Introduction To The Toric Code

Hongyue Li

Abstract

In this paper, we introduce the notion of quantum error correction and analyze the error model of the toric code via a correspondence with the random-bond Ising model and random-plaquette gauge model. Finally, we use computer simulations to estimate the accuracy threshold of toric code under this error model.

1 Quantum Error Correction

When we send a signal in a noisy channel, it may be damaged. To make sure the signal is sent, we can try to use some code that has some redundancy and is able to correct errors. For example, we want to send bits 0 or 1 through a classical channel, but this channel has a probability p of flipping the bit. Thus our information may be destroyed with this probability p. To reduce the probability of not sending information successfully, we need an error correction protocol that can encode and decode the message to make it more robust to errors. For example, we can use the repetition code.

Example 1 (Repitition Code)**.** *We duplicate the message by making k copies 00...0/11...1 for each bit 0/1. For large enough k the probability that more than half of 0 is flipped is very low (it decays exponentially with k). Thus we can decode the original message 0, 1 by majority vote, with an exponentially small error.*

The idea of quantum error correction is similar. A quantum channel is generally noisy because of the coupling with the environment. Denote the quantum channel describing such errors as E. Quantum information is generally not preserved by a quantum channel for a generic input. Error correction is possible if we restrict the input. If we define an encoding map C and a decoding map D, the condition of error correction is

$$D \circ E \circ C = I$$

The image of encoding map C defines a subspace of the Hilbert space, called the code subspace.

Definition 1. *The Pauli matrices are* $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. *Let* P^n *denote the n-fold tensor product of the Pauli operators,*

$$P_n =< X_i, Z_j, 1 \leq i, j \leq n > .$$

Example 2 (Bit-Flip Code and Phase-Flip Code)**.** *The repetition code works in a classical channel, because classical bits are easy to measure and to repeat. This approach does not work for a quantum channel in which, due to the no-cloning theorem, it is not possible to repeat a single qubit three times.*

To prevent a single bit-flip error X, we use the encoding: $|0\rangle \to |000\rangle, |1\rangle \to |111\rangle$.

To prevent a single phase-flip error Z, we use the encoding: $|0\rangle \to |+ ++\rangle, |1\rangle \to |- --\rangle$*, here* $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), |-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$.

1.1 Stabilizer Code

A stabilizer code is a quantum code defined by a stabilizer group S, which is an Abelian subgroup of the Pauli group on n qubits. We require that $-I \notin S$. The codespace C is a subspace of the Hilbert space $(\mathbb{C}^2)^{\otimes n}$ spanned by the +1 eigenvectors of the stabilizers in S. The number of logical qubits is

$$k := \log \dim C = n - \#(\text{independent generators of S})$$

Example 3 (The Shor Code)**.** *The Shor code can protect against an arbitrary single-qubit bit-flip error or phase-flip error. The encoding is:*
$|0\rangle_L = \frac{1}{2^{3/2}} (|000\rangle + |111\rangle) (|000\rangle + |111\rangle) (|000\rangle + |111\rangle)$,
$|1\rangle_L = \frac{1}{2^{3/2}} (|000\rangle - |111\rangle) (|000\rangle - |111\rangle) (|000\rangle - |111\rangle)$
The stabilizer group S is generated by

$$< Z_1Z_2, Z_2Z_3, Z_4Z_5, Z_5Z_6, Z_7Z_8, Z_8Z_9, X_1X_2X_3X_4X_5X_6, X_4X_5X_6X_7X_8X_9 >$$

and consists of 2^8 *terms in total. With 9 physical qubits and 8 generators, the Shor code encodes 1 logical qubit.*

