$ and $B = |n\rangle \setminus A$. We shall write

$$m_A = (m_1, \ldots, m_t), \quad U_A = \prod_{u \in A} e^{i \eta_u Z_u},$$

and

$$\Pi^x_A = \prod_{u \in A} \frac{1}{2} (I + m_u X_u).$$

Then $p^x_t(m_A) = \langle \psi_L | U_A^\dagger \Pi^x_A U_A |\psi_L\rangle$. Encoding each qubit into the $C_4$-code as discussed in Section III gives

$$p^x_t(m_A) = \langle \bar{\psi}_L | U_A^\dagger \Pi^x_A U_A |\bar{\psi}_L\rangle.$$  \hspace{1cm} (30)

The Majorana representation of the logical state $\bar{\psi}_L$ defined in Lemma 3 gives

$$|\bar{\psi}_L\rangle = \sqrt{\gamma} \prod_{u \in V} \frac{1}{2} (I + S_u |\bar{\psi}_L\rangle \otimes |\phi_{\text{link}}\rangle)$$  \hspace{1cm} (31)

where $\gamma$ is a normalizing coefficient depending only on $n$, $\phi_{\text{link}}$ is a product state defined in Eq. (13), and $|\bar{\psi}_L\rangle$ is the basis state of modes $c_1, c_2, c_3, c_4$ stabilized by $ic_1 c_2$ and $ic_3 c_4$ (recall that we have chosen $|\psi_L\rangle = |+L\rangle$). The state $|\bar{\psi}_L\rangle \otimes |\phi_{\text{link}}\rangle$ can be prepared using FLO gates since it is stabilized by two-mode operators $L_c, ic_1 c_2$ and $ic_3 c_4$. To simplify the notation, we shall absorb the pairs $ic_1 c_2$ and $ic_3 c_4$ into $\phi_{\text{link}}$. Accordingly, below we assume that $\phi_{\text{link}}$ is a state of $4n$ modes defined as

$$|\phi_{\text{link}}\rangle \langle \phi_{\text{link}}\rangle = \frac{1}{2} (I + ic_1 c_2) \frac{1}{2} (I + ic_3 c_4) \prod_{c \in E} \frac{1}{2} (I + L_c).$$  \hspace{1cm} (32)

Plugging Eq. (31) into Eq. (30) gives

$$p^x_t(m_A) = \langle \phi_{\text{link}} | U_A^\dagger \Pi^x_A U_A |\phi_{\text{link}}\rangle,$$  \hspace{1cm} (33)
where \( \Omega = \prod_{u \in A} \frac{1}{2} (I + S_u) \) is the projector onto the codespace of the \( C_4 \)-code. Write \( \Omega = \Omega_A \Omega_B \), where

\[
\begin{align*}
\Omega_A &= \prod_{u \in A} \frac{1}{2} (I + S_u) \quad \text{and} \\
\Omega_B &= \prod_{u \in B} \frac{1}{2} (I + S_u).
\end{align*}
\]

Since \( \overline{U}_A \) and \( \overline{\Pi}^x_A \) commute with \( \Omega_A \), one gets

\[
P^x_t(m_A) = \gamma \langle \phi_{\text{link}} | \overline{U}_A (\overline{\Pi}^x_A \Omega_A) \overline{U}_A (I + S_B) | \phi_{\text{link}} \rangle.
\]  
(34)

Assume first that \( B \neq \emptyset \). Expand the projector \( \Omega_B \) as

\[
\Omega_B = 2^{-|B|} \sum C \subseteq B S_C, \quad S_C = \prod_{u \in C} S_u.
\]

Consider a link operator \( L_e \) associated with some edge \( e \) such that both endpoints of \( e \) belong to \( B \). Such a link operator commutes with all operators acting on \( A \). Note that \( L_e \) commutes with all vertex stabilizers \( S_u \) except for the two stabilizers located at the endpoints of \( e \). It follows that \( L_e S_C L_e = -S_C \) if exactly one endpoint of \( e \) belongs to \( C \). Furthermore, since \( \phi_{\text{link}} \) is stabilized by all link operators one infers that \( \langle \phi_{\text{link}} | O_A S_C | \phi_{\text{link}} \rangle = 0 \) for any operator \( O_A \) unless \( C = \emptyset \) or \( C = B \). Here we used the fact that \( B \) is a connected set for any \( t \), see Fig. 7.

Substituting the expansion of \( \Omega_B \) into Eq. (34) and using the above observation gives

\[
P^x_t(m_A) = \gamma' \langle \phi_{\text{link}} | \overline{U}_A (\overline{\Pi}^x_A \Omega_A) \overline{U}_A (I + S_B) | \phi_{\text{link}} \rangle,
\]  
(35)

where \( \gamma' = 2^{-|B|} \gamma \). Next we note that \( \phi_{\text{link}} \) is stabilized by \( S_A S_B \) since the latter coincides with the product of all link operators \( L_e \) (including \( ic_1 c_2 \) and \( ic_3 c_4 \)) and \( \phi_{\text{link}} \) is stabilized by any link operator. Thus one can replace \( S_B \) by \( S_A \) in Eq. (35). However, since \( S_A \) commutes with \( \overline{U}_A \) and \( \gamma \Omega_A = \Omega_A \), we arrive at

\[
P^x_t(m_A) = 2 \gamma' \langle \phi_{\text{link}} | \overline{U}_A (\overline{\Pi}^x_A \Omega_A) \overline{U}_A (I + S_B) | \phi_{\text{link}} \rangle.
\]  
(36)

