Developments in experimental quantum error correction

Eric Huang, Thomas Steckmann, Maryam Muassar, Suying Liu

April 2024

Abstract

Experimental demonstration of fault-tolerance is a key ingredient to useful quantum computation. Since the fundamentals of the theory of quantum fault-tolerance have been posed, many experimental efforts have led to significant milestones in the long path towards fault-tolerance. In this review, we revisit a simple criterion for how experimentalists may claim to have demonstrated fault-tolerance, reviewing the most basic post-selected five-qubit protocol before summarizing the advances in scaling up surface codes in superconducting qubits by Google, outlining the logical Bell pair and GHZ state preparation in Rydberg neutral atom arrays by the Harvard collaboration and finally discussing the fault-tolerant Bell state sampling in the trapped ion system by Quantinuum.

Contents

1 Introduction

In order to build a useful quantum computer, we need a way to mitigate and correct any errors that occur either due to faults occurring during the computation, or faults that occur due to the system's interaction with the environment. The threshold theorem comes to our rescue by stating that if we can get the error rate in our devices below the threshold p_T, it is possible to perform a fault tolerant simulation of a circuit such that the output distribution is within reasonable statistical distance of the output from an ideal circuit. This has triggered a race of sorts between different platforms to reach the threshold as soon as possible, with neutral atom arrays, ion traps and superconducting qubits being the notable contenders. Each platform is at a different stage, from the Google experiment demonstrating a logical memory, to Quantinuum beginning to do early stage fault tolerance experiments. To our knowledge, this is a summary of the most exciting developments as of late April 2024, but it is likely that things will change in the coming months.

2 Experimental fault-tolerance criterion

The goal of the field is to build a fault-tolerant quantum computer. Experiments attempting to demonstrate fault-tolerance may be performed over a wide variety of physical systems, which necessitates a uniform way of comparing their results and performance.

Gottesman 2016 [1] poses the following important question: When can an experimentalist claim to have demonstrated fault-tolerance? The answer proposed was that this may be claimed when the experimentalist shows that for all circuits in a family of circuits, the error rate of the output of an encoded circuit is lower than that for any attempt to run it unencoded on the hardware.

Specifically, it was proposed to restrict our interest to *complete quantum circuits*, that is, quantum circuits which start from a small well-defined set of product states such as $|0\rangle^{\otimes n}$, apply gates from a gate set, and perform measurements to give a classical output. In this sense, the input being the specification of the circuit C to run is classical, and the output being a bit string $j \in \mathbb{Z}_2^n$ is also classical. Even in the ideal case of no faults and errors, the output is in general stochastic and must be characterised by a probability distribution $\{p_j\}$. For the small circuits likely to be tested in the early fault-tolerance experiments, this distribution may be exactly computed.

If directly implementing the circuit in hardware using without a fault-tolerant protocol results in an output distribution $\{q_j\}$, then we may quantify how far this distribution strays from the ideal distribution by a *total variation distance*

$$P_u(C) = \frac{1}{2} \sum_j |q_j - p_j| \,. \tag{1}$$

This statistical distance for ranges from the perfect case of 0 to the extreme of 1 being the worst-case.

As the same circuit may be implemented using different subsets of qubits and gates on a device, of which some are of a superior quality over others, we are interested in P_u reflecting the *best* possible implementation using the best qubits and the best gates for the circuit.

A fault-tolerant protocol is designed to protect against errors and faults using a toolkit of fault-tolerant *gadgets*. This may involve running an encoded version of the ideal circuit on potentially faulty hardware holding logical data qubits in a large quantum error correcting code, with the corresponding logical state initialized by state preparation gadgets, gates performed by logical gate gadgets, and outputs obtained by measurement gadgets. But more generally, the circuit is encoded into a much larger circuit that gets run to produce more output bits than before. These outputs may be interpreted with a *decoding scheme* to recover the corresponding output of the same form as the unencoded circuit output. If there are no faults then these decoded bits should conform to the same statistical distribution as that of the ideal circuit. The quality of the output from this encoded circuit is similarly quantified by a statistical distance $P_e(C)$ that characterizes the deviation of this output distribution from that of the ideal case. This is illustrated in Figure 1.

The criterion by which an experimentalist may claim that an experiment has demonstrated fault-tolerance is that

$$P_e(C) < P_u(C) \tag{2}$$

for all possible circuits C. That is, an experimental protocol is fault-tolerant if for all possible circuits, the encoded circuit produces a classical output distribution that is closer to the ideal distribution than that obtained from running the circuit unencoded in the best way possible. Naturally, it is only fair that the experiments for both cases were performed on the same hardware running under comparable conditions.

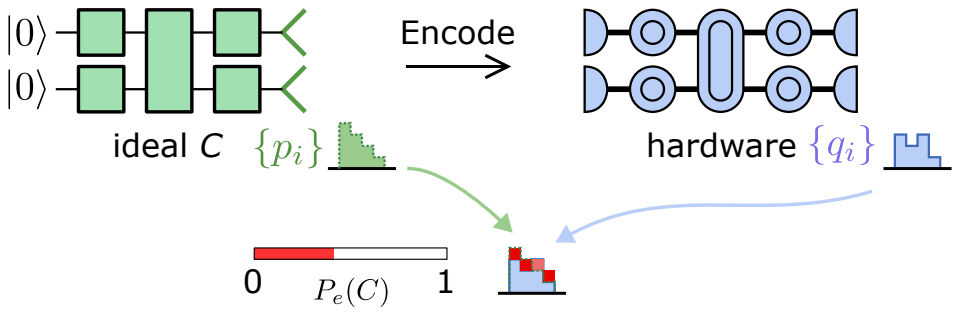

Figure 1: The total variation distance $P_e(C)$ between the output distribution of the ideal circuit $\{p_i\}$ and the output distribution of the encoded hardware circuit after interpretation $\{q_i\}$ quantifies the fault-tolerance of the implementation.

2.1 Concessions

The above criterion as stated may be found rather strict in practice and difficult to achieve for several reasons. As such, some concessions are made that allow for early experimentalists to make reasonable claims of fault-tolerance for their systems.

Firstly, as the number of circuits grows exponentially in the number of qubits and the depth of the circuit, it is not realistic to require the testing of all possible circuits. Thus a smaller family of circuits must be chosen, but the family should be sufficiently diverse to capture all the edge cases, since some circuits may harbour pathological features that cause errors to accumulate and conspire coherently at the detriment of the outputs. Although it was conceded that the choice of the family of circuits to test is a subjective matter, it is desirable that a *universal family* be tested, which is one that can be composed to implement any arbitrary quantum algorithm.

Secondly, it is appreciated that early systems may not have the full capabilities required to perform full error correction cycles, such as mid-circuit measurements with qubit resets, and applying gates conditional on those intermediate outcomes. Thus, for such early demonstrations of fault-tolerance, post-selection based on measurement outcomes is permitted. The justification was that post-selection is an integral step in fault-tolerant state preparation protocols for initialization, where it is sometimes referred to as pre-selection in

such a context. This is acceptable provided the non-rejected runs are still numerous enough to yield sufficient statistics, which would be the case for early small-scale experiments.

Lastly, since we are collecting statistics of random outputs, it is important that the appropriate statistical tests are done to distinguish between the distributions and confidently claim that the encoded circuit does indeed produce a more ideal distribution than the unencoded one.

2.2 Minimalist demonstration

In light of such a simple criterion, the same paper also proposed a minimalist experiment to demonstrate fault-tolerance using only 5 physical qubits. In particular, it uses the four-qubit $[[4, 2, 2]]$ quantum error detecting code and one ancilla qubit.

The four-qubit error-detecting code is a stabilizer code with the following two stabilizer generators and two pairs of logical operators

$$\begin{array}{rcccc} M_1 & = & X & X & X & X \\ M_2 & = & Z & Z & Z & Z \\ \hline \overline{X}_1 & = & X & I & X & I \\ \overline{Z}_1 & = & Z & Z & I & I \\ \overline{X}_2 & = & X & X & I & I \\ \overline{Z}_2 & = & Z & I & Z & I \end{array} \tag{3}$$

The code also supports two additional transversal Clifford gates

$$\overline{\mathrm{SWAP}}\,\overline{H}_1\overline{H}_2 = H \otimes H \otimes H \otimes H \tag{4}$$

$$\overline{CZ}\,\overline{Z}_1\overline{Z}_2 = S \otimes S \otimes S \otimes S, \tag{5}$$

which allows for more interesting 2-qubit logical circuits.

Its logical $|\overline{00}\rangle$ state is a GHZ state and has the useful property that all codewords are superpositions of only even-weight states

$$|\overline{00}\rangle = |0000\rangle + |1111\rangle \tag{6}$$

$$|\overline{01}\rangle = |1100\rangle + |0011\rangle \tag{7}$$

$$|\overline{10}\rangle = |1010\rangle + |0101\rangle \tag{8}$$

$$|\overline{11}\rangle = |0110\rangle + |1001\rangle\,. \tag{9}$$

This allows for the logical state to be measured fault-tolerantly in the Z basis using post-selection, simply by measuring all 4 physical qubits in the Z basis to obtain a classical bit string, and matching it with the appropriate code word above, while discarding any outcome of odd weight. This measurement scheme thus tolerates up to 1 fault as promised.

