To Learn a Mocking-Black Hole

Lorenzo Leone,1,* Salvatore F.E. Oliviero,1 Stefano Piemontese,1 Sarah True,1 and Alioscia Hamma1,2

1Physics Department, University of Massachusetts Boston, 02125, USA
2Dipartimento di Fisica Ettore Pancini, Università degli Studi di Napoli Federico II, Via Cintia 80126, Napoli, Italy

In a seminal paper[1], Hayden and Preskill showed that information can be retrieved from a black hole that is sufficiently scrambling, assuming that the retriever has perfect control of the emitted Hawking radiation and perfect knowledge of the internal dynamics of the black hole. In this paper, we show that for t – doped Clifford black holes - that is, black holes modeled by random Clifford circuits doped with an amount t of non-Clifford resources - an information retrieval decoder can be learned with fidelity scaling as exp(−αt) using quantum machine learning while having access only to out-of-time-order correlation functions. We show that the crossover between learnability and non-learnability is driven by the amount of non-stabilizer present in the black hole and sketch a new approach to quantum complexity.

Introduction.— The onset of chaos is at the root of the explanation of thermalization in closed quantum systems[2–6]. Although the precise definition of quantum chaos remains elusive, one can usefully refer to it as a bundle of features comprising information scrambling[7–11], complex pattern of entanglement[12–16], universal behavior of out-of-time-order correlation functions (OTOCs)[17–28] and quantum dynamics that can be efficiently modeled by random unitary operators[19, 29–32]. Information scrambling is the quantum analogue of the butterfly effect, signaling that local disturbances are spread through operator growth[33–35].

In the context of black hole physics, one wonders whether information is destroyed by a black hole or can be recovered from Hawking radiation as the black hole evaporates. It is commonly assumed that a black hole thermalizes information quickly[2, 3, 36] and that information is also rapidly spread across every part of the system, making the black hole a fast scrambler[37–40]. Moreover, one assumes evolution to be described by a unitary operator $U$. Under these assumptions, Hayden and Preskill showed that the Hawking radiation releases information very quickly, and that one can therefore recover any information initially dumped into the black hole with just a slight overhead of information read-out from the subsequent Hawking radiation[1]. This remarkable result hinges on the fact that it is the very scrambling dynamics of the black hole that allows information to be transferred to the Hawking radiation. Yoshida and Kitaev showed how this information can be retrieved by an observer with perfect knowledge of both the initial state of the black hole and its unitary dynamics $U[41]$.

Obtaining perfect knowledge of a black hole’s internal structure and dynamics sounds like an impossibly daunting task. Could we perhaps learn $U$ by feeding the Hawking radiation into a suitable quantum machine learning (QML) algorithm[42–62]? Such an approach seems promising at first glance, but quickly becomes futile in practice; in order for the recovery algorithm to work, the black hole must be scrambling, but the supposed complexity of scrambling dynamics hinders our ability to learn about its details. Indeed, extensive analysis[63–71] of the barren plateau problem—i.e. system size-exponential vanishing of cost function gradients in variational quantum algorithms—has shown that no QML training protocol can effectively learn $U$ if it is scrambling.

And yet, even so things are not hopeless. We cannot learn $U$, but we can train a quantum circuit to decode it by learning a model unitary $V$ that is good enough to unscramble the Hawking radiation, recover the information tossed into the black hole, and perform teleportation. Despite being very different from the original scrambler $U$, this so-called mocking black hole $V$ is optimized to perform the desired tasks.

In this paper, we show that if a black hole is modeled by a unitary $U_t$ consisting of an element of the Clifford group with a doping $t$ of non-Clifford resources, one can use a cost function directly obtained from the OTOCs to learn a mocking $V$ capable of recovering information from the black hole with fidelity $\mathcal{F}(V) \sim \exp(-\alpha t)$. This is possible because, while the Clifford group is a good scrambler[72–74], it does not produce a complex pattern of entanglement across the subsystems[75, 76]. We also show that, with the same technique, one can teleport a state with zero initial knowledge of black hole dynamics.

Decoding a Black Hole.— Let us start by reviewing the decoding algorithm for information retrieval from a scrambling black hole[1, 41]. After a long-time evaporation process, the state of the black
hole $B$ is maximally entangled with the already emitted Hawking radiation $B'$, which can be accessed by the observer Bob. This state is described by an EPR pair $|BB'\rangle$. Alice possesses a quantum state $A$ maximally entangled with a reference state $R$, that is an EPR pair $|RA\rangle$. The maximal entanglement between $A$ and $R$ quantifies the information about $A$ possessed by $R$. This information is lost when the black hole destroys correlations between $A$ and $R$.

The initial state of the whole system is thus described by $|RA\rangle|BB'\rangle$, where $|\Lambda\Lambda'\rangle := (2^n)^{-1/2}\sum_{i} |i\rangle_{\Lambda} |i\rangle_{\Lambda'}$ where $\Lambda, \Lambda'$ are two sets of $n$ qubits and $\Pi_{\Lambda\Lambda'}$ the corresponding projector. The black hole dynamics are modeled by a unitary $U = I_{RB'} \otimes U_{AB}$ which must be random enough to scramble information. After the evolution $U$, the state of the entire system is given by the pure state $|\Psi_{RB'CD}\rangle = I_{RB'} \otimes U_{AB} |RA\rangle|BB'\rangle$ in $RB'CD$, see Fig.2. Notice that the number $n$ of qubits initially internal to the black hole obeys $n = n_A + n_B = n_C + n_D$, so the Hilbert space of the black hole interior is shrinking as more Hawking radiation is emitted.

