Tight Bounds on Pauli Channel Learning without Entanglement

Senrui Chen,[1,*] Changhun Oh,[1,2] Sisi Zhou,[3,4] Hsin-Yuan Huang,[3,5] and Liang Jiang[1,†]
[1]*Pritzker School of Molecular Engineering, The University of Chicago, Chicago, Illinois 60637, USA*
[2]*Department of Physics, Korea Advanced Institute of Science and Technology, Daejeon 34141, Korea*
[3]*Institute for Quantum Information and Matter, California Institute of Technology, Pasadena, California 91125, USA*
[4]*Perimeter Institute for Theoretical Physics, Waterloo, Ontario N2L 2Y5, Canada*
[5]*Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*

(Received 7 December 2023; revised 8 April 2024; accepted 15 April 2024; published 1 May 2024)

Quantum entanglement is a crucial resource for learning properties from nature, but a precise characterization of its advantage can be challenging. In this Letter, we consider learning algorithms without entanglement to be those that only utilize states, measurements, and operations that are separable between the main system of interest and an ancillary system. Interestingly, we show that these algorithms are equivalent to those that apply quantum circuits on the main system interleaved with mid-circuit measurements and classical feedforward. Within this setting, we prove a tight lower bound for Pauli channel learning without entanglement that closes the gap between the best-known upper and lower bound. In particular, we show that $\Theta(2^n\varepsilon^{-2})$ rounds of measurements are required to estimate each eigenvalue of an n-qubit Pauli channel to ε error with high probability when learning without entanglement. In contrast, a learning algorithm with entanglement only needs $\Theta(\varepsilon^{-2})$ copies of the Pauli channel. The tight lower bound strengthens the foundation for an experimental demonstration of entanglement-enhanced advantages for Pauli noise characterization.

DOI: 10.1103/PhysRevLett.132.180805

Entanglement lies at the heart of quantum information science and technology, providing significant advantages over classical information processing in computation [1], communication [2,3], metrology [4–6], and many other aspects. A recent line of research uses information-theoretic tools to obtain rigorous and exponential quantum advantages in learning [7–15]. It is shown, both theoretically and experimentally, that quantum resources can bring significant speedup for learning certain properties from the nature, e.g., learning expectation values of many observables for a quantum state [7]. However, the connection between these quantum advantages with specific quantum resources, e.g., quantum entanglement, is far from clear. This problem is prominent for learning properties from quantum channels, where there are many different ways of defining a “quantum-enhanced” experiments [8,10,12,14], depending on whether one allows ancillary systems, concatenation of channels, mid-circuit controls, etc. A scenario that precisely captures the role of quantum entanglement in learning is under exploration.

Apart from studying learning schemes from a quantum resource-theoretic perspective [16], one can also take an operational approach. Specifically, a class of quantum operations known as mid-circuit measurement and classical feedforward have drawn increasing attention recently. While they are important building blocks of fault-tolerant quantum computation [17,18] and have found applications in recent experiments [19,20], a framework that allows systematic study of the effectiveness of mid-circuit measurements and classical feedforward in learning has yet to be established.

Turning to concrete learning tasks, a class of quantum channels that has drawn particular interest is the Pauli channel, which is defined to be a stochastic mixture of (multiqubit) Pauli operations [1]. The Pauli channel is not only a basic model in quantum information theory, but also plays a crucial role in characterizing noisy quantum systems, with applications in quantum benchmarking [21–23], quantum noise mitigation [24–26], quantum error correction [27], etc. Techniques such as randomized compiling can engineer general quantum noise into Pauli channel under realistic assumptions [28,29]. A prerequisite for many of these applications is to learn an unknown Pauli channel. Therefore, it is natural to study the protocols and limitations of Pauli channel learning. There have been several recent works exploring this direction [14,30–32]. Specifically, Ref. [14] studies the sample complexity of Pauli channel estimation using information-theoretic methods, and shows an exponential separation between using and not using ancilla for learning every eigenvalue of an n-qubit Pauli channel to $\pm\varepsilon$ precision. However, as shown by the current work, the ancilla-free lower bound given there, $\Omega(2^{n/3})$, was not tight. While an ideal ancilla-assisted protocol has sample complexity $\Theta(1/\varepsilon^2)$ [33], real-world imperfections such as state preparation and measurement (SPAM) noise can introduce a weak exponential sampling overhead