Suppose that the errors affecting an encoded quantum state are a subset $\mathcal{E} \in P^n$. Because $\mathcal{E}$ and $\mathcal{S}$ are both subsets of P^n, an error $E \in \mathcal{E}$ that affects an encoded quantum state either commutes or anticommutes with any particular element g in $\mathcal{S}$. The error E is correctable if it anticommutes with an element g in $\mathcal{S}$. An anticommuting error E is detectable by measuring each element g in $\mathcal{S}$ and computing a syndrome $\mathbf{r}$ identifying E. The syndrome is a binary vector $\mathbf{r}$ with length $n-k$ whose elements identify whether the error E commutes or anticommutes with each $g \in \mathcal{S}$. An error E that commutes with every element g in $\mathcal{S}$ is correctable if and only if it is in $\mathcal{S}$. It corrupts the encoded state if it commutes with every element of $\mathcal{S}$ but does not lie in $\mathcal{S}$. So we compactly summarize the stabilizer error-correcting conditions: a stabilizer code can correct any errors E_1, E_2 in $\mathcal{E}$ if

$$E_1^\dagger E_2 \notin \mathcal{Z}(\mathcal{S})$$

or

$$E_1^\dagger E_2 \in \mathcal{S},$$

where $\mathcal{Z}(\mathcal{S})$ is the centralizer of $\mathcal{S}$ (i.e., the subgroup of elements that commute with all members of $\mathcal{S}$).

For every stabilizer code with the stabilizer group $\mathcal{S}$, we can define a stabilizer Hamiltonian

$$H = -\sum_{S \in \mathcal{S}} S,$$

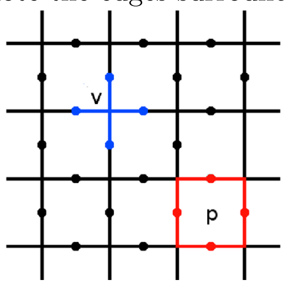
where the sum is over the set of stabilizer generators. The ground states of the Hamiltonian H correspond to the codespace C.

2 Toric Code

The toric code (or more generally, the surface code) is defined on a 2D surface with vertices V, edges E, and faces F, with a spin-$\frac{1}{2}$ degree of freedom located on each edge. Stabilizer operators are defined on the spins around each vertex v and plaquette (or face) p of the lattice as follows,

$$X(v) = \prod_{i \in v} X_i, \quad Z(p) = \prod_{i \in p} Z_i.$$

Here we use $i \in v$ to denote the edges touching the vertex v, and $i \in p$ to denote the edges surrounding the plaquette/face p.



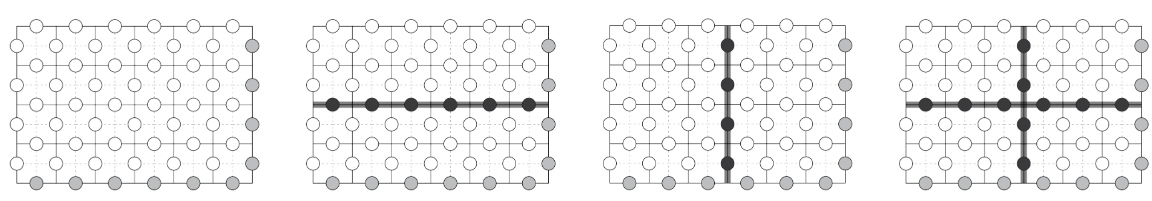
The 2D toric code is a topological code, i.e. it can be defined on a lattice and its stabilizer generators are geometrically local. For the toric code we have $|E|$ physical qubits, $|V| + |F|$ stabilizer generators and two relations among them $\prod_{v \in V} X(v) = I, \prod_{f \in F} Z(f) = I$. The Euler characteristic $\chi = |V| - |E| + |F| = 2 - 2g$, so the number of logical qubits encoded in the toric code is 2g, where g denotes the genus of the manifold on which we define the codes.

The Hamiltonian of the toric code is

$$H = -\sum_{v \in V} X(v) - \sum_{f \in F} Z(f)$$

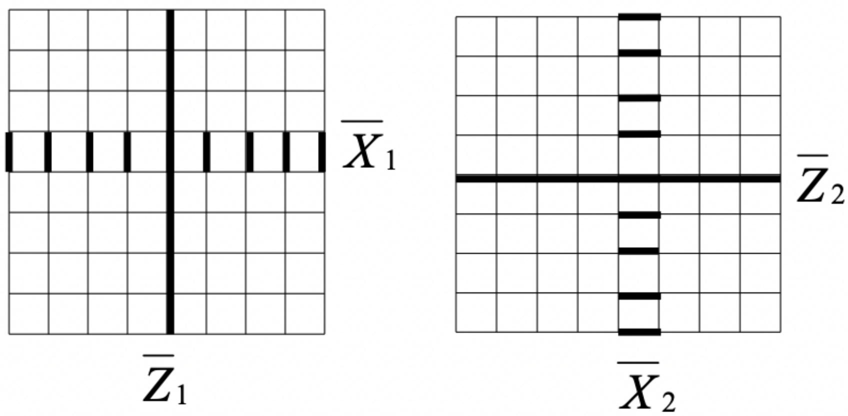
The ground state space of this Hamiltonian is the stabilizer space of the code.

