Leakage Suppression in the Toric Code

Martin Suchara
IBM T. J. Watson Research Center
1101 Kitchawan Road
Yorktown Heights, New York 10598
Email: msuchar@us.ibm.com

Andrew W. Cross
IBM T. J. Watson Research Center
1101 Kitchawan Road
Yorktown Heights, New York 10598
Email: awcross@us.ibm.com

Jay M. Gambetta
IBM T. J. Watson Research Center
1101 Kitchawan Road
Yorktown Heights, New York 10598
Email: jay.gambetta@us.ibm.com

***Abstract*—Quantum codes excel at correcting local noise but fail to correct leakage faults that excite qubits to states outside the computational space. Aliferis and Terhal have shown that an accuracy threshold exists for leakage faults using gadgets called leakage reduction units (LRUs). However, these gadgets reduce the threshold and increase experimental complexity, and the costs have not been thoroughly understood. We explore a variety of techniques for leakage resilience in topological codes. Our contributions are threefold. First, we develop a leakage model that is physically motivated and efficient to simulate. Second, we use Monte-Carlo simulations to survey several syndrome extraction circuits. Third, given the capability to perform 3-outcome measurements, we present a dramatically improved syndrome processing algorithm. Our simulations show that simple circuits with one extra CNOT per qubit reduce the accuracy threshold by less than a factor of 4 when leakage and depolarizing noise rates are comparable compared to a scenario without leakage. This becomes a factor of 2 when the decoder uses 3-outcome measurements. Finally, we make the surprising observation that for physical error rates less than 2×10^{-4}, placing LRUs after every gate may achieve the lowest logical error rate. We expect that the ideas may generalize to other topological codes.**

I. Introduction

Large-scale quantum computers require a fault-tolerant architecture based on quantum error-correcting codes. Topological error-correcting codes [1], [2] stand out due to their favorable properties such as local check operators, simple syndrome extraction circuits [3], and flexible fault-tolerant logic based on transversal gates, code deformation [4], [5], or lattice surgery [6]. These properties endow topological codes with a high accuracy threshold, estimates of which vary from 0.67% [7] to above 1% [8].

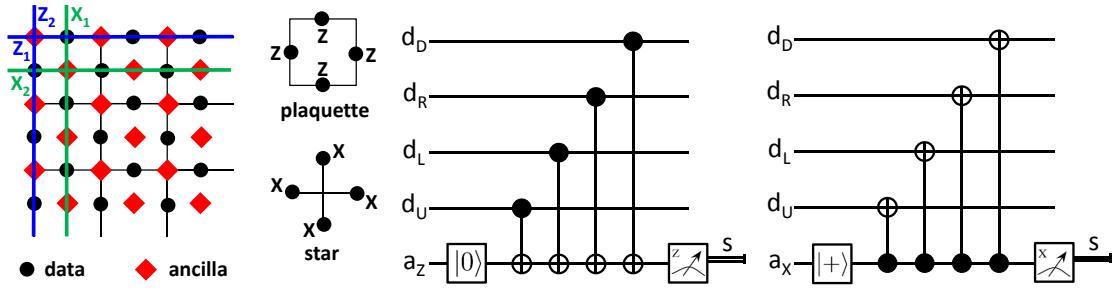
These threshold estimates assume a depolarizing noise model that approximates realistic noise. However, the model does not include so-called leakage faults that map quantum states out of the 2-dimensional qubit subspace and into a higher-dimensional Hilbert space, a behavior exhibited by many physical realizations of qubits. Leakage errors can be mitigated by constructing special quantum circuits that convert leakage errors into "regular" errors and may or may not simultaneously raise a flag to indicate the location of the leakage error. Gadgets that detect leakage convert it into a located loss error that is easier to correct [9]–[11]. Leakage reduction units (LRUs) convert leakage errors into regular errors but do not give an indication that leakage has occurred [12], [13]. Leakage mitigation with LRUs was rigorously analyzed by Aliferis and Terhal [12], who showed existence of a threshold to leakage in concatenated codes. Subsequent work has shown that circuits can be further simplified while remaining fault-tolerant [14].

