Error correction phase transition in noisy random quantum circuits

Jon Nelson*,[1,2] Joel Rajakumar*,[1,2] and Michael J. Gullans[1,2]

[1] *Joint Center for Quantum Information & Computer Science, University of Maryland and NIST*
[2] *Department of Computer Science, University of Maryland**

In this work, we study the task of encoding logical information via a noisy quantum circuit. It is known that at superlogarithmic-depth, the output of any noisy circuit without reset gates or intermediate measurements becomes indistinguishable from the maximally mixed state, implying that all input information is destroyed. This raises the question of whether there is a low-depth regime where information is preserved as it is encoded into an error-correcting codespace by the circuit. When considering noisy random encoding circuits, our numerical simulations show that there is a sharp phase transition at a critical depth of order p^{-1}, where p is the noise rate, such that below this depth threshold quantum information is preserved, whereas after this threshold it is lost. Furthermore, we rigorously prove that this is the best achievable trade-off between depth and noise rate for any noisy circuit encoding a constant rate of information. Thus, random circuits are optimal noisy encoders in this sense.

I. INTRODUCTION

Random quantum circuits are known to exhibit a variety of phase transitions when subjected to intermediate nonunitary operations such as measurement or noise. The main feature that many of these effects have in common is that when the rate of these nonunitary operations exceeds a critical threshold, the circuit is no longer able to prevent quantum information from escaping. A prominent example is the measurement-induced phase transition in monitored random circuits [1–3], which has been directly linked to a transition in the ability of the circuit to retain memory of its input state [4]. Similarly, Weinstein et al. [5] showed that randomly swapping qubits into the environment leads to a phase transition at which point the initial state's information becomes entirely lost to the environment. In a related study, Qian et al. [6] identified a phase transition in coherent information driven by the interplay between single-qubit noise and the rate at which fresh ancilla qubits are introduced.

Most of these works focus on phase transitions that are independent of the circuit depth. In contrast, we study random circuits subject to mid-circuit noise, where the effects of noise compound with increasing depth. This cumulative behavior gives rise to a rich phase diagram that depends jointly on noise strength and circuit depth, which we fully characterize through numerical simulations and matching analytical bounds.

For sufficiently low depth, [7] designed a protocol that recovers logical information from the output of low-depth noisy random quantum circuits. However, it is known that at $\omega(\log n)$ [8] depth, the output of any noisy circuit converges to the maximally mixed state implying that all input information is destroyed [9, 10]. This suggests that as the depth is scaled up there must exist some transition between a regime where the input quantum state is recoverable to a regime where it is destroyed by the accumulation of noise. We numerically study this question for the case of circuits with gates drawn from either the Clifford or Haar ensemble and discover that a phase transition does indeed exist at a critical depth of $d^* = \Theta(1/p)$, where p is the noise strength. Our numerics show that below this depth threshold, the majority of the logical information is still preserved by the circuit, but after this threshold the majority of this information is lost to the environment. This expands upon the scope of [5], which shows a separate depth-independent critical noise rate that lies above ours, beyond which *all* of the logical information is lost in finite depth.

Furthermore, we analytically prove that our observed above-threshold behavior holds for *any* noisy circuit and not just the random ones we consider in our numerics. In particular, we prove that for $d > d^*$ the coherent information of the noisy circuit is upper bounded strictly below zero. Our numerics thus imply that random circuits are the best possible noisy encoders since they can preserve information all the way up to this depth upper bound [11].

The optimality of random circuits for protecting information is not unprecedented. Random encoding circuits have shown great promise given their desirable coding parameters [12–15]. For example, Hayden *et al.* demonstrated that encoding with Haar random unitaries achieves the quantum channel capacity [15]. Furthermore, Gullans *et al.* numerically showed that even $O(\log n)$-depth random Clifford circuits are sufficient to saturate the capacity of the erasure channel [14]. Intuitively, these random encoding circuits achieve such high performance due to their ability to scramble quantum information quickly [16]. However, these results often assume that the logical information can be perfectly encoded before noise can act on the state. In reality, noise can act throughout the encoding circuit, which can destroy logical information before it is fully encoded. The challenge with this more realistic setting is that the fast-scrambling nature of random circuits not only encodes quantum information more quickly but also spreads er-

* These authors contributed equally to this work.
email: nelson1@umd.edu, jrajakum@umd.edu

rors faster throughout the circuit. At first glance, it is unclear which of these two competing effects wins out. Our results clarify this picture.

II. SUMMARY OF RESULTS

In this work, we consider circuits that act on n qubits where k are considered to be logical qubits and the rest are ancilla qubits initialized to $|0\rangle^{n-k}$. The encoding rate is denoted by $r := k/n$. We first study brickwork circuits where 2-qubit Clifford gates are drawn randomly and applied to qubits on a 2D lattice. After each layer of gates, a round of depolarizing noise is applied. In order to characterize the ability to recover the original logical information, we then allow for a perfect syndrome measurement after the noisy circuit is applied. We numerically estimate the coherent information of this channel for various depths and noise rates, where the coherent information characterizes how much of the logical information is recoverable at the output of the channel and is a lower bound on the single-shot quantum channel capacity [17, 18]. Using an encoding rate of $r = 1/8$, our results show that when pd is larger than approximately 0.6, the coherent information of the channel approaches its minimum as the system size increases. On the other hand, when the product of depth and noise strength is below this threshold, then the coherent information instead approaches a finite positive value implying that the original logical information is partially recoverable.

Next, we consider a family of circuits where 2-qubit gates are chosen from the Haar measure and applied to disjoint random pairs of qubits in each layer. Then, a round of either depolarizing errors or amplitude damping errors are applied. As in the previous case, we are also interested in the coherent information of this channel as a function of noise strength and depth. However, Haar random circuits are not tractable to classically simulate so we must take a different approach to conduct the numerics. For this, we consider a standard approximation to the coherent information, which Gullans *et al.* refer to as the log-average purity [14]. Although this is a heuristic approximation, it is expected that the fluctuations in the coherent information over the random circuits is small enough such that the log-average purity is a good approximation [19–21]. This quantity can be computed using standard calculations using the second moment operator of the Haar measure. Our numerics once again show a phase transition at a critical depth of $d^* = O(1/p)$.

Finally, we prove a rigorous upper bound on the coherent information for any noisy circuit (any gateset on any architecture). Informally, we prove the two following statements. First, we show that when $d > p^{-1}\log(1+1/r)$ then the coherent information, I_c, is strictly less than zero implying that the logical information is far from recoverable by the noisy circuit. Letting r be a constant this bound shows that no noisy circuit can preserve quantum information for $d \geq d^*$ where $d^* = O(1/p)$, giving analytical justification for the above-threshold behavior observed in our numerics. Second, we show that if $r > (2e^p - 1)^{-1}$, then the coherent information is also bounded below zero. This gives an upper bound on how high the encoding rate can be before it is impossible to recover the information after applying the noisy channel.

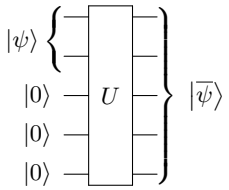


FIG. 1. Depiction of a general encoding circuit. In this case $|\psi\rangle$ is the k-qubit unencoded logical state. $|\overline{\psi}\rangle$ is the n-qubit encoded logical state.

III. BACKGROUND

A quantum error-correcting code maps a smaller 2^k-dimensional logical Hilbert space into a larger 2^n-dimensional physical Hilbert space. Any unitary on n qubits can be interpreted as such a mapping by allowing k of the n input qubits to represent the logical state and treating the rest of the input qubits as ancillas. This map is depicted in fig. 1. For instance, if the encoding unitary U is a Haar random unitary, then the codespace is a Haar random subspace.