Meanwhile, there are also gadgets for the logical state preparation of $|\overline{00}\rangle$, $|\overline{0+}\rangle$ and $|\overline{00}\rangle + |\overline{11}\rangle$ that can tolerate up to 1 fault. $|\overline{00}\rangle$ and $|\overline{0+}\rangle$ may be prepared from a the $|0\rangle^{\otimes n}$ product state without trouble, for instance

$$|\overline{0+}\rangle = |\overline{00}\rangle + |\overline{01}\rangle \tag{10}$$

$$= |0000\rangle + |1111\rangle + |1100\rangle + |0011\rangle \tag{11}$$

$$= (|00\rangle + |11\rangle) \otimes (|00\rangle + |11\rangle) \tag{12}$$

$$= \mathrm{CNOT}_{12} H_1 |00\rangle \otimes \mathrm{CNOT}_{34} H_3 |00\rangle \tag{13}$$

$$= \mathrm{CNOT}_{12}\mathrm{CNOT}_{34} H_1 H_3 |0000\rangle\,. \tag{14}$$

However, preparing the logical $|\overline{00}\rangle$ state, which is a GHZ state, requires the ancilla qubit to check it was correctly prepared. A way of doing this is shown in Figure 2, where the run is discarded if the ancilla is measured as $|1\rangle$, an instance of pre-selection. After the state is prepared, the encoded circuit U may be run.

As it is merely capable of detecting errors, we are unable to flush out the errors with error correction gadgets, so we are limited in the depth of circuits we can perform, for which 10 to 100 was suggested as

reasonable limits. It was hoped that despite these limitations and with postselection, detectable differences in performance between the encoded circuit and an unencoded circuit would be found to satisfy the criterion.

The specific subfamily of circuits S recommended to be tested was parameterized by

1. A maximum depth T,
2. A maximum periodicity p, and
3. The number of circuit types r of each depth or periodicity

to be constructed in the following manner.

1. For each depth $t = 0, 1, \ldots, T$, choose r random circuits. If r is greater than the number of possible circuits, then include them all.
2. For each period $q = 1, 2, \ldots, q$, choose r random circuits of depth q each of which is repeated $\lfloor T/q \rfloor$ times.

This subfamily would have at most $r(T + 1) + rT(\ln p + 1)$ circuits. In addition, one may also construct families for each initial state preparable.

2.2.1 Other protocols

While the minimalist protocol based on the $[[4, 2, 2]]$ code was described in detail in the paper, other small codes could also similarly be used to demonstrate satisfying the criterion. The resource requirements of the other codes could also be fairly minimal, such as by using flag qubits in the five-qubit code, which would require only 6 qubits [2]. The efforts by Quantinuum [3] to show fault-tolerance using the 7-qubit code under a similar but modified criterion will also be reviewed in this report in detail. The demonstration by Google by scaling up surface codes uses a markedly different approach, looking at reduction in error rate in storage rather than with random circuits [4].

It should be noted that the criterion as proposed was merely intended as a guideline for early experimental demonstrations of fault-tolerance, and not necessarily practical to use for large-scale fault-tolerance. This is because there is no way to efficiently compute the ideal distribution $\{p_i\}$ in general when the circuits are non-Clifford and thus not amenable to stabilizer-circuit simulation.

As such, it was emphasized in the paper that the goal of such experiments should be to teach us something new about the physics of these systems rather than merely demonstrate reaching a milestone.

2.3 Hierarchy of fault-tolerance milestones

The basic demonstration of fault-tolerance is merely one milestone required in the path to fault-tolerant quantum computing at a scale that would be useful for executing algorithms that are believes to have a quanutm speedup.

In general, the following milestones are expected for any platform in general.

1. Build a qubit with low error rates in state preparation, measurement and storage, *i.e.* coherence times much longer than the gate times.
2. Build a system of many qubits with low entangling gate error rates.
3. Demonstrate a fault-tolerant gadget for logical state preparation.
4. Demonstrate break-even fault-tolerance in storage.
5. Show exponential suppression in storage error rate with increasing size. The Google surface code experiment aims to show this [4], but it is uncertain whether scaling up further would continue to reduce the error rates exponentially.
6. Demonstrate Clifford gate fault-tolerant gadgets reduce error rates.

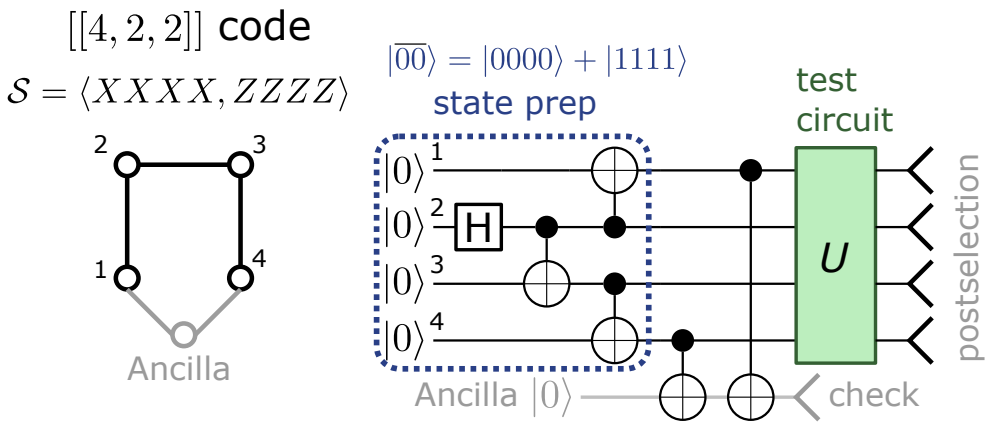

Figure 2: (left) The 5-qubit fault-tolerance experiment requires four qubits encoded in a $[[4, 2, 2]]$ quantum error-detecting code plus one ancilla qubit with the connectivity as shown in the graph. (right) The circuit for fault-tolerantly preparing the logical $|\overline{00}\rangle$ state. The run is discarded if the ancilla is measured to be $|1\rangle$. A circuit U is selected from the family of circuits to test. There is also postselection on the final outputs, with odd-weight outputs being discarded, and decoding by looking up the matching codewords in equations (6-9).

7. Demonstrate fault-tolerant encoded circuits have lower error rates for a family of circuits. This is the criterion in [1] and the Quantinuum experiment [5] is done in is spirit this but is restricted to preparing logical Bell states.

8. Demonstrate magic state distillation and injection to implement logical non-Clifford gate.

9. Demonstrate exponential suppression of errors far below threshold in large non-Clifford circuits.

3 Platforms, codes, gadgets used

Now that we have set the scene for the technological milestones on the path to fault-tolerance, let us summarize the state of the art papers that aim to achieve this. Before we dive into the details of each platform, we provide a quick comparison of the platforms and their parameters in table 1,

Metric	Harvard	Google	Quantinuum
Architecture	Rydberg Array	Super Conducting	Trapped Ion
Total Qubits	288	72	32
T1	$4s$	$20\mu s$	$> 60s$
T2	$> 1s$	$30\mu s$	$> 1s$
Native Gates	X,Y,Z,ZZ	H,CZ	ZZ
Connectivity	Arbitrary	Nearest neighbor	All-to-all
2-qubit Gate timing	$400ns$	34 ns	$\approx 100\mu s$
1q Gate errors	$0.1 - 0.4\%$	0.1%	0.0025%
2q Gate errors	0.5%	0.3%	0.14(1)%
SPAM errors	0.4%	Not reported	0.15%

Table 1: Summary of device parameters

4 Google

4.1 Hardware and design

The Google Sycamore processor is based on a layout of tunable superconducting qubits coupled to each other via tunable couplers. As an abstraction, each transmon can be considered as a two level system (since it is an anharmonic oscillator and other levels have different level spacing). Each transmon qubit has two controls on it: the magnetic flux control which changes the frequency and the control on the microwave drive which excites the qubits [6]. Thermal effects can easily decohere these qubits, so normally they are kept in a dilution refrigerator to maintain temperatures at a few mK, while all the control circuitry is placed outside, allowing to tune coupling between qubits as well as measure them all simultaneously. Single qubit gates are executed by applying microwave pulses which keeping the couplers off, but are still affected by two-level system defects: stray interactions whose effects are included in the single qubit gate fidelity. Even when the coupling is turned off, when gates are run simultaneously there is a chance of causing some residual coupling which introduces **crosstalk**. This is particularly noticeable for two-qubit gates: while individual two qubit gates run with an error rate of 0.3%, it jumps up to 0.6% when gates are operating simultaneously. The qubit readout is also benchmarked, and does not incur an error due to simultaneous measurement.