Using the identity

\[
\overline{\Pi}^x_A \Omega_A = \prod_{u \in A} \frac{1}{2} (I + m_a \overline{X}_u) \frac{1}{2} (I + m_a \overline{X}_u S_u)
\]

one finally gets

\[
P^x_t(m_A) = 2^{t+1-n} \gamma \| G_{2a} G_{2a-1} \cdots G_2 G_1 \phi_{\text{link}} \|^2
\]  
(37)

where \( G_{2a-1}, G_{2a} \) are operators acting non-trivially only on the subset of modes \( \Gamma_a \), namely

\[
G_{2a-1} = \frac{1}{2} (I + m_a \overline{X}_a) e^{\eta_a} Z_a,
\]

and

\[
G_{2a} = \frac{1}{2} (I + m_a \overline{X}_a S_a).
\]

Noting that \( \overline{Z}_a, \overline{X}_a \) and \( \overline{X}_a S_a \) have the form \( i c_p c_q \) for some \( p, q \in \Gamma_a \), see Eqs. (7,8), one concludes that Eq. (37) includes only FLO gates.

The above arguments also apply to the case \( B = \emptyset \), that is, \( t = n \). The only difference is that now Eq. (35) has no term \( S_B \) and thus Eq. (37) becomes

\[
p^x_t(m) = \gamma \| G_{2a} G_{2a-1} \cdots G_2 G_1 \phi_{\text{link}} \|^2.
\]  
(38)

Let us discuss the cost of sampling \( m \) from \( p^x_t(m) \). First, observe that any single-qubit marginal state of \( \psi_L \) is maximally mixed since the surface code has no stabilizers of weight one. Since the same is true about the state \( U | \psi_L \rangle \), one can pick the first measurement outcome \( m_1 \) at random from the uniform distribution. Suppose we have already sampled \( m_1, \ldots, m_{t-1} \) and the simulator’s current state is

\[
| \phi_{t-1} \rangle = G_{2t-2} G_{2t-3} \cdots G_2 G_1 | \phi_{\text{link}} \rangle.
\]  
(39)

Plugging Eqs. (37,38) into Eq. (29) gives

\[
p^x_{t}(m_t | m_1, \ldots, m_{t-1}) = \frac{\gamma t \| G_{2t} G_{2t-1} \phi_{t-1} \|^2}{\| \phi_{t-1} \|^2},
\]  
(40)

where \( \gamma_t = 2 \) for \( t < n \) and \( \gamma_n = 1 \). The conditional probability of the outcome \( m_t = 1 \) can be computed by simulating two more FLO gates \( G_{2t+1}, G_{2t} \) starting from the state \( \phi_{t-1} \) which takes time \( O(n^2) \). Once the conditional probability is computed, one can sample \( m_t \) by tossing a suitably biased coin. This produces the next syndrome \( m_t \) together with the next state \( \phi_t \). After \( n \) iterations one gets the desired sample \( m \) from \( p^x_t(m) \). Since the above algorithm uses \( O(n) \) FLO gates, the simulation runtime scales as \( O(n^3) \).

As before, we can reduce the runtime to \( O(n^2) \) by exploiting the fact that the initial state \( \phi_{\text{link}} \) has a product form and by observing that once a pair of modes have been measured, it can be removed from the simulator. Indeed, consider some intermediate step \( t \) and let \( j \) be the column of the lattice that contains \( t \)-th qubit. Then all modes in the column \( j \) are now available to be loaded into the simulator. Likewise, all modes in the columns \( j+2, \ldots, d \) can be grouped into entangled pairs located on edges such that each pair is stabilized by the respective link operator \( L_e \). Such entangled pairs do not need to be loaded into the simulator. Thus at any given time step the number of “active” modes that needs to be simulated is only \( O(n^{1/2}) \). Accordingly, the cost of simulating a single FLO gate is at most \( O(n) \). Since the number of gates is \( O(n) \), the total simulation cost is \( O(n^2) \).

**Remark 1:** The same reasoning shows that the probability \( p^x_t(m) \) of any given outcome \( m \) can be computed up to the normalizing coefficient \( \gamma \) in time \( O(n^2) \) by simulating the FLO circuit defined in Eq. (38). Furthermore, the normalizing coefficient \( \gamma \) depends only on \( n \) (but not on the rotation angles \( \eta_a \)).

**Remark 2:** Below we shall also use a slightly modified version of the above algorithm where the initial state \( \psi_L \) is an eigenvector of the logical-\( Y \) operator. The modified version is exactly the same as above except that the
initial state $\phi_{\text{link}}$ in Eq. (32) is defined as
\[
|\phi_{\text{link}}\rangle = \frac{1}{2}(I - ic_1 c_3)\frac{1}{2}(I + ic_2 c_4) \prod_{e \in E} \frac{1}{2}(I + L_e)
\]
This corresponds to initializing the unpaired modes in the logical-Y state.

B. Computing the logical rotation angle

It remains to show how to compute the logical rotation angle $\theta_s$ for a given syndrome $s$. Let us first initialize the logical qubit in the X-basis, that is, we choose $|\psi_L\rangle = |+_L\rangle$. Let $\phi_s$ be the final logical state defined in Eq. (26).