The information possessed by Alice in her journal would be recovered by Bob if he could process the Hawking radiation and end up with a state maximally entangled with another reference state $R'$ in his possession. In the case of Alice’s state being a pure state $|\psi_A\rangle$, this would amount to teleporting $|\psi_A\rangle$ from $A$ to $R'$. As [1] showed, the existence of such a recovery procedure is contingent on the unitary $U$ being scrambling enough. Because of unitarity, correlations can only be transferred, so if $U$ is scrambling enough to destroy correlations between $R$ and $C$, these must show up somewhere else and can be decoded by a suitable $V^\ast$. The information shared between the qubits of Alice $R$ and the internal degrees of freedom of the black hole $C$ is given by the mutual information $I(R|C) := S(\rho_B) + S(\rho_C) - S(\rho_{BC})$. One can easily see [7] that for the state $|\Psi\rangle_{RB'CD}$ one has


\[ I(R(C)) = n_A + n_C - S(\rho_{RC}) \]

Define the OTOC as \( \Omega(U) := d^{-1} \langle tr(P_A P_D(U) P_A P_D(U)) \rangle_{P_A, P_D} \), then, the mutual information can be related to the OTOCs \( \Omega(U) \) by [7]

\[ S_2(\rho_{RC}) = -\log \frac{d_A}{d_C} \Omega(U) \]  

(1)

In the above formulae, \( P_{\alpha_n} \) is the Pauli group on \( n_A \) qubits with \( \alpha \in \{A, B, C, D\} \), \( P_D(U) \equiv U^i P_D U^i \) and \( \langle \cdot \rangle_{P_A} \equiv \frac{1}{d_A} \sum_{P \in P_{\alpha_A}} \langle \cdot \rangle \) represents the group average.

A unitary \( U \) is said to be scrambling iff \( \Omega(U) \simeq d_A^{-2} + d_B^{-2} - (d_A d_D)^{-2} \), where \( \simeq \) means up to an order \( d^{-1} \). In the limit of \( d_A \ll d_B \) and \( d_C \ll d_D \), the scrambling dynamics will therefore imply that \( I(R(C)) \simeq 0 \), thus resulting in the destruction of correlations between \( R \) and \( C \).

Let us see how Bob is able to recover the information initially possessed by Alice. Bob possesses an EPR pair \( A'R' \) and applies a unitary \( V^* \) to the old Hawking radiation \( B' \) and one half of his pair \( A' \). Then, by reading the additional Hawking radiation \( D \), effectively entangling it in another EPR pair \( DD' \), Bob projects by \( \Pi_{DD'} \) onto \( DD' \) to obtain a final state

\[ |\Psi_{out}(t)\rangle = \frac{1}{\sqrt{d}} \Pi_{DD'} V_{B'A}^U A_B |RA\rangle |BB'\rangle |A'R'\rangle, \]

where \( P_{out} \) is a normalization [77]. The recovery algorithm is successful if one obtains a maximally entangled pair between \( R \) and \( R' \) that is factorized from the rest. In this way, the information about \( A \) originally in the hands of Alice has been successfully transferred to Bob [1]. In other words, the final state must read \( |\Psi_{out}(t)\rangle \simeq |RR'\rangle \otimes |rest\rangle_{CC} \otimes |DD'\rangle \).

Success of the algorithm can be established by computing the fidelity \( F := \text{tr}(\Psi_{out}(t) \Pi_{RR'}) \) where

\[ |\Psi_{out}(t)\rangle \equiv \text{tr}_{CC} \Pi_{DD'} |\Psi_{out}(t)\rangle. \]

Using techniques similar to those in [41, 78, 79], we obtain [77]:

\[ F(V) = \frac{1}{d_A} \frac{\langle tr(P_D(U)P_D(V)) \rangle_{P_D}}{\langle tr(P_D(U)P_{PA} P_D(V) P_{PA}) \rangle_{P_A, P_D}} \]

(2)

In the above expression, the role of \( V \) is to mock the behavior of the black hole modeled by \( U \). If \( V = U \), the behavior of \( U \) is obviously replicated perfectly. Indeed, the result in [41] shows that, in the ideal case \( V = U \), the fidelity \( F(V) \) reads

\[ F(U) = \frac{1}{d_A^2} \Omega(U) \]

and one can see that if the black hole is indeed scrambling, one obtains a fidelity \( F(U) = 1 - O(d_A^2/d_B^2) \). One can easily see [78] that

\[ \text{tr}(P_D(U)P_A P_D(V)P_A) = \langle \Psi_{out}(t) \rangle_{P_A} P_D \otimes P_D \]

where \( |\Psi_{out}(U, V)\rangle := \langle uu \otimes vv | P_A \otimes 1_{RR'} |BB'\rangle |RA\rangle |BB'\rangle |A'R'\rangle \). Thus, \( F(V) \) is a quantity that can be recast as a ratio of expectation values of local observables with supports on accessible parts of the system.

**Learning the Black Hole decoder.**— Can we learn the Black hole by observing its Hawking radiation by some quantum machine learning algorithm? The presence of barren plateaus [63–70] seems to forbid any kind of learning if the black hole is scrambling enough, which is also the condition that allows information retrieval.

The main result of this paper is that - under suitable conditions - one can nevertheless learn a decoder \( V \) that mocks the behavior of the black hole sufficiently enough to recover information.

\[ \text{Figure 3: Average over 200 realizations of the fidelity } F(V) \text{ at the end of the Metropolis Algorithm as a function of the doping } t, \text{ fit to } A\exp(-\alpha t) + B \text{ ( } \alpha = 0.167, 0.129 \text{ for } N_C = 1, 2 \text{ respectively). Each realization is performed for a newly generated black hole unitary } U. \text{ Inset: Same for the teleportation fidelity } F_\psi, \text{ with the same values for } \alpha \text{ [77].} \]

Consider the case that the black hole is modeled by a \( t \)-doped random Clifford circuit \( U \in C_t := \{ C_1 C_2 C_1 \cdots T_2 C_2 T_1 | C_k \in C(2^n) \} \) with \( T_k \) being single qubit \( T \)-gates applied on random qubits [75]. Such operators are scrambling enough [75, 80] for every \( t \). However, for \( t = 0 \) - i.e. random Clifford operators - a learning protocol based on a Metropolis algorithm is possible [12, 81, 82]. Moreover, it has been recently shown that low \( t \)-doping random Clifford circuits can be efficiently disentangled by annealing through a suitable Metropolis algorithm [16]. In the following, we show a similar Metropolis algorithm based on the cost function \( F(V) \equiv (1 + c(U, V))^{-1} \) to learn the mocking operator \( V \).