Now we restrict our attention to the toric code defined on a torus $T^2 = S^1 \times S^1$ lattice of side length L only, so there are $2L^2$ physical qubits which encode 2 logical qubits. The four orthogonal ground states $|00\rangle, |10\rangle, |01\rangle, |11\rangle$ correspond to a, b, c, d in the figure below.



(a) (b) (c) (d)

Then a string of X operators looped around the horizontal dimension maps $|00\rangle \leftrightarrow |10\rangle$, and a loop of Z operators in the vertical direction maps $|00\rangle \leftrightarrow |00\rangle$ and $|10\rangle \leftrightarrow -|10\rangle$, therefore we can identify these strings as logical Pauli operators $\bar{X}_1$ and $\bar{Z}_1$ acting on the first logical qubit. Similarly, the remaining loops of X and Z operators act as logical operators $\bar{X}_2, \bar{Z}_2$ on the second qubit. (Note that the action of these logical operators does not depend on a specific path as long as that path winds around the torus.)



3 Error Model of Toric Code

The unique nature of the topological codes, such as the toric code, is that stabilizer violations can be interpreted as quasiparticles. Specifically, if the code is in a state $|\phi\rangle$ such that,

$$X(v)|\phi\rangle = -|\phi\rangle,$$

a quasiparticle known as an e-anyon can be said to exist on the vertex v. Similarly, violations of the $Z(f)$ are associated with so-called m-anyons on the plaquettes. The stabilizer space therefore corresponds to the anyonic vacuum. Single spin errors cause pairs of anyons to be created and transported around the lattice.

When errors create an anyon pair and move the anyons, one can imagine a path connecting the two composed of all links acted upon. If the anyons then meet and are annihilated, this path describes a loop. If the loop is topologically trivial, it has no effect on the stored information. The annihilation of the anyons, in this case, corrects all of the errors involved in their creation and transport. But a cycle could be homologically nontrivial. A product of Z's corresponding to a nontrivial cycle commutes with the code stabilizer (because it is a cycle), but is not contained in the stabilizer (because the cycle is nontrivial). Therefore, while this operator preserves the code subspace, it acts nontrivially on encoded quantum information. The errors, in this case, are therefore not corrected but consolidated.

Consider the noise model for which bit errors X and phase errors Z occur independently on each spin, both with probability p. When p is low, this will create sparsely distributed pairs of anyons that have not moved far from their point of creation. Correction can be achieved by identifying the pairs that the anyons were created in (up to an equivalence class), and then re-annihilating them to remove the errors. As p increases, however, it becomes more ambiguous as to how the anyons may be paired without risking the formation of topologically non-trivial loops. This gives a threshold probability p_c, under which the error correction will almost certainly succeed.

3.1 The Statistical Physics of Error Model

In this section, we will show that assuming no measurement error, the toric code error model corresponds to the random-bond Ising model, and with measurement error, it corresponds to the random-plaquette gauge model.

We will formulate an order parameter that distinguishes two phases of a quantum memory: an "ordered" phase in which reliable storage is possible, and a "disordered phase" in which errors unavoidably afflict the encoded quantum information.

3.1.1 Random-Bond Ising Model

The two-dimensional random-bond Ising model (RBIM) is an Ising spin system on a torus lattice, with a variable $s_i = \pm 1$ residing at each lattice site i and a quenched random variable $\tau_{ij} = \pm 1$ at each edge ij, $\text{Prob}(\tau_{ij} = -1) = p$, where p is referred to as the concentration of antiferromagnetic bonds. Its Hamiltonian is

$$H = -J \sum_{\langle ij \rangle} \tau_{ij} s_i s_j.$$

Its partition function

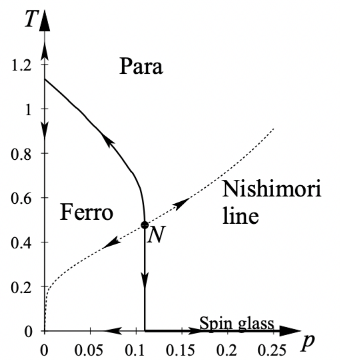
$$Z(K, p) = \sum_{\{s_i\}} \exp(K \sum_{\langle ij \rangle} \tau_{ij} s_i s_j),$$

here $K = \beta J = \frac{1}{k_B T} J$.