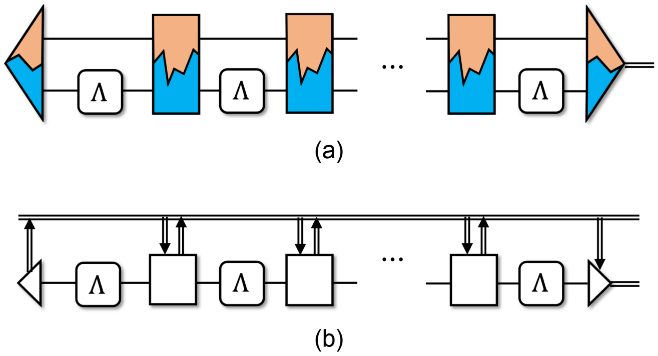

FIG. 1. (a) Separable schemes. The two different colors indicate the operations are separable. (b) Classical-memory-assisted schemes. The double line represents classical registers. The square box represents adaptively chosen (arrows coming from the classical registers) quantum instruments (outcomes sent to the classical registers). We will show the two schemes are equivalent in terms of sample complexity, so we call both entanglement-free schemes.

(see Fig. 2). Therefore, to establish an advantage with moderate system sizes and realistic levels of imperfection, it is highly desirable to tighten the ancilla-free lower bound.

In this Letter, we introduce a class of learning schemes that do not exploit entanglement between the main system and the ancillary system, from a resource-theoretic perspective. We also introduce a class of schemes that describe quantum circuits assisted with mid-circuit measurement and classical feedforward from an operational perspective (see Fig. 1). Perhaps surprisingly, we show the two schemes are equivalent in terms of sample complexity for any learning tasks. This provides a new operational interpretation for quantum entanglement as a resource [34–36]. We then show that information-theoretic methods can be used to prove sample complexity lower bound in the above scenario. For the task of Pauli channel learning mentioned above, we obtain a tight lower bound of $\Omega(2^n/\varepsilon^2)$, closing the cubic gap with the known upper bound, and providing a tight exponential separation with the entanglement-assisted bound of $\Theta(1/\varepsilon^2)$ [14]. We show that this separation persists even if the entanglement-assisted scheme suffers from a reasonable amount of realistic noise. Finally, we show how our results impose a limit on the efficiency of characterizing gate-dependent Pauli noise channels.

Setup.—Consider the task of learning properties of an n-qubit quantum channel Λ from certain family by querying multiple copies of the channel. We start by defining entanglement-free learning schemes. We introduce the class of "separable schemes," where a main system $\mathcal{H}_S$ and an ancillary system $\mathcal{H}_A$ are given. The main system is where Λ acts on and has a fixed dimension of 2^n, while the ancillary system can be arbitrarily large. Now, a separable scheme allows interleaving copies of Λ on $\mathcal{H}_S$ with any processing operations (including state preparation, measurement, and quantum channels) on $\mathcal{H}_S \otimes \mathcal{H}_A$, with the only restriction being that all the processing operations are separable [37] across $\mathcal{H}_A$ and $\mathcal{H}_S$. A schematic for separable schemes is shown in Fig. 1(a). The separable operations are known to be the largest set of operations that do not generate entanglement from unentangled states (even when acting only on a subspace) [37], and is thus a suitable model for entanglement-free strategies from a quantum resource-theoretic [16] point of view. Furthermore, separable operations contain as a subset other physically motivated classes of operations like local operation and classical communications [38,39], which is also a standard choice of free operation in the resource theory of entanglement [34–36]. A lower bound on sample complexity for the former implies a lower bound for the latter.

Besides separable schemes, we introduce another operationally motivated class of schemes called "classical-memory-assisted schemes." Here, one can only access the main system $\mathcal{H}_S$ (with a fixed dimension of 2^n) and arbitrarily many classical registers. The allowed operations are to interleave copies of Λ with adaptively chosen "quantum instruments," which are defined as quantum channels associated with outputs to classical registers. By "adaptively" we mean the quantum instruments can be chosen according to the classical registers. A schematic is given in Fig. 1(b). One can think of such schemes as quantum circuits assisted by mid-circuit measurement and classical feedforward control. We remark that the classical-memory-assisted schemes include the ancilla-free concatenating scheme introduced in [14] as a special case, which can describe most randomized benchmarking [40,41] type protocols. The lower bounds obtained in this work thus also hold for those protocols.

While the above two schemes are introduced with different motivations, perhaps surprisingly, they are equivalent in terms of sample complexity. We have the following result.

Proposition 1.—For any separable scheme A, there exists a classical-memory-assisted scheme B that generates the same outcome distribution as A for any underlying Λ using the same number of copies and vice versa.