From the quantum machine learning point of view, we describe the procedure as follows. To learn \( V \), we first train our estimator \( c(U, V) \) by preparing a number \( O(\alpha^2) \) of copies of a known
state. This state is used as the input of a black hole. Starting with the identity $V_0 = I$, one tries the cost function $c(U,V)$ and optimizes it by local modifications using a Clifford gate $g$, so that $V_0 \mapsto gV_0$. Acceptance of the gate depends on improving the cost function $c(U,V)$ or lowering with Boltzmann probability parametrized by $\beta$.[77] The many copies of the training journal tossed in the black hole (Fig.1) correspond to the steps of the Metropolis algorithm needed. After the training, one has settled on a $V$ that minimizes the cost function $c(U,V)$ with the number of steps allowed. One can then use $V$ in the recovery algorithm, and recover information about a new journal tossed into $U$ with fidelity $F(V)$.

The numerical simulation of the whole algorithm is computationally very expensive. In order to perform the numerical simulations on a smaller space, we can exploit the fact that we are working in the scenario $d_C \ll d_D$. By averaging over the Pauli group on $C$ instead of $D$, one can in fact prove[77] that the cost function $c(U,V)$ can be computed as

$$c(U,V) = \frac{\sum_{P_C, P_A \neq 1_A} | \text{tr}((U^\dagger P_C V P_A))|^2}{\sum_{P_C} | \text{tr}((U^\dagger P_C V))|^2} \tag{3}$$

The simulations are conducted for a system of $n = 10$ qubits initialized in the state $|0\rangle^\otimes n$, with $n_A = 2$ and $n_C = 1, 2$. $U$ is a random Clifford circuit consisting of $t$ layers each with $O(n^2)$ local gates CNOT, $H$, $P$ interspersed by a single $T$ gate per layer[16, 75].

The results are shown in Fig.3. As we can see, for $t = 0$ a mocking operator $V$ for a black hole modeled by a pure random Clifford circuit can be learned with perfect fidelity. As the number $t$ of non-Clifford resources increases, the fidelity for the mocking operator $V$ decreases exponentially in $t$. Notice that the fidelity also depends on the size $n_C$ of the interior of the black hole. The fact that we are not learning the operator $U$ can be checked by comparing $V$ with $U$. Indeed, this would be forbidden by the barren plateau result found in[63]. We find that $d^{-1} |\text{tr} U^\dagger V|^2 < 0.04$ for every value of $t$, including $t = 0$. It is important to remark that we only try to reconstruct the mocking $V$ using merely Clifford resources, in spite of the fact that the original $U$ also contains non-Clifford resources. Although searching for $V$ within the Clifford group means we cannot reconstruct perfectly $U$, the learning algorithm fails with the addition of non-Clifford resources. For instance, for $n_T = 6, n_A = 2, n_C = 1$, the fidelity obtained using only Clifford gates attains a value of .53 versus $A$ when allowing the inclusion of $T$ gates. Universal resources are powerful, but pollute the algorithm[16].

The same recovery algorithm of [78] can be used to employ the recovery unitary $V$ to perform quantum teleportation between Alice and Bob. We now show that the previous learning of the mocking $V$ can be efficiently used to perform teleportation without having any previous knowledge of $U$. Let $|\psi_A\rangle$ be the pure state of Alice to be teleported. The fidelity to teleport a state $|\psi_A\rangle$ can be computed as[77]

$$F_\psi = \frac{\sum_{P_C} \text{tr}[|\psi_A\rangle \langle \psi_A| \text{tr}_B(U^\dagger P_C U)]}{\sum_{P_C} \text{tr}[|\psi_A\rangle \langle \psi_A| \text{tr}_B(U^\dagger P_C U) \text{tr}_B(U^\dagger P_C V)]} \tag{4}$$

In the inset of Fig.3 we plot the value $F_\psi$ as a function of the doping $t$. Again, perfect teleportation is achieved for a Clifford black hole ($t = 0$) while the fidelity $F_\psi$ decreases as non-Clifford resources are employed.

Quantum Complexity.— We have seen that the Clifford black hole can thus be perfectly learned - again, in the sense that one can learn the mocking operator $V$ - while this learning becomes less and less reliable with the injection of non-Clifford resources. Why is that? This is another instance of the fact[75, 83] that quantum complexity arises when scrambling (that is, efficient entanglement) is combined with non-stabilizerness, or magic[83–87]: the resource that is at the root of quantum advantage for quantum computers and the non-simulability of generic quantum systems by classical computers[88, 89].

While Clifford circuits are just as efficient in scrambling as a Haar-random unitary, they do not create a complex pattern of entanglement, and the fluctuations of entanglement are very different[75]. Quantum complexity is driven by the conspiracy of entanglement and non-stabilizerness (magic), or, in other words, by the complexity of entanglement[12, 13]. In particular, the ensemble fluctuations of the OTOCs defined as $\Delta \Omega_t := \langle \Omega(U)^2 \rangle_{U \in C_t} - \langle \Omega(U) \rangle_{U \in C_t}^2$ behave very differently, and it is not surprising that they are governed by the 8-OTOC, which probes more fine grained properties of scrambling[75], see [77]:

$$\Delta \Omega_t \approx \frac{1}{d_A d_D^3} \left[ (\text{otocs}(U))_{U \in C_t} + O(d^{-2}) \right] \tag{5}$$

Using the techniques introduced in [75, 76], one can compute Eq. (5) and find $\Delta \Omega_t \approx d_A^2 d_D^3 (\tfrac{1}{4})^t$, which interpolates between $\Delta \Omega_t = O(d_A^2)$ and $O(d_D^3)$ for $t = O(1), O(n_D)$, respectively. The relative fluctuations for the cost function $F^{-1}$ for $U = V$ are very therefore small, revealing a bar-
The internal dynamics of the black hole cannot be resolved even by a quantum machine learning algorithm, so information retrieval from a black hole through Hawking radiation seems hopeless. However, we have shown that one can learn a mocking-unitary that is capable of unscrambling and decoding the Hawking radiation by a Metropolis algorithm based on the observation of out-of-time-order correlation functions. The learning is possible for black holes that can be modeled by slightly polluted Clifford circuits. Highly polluted black holes cannot be learned. This result illustrates the crossover from simpler to complex quantum behavior and demonstrates how the onset of quantum chaos is driven by the conjunction of entanglement and non-stabilizer states. In view of the results on quantum certification in [90], it would be interesting to show that the same intractability also shows up in the amount of measurements needed to reliably measure the OTOCs in order to effectively perform information retrieval.