In this particular setup, we have a 72 qubit superconducting device supporting a 49 qubit distance-5 rotated surface code (25 data qubits + 24 measure qubits), which narrowly outperforms its subset of 17 qubit distance-3 code. Note that in the rotated surface code, we also have two-mode stabilizer generators.

4.2 Initialization, syndrome extraction and readout

This paper [4] began a first step towards fault tolerance, and hence they do not do any fault tolerant gates in this paper, rather the focus is on comparing the performance between distance-5 and distance-3 codes, hence this code can be thought of as a **logical memory**. To initialize, all the data qubits are prepared in the $|0\rangle$ or $|1\rangle$ state, which corresponds to an eigenstate of the Z stabilizers. In the first cycle, the X stabilizers are measured, projecting the code into an entangled state with the logical operators X_L and Z_L shown in Figure 3. In the subsequent cycles, the syndrome is extracted by measuring the values of the X

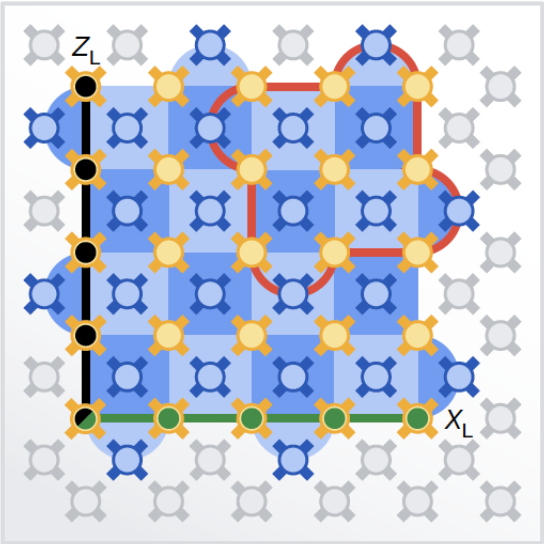

Figure 3: Rotated surface code that has the same geometry of qubits as the Sycamore chip. Yellow qubits are data qubits, blue qubits are measure qubits. Logical X and Z are also shown.

and Z stabilizers by applying the appropriate sequence of H and CZ gates. At every time step, the starting state for the stabilizer measurement qubit is $|+\rangle$ and four CZ gates are applied. For an odd number of errors, we obtain a syndrome. Note that at all times, each qubit participates in four CZ gates, and at the end of

the cycle, dynamical decoupling is applied to the data qubits to mitigate cross-talk while measure qubits are measured and reset. In the final cycle, they measure the data qubits in the Z basis to give both parity and measurement of the logical state.

4.3 Error detection and correction

Measurements of stabilizers allows a way to know if an error has occurred. Apart from correction, they can also be used for error detection. For the surface code, they notice a 12 percent increase in the error rate over 25 cycles for the distance 5 code compared to an 8 percent increase for the distance 3 code. They attribute this effect to states **leaking** into the non-computational basis states (which correspond to the higher excited states of the transmon); since the distance 5 code has more gates operating on it, has more chances to have stray interactions and thus leak out.

4.4 Physical noise

Because the process is about 10 times slower then two-qubit gates, readout and reset operations introduce significant idling errors. To simulate the errors, they start with the depolarising channel, include higher order effects such as leakage, and account for short T_1 and T_2 times. They construct a Pauli+ model, which is a generalization of the Pauli frame tracking since it decomposes the noisy gates into ideal Clifford channels and a generalized Pauli channel.

They also discuss correlated errors, which they describe separately as **timelike** and **spacelike** pairs. In a timelike pair, measurement and reset errors are caught by the same stabilizer in two different rounds. In a space-like pair, imagine a Z error that occurs on the data, and there are a number of gates in between, like the CZ gate and Hadamard, that propagate the error to additional check qubits, so the same error is propagated to multiple qubits in the same slice of the circuit.

Of course, there can always be combination of both space and timelike correlations. It is important to understand and characterize such errors because otherwise the decoder can interpret two correlated errors to be triggered by independent sources and cause a false correction to occur.

4.5 Decoding

A belief matching decoder is used, which is a combination of a belief propagation and minimum weight perfect matching algorithm. The decoder takes as input the error hypergraph, and using belief propagation, it updates the weight of the error probabilities on the hyperedges. The updated hypergraph is then decomposed into an X and Z type hypergraph, which is then run through a MWPM decoder to find the most likely error events. They also describe a tensor network decoder which is built from a modification of a maximum likelihood decoder, which while optimal, is much slower.

4.6 Key results

- As shown in the plot in Figure 4, the logical error rate for distance-5 codes are markedly better than distance 3-codes, after averaging them over four blocks. Note that there is no post selection for high energy events. In particular, the logical error per cycle is $\epsilon_5 = (2.914 \pm 0.016\%)$ for the distance 5 code, compared to $\epsilon_3 = (3.028 \pm 0.023)\%$ for distance-3 codes. This result is important as a first step in experimentally showing that error correction helps to decrease error rates as distance increases. However, since the physical error rates are barely below threshold, it is expected that improving the physical error rates will considerably improve the performance of correction.

- Another important observation here is that as known physical error rates increase, the logical error rate may not necessarily improve as the distance gets bigger. They also therefore experiment with a distance-25 bit-flip repetition code, which allows to achieve a logical error per cycle of $(1.7+0.3)\times10^{-6}$. This is significantly better than the XZZX surface codes, and also shows a proof of concept that low logical error rates are possible for superconducting devices. This also exposes a previously unknown source of error, which is high energy bursts of cosmic radiation which can temporarily cause highly

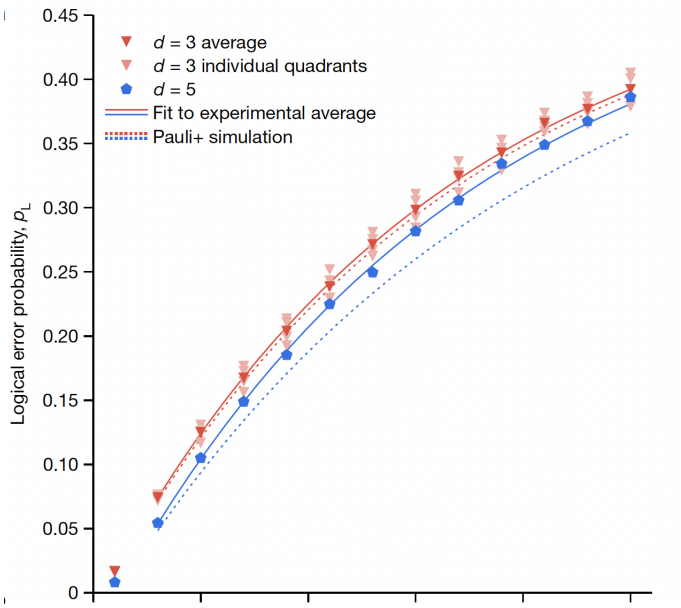

Figure 4: Plot of logical error probability over 25 quantum error correction cycles. $d = 5$ code consistently outperforms distance-3 code. There is no post selection on highly correlated events such as cosmic ray events or leakage.

correlated errors. These cosmic rays will be important to study and mitigate for superconducting circuits going forward.

5 Harvard and collaborations

Harvard, UMD, and MIT also contributed to showcasing fault-tolerance using neutral atom systems, in the paper titled "Logical quantum processor based on reconfigurable atom arrays" [7]. The paper explores three key aspects demonstrating the capabilities of the logical Rydberg processor. Firstly, they exhibit the Quantum Error Correction (QEC) property, showcasing improved performance as the code distance scales up, demonstrated through the preparation of logical Bell pairs using the surface code. Next, they illustrate the implementation of a fault-tolerant logical circuit to prepare GHZ states utilizing the [[7,1,3]] color code. Lastly, they utilize 3D code blocks of dimensions [[8,3,2]] to actualize computationally challenging sampling circuits equivalent to IQP circuits, using a total of 48 logical qubits entangled with hypercube connectivity. This section overviews the hardware of the neutral atom platform, highlighting its advantages in fault-tolerance, and discusses the first two significant demonstrations, providing explanations of the experimental details and key results.

5.1 Physical Parameters

In the previous section, we have presented most of the key parameters of the hardware in the Table 1. Here we will illustrate the most important factors related to the demonstration experiments in more detail.

Rydberg atoms are atoms in a highly excited state, where the principle quantum number can range from 10 to several hundreds. Therefore, they have a long coherence time and strong interaction. The Lukin group [8] realized a 51-qubits quantum simulator in 2017 by using Rydberg atoms as the building block for physical qubits and trapping Rydberg atoms inside tightly focused laser beams to individually control them. In Figure 5, the Rydberg atoms (blue) are trapped individually by the optical tweezers (red) and coupled together by lasers (orange). The ground state $|g\rangle$ is used as the $|0\rangle$ state and the Rydberg state $|r\rangle$ is used as the $|1\rangle$ state with Ω being the Rabi frequency which characterizes the strength of the interaction between the atom and the laser field. With the optical tweezers, they are capable of cleverly trapping the atom and

moving them around spatially which creates a the quantum simulator for an Ising-type quantum spin model.