When these encoding circuits are subject to mid-circuit noise, the logical information can be partially destroyed. To quantify the error-protecting capability of these encoding circuits under mid-circuit noise, we study their coherent information. We closely follow [22] and adopt their notation for consistency. We first define the encoding channel, which maps states in the logical Hilbert space $\mathcal{H}_A$ to states in the physical Hilbert space $\mathcal{H}_B$, as follows:

$$\mathcal{E}^{A\to B}(\rho) = \tilde{\mathcal{C}}(\rho \otimes |0\rangle\langle 0|^{\otimes n-k}) \tag{1}$$

where $\tilde{\mathcal{C}}$ denotes the noisy circuit. To evaluate the noisy circuit's ability to transmit logical information, we would like to understand whether it is feasible to recover the original logical codestate–in particular, does there exist a decoding channel $\mathcal{D}^{B\to A}$ such that for any $|\psi\rangle$, $\mathcal{D}\circ\mathcal{E}(|\psi\rangle\langle\psi|) = |\psi\rangle\langle\psi|$. This can be quantified by first introducing a reference system R and considering the maximally entangled state between R and the logical qubits A denoted as follows

$$|\Phi\rangle_{RA} = 2^{-k/2}\sum_{i=1}^{2^k}|i,i\rangle\,, \tag{2}$$

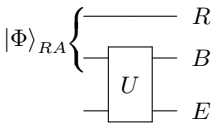


FIG. 2. Diagram for the Stinespring dilation of the noisy circuit acting on one half of the maximally entangled state. Importantly, the coherent information of this operation characterizes how well quantum information can be recovered from the output of the noisy circuit.

where 2^k is the dimension of $\mathcal{H}_A$. For notational purposes we also denote the density matrix of the maximally entangled state as $\Phi_{RA} := |\Phi_{RA}\rangle\langle\Phi_{RA}|$. The quantity of interest is then the coherent information of the noisy encoding channel $\mathcal{E}$ acting on the logical qubits of this maximally entangled state, which we denote as $I_c(\mathbb{I}/2^k, \mathcal{E})$, where $\mathbb{I}/2^k$ is the reduced state of Φ_{RA} on A. Before formally defining the coherent information, it will be helpful to consider the Stinespring dilation of $\mathcal{E}^{A\to B}$ and denote it as $U^{A\to BE}$ where E represents the environment. Finally, the output state of this dilation is denoted by $|\phi\rangle_{RBE}$. This purified model of our setting is depicted in fig. 2. With this setup in place, we can now define the coherent information in terms of entropies of reduced density matrices of $|\phi\rangle\langle\phi|_{RBE}$:

$$I_c(\mathbb{I}/2^k, \mathcal{E}) = S(B) - S(E), \tag{3}$$

which we abbreviate by I_c.

First, note that this quantity is essentially equivalent to the quantum mutual information between R and B with respect to $|\phi\rangle\langle\phi|_{RBE}$:

$$\begin{aligned} S(B) - S(E) &= S(B) - S(RB) && (4)\\ &= I(R:B) - S(R) && (5)\\ &= I(R:B) - k && (6)\end{aligned}$$

Therefore, it is exactly the mutual information between the reference system R and the physical qubits B, shifted by k. Since $0 \le I(R:B) \le 2k$, we have that $-k \le I_c \le k$. It can be shown that if the upper bound is saturated then any logical codestate can be perfectly recovered from the noisy channel. The converse is also true in that if the purification of the maximally entangled state can be recovered, then $I_c = k$. On the other hand, it can also be shown that when $I_c = -k$ no logical information can be transmitted. Thus, the coherent information is a useful quantity for understanding how well logical information can be preserved under the noisy encoding channel. For self-containment, we present a proof sketch for these statements in section A 1, which continues to follow John Preskill's notes [22].

Another motivation for studying the coherent information is that it is closely related to the quantum channel capacity by the formula [17, 18]:

$$\mathcal{Q}(\mathcal{N}) = \lim_{N\to\infty} \frac{1}{N} \max_{\rho} I_c(\rho, \mathcal{N}^{\otimes N}). \tag{7}$$

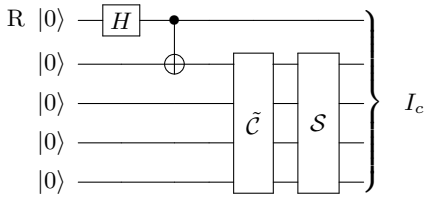


FIG. 3. Diagram of our numerical simulation. Qubit 1 represents the reference qubit and is prepared in a Bell pair with qubit 2, which represents the logical qubit (here, there is only one logical qubit and so $k = 1$). Qubits 3-5 represent the ancilla, or stabilizer, qubits. The noisy Clifford circuit is applied to the data qubits (qubits 2-5) which is denoted by $\tilde{\mathcal{C}}$. Next, the stabilizer generators of the code are measured perfectly which is denoted by $\mathcal{S}$. Finally, the coherent information of this channel is calculated.

IV. 2D RANDOM CLIFFORD CIRCUITS

We begin by studying random 2D Clifford circuits with mid-circuit depolarizing noise after every gate, where the single-qubit depolarizing channel is defined as follows:

$$\mathcal{D}_p(\rho) = (1-p)\rho + \frac{p}{3}X\rho X^\dagger + \frac{p}{3}Y\rho Y^\dagger + \frac{p}{3}Z\rho Z^\dagger \tag{8}$$

It can be seen that these 2D Clifford circuits encode input states into stabilizer codes as follows. First, we refer to the ancilla qubits as "stabilizer input qubits". Since these are initialized to $|0\rangle^{\otimes n-k}$ they are stabilized by the generators Z_i where i indexes each input qubit. Assuming there is no noise, the output of the circuit will be stabilized by $UZ_iU^\dagger$ where U represents the Clifford circuit. Thus, the stabilizer code is defined by the generators $UZ_iU^\dagger$ for each stabilizer input qubit i.

a. Computational Method. As discussed in the previous section, our goal will be to estimate the average coherent information of noisy random Clifford circuits acting on half of the maximally entangled state. For the case of Clifford gates, we can calculate this efficiently since the maximally entangled state is a stabilizer state (it is the tensor product of k Bell pairs which can each be prepared using Hadamards and CNOT gates), and so the entire procedure can be simulated using the Gottesman-Knill theorem extended to mixed states [23, 24]. After applying the noisy circuit we additionally allow for a perfect syndrome measurement of the stabilizer generators, which projects the state back onto the codespace. Finally, the coherent information is calculated by computing the entropy of the output state of this protocol. An example of this entire procedure is depicted in fig. 3.

b. Numerical Setup. For each depth between depth 10 and 24, the simulation is repeated while sweeping over a wide range of error rates. In addition, the system size is varied as $n = 16, 24, 144, 256$, and 400. Furthermore, a code rate of $1/8$ is chosen, and logical qubits are placed evenly throughout the 2D grid. For each choice of depth, error rate, and system size, we collect 1000 samples and compute the average coherent information.

c. Results. It is evident from fig. 4a that there is a clear crossing point where the coherent information for each system size all converge to zero at a critical error rate. In particular, it can be seen that before this critical error rate, the coherent information is positive and continues to increase with larger system size. Therefore, this parameter regime is in a phase in which error correction is possible. However, after the critical error rate, the coherent information is negative and increasing the system size leads to even lower coherent information. This indicates that the circuit is now in a phase in which error correction is no longer possible. To further investigate this phase transition, a scaling collapse is plotted which uses the following dimensionless scaling ansatz, whose form is motivated by the statistical mechanics of phase transitions [25]:

$$I_c = f(n^{\lambda}(p - p^*)) \tag{9}$$

We can estimate the critical exponent λ and the critical error rate p^* by approximating $f(x)$ as $f(x) = A + Bx + Cx^2$ using its Taylor expansion around $x = 0$. Fitting these five parameters, A, B, C, p^*, and λ, to the data results in an estimate for p^* and λ, which can then be used to plot I_c with respect to our dimensionless ansatz. From fig. 4b it is apparent that our scaling ansatz faithfully parameterizes the behavior of I_c since all of the data points fall on a single straight line regardless of the system size.

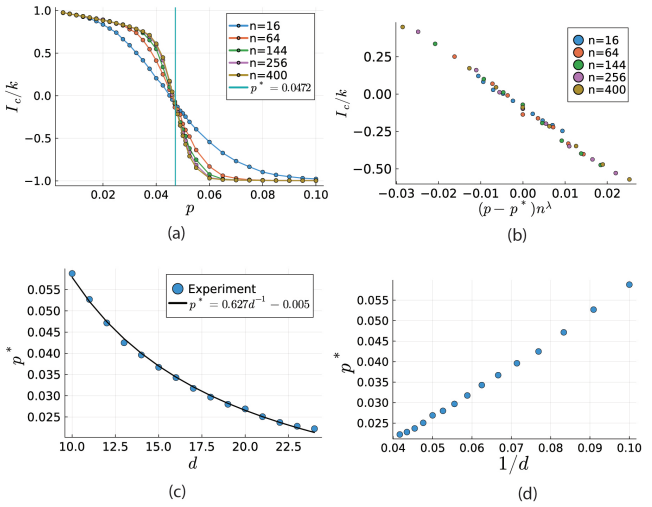


FIG. 4. Numerical results for 2D random brickwork Clifford circuits with periodic boundary conditions subjected to mid-circuit depolarizing noise. In all cases an encoding rate of $r = 1/8$ is used. (a) Coherent information per logical qubit is displayed with respect to a range of error rates p between 0.005 and 0.1. System size is scaled from $n = 16$ and $n = 400$. The circuit depth is fixed at $d = 12$. The critical error rate is depicted by the vertical line which occurs at $p^* = 0.0472$. (b) A scaling collapse for the data in (a) is plotted for error rates near the critical point. The scaling ansatz uses the estimated critical error rate $p^* = 0.0472$ and critical exponent $\lambda = 0.616$. (c) The plots in (a) and (b) are repeated for circuit depths d ranging from 10 to 24. Linear regression on the inverse depth is used to estimate the relationship that $p^* = 0.627d^{-1} - 0.005$. (d) Finally, the critical error rate p^* is plotted with respect to the inverse depth.