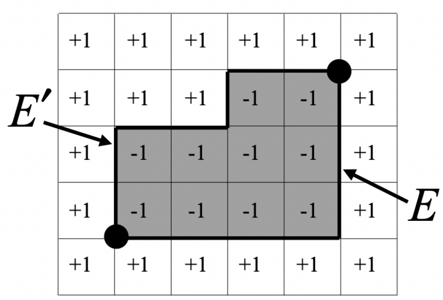
The 'Nishimori line' is

$$e^{-2K} = \frac{p}{1-p}.$$

We fix K to be on this line.



For the error model of the toric code, we assume that there is no measurement error and qubit errors arise independently on each edge with probability p. Denote the set of edges where an error occurs by E. We observe the syndrome $S = \partial E$, and say our correction is E'. Now if the symmetric difference $E' \triangle E$ is trivial, our error recovery is successful. Since E and E' have the same boundary S, we can write $E' = E \triangle C$, where C is a union of cycles, $\partial C = \varnothing$. Say C belongs to the homology class h. Since we only observe $S = \partial E$, we compute the probability $prob(h|S)$ for all four homology classes and pick the one with the highest conditional probability as our recovered C' and let $E' \subset C'$ be segments of the cycles that are separated by S.



Now the correspondence between the error model of the toric code and the random-bond Ising model is clear. We view the error $E = \{(i,j) : \tau_{ij} = -1\}$, our recovered error $E' = \{(i,j) : \tau_{ij} s_i s_j = -1\}$, and the cycle $C = E \triangle E' = \{(i,j) : s_i s_j = -1\}$, note that we always have $\partial C = \varnothing$.

The model has enhanced symmetry properties along the Nishimori line. The antiferromagnetic bond chain E and the excited bond chain E' are generated by sampling the same probability distribution, subject to the constraint that both chains have the same boundary points.

$Prob(E) = Prob(\tau) = \prod_l (1-p) \prod_{\tau_l=-1} \frac{p}{1-p} \propto \prod_{l\in E} \frac{p}{1-p}$
$Prob(E') \propto \prod_{\tau_{ij}s_is_j=-1} e^{-2K} \propto \prod_{l\in E'} \frac{p}{1-p}$

So the random-bond Ising model along the Nishimori line corresponds exactly to the error model of the toric code.

3.1.2 Random Plaquette Gauge Model

The three-dimensional random-plaquette gauge model (RPGM) is defined on a 3D lattice $T^2 \times I = S^1 \times S^1 \times I$, here $I = [0, 1]$, with a variable $u_l = \pm 1$ residing at each edge and a quenched random variable $\tau_p = \pm 1$ at each plaquette, $\mathrm{Prob}(\tau_p = -1) = p$. Its Hamiltonian is

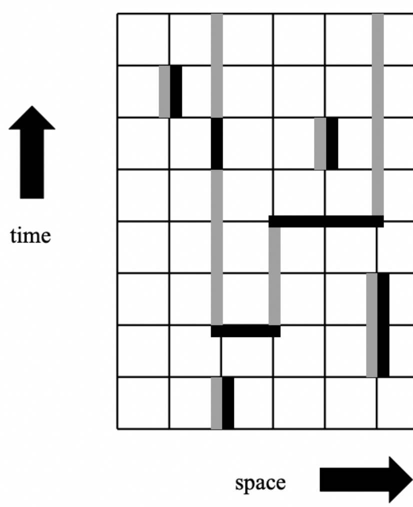
$$H = -J \sum_p \tau_p U_p,$$

here $U_p = \prod_{l\in p} u_l$.