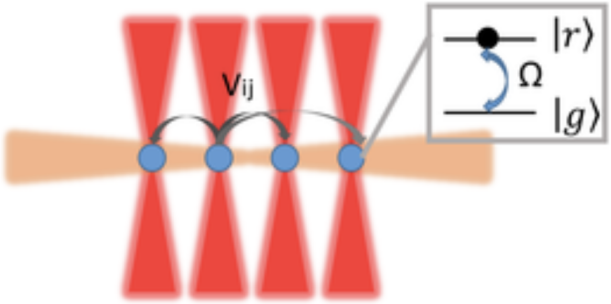

Figure 5: Rydberg atoms (blue), are held in place by tweezers (red) and interacted by a laser (orange).[8]

Since Rydberg atoms have long coherence times and techniques to control them such as laser cooling and optical tweezers can easily generalized, the potential of Rydberg quantum simulators has been extended in 2024 with the system scaling up to 288 qubits [7]. The main capabilities of the system can be summarized as

1. High two-qubit gate fidelity,
2. Arbitrary connectivity,
3. Fully programmable single-qubit rotations, and
4. Mid-circuit measurements.

Besides, there are several other outstanding features of the 288-qubit system which help with demonstrating the QEC property or running Fault-Tolerant (FT) algorithms. The first feature is its *zone architecture*. As shown in Figure 6, where the atoms are divided into three zones: a storage zone, an entangling zone and a readout zone. The division of different zones highly avoids cross walk while the readout zone makes it possible to do mid-circuit measurement without disturbing the coherence of the computational qubits still in operation. The second is logical level-control. More specifically, instead of shooting laser pulses to individual physical qubits, the lasers can be applied to the code block which represents the logical qubits. This feature allows the system to be easily scaled up without increasing the number of controls. Additionally, the Rydberg atoms have low idling error which leads to 1% logical decoherence per additional encoding step. Those nice properties make the Rydberg quantum simulator a good candidate to demonstrate fault-tolerance.

5.2 Demonstration of QEC property

Similarly to the Google experiment, the Harvard collaboration also demonstrated a quantum memory of logical Bell pairs with surface codes. According to the threshold theorem, as long as the physical error rate is below the threshold, the performance of the encoding circuits will improve with the code-distance. The goal of experiments is to show that the logical error is reduced with larger code distances of the surface code.

5.2.1 Experiment details and results

As discussed in the above section, Rydberg atoms have long coherence times ($T_2 > 1s$), which can be increased to the scale of ten to hundreds of seconds by an assembled register of individually-controlled qubits[9]. Also, the storage zone allows the physical qubits being safely and idly stored for a long period of time without requiring active correction like repeated stabilizer measurements. Thus, improving the coherence of the logical qubits will not dramatically improve the performance of the encoding circuits, but improving the fidelity of the entangling gates will yield larger gains. Thus the experiments mainly test the improvement of the transversal CNOT gate as the code distance is scaled up. In particular, the code family used here is the rotated surface codes with code parameters $[[d^2, 1, d]]$. The experimental circuits and a flowchart summarizing the protocol are shown in Figure 7. We shall continue to discuss these in detail in teh following parts.

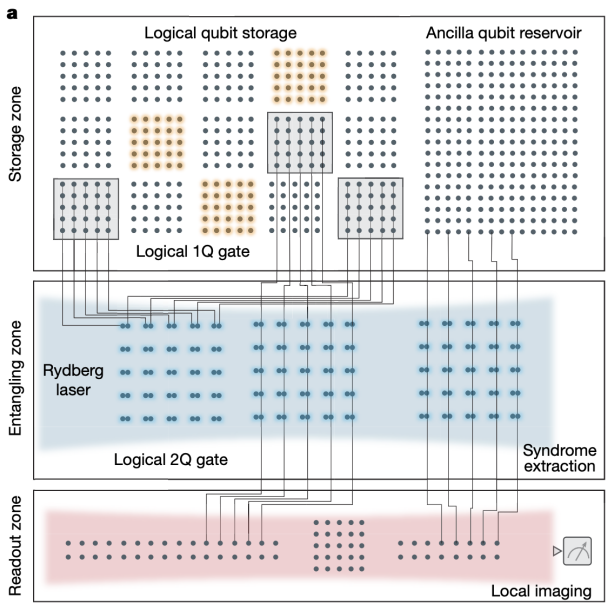

Figure 6: The zoned architecture of the system with 3 zones [7].

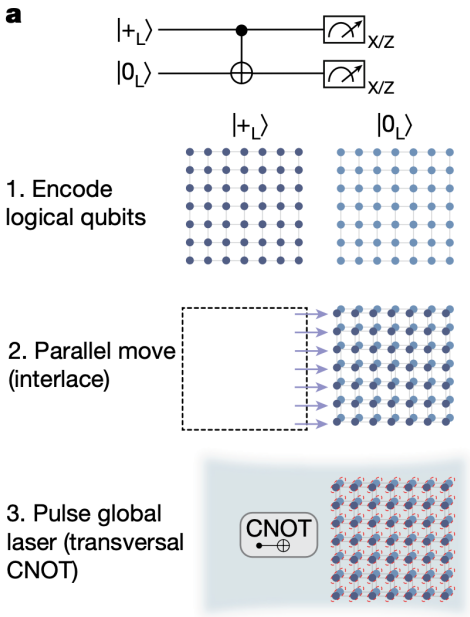

Figure 7: Logical Bell Pair Experiment [7].

Initialization In order to initialize the Bell state, we first need to prepare $|+\rangle_L$ and $|0\rangle_L$ with two surface codes. The first code block starts with all physical qubits in state $|+\rangle$ to prepare an eigenstate of X_L and all X stabilizers, and then measures all Z stabilizer generators with ancilla qubits. The second code block prepares $|0\rangle_L$ in the similar way by exchanging X and Z stabilizers.

Transversal CNOT The logical CNOT is transversal for surface code and the control pulse can be applied globally due to the entangling zone architecture. After parallel movement of the atoms into the entangling zone, then the global Rdyberg excitation laser is pulsed to realize the physical CNOT gate on each pair of

the blocks. The process allows a high-fidelity and transversal CNOT gate in a single parallel step.

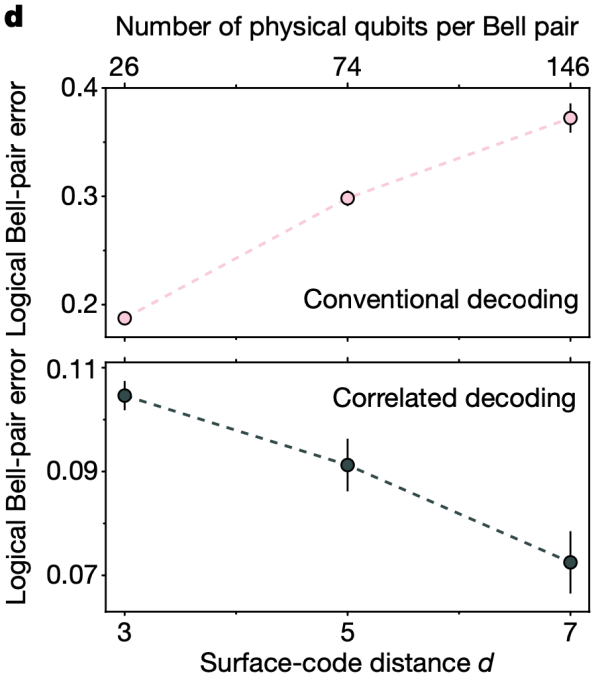

Figure 8: Logical Bell pair experiment results [7].

Projective Readout After the transversal CNOT, projective measurement is used to evaluate the logical Bell state stabilizer $X_L^1 X_L^2$ and $Z_L^1 Z_L^2$. However, since there is only a single round of syndrome measurement, the state preparation step is not fault-tolerant (labelled nFT). More specifically, if there is no error in the data blocks we should get the trivial syndrome, but a measurement error for any of the ancilla will incurs an error syndrome. Then by the correction step, we will introduce errors for the initialization step. Moreover, the CNOT gate will propagate X errors from the control to target and Z errors from the target to the control. Therefore, the nFT initialization will influence the measurement results of logical Bell state stabilizers. The effect can also be seen from the Figure 8 with conventional decoding (minimum-weight perfect matching within each individual code block), i.e. the logical Bell pair has a larger error when scaling up the code-distance or system size. This is because the nFT error in the initialization step accumulates and propagates through transversal gate as the system grows. In order to suppress the effect of nFT errors, another trick called "correlated Decoding" is introduced by the authors. The idea of correlated decoding comes from the fact that the error after propagating through a transversal CNOT gate has strong correlation. Considering the control and target code blocks together by a joint decoding graph with edges and hyperedges (link between blocks) can partially mitigate the error introduced by the nFT initialization. The results with correlated decoding are shown in Figure 8, where we can see that the logical Bell pair has a smaller error with increasing code-distance. This is also suggests that with the correlated decoding trick, they surpass the threshold for this specific circuit. However, in practical usage, the transversal CNOT gate should be used with many rounds of error syndrome extraction which is expected to have a lower threshold than this single demonstration.