V. ALL-TO-ALL HAAR RANDOM CIRCUITS

Next, we conduct similar numerics for the case of Haar random circuits with all-to-all connectivity. This is not feasible by direct simulation but can be done using a standard second moment approximation [19–21].

a. Computational Method. In the definition of coherent information, we first replace the von Neumann entropy with the Renyi-2 entropy: $S_2(A) = -\log \operatorname{Tr} \rho_A^2$. This gives:

$$I_{c,2} = S_2(B) - S_2(RB) \tag{10}$$

$$= -\log\left(\operatorname{Tr} \rho_B^2\right) + \log\left(\operatorname{Tr} \rho^2\right), \tag{11}$$

where $\rho = (\mathrm{Id}_R \otimes \mathcal{E})(|\Phi_{RA}\rangle\langle\Phi_{RA}|)$ represents the marginal state on RB of $|\phi_{RBE}\rangle$. Next, we average over the random circuit instances and move the expectation over circuit instances $\mathcal{C}$ inside of the log.

$$\mathop{\mathbf{E}}_{\mathcal{C}} I_{c,2} = -\mathop{\mathbf{E}}_{\mathcal{C}} \log\left(\operatorname{Tr} \rho_B^2\right) + \mathop{\mathbf{E}}_{\mathcal{C}} \log\left(\operatorname{Tr} \rho^2\right) \tag{12}$$

$$\to I_p = -\log\left(\mathop{\mathbf{E}}_{\mathcal{C}} \operatorname{Tr} \rho_B^2\right) + \log\left(\mathop{\mathbf{E}}_{\mathcal{C}} \operatorname{Tr} \rho^2\right), \tag{13}$$

Notice that this expression is now a function of second moment operators of the Haar measure, which are tractable to compute in the special case where the gates have all-to-all connectivity. In addition, it is possible to perform these calculations for any single-qubit noise channels and so we present results for both depolarizing noise and amplitude damping noise. The single-qubit depolarizing channel is defined in eq. (8). The amplitude damping channel is defined as follows:

$$\mathcal{A}_p(\rho) = A_0 \rho A_0^\dagger + A_1 \rho A_1^\dagger \tag{14}$$

where

$$A_0 = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1-p} \end{bmatrix} \tag{15}$$

and

$$A_1 = \begin{bmatrix} 0 & \sqrt{p} \\ 0 & 0 \end{bmatrix} \tag{16}$$

To follow the full calculation for both choices of noise see section A 2.

b. Numerical Setup. Once again, we perform the numerics over a wide range of depths and noise strengths. As in the Clifford case, we provide a plot of the scaling collapse with respect to a dimensionless scaling ansatz. In our numerics, we again consider a rate of $r = 1/8$ and consider system sizes of $n = 10$ to $n = 50$.

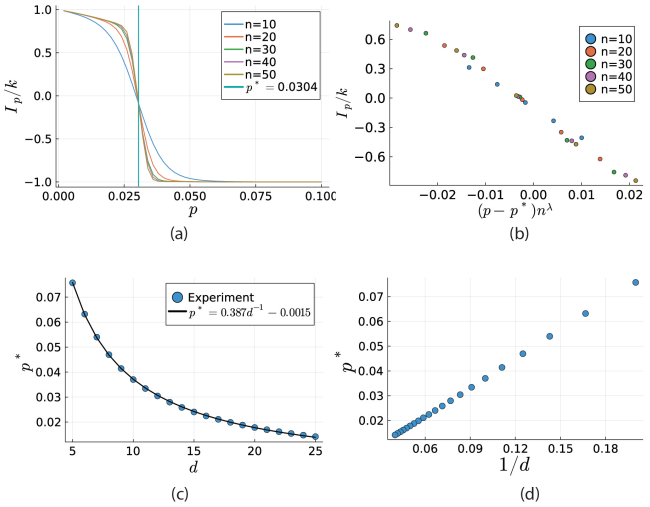


FIG. 5. Numerical results for all-to-all Haar random circuits subjected to mid-circuit depolarizing noise. In all cases an encoding rate of $r = 1/8$ is used. (a) Log-average purity per logical qubit is displayed with respect to a range of error rates p between 0.002 and 0.1. System size is scaled from $n = 10$ and $n = 50$. The circuit depth is fixed at $d = 12$. The critical error rate is depicted by the vertical line which occurs at $p^* = 0.0304$. (b) A scaling collapse for the data in (a) is plotted for error rates near the critical point. The scaling ansatz uses the estimated critical error rate $p^* = 0.0304$ and critical exponent $\lambda = 0.467$. (c) The plots in (a) and (b) are repeated for circuit depths d ranging from 5 to 25. Linear regression on the inverse depth is used to estimate the relationship that $p^* = 0.387d^{-1} - 0.0015$. (d) Finally, the critical error rate p^* is plotted with respect to the inverse depth.

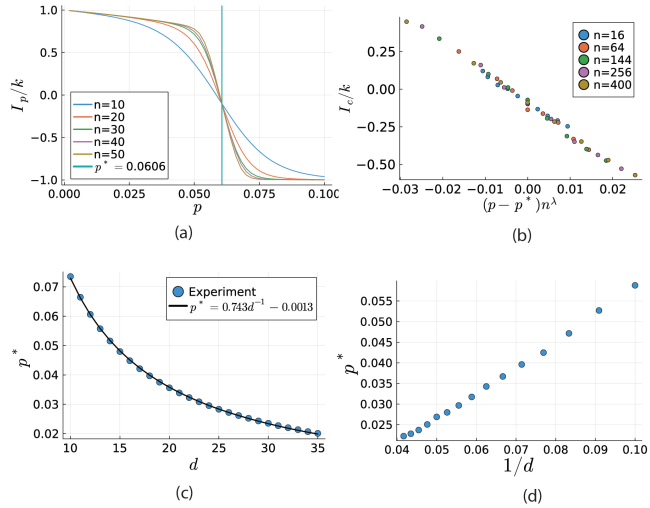


FIG. 6. Numerical results for all-to-all Haar random circuits subjected to mid-circuit amplitude damping noise. In all cases an encoding rate of $r = 1/8$ is used. (a) Log-average purity per logical qubit is displayed with respect to a range of error rates p between 0.002 and 0.1. System size is scaled from $n = 10$ and $n = 50$. The circuit depth is fixed at $d = 12$. The critical error rate is depicted by the vertical line which occurs at $p^* = 0.0606$. (b) A scaling collapse for the data in (a) is plotted for error rates near the critical point. The scaling ansatz uses the estimated critical error rate $p^* = 0.0606$ and critical exponent $\lambda = 0.525$. (c) The plots in (a) and (b) are repeated for circuit depths d ranging from 10 to 35. Linear regression on the inverse depth is used to estimate the relationship that $p^* = 0.743d^{-1} - 0.0013$. (d) Finally, the critical error rate p^* is plotted with respect to the inverse depth.

c. Results. We present the results for the case of depolarizing noise in fig. 5 and the results for amplitude damping noise in fig. 6. When depth is fixed it can be seen that the coherent information once again undergoes a phase transition as the error rate is scaled. Note that the same would be true if we instead fixed the error rate and scaled the depth. In the case of depolarizing noise, the critical point occurs at approximately $p^*d^* \sim 0.39$ while in the amplitude damping case, the critical point occurs at $p^*d^* \sim 0.74$. This suggests that the circuits are more robust to amplitude damping noise than depolarizing. In fact, for a given error rate, information can be preserved for almost twice as long for amplitude damping noise as opposed to depolarizing noise. Interestingly, both noise models have similar critical exponents: $\lambda = 0.467$ in the depolarizing case and $\lambda = 0.525$ in the amplitude damping case. Comparing the 2D numerics in the prior section to the all-to-all case, it is worthwhile to note that for depolarizing noise the 2D circuits have a higher critical point of $p^*d^* \sim 0.627$ compared to $p^*d^* \sim 0.39$ in the all-to-all case. Thus, the extra degrees of interaction in the all-to-all case seems to have spread the noise more quickly rather than encode the information faster. The 2D case also results in a higher critical exponent of $\lambda = 0.616$.