Acknowledgments.— The authors acknowledge support from NSF award number 2014000. The work of L.L. and S.F.E.O. was supported in part by College of Science and Mathematics Dean’s Doctoral Research Fellowship through fellowship support from Oracle, project ID R20000000025727.

Lorenzo.Leone001@umb.edu

1. P. Hayden and J. Preskill, Black holes as mirrors: quantum information in random subsystems, Journal of High Energy Physics 2007(09), 120 (2007), doi: 10.1088/1126-6708/2007/09/120
2. S. Lloyd, Black Holes, Demons and the Loss of Coherence: How complex systems get information, and what they do with it, Ph.D. thesis, Rockefeller University (1988).
3. M. Srednicki, Chaos and quantum thermalization, Physical Review E 50, 888 (1994), doi: 10.1103/PhysRevE.50.888.
4. N. Linden, S. Popescu et al., Quantum mechanical evolution towards thermal equilibrium, Physical Review E 79, 061103 (2009), doi: 10.1103/PhysRevE.79.061103.
5. S. Popescu, A. J. Short and A. Winter, Entanglement and the foundations of statistical mechanics, Nature Physics 2(11), 754 (2006), doi: 10.1038/nphys444.
6. M. Rigol and L. F. Santos, Quantum chaos and thermalization in gapped systems, Physical Review A 82, 011604 (2010), doi: 10.1103/PhysRevA.82.011604.
7. P. Hosur, X.-L. Qi et al., Chaos in quantum channels, Journal of High Energy Physics 2016(2), 4 (2016), doi: 10.1007/JHEP02(2016)004.
8. D. Ding, P. Hayden and M. Walter, Conditional mutual information of bipartite unitaries and scrambling, Journal of High Energy Physics 2016(12), 145 (2016), doi: 10.1007/JHEP12(2016)145.
9. W. Brown and O. Fawzi, Scrambling speed of random quantum circuits, doi: 10.48550/ARXIV.1210.6644 (2012).
10. B. Yan, L. Cincio and W. H. Zurek, Information scrambling and Loschmidt echo, Physical Review Letters 124, 160603 (2020), doi: 10.1103/PhysRevLett.124.160603.
11. A. Touil and S. Dettmer, Quantum scrambling and the growth of mutual information, Quantum Science and Technology 5(3), 035005 (2020), doi: 10.1088/2058-9565/ab8ebb.
12. C. Chamon, A. Hamma and E. R. Muccillo, Emergent irreversibility and entanglement spectrum statistics, Physical Review Letters 112, 240501 (2014), doi: 10.1103/PhysRevLett.112.240501.
13. Z.-C. Yang, A. Hamma et al., Entanglement complexity in quantum many-body dynamics, thermalization, and localization, Physical Review B 96, 020408 (2017), doi: 10.1103/PhysRevB.96.020408.
14. Z.-W. Liu, S. Lloyd et al., Generalized entanglement entropies of quantum designs, Physical Review Letters 120, 130502 (2018), doi: 10.1103/PhysRevLett.120.130502.
15. Z.-W. Liu, S. Lloyd et al., Entanglement, quantum randomness, and complexity beyond scrambling, Journal of High Energy Physics 2018(7), 41 (2018), doi: 10.1007/JHEP07(2018)041.
16. S. True and A. Hamma, Transitions in entanglement complexity in random circuits, doi: 10.48550/ARXIV.1802.02648 (2018).
17. A. I. Larkin and Y. N. Ovchinnikov, Quasiclassical method in the theory of superconductivity, J. Exp. Theor. Phys. 28, 1200 (1969).
18. A. Kitaev, Hidden correlations in the Hawking radiation and thermal noise, In Talk given at the Fundamental Physics Prize Symposium, vol. 10 (2014).
19. D. A. Roberts and B. Yoshida, Chaos and complexity by design, Journal of High Energy Physics 2017(4), 121 (2017), doi: 10.1007/JHEP04(2017)121.
20. D. J. Luitz and Y. Bar Lev, Information propagation in isolated quantum systems, Physical Review B 96, 020406 (2017), doi: 10.1103/PhysRevB.96.020406.
21. K. Hashimoto, K. Murata and K. Yoshii, Out-of-time-order correlators in quantum mechanics, Journal of High Energy Physics 2017(10), 138 (2017), doi: 10.1007/JHEP10(2017)138.
22. N. Yunger Halpern, Jarzynski-like equality for the out-of-time-ordered correlator, Physical Review A 95, 012120 (2017), doi: 10.1103/PhysRevA.95.012120.
23. I. García-Mata, M. Saraceno et al., Chaos signatures in the short and long time behavior of the out-of-time or-
[24] T. Rakovszky, F. Pollmann and C. W. von Keyserlingk, Diffusive hydrodynamics of out-of-time-ordered correlators with charge conservation, Physical Review X 8, 031058 (2018), doi:10.1103/PhysRevX.8.031058

[25] C.-J. Lin and O. I. Morrunich, Out-of-time-ordered correlators in a quantum Ising chain, Physical Review B 97, 144304 (2018), doi:10.1103/PhysRevB.97.144304

[26] E. M. Fortes, J. Garcia-Mata et al., Gauging classical and quantum integrability through out-of-time-ordered correlators, Physical Review E 100, 042201 (2019), doi:10.1103/PhysRevE.100.042201

[27] W. G. Brown, Random Quantum Dynamics: from Random Quantum Circuits to Quantum Chaos, Ph.D. thesis, Dartmouth College (2010).