The formal definitions of both schemes, rigorous statement of Proposition 1, and the proof are given in the Supplemental Material (SM), Sec. II [42]. Thanks to Proposition 1, we see that both schemes capture the power of learning without entanglement and can be treated interchangeably when studying the sample complexity. In the remaining part of this Letter, we will focus on the classical-memory-assisted scheme, which has a clearer operational meaning. Specifically, there can be two different notions of complexity: (1) the sample complexity N_{samp}, which is the number of copies of Λ, and (2) the number of measurements N_{meas}, which is the number of quantum instruments with nontrivial measurement. Clearly, $N_{\mathrm{samp}} \geq N_{\mathrm{meas}}$, as one can concatenate multiple copies of Λ and only make one measurement, just like in randomized benchmarking. The lower bound we derive later will hold for N_{meas}.

Bounds on Pauli channel learning.—Having set up the formalism, we now study the specific problem of Pauli channel learning. An n-qubit Pauli channel Λ has the following two equivalent forms:

$$\Lambda(\rho) \coloneqq \sum_{b\in\mathsf{P}^n} p_b P_b \rho P_b = \frac{1}{2^n}\sum_{a\in\mathsf{P}^n} \lambda_a P_a \mathrm{Tr}[P_a\rho], \quad (1)$$

where $\mathsf{P}^n = \{I, X, Y, Z\}^{\otimes n}$ is the n-qubit Pauli group (modulo phase), $\{p_b\}_b$ is the Pauli error rates, and $\{\lambda_a\}_a$ is the Pauli eigenvalues [30]. Note that Pauli eigenvalues are also known as Pauli fidelities, which have been useful in quantum benchmarking [21,22,29,45], quantum error mitigation [24,25,46], etc. The task we consider is to learn each of the Pauli eigenvalues λ_a to additive precision ε with high success probability. More precisely, we have the following result.

Theorem 1.—If there exists an entanglement-free scheme that, for any n-qubit Pauli channel Λ, outputs an estimator $\hat{\lambda}_a$ such that $|\hat{\lambda}_a - \lambda_a| \le \varepsilon \le 1/6$ with probability at least $2/3$ for any $a \in \mathsf{P}^n$, after making N rounds of measurement, then $N = \Omega(2^n/\varepsilon^2)$.

This matches the known upper bound of $O(2^n/\varepsilon^2)$ based on minimal stabilizer covering [14,30], solving an open problem raised therein. Note that the task we consider is to estimate each λ_a with $2/3$ success probability *individually* rather than *simultaneously*. For the latter task, there will be an additional factor of n in the upper bound, but our lower bound still holds and is tight up to this logarithmic factor. Combined with the bound of $\Theta(1/\varepsilon^2)$ using entanglement-assisted scheme [14], this gives a tight exponential separation for learning with and without entanglement in the task of Pauli channel learning. Another noteworthy feature is that our lower bound has a moderate constant factor. For example, with $\varepsilon \le 0.1$ and $n \ge 5$ we have $N \ge 0.01 \times 2^n/\varepsilon^2$.

To highlight the experimental relevance of our result, in Fig. 2 we plot our lower bound of Theorem 1, the best previously known ancilla-free lower bound from [14], and the upper bound from an entanglement-assisted scheme studied in [14] with noisy Bell states preparation (see SM. Sec. V [33] for details). Figure 2 clearly indicates that our improved lower bound is crucial for demonstrating the entanglement-enabled advantages with a moderate number of qubits and fidelity. For example, with Bell pair fidelity 95%, the previous lower bound needs at least 85 qubits to start seeing any separation, while our improved lower bound needs as few as 25 qubits to obtain a factor of 10^5 advantages in sample complexity; With Bell pair fidelity below 90%, only our improved lower bound is able to obtain any separation.

Proof sketch of Theorem 1.—We extend the framework for proving exponential separation between learning with and without quantum memory [7,8]. The key idea is known as the Le Cam's two-point method [47] that reduces learning to hypothesis testing. Specifically, we first construct two hypotheses of Pauli channels (or, mixture of Pauli channels) that are close to each other. By assumption, a learning scheme can distinguish the two hypotheses with good probability. Consequently, the total variation distance (TVD) between the outcome probability distribution generated by the scheme under the two hypotheses needs to be at least constantly large. Therefore, if we can upper bound the contributions to the TVD from each measurement to be exponentially small, we will obtain an exponential lower bound on the number of measurements. However, the existing techniques for upper bounding TVD [7,8] do not carry over when channel concatenation is allowed, let alone mid-circuit measurements. Our technique to address this issue is to establish a recurrence relation on the mid-circuit states between each measurement step, and to upper bound the growth of TVD via mathematical induction. The full proof is presented in SM, Sec. III [33]. ■