VI. WORST-CASE CIRCUITS

In order to support these numerical results, we analytically prove a rigorous upper bound on the coherent information of *any* noisy circuit, which asymptotically matches our numerics for noisy random circuits. Note these bounds apply to circuits on any architecture and even allow for arbitrary non-local gates. This suggests that noisy random circuits achieve the optimal asymptotic trade-off between error rate and depth.

Theorem VI.1. *Let $\tilde{\mathcal{C}}$ be an arbitrary quantum circuit acting on n qubits, which has d layers of interspersed depolarizing noise of strength p acting independently on each qubit. Then the coherent information of the encoding channel which encodes a k-qubit quantum state ζ as $\mathcal{E}(\zeta) = \tilde{C}(\zeta \otimes |0\rangle\langle 0|^{n-k})$, is bounded as follows:*

$$I_c(\frac{\mathbb{I}}{2^k}, \mathcal{E}) \leq e^{-pd}(n + k) - k \tag{17}$$

We obtain this bound using relative entropy convergence under tensor products of depolarizing channels [9, 26]. The proof is deferred to section A 3.

Let us denote the encoding rate as $r := k/n$. This implies that when $d > p^{-1}\log(1 + 1/r)$, then $I_c < 0$. Assuming r is some constant, this proves that the coherent

information is negative and thus the logical information is unrecoverable for any noisy circuit after a depth threshold of $d^* = O(p^{-1})$. Furthermore, setting $d = 1$ shows that $r < (e^p - 1)^{-1}$ in order for $I_c \geq 0$. This provides a bound on the largest allowable encoding rate that is still recoverable after any depth.

It is worthwhile to sanity check this bound against known fault-tolerance results. In particular, Aharonov *et al.* [10] give a protocol (that is closely related to [27]) that is robust to errors and uses at least $k2^d$ ancilla qubits. This corresponds to a rate of $r \leq 2^{-d}$. Plugging this into our bound shows that the coherent information of their circuit becomes negative after a depth of $p^{-1}\log(1 + 1/r) \geq p^{-1}d > d$. Thus, their protocol uses just enough ancilla qubits to avoid our bounds.

VII. DISCUSSION

Our results suggest several avenues for future work. For instance, one immediate goal is to prove the existence of our observed phase transition analytically. Similar results have been achieved for low-depth random-circuit codes in the code capacity setting [28] and so it is possible that related techniques could be extended to our setting of random circuits subject to mid-circuit depolarizing noise.

Next, our work has tightly characterized the tradeoff between depth and error rate, but it remains to understand how encoding rate fits into this picture. For instance, up until what depth does the circuit preserve a vanishing encoding rate? It would also be interesting to investigate how our results change as the circuit connectivity is varied between 2D and all-to-all architectures (e.g. varying the degree of the connectivity graph). We hope to further expand our phase diagram to include these factors.

Finally, a constant-depth phase transition in quantum circuits with depolarizing noise has been previously observed in a different setting of computational complexity. In particular, the output distributions of noisy IQP and Clifford-magic circuits [29, 30] can be efficiently classically sampled above a critical depth of approximately $d^* = \tilde{O}(1/p)$ (omitting log factors), whereas below this critical depth, the task of sampling is outside the polynomial hierarchy [31], indicating there is no efficient classical algorithm. These prior works exploit a noise percolation phenomenon beyond the critical depth, which cuts the output state into many small, unentangled pieces. Our work shows a similar phenomenon: namely, we show an inability to effectively encode and protect quantum information past the critical depth threshold, which also implies a breakdown of long-range entanglement. It would be interesting to see whether this connection could be made more rigorous.

ACKNOWLEDGEMENTS

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Accelerated Research in Quantum Computing, Fundamental Algorithmic Research toward Quantum Utility (FAR-Qu). We thank Dominik Hangleiter, Zhi-Yuan Wei, Daniel Malz and Alexey Gorshkov for helpful discussions. This material is based upon work supported in part by the NSF QLCI award OMA2120757. This work was performed in part at the Kavli Institute for Theoretical Physics (KITP), which is supported by grant NSF PHY-2309135. JN is supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE 2236417.

[1] Y. Li, X. Chen, and M. P. A. Fisher, Phys. Rev. B **100**, 134306 (2019).
[2] B. Skinner, J. Ruhman, and A. Nahum, Phys. Rev. X **9**, 031009 (2019).
[3] Y. Li, X. Chen, and M. P. A. Fisher, Phys. Rev. B **98**, 205136 (2018).
[4] M. J. Gullans and D. A. Huse, Physical Review X **10** (2020), 10.1103/physrevx.10.041020.
[5] Z. Weinstein, S. P. Kelly, J. Marino, and E. Altman, Physical Review Letters **131** (2023), 10.1103/physrevlett.131.220404.
[6] D. Qian and J. Wang, "Coherent information phase transition in a noisy quantum circuit," (2024), arXiv:2408.16267 [quant-ph].
[7] J. Nelson, G. Bentsen, S. T. Flammia, and M. J. Gullans, Phys. Rev. Res. **7**, 013040 (2025).
[8] In this paper, we adopt the standard asymptotic notation O, Ω, Θ, o, and ω to describe relationships between the growth rates of two nonnegative functions $f(n)$ and $g(n)$ as $n \to \infty$. We write $f = O(g)$ when $\lim_{n\to\infty} f(n)/g(n) < \infty$, which is equivalent to stating that $g = \Omega(f)$. If $\lim_{n\to\infty} f(n)/g(n) = 0$, we denote this by $f = o(g)$, or equivalently $g = \omega(f)$. The notation $f = \Theta(g)$ indicates that both $f = O(g)$ and $f = \Omega(g)$ are satisfied. When a tilde is used, as in $\tilde{\Theta}$, it signifies that multiplicative polylogarithmic factors in n are omitted from the scaling.
[9] A. Müller-Hermes, D. Stilck França, and M. M. Wolf, Journal of Mathematical Physics **57** (2016), 10.1063/1.4939560.
[10] D. Aharonov, M. Ben-Or, R. Impagliazzo, and N. Nisan, "Limitations of noisy reversible computation," (1996), arXiv:quant-ph/9611028 [quant-ph].
[11] Optimal, here, is with respect to the asymptotic relationship between depth and noise strength and does not take into account other parameters such as encoding rate.
[12] W. Brown and O. Fawzi, in *2013 IEEE International Symposium on Information Theory* (2013) pp. 346–350.
[13] W. Brown and O. Fawzi, Communications in mathematical physics **340**, 867 (2015).

[14] M. J. Gullans, S. Krastanov, D. A. Huse, L. Jiang, and S. T. Flammia, Phys. Rev. X **11**, 031066 (2021).

[15] P. M. Hayden, M. Horodecki, A. J. Winter, and J. T. Yard, Open Syst. Inf. Dyn. **15**, 7 (2007).

[16] W. Brown and O. Fawzi, "Scrambling speed of random quantum circuits," (2013), arXiv:1210.6644 [quant-ph].

[17] S. Lloyd, Physical Review A **55**, 1613–1622 (1997).

[18] I. Devetak and P. W. Shor, "The capacity of a quantum channel for simultaneous transmission of classical and quantum information," (2004), arXiv:quant-ph/0311131 [quant-ph].

[19] A. Nahum, J. Ruhman, S. Vijay, and J. Haah, Phys. Rev. X **7**, 031016 (2017).

[20] A. Nahum, S. Vijay, and J. Haah, Phys. Rev. X **8**, 021014 (2018).

[21] T. Zhou and A. Nahum, Physical Review B **99** (2019), 10.1103/physrevb.99.174205.

[22] J. Preskill, "Physics219 caltech lecture notes," .

[23] D. Gottesman, "The heisenberg representation of quantum computers," (1998), arXiv:quant-ph/9807006 [quant-ph].