Most prior studies of quantum codes do not consider errors due to leakage, and the the impact of leakage on the performance of error-correcting codes has not been thoroughly understood, particularly in the surface code. Topological codes are known to have a 50% threshold to idealized erasure errors [15], [16]. However, this idealized model is far from realistic. Moreover, the erasure correction strategy [15], [17] measures check operators that each act over a relatively large neighborhood, and this is challenging to do with a fixed planar array of qubits [18]. In less idealized leakage models, leakage errors do not correspond to simple erasures since they may remain undetected or are only imprecisely located. Leakage of this form has been studied for a quantum repetition code [19] and for specific types of gates applied to superconducting qubits [20], [21]. Despite this work, the impact of leakage on the accuracy threshold of topological codes remains unclear. To address this problem, we systematically study several leakage reducing circuits and develop a new syndrome processing strategy that further enhances the benefits of these circuits.

II. Toric Code

The toric code [1] is the prototypical example of topological stabilizer codes. It is defined on a d by d square array where d is the code distance. The left-right and top-bottom boundaries of the array are associated. Vertices of the array are connected to form a graph where each of the n edges carries a physical qubit, called a code or data qubit. Each vertex and each face carries a qubit used for error-correction, called a syndrome or ancilla qubit (see Fig. 1a). The stabilizer S of the code is generated by a set of check operators $\{A_v\}$ and $\{B_f\}$ that belong to the n-qubit Pauli group and are attached to each vertex v and face f of the graph. These operators are tensor products of single qubit Pauli operators X and Z. The vertex (or star) operators $A_v = \bigotimes_{\varepsilon \in N(v)} X_\varepsilon$ are X-type checks that apply an X to the qubits on the four edges $N(v)$ incident on vertex v. The face (or plaquette) operators $B_f = \bigotimes_{\varepsilon \in N(f)} Z_\varepsilon$ are Z-type checks that apply a Z to the qubits on the four edges $N(f)$ on the boundary of face f. The toric code encodes a pair of logical qubits. Representatives for each class of logical Pauli X and Z operators are shown in Fig. 1a.

The check operators are measured simultaneously [3] using circuits shown in Fig. 1b and Fig. 1c. In the first step, all plaquette ancillas are prepared in the $|0\rangle$ state and all site ancillas in the $|+\rangle$ state. In the next four steps, CNOT gates act between each ancilla and the data qubit above, left, right, and below the ancilla, in that order. This corresponds to the gate order used in [7], [8]. Finally, the plaquette ancillas are



(a) (b) (c)

Fig. 1. (a) The toric code has a natural two-dimensional layout on the surface of a torus. Qubits function as either data qubits, used to store the encoded quantum state, or ancilla qubits, used to measure check operators (stabilizers) of the quantum code. Z-type check operators associate to faces (plaquettes) and X-type check operators associate to vertices (stars). Each check operator involves four data qubits and is measured using an ancilla qubit. The torus encodes a pair of qubits whose representative logical Pauli operators X_1, Z_1, X_2, and Z_2 are shown. (b,c) These circuits measure (b) plaquette and (c) star operators using four CNOT gates together with preparations and measurements. In each circuit, data qubits $d_{D,R,L,U}$ interact with an ancilla qubit $a_{z/x}$ that is prepared in the state $|0\rangle$ or $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. This ancilla is then measured in the basis of eigenstates of Z or X, respectively.

measured in the Z eigenbasis and the star ancillas in the X eigenbasis. These six steps constitute an error-correction cycle. Since syndrome measurements may not be reliable, $O(d)$ error-correction cycles are performed to improve confidence in the syndrome. A syndrome history consists of these $O(d)$ syndrome measurement outcomes each with $2d^2$ bits.