5.3 Demonstration of fault-tolerant state preparation of logical GHZ states

Another important demonstration done in the paper [7] at the algorithmic level is to use a fault-tolerant Clifford circuit to prepare the logical GHZ state. In the experiment, the $[[7, 1, 3]]$ color code is used as it can do the Hadamard gate H, the $\pi/2$ phase gate S and the CNOT gate transversally. The goal here is to show fault-tolerance in the sense that the encoded version of the circuit has higher fidelity.

5.3.1 Experiment details

Creation of the GHZ state is the test case. The FT circuits to create the logical GHZ state are presented in Figure 9. Each circuit is divided into two parts, the FT state preparation for $|0\rangle_L$ and the FT creation of the GHZ state. In the first part, all ten logical qubits are encoded in the $[[7, 1, 3]]$ code by the encoding circuit presented in the left part of Figure 9. Five of them are computational logical qubits and the other five are the ancillas which are used to detect errors on the $|0\rangle_L$ via transversal CNOT gates. After the state is prepared, the five ancillary logical qubits are moved to the storage zone. The remaining five computational logical qubits are used to prepare the logical GHZ state. Note that here in the circuit, there are actually four of them used to make a 4-qubit GHZ state. However, the circuits can also be modified make a 5-qubit GHZ state, but they only used four as a demonstration to allow the convenience of doing full tomography. Finally, logical Clifford rotations are used for direct fidelity estimation and tomography.

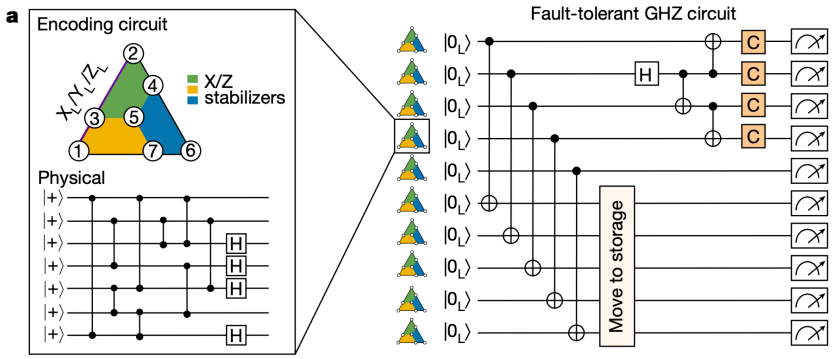

Figure 9: Circuit to preparte the logical GHZ state in the Steane code [7]

5.3.2 Key results

State Initailization An error detection scheme is applied to the logical $|0\rangle_L$ state initialization. By post-selecting on the ancilla logical flags, the average logical fidelity on the five computational qubits of the initialization is $99.91 + 0.04\%$, as seen in Figure 10. We notice that the logical FT (with CNOT) case where error detection is done can exceed both the physical qubit $|0\rangle$ initialization fidelity (99.34%) and physical two-qubit gate fidelity 99.5% [10]. In this step, we can see the encoded circuits to prepare the zero state have better performance than the physical unencoded version. Thus, it actually achieves fault-tolerance on these small circuits to prepare the zero state with error detection.

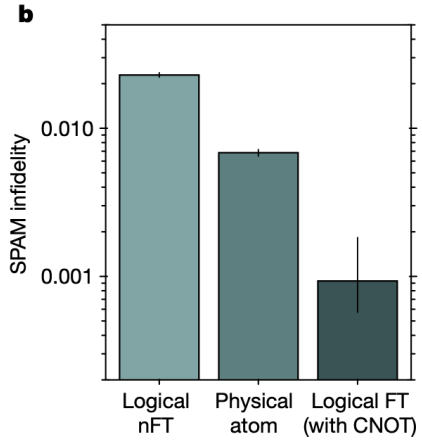

Figure 10: Results for preparing the logical zero state under various post-selection schemes [7].

Logical GHZ fidelity The second part of the circuit is to use four logical zero states to create the GHZ states. The results are shown in the Figure 11. There are three different scenarios presented in Figure11(c). The nFT (non-Fault-Tolerant) experiment is with encoded logical qubits and transversal gates but without any post-selection. The FT (Fault-Tolerant) result refers to the experiment with error-detection on the logical flags and finally the EDFT (Error Detection Fault Tolerant) is the case where full error detection is applied by postselecting on flags and stabilizers of the computation logical qubits. The nFT case only has the logical GHZ fidelity being around 55%, which makes sense as there is no error detection step and errors accumulate as more qubits are added into the system. The FT case reaches the fidelity 72% which is higher than the nFT case. Theoretically, the logical GHZ state should have a failure probability scaling as p^2 if the physical error is p. However, the error accumulated during the operation of the circuits and the spreading of the errors through transversal gates effectively increases p and limits the fidelity to 72%. Therefore, we can not see that the logical error is smaller than the physical error rate here, thereby failing to demonstrate FT. But for the EDFT case, postselecting on all the stabilizers of the computational logicals being correct will lead to a very high fidelity (99.85%), but of course at the high cost of postselection.

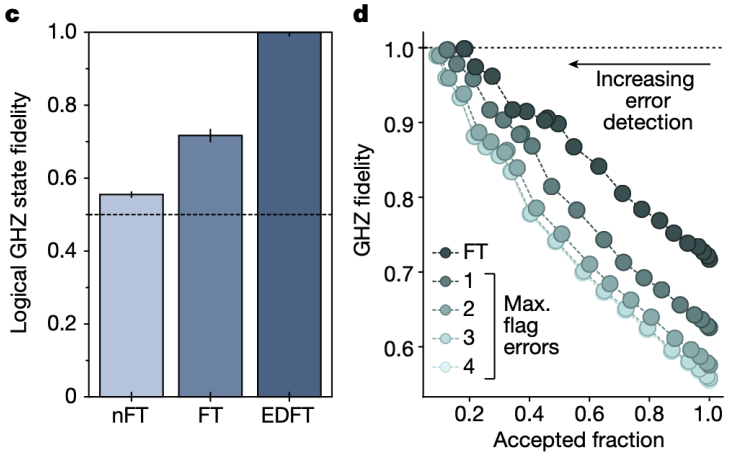

Figure 11: Fidelities for preparing the logical GHZ state under various post-selection schemes [7].

Partial postselection As discussed in the previous paragraph, if we want to reach to the highest fidelity, the overhead of post-selection should also be considered. Here the trade-off is studied by a partial postselection idea. The idea comes from the fact that not all nontrivial syndromes will lead to algorithmic failure with equal probability, so we can only discard those which have higher probability to give us the wrong logical states according to the weight of the correlated matching in the whole algorithm. Figure 11(d) gives us the sense of this trade-off, where we can notice that the if the strictness of error detection increases, the GHZ state fidelity also increases. Discarding 50% of the experiment repetitions will lead to a logical GHZ state fidelity of around 90%.

6 Quantinuum: Logical Bell State Preparation

In the recent paper "Demonstration of logical qubits and repeated error correction with better-than-physical error rates" from authors at Microsoft and Quantinuum [5], logical Bell state-preparation below physical error rates was demonstrated using the Steane CSS [[7,1,3]] Color code. Furthermore, the paper demonstrated repeated rounds of error correction for a family of single qubit Clifford and `CX` gates within a single block of a newly developed [[12,2,4]] so-called Carbon code. In the later experiment, the logical error rates are better than the physical. In this section, we outline the key details of the experiment, the quantum computer used, and the error correcting gadgets demonstrated in the work. We focus primarily on the first experiment (preparing a logical Bell State using the Steane code) as details on the gadgets and properties of the Carbon

code are limited and the full theory paper has yet to be released [11]. We conclude by discussing the implications and limitation of the work.

6.1 Hardware Summary

The Quantinuum H2-1 processor is a trapped ion based system which encodes computational states in the hyperfine clock states of the $^{171}Yb^{+}$ ion, with programmable single qubit gates and two-qubit gates arising from pair-wise Mølmer-Sørenson (MS) type interactions which are wrapped in single qubit gates to create the base level $U_{ZZ}(\theta) = exp(-i\frac{\theta}{2}ZZ)$ gate. The processor uses a quantum charge coupled device (QCCD) architecture in which individual ions are shuttled between storage and computational zones in the trap, of which there are 4 in the device. The system can perform up to 4 parallel two-qubit gates at a time with no hardware geometry requirements. A particular advantage of the device is the scalability of the architecture, although 2D shuttling (beyond 1D shuttling used here) will likely be required for future devices. Parallel gates are also currently limited by the number of interaction zones which can be increased for future generations [12].