Its partition function

$$Z(K, p) = \sum_{\{u\}} \exp(K \sum_p \tau_p U_p).$$

For the error model of the toric code, we now assume that the measurement error at each site happens independently with probability q=p and qubit errors arise independently on each edge with probability p.



In this figure, an error history is shown together with the syndrome history that it generates, for the toric code. For clarity, the three-dimensional history of the two-dimensional code block has been compressed into two dimensions. Qubits reside on plaquettes, and four-qubit check operators are measured at each vertical link. Links where errors have occurred are darkly shaded, and links, where the syndrome is nontrivial, are lightly shaded. Errors on horizontal links indicate where a qubit flipped between successive syndrome measurements, and errors on vertical links indicate where the syndrome measurement was wrong. Vertical links that are shaded both lightly and darkly are locations where a nontrivial syndrome was found erroneously. The chain S of lightly shaded links (the syndrome) and the chain E of darkly shaded links (the errors) both have the same boundary.

Now we describe the correspondence between the error model of the toric code with measurement errors and the dual lattice of the random-plaquette gauge model along the Nishimori line. We let $E = \{p : \tau_p = -1\}$ be the set of wrong sign plaquettes that corresponds to the dual of the error chain that contains both qubit and measurement errors, $E' = \{p : \tau_p U_p = -1\}$ be the recovered chain. Error recovery is successful if $C = E \triangle E' = \{p : U_p = -1\}$ is trivial (Note that C is always a union of closed surfaces since $U_p = \prod_l u_l$).

3.2 Error Recovery

In theory, given the observed syndrome, we can compute the free energy/probability of each homology class, and choose the one with minimum free energy/maximum probability. In the ordered phase, as the volume goes to infinite, the chain of minimum energy will be in the homology class with minimum free energy almost surely[1]. Therefore in practice, we can simply decode the syndrome by minimizing the energy/maximizing the probability of a single chain. This can be achieved by calculating a minimum length chain the boundary of which is the vertices in the syndrome, which can be reduced to a minimum-weight perfect matching problem. We recover the code by applying the chain to our observed lattice and calculate the logical qubits.

We present an algorithm that solves this problem. In practice there is a python package called 'pymatching' that has an efficient implementation of this procedure.

Algorithm 1: Blossom algorithm

Data: Input data: Syndrome $S = \{v_i\}$ on the lattice

Result: Output result: Paths P whose boundary is the syndrome, $\partial P = S$

1 Calculate the pairwise distances of all vertices in the syndrome S: $w_{ij} = \mathrm{dist}(v_i, v_j)$. Let E be the set of all pairs (i, j).

2 Formulate an integer programming problem:

$$\begin{aligned} \text{Minimize} \quad & \sum_{e \in E} w_e x_e \\ \text{subject to} \quad & \sum_{v \in e} x_e = 1 \quad \forall v \in S \\ & x_e \in \{0, 1\} \quad \forall e \in E. \end{aligned}$$

3 This can be relaxed to a linear program:

$$\begin{aligned} \text{Minimize} \quad & \sum_{e \in E} w_e x_e \\ \text{subject to} \quad & \sum_{v \in e} x_e = 1 \quad \forall v \in S \\ & \sum_{e \in \delta(o)} x_e \geq 1 \quad \forall o \in O \\ & x_e \geq 0 \quad \forall e \in E. \end{aligned}$$

Here we denote the boundary edges of some subset of the nodes $U \subseteq V$ by $\delta(U) := \{(u, v) \in E \mid u \in U, v \in V \setminus U\}$, and let O be the set of all subsets of V of odd cardinality at least three, i.e., $O := \{o \subseteq V : |o| > 1, |o| \bmod 2 = 1\}$.

4 We consider the dual to this linear program and starts with an empty matching and a feasible dual solution, and iteratively increases the cardinality of the matching and the value of the dual objective while ensuring the dual problem constraints remain satisfied. Eventually, we will have a pair of feasible solutions to the primal and dual problem satisfying the complementary slackness conditions at which point we know we have a perfect matching M of minimal weight.