Define logical amplitudes
\[
A_s^+ = \langle +_L|\phi_s\rangle = \cos(\theta_s)
\]
and
\[
A_s^- = \langle +_L|Z_L|\phi_s\rangle = i \sin(\theta_s).
\]
Here we used Lemma 4. Then
\[\tan^2(\theta_s) = \left| \frac{A_s^+}{A_s^-} \right|^2. \tag{44}\]
Using Eq. (26) and the identity $\Pi_s = C_s \Pi_0 C_s$ one gets
\[\tan^2(\theta_s) = \frac{|\langle +_L|Z_L C_s U|+_L\rangle|^2}{|\langle +_L|C_s U|+_L\rangle|^2}. \tag{45}\]
Let $U_+ = C_s U$ and $U_- = Z_L C_s U$. By definition, $U_{\pm}$ are products of single-qubit Z rotations:
\[U_{\pm} = \prod_{u=1}^n e^{i \eta_u} z_u. \]
Let us expand the logical state $|+_L\rangle$ in the Z-basis:
\[|+_L\rangle = |\mathcal{L}\rangle^{-1/2} \sum_{x \in \mathcal{L}} |x\rangle,
\]
where $\mathcal{L}$ is the set of basis states $x \in \{0,1\}^n$ that obey Z-stabilizers of the surface code (we do not need an explicit formula for $\mathcal{L}$). Since $U_{\pm}$ is diagonal in the Z-basis,
\[\langle +_L|U_{\pm}|+_L\rangle = 2^{n/2} |\mathcal{L}|^{-1/2} \langle \oplus^n |U_{\pm}|+_L\rangle. \tag{46}\]
Substituting this into Eq. (44) gives
\[\tan^2(\theta_s) = \frac{|\langle \oplus^n |U_-|+_L\rangle|^2}{|\langle \oplus^n |U_+|+_L\rangle|^2} = \frac{p_-}{p_+}. \tag{47}\]
Since $U_{\pm}$ is a tensor product of Z-rotations, $p_{\pm}$ is a special case of the probability $p^Z(m)$ defined in Eq. (28) with $m_u = 1$ for all $u$. We have already shown that one can compute $\gamma^{-1}p_{\pm}$ in time $O(n^2)$, where $\gamma$ depends only on $n$, see Remark 1 at the end of Section V A. Thus $\tan^2(\theta_s) = \gamma^{-1}p_{\pm}$ can be computed in time $O(n^2)$.

Next let us initialize the logical qubit in the $Y$-basis state: $|\psi_L\rangle = |Y_L\rangle$, where $|Y\rangle \equiv (|0\rangle + |1\rangle)/\sqrt{2}$. Define logical amplitudes
\[B_s^+ = \langle +_L|\phi_s\rangle = e^{i\pi/4}(\cos(\theta_s) + \sin(\theta_s)) \tag{48}\]
and
\[B_s^- = \langle +_L|Z_L|\phi_s\rangle = e^{-i\pi/4}(\cos(\theta_s) - \sin(\theta_s)). \tag{49}\]
The same arguments as above show that
\[\tan^2(\theta_s - \pi/4) = \frac{|\langle \oplus^n |U_-|Y_L\rangle|^2}{|\langle \oplus^n |U_+|Y_L\rangle|^2} = \frac{q_-}{q_+}. \tag{50}\]
Since $U_{\pm}$ is a product of Z-rotations, $q_{\pm}$ is a special case of the probability $p^Y(m)$ defined in Eq. (28) with $m_u = 1$ for all $u$ and the initial state $\psi_s$ chosen as a logical $Y$-basis state. We have already shown that one can compute $q_{\pm}$ in time $O(n^2)$ see Remarks 1,2 at the end of Section V A. Thus $\tan^2(\theta_s - \pi/4) = q_-/q_+$ is computable in time $O(n^2)$. Combining Eqs. (47,50) gives
\[\cos(2\theta_s) = \frac{p_+ - p_-}{p_+ + p_-} \quad \text{and} \quad \sin(2\theta_s) = \frac{q_+ - q_-}{q_+ + q_-}. \tag{51}\]
This determines the logical rotation angle modulo $\pi$.

VI. NUMERICAL RESULTS

We implemented the algorithms described above for translation-invariant coherent noise and surface codes with distance $5 \leq d \leq 49$. The smallest distance $d = 3$ was skipped because of strong finite-size effects (note that the considered surface codes are only defined for odd values of $d$). We used the maximum distance $d = 37$ for storage simulations and $d = 49$ for state preparation simulations. The logical error rate $P_L$ was estimated by the Monte Carlo method with at least 50,000 syndrome samples per data point. (The only exception is Fig. 12 where we used 5,000 syndrome samples per data point.)

A. Numerical results for storage

Consider first the protocol A for storage of a logical state in the presence of Z-rotation errors. In this section we only consider translation-invariant errors of the form $\exp(i\theta Z)^{\oplus n}$, where $\theta$ is the only noise parameter. Recall that we define the logical error rate as
\[P_L = 2 \sum_s p(s)|\sin\theta_s|, \tag{52}\]
where $p(s)$ is the probability of observing a syndrome $s$ and $\theta_s$ is the logical rotation angle conditioned on the
syndrome, see Lemma 4 in Section V. To motivate this definition, consider a conditional logical channel
\[ \Lambda_s(\rho) = e^{i\theta_s Z} \rho e^{-i\theta_s Z} \]  
(53)
that describes the residual logical error for a given syndrome \( s \). Let \( \| : \|_\diamond \) denote the diamond-norm \cite{40} on the space of quantum channels and \( \text{id} \) be the single-qubit identity channel. The identity \( \|\Lambda_s - \text{id}\|_\diamond = 2|\sin \theta_s| \) shows that \( P^L \) coincides with the average diamond-norm distance between the conditional logical channel and the identity channel,
\[ P^L = \sum_s p(s)\|\Lambda_s - \text{id}\|_\diamond. \]
Using the symmetries of the surface code one can easily check that \( P^L \) is invariant under flipping the sign of \( \theta \). Accordingly, it suffices to simulate \( \theta \geq 0 \).