Each computational $^{171}Yb^{+}$ ion is paired with a $^{138}Ba^{+}$ ion for sympathetic cooling. MS gates make use of the normal mode oscillations between charge-coupled ions in a trap (phonon modes), and require near-zero occupation of these modes for high-fidelity gates. Sympathetic side-band cooling allows for cooling the ion chains of either 2 or 4 ions via laser interactions with only the Barium ions, which do not contain quantum information and thus do not disturb the information stored in the Yb hyper-fine states. Sympathetic cooling is an essential component of the hardware, as it allows faster ion transport and many ion-chain merge/split operations throughout the circuit (in contrast to single-chain single-species ion traps).

A key limiting factor in the H2-1 system is the long time required per experimental shot. In test circuits, the gate-execution time represents 1-2% of the wall-time budget, while the time required for transport and cooling are around 59% and 29% respectively, depending on the specific gates and connections required by a given circuit. In many cases, the actual time per shot is over a second, which represents a significant overhead cost in executing circuits. This clock-speed drawback is mitigated by the high fidelity gates permitted by the system – the first system has better than 99.9% fidelity across all gates (reported separately).

6.2 [[7,1,3]] Code and Gadgets

We focus only on the details of the gadgets used in the [[7,1,3]] CSS Color code experiment (which we will refer to as the Color code going forward). This particular codes has a few nice advantages. It is a CSS code and thus permits a natural measurement gadget; it also permits transversal H and `CX` gates, which are necessary for the Bell state preparation circuit (contrast with the [[5,1,3]] code which does not permit a two-qubit transversal Clifford gate); and it is well studied, resulting in optimized fault-tolerant gadgets using minimal ancilla qubits. As a result of the syndrome and state-preparation gadget improvements, it only requires 7 physical qubits and 3 ancilla qubits per logical qubits. The experiment then requires only 20 physical qubits, which is below the maximum capacity of the H2-1 system of 32 qubits.

6.2.1 State Preparation

The state-preparation gadget requires only a single flag ancilla which is used to verify the success of the state-preparation of $|\bar{0}\rangle$. The full circuit and gadget is outlined in Goto (2016) [13], and we summarize the key points here, referencing Figure 12. There are two key considerations which allow for using a single stabilizer measurement to verify the state preparation. The circuit consists of three notable regions, marked as (a-c). Sections a and b are a non-fault tolerant circuit used to prepare $|\bar{0}\rangle$, and as pointed out by Goto 2016, the stabilizers describing the state after the gates in section (a) are given in table 2, and have the useful benefit of converting any weight two errors which may have been introduced by the gates into equivalent weight 1 errors. Thus, only the gates in section (b) are able to introduce weight two errors. Additionally, the final state $|\bar{0}\rangle$ is not sensitive to weight two Z type errors due to the degeneracy of the code – each Z error can be converted to an equivalent weight 1 or identity error on the state. As there are two `CX` gates in section (b), and we are not worried about correlated Z errors. Thus we only need to verify that no weight two X errors were introduced by the `CX` gates. Checking the Color code stabilizer $Z_0Z_5Z_6$ suffices to verify both possible correlated errors X_1X_5, and X_4X_6. Success in the state-preparation corresponds to measuring

$Z_0Z_1Z_2$	$X_0X_1X_5$	$X_0X_2X_6$	$X_3X_5X_6$
Z_1Z_4	Z_3Z_5	$Z_2Z_3Z_6$	

Table 2: The stabilizers describing the state after the first 6 entangling gates in the logical $|\bar{0}\rangle$ preparation circuit.

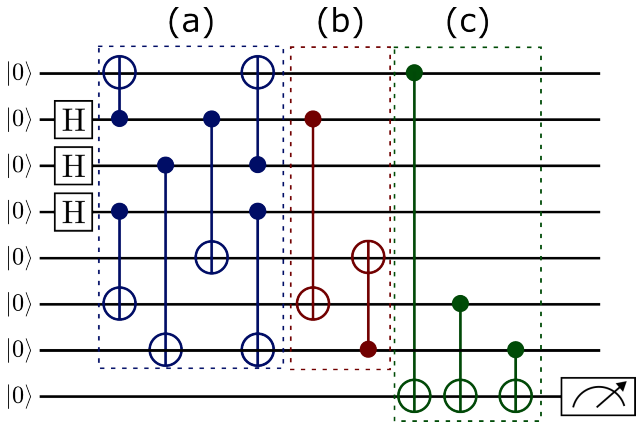

Figure 12: The state preparation gadget used in the $|\bar{0}\rangle$ gadget. The final qubit is the ancilla used to extract the $Z_0Z_5Z_6$ stabilizer measurement, and sections a-c are marked

0 in the check qubit. In the Quantinuum experiment, they repeat the state-preparation procedure up to 3 times (fully resetting the circuit between encoding attempts). The logical SPAM errors are not reported in this particular paper, but were reported as 99.984(9)% after 3 rounds for an earlier generation of the hardware [14].

6.2.2 Syndrome Measurement

Figure 13 outlines the specific syndrome measurement gadget used in the experiment. Originally formulated in Reichardt (2020) and expanded on in Ryan-Anderson (2021)[14, 15], the gadget would ideally use adaptive circuits to extract an accurate syndrome. Using three ancilla qubits, three syndromes are simultaneously extracted in a procedure that flags any errors on the ancilla qubits which might spread into a weight two error in the code-block (hook errors). In the adaptive version of the gadget, the simultaneous syndrome gadgets are applied and checked for hook errors. In the event that an error is not flagged, the second round of 3 syndromes are extracted. If a measurement round results in a flag, unflagged stabilizers are measured. The fault-tolerance of the gadget relies on detecting possible errors during the syndrome extraction using the flag qubits. When a flag is found, the single permitted error during the QEC gadget has been detected and the remaining syndromes can be directly extracted without additional fault-tolerant consideration.

In the Quantinuum experiment, the target is to prepare a Bell state fault-tolerantly. As such, they consider the Bell-state as a resource state in the same vein as the $|\bar{0}\rangle$ state-preparation gadget prepares a resource state. That is, error correction is not needed during the state-preparation, only error detection. Post-selecting on non-trivial syndromes provides the error detection, and unflagged syndromes are not needed in the QEC gadget as they are only necessary for active error correction mid-circuit. Thus, only the two circuits outlined in Figure 13 are executed, and post-selection occurs on non-trivial syndromes (which includes post-selecting on flags).

6.2.3 Readout

Readout of the Color code is performed using standard CSS readout and decoding, which does not require any additional ancilla qubits. As this topic was covered in the course, we do not cover it here except to note that it does not require any additional ancilla.

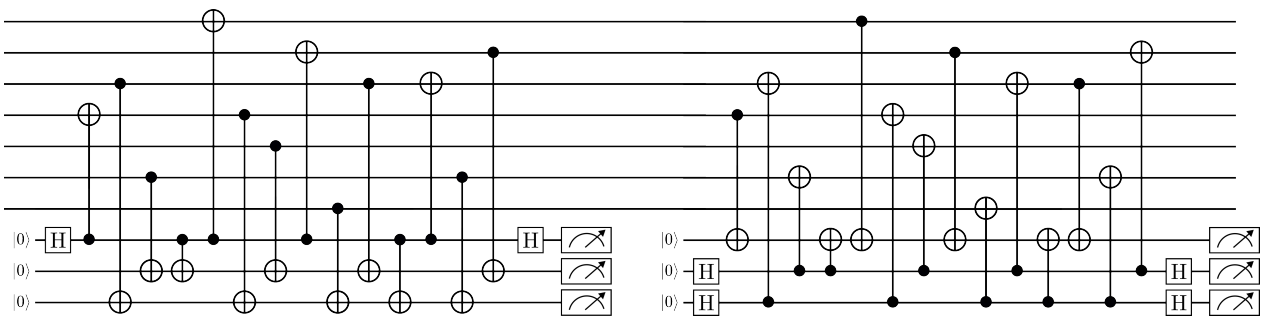

Figure 13: The full flagged syndrome extraction circuit. This particular circuit is not fault-tolerant, but is used instead for pre-selection in the Bell state preparation circuit. Each half of the circuit measures three syndromes in parallel [15].

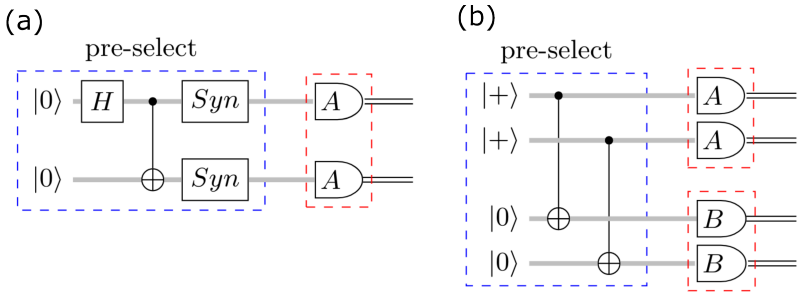

Figure 14: (a) Logical circuit used for the Color code circuits. (b) Logical circuit used to prepare two Bell states in the Carbon code. Blue dotted lines surround the "pre-selection" region in which non-trivial flagged-syndromes are rejected. Red dotted lines surround the error corrected region of the circuit (readout). Figured adapted from [5].