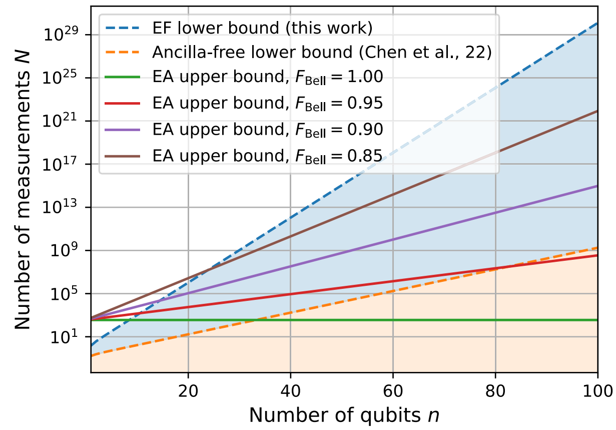

FIG. 2. Sample complexity for Pauli channel learning. The task is to estimate any Pauli fidelity to $\varepsilon = 0.1$ additive precision with at least $2/3$ success probability. The dashed lines represent our entanglement-free (EF) lower bound of Theorem 1 and the ancilla-free lower bound from [14]. The solid lines represent the sample complexity upper bound calculated from an entanglement-assisted (EA) scheme with noisy Bell state and measurements. For simplicity, we assume the state preparation suffers from depolarizing noises, so that each noisy 2-qubit Bell pair $\tilde{\rho}_{\mathrm{Bell}}$ has fidelity $F_{\mathrm{Bell}} = \langle\Psi_+|\tilde{\rho}_{\mathrm{Bell}}|\Psi_+\rangle$. The colored region indicates entanglement(ancilla)-enabled advantages.

Bounds on learning identifiable parameters.—In practice, Pauli channels are often used to model noise affecting Clifford gates [24,28,29]. One issue for learning *gate-dependent* noise channel is that, because of the existence of SPAM error, certain parameters of the noise channel might become nonidentifiable (or "unlearnable"), meaning that they cannot be identified independently from the noisy SPAM [48–52]. Specifically, for Clifford gate-dependent Pauli noise channel, a complete characterization of the learnable parameters is given in [52], which shows that some Pauli eigenvalues λ_a cannot be identified SPAM independently, but the geometric mean of certain set of eigenvalues, $(\prod_{a\in S}\lambda_a)^{1/|S|}$, can be. This is consistent with the existing noise learning protocols that can characterize gate-dependent Pauli noise SPAM robustly up to some degeneracy [21,22,29]. Our goal here is to find a lower bound for these scenarios. More precisely, we hope to address the following question: what is the sample complexity to learn the identifiable parameters rather than the whole Pauli channel?.

We will focus on a further simplified task: given a partition of n-qubit Pauli operators into some disjoint sets, $\{S_i\}_i$. The task is to learn the geometric mean of the Pauli eigenvalues within each S_i to additive precision ε with high probability. We denote the maximal cardinality among all S_i by C. As a motivating example, for the Pauli noise associated with a CNOT gate, it is shown [52] that the Pauli fidelity $\lambda_{XI}, \lambda_{XX}$ cannot be identified individually, but the geometric mean $\sqrt{\lambda_{XI}\lambda_{XX}}$ can. All the other learnable parameters are also geometric mean of up to two Pauli fidelities (related to the fact that $\mathrm{CNOT}^2 = \mathbb{1}$). Note that, there can be more learnable parameters than those decided by a partition (e.g., $\sqrt{\lambda_{XI}\lambda_{XX}}, \sqrt{\lambda_{YI}\lambda_{YX}}, \sqrt{\lambda_{XI}\lambda_{YX}}$ are three independent learnable parameters), but this simplified task is sufficient to give a sample complexity lower bound. We have the following result.