Our numerical results for the logical error rate are presented in Fig. 8. The data suggests that the quantity \( P^L \) decays exponentially in the code distance \( d \) for \( \theta < \theta_0 \), where
\[ 0.08\pi \leq \theta_0 \leq 0.1\pi \]  
(54)
can be viewed as an error correction threshold. We observe the exponential decay of \( P^L \) as a function of \( d \) in the sub-threshold regime.

Although the logical error rate \( P^L \) is a meaningful measure of how well the initial logical state is preserved, it provides no insight into the structure of residual logical errors. Algorithm A gives us a unique opportunity to investigate the logical-level noise since it outputs both the syndrome \( s \) and the logical rotation angle \( \theta_s \) conditioned on \( s \). Fig. 9 shows the empirical probability distribution of \( \theta_s \) obtained by sampling \( 10^6 \) syndromes \( s \) for the physical \( Z \)-rotation angle \( \theta = 0.08\pi \) (which we expect to be slightly below the threshold). We compare the cases \( d = 9 \) and 25. In both cases the distribution has a sharp peak at \( \theta_s = 0 \) (equivalent to \( \theta_s = \pi \)). This peak indicates that error correction almost always succeeds in the considered regime. For ease of visualization, we truncated the peak at \( \theta_s = 0 \) on the histograms. It can be seen that increasing the code distance has a dramatic effect on the distribution of \( \theta_s \). The distance-9 code has a broad distribution of \( \theta_s \) meaning that the logical-level noise retains a strong coherence. On the other hand, the distance-25 code has a sharply peaked distribution of \( \theta_s \) with a peak at \( \theta_s = \pi/2 \) which corresponds to the logical Pauli error \( Z_L \). Such errors are likely to be caused by “ambiguous” syndromes \( s \) for which the minimum weight matching decoder makes a wrong choice of the Pauli correction \( C_s \). We conclude that as the code distance increases, the logical-level noise can be well approximated by random Pauli errors even though the physical-level noise is coherent.

To investigate this effect more systematically, it is desirable to have a metric quantifying the degree of coherence present in the logical-level noise. To this end let us consider the twirled version of the logical channel \( \Lambda_s \),
\[ \Lambda_s^{\text{twirl}}(\rho) = (1 - \epsilon_s)\rho + \epsilon_s Z\rho Z, \quad \epsilon_s \equiv \sin^2 \theta_s, \]
and the corresponding logical error rate
\[ P^L_{\text{twirl}} = \sum_s p(s)\|\Lambda_s^{\text{twirl}} - \text{id}\|_\diamond = 2 \sum_s p(s)\sin^2 \theta_s. \]  
(55)
Comparison of Eqs. (52,55) reveals that \( P^L \geq P^L_{\text{twirl}} \) with the equality iff the distribution of \( \theta_s \) has all its weight.
FIG. 9. These histograms show the empirical probability distribution of logical rotation angles $\theta_s$ for the code distance $d = 9$ (left) and $d = 25$ (right). The histograms use the same noise parameter $\theta = 0.08 \pi$. For ease of visualization, we truncated the main peak at $\theta_s = 0$.

FIG. 10. Coherence ratio $P^L/P^L_{\text{twirl}}$ for the conditional logical channel (left) and for the average logical channel (right). In both cases increasing the code distance makes the logical-level noise less coherent.

on $\{0, \pi/2\}$, that is, when the logical noise is incoherent. It is therefore natural to measure coherence of the logical noise by the ratio $P^L/P^L_{\text{twirl}}$. This “coherence ratio” is plotted as a function of $\theta$ on Fig. 10(a). The data indicates that the coherence ratio decreases for increasing system size approaching one for large code distances. This further supports the conclusion that the logical noise has a negligible coherence. Finally, in Fig. 10(b), we show the analogous quantity for the average logical noise channel [20] defined as

$$\Lambda = \sum_s p(s) \Lambda_s.$$  

This average channel provides an appropriate model for the logical-level noise if the environment has no access to the measured syndrome. This may be relevant, for instance, in the quantum communication settings where noise acts only during transmission of information. Thus
one can alternatively define the coherence ratio as

\[ P_L / P_{\text{twirl}} = \frac{\| \Lambda - \text{id} \|_\infty}{\| \Lambda^{\text{twirl}} - \text{id} \|_\infty}, \]

where \( \Lambda^{\text{twirl}} \) is the Pauli-twirled version of \( \Lambda \). In our case \( \Lambda(\rho) = (1 - \epsilon)\rho + \epsilon Z\rho Z + i\delta(Z\rho - \rho Z) \), where \( \epsilon = \sum_s p(s) \sin^2(\theta_s) \) and \( \delta = \sum_s p(s) \sin(2\theta_s)/2 \), see Eq. (53). A simple calculation yields

\[ \| \Lambda - \text{id} \|_\infty = 2\sqrt{\epsilon^2 + \delta^2} \quad \text{and} \quad \| \Lambda^{\text{twirl}} - \text{id} \|_\infty = 2\epsilon. \]

The coherence ratio of the average logical channel is plotted as a function \( \theta \) on Fig. 10(b). It provides a particularly strong evidence that in the limit of large code distances, coherent physical noise gets converted into incoherent logical noise.