6.3 Experiment results

The Microsoft and Quantinuum experiment using the Color code demonstrated the fault-tolerant preparation of a logical Bell state with logical error rates lower than the physical error rates. The circuit, outlined in Figure 14, prepares the logical state $|\bar{0}\bar{0}\rangle$, applies transversal H and `CX` gates, post-selects on one round of syndrome measurements, and performs logical measurement in the X, Y, and Z Pauli bases. The Bell state fidelity is determined by the parity of the measurements. The circuit is operated in both "pre-selection" and "post-selection" modes, rejecting non-trivial syndromes during the circuit execution in the former case and rejecting all non-trivial syndromes in both readout and syndrome measurement in the later case. In the former case, approximately 75% of shots are accepted, while in the full error-detection operation mode approximately 72% of total shots are accepted.

The key result of the paper is the statistically significant separation in fidelity for each of the two modes for the logical circuit and the physical circuit using two qubits. The error rates for the physical circuit were reported at $0.50^{+0.03\%}_{-0.03\%}$, while the pre-selected and error detected circuits had fidelities of $0.05^{+0.04\%}_{-0.03\%}$, and $0.001^{+0.013\%}_{-0.001\%}$, respectively. These results represent gains of 9.8 and 500 over the physical error rates.

6.4 Carbon Code

The same work reports results using a CSS [[12,2,4]] Carbon code, which has yet to be reported in the literature [11]. As there are few details on the code, we will not dwell on the details of the experiment except to report the broad characteristics and results. The Carbon code permits transversal H and `CX` gates, CSS type measurements, and a Knill measurement based scheme for syndrome measurement. Two separate experiments were performed: the first involved the preparation of two Bell states across two Carbon code blocks utilizing transversal gates between the blocks, while the second comprised repeated rounds of two-qubit

Circuit Error rates	1 round	2 rounds	3 rounds
Physical circuit	$0.35^{+0.09}_{-0.08}\%$	$0.8^{+0.1}_{-0.1}\%$	$1.2^{+0.2}_{-0.1}\%$
Logical circuit	$0.03^{+0.06}_{-0.03}\%$	$0.4^{+0.2}_{-0.1}\%$	$0.8^{+0.3}_{-0.3}\%$

Table 3: Circuit error rates

gates and syndrome measurements within a single Carbon code block [5].

The first experiment, preparing Bell States, performed similarly to the Color code experiment in terms of physical, pre-selected, and post-selected error rates of $0.8^{+0.1}_{-0.1}\%$, $0.17^{+0.07}_{-0.06}\%$, and $0.001^{+0.015\%}_{-0.001\%}$ with gains of 4.7 and 800 over the physical circuit, respectively.

The second experiment required a single block of the Carbon code, in addition to an ancilla block of 12 qubits used for a Knill-type error correction gadget. The circuit consisted of up to 3 rounds of random Clifford circuits using each using two gates from C_0X_1, C_1X_0, or `SWAP` gates. Thus, the logical circuit over 3 rounds of error correction contained up to 6 two-qubit gates and 12 single qubit Clifford gates, counting `SWAP` gates as a single two-qubit gate. The motivation for the particular circuit is that `CX` and `SWAP` gates can be performed within the Carbon code block using only qubit permutations, which are high fidelity operations with a routing based architecture. Thus, no transversal `CX` gates are needed in the particular logical circuit, in contrast to the Bell-state preparation experiment which applied transversal `CX` gates between two blocks.

As the code is distance 4, all weight 1 errors can be corrected, while weight 2 errors can be detected. Additionally, some weight 2 errors can be decoded correctly with high-likelihood. The circuit operates in a hybrid error-correction, error-detection mode in which high likelihood errors are decoded and corrected, while syndromes which are difficult to decode with high likelihood are post-selected on.

The experiment resulted in statistically significant improvements over the physical error rates for 1 and 2 rounds of error correction, while the error rate for 3 rounds overlaps slightly with the physical error rate distribution due to limited shots in the logical circuit operation mode. In the physical error rate case, the circuits have errors increasing by approximately 0.4% per round – the logical circuit has a similar error per round, but a much lower error rate after the first round, as described in Table 3.

6.5 Discussion

These experiments on the Quantinuum H2-1 system represent the strongest experiments in full error correction toward implementing Gottesman's 2016 proposal for demonstrating error correction. Indeed, error correction via error detection, as originally proposed, has been fully realized for families of multi-qubit Clifford circuits. The stronger condition of active error correction with physical error rates above logical error rates has nearly been realized, but cannot yet be reported with full confidence. Microsoft and Quantinuum's experiments in Bell-state preparation via pre-selection demonstrate robust state-preparation and the application of certain gadgets, but rely heavily on error detection, which is expected to have much stronger performance over error correction on the same QECC. The Carbon code experiment using a family of two-qubit Clifford circuits is a much stronger case for full fault-tolerance, with a few limitations. The first limitation is that the circuits also rely on error detection and pre-selection to increase the fidelities in a way that is not scalable to larger circuits. However, they do demonstrate multiple rounds of error correction with better than physical error rates, and the hybrid correction/detection scheme used for the distance 4 code is fundamentally more scalable than pure post-selection. Second, the Carbon code experiment relies on lower-error, permutation-based, within-block gates which will not scale to experiments beyond two-qubits. A much more convincing experiment would demonstrate a full family of multi-qubit circuits that is fully scalable to larger numbers of qubits by including transversal gates as well. Unfortunately, due to qubit limitations, such an experiment was not possible on the H2-1 system.

Another key challenge with Quantinuum's experiments is the need to optimize around limited numbers of qubits, which additionally prevents testing larger distance codes as what has been done in the Google and Harvard experiments on surface codes. With the same gate fidelity and larger numbers of qubits, either with future iterations of the H2 architecture or the future H3 architecture using 2D ion grids and routing lattices, unambiguous experiments demonstrating logical circuits may be demonstrated. While remarkable developments in demonstrating the building blocks of error correction has been made, a few key elements

remain to be further demonstrated.

7 Quantinuum: quantum teleportation of a logical state

In [3] , Quantinuum establishes a use case for early fault tolerance experiments by showing quantum teleportation across three logical qubits. They perform gates two different ways: via lattice surgery on color codes, and using transversal gates. They use the Steane [[7,1,3]] code to do so, and since the Steane code is a CSS code, it naturally hosts a set of transversal gates.

7.1 Physical design

The task is to send a qubit state $|\Psi\rangle$ on the first qubit to the third qubit. A Bell state $\frac{1}{2}(|00\rangle + |11\rangle)$ is initialized for the second and third qubits to share. After entangling the first and the second qubits, joint measurements are performed which correspond to measuring $X \otimes X$ and $Z \otimes Z$ on both qubits. The basic teleportation circuit is given in Figure 15.

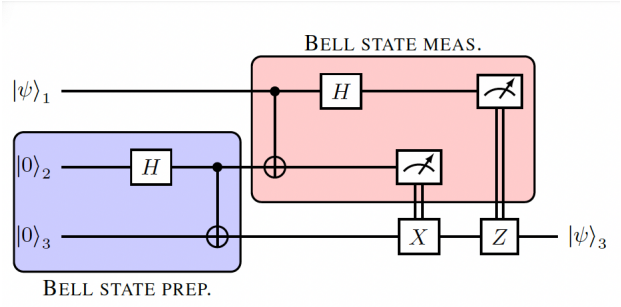

Figure 15: The general teleportation circuit

The experimenters start by performing this circuit at the physical level to establish a physical performance baseline, and repeat and average this across four different zones in their device (measurement and gating zones in the H2 trap). They send a set of six informationally complete states across this circuit and measure state fidelity of the states as $F_a = \frac{1}{6}\sum p_\Psi$. They also measure process fidelities as $F_p = ((d+1)F_a - 1)/d$. This process does not distinguish different sources of error as the focus is on different protocols and their fidelity.

At the physical level, the process fidelity comes out to be $0.9895^{\pm 3}$. They do note that state fidelity varies across different choices of states and justify it as their devices' sensitivities to different states.

7.2 Logical protocols

7.2.1 Transversal circuits

The first teleportation protocol is carried using a transversal gate set. While a transversal gate set does not propagate one fault to many others, to prevent the accumulation of probabilistic noise, they use QEC gadgets between layers of transversal gates. They also justify that since their gates are not exactly at threshold, it is also useful to have QEC gadgets between logical gates to benchmark performances more accurately. It is worth noting that their physical CNOT perform at 10^{-3} and their logical CNOTs are near a pseudothreshold. The circuit for the transversal teleportation is given in Figure 16. The above circuit is encoded using the $[[7, 1, 3]]$ code.