Theorem 2.—Given a partition of the n-qubit Pauli group, $\{S_i\}_i$, with maximum cardinality $C \coloneqq \max_i |S_i|$, define the geometrically averaged Pauli fidelity

$$\bar{\lambda}_{S_i} \coloneqq \mathrm{sgn}\left(\prod_{a\in S_i}\lambda_i\right)\prod_{a\in S_i}|\lambda_i|^{1/|S_i|}.$$

If there exists an entanglement-free scheme that, for any n-qubit Pauli channel Λ, outputs an estimator $\hat{\lambda}_{S_i}$ such that $|\hat{\lambda}_{S_i} - \bar{\lambda}_{S_i}| \le \varepsilon \le 1/6C$ with probability at least $2/3$ for any S_i, after making N rounds of measurements, then $N = \Omega(2^n\varepsilon^{-2}C^{-2})$.

Many multiqubit Clifford gates of practical interest have a polynomial (e.g., permutation gate) or constant (e.g., parallel CNOTs) order. That is, applying the gate a polynomial or constant number of times yields identity. This means their learnable parameters are groups of at most polynomial or constant many Pauli eigenvalues. Our result shows that there is still an exponential sample complexity lower bound for entanglement-free learning schemes in such cases.

The techniques for proving Theorem 2 is very similar to those for Theorem 1. Here, we just need to construct a different family of Pauli channels that can be distinguished by only looking at the geometrically averaged Pauli eigenvalues within each S_i. The other steps will carry over. One may notice that Theorem 1 can be viewed as a corollary of Theorem 2. We decide to present them separately for clarity. The proof is given in SM, Sec. IV [33].

Discussion.—In this Letter, we introduce two classes of learning schemes to capture the notion of learning without entanglement. One is quantum resource-theoretic, using only entanglement nongenerating operations between system and ancilla. The other is operational, describing quantum circuits assisted by mid-circuit measurement and classical feedforward. Both schemes are shown to be equivalent in terms of sample complexity. We then prove a tight lower bound for Pauli channel learning within this model. Our results extend existing proof techniques [7,8] and improve upon the best-known lower bounds in the literature [14]. We also generalize our bounds for practical quantum noise characterization settings.

Our scenario differs from existing frameworks of learning with or without quantum memory, which we briefly review below. For learning properties of quantum states, learning with quantum memory (or quantum-enhanced learning) usually means one can perform collective measurement on multiple copies of the state, while learning without quantum memory (or conventional learning) means only measurements on individual copies are allowed, though adaptivity is often allowed. Examples include [7–11]. For learning properties of quantum channels, there have been multiple definitions of learning without memory: Refs. [10,12] require the scheme to have no ancilla nor concatenation (i.e., sequentially applying the channel of interest); Refs. [7,8,13], etc., allow ancilla but not concatenation. In contrast, Ref. [14] studies schemes with concatenation but without ancilla. The scenario in the current work is strictly more general than Ref. [14] as we allow mid-circuit processing with quantum instruments instead of only quantum channels, and we justify our definition by connecting to the resource theory of entanglement. It is interesting to explore other learning tasks that admit a separation in our definition.

The problem of Pauli channel learning has been studied with different figures of merit. For Pauli error rates, Ref. [30] provides an ancilla-free protocol that learns the Pauli error rates to precision ε in l_2 distance with $\tilde{O}(2^n/\varepsilon^2)$ samples, which implies an upper bound of $\tilde{O}(2^{3n}/\varepsilon^2)$ for learning in l_1 distance. Reference [31] shows that $\tilde{O}(\log n/\varepsilon^2)$ samples are sufficient to learn the Pauli error rates to precision ε in l_∞ distance without using ancilla. In the case in which ancilla is allowed, one can use Bell states and Bell measurements to directly sample from the Pauli error rates, which implies an $O(2^{2n}/\varepsilon^2)$ upper bound for learning in l_1 distance [14,31]. For Pauli eigenvalues, Ref. [14] gives a family of k-qubit ancilla-assisted protocols using $O(n2^{n-k}/\varepsilon^2)$ samples for learning in l_∞ distance, for $0 \le k \le n$. In terms of the lower bounds, Ref. [14] focuses on learning Pauli eigenvalues to constant error in l_∞ distance, and in particular obtains a lower bound $\Omega(2^{n/3})$ for the number of measurements for any ancilla-free schemes with concatenation. The current work improves this lower bound to be tight (for more general schemes). Another recent work studies learning Pauli error rate to error ε in l_1 distance [32]. For ancilla-free schemes with adaptivity, they obtain $\Omega(2^{2n}/\varepsilon^2)$ for general case and $\Omega(2^{2.5n}/\varepsilon^2)$ when ε is exponentially small in n. They allow concatenation with *unital* processing channels and the lower bound holds for the number of measurements. The results of the current work and Ref. [32] do not imply each other [53]. Whether a tighter lower bound can be established with other figures of merit remains an open problem.