[28] W. G. Brown and L. Viola, Convergence rates for arbitrary statistical moments of random quantum circuits, Physical Review Letters 104, 250501 (2010), doi:10.1103/PhysRevLett.104.250501

[29] H. Gharibyan, M. Hanada et al., Onset of random matrix behavior in scrambling systems, Journal of High Energy Physics 2018(7), 124 (2018), doi:10.1007/JHEP07(2018)124

[30] A. Nahum, S. Vijay and J. Haah, Operator spreading in random unitary circuits, Physical Review X 8, 021014 (2018), doi:10.1103/PhysRevX.8.021014

[31] L. D’ Alessio, Y. Kafri et al., From quantum chaos and eigenstate thermalization to statistical mechanics and thermodynamics, Advances in Physics 65(3), 239 (2016), doi:10.1080/00018732.2016.1198134

[32] Y. Sekino and L. Susskind, Fast scramblers, Journal of High Energy Physics 2008(10), 065 (2008), doi:10.1088/1126-6708/2008/10/065

[33] N. Lashkari, D. Stanford et al., Towards the fast scrambling conjecture, Journal of High Energy Physics 2013(4), 22 (2013), doi:10.1007/JHEP04(2013)022

[34] S. H. Shenker and D. Stanford, Multiple shocks, Journal of High Energy Physics 2014(12), 46 (2014), doi:10.1007/JHEP12(2014)046

[35] J. Maldacena, S. H. Shenker and D. Stanford, A bound on chaos, Journal of High Energy Physics 2016(8), 106 (2016), doi:10.1007/JHEP08(2016)106

[36] A. Arrasmith, L. Cincio et al., Quantum autoencoders for efficient compression of quantum data, Quantum Science and Technology 2(4), 045001 (2017), doi:10.1088/2058-9565/aa8072

[37] R. LaRose, A. Tikku et al., Variational quantum state diagonalization, npj Quantum Information 5(1), 57 (2019), doi:10.1038/s41534-019-0167-6

[38] A. Arrasmith, L. Cincio et al., Variational consistent histories as a hybrid algorithm for quantum foundations, Nature Communications 10(1), 3438 (2019), doi:10.1038/s41467-019-11417-0

[39] M. Cerezo, A. Perochnik et al., Variational Quantum Fidelity Estimation, Quantum 4, 248 (2020), doi:10.22331/q-2020-03-26-248

[40] K. Sharma, S. Khatri et al., Noise resilience of variational quantum compiling, New Journal of Physics 22(4), 045006 (2020), doi:10.1088/1367-2630/ab584c

[41] B. Yoshida and A. Kitaev, Efficient decoding for the Hayden-Preskill protocol, doi:10.48550/ARXIV.1710.03563 (2017).

[42] J. Biamonte, P. Wittek et al., Quantum machine learning, Nature 549(7671), 195 (2017), doi:10.1038/nature23474

[43] C. Ciliberto, M. Herbster et al., Quantum machine learning: a classical perspective, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences 474(2209), 20170551 (2018), doi:10.1098/rspa.2017.0551

[44] V. Dunjko and H. J. Briegel, Machine learning & artificial intelligence in the quantum domain: a review of recent progress, Reports on Progress in Physics 81(7), 074001 (2018), doi:10.1088/1361-6633/aab406

[45] S. Lloyd and C. Weedbrook, Quantum generative adversarial learning, Physical Review Letters 121, 040502 (2018), doi:10.1103/PhysRevLett.121.040502

[46] K. Poland, K. Beer and T. J. Osborne, No free lunch for quantum machine learning, doi:10.48550/ARXIV.2003.14103 (2020).

[47] S. Lloyd, M. Schuld et al., Quantum embeddings for machine learning, doi:10.48550/ARXIV.2001.05622 (2020).

[48] S. Lloyd, M. Schuld et al., Quantum embeddings for machine learning, doi:10.48550/ARXIV.2001.05622 (2020).

[49] J. R. McClean, J. Romero et al., The theory of variational hybrid quantum-classical algorithms, New Journal of Physics 18(2), 023023 (2016), doi:10.1088/1367-2630/18/2/023023.

[50] J. Romero, J. P. Olson and A. Aspuru-Guzik, Quantum autoencoders for efficient compression of quantum data, Quantum Science and Technology 2(4), 045001 (2017), doi:10.1088/2058-9565/aa8072

[51] R. LaRose, A. Tikku et al., Variational quantum state diagonalization, npj Quantum Information 5(1), 57 (2019), doi:10.1038/s41534-019-0167-6

[52] A. Arrasmith, L. Cincio et al., Variational consistent histories as a hybrid algorithm for quantum foundations, Nature Communications 10(1), 3438 (2019), doi:10.1038/s41467-019-11417-0

[53] M. Cerezo, A. Perochnik et al., Variational Quantum Fidelity Estimation, Quantum 4, 248 (2020), doi:10.22331/q-2020-03-26-248

[54] K. Sharma, S. Khatri et al., Noise resilience of variational quantum compiling, New Journal of Physics 22(4), 045006 (2020), doi:10.1088/1367-2630/ab584c

[55] C. Bravo-Prieto, R. LaRose et al., Variational quantum linear solver, doi:10.48550/ARXIV.1909.05820 (2019).

[56] M. Cerezo, K. Sharma et al., Variational quantum state eigensolver, doi:10.48550/ARXIV.2004.01372 (2020).

[57] B. Commeau, M. Cerezo et al., Variational hamiltonian diagonalization for dynamical quantum simulation, doi:10.48550/ARXIV.2009.02559 (2020).

[58] Y. Li and S. C. Benjamin, Efficient variational quantum simulator incorporating active error minimization, Phys. Rev. X 7, 021050 (2017), doi:10.1103/PhysRevX.7.021050.

[59] X. Yuan, S. Endo et al., Theory of variational quantum
S. F. Oliviero, L. Leone and A. Hamma, 

K. Sharma, M. Cerezo et al., Trainability of dissipative perception-based quantum neural networks, doi:10.48550/ARXIV.2005.12438 (2020).

T. Volkoff and P. J. Coles, Large gradients via correlation in random parameterized quantum circuits, Quantum Science and Technology 6(2), 025008 (2021), doi:10.1088/2058-9565/abd891.

K. Endo, T. Nakamura et al., Quantum self-learning monte carlo and quantum-inspired fourier transform sampler, Phys. Rev. Research 2, 043442 (2020), doi:10.1103/PhysRevResearch.2.043442.