The corrective operation is determined by a decoding algorithm that processes the syndrome history. The classic approach [3] processes plaquette and star syndromes independently to find bit and phase error corrections. A separate decoding graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is constructed [22] for the two cases. The graph is built by translating the unit cell shown in Fig. 2a. The edges $\mathcal{E}$ correspond to error events in the error-correction circuits. The horizontal edges b and d represent qubit errors, the vertical edge a represents measurement errors and the diagonal edges c, e, and f represent correlated errors. The weight of each edge approximates the negative logarithm of the probability of the corresponding error. Each nontrivial syndrome in the syndrome history introduces a defect in the decoding graph. A corrective operation that is likely to succeed is obtained by matching pairs of defects using the minimum weight matching algorithm [23] and correcting the bit or phase flip errors on chains of qubits between the matched pairs of defects.

III. Leakage Errors

Leakage errors occur because qubits are not idealized two-level systems but instead occupy a larger physical Hilbert space. Interactions with the system may populate states in that larger physical space. Here we define a simple stochastic leakage model that captures key aspects of leakage and yet remains amenable to stabilizer simulation. An interested reader may find justification of our choices in [24].

Our stochastic model assumes that each subsystem is a 3-level system (a qutrit). Our model can be simulated by maintaining and propagating a label (I, X, Y, Z or L) for each qubit, where "L" denotes that the qubit leaked. A key benefit of our label-based model is that full-blown stabilizer simulation is not needed for the study of fault-tolerant quantum error-correction and it suffices to track the Pauli error for each qubit. Discrete leakage events occur independently with probability $p_\uparrow$ on each output qutrit of an ideal gate. There is a relaxation process with probability $p_\downarrow$, analogous to amplitude damping, that acts independently on each output qutrit of an ideal gate. Another important aspect of our model concerns two-qubit gates. If one of the interacting qubits is leaked, the other qubit is depolarized. Noisy two-qubit gates are followed by independent excitation and relaxation maps on each output qutrit.

The model has parameters p (depolarizing error probability), q (syndrome measurement error probability, we set $q = p$), $p_\uparrow$ (probability of excitation outside of the computational space), and $p_\downarrow$ (probability of relaxation back to the computational space). For convenience, we also define the relative excitation rate $r = p_\uparrow/p$ and relaxation rate $s = p_\downarrow/p$. Elementary quantum gates behave as follows. Idle qubits depolarize with probability p, where we apply one of X, Y, or Z uniformly at random. They undergo free evolution and do not leak. However, an idle leaked qubit may relax to the computational space with probability $p_\downarrow$. State preparation succeeds with probability $1-p$, otherwise the orthogonal state is prepared. With probability $p_\uparrow$ the prepared qubit leaks. If the input qubit is in the computational space, measurements report the incorrect outcome with probability $q = p$ and are otherwise correct. When the input qubit is leaked, we consider two scenarios: either the measurement cannot distinguish higher levels and reports "1", or the measurement has a third outcome "L". Noisy CNOT gates suffer joint depolarizing noise with probability p, where a non-identity two-qubit Pauli error occurs uniformly at random. If one of the inputs to the CNOT is leaked, the other non-leaked qubit is completely depolarized, i.e. one of the 4 single-qubit Pauli operators is applied uniformly to the non-leaked qubit. Finally, each output qubit leaks with probability $p_\uparrow$ and relaxes with probability $p_\downarrow$.

Unlike the model in [19], we assume that qubits reach a steady state leakage distribution at the beginning of any simulation, which more faithfully models conditions of a long computation. Consider error-correction circuits in the standard toric code (Fig. 1b and Fig. 1c). Unlike ancillas, data qubits are never re-initialized, and therefore they gradually accumulate leakage according to the transition probabilities $p_\uparrow$ and $p_\downarrow$ until equilibrium is reached. It is easy to show from direct calculation of eigenstates of the transition matrix that the equilibrium