5 **return** *The set of paths P that corresponds to the matching M.*

3.3 Theoretical Accuracy Threshold

A toric code is a Calderbank-Shor-Steane (CSS) code, meaning that each stabilizer generator is either a tensor product of X's or a tensor product of Z's. If X errors and Z errors each occur with probability p, then it is known that CSS codes exist with asymptotic rate $R = k/n$ (where n is the block size and k is the number of encoded qubits) such that error recovery will succeed with probability arbitrarily close to one, where $R = 1 - 2H_2(p)$ here $H_2(p) := -p \log_2 p - (1-p)\log_2(1-p)$ is the binary Shannon entropy. This rate hits zero when p has the value $p_c \simeq 0.110028$.

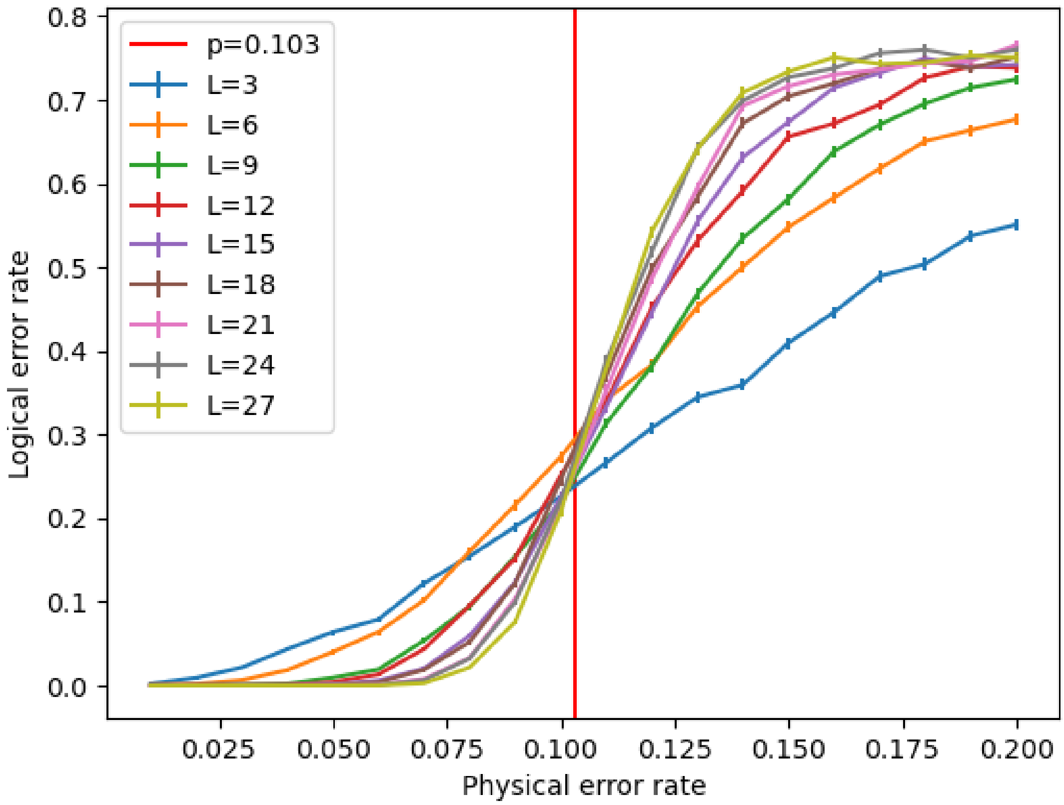
4 Experiment

4.1 Experiment 1 (No Measurement Error)

We first estimate the threshold of the toric code under an independent noise model with perfect syndrome measurements. Since we assume the X and Z errors are independent, we only consider decoding Z-type errors with X-type stabilizers. This corresponds to the value of the critical point of the 2D random-bond Ising model on the Noshimori line.

We first generate errors on each site independently with probability p on each edge, then apply the X-stabilizers to obtain the syndrome, and finally use the minimum weight matching algorithm to reconstruct the error chain. We compare the recovered logical qubits with the true logical qubits and declare a success/failure if they are equal/not equal.

We perform a Monte Carlo simulation with 5000 iterations, for each 2D lattice of side length $L \in \{3, 6, 9, ..., 27\}$ and for error probability from $p = 0.01$ to $p = 0.2$.