Z. Holmes, A. Arrasmith et al., Barren plateaus preclude learning scramblers, Phys. Rev. Lett. 126, 190501 (2021), doi:10.1103/PhysRevLett.126.190501.

J. R. McClean, S. Boixo et al., Barren plateaus in quantum neural network training landscapes, Nature Communications 9(1), 4812 (2018), doi:10.1038/s41467-018-07090-4.

M. Cerezo, A. Sone et al., Cost function dependent barren plateaus in shallow parameterized quantum circuits, Nature Communications 12(1), 1791 (2021), doi:10.1038/s41467-021-21728-w.

S. Wang, E. Fontana et al., Noise-induced barren plateaus in variational quantum algorithms, Nature Communications 12(1), 6961 (2021), doi:10.1038/s41467-021-27045-6.

M. Cerezo and P. J. Coles, Higher order derivatives of quantum neural networks with barren plateaus, Quantum Science and Technology 6(3), 035006 (2021), doi:10.1088/2058-9565/abf51a.

A. Arrasmith, M. Cerezo et al., Effect of barren plateaus on gradient-free optimization, Quantum 5, 558 (2021), doi:10.22331/q-2021-10-05-558.

E. Grant, L. Wossnig et al., An initialization strategy for addressing barren plateaus in parameterized quantum circuits, Quantum 3, 214 (2019), doi:10.22331/q-2019-12-09-214.

E. R. Anschuetz and B. T. Kiani, Beyond barren plateaus: Quantum variational algorithms are swamped with traps, doi:10.48550/ARXIV.2205.05786 (2022).

R. J. García, C. Zhao et al., Barren plateaus from learning scramblers with local cost functions, doi:10.48550/ARXIV.2205.06679 (2022).

A. J. Scott, Optimizing quantum process tomography with unitary 2-designs, Journal of Physics A: Mathematical and Theoretical 41(5), 055308 (2008), doi:10.1088/1751-8113/41/5/055308.

H. Zhu, Multiqubit Clifford groups are unitary 3-designs, Physical Review A 96, 062336 (2017), doi:10.1103/PhysRevA.96.062336.

H. Zhu, R. Kueng et al., The clifford group fails gracefully to be a unitary 4-design, doi:10.48550/ARXIV.1609.08172 (2016).

L. Leone, S. F. E. Oliviero et al., Quantum Chaos is Quantum, Quantum 5, 453 (2021), doi:10.22331/q-2021-05-04-453.

S. F. Oliviero, L. Leone and A. Hamma, Transitions in entanglement complexity in random quantum circuits by measurements, Physics Letters A 418, 127721 (2021), https://doi.org/10.1016/j.physleta.2021.127721.

See Supplemental Material for formal proofs, which includes[91–96].

B. Yoshida and N. Y. Yao, Disentangling scrambling and decoherence via quantum teleportation, Physical Review X 9, 011006 (2019), doi:10.1103/PhysRevX.9.011006.

R. Li and J. Wang, Hayden-preskill protocol and decoding hacking radiation at finite temperature (2021), doi:10.48550/ARXIV.2108.09144.

B. Yoshida, Recovery algorithms for clifford hayden-preskill problem (2021), doi:10.48550/ARXIV.2106.15628.

D. Shaffer, C. Chamon et al., Irreversibility and entanglement spectrum statistics in quantum circuits, Journal of Statistical Mechanics: Theory and Experiment 2014(12), P12007 (2014), doi:10.1088/1742-5468/2014/12/p12007.

S. Zhou, Z.-C. Yang et al., Single T gate in a Clifford circuit drives transition to universal entanglement spectrum statistics, Science Physics 9, 87 (2020), doi:10.21468/SciPostPhys.9.6.087.

L. Leone, S. F. E. Oliviero and A. Hamma, Stabilizer Rényi entropy, Physical Review Letters 128, 050402 (2022), doi:10.1103/PhysRevLett.128.050402.

E. T. Campbell and D. E. Browne, Bound states for magic state distillation in fault-tolerant quantum computation, Physical Review Letters 104, 030503 (2010), doi:10.1103/PhysRevLett.104.030503.

V. Veitch, S. A. H. Mousavian et al., The resource theory of stabilizer quantum computation, New Journal of Physics 16(1), 013009 (2014), doi:10.1088/1367-2630/16/1/013009.

E. T. Campbell, H. Anwar and D. E. Browne, Magic-state distillation in all prime dimensions using quantum reed-muller codes, Physical Review X 2, 041021 (2012), doi:10.1103/PhysRevX.2.041021.

M. Howard and E. Campbell, Application of a resource theory for magic states to fault-tolerant quantum computing, Physical Review Letters 118, 090501 (2017), doi:10.1103/PhysRevLett.118.090501.

E. Knill and R. LaFlamme, Theory of quantum error-correcting codes, Physical Review A 55, 900 (1997), doi:10.1103/PhysRevA.55.900.

S. Aaronson and D. Gottesman, Improved simulation of stabilizer circuits, Physical Review A 70, 052328 (2004), doi:10.1103/PhysRevA.70.052328.

L. Leone, S. F. E. Oliviero and A. Hamma, Magic hinders quantum certification, doi:10.48550/ARXIV.2204.09295 (2022).

B. Yoshida, Decoding the entanglement structure of monitored quantum circuits (2021), doi:10.48550/ARXIV.2109.08691.

V. Strassen, Gaussian elimination is not optimal, Numerische Mathematik 13(4), 354 (1969), doi:10.1007/BF02165411.

J. Alaman and V. V. Williams, A refined laser method and faster matrix multiplication, doi:10.48550/ARXIV.2010.05846 (2020).