Our results have implications in practical quantum noise characterization tasks. On the one hand, they set up an exponential barrier for any entanglement-free Pauli channel learning protocols [21,30] without additional assumptions on the Pauli noise model. The barrier persists even if one only aims at learning the SPAM independently identifiable part of the noise channel. On the other hand, this motivates the development of an entanglement-assisted noise characterization protocol, which was pioneered in Ref. [14]. It is shown there that an entanglement-assisted protocol can learn the Pauli eigenvalues efficiently and SPAM robustly, given access to good quantum memory. We believe the tight bounds obtained in the current work will strengthen the foundation for experimentally demonstrating the advantage of entanglement in this noise characterization task.

Finally, it is interesting to explore a deeper connection between the resource theory of entanglement [34–36] and quantum learning theory. Specifically, since the tight bounds for Pauli channel learning with no entanglement and arbitrary entanglement have been settled, it is natural to ask about learning with a certain amount of entanglement. There could be different ways of defining the entanglement cost of a learning scheme, one of which is to allow a bounded amount of ancillary qubits [8,14]. An upper bound of $\tilde{O}(2^{n-k}/\varepsilon^2)$ for k-ancillary-qubit-assisted scheme is known and has proven optimal for some restricted classes of schemes [14], but a general answer to this question is yet to be found.

Note added.—During the completion of this manuscript, we became aware of an independent and contemporaneous work [54] that obtains, among other results, a tight lower bound of $\Omega(2^n/\varepsilon^2)$ on the number of measurements for learning every Pauli eigenvalue to ε precision with ancilla-free concatenating schemes. Their proof is based on a different technique, which leads to different features in their results compared to ours. On the one hand, the ε in their bound can be any value within (0, 1], while ours is restricted to $\varepsilon \in (0, 1/6]$. On the other hand, our bound has a better constant factor and applies for the more general setting of classical-memory-assisted schemes.

We thank Matthias Caro, Steve Flammia, John Preskill, and Alireza Seif for helpful discussions. We thank Sitan Chen and Weiyuan Gong for communicating with us about their independent and contemporaneous results. S. C., C. O., and L. J. acknowledge support from the ARO (W911NF-23-1-0077), ARO MURI (W911NF-21-1-0325), AFOSR MURI (FA9550-19-1-0399, FA9550-21-1-0209), NSF (OMA-1936118, ERC-1941583, OMA-2137642), NTT Research, and the Packard Foundation (2020-71479). This material is based upon work supported by the U.S. Department of Energy, Office of Science, National Quantum Information Science Research Centers. S. C. thanks Caltech IQIM for hospitality, where part of this work was completed. S. Z. acknowledges funding provided by the Institute for Quantum Information and Matter, an NSF Physics Frontiers Center (NSF Grant PHY-1733907) and Perimeter Institute for Theoretical Physics, a research institute supported in part by the Government of Canada through the Department of Innovation, Science and Economic Development Canada and by the Province of Ontario through the Ministry of Colleges and Universities. H. H. is supported by a Google Ph.D. fellowship and a MediaTek Research Young Scholarship. H. H. acknowledges the visiting associate position at Massachusetts Institute of Technology.

*Corresponding author: csenrui@uchicago.edu
†Corresponding author: liang.jiang@uchicago.edu

[1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition*, 10th ed. (Cambridge University Press, USA, 2011), 10.1017/CBO9780511976667.
[2] N. Gisin and R. Thew, Quantum communication, Nat. Photonics **1**, 165 (2007).
[3] H. J. Kimble, The quantum internet, Nature (London) **453**, 1023 (2008).
[4] V. Giovannetti, S. Lloyd, and L. Maccone, Quantum metrology, Phys. Rev. Lett. **96**, 010401 (2006).
[5] V. Giovannetti, S. Lloyd, and L. Maccone, Advances in quantum metrology, Nat. Photonics **5**, 222 (2011).
[6] E. Polino, M. Valeri, N. Spagnolo, and F. Sciarrino, Photonic quantum metrology, AVS Quantum Sci. **2**, 024703 (2020).
[7] H.-Y. Huang, M. Broughton, J. Cotler, S. Chen, J. Li, M. Mohseni, H. Neven, R. Babbush, R. Kueng, J. Preskill *et al.*, Quantum advantage in learning from experiments, Science **376**, 1182 (2022).
[8] S. Chen, J. Cotler, H.-Y. Huang, and J. Li, Exponential separations between learning with and without quantum memory, in *Proceedings of the 2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)* (IEEE, New York, 2022), pp. 574–585.
[9] S. Bubeck, S. Chen, and J. Li, Entanglement is necessary for optimal quantum property testing, in *Proceedings of the 2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)* (IEEE, New York, 2020), pp. 692–703.
[10] H.-Y. Huang, R. Kueng, and J. Preskill, Information-theoretic bounds on quantum advantage in machine learning, Phys. Rev. Lett. **126**, 190505 (2021).
[11] S. Chen, B. Huang, J. Li, A. Liu, and M. Sellke, Tight bounds for state tomography with incoherent measurements, arXiv:2206.05265.