Finally, let us compare logical error rates \( P_L \) computed for coherent physical noise \( \mathcal{N}(\rho) = e^{i\theta Z} \rho e^{-i\theta Z} \) and its Pauli-twirled version \( \mathcal{N}_{\text{twirl}}(\rho) = (1 - \epsilon)\rho + \epsilon Z\rho Z \) with \( \epsilon = \sin^2(\theta) \). Applying the Pauli twirl at the physical level amounts to ignoring the coherent part of the noise. Let \( P_L(\mathcal{N}_{\text{twirl}}) \) be the logical error rate corresponding to \( \mathcal{N}_{\text{twirl}} \). The plot of \( P_L(\mathcal{N}_{\text{twirl}}) \) and the ratio \( P_L/P_L(\mathcal{N}_{\text{twirl}}) \) are shown on Fig. 11. It can be seen that applying the Pauli twirl approximation to the physical noise gives an accurate estimate of the error threshold but significantly underestimates the logical error probability in the sub-threshold regime. We conclude that coherence of noise may have a profound effect on the performance of large surface codes in the sub-threshold regime which is particularly important for quantum fault-tolerance.

B. Numerical results for state preparation

Next consider the protocol B for preparing the logical basis state \(|\pm_L\rangle\) by performing syndrome measurements on the initial product state

\[ \langle \exp(-i\varphi X) \rangle \langle \exp(i\theta Z) \rangle^\otimes n. \]

Here \( \varphi, \theta \in [0, \pi) \) are noise parameters. The ideal protocol corresponds to \( \theta = 0 \). Define the logical error rate \( P^L \) as the average trace-norm distance between the final logical state \( |\phi_\phi\rangle \langle \phi_\phi| \) and the target state \(|\pm_L\rangle \langle \pm_L|\).

Equivalently,

\[ P^L = 2^{1/2} \sum_s p(s) \sqrt{1 - \langle \phi_\phi | X_L | \phi_\phi \rangle}. \]

Since the considered noise model generates correlations between \( X \)- and \( Z \)-syndromes, we opted not to use the minimum-weight matching decoder (which treats \( X \)- and \( Z \)-syndromes independently). Instead, we used a simplified decoder that chooses a Pauli correction \( C_s \) such that the final logical state \( \phi_\phi \) always obeys \( \langle \phi_\phi | X_L | \phi_\phi \rangle \geq 0 \). The simplified decoder is optimal in the sense that it minimizes the logical error rate \( P^L \) under the constraint that \( C_s \) is a Pauli operator. Note that \( 0 \leq P^L \leq \sqrt{2} \) for all noise parameters. The symmetries of the surface code imply that \( P^L \), considered as a function of \( \theta \) and \( \varphi \), is invariant under transformations

\[ \theta \leftrightarrow \theta + \pi/2, \quad \varphi \leftrightarrow \varphi + \pi/2, \quad \theta \leftrightarrow -\theta, \quad \varphi \leftrightarrow -\varphi. \]

Thus it suffices to simulate the region \( 0 \leq \theta, \varphi \leq \pi/4 \).
Our numerical results spell good news for quantum engineers pursuing surface code realizations: thresholds for state preparation and storage are reasonably high, suggesting that coherent noise is not as detrimental as one could expect from the previous studies. The numerical investigation of the logical-level noise gives rise to a conceptually appealing conjecture: error correction converts coherent physical noise to incoherent logical noise (for large code sizes). Whether this is an artifact of the considered error correction scheme or manifestation of a more general phenomenon is an interesting open question.

Although we simulated only translation-invariant noise models, all our algorithms apply to more general qubit-dependent noise. This enables numerical study of recently proposed state injection protocols [41], e.g. preparation of logical magic states, in the presence of coherent errors. Another possible application could be testing the so-called disorder assisted error correction method [42–44] where artificial randomness introduced in the code parameters suppresses coherent propagation of errors due to the Anderson localization phenomenon. We leave as an open question whether our algorithms can be extended to more general coherent noise models such as those including systematic cross-talk errors.

ACKNOWLEDGMENTS

The authors thank Steven Flammia, Jay Gambetta, and David Gosset for helpful discussions. RK is supported by the Technical University of Munich – Institute for Advanced Study, funded by the German Excellence Initiative and the European Union Seventh Framework Programme under grant agreement no. 291763. He acknowledges partial support by the National Science Foundation under Grant No. NSF PHY-1125915, and thanks the Kavli Institute for Theoretical Physics for their hospitality. NP acknowledges support by TUM-PREP. SB acknowledges support from the IBM Research Frontier Institute and from Intelligence Advanced Research Projects Activity (IARPA) under contract W911NF-16-0114.