distribution for data qubits is given by $p_{eq} \approx \frac{4p_\uparrow}{4p_\uparrow+6p_\downarrow}$ where the factor 4 is due to the four $CNOT$ gates, and the factor 6 is due to the four $CNOT$s and two idle time steps. The second eigenvalue of the transition matrix is $(1-p_\downarrow)^6(1-p_\uparrow)^4$, so the non-equilibrium component for each qubit decays as $\exp(-\zeta n)$ where $\zeta = -6\ln[(1-p_\downarrow)(1-p_\uparrow)]$. This suggests that starting from an equilibrium distribution is not overly pessimistic. Under the reasonable assumption $p_\uparrow \approx p_\downarrow$, the fraction of leaked data qubits reaches 40% at which point the threshold is far exceeded and computation without employing some leakage reduction technique is impossible.

IV. Leakage Reduction

Leakage reduction is a process that converts leakage errors into regular errors that can be corrected by quantum codes. Prompt leakage reduction is desirable as a single leaked qubit can cause many other errors (e.g., a leaked data qubit can damage ancillas which then spread more errors to other data qubits). A circuit that converts leakage on a single qubit into a regular error is called a leakage reduction unit (LRU). LRUs satisfy two properties [12]: (a) they apply identity to unleaked inputs, and (b) they output some computational state if the input is leaked. Fig. 3 shows an LRU based on one-bit teleportation [25]. We use LRUs as building blocks to construct several leakage suppressing circuits that we describe next. Our discussion here omits some technical details that can be found in [24].

A. Leakage Reduction without Leakage Detection

We build several circuits that represent a tradeoff between circuit complexity and leakage reduction effectiveness. We list circuits from the most complicated to the simplest one:

Full-LRU: appends a LRU after each gate occurrence in the syndrome extraction circuit (as in [12]); see Fig 4a. Leakage is removed immediately after each gate but many additional gates and qubits are needed (16 additional gates in 4 LRUs per data and ancilla qubit per cycle). No LRU is applied to ancillas prior to measurements.

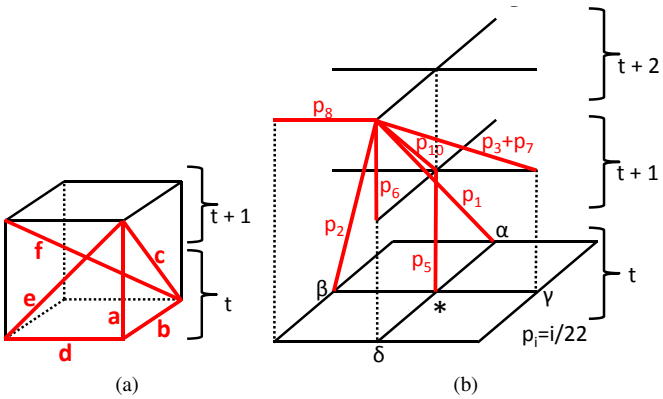


Fig. 2. (a) The unit cell of the Standard decoding graph has six edges with distinct weights corresponding to the probability of the underlying error. The edges that appear are determined by the error-correction circuit. (b) The X error decoding graph of the HL decoder for the Quick circuit. Upon leakage detection at time t on the plaquette ancilla *, the weights of the depicted edges are replaced by probability p_i.

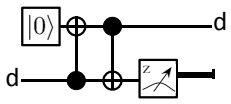


Fig. 3. A leakage reduction unit (LRU) implemented by one-bit teleportation. If the original qubit at the bottom is in the computational space, the data is teleported to the new qubit at the top. If the original qubit is leaked, the new qubit is depolarized but remains in the computational space.

Partial-LRU: uses LRUs only once per error correction cycle. LRUs act on each data qubit while the ancillas are measured; see Fig. 4b. This strategy removes leakage less frequently but only uses 4 additional gates (in 1 LRU) per data qubit per cycle. Ancillas have no LRUs as they are periodically measured and reinitialized.

Quick: is a circuit that further reduces the frequency of leakage reduction; see Fig. 4c. At the end of each cycle, each data qubit is swapped with an ancilla. A similar construction in [21] reduced leakage in a superconducting qubit model. We choose to swap the ancilla with the data qubit d_D immediately below the ancilla. Three $CNOT$ gates implement a $SWAP$ and, due to gate cancellation, there is only one additional $CNOT$ gate per data qubit compared to the standard circuit. Each physical qubit is measured and reset every other cycle, so qubits do not remain leaked for many cycles. A key benefit is simplicity and no need for extra ancillas.

No LRU: is the standard circuit depicted in Fig. 1b, 1c. The circuit is not suitable for error correction in the presence of leakage but serves as a reference point.

The *Standard decoder* is an algorithm that processes syndromes produced by leakage reducing circuits. For each circuit we adjust the edge weights in the decoding graph (Fig. 2a) so that the decoder based on minimum weight matching finds a correction from among the most likely errors consistent with the syndrome.

The prior probability of edge ε is $p(\varepsilon) = \sum_j p_j(\varepsilon)$ where $p_j(\varepsilon)$ is the probability that a fault occurs at location j and yields syndrome $\partial\{\varepsilon\}$. The sum is taken over all locations in the circuit. For small error rates $p(\varepsilon)$ approximates the probability that ε carries an error. The edge weights are then given by $-\log p(\varepsilon)$. Edge weights can be approximated by counting single-location faults [8]. We list numeric values of the edge weights in [24].

B. Heralded Leakage Reduction

So far we assumed that measurements cannot distinguish a leaked qubit from a qubit in the state $|1\rangle$. We now assume that measurements can produce a third outcome "L". For simplicity we also assume that "L" is output if and only if the measured qubit is leaked. We modify the decoders to use this additional information. This requires no change to the quantum circuits, but leads to significantly improved error decoding.

The *Heralded Leakage (HL) decoder* ingests results from 3-outcome measurements. This is advantageous because leakage is likely to cause regular errors on specific locations around the leaked qubit, and the decoder is then able to match defects using correct prior probabilities in the space-time neighborhood of the "L" event in the decoding graph. Such

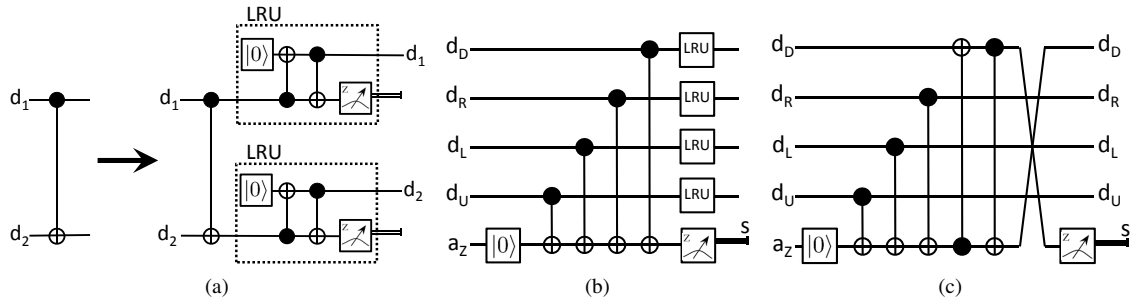


Fig. 4. (a) The Full-LRU circuit replaces each gate with the gate followed by LRUs on each output qubit. (b) In the partial-LRU scheme, we insert an LRU on each data qubit at the end of each error-correction cycle. (c) The Quick scheme uses a circuit that is equivalent to those in Fig. 1. The final CNOT gate acting on the ancilla and the d_U data qubit has been replaced by a CNOT followed by a SWAP. This simplifies to a pair of CNOT gates.

events do not indicate exactly which gate suffered a leakage fault, since "L" events can only be detected by measuring a qubit, but the decoder learns that the qubit leaked at some location between its initialization and measurement.

Conditioned on the observed "L" events, the HL decoder generates a new decoding graph by modifying the Standard decoding graph (shown in Fig. 2a). Suppose a set of these events $\mathcal{L}$ occurs. We can construct a conditional decoding graph $\mathcal{G}_{\mathcal{L}}$ whose edges ε again correspond to independent faults and whose edge weights are computed from conditional probabilities $p(\varepsilon|\mathcal{L})$. These probabilities are computed similarly as in Section IV-A. $\mathcal{G}_{\mathcal{L}}$ is not translation invariant and contains additional edges not present in Fig. 2a. There will be low-weight edges in $\mathcal{G}_{\mathcal{L}}$ associated to highly probable errors on the qubits that interacted with leaked qubits. The added edges and their weights are specific to each leakage reduction circuit. Fig. 2b shows example of the additional edges added to the X error decoding graph for the Quick circuit.

To compute the edge weights in the conditional decoding graph, we proceed as follows, examining each "L" event on qubit q independently. We consider each of the n fault locations between and including q's initialization and measurement. The probability that q is leaked when interacting at the i-th location is approximately i/n. An X and Z error each occurs on the qubit that interacts with q at location i with probability $p_i = i/2n$. We consider each of these errors separately, find the defect pair they cause, and modify the weight of the edge connecting this defect pair.

V. Simulation Results

We simulated the No LRU, Partial-LRU, Full-LRU, and Quick leakage reduction strategies with the Standard and HL decoder. Fault-tolerant error-correction at code distance d requires $O(d)$ cycles of syndrome measurements so we simulate d cycles. Each qubit is initially assigned the L label with probability equal to the equilibrium leakage rate for that particular qubit. Results were generated by Monte Carlo simulations repeated for at least $10,000$ iterations and at least $1,000$ failures for each configuration. We used about 30,000 CPU-hours on an IBM Blue Gene/Q supercomputer.

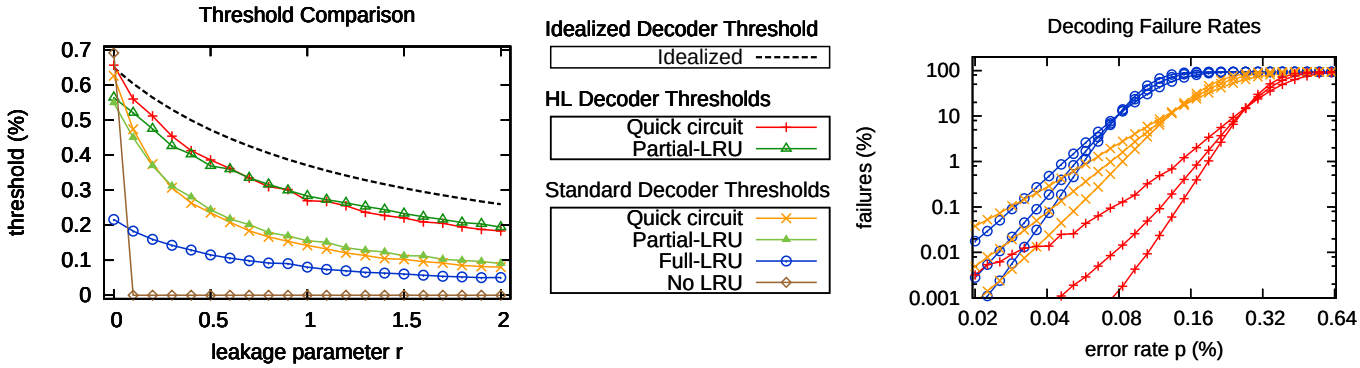
Fig. 5a explores the accuracy thresholds for the Standard and HL decoders as a function of the amount of leakage (the relative excitation rate r). The plot was obtained by recording the failure rates for code distance $d = 7$ and $d = 9$ and by varying r in increments of 0.1 and p in increments of 0.01%. We chose a relative relaxation rate $s = 1$. For each parameter choice we found the crossover of the failure rates for the two code distances, which approximates the threshold.

Fig. 5a shows that in the regime with no leakage ($r = 0$), simpler syndrome extraction circuits have higher thresholds. For example, the No LRU circuit has a threshold of about 0.70% whereas the Full-LRU circuit has a threshold of about 0.22%. The Quick and Partial-LRU circuits give an almost identical threshold. The Quick circuit benefits from having fewer gates and qubits, whereas the Partial-LRU benefits from more frequent qubit reinitializations and therefore more effective leakage suppression. Finally, the HL decoder significantly improves the threshold compared to the Standard decoder.

To understand the decrease of the threshold with increasing leakage in Fig. 5a we consider an idealized scenario where all leakage errors are immediately corrected. Leaked qubits are replaced with a depolarized qubit, causing error X, Y, Z, or I, which is equivalent to a regular error with probability $\beta = 3/4$. Suppose that physical errors occur at an effective rate $\tilde{p} = (1 + \beta r)p$ where the first term is due to regular depolarizing errors and the second one due to leakage. After some algebra we can show that the threshold is approximately $\frac{\alpha}{1+\beta r}$ where α is a constant equal to the numeric value of the threshold without leakage. We plot the threshold of the "idealized decoder" in Fig. 5a by assuming $\beta = 3/4$ and choosing α such that we intersect the threshold of the Quick circuit at $r = 0$. We expect that any leakage reducing circuit will need at least one additional physical gate like the Quick circuit. A Standard decoder's threshold will not exceed that of the idealized one, but a decoder that uses results from 3-outcome measurements could.

We observed [24] that the thresholds do not depend much on the relaxation rate s, likely because relaxation is much slower than the leakage reduction occurring in the circuits.

To study the effectiveness of the decoders, we chose $r = s = 1$ and recorded the decoding failure rates as a function of p. Recall that $p_\uparrow = rp$ and $p_\downarrow = sp$. In Fig. 5b we show failure rates for the Full-LRU circuit and the Quick circuit with the Standard and HL Decoder. The thresholds occur at the crossing points of each set of curves. When p is well below threshold, the success rate of the Full-LRU circuit improves faster than for the Quick circuit with Standard decoder, which suggests that for low enough physical error rates the Full-LRU circuit will perform best.



(a) (b)

Fig. 5. (a) Summary of thresholds for the HL and Standard decoders. The HL decoder offers a substantially improved threshold. (b) Comparisons of decoding failure probabilities for the Full-LRU circuit and the Quick circuit with the Standard and HL decoder.

VI. Conclusion

In this work we defined a simple noise model that captures key aspects of leakage and is tractable to simulate. We used the model to systematically survey the performance of various circuits that suppress leakage in the toric code. We considered a scenario where measurements can distinguish leaked and non-leaked qubits, and another in which they cannot. For each of these cases we designed an optimized decoding algorithm. Our simulation results show that the small, simple circuits effectively suppress leakage and reduce the accuracy threshold by less than a factor of 2 with our decoder that uses information about leakage detection events. We also present evidence that large complex circuits that reduce leakage after every gate may achieve the lowest logical error rates for physical error rates below 2×10^{-4} despite having low accuracy threshold. Future work could explore logical error rates of these circuits at higher distances and at error rates far below threshold, perhaps by applying the splitting method [22].

Acknowledgment

We thank Sergey Bravyi, Ken Brown, Oliver Dial, David DiVincenzo, Easwar Magesen, Graeme Smith, and John Smolin for helpful discussions, Chris Lirakis and Mark Ritter for their support, and Karen Bard for making Blue Gene available to us. AWC and JMG acknowledge support from ARO under contract *W911NF-14-1-0124*.

References

[1] A. Kitaev, "Fault-tolerant quantum computation by anyons," *Annals Phys.*, vol. 303, pp. 2–30, 2003.

[2] S. Bravyi and A. Kitaev, "Quantum codes on a lattice with boundary," *arXiv:quant-ph/9811052*, 1998.

[3] E. Dennis, A. Kitaev, A. Landahl, and J. Preskill, "Topological quantum memory," *J. Math. Phys.*, vol. 43, pp. 4452–4505, 2002.