[24] S. Aaronson and D. Gottesman, Physical Review A **70** (2004), 10.1103/physreva.70.052328.

[25] M. E. Fisher, Rev. Mod. Phys. **70**, 653 (1998).

[26] J. Nelson, J. Rajakumar, and M. J. Gullans, "Limitations of noisy geometrically local quantum circuits," (2025), arXiv:2510.xxxxx [quant-ph].

[27] D. Aharonov and M. Ben-Or, "Fault-tolerant quantum computation with constant error rate," (1999), arXiv:quant-ph/9906129 [quant-ph].

[28] G. Liu, Z. Du, Z.-W. Liu, and X. Ma, "Approximate quantum error correction with 1d log-depth circuits," (2025), arXiv:2503.17759 [quant-ph].

[29] J. Rajakumar, J. D. Watson, and Y.-K. Liu, "Polynomial-time classical simulation of noisy iqp circuits with constant depth," in *Proceedings of the 2025 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)* (Society for Industrial and Applied Mathematics, 2025) p. 1037–1056.

[30] J. Nelson, J. Rajakumar, D. Hangleiter, and M. J. Gullans, "Polynomial-time classical simulation of noisy circuits with naturally fault-tolerant gates," (2024), arXiv:2411.02535 [quant-ph].

[31] K. Fujii and S. Tamate, Sci Rep **6** (2016), 10.1038/srep25598, arXiv:1406.6932.

[32] B. Ware, A. Deshpande, D. Hangleiter, P. Niroula, B. Fefferman, A. V. Gorshkov, and M. J. Gullans, "A sharp phase transition in linear cross-entropy benchmarking," (2023), arXiv:2305.04954 [quant-ph].

[33] A. M. Dalzell, N. Hunter-Jones, and F. G. S. L. Brandão, Commun. Math. Phys. **405**, 78 (2024), arXiv:2111.14907.

[34] M. Wilde, "markwilde.com," `http://www.markwilde.com/teaching/2015-fall-qit/lectures/lecture-19.pdf`, [Accessed 11-09-2025].

Appendix A: Omitted proofs and calculations

1. Operational Interpretation of Coherent Information

Fact A.1. *Let $\mathcal{E}^{A\to B}$ be a channel from $\mathcal{H}_A$ to $\mathcal{H}_B$ where $\mathcal{H}_A$ has dimension 2^k. If $I_c(\mathbb{I}_A/2^k, \mathcal{E}) = k$ if and only if there exists a decoding channel $\mathcal{D}$ such that for any $|\psi\rangle \in \mathcal{H}_A$:*

$$\mathcal{D} \circ \mathcal{E}(|\psi\rangle\langle\psi|) = |\psi\rangle\langle\psi| \tag{A1}$$

Proof. The provided proof sketch follows the notes of [22] and adopts similar notation for consistency. As in the main text, we consider the purification of $\mathbb{I}_A/2^k$ to be $|\Phi\rangle_{RA} = \sum_{i=1}^{2^k} |i,i\rangle\langle i,i|$, the channel dilation of $\mathcal{E}$ to be $U^{A\to BE}$, and the purification of the output to be $|\phi_{RBE}\rangle$. We start by showing that $I_c = k$ implies that the reference and environment of $|\phi_{RBE}\rangle$ are decoupled.

$$I_c = S(B) - S(E) = S(RE) - S(E) = k \tag{A2}$$
$$\implies S(RE) = k + S(E) \tag{A3}$$
$$\implies S(RE) = S(R) + S(E) \tag{A4}$$

Therefore, the marginal state of $|\phi\rangle_{RBE}$ on RE must be a product state. First let us purify the marginal states on R and E and denote these as $|\psi\rangle_{RB_1}$ and $|\chi\rangle_{B_2E}$ which together give the purification of the marginal state on RE as $|\psi\rangle_{RB_1} \otimes |\chi\rangle_{B_2E}$. Next, using the fact that all purifications differ by a unitary transformation on the purifying register, we have that $|\phi\rangle_{RBE} = W_B(|\psi\rangle_{RB_1} \otimes |\chi\rangle_{B_2E})$ where W_B is a unitary matrix that acts only on $B = B_1 \cup B_2$. Finally, using this fact once more, we have that there exists a unitary V_{B_1} such that $|\Phi\rangle_{RA} = V_{B_1} |\psi\rangle_{RB_1}$. Therefore, we can let our decoding map be $\mathcal{D} = V_{B_1} W_B^\dagger$, which exactly recovers the maximally entangled state. Finally, we can show that recovering the purification of the maximally entangled state implies the ability to recover any logical codestate. This can be done using the relative-state method (section 3.3 of [22]) as shown below.

First, let $|\varphi\rangle_A$ be some logical codestate. We can write this as:

$$|\varphi\rangle_A = \sum_i \varphi_i |i\rangle_A = \sqrt{2^n} \sum_i \varphi_i \left({}_R\langle i|\Phi\rangle_{RA}\right) = \sqrt{2^n}\, {}_R\langle\varphi^*|\Phi\rangle_{RA} \tag{A5}$$

Next, using linearity we have:

$$\mathcal{D} \circ \mathcal{E}(|\varphi\rangle\langle\varphi|_A) = 2^n\, {}_R\langle\varphi^*|(\mathrm{Id}_R \otimes D \circ \mathcal{E})(|\Phi_{RA}\rangle\langle\Phi_{RA}|)\, |\varphi^*\rangle_R \tag{A6}$$
$$= 2^n\, {}_R\langle\varphi^*|\Phi_{RA}\rangle\, \langle\Phi_{RA}|\varphi^*\rangle_R \tag{A7}$$
$$= |\varphi\rangle\langle\varphi|_A \tag{A8}$$

The converse is also true in that if the purification of the maximally entangled state can be recovered, then $I_c = k$. This follows by noting that $I_c(\rho_A, \mathcal{D} \circ \mathcal{E}) = k$ since the reference and environment are necessarily decoupled if the reference system is successfully purified into the channel's output. Next, the statement follows by the quantum data processing inequality, which states that a channel cannot increase the coherent information.

□

Fact A.2. *Let $\mathcal{E}^{A\to B}$ be a channel from $\mathcal{H}_A$ to $\mathcal{H}_B$ where $\mathcal{H}_A$ has dimension 2^k. If $I_c(\mathbb{I}_A/2^k, \mathcal{E}) = -k$, then for any decoding channel $\mathcal{D}$ and any two codestates $|\psi_1\rangle$ and $|\psi_2\rangle$,*

$$\mathcal{D} \circ \mathcal{E}(|\psi_1\rangle\langle\psi_1|) = \mathcal{D} \circ \mathcal{E}(|\psi_2\rangle\langle\psi_2|) \tag{A9}$$

i.e. no logical information can be transmitted.

Proof. This is seen by observing that when $I_c = -k$ then $I(R:B) = 0$ even after any decoding map is applied due to the quantum data processing inequality. Let $\rho_{RB} = (\mathrm{Id}_R \otimes \mathcal{D} \circ \mathcal{E})(|\Phi_{RA}\rangle\langle\Phi_{RA}|)$. Since $I(R:B) = 0$ we have that $\rho_{RB} = \rho_R \otimes \rho_B = \frac{\mathbb{I}_R}{2^n} \otimes \rho_B$. Then, we have that for any logical state $|\varphi\rangle\langle\varphi|_A$:

$$\mathcal{D} \circ \mathcal{E}(|\varphi\rangle\langle\varphi|_A) = 2^n\, {}_R\langle\varphi^*|(\mathrm{Id}_R \otimes \mathcal{D} \circ \mathcal{E})(|\Phi_{RA}\rangle\langle\Phi_{RA}|)\, |\varphi^*\rangle_R \tag{A10}$$
$$= 2^n\, {}_R\langle\varphi^*|\frac{\mathbb{I}_R}{2^n} \otimes \rho_B\, |\varphi^*\rangle_R \tag{A11}$$
$$= \rho_B \tag{A12}$$

Therefore, no matter what the logical state is, the output will always be the same, and so no information can be recovered. □

2. Calculating the Log-Average Purity

Recall that the log-average purity is defined as:

$$I_p = -\log\left(\mathop{\mathbf{E}}_{\mathcal{C}} \operatorname{Tr} \rho_B^2\right) + \log\left(\mathop{\mathbf{E}}_{\mathcal{C}} \operatorname{Tr} \rho^2\right) \tag{A13}$$

We can rewrite the right-hand side of this equation in a way that is easier to calculate by first introducing a second copy of the state giving $\rho \otimes \rho$. Next, we introduce the notation $\mathcal{S}$, which denotes the swap operator between two copies of a qubit, and $\mathcal{I}$, which denotes the identity operator acting on two copies of a qubit. We further define $\mathbb{S} := \mathcal{S}^{\otimes n}$ and $\mathbb{S}_A := \mathcal{I}^{\otimes |\bar{A}|} \otimes \mathcal{S}^{\otimes |A|}$. Using this notation, we have the following identity:

$$\operatorname{Tr} \rho_C^2 = \operatorname{Tr} \mathbb{S}_C (\rho \otimes \rho) \tag{A14}$$

for an arbitrary region C. Applying this to eq. (13), we get:

$$I_p = -\log\left(\mathop{\mathbf{E}}_{\mathcal{C}} \operatorname{Tr} \mathbb{S}_B \rho^{\otimes 2}\right) + \log\left(\mathop{\mathbf{E}}_{\mathcal{C}} \operatorname{Tr} \mathbb{S} \rho^{\otimes 2}\right) \tag{A15}$$

Recall that $\rho = (\mathrm{Id}_R \otimes \mathcal{E})(|\Phi_{RA}\rangle\langle\Phi_{RA}|)$. Plugging this in, we can make further rearrangements as follows:

$$-\log\left(\mathop{\mathbf{E}}_{\mathcal{C}} \operatorname{Tr} \mathbb{S}_B \rho^{\otimes 2}\right) = -\log\left(\mathop{\mathbf{E}}_{\mathcal{C}} \operatorname{Tr} \mathbb{S}_B (\mathrm{Id}_R \otimes \mathcal{E})^{\otimes 2} (|\Phi_{RA}\rangle\langle\Phi_{RA}|^{\otimes 2})\right) \tag{A16}$$

$$= -\log\left(\operatorname{Tr} \mathbb{S}_B \mathop{\mathbf{E}}_{\mathcal{C}} \mathcal{E}^{\otimes 2} \left(\frac{\mathcal{I}}{4}\right)^{\otimes k}\right) \tag{A17}$$

And similarly for the second term:

$$\log\left(\mathop{\mathbf{E}}_{\mathcal{C}} \operatorname{Tr} \mathbb{S} \rho^{\otimes 2}\right) = \log\left(\mathop{\mathbf{E}}_{\mathcal{C}} \operatorname{Tr} \mathbb{S} (\mathrm{Id}_R \otimes \mathcal{E})^{\otimes 2} (|\Phi_{RA}\rangle\langle\Phi_{RA}|^{\otimes 2})\right) \tag{A18}$$

$$= \log\left(\operatorname{Tr} \mathbb{S}_B \mathop{\mathbf{E}}_{\mathcal{C}} \mathcal{E}^{\otimes 2} \left(\frac{\mathcal{S}}{4}\right)^{\otimes k}\right) \tag{A19}$$

$$= \log\left(\operatorname{Tr} \mathbb{S}_B \mathop{\mathbf{E}}_{\mathcal{C}} \mathcal{E}^{\otimes 2} \left(\frac{\mathcal{S}}{2}\right)^{\otimes k}\right) - k \tag{A20}$$

Notice that in the above expression we normalize $\mathcal{I}$ and $\mathcal{S}$ by their respective traces. We begin our calculation of $\mathcal{E}^{\otimes 2}((\frac{\mathcal{I}}{4})^{\otimes k})$ by applying two copies of the first layer of $\mathcal{C}$ to $(\frac{\mathcal{I}}{4})^{\otimes k} \otimes (|0\rangle\langle 0|^{\otimes 2})^{\otimes n-k}$. We let the first layer of gates be single-qubit gates drawn from the Haar measure. Notice that introducing this layer does not change the expectation since the Haar measure is both left and right invariant over the unitary group. Let us denote the second moment operator for the single-qubit Haar random gate acting on qubit i as

$$\mathcal{M}_i(O) = \mathop{\mathbf{E}}_{V \sim \mathcal{H}} (V \otimes V) O (V \otimes V)^\dagger \tag{A21}$$

where O is an operator on both copies of qubit i. The key insight is that this average, also known as the second moment operator of the Haar measure, projects any operator onto the symmetric subspace, which is spanned by tensor products of $\mathcal{S}/2$ and $\mathcal{I}/4$. In fact, every remaining operation leaves our state in this symmetric subspace. This is still intractable, however, since this space is exponentially large. In order to simplify the calculation further, we follow [32], which considers an all-to-all circuit geometry where qubits are randomly permuted between layers of gates. We also adopt much of the notation from [32] for consistency. Now, after symmetrizing over permutations, our state is constrained to a subspace spanned by $n+1$ basis operators, which are defined by taking the average over all tensor products of the swap and identity operators with the same Hamming weight of total swap operators. More specifically, we denote the basis states by $\{|S\rangle\}$ where $S \in [0, n]$ specifies the Hamming weight,

$$|S\rangle = \frac{1}{\binom{n}{S}} \sum_{\sigma \in \{0,1\}^n \text{ s.t. } |\sigma| = S} \left(\frac{\mathcal{I}}{4}\right)^{1-\sigma_i} \left(\frac{\mathcal{S}}{2}\right)^{\sigma_i} \tag{A22}$$

With this basis, we can formulate our problem as applying a series of transfer matrices [32, 33]. Deferring the calculation details to the proof of Fact A.3, we have that the state after the first layer of single-qubit gates and random permutations is the following:

$$\mathop{\mathbf{E}}_{P\sim S_n} P\circ\mathcal{M}_n\circ\cdots\circ\mathcal{M}_1(\left(\frac{\mathcal{I}}{4}\right)^{\otimes k}\otimes(|0\rangle\langle 0|^{\otimes 2})^{\otimes n-k}))=\frac{1}{3^{n-k}}\sum_{S=0}^{n-k}\binom{n-k}{S}2^{n-k-S}\,|S\rangle \tag{A23}$$

$$\mathop{\mathbf{E}}_{P\sim S_n} P\circ\mathcal{M}_n\circ\cdots\circ\mathcal{M}_1(\left(\frac{\mathcal{S}}{2}\right)^{\otimes k}\otimes(|0\rangle\langle 0|^{\otimes 2})^{\otimes n-k}))=\frac{1}{3^{n-k}}\sum_{S=0}^{n-k}\binom{n-k}{S}2^{n-k-S}\,|S+k\rangle \tag{A24}$$

where S_n is the n-element permutation group.

Next, we must show how to update this state after each layer of 2-qubit gates and noise channels. To do this, it is sufficient to show the action of each layer on each basis state $|S\rangle$. Let us denote the second moment operator for the two-qubit Haar random gate acting on qubits i and j as

$$\mathcal{M}_{i,j}(O)=\mathop{\mathbf{E}}_{V\sim\mathcal{H}}(V\otimes V)O(V\otimes V)^\dagger \tag{A25}$$

where V is a two-qubit gate acting on qubits i and j and O is an operator on both copies of qubits i and j..

Then we define the transfer matrix $M_{S',S}$, which corresponds to the action of a layer of 2-qubit gates followed by a random permutation:

$$\mathop{\mathbf{E}}_{P\in S_n} P\circ\mathcal{M}_{n-1,n}\circ\cdots\circ\mathcal{M}_{1,2}(|S\rangle)=\sum_{S'=0}^{n}M_{S',S}\,|S'\rangle \tag{A26}$$

It can be shown (see Fact A.4) that $M_{S',S}$ is defined as follows:

$$M_{S',S}=\frac{1}{\binom{n}{S}}\sum_{n_0,n_1,n_2}2^{n_1}\binom{n/2}{n_0,n_1,n_2}\delta_{S,n_1+2n_2}\left[\sum_{a,b}\binom{n_1}{a,b}\left(\frac{4}{5}\right)^a\left(\frac{1}{5}\right)^b\delta_{S',2b+2n_2}\right] \tag{A27}$$

Finally, we must give the transfer matrix associated with applying a layer of single-qubit noise channels. In this case we apply random single-qubit gates right after the noise channel, which has no physical effect due to the left-right invariance of the Haar measure. The transfer matrix can be defined as

$$N_{S',S}=\sum_{a=\max(0,S-S')}^{\min(S,N-S')}\binom{S}{a}\binom{N-S}{S'-(S-a)}\delta^{S'-(S-a)}\gamma^a(1-\delta)^{N-S'-a}(1-\gamma)^{S-a} \tag{A28}$$

where δ and γ are parameters of the noise channel that depend on the noise rate. See Fact A.5 for further details in the derivation of this transfer matrix. For the depolarizing channel defined in eq. (8), $\delta=0$ and $\gamma=1-(1-\frac{4p}{3})^2$. Setting $\delta=0$, the transfer matrix in eq. (A28) simplifies to:

$$N_{S',S}=\binom{S}{S'}(1-\gamma)^{S'}\gamma^{S-S'} \tag{A29}$$

where $N_{S',S}=0$ if $S<S'$. For the damping channel defined in eq. (14), $\delta=\frac{p^2}{3}$ and $\gamma=\frac{2}{3}(2p-p^2)$.