[12] D. Aharonov, J. Cotler, and X.-L. Qi, Quantum algorithmic measurement, Nat. Commun. **13**, 887 (2022).

[13] M. C. Caro, Learning quantum processes and hamiltonians via the pauli transfer matrix, arXiv:2212.04471.

[14] S. Chen, S. Zhou, A. Seif, and L. Jiang, Quantum advantages for Pauli channel estimation, Phys. Rev. A **105**, 032435 (2022).

[15] S. Chen, J. Cotler, H.-Y. Huang, and J. Li, The complexity of NISQ, Nat. Commun. **14**, 6001 (2023).

[16] E. Chitambar and G. Gour, Quantum resource theories, Rev. Mod. Phys. **91**, 025001 (2019).

[17] P. W. Shor, Fault-tolerant quantum computation, in *Proceedings of 37th Conference on Foundations of Computer Science* (IEEE, New York, 1996), pp. 56–65.

[18] D. Gottesman, Theory of fault-tolerant quantum computation, Phys. Rev. A **57**, 127 (1998).

[19] M. Iqbal, N. Tantivasadakarn, T. M. Gatterman, J. A. Gerber, K. Gilmore, D. Gresh, A. Hankin, N. Hewitt, C. V. Horst, M. Matheny *et al.*, Topological order from measurements and feed-forward on a trapped ion quantum computer, arXiv:2302.01917.

[20] K. Singh, C. Bradley, S. Anand, V. Ramesh, R. White, and H. Bernien, Mid-circuit correction of correlated phase errors using an array of spectator qubits, Science **380**, 1265 (2023).

[21] A. Erhard, J. J. Wallman, L. Postler, M. Meth, R. Stricker, E. A. Martinez, P. Schindler, T. Monz, J. Emerson, and R. Blatt, Characterizing large-scale quantum computers via cycle benchmarking, Nat. Commun. **10**, 5347 (2019).

[22] A. Carignan-Dugas, D. Dahlen, I. Hincks, E. Ospadov, S. J. Beale, S. Ferracin, J. Skanes-Norman, J. Emerson, and J. J. Wallman, The error reconstruction and compiled calibration of quantum computing cycles, arXiv:2303.17714.

[23] R. Harper, S. T. Flammia, and J. J. Wallman, Efficient learning of quantum noise, Nat. Phys. **16**, 1184 (2020).

[24] E. Van Den Berg, Z. K. Minev, A. Kandala, and K. Temme, Probabilistic error cancellation with sparse Pauli–Lindblad models on noisy quantum processors, Nat. Phys. **19**, 1116 (2023).

[25] S. Ferracin, A. Hashim, J.-L. Ville, R. Naik, A. Carignan-Dugas, H. Qassim, A. Morvan, D. I. Santiago, I. Siddiqi, and J. J. Wallman, Efficiently improving the performance of noisy quantum computers, arXiv:2201.10672.

[26] Y. Kim, A. Eddins, S. Anand, K. X. Wei, E. Van Den Berg, S. Rosenblatt, H. Nayfeh, Y. Wu, M. Zaletel, K. Temme *et al.*, Evidence for the utility of quantum computing before fault tolerance, Nature (London) **618**, 500 (2023).

[27] D. K. Tuckett, S. D. Bartlett, and S. T. Flammia, Ultrahigh error threshold for surface codes with biased noise, Phys. Rev. Lett. **120**, 050505 (2018).

[28] J. J. Wallman and J. Emerson, Noise tailoring for scalable quantum computation via randomized compiling, Phys. Rev. A **94**, 052325 (2016).

[29] A. Hashim, R. K. Naik, A. Morvan, J.-L. Ville, B. Mitchell, J. M. Kreikebaum, M. Davis, E. Smith, C. Iancu, K. P. O'Brien, I. Hincks, J. J. Wallman, J. Emerson, and I. Siddiqi, Randomized compiling for scalable quantum computing on a noisy superconducting quantum processor, Phys. Rev. X **11**, 041039 (2021).