Appendix A: Proof of Lemma 4

We have \( C_s = Z(h_s) \) for some \( h_s \in \{0,1\}^n \). Using the identities \( \Pi_s = C_s \Pi_0 C_s \) and \( \Pi_0 \psi_L = \psi_L \) one gets

\[
p(s) = \| \Pi_0 C_s U \Pi_0 \psi_L \|^2.
\]

(A1)

Expanding the error \( U \) in the Pauli basis gives

\[
U = \sum_{g \in \{0,1\}^n} \alpha_g i^{\mid g \mid} Z(g)
\]
for some real coefficients \(\alpha_g\). Here \(|g|\) denotes the Hamming weight of a string \(g\). Thus

\[
\Pi_0 C_s U \Pi_0 = \sum_{g \in \{0,1\}^n} \alpha_g |g|\Pi_0 Z(g \oplus h_s) \Pi_0.
\]

Note that \(\Pi_0 Z(f) \Pi_0 = 0\) unless \(Z(f)\) or \(Z_L Z(f)\) is a stabilizer of the surface code. Let \(A \subseteq \mathbb{F}_2^n\) be a linear subspace spanned by \(Z\)-stabilizers (considered as binary vectors) and let \(Z_L = Z(l)\) for some \(l \in \{0,1\}^n\). Then

\[
\Pi_0 C_s U \Pi_0 = a_s \Pi_0 + b_s \Pi_0 Z_L,
\]

where

\[
\alpha_s = \sum_{g \in A \oplus h_s} \alpha_g |g| \quad \text{and} \quad b_s = \sum_{g \in A \oplus h_s \oplus l} \alpha_g |g|.
\]

Here \(\oplus\) denotes addition of binary strings modulo two. Suppose first that \(h_s\) has even weight. Since any element of \(A\) has even weight, the sum that defines \(a_s\) runs over even-weight vectors \(g\), that is, \(a_s\) is real. Likewise, since \(l\) has odd weight, the sum that defines \(b_s\) runs over odd-weight vectors \(g\), that is, \(b_s\) is imaginary. Define \(q_s = |a_s|^2 + |b_s|^2\). Note that \(q_s\) does not depend on \(\psi_L\). One arrives at

\[
\Pi_0 C_s U \Pi_0 = \sqrt{q_s} \cdot \Pi_0 \exp [i \theta_s Z_L],
\]

where \(\theta_s \in [0,2\pi)\) is chosen such that \(a_s + b_s = \sqrt{q_s} e^{i \theta_s}\).

Since \(\exp (i \pi Z_L) = -I\) and overall phase factors do not matter, one can assume \(\theta_s \in [0,\pi]\). Substituting Eq. (A2) into Eq. (A1) gives \(p(s) = q_s\), that is, \(p(s)\) does not depend on \(\psi_L\), as claimed. Substituting \(\Pi_s = C_s^{\dagger} \Pi_0 C_s\) into Eq. (26), noting that \(C_s^{\dagger} = I\), and using Eq. (A2) with \(q_s = p(s)\) proves Eq. (27).

The case when \(h_s\) has odd weight is completely analogous, except that now \(b_s\) is real and \(a_s\) is imaginary. Choose \(\theta_s \in [0,2\pi)\) such that \(a_s + b_s = i \sqrt{q_s} e^{i \theta_s}\). Then \(\theta_s\) obeys Eq. (A2) with an extra factor of \(i\) on the right-hand side. The rest of the proof is exactly as above.

### Appendix B: Fermionic linear optics

In this appendix we state some necessary facts on simulation of fermionic linear optics. The material of this section is based on Refs. [37–39]. Let \(P_n\) be the group generated by single-qubit Pauli operators \(X, Y, Z\) with
For $2 \leq j \leq n$. They obey commutation rules
\[ c_p c_q = -c_q c_p \quad \text{for } p \neq q. \]  
(B2)

More generally, given a bit string $x \in \{0,1\}^{2n}$, define a Majorana monomial $c(x) = c_1^x c_2^x \cdots c_{2n}^x$. Then
\[ c(x)c(y) = (-1)^{|x \cdot y|}c(y)c(x) \]  
(B3)

whenever at least one of the strings $x,y$ has even weight.

Any $n$-qubit operator can be uniquely expressed as a linear combination of the $4^n$ Majorana monomials $c(x)$.

Suppose $\varrho$ is a (mixed) $n$-qubit state. The covariance matrix of $\varrho$ is a real anti-symmetric matrix $\varrho$ of size $2n \times 2n$ defined by
\[
\varrho_{p,q} = \begin{cases} 
\text{Tr}(i c_p c_q \varrho) & \text{if } p \neq q \\
0 & \text{if } p = q
\end{cases} 
(B4)
\]

A state $\varrho$ is called Gaussian if $\varrho$ is a linear combination of only even-weight Majorana monomials $c(x)$, and the expectation value of $c(x)$ on $\varrho$ can be computed from the covariance matrix $\varrho$ using Wick’s theorem. For example, we require that
\[
-\text{Tr}(c_p c_q c_r c_s \varrho) = \varrho_{p,q} \varrho_{r,s} - \varrho_{p,r} \varrho_{q,s} + \varrho_{p,s} \varrho_{q,r}
\]  
(B5)

for all $p \neq q \neq r \neq s$. More generally, we require that
\[
\text{Tr}(i^{x_1} c(x) \varrho) = \text{Pf}(\varrho_{[x]})
\]  
(B6)

for all even-weight $x \in \{0,1\}^{2n}$, where $\varrho_{[x]}$ is a submatrix of $\varrho$ including only rows and columns $p$ with $x_p = 1$ and $\text{Pf}$ denotes the Pfaffian. In the present paper we only use the special case Eq. (B5). To summarize, a Gaussian state can be fully specified by its covariance matrix.