Finally, after applying the transfer matrices for d layers of gates, permutations, and noise channels, we have obtained a decomposition of $\mathbf{E}_{\mathcal{C}}\,\mathcal{E}^{\otimes 2}\left(\frac{\mathcal{I}}{4}\right)^{\otimes k}$ and $\mathbf{E}_{\mathcal{C}}\,\mathcal{E}^{\otimes 2}\left(\frac{\mathcal{S}}{2}\right)^{\otimes k}$ in basis vectors spanned by $|S\rangle$. We can then efficiently compute $\operatorname{Tr}\mathbb{S}\,\mathbf{E}_{\mathcal{C}}\,\mathcal{E}^{\otimes 2}\left(\frac{\mathcal{I}}{4}\right)^{\otimes k}$ and $\operatorname{Tr}\mathbb{S}\,\mathbf{E}_{\mathcal{C}}\,\mathcal{E}^{\otimes 2}\left(\frac{\mathcal{S}}{2}\right)^{\otimes k}$ using the fact that

$$\operatorname{Tr}\mathbb{S}\,|S\rangle=2^{2S-n} \tag{A30}$$

and linearity of trace.

Fact A.3.

$$\mathop{\mathbf{E}}_{P\sim S_n} P\circ\mathcal{M}_n\circ\cdots\circ\mathcal{M}_1(\left(\frac{\mathcal{I}}{4}\right)^{\otimes k}\otimes(|0\rangle\langle 0|^{\otimes 2})^{\otimes n-k}))=\frac{1}{3^{n-k}}\sum_{S=0}^{n-k}\binom{n-k}{S}2^{n-k-S}\,|S\rangle \tag{A31}$$

$$\mathop{\mathbf{E}}_{P\sim S_n} P\circ\mathcal{M}_n\circ\cdots\circ\mathcal{M}_1(\left(\frac{\mathcal{S}}{2}\right)^{\otimes k}\otimes(|0\rangle\langle 0|^{\otimes 2})^{\otimes n-k}))=\frac{1}{3^{n-k}}\sum_{S=0}^{n-k}\binom{n-k}{S}2^{n-k-S}\,|S+k\rangle \tag{A32}$$

where S_n is the n-element permutation group and $\mathcal{M}_i$ is defined as in eq. (A21).

Proof. The Haar average of an operator O is defined as follows:

$$\mathop{\mathbf{E}}_{V\sim\mathcal{H}}(V\otimes V)O(V\otimes V)^\dagger=a\mathcal{I}+b\mathcal{S}, \tag{A33}$$

where

$$a=\frac{\operatorname{Tr}O-\frac{1}{2}\operatorname{Tr}O\mathcal{S}}{3} \tag{A34}$$

and

$$b=\frac{\operatorname{Tr}O\mathcal{S}-\frac{1}{2}\operatorname{Tr}O}{3} \tag{A35}$$

The only relevant input operators are $|0\rangle\langle 0|^{\otimes 2}$, $\mathcal{S}$, $\mathcal{I}$. We have

$$\mathop{\mathbf{E}}_{V\sim\mathcal{H}}(V\otimes V)\,|0\rangle\langle 0|^{\otimes 2}\,(V\otimes V)^\dagger=\mathcal{I}/6+\mathcal{S}/6, \tag{A36}$$

$$\mathop{\mathbf{E}}_{V\sim\mathcal{H}}(V\otimes V)\mathcal{S}(V\otimes V)^\dagger=\mathcal{S}, \tag{A37}$$

$$\mathop{\mathbf{E}}_{V\sim\mathcal{H}}(V\otimes V)\mathcal{I}(V\otimes V)^\dagger=\mathcal{I}, \tag{A38}$$

Now averaging this uniformly over qubit permutations we get:

$$\frac{1}{n!}\sum_{P\in S_n}P\circ M_n\circ\cdots\circ\mathcal{M}_1(\left(\frac{\mathcal{I}}{4}\right)^{\otimes k}\otimes|0\rangle\langle 0|^{\otimes n-k})=\frac{1}{3^{n-k}}\sum_{S=0}^{n-k}\binom{n-k}{S}2^{n-k-S}\,|S\rangle \tag{A39}$$

where S_n is the n-element permutation group. A similar calculation shows the following as well:

$$\frac{1}{n!}\sum_{P\in S_n}P\circ M_n\circ\cdots\circ\mathcal{M}_1(\left(\frac{\mathcal{S}}{2}\right)^{\otimes k}\otimes|0\rangle\langle 0|^{\otimes n-k})=\frac{1}{3^{n-k}}\sum_{S=0}^{n-k}\binom{n-k}{S}2^{n-k-S}\,|S+k\rangle \tag{A40}$$

□

Fact A.4.

$$M_{S',S}=\frac{1}{\binom{n}{S}}\sum_{n_0,n_1,n_2}2^{n_1}\binom{n/2}{n_0,n_1,n_2}\delta_{S,n_1+2n_2}\left[\sum_{a,b}\binom{n_1}{a,b}\left(\frac{4}{5}\right)^a\left(\frac{1}{5}\right)^b\delta_{S',2b+2n_2}\right] \tag{A41}$$

where M is defined as in eq. (A26)

Proof sketch. It is helpful to first note the 2-qubit transfer matrix M is given by

$$M=\begin{bmatrix}1&4/5&4/5&0\\0&0&0&0\\0&0&0&0\\0&1/5&1/5&1\end{bmatrix} \tag{A42}$$

where the columns and rows are indexed by $\mathcal{I}\otimes\mathcal{I}$, $\mathcal{I}\otimes\mathcal{S}$, $\mathcal{S}\otimes\mathcal{I}$, and $\mathcal{S}\otimes\mathcal{S}$. This can be computed using eq. (A33). In the above definition of M, n_0, n_1, n_2 represent the number of gates which act on a tensor product of 0, 1, or 2 swap operators. Among the gates that encounter 1 swap operator (there are n_1 such gates), a and b represent the number of these gates that output a tensor product of 0 or 2 swap operators. □

Fact A.5.

$$N_{S',S} = \sum_{a=\max(0,S-S')}^{\min(S,N-S')} \binom{S}{a} \binom{N-S}{S'-(S-a)} \delta^{S'-(S-a)} \gamma^a (1-\delta)^{N-S'-a} (1-\gamma)^{S-a} \tag{A43}$$

Proof sketch. The single-qubit transfer matrix of a noise channel followed by a random single-qubit gate takes the general form of

$$N = \begin{bmatrix} 1-\delta & \gamma \\ \delta & 1-\gamma \end{bmatrix} \tag{A44}$$

where the rows and columns are indexed by $\mathcal{I}$ and $\mathcal{S}$. δ and γ can both be calculated by applying eq. (A33) for the given noise channel. For a given qubit, there are four possibilities allowed by this transfer matrix: $\mathcal{I} \to \mathcal{I}$, $\mathcal{I} \to \mathcal{S}$, $\mathcal{S} \to \mathcal{I}$, and $\mathcal{S} \to \mathcal{S}$. In eq. (A43), a represents the number of qubits that undergo the transition $\mathcal{S} \to \mathcal{I}$. Notice that if $S > S'$ then there must be at least $S - S'$ qubits that experience $\mathcal{S} \to \mathcal{I}$ which is why a takes this as its minimum value. In addition, the output contains at most $N - S'$ $\mathcal{I}$ operators. Furthermore, there are only S qubits that contain $\mathcal{S}$ operator. Therefore a cannot exceed $N - S'$ or S which is why these form the upper limits of the sum. Now, the remaining quantities can be deduced from a as follows. There are $S - a$ qubits experiencing the transition $\mathcal{S} \to \mathcal{S}$, $S' - (S - a)$ experiencing the transition $\mathcal{I} \to \mathcal{S}$, and $N - S' - a$ qubits experiencing the transition $\mathcal{I} \to \mathcal{I}$. Accounting for the degeneracies and including the appropriate factor for each transition type results in the desired equation. □