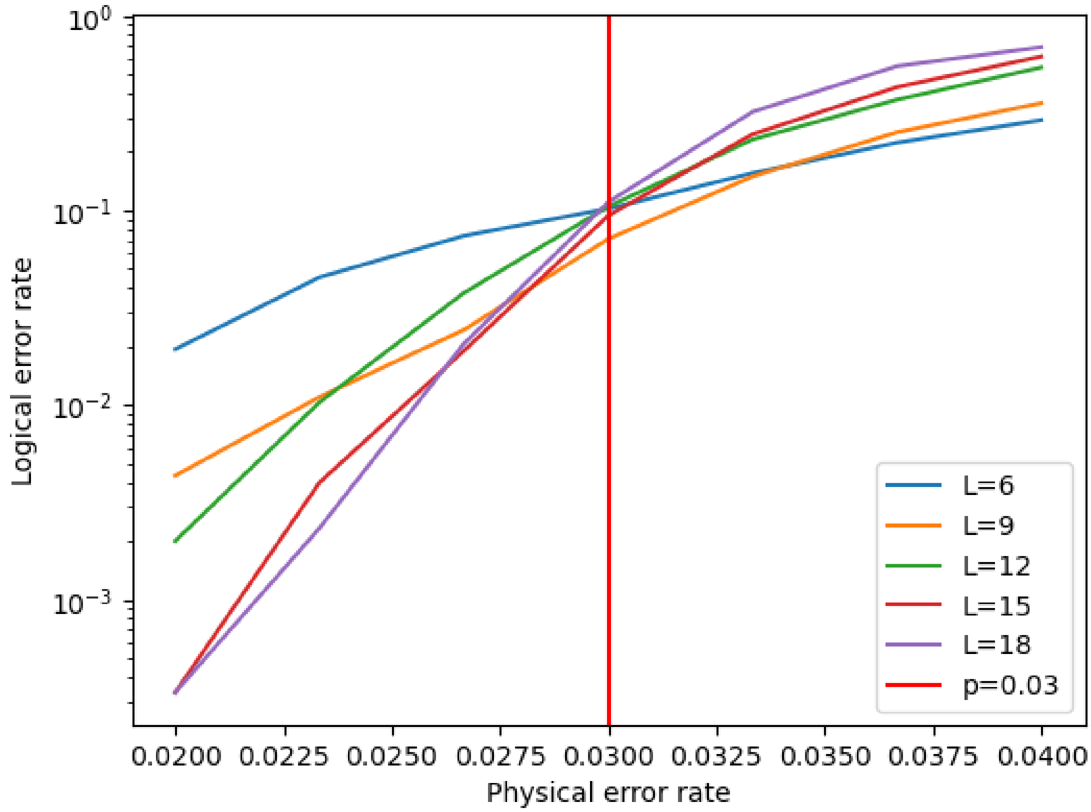


Based on the figure, we observe a phase transition around p=0.11, which agrees with [1].

4.2 Experiment 2 (With Measurement Error)

Now we assume the measurement error is also p. This corresponds to the value of the critical point of the 3D random-plaquette gauge model on the Noshimori line. We generate errors independently with probability p on each edge of the 3D lattice, and all other procedures are the same as those in the previous experiment.

We perform a Monte Carlo simulation with 3000 iterations, for each 3D lattice of side length $L \in \{6, 9, 12, 15, 18\}$ and for error probability from $p = 0.02$ to $p = 0.04$.



Based on the figure, we observe a phase transition around p=0.03, which agrees with [2].

The discrepancy between the results obtained by the two experiments is expected[1].

5 Conclusion

In this paper, we first introduced the notion of quantum error correction. Then we studied the toric code, described a correspondence between the error model of the toric code with the random-bond Ising model and the random-plaquette gauge model, and used computer simulation to estimate the accuracy threshold of toric code.

6 Appendix

Our code is available here on github:
`https://github.com/ordinarylhy/Physics212Project`.

References

[1] Eric Dennis, Alexei Kitaev, Andrew Landahl, and John Preskill. Topological quantum memory. *Journal of Mathematical Physics*, 43(9):4452–4505, aug 2002.

[2] Chenyang Wang, Jim Harrington, and John Preskill. Confinement-higgs transition in a disordered gauge theory and the accuracy threshold for quantum memory. *Annals of Physics*, 303(1):31–58, January 2003.