[4] R. Raussendorf and J. Harrington, "Fault-tolerant quantum computation with high threshold in two dimensions," *Phys. Rev. Lett.*, vol. 98, no. 190504, 2007.

[5] R. Raussendorf, J. Harrington, and K. Goyal, "Topological fault-tolerance in cluster state quantum computation," *New. J. Phys.*, vol. 9, pp. 199–219, 2007.

[6] C. Horsman, A. Fowler, S. Devitt, and R. V. Meter, "Surface code quantum computing by lattice surgery," *New J. Phys.*, vol. 14, no. 123011, 2012.

[7] A. M. Stephens, "Fault-tolerant thresholds for quantum error correction with the surface code," *Phys. Rev. A*, vol. 89, no. 022321, 2014.

[8] D. S. Wang, A. G. Fowler, and L. C. Hollenberg, "Surface code quantum computing with error rates over 1%," *Phys. Rev. A*, vol. 83, no. 020302, 2011.

[9] D. Gottesman, "Stabilizer codes and quantum error correction," *Caltech Ph.D. thesis*, 1997.

[10] J. Preskill, "Reliable quantum computers," *Proc. Roy. Soc. Lond.*, vol. A454, pp. 385–410, 1998.

[11] E. Knill, R. Laflamme, and W. Zurek, "Resilient quantum computation: error models and thresholds," *Proc. Roy. Soc. A*, vol. 454, no. 1969, pp. 365–384, 1998.

[12] P. Aliferis and B. M. Terhal, "Fault-tolerant quantum computation for local leakage faults," *Quant. Inf. Comp.*, vol. 7, no. 1, pp. 139–156, 2007.

[13] B. Fong and S. Wandzura, "Universal quantum computation and leakage reduction in the 3-qubit decoherence free subsystem," *Quant. Inf. Comp.*, vol. 11, no. 11–12, pp. 1003–1018, 2011.

[14] B. Fortescue, S. Nawaf, and M. Byrd, "Fault-tolerance against loss for photonic FTQEC," *arXiv:1405.1766*, 2014.

[15] T. M. Stace, S. D. Barrett, and A. C. Doherty, "Thresholds for topological codes in the presence of loss," *Phys. Rev. Lett.*, vol. 102, no. 200501, 2009.

[16] K. Fujii and Y. Tokunaga, "Error and loss tolerances of surface codes with general lattice structures," *Phys. Rev. A*, vol. 86, no. 020303, 2012.

[17] S. D. Barrett and T. M. Stace, "Fault tolerant quantum computation with very high threshold for loss errors," *Phys. Rev. Lett.*, vol. 105, no. 200502, 2010.

[18] J. Chow *et al.*, "Implementing a strand of a scalable fault-tolerant quantum computing fabric," *Nature Comm.*, vol. 5, no. 4015, 2014.

[19] A. G. Fowler, "Coping with qubit leakage in topological codes," *Phys. Rev. A*, vol. 88, no. 042308, 2013.

[20] J. Ghosh, A. G. Fowler, J. M. Martinis, and M. R. Geller, "Understanding the effects of leakage in superconducting quantum-error-detection circuits," *Phys. Rev. A*, vol. 88, no. 062329, 2013.

[21] J. Ghosh and A. G. Fowler, "A leakage-resilient scheme for the measurement of stabilizer operators in superconducting quantum circuits," *arXiv:1406.2404 [quant-ph]*, 2014.

[22] S. Bravyi and A. Vargo, "Simulation of rare events in quantum error correction," *Phys. Rev. A*, vol. 88, no. 062308, 2013.

[23] J. Edmonds, "Paths, trees, and flowers," *Canad. J. Math*, vol. 17, pp. 449–467, 1965.

[24] M. Suchara, A. Cross, and J. Gambetta, "Leakage suppression in the toric code," *arXiv:1410.8562*, 2014.

[25] X. Zhou, D. Leung, and I. Chuang, "Methodology for quantum logic gate constructions," *Phys. Rev. A*, vol. 62, no. 052316, 2000.