3. Proof of Theorem VI.1

Our proof strategy is to relate coherent information to a *relative entropy distance* which is defined as $D(\rho||\zeta) := \operatorname{Tr} \rho(\log \rho - \log \zeta)$, where we can apply the following convergence result for noisy quantum circuits,

Fact A.6 (Follows from Lemma 3.2 of [26])**.** *Let X and Y be quantum systems, and let ρ_{XY} denote the joint state obtained as follows: start from any input state on XY, apply an arbitrary quantum circuit acting only on X, which has d layers of interspersed depolarizing noise of strength p acting independently on each qubit in X. Then,*

$$D(\rho_{XY} \| \frac{\mathbb{I}_X}{2^{|X|}} \otimes \rho_Y) \leq (1-p)^d |XY| \tag{A45}$$

Theorem A.7 (Restatement of Theorem VI.1)**.** *Let $\tilde{\mathcal{C}}$ be an arbitrary quantum circuit acting on n qubits, which has d layers of interspersed depolarizing noise of strength p acting independently on each qubit. Then the coherent information of the encoding channel which encodes a k-qubit quantum state ζ as $\mathcal{E}(\zeta) = \tilde{C}(\zeta \otimes |0\rangle\langle 0|^{n-k})$, is bounded as follows:*

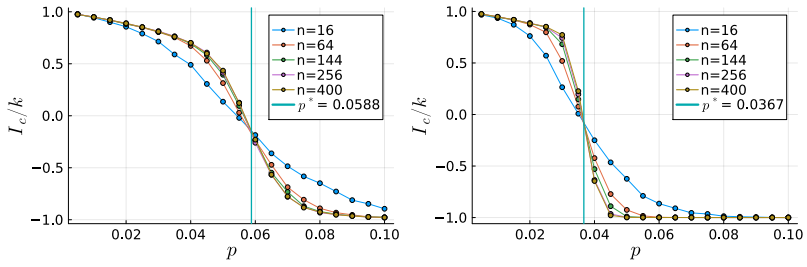
$$I_c(\frac{\mathbb{I}}{2^k}, \mathcal{E}) \leq e^{-pd}(n+k) - k \tag{A46}$$

Proof. We will consider the Hilbert spaces corrresponding to A, B, R, E and entropies of the reduced density matrices of $|\phi\rangle\langle\phi|_{RBE}$, as defined in section III. We will denote $\rho = \operatorname{Tr}_E(|\phi\rangle\langle\phi|_{RBE}) = (\mathrm{Id}_R \otimes \mathcal{E})(|\Phi_{RA}\rangle\langle\Phi_{RA}|)$. We also use the known identity that $S(B|R) = |B| - D(\rho_{RB} \| \rho_R \otimes \frac{\mathbb{I}_B}{2^{|B|}})$ [34].

$$\begin{aligned}
I_c &= S(B) - S(E) \\
&= S(B) - S(RB) \\
&= S(B) - S(B|R) - S(R) \\
&\leq n - S(B|R) - k && \text{(since } \rho_R \text{ is maximally mixed)} \\
&= n - (n - D(\rho_{RB} \| \rho_R \otimes \frac{\mathbb{I}_B}{2^{|B|}})) - k \\
&= D(\rho_{RB} \| \rho_R \otimes \frac{\mathbb{I}_B}{2^{|B|}}) - k \\
&= (1-p)^d (n+k) - k \\
&\leq e^{-pd}(n+k) - k
\end{aligned}$$

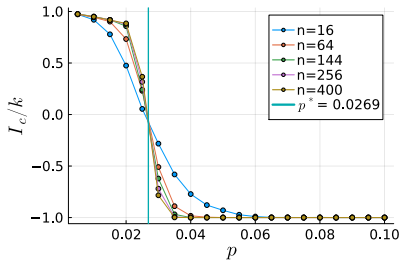
□

Appendix B: Supplementary Plots

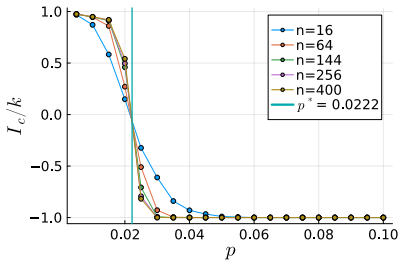


(a) $d = 10$

(b) $d = 15$

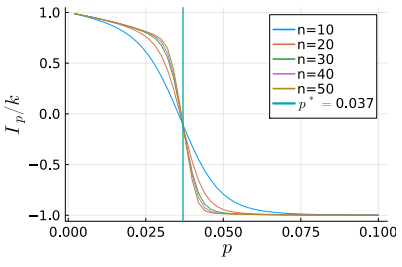


(c) $d = 20$

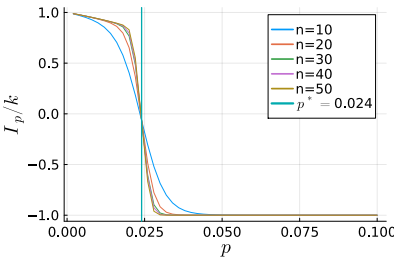


(d) $d = 24$

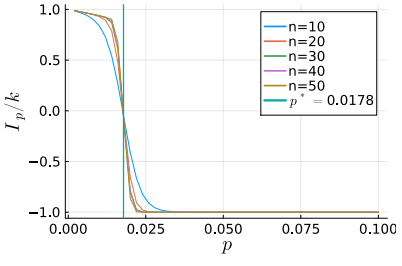
FIG. 7. I_c/k vs p for 2D random Clifford circuits at various depths $d = 10, \ldots, 24$ where many depth values are omitted to conserve space. The estimated critical error rates for each of these plots are used to generate fig. 4.



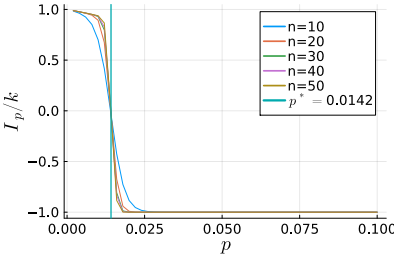
(a) $d = 10$



(b) $d = 15$

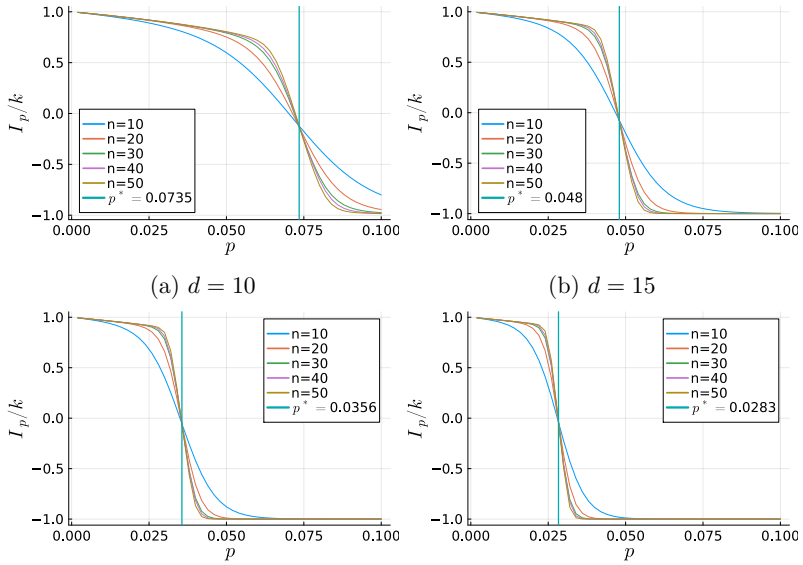


(c) $d = 20$



(d) $d = 25$

FIG. 8. I_p/k vs p for Haar random circuits with depolarizing noise at various depths $d = 5, \ldots, 25$ where many depth values are omitted to conserve space. The estimated critical error rates for each of these plots are used to generate fig. 5.



(a) $d = 10$ (b) $d = 15$

(c) $d = 20$ (d) $d = 25$

FIG. 9. I_p/k vs p for Haar random circuits with amplitude damping noise at various depths $d = 10, \ldots, 35$ where many depth values are omitted to conserve space. The estimated critical error rates for each of these plots are used to generate fig. 6.