[30] S. T. Flammia and J. J. Wallman, Efficient estimation of Pauli channels, ACM Trans. Quantum Comput. **1**, 1 (2020).

[31] S. T. Flammia and R. O'Donnell, Pauli error estimation via population recovery, Quantum **5**, 549 (2021).

[32] O. Fawzi, A. Oufkir, and D. S. França, Lower bounds on learning pauli channels, arXiv:2301.09192.

[33] There is an additional factor of n in the upper bound of [14] because the task there is to estimate all λ_a simultaneously with high probability. Here we only require estimating each λ_a individually with high probability. One can derive the former using the latter via the union bound and paying a factor of n in sample complexity.

[34] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher, Concentrating partial entanglement by local operations, Phys. Rev. A **53**, 2046 (1996).

[35] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Mixed-state entanglement and quantum error correction, Phys. Rev. A **54**, 3824 (1996).

[36] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight, Quantifying entanglement, Phys. Rev. Lett. **78**, 2275 (1997).

[37] J. I. Cirac, W. Dür, B. Kraus, and M. Lewenstein, Entangling operations and their implementation using a small amount of entanglement, Phys. Rev. Lett. **86**, 544 (2001).

[38] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels, Phys. Rev. Lett. **70**, 1895 (1993).

[39] E. Chitambar, D. Leung, L. Mančinska, M. Ozols, and A. Winter, Everything you always wanted to know about locc (but were afraid to ask), Commun. Math. Phys. **328**, 303 (2014).

[40] J. Emerson, R. Alicki, and K. Życzkowski, Scalable noise estimation with random unitary operators, J. Opt. B **7**, S347 (2005).

[41] E. Knill, D. Leibfried, R. Reichle, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland, Randomized benchmarking of quantum gates, Phys. Rev. A **77**, 012307 (2008).

[42] See Supplemental Materials at http://link.aps.org/supplemental/10.1103/PhysRevLett.132.180805, which includes Refs. [43,44], for additional definitions and detailed proofs.

[43] M.-D. Choi, Completely positive linear maps on complex matrices, Linear Algebra Appl. **10**, 285 (1975).

[44] W. Hoeffding, Probability inequalities for sums of bounded random variables, *The Collected Works of Wassily Hoeffding* (Springer, New York, 1994), p. 409, 10.1007/978-1-4612-0865-5_26.

[45] S. T. Flammia, Averaged Circuit Eigenvalue Sampling,, in *17th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2022). Leibniz International Proceedings in Informatics (LIPIcs)* (Schloss Dagstuhl—Leibniz-Zentrum für Informatik, 2022), 10.4230/LIPIcs.TQC.2022.4.

[46] S. Chen, W. Yu, P. Zeng, and S. T. Flammia, Robust shadow estimation, PRX Quantum **2**, 030348 (2021).

[47] L. LeCam, Convergence of estimates under dimensionality restrictions, Ann. Stat. **1**, 38 (1973).
[48] R. Blume-Kohout, J. King Gamble, E. Nielsen, J. Mizrahi, J. D. Sterk, and P. Maunz, Robust, self-consistent, closed-form tomography of quantum logic gates on a trapped ion qubit, arXiv:1310.4492.
[49] T. Proctor, K. Rudinger, K. Young, M. Sarovar, and R. Blume-Kohout, What randomized benchmarking actually measures, Phys. Rev. Lett. **119**, 130502 (2017).
[50] E. Nielsen, J. K. Gamble, K. Rudinger, T. Scholten, K. Young, and R. Blume-Kohout, Gate set tomography, Quantum **5**, 557 (2021).
[51] H.-Y. Huang, S. T. Flammia, and J. Preskill, Foundations for learning from noisy quantum experiments, arXiv:2204.13691.
[52] S. Chen, Y. Liu, M. Otten, A. Seif, B. Fefferman, and L. Jiang, The learnability of pauli noise, Nat. Commun. **14**, 52 (2023).
[53] We remark that a lower bound of $\Omega(2^{3n}/\varepsilon^2)$ for learning Pauli error rates in l_1 distance would imply a bound of $\Omega(2^n/\varepsilon^2)$ for learning Pauli eigenvalues in l_∞ distance via the Parseval-Plancherel identity and relations of l_p norm.
[54] S. Chen and W. Gong, Futility and utility of a few ancillas for pauli channel learning, arXiv:2309.14326.