It is well known that FLO gates defined in Section II preserve the class of Gaussian states. Thus a quantum circuit composed of FLO gates acting on some initial Gaussian state can be efficiently simulated if we know how to update the covariance matrix under the action of each gate. These update rules are stated below.

Consider a pair of modes $p,q \in \{1,2n\}$ and let
\[
\lambda = (1/2)(I + ic_p c_q).
\]  
(B7)

Note that $\lambda$ is a projector. It describes a post-selective parity measurement for the pair of modes $p,q$ with the measurement outcome $+1$ (note that the outcome $-1$ can be obtained simply by exchanging $p$ and $q$).

**Fact 1.** If $\varrho$ is a Gaussian state with covariance matrix $\varrho$ then $\varrho' = \varrho \lambda / \text{Tr}(\lambda \varrho)$ is a Gaussian state with a covariance matrix $\varrho'$ that can be computed in time $O(n^2)$ by the following algorithm:

```plaintext
function MEASURE(M,p,q)
    $\lambda \leftarrow (1/2)(1 + M_{p,q})$
    $\lambda \leftarrow (1/2)(1 + M_{p,q})$
    $\lambda \leftarrow (1/2)(1 + M_{p,q})$
    if $\lambda \neq 0$
        $K \leftarrow p$-th column of $M$
        $L \leftarrow q$-th column of $M$
        $M' \leftarrow M + (2\lambda)^{-1}(KLT - LKT)$
        Set to zero rows and columns $p,q$ of $M'$
        $M'_{p,q} \leftarrow 1$
        $M'_{q,p} \leftarrow 1$
        return ($\lambda, M'$)
    end if
end function
```

We note that the final matrix $M'$ has a block structure such that the modes $p,q$ are not coupled to any other modes. Thus one can remove rows and columns $p,q$ from $M'$ without losing any information.

Next let us discuss two-mode rotations
\[
U = \exp (\gamma c_p c_q).
\]  
(B8)

Here $\gamma \in [0,\pi]$ is the rotation angle. One can check that $U$ is a unitary operator such that the action of $U$ in the Heisenberg picture is
\[
U^1 c_p U = \cos (2\gamma) c_p + \sin (2\gamma) c_q,
\]
\[
U^1 c_q U = -\sin (2\gamma) c_p + \cos (2\gamma) c_q,
\]
\[
U^1 c_r U = c_r, \quad \text{if } r \notin \{p,q\}.
\]  
(B9)

We shall describe the transformation Eq. (B9) by an orthogonal matrix $R \in SO(2n)$ such that
\[
U^1 c_t U = \sum_{s=1}^{2n} R_{t,s} c_s, \quad 1 \leq t \leq 2n.
\]

**Fact 2.** If $\varrho$ is a Gaussian state with a covariance matrix $\varrho$ then $\varrho' = U^\dagger \varrho U$ is a Gaussian state with a covariance matrix $\varrho'$ that can be computed in time $O(n)$.

Finally, the class of Gaussian states is closed under the tensor product operation.

**Fact 3.** Let $\varrho_i$ be a Gaussian state of $n_i$ qubits with a covariance matrix $\varrho_i$. Here $i = 1,2$. Then $\varrho_1 \otimes \varrho_2$ is a Gaussian state of $n_1 + n_2$ qubits with a covariance matrix
\[
\varrho_1 \otimes \varrho_2 \equiv \begin{bmatrix} M_1 & 0 \\ 0 & M_2 \end{bmatrix}.
\]  
(B10)

All our algorithms work by simulating a sequence of two-mode parity measurements and two-mode rotations starting from a certain simple initial Gaussian state. To describe these initial states we need two more facts.

**Fact 4.** Let $\psi$ be a pure single-qubit state with a Bloch vector $\vec{b} = (b^x,b^y,b^z)$. Let $\vec{c}$ be the C4-encoding of $\psi$ defined in Eqs. (7,8). Then $\psi$ is Gaussian with a covariance
states can be defined by an abelian group of Pauli stabilizers. Namely, suppose \( \sigma : [2n] \to [2n] \) is a permutation.

**Fact 5.** There exists a unique Gaussian state \( \rho = |\phi\rangle\langle\phi| \) such that \( i\sigma_{(2j-1)}\sigma_{(2j)} |\phi\rangle = |\phi\rangle \) for all \( j = 1, \ldots, n \). The state \( \rho \) has a covariance matrix

\[
M_{r,s} = \sum_{j=1}^{n} \delta_{r,\sigma(2j-1)}\delta_{s,\sigma(2j)} - \delta_{r,\sigma(2j)}\delta_{s,\sigma(2j-1)}.
\]  

(B12)

The above facts are sufficient to simulate any quantum circuit composed of FLO gates, as defined in Section II, and provide all details necessary for implementation of our algorithms \( A \) and \( B \).