B. Collins, Moments and cumulants of polynomial random variables on unitary groups, the Itzykson-Zuber integral, and free probability, International Mathematics Research Notices 2003(17), 953 (2003), doi:10.1155/S107379280320917X.
SUPPLEMENTAL MATERIAL

Appendix A: Proof of Eqs. (2) and (4)

To prove Eq. (2) we use the well-known techniques introduced in [41] and then reviewed in [78–80, 91]. The main idea is to represent states in the Hilbert space $\mathcal{H}_R \otimes \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_B' \otimes \mathcal{H}_A' \otimes \mathcal{H}_R'$ as graphs. Starting with

$$|\Psi_{\text{out}}\rangle = \frac{1}{\sqrt{P_{\text{out}}}} \Pi_{D^D} V_{B'A'} U_{AB} |AB\rangle |BB'\rangle |A'R'\rangle \equiv \frac{1}{\sqrt{P_{\text{out}}}} \Pi_{D^D} |\chi\rangle$$

we compute the fidelity with the state $|RR'\rangle$ by contracting graphs:

$$\mathcal{F}(V) P_{\text{out}} = \frac{1}{d_A^2} \left( \text{tr} P_D(U) P_D(V) \right)_{PD}$$

(A1)
where the normalization $P_{\text{out}}$ is still unknown. The computation of $P_{\text{out}}$ can be derived from the normalization of $|\Psi_{\text{out}}\rangle$ graphically as:

\[
P_{\text{out}} \equiv \langle \chi | \Pi_{DD'} \otimes \Pi_{RR'CC'} | \chi \rangle \equiv \frac{1}{d} \langle \text{tr}(P_D(U_P A_P) P_D(V_P A_P)) \rangle_{P_A, P_D}
\]

(A2)

Putting it all together leads to Eq. (2). To prove Eq. (4) we can proceed in a similar way, although the setup is slightly different. First, consider the initial state (output of the black hole evolution) to be $|\Phi_{B'CD'}\rangle := U_{AB} |\psi\rangle_A |BB'\rangle$. Bob first applies the decoder unitary $V^*_B A'$ on the state $|\Phi_{B'CD'}\rangle A'R'$ and then project it on $\Pi_{DD'}$. The output state reads:

\[
|\Phi_{\text{out}}\rangle = \frac{1}{\sqrt{P_{\psi}}} \Pi_{DD'} V^*_B A' |\Phi_{B'CD}\rangle A'R' = \frac{1}{\sqrt{P_{\psi}}} \langle \text{tr}[\psi A \text{tr}B(V^*_B U_P C_U)] \text{tr}[\psi A \text{tr}B(U^*_B P_C V_U)] \rangle_{P_C}
\]

(A3)

where $P_{\psi}$ is a normalization factor. Using the graphical representation of $|\Phi_{\text{out}}\rangle$, one computes:

\[
\mathcal{F}_{\psi} P_{\psi} = \frac{1}{d_B} \left( \langle \text{tr}[\psi A \text{tr}B(V^*_B P_C U)] \rangle \text{tr}[\psi A \text{tr}B(U^*_B P_C V)] \right)_{PC}
\]

(A4)
and

\[ P_0 \equiv \begin{array}{c|ccc|c} \hline \psi \\ \hline A & B & D & A* \\ U^\dagger & C & D & \hline \end{array} \]

\[ = \frac{1}{d_B} \left( \text{tr}[\psi_A \text{tr}_B(V_U^\dagger P_C U) \text{tr}_B(U^\dagger P_C V_U)] \right)_{P_C} \quad (A5) \]

taking the ratio between Eq. (A4) and Eq. (A5) one obtains Eq. (4).

**Appendix B: Metropolis algorithm**

First let us write Eq. (2) in a smaller space, proving Eq. (3). Consider the denominator of Eq. (2); using \( \text{tr}(O_2) \equiv \text{tr}(T O_{\otimes 2}) \) where \( T \) is the swap operator defined on \( \mathcal{H}_{\otimes 2} \), we can rewrite this term as:

\[ \langle \text{tr}(P_D(U)P_D(V)) \rangle_{P_D,P_D} = \frac{1}{d_D d_A} \text{tr}(V^\dagger \otimes U^\dagger T_C U \otimes V T_A) = \frac{1}{d_D d_A} \sum_{P_C,P_A} |\text{tr}(U^\dagger P_C V)|^2 \quad (B1) \]

using the fact that \( \sum_{P_\Lambda} P_\Lambda^{\otimes 2} = d_\Lambda T_\Lambda \) where \( T_\Lambda \) is the swap operator having support on \( \Lambda \) and \( TT_\Lambda = T_\bar{\Lambda} \) where \( \bar{\Lambda} \) is the complement of \( \Lambda \). Similarly for the numerator of Eq. (2):

\[ \langle \text{tr}(P_D(U)P_D(V)) \rangle_{P_D} = \frac{1}{d} \sum_{P_C} |\text{tr}(U^\dagger P_C V)|^2 \quad (B2) \]

Putting it all together, we find:

\[ \mathcal{F}(V) = (1 + c(U, V))^{-1} \quad (B3) \]

where \( c(U, V) \) is defined in Eq.(3). Note that \( c(U, V) \geq 0 \) with \( c(U, V) = 0 \) if and only if \( \mathcal{F}(V) = 1 \), making \( c(U, V) \) the best candidate for a cost function. Let us describe the algorithm to numerically find an optimal recovery mocking unitary \( V \):

**Algorithm 1 Training Algorithm**

1: \( U \leftarrow \) random doped Clifford circuit \( \mathcal{C}_t \)
2: \( \mathcal{F}_{\text{max}} \leftarrow \mathcal{F}(U) \)
3: \( V \leftarrow I \)
4: \( \mathcal{F} \leftarrow \mathcal{F}(V) \)
5: while \( \mathcal{F} < \mathcal{F}_{\text{max}} \) do
6: \( \mathcal{V}_{\text{old}} \leftarrow V \)
7: \( \mathcal{F}_{\text{old}} \leftarrow \mathcal{F} \)
8: \( g \leftarrow \) random Clifford gate
9: \( V \leftarrow gV \)
10: \( \mathcal{F} \leftarrow \mathcal{F}(V) \)
11: if \( \mathcal{F} < \mathcal{F}_{\text{old}} \) then
12: \( r \in [0,1] \)
13: if \( r < \exp \left[-\beta( \mathcal{F}_{-1} - \mathcal{F}_{\text{old}}^{-1}) \right] \) then
14: \( \text{Undo change: } V \leftarrow \mathcal{V}_{\text{old}} \)
15: \( \mathcal{F} \leftarrow \mathcal{F}_{\text{old}} \)
16: end if
17: end if
18: end while
A constant cooling schedule is employed with $\beta = 250$. In each iteration of the \textit{while} loop, several $2^n \times 2^n$ matrix multiplications must be performed to obtain the operator $UP_UV^\dagger P_A$. These operations quickly become computationally expensive as we increase the system size; even the fastest algorithms for this task \cite{92, 93} would require a computation time that grows exponentially in $n$. Furthermore, to avoid indefinite runtimes when the gradient of $\mathcal{O}(U, V)$ becomes small, we halt the algorithm after $T_{\max}$ steps. The value of $T_{\max}$ used is $100n^2$ when $N_C = 1$ and $300n^2$ when $N_C = 2$, chosen empirically by observing the algorithm’s runtime when $t = 0$. Note that the time requirement grows with the number of qubits as $T_{\max} \propto n^2$, as seen also in \cite{12, 16, 81}. Therefore, the numerical analysis in this work is performed for systems of 10 qubits, as larger system sizes would require exponentially larger computational times.