It is worth noting the similarities and differences between Figure 15 and Figure 16. As we learnt in class, the circuit is replaced by a fault tolerant circuit in the following way: states are replaced by logical states, gates are replaced by gadgets, and every gadget is ideally interspersed with QEC gadgets. However, here they only have a QEC gadget in the encoding stage, and the reason for this is that since QEC gadgets themselves can introduce errors, they are used more as a benchmark on performance with and without QEC rather than as a gadget in a fully fault tolerant circuit.

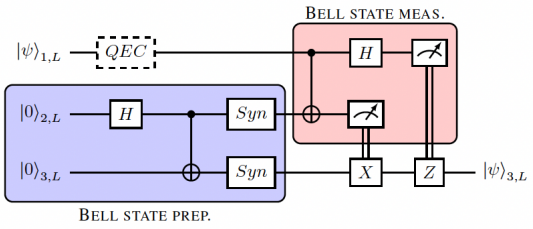

Figure 16: The logical teleportation circuit. The QEC represents a round of QEC followed by two rounds of adaptive syndrome extraction, while *Syn* represent flagged syndrome extraction used to post select on preparations with a non-trivial syndrome

State preparation:

For the first layer, we need to encode states in a fault tolerant manner, hence all three states in Figure 3 are initialized to their respective logical states using a fault tolerant encoder. For qubits 2 and 3, one can measure $\bar{Z}$ on both qubits in a non-destructive manner. Next transversal Hadamard and CNOT gates are performed, followed by syndrome extraction Syn using flag circuitry. Using flag circuitry allows to check for errors while extracting syndromes, as described above, and the ones with non-trivial syndromes are post-selected away. Note that this first block (purple block) corresponds to creating a resource state, and we assume we have access to **postselection** in this step so we can prepare the required resource states offline by repeating verification and syndrome extraction as many times as possible, and to as high of a fidelity as possible. However, note that even in this Bell state preparation gadget, it is not possible to detect weight-3 errors.

The state $|\Psi\rangle$ is similarly encoded using a FT encoder, and is repeated three times (1 round of full QEC, and two rounds of flagged syndrome extraction). There is also no post selection for the state on qubit 1. They also do two versions of the same experiment, one with a QEC gadget after the logical state preparation, dubbed as 1QEC, and one without, named 0QEC. The QEC gadget is the full adaptive syndrome extraction circuit, also described above, along with one round of active error correction. However, they do note that adding the QEC gadget causes the process fidelity to decrease from $0.989^{\pm 1}$ to $0.975^{\pm 2}$, which suggests that QEC gadget is one of the noisier components in the circuit.

State measurement:

The goal is to apply Bell measurement on qubits one and two, and then send classical information to qubit three so it can be updated. To measure the third qubit, they apply a single qubit transversal gate to measure in a particular basis. Overall, decoding is done four times in this circuit: once during the QEC gadget and then once per qubit at the end of the circuit when they are measured destructively.

In order to perform gates, they use two different methods, one being applying gates transversally, while the other is to perform lattice surgery. Note that the $[[7, 1, 3]]$ code is a color code:

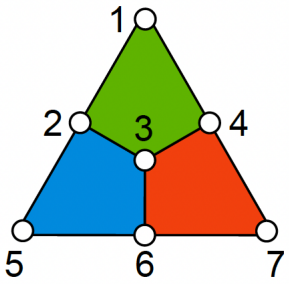

Figure 17: [[7,1,3]] code as a triangular color code. The logical representatives are operators on the boundary $X_1X_3X_5$ and $Z_1Z_3Z_5$.

7.2.2 Lattice surgery

In order to perform gates, they measure operators on the boundary in such a way that it corresponds to M_{XX} and M_{ZZ} (this corresponds to measuring $\overline{X}$ on two logical qubits for separate code blocks). In order to do teleportation, a joint $\overline{X_2X_3}$ is measured, along with flag error correction to catch hook errors. The next step is to measure $\overline{Z_1Z_2}$, again with a flag circuit.

The lattice surgery circuit has a process fidelity of $0.851^{\pm 9}$, and after post selection, it gives a process fidelity of $0.962^{\pm}5$. While everything else in the lattice surgery was kept to be the same as the transversal experiment, the lattice surgery circuit was longer than the transversal one, which presents more opportunity to incur errors. The circuit to implement teleportation via lattice surgery is given is Figure 18. They perform

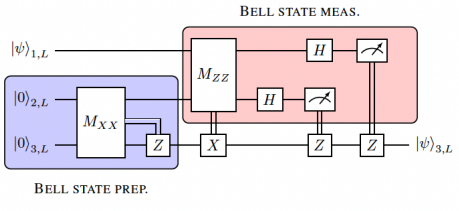

Figure 18: Logical teleportation via lattice surgery

another version of the same circuit but without preparing a Bell resource state, and they accomplish this using one-bit teleportation.

8 Conclusion

The minimum requirements for running a fault-tolerant algorithm, as proposed by Gottesman using error detection schemes, has been demonstrated in both the Quantinuum and Harvard experiments. In each case, post-selection improves circuit fidelity in meaningful amounts and brings circuit fidelity above the physical fidelities. However, there remain stronger conditions for scalable quantum error correction — requiring active error correction, classical feedback, scalable codes and gates, and reduced dependence on error detection and the exponential overhead in shot counts associated with pure error detection — which have yet to be simultaneously realized. However, despite the limitations of the experiments discussed above, they have demonstrated remarkable accomplishments; The Harvard experiment using hypercube block IQP circuits demonstrated a family of a classically hard, but not universal circuits (the Quantinuum experiments are restricted to Clifford circuits), while Quantinuum's H2-1 platform is the only system in which error correction with below physical error rates has been demonstrated.

We are at an exciting time in the development of quantum computers. Error correction schemes can now be implemented with fidelities and scales which prevent efficient classical simulation, and the noise at the logical level can be explored. Furthermore, as systems continue to improve in gate fidelities, connectivity, and qubit counts, we will see full unambiguous demonstrations of logical circuits with logical error rates below physical error rates across larger families of circuits, ultimately targeting full universal circuits beyond IQP or Clifford circuit families.

9 Contributions

Eric Huang worked on section 2, Maryam Mudassar worked on sections 4 and 7, Suying Liu worked on section 5 and Thomas Steckmann worked on section 6. We worked on the presentation together as well as some editing of the document.

References

[1]D. Gottesman, *Quantum fault tolerance in small experiments*, Oct. 2016 (cit. on pp. 2, 6).

[2]R. Chao and B. W. Reichardt, "Fault-tolerant quantum computation with few qubits", npj Quantum Information **4**, 1–8 (2018) (cit. on p. 5).

[3]C. Ryan-Anderson et al., "High-fidelity and Fault-tolerant Teleportation of a Logical Qubit using Transversal Gates and Lattice Surgery on a Trapped-ion Quantum Computer", (2024) (cit. on pp. 5, 19).

[4]R. Acharya et al., *Suppressing quantum errors by scaling a surface code logical qubit*, 2022 (cit. on pp. 5, 7).

[5]M. P. da Silva et al., *Demonstration of logical qubits and repeated error correction with better-than-physical error rates*, Apr. 2024 (cit. on pp. 6, 14, 17, 18).

[6]F. Arute et al., "Quantum supremacy using a programmable superconducting processor", Nature **574**, 505–510 (2019) (cit. on p. 7).

[7]D. Bluvstein et al., "Logical quantum processor based on reconfigurable atom arrays", Nature **626**, 58–65 (2024) (cit. on pp. 9–14).

[8]H. Bernien et al., "Probing many-body dynamics on a 51-atom quantum simulator", Nature **551**, 579–584 (2017) (cit. on pp. 9, 10).

[9]K. Barnes et al., "Assembly and coherent control of a register of nuclear spin qubits", Nature Communications **13**, 2779 (2022) (cit. on p. 10).

[10]S. J. Evered et al., "High-fidelity parallel entangling gates on a neutral-atom quantum computer", Nature **622**, 268–272 (2023) (cit. on p. 13).

[11]A. e. a. Paetznick, To be published (2024) (cit. on pp. 15, 17).

[12]S. A. Moses et al., "A Race Track Trapped-Ion Quantum Processor", Physical Review X **13**, 041052 (2023) (cit. on p. 15).

[13]H. Goto, "Minimizing resource overheads for fault-tolerant preparation of encoded states of the Steane code", Scientific Reports **6**, 19578 (2016) (cit. on p. 15).

[14]C. Ryan-Anderson et al., "Realization of Real-Time Fault-Tolerant Quantum Error Correction", Physical Review X **11**, 041058 (2021) (cit. on p. 16).

[15]B. W. Reichardt, "Fault-tolerant quantum error correction for Steane's seven-qubit color code with few or no extra qubits", Quantum Science and Technology **6**, 015007 (2020) (cit. on pp. 16, 17).