---

[1] R. Barends, J. Kelly, A. Megrant, A. Veitia, D. Sank, E. Jeffrey, T. White, J. Mutus, A. Fowler, B. Campbell, et al., Nature 508, 500 (2014).
[2] J. Kelly, R. Barends, A. Fowler, A. Megrant, E. Jeffrey, T. White, D. Sank, J. Mutus, B. Campbell, Y. Chen, et al., Nature 519, 66 (2015).
[3] A. Corcoles, E. Magesan, S. Srinivasan, A. Cross, M. Steffen, J. Gambetta, and J. Chow, Nature Communications 6, 6979 (2015).
[4] N. Ofek, A. Petrenko, R. Heeres, P. Reinhold, Z. Leghtas, B. Vlastakis, Y. Liu, L. Frunzio, S. Girvin, L. Jiang, et al., arXiv preprint arXiv:1602.04768 (2016).
[5] M. Takita, A. D. Corcoles, E. Magesan, B. Abdo, M. Brink, A. Cross, J. M. Chow, and J. M. Gambetta, Phys. Rev. Lett. 117, 210505 (2016).
[6] A. Kitaev, Annals of Physics 303, 2 (2003).
[7] S. Bravyi and A. Y. Kitaev, arXiv preprint quant-ph/9811052 (1998).
[8] E. Dennis, A. Kitaev, A. Landahl, and J. Preskill, Journal of Mathematical Physics 43, 4452 (2002).
[9] R. Raussendorf and J. Harrington, Phys. Rev. Lett. 98, 190504 (2007).
[10] A. G. Fowler, A. M. Stephens, and P. Groszkowski, Phys. Rev. A 80, 052312 (2009).
[11] D. Gottesman, PhD thesis, California Institute of Technology, arXiv preprint arXiv:9705052 (1997).
[12] A. G. Fowler, Phys. Rev. Lett. 109, 180502 (2012).
[13] D. S. Wang, A. G. Fowler, and L. C. L. Hollenberg, Phys. Rev. A 83, 020302 (2011).
[14] M. Gutiérrez, L. Svec, A. Vargo, and K. R. Brown, Phys. Rev. A 87, 030302 (2013).
[15] Y. R. Sanders, J. J. Wallman, and B. C. Sanders, New Journal of Physics 18, 012002 (2016).
[16] R. Kueng, D. M. Long, A. C. Doherty, and S. T. Flammia, Phys. Rev. Lett. 117, 170502 (2016).
[17] J. Wallman, C. Granade, R. Harper, and S. T. Flammia, New Journal of Physics 17, 113020 (2015).
[18] D. Puzzuoli, C. Granade, H. Haas, B. Criger, E. Magesan, and D. G. Cory, Phys. Rev. A 89, 022306 (2014).
[19] E. Magesan, D. Puzzuoli, C. E. Granade, and D. G. Cory, Phys. Rev. A 87, 012324 (2013).
[20] B. Rahn, A. C. Doherty, and H. Mabuchi, Phys. Rev. A 66, 032304 (2002).
[21] J. Fern, J. Kempe, S. N. Simic, and S. Sastry, IEEE Transactions on Automatic Control 51, 418 (2006).
[22] D. Greenbaum and Z. Dutton, arXiv preprint arXiv:1612.03908 (2016).
[23] M. Li, M. Gutierrez, S. E. David, A. Hernandez, and K. R. Brown, arXiv preprint arXiv:1702.01155 (2017).
[24] Y. Tomita and K. M. Svore, Phys. Rev. A 90, 062320 (2014).
[25] J. P. Barnes, C. J. Trout, D. Lucarelli, and B. D. Clader, Phys. Rev. A 95, 062338 (2017).
[26] C. Chamberland, J. Wallman, S. Beale, and R. Laflamme, Phys. Rev. A 95, 042332 (2017).
[27] A. S. Darmawan and D. Poulin, Phys. Rev. Lett. 119, 040502 (2017).
[28] Y. Suzuki, K. Fujii, and M. Koashi, arXiv preprint arXiv:1703.03671 (2017).
[29] X.-G. Wen, Phys. Rev. Lett. 90, 016803 (2003).
[30] H. Bombin and M. A. Martin-Delgado, Phys. Rev. A 76, 012305 (2007).
[31] A. G. Fowler, A. C. Whiteside, and L. C. L. Hollenberg, Phys. Rev. Lett. 108, 180501 (2012).
[32] J. Emerson, M. Silva, O. Moussa, C. Ryan, M. Laforest, J. Baugh, D. G. Cory, and R. Laflamme, Science 317, 1893 (2007).
[33] M. Silva, E. Magesan, D. W. Kribs, and J. Emerson, Phys. Rev. A 78, 012347 (2008).
[34] A. Kitaev, Annals of Physics 321, 2 (2006), january Special Issue.
[35] B. M. Terhal, F. Hassler, and D. P. DiVincenzo, Phys. Rev. Lett. 108, 260504 (2012).
[36] E. Knill, arXiv preprint quant-ph/0108033 (2001).
[37] S. Bravyi, Quant. Inf. Comp. 5, 216 (2004).
[38] B. M. Terhal and D. P. DiVincenzo, Phys. Rev. A 65, 032325 (2002).
[39] S. Bravyi and R. Koenig, Quant. Inf. Comp. 12, 925 (2012).
[40] A. Y. Kitaev, Russian Mathematical Surveys 52, 1191 (1997).
[41] J. Lodyga, P. Mazurek, A. Grudka, and M. Horodecki, Scientific reports 5, 8975 (2015).
[42] J. R. Wootton and J. K. Pachos, Phys. Rev. Lett. 107, 030503 (2011).
[43] C. Stark, L. Pollet, A. Imamoglu, and R. Renner, Phys. Rev. Lett. 107, 030504 (2011).
[44] S. Bravyi and R. König, Communications in Mathematical Physics.
ical Physics 316, 641 (2012).