The parameters for the results fit to $A \exp(-\alpha t) + B$ shown in Fig.3 are given below.

| $\alpha$ | $N_C = 1$ | $N_C = 2$ |
|----------|-----------|-----------|
| $A$      | 0.7243    | 0.8483    |
| $B$      | 0.2757    | 0.1517    |

| $\mathcal{F}_f^{(U,V)}$ | $A$ | $B$ |
|-------------------------|-----|-----|
| $\alpha$               | 0.5860 | 0.6794 |
| $\mathcal{F}_A^{(U,V)}$ | 0.4140 | 0.3206 |

Appendix C: The complexity of scrambling: the 8–OTOC

In this section we present a single correlation function which tells us whether the algorithm is going to fail or not, i.e. the 8–point out of time order correlator, proving Eq. (5). First, let us write the OTOC $\Omega(U) = \langle \text{tr}(P_A P_D(U) P_A P_D(U)) \rangle_{P_A, P_D}$ as:

$$\Omega(U) = \frac{1}{d_A^2} \frac{1}{d_D^2} + \frac{1}{d_A^2 d_D^2} \sum_{P_A, P_D \neq 1} \frac{1}{d} \text{tr}(P_A P_D(U) P_A P_D(U))$$  \hspace{5em} (C1)

we thus can alternately define a scrambling unitary $U$ as one such that:

$$f(U) = \frac{1}{d_A^2 d_D^2} \sum_{P_A, P_D \neq 1} \text{oto}_{4}(P_A, P_D(U)) = O(d^{-1})$$  \hspace{5em} (C2)

where we defined $\text{oto}_{4}(P_A, P_D(U)) := \frac{1}{d} \text{tr}(P_A P_D(U) P_A P_D(U))$. Note that for $U \in \mathcal{C}(2^n)$ being a Clifford operator $\text{oto}_{4}(P_A, P_D(U)) = \pm 1$ for any choice of $P_A, P_D$, but $f(U) = O(d^{-1})$ still holds for Clifford unitaries, i.e. Clifford unitaries are scramblers. Note also that

$$\langle f(U) \rangle_U = \frac{(d_A^2 - 1)(d_D^2 - 1)}{d_A^2 d_D^2} \langle \text{oto}_{4}(P_1, P_2(U)) \rangle_U$$  \hspace{5em} (C3)

for any $P_1, P_2$ chosen to be two non-identity Pauli operators and $U$ being a Clifford or $t$–doped Clifford operator. Now let us look at the fluctuations in the ensemble of $t$–doped Clifford operators $U \in \mathcal{C}_t$. Direct calculation leads to:

$$\Delta \Omega_t = \langle f(U)^2 \rangle_U - \langle f(U) \rangle_U^2 = \langle f(U)^2 \rangle_U + O(d^{-2})$$  \hspace{5em} (C4)

the second equality follows from the definition of scrambling unitaries – i.e. $\langle f(U) \rangle_U = O(d^{-1})$ cfr. Eq. (C2) – and it holds for any $t \in \mathbb{N}$. Thus, turn to analyze $\langle f(U)^2 \rangle_U$:

$$\langle f(U)^2 \rangle_U = \frac{(d_A^2 - 1)(d_D^2 - 1)}{d_A^2 d_D^2} \langle \text{oto}_{4}(P_1, P_2(U))^2 \rangle_U + \langle R(U) \rangle_U$$  \hspace{5em} (C5)
\[ R(U) := \frac{1}{d_A^4 d_D^4} \sum_{P_A, P_D, P'_A, P'_D \neq \mathbb{1}} [1 - \delta(P_A = P'_A)\delta(P_D = P'_D)] \times \text{otoc}_4(P_A, P_D(U)) \times \text{otoc}_4(P'_A, P'_D(U)) \]  

(C6)

Using standard techniques of the Haar measure over groups [75, 94–96] one proves that for any \( t \in \mathbb{N} \) we have \( \langle R(U) \rangle = O(d^{-2}) \). Lastly, using the fact that the Pauli group forms a 1-design [19], it is easy to prove [19] the following:

\[ \text{otoc}_4(P_1, P_2(U))^2 = \text{otoc}_8(U) \]  

(C7)

where we defined the following 8-point out of time order correlation function:

\[ \text{otoc}_8(U) := \frac{1}{d^4} \langle \text{tr}(P_1 P_2(U) P_1 P_2(U) P_1 P_2(U) P_1 P_2(U) P_1 P_2(U)) \rangle_{P \in \mathcal{P}_n} \]  

(C8)

where \( P_1 \) and \( P_2 \) are any non-identity Pauli operators. We can finally write the following formula:

\[ \Delta \Omega_t = \frac{(d_A^2 - 1)(d_D^2 - 1)}{d_A^4 d_D^4} \langle \text{otoc}_8(U) \rangle_U + O(d^{-2}) \]  

(C9)

in the limit of \( d_D \gg d_A \gg 1 \), one arrives to Eq. (5). Thus the ensemble fluctuations of the 4-point OTOC are proportional to the ensemble the average of an 8-point correlation function, which probes more fine grained properties of scrambling, i.e. the complexity of scrambling. We thus define the complexity of scrambling produced by some unitary operator \( U \) as the behavior of the 8-point correlation function defined in Eq. (C8).