Learning quantum states and unitaries of bounded gate complexity

Haimeng Zhao*, Laura Lewis*, Ishaan Kannan*, Yihui Quek, Hsin-Yuan Huang, and Matthias C. Caro

The tasks of general *state* and *process tomography* – that is, determining an unknown quantum state/unitary from copies of it/queries to it – are practically ubiquitous [1–13] but exponentially costly [14–16]. Despite this theoretical obstacle, there is hope that efficient tomography remains within reach, since the vast majority of states and unitaries are not of immediate physical interest [17–19]. Indeed, practitioners may be able to leverage *prior knowledge* about the unknown state or process. Previous works have demonstrated efficient learning if the unknown state is known to be a stabilizer [20–23], of limited T-gate count [24–27], or the output of a shallow circuit [28]; or if the unknown process is a local Pauli noise channel [29].

Our work centers around a prior that is fundamental in physics and timely to the current state of quantum technology, albeit underexplored in tomography: gate complexity. While the vast majority of many-body states and unitaries have exponential gate complexity, those that can even be prepared "in Nature" [17–19] are likely to have at most polynomial gate complexity. This raises the following question:

What is the complexity of learning states/unitaries with bounded gate complexity?

We first study the *sample/query complexity* of learning an n-qubit state/unitary implemented by G two-qubit gates: the number of samples collected from the system or queries to the process required to output an ϵ-accurate classical description of the unknown object. Previously, [30] showed that the state learning task can be accomplished with sample complexity $\tilde{\mathcal{O}}(nG^2/\epsilon^4)$. However, it is not known if this is the optimal sample complexity for the state learning task. In this work, we fully resolve this problem by proving a matching upper and lower bound (up to log factors) of $\tilde{\Theta}(G/\epsilon^2)$ for learning n-qubit states generated by G gates to trace distance ϵ. For learning unitaries, we establish a query complexity scaling linear in the gate complexity G and independent of the system size n, which is optimal in the scaling of G (up to log factors). Turning to the question of *computational complexity*, we demonstrate that any quantum learning algorithm requires computational time scaling exponentially in G.

Our results pose fresh implications in other subfields of quantum information. Firstly, the Brown-Susskind conjecture [31–36] states that the complexity of a generic local quantum circuit grows linearly with the number of gates for an exponentially long time, holographically dual to the steady growth of a wormhole's volume in the bulk theory. This conjecture has recently been confirmed [37, 38] with "complexity" understood as (exact) "circuit complexity" [35]. Our work suggests that this conjecture may also be true for an alternative notion of complexity – that of learning the quantum circuit. Secondly, our work completely characterizes the complexity of inferring a classical circuit description of a quantum circuit from limited copies of its output state: yet such a description holds the key to computing arbitrarily many more properties of this state (via classical simulation algorithms) than would have been possible by using the limited state copies directly. We thus expect our results to shed light on the link between classical simulability and learnability of quantum states [39, 40]. They also provide a *learning* perspective on the celebrated recent notion of state complexity classes [41, 42], which ask what states can be synthesized by polynomial-space quantum circuits.

I. LEARNING QUANTUM STATES AND UNITARIES

Our sample complexity results are summarized in Table I. We also establish the exponential-in-G growth of computational complexity, assuming that Ring Learning with Errors [43] (RingLWE) cannot be solved in subexponential time with a quantum computer (Theorems 2 and 6).

Sample complexity	State (Thm 1)	Unitary (average-case) (Thm 4)	Unitary (worst-case) (Thm 3)
Upper bound	$\tilde{\mathcal{O}}\left(G/\epsilon^2\right)$	$\tilde{\mathcal{O}}\left(G\min\left\{1/\epsilon^2, \sqrt{2^n}/\epsilon\right\}\right)$	$\tilde{\mathcal{O}}\left(2^n G/\epsilon\right)$
Lower bound	$\tilde{\Omega}\left(G/\epsilon^2\right)$	$\Omega\left(G/\epsilon\right)$	$\Omega\left(2^{\min\{G/(2C), n/2\}}/\epsilon\right)$

Table I. **Sample complexity of learning n-qubit states and unitaries with circuit complexity G.** The learning accuracy ϵ is measured in trace distance for states, root mean squared trace distance for average case unitary learning, and diamond distance for worst case. Here, $C > 0$ is some universal constant.

a. ***Learning quantum states.*** First, we consider learning quantum states of bounded complexity. Let $|\psi\rangle = U|0\rangle^{\otimes n}$ be an n-qubit pure state generated by a unitary U consisting of G two-qubit gates. Given N copies of $\rho \triangleq |\psi\rangle\langle\psi|$, we aim to learn a classical circuit description $\hat{\rho}$ of ρ that is ϵ-close to ρ in trace distance: $\mathrm{d}_{\mathrm{tr}}(\hat{\rho}, \rho) = \|\hat{\rho} - \rho\|_1/2 < \epsilon$. We fully characterize the sample complexity in the following theorem.

Theorem 1 (State learning)**.** *$N = \tilde{\Theta}\left(G/\epsilon^2\right)$ copies of an n-qubit pure state with circuit complexity G are necessary and sufficient to learn it to within ϵ trace distance with high probability.*

* These authors contributed equally to this work.

This result also resolves an open question from [44] and improves [30] to achieve the optimal scaling. In the regime where $G = \mathcal{O}(\mathsf{poly}(n))$, the sample complexity of our algorithm improves substantially over the sample complexity $N = \mathcal{O}(2^n)$ of the sample-optimal result for arbitrary pure quantum states [14, 15]. Note that we do not need to know or have access to the unitary U which generates the unknown state $|\psi\rangle$.

While our algorithm to learn the unknown quantum state $|\psi\rangle$ is sample-efficient and sample-optimal, it is computationally inefficient. We prove that this cannot be avoided in general: any quantum algorithm that learns $|\psi\rangle$ given access to copies of this state must use time exponential-in-G, under the commonly-believed cryptographic assumption that RingLWE [43] cannot be solved by a quantum computer in sub-exponential time. This imposes strong computational complexity limitations on learning even comparatively simple states, in stark contrast with our sample complexity results. Meanwhile, we show that the learning task is computationally-efficiently solvable for $G = \mathcal{O}(\log n)$, implying a transition point of computational efficiency. Previous work [45, 46] arrives at related hardness results for polynomial gate complexity $G = \mathrm{poly}(n)$, but our detailed analysis allows us to sharpen the computational lower bound and obtain this transition point.

Theorem 2 (Computational hardness of state learning)**.** *Any quantum algorithm that learns an n-qubit state with circuit complexity G to within ϵ trace distance requires $\exp(\Omega(\min(G, n)))$ time, assuming the quantum sub-exponential hardness of RingLWE. Meanwhile, for $G = \mathcal{O}(\log n)$, an efficient learning algorithm exists.*

b. ***Learning quantum unitaries.*** Next, we consider learning unitaries of bounded complexity. Let U be a unitary consisting of G two-qubit gates. Given query access to the unitary U, we aim to learn a classical circuit description $\hat{U}$ of U that is ϵ-close to U. A natural distance metric analogous to the trace distance for states is the diamond distance $\mathrm{d}_\diamond(U, V) = \max_\rho \|(U \otimes I)\rho(U \otimes I)^\dagger - (V \otimes I)\rho(V \otimes I)^\dagger\|_1$. We find that in this worst-case learning task, a number of queries exponential in G is necessary.

Theorem 3 (Worst-case unitary learning)**.** *Any quantum algorithm learning an n-qubit unitary with circuit complexity G in diamond distance with high probability must use at least $\Omega\left(2^{\min\{G/(2C), n/2\}}/\epsilon\right)$ queries, where $C > 0$ is a universal constant. Meanwhile, there exists such an algorithm using $\tilde{\mathcal{O}}(2^n G/\epsilon)$ queries.*

Having established this no-go theorem for worst-case learning, we turn to a more realistic average-case learning alternative. Here, the accuracy is measured using the root mean squared trace distance between output states over Haar-random inputs, $\mathrm{d}_{\mathrm{avg}}(U, V) = \sqrt{\mathbb{E}_{|\psi\rangle}[\mathrm{d}_{\mathrm{tr}}(U|\psi\rangle, V|\psi\rangle)^2]}$. This metric characterizes the average error when testing the learned unitary on randomly chosen inputs. We find that, similarly to the state learning task, linear-in-G many queries are both necessary and sufficient in this case.

Theorem 4 (Average-case unitary learning)**.** $N = \tilde{\mathcal{O}}\left(G \min\{1/\epsilon^2, \sqrt{2^n}/\epsilon\}\right)$ *queries are sufficient to learn an n-qubit unitary with circuit complexity G to ϵ root mean squared trace distance with high probability. Meanwhile, at least $\Omega\left(G/\epsilon\right)$ queries to the unitary, its inverse, or the controlled versions are necessary.*

The similar linear-in-G sample/query complexity in Theorems 1 and 4 hints at a common underlying source of complexity. However, in contrast to state learning, unitary learning comes with two natural such sources: (1) to readout input and output states, and (2) to learn the mapping from inputs to outputs. Our results suggest that the former may encapsulate the central difficulty whereas the latter may be easy. This seemingly contradicts recent quantum no free lunch theorems [47, 48], which state that $\Omega(2^n)$ samples are required to learn a generic unitary even from classically described input-output state pairs. To resolve this paradox, we reformulate the quantum no free lunch theorem from an information-theoretic perspective.

Theorem 5 (Learning with classical descriptions)**.** $\mathcal{O}(2^n/r)$ *classically described samples with mixed (entangled) input states of (Schmidt) rank r are sufficient to learn any n-qubit unitary to any accuracy with high probability. Moreover, any such algorithm that is robust to noise needs at least $\Omega(2^n/r)$ samples.*

Similarly to state learning, our average-case unitary learning algorithm is not computationally efficient. This is again inevitable, and the same is true for worst-case unitary learning. Moreover, the hard instances we construct are implementable with Clifford+T circuits with $\tilde{\omega}(\log n)$ T gates [24], so, together with Theorem 2, this gives a negative answer to an open question (the fifth question) in the survey [39].

Theorem 6 (Computational hardness of unitary learning)**.** *Any quantum algorithm that learns an n-qubit unitary with circuit complexity G requires $\exp(\Omega(\min(G, n)))$ time, assuming the quantum sub-exponential hardness of RingLWE. Meanwhile, for $G = \mathcal{O}(\log n)$, an efficient learning algorithm exists.*

Apart from learning quantum states and dynamics themselves, a more classically minded learner may care more about learning classical functions resulting from quantum processes. We define these *physical functions* $f(x) : [0, 1]^\nu \to \mathbb{R}$ in three steps: (1) a fixed state preparation procedure that can depend on x; (2) a unitary evolution consisting of G tunable two-qubit gates and arbitrary fixed unitaries that can depend on x, arranged in a circuit structure; (3) the measurement of a fixed observable, whose expectation is the function output. Despite the generality of this setup, we find that certain well-behaved functions are actually not physical: they cannot be efficiently approximated or learned via physical functions. This reveals a fundamental limitation on the functional expressivity of both nature and practical quantum machine learning models [49–52].

Theorem 7 (Approximating and learning with physical functions). *At least $G \geq \tilde{\Omega}(1/\epsilon^{\nu/2})$ gates and $N \geq \Omega(1/\epsilon^{\nu})$ samples are needed to approximate and learn arbitrary 1-bounded and 1-Lipschitz $\mathbb{R}$-valued functions on $[0,1]^{\nu}$ to accuracy ϵ in $\|\cdot\|_{\infty}$ with high probability using physical functions.*

II. PROOF IDEAS

Sample complexity upper bounds. We prove the upper bounds in Theorems 1 and 4 by explicitly constructing learning algorithms using a hypothesis selection protocol [53] based on classical shadow tomography [54]. Specifically, we construct a covering net $\mathcal{N}$ over the set of states/unitaries with bounded circuit complexity G such that for any such state/unitary, there exists a candidate in the covering net that is ϵ-close in trace/average case distance to it, respectively. Then it remains to select and output the candidate with the smallest trace distance/average case distance to our unknown object. To do this, we utilize classical shadows with random Clifford measurements to estimate all distances simultaneously to ϵ error using $\mathcal{O}(\log|\mathcal{N}|/\epsilon^2) \leq \tilde{\mathcal{O}}(G/\epsilon^2)$ copies/queries. The last inequality follows because we are able to prove that $|\mathcal{N}| \leq e^{\tilde{\mathcal{O}}(G)}$. Finally, we output the candidate with the smallest distance.

The above strategy leads to a sample complexity that depends logarithmically on n, which is undesirable when G is smaller than $n/2$ (i.e. when some of the qubits are, in fact, never influenced by the circuit). We improve upon this by first performing a junta learning step [55] to identify which of the qubits are acted on non-trivially. After identifying the non-trivial qubits, we perform a measure-and-postselect step. This allows us to construct a covering net only over the qubits acted upon non-trivially, whose cardinality no longer depends on n (thus removing the n dependence in the sample complexity). We subsequently perform hypothesis selection as before but over the members of this covering net.

Furthermore, for unitary learning, we improve the ϵ dependence to the Heisenberg scaling $\tilde{\mathcal{O}}(1/\epsilon)$ via a bootstrap method [16], using the above learner as a sub-routine. We iteratively refine our learning outcome $\hat{U}$ by performing hypothesis selection over a covering net of $(U\hat{U}^{\dagger})^p$, with p increasing exponentially as the iteration proceeds. Although the circuit complexity of $(U\hat{U}^{\dagger})^p$ grows with p, a covering net with p-independent cardinality can be constructed based on the one-to-one correspondence with U. However, unlike the diamond distance learner in [16], which has fine control over every eigenvalue of the unitaries, our average-case learner only has average control over eigenvalues. Thus for the bootstrap to work (i.e., for the learning error to decrease with increasing p), the average-case learner has to work in an exponentially small error regime, which results in a dimensional factor in the final sample complexity $\tilde{\mathcal{O}}(\sqrt{2^n}G/\epsilon)$.

Sample complexity lower bounds. We prove the lower bounds in Theorems 1 and 4 by reduction to a distinguishing task: if we can approximately learn states/unitaries, then we can use this learning algorithm to distinguish a set of states/unitaries that are far apart from each other. Hence a lower bound on the sample complexity of distinguishing states/unitaries in a packing net implies a lower bound for learning.

For state learning, we construct a packing net $\mathcal{M}$ of the set of $(\log_2 G)$-qubit states. These states have circuit complexity $\sim G$ because $\mathcal{O}(2^k)$ two-qubit gates can implement any pure k-qubit states [56]. We prove that $|\mathcal{M}| \geq e^{\Omega(G)}$, which means that to distinguish these states, one has to gather $\Omega(\log|\mathcal{M}|) \geq \Omega(G)$ bits of information. Meanwhile, Holevo's theorem [57] asserts that the amount of information carried by each sample is upper bounded by $\tilde{\mathcal{O}}(\epsilon^2)$ [58]. Hence, we need at least $\tilde{\Omega}(G/\epsilon^2)$ copies of the unknown state.

Similarly, for unitary learning, we construct a packing net by stacking all the gates into $\log_4 G$ qubits, using the fact that $\mathcal{O}(4^k)$ two-qubit gates can implement any k-qubit unitaries [59]. Lacking an analog of Holevo's theorem for unitary queries, we turn to a recently established bound on the success probability of unitary discrimination [60] and obtain an $\Omega(G)$ sample complexity lower bound for constant ϵ. To incorporate the ϵ dependence, we follow [16] and map the problem into a fractional query problem. We show that with N queries, we can use the learning algorithm to simulate [61, 62] an $\mathcal{O}(\epsilon N)$ query algorithm that solves the above constant-error distinguishing problem. This gives us the desired $\Omega(G/\epsilon)$ lower bound.

Computational hardness. We prove the computational complexity lower bounds in Theorems 2 and 6 by reduction to the task of distinguishing pseudorandom states/functions [63, 64] from truly random states/-functions. The hardness of this task relies on the computational assumption that Ring Learning with Errors (RingLWE) cannot be solved efficiently with a quantum computer [65]. In particular, if we consider learning a pseudorandom state or a unitary implementing a pseudorandom function, then an efficient learner implies an efficient distinguisher from Haar-random states or truly random functions, thus contradicting the pseudorandomness assumption. The circuit complexity at which this computational hardness kicks in is the complexity of the circuit required to implement pseudorandom states/functions. Indeed, we show that the pseudorandom functions constructed in [65] based on the hardness of RingLWE can be implemented with $G = \mathcal{O}(\mathsf{poly}(n))$ gates and depth $\mathcal{O}(\mathsf{polylog}(n))$, Then, we can also construct pseudorandom quantum states from these pseudorandom functions [66], which can be implemented with $G = \mathcal{O}(\mathsf{poly}(n))$ gates and depth $\mathcal{O}(\mathsf{polylog}(n))$. We can boost this to our exponential computational complexity lower bound by assuming the sub-exponential hardness of RingLWE rather than just polynomial hardness. The efficient learning algorithm at $G = \mathcal{O}(\log n)$ follows by junta learning [55] and standard tomography methods.

[1] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone. Quantum metrology. *Physical review letters*, 96(1):010401, 2006.

[2] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone. Advances in quantum metrology. *Nature photonics*, 5(4):222–229, 2011.

[3] Christian Kokail, Rick van Bijnen, Andreas Elben, Benoît Vermersch, and Peter Zoller. Entanglement Hamiltonian tomography in quantum simulation. *Nature Physics*, 17(8):936–942, 2021.

[4] Jose Carrasco, Andreas Elben, Christian Kokail, Barbara Kraus, and Peter Zoller. Theoretical and experimental perspectives of quantum verification. *PRX Quantum*, 2(1):010102, 2021.

[5] Masoud Mohseni, Ali T Rezakhani, and Daniel A Lidar. Quantum-process tomography: Resource analysis of different strategies. *Physical Review A*, 77(3):032322, 2008.

[6] Jeremy L O'Brien, Geoff J Pryde, Alexei Gilchrist, Daniel FV James, Nathan K Langford, Timothy C Ralph, and Andrew G White. Quantum process tomography of a controlled-not gate. *Physical review letters*, 93(8):080502, 2004.

[7] Andrew James Scott. Optimizing quantum process tomography with unitary 2-designs. *Journal of Physics A: Mathematical and Theoretical*, 41(5):055308, 2008.

[8] Isaac L Chuang and Michael A Nielsen. Prescription for experimental determination of the dynamics of a quantum black box. *Journal of Modern Optics*, 44(11-12):2455–2467, 1997.

[9] Ryan Levy, Di Luo, and Bryan K Clark. Classical shadows for quantum process tomography on near-term quantum computers. *arXiv preprint arXiv:2110.02965*, 2021.

[10] Seth T Merkel, Jay M Gambetta, John A Smolin, Stefano Poletto, Antonio D Córcoles, Blake R Johnson, Colm A Ryan, and Matthias Steffen. Self-consistent quantum process tomography. *Physical Review A*, 87(6):062119, 2013.

[11] Robin Blume-Kohout, John King Gamble, Erik Nielsen, Kenneth Rudinger, Jonathan Mizrahi, Kevin Fortier, and Peter Maunz. Demonstration of qubit operations below a rigorous fault tolerance threshold with gate set tomography. *Nature communications*, 8(1):14485, 2017.

[12] Hsin-Yuan Huang, Steven T Flammia, and John Preskill. Foundations for learning from noisy quantum experiments. *arXiv preprint arXiv:2204.13691*, 2022.

[13] Zhenyu Cai, Ryan Babbush, Simon C Benjamin, Suguru Endo, William J Huggins, Ying Li, Jarrod R McClean, and Thomas E O'Brien. Quantum error mitigation. *arXiv preprint arXiv:2210.00921*, 2022.

[14] Jeongwan Haah, Aram W Harrow, Zhengfeng Ji, Xiaodi Wu, and Nengkun Yu. Sample-optimal tomography of quantum states. *IEEE Transactions on Information Theory*, 63(9):5628–5641, 2017.

[15] Ryan O'Donnell and John Wright. Efficient quantum tomography. In *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, pages 899–912, 2016.

[16] Jeongwan Haah, Robin Kothari, Ryan O'Donnell, and Ewin Tang. Query-optimal estimation of unitary channels in diamond distance. *arXiv preprint arXiv:2302.14066*, 2023.

[17] David Poulin, Angie Qarry, Rolando Somma, and Frank Verstraete. Quantum simulation of time-dependent Hamiltonians and the convenient illusion of Hilbert space. *Physical review letters*, 106(17):170501, 2011.

[18] Fernando GSL Brandão, Wissam Chemissany, Nicholas Hunter-Jones, Richard Kueng, and John Preskill. Models of quantum complexity growth. *PRX Quantum*, 2(3):030316, 2021.

[19] Thomas Barthel and Jianfeng Lu. Fundamental limitations for measurements in quantum many-body systems. *Physical Review Letters*, 121(8):080406, 2018.

[20] Scott Aaronson and Daniel Gottesman. Improved simulation of stabilizer circuits. *Phys. Rev. A*, 70(5):052328, 2004.

[21] Ashley Montanaro. Learning stabilizer states by bell sampling. *arXiv preprint arXiv:1707.04012*, 2017.

[22] Richard A Low. Learning and testing algorithms for the clifford group. *Physical Review A*, 80(5):052314, 2009.

[23] Andrea Rocchetto. Stabiliser states are efficiently pac-learnable. *Quantum Info. Comput.*, 18(7–8):541–552, June 2018.

[24] Ching-Yi Lai and Hao-Chung Cheng. Learning quantum circuits of some t gates. *IEEE Transactions on Information Theory*, 68(6):3951–3964, 2022.

[25] Sabee Grewal, Vishnu Iyer, William Kretschmer, and Daniel Liang. Efficient learning of quantum states prepared with few non-clifford gates. *arXiv preprint arXiv:2305.13409*, 2023.

[26] Sabee Grewal, Vishnu Iyer, William Kretschmer, and Daniel Liang. Efficient learning of quantum states prepared with few non-clifford gates ii: Single-copy measurements. *arXiv preprint arXiv:2308.07175*, 2023.

[27] Dominik Hangleiter and Michael J Gullans. Bell sampling from quantum circuits. *arXiv preprint arXiv:2306.00083*, 2023.

[28] Cambyse Rouzé and Daniel Stilck França. Learning quantum many-body systems from a few copies. *arXiv preprint arXiv:2107.03333*, 2021.

[29] Cambyse Rouzé and Daniel Stilck França. Efficient learning of the structure and parameters of local pauli noise channels. *arXiv preprint arXiv:2307.02959*, 2023.

[30] Scott Aaronson. Shadow tomography of quantum states. In *STOC*, pages 325–338, 2018.

[31] Adam R Brown and Leonard Susskind. Second law of quantum complexity. *Physical Review D*, 97(8):086015, 2018.

[32] Leonard Susskind. Black holes and complexity classes. *arXiv preprint arXiv:1802.02175*, 2018.

[33] Leonard Susskind. Computational complexity and black hole horizons. *Fortschritte der Physik*, 64(1):24–43, 2016.

[34] Adam Bouland, Bill Fefferman, and Umesh Vazirani. Computational pseudorandomness, the wormhole growth paradox, and constraints on the ads/cft duality. *arXiv preprint arXiv:1910.14646*, 2019.

[35] Douglas Stanford and Leonard Susskind. Complexity and shock wave geometries. *Physical Review D*, 90(12):126007, 2014.

[36] Adam R Brown, Daniel A Roberts, Leonard Susskind, Brian Swingle, and Ying Zhao. Complexity, action, and black holes. *Physical Review D*, 93(8):086006, 2016.

[37] Jonas Haferkamp, Philippe Faist, Naga BT Kothakonda, Jens Eisert, and Nicole Yunger Halpern. Linear growth of quantum circuit complexity. *Nature Physics*, 18(5):528–532, 2022.

[38] Zhi Li. Short proofs of linear growth of quantum circuit complexity. *arXiv preprint arXiv:2205.05668*, 2022.

[39] Anurag Anshu and Srinivasan Arunachalam. A survey on the complexity of learning quantum states. *arXiv preprint arXiv:2305.20069*, 2023.

[40] M. Hinsche, M. Ioannou, A. Nietner, J. Haferkamp, Y. Quek, D. Hangleiter, J.-P. Seifert, J. Eisert, and R. Sweke. One t gate makes distribution learning hard. *Phys. Rev. Lett.*, 130:240602, Jun 2023.

[41] Gregory Rosenthal and Henry Yuen. Interactive proofs for synthesizing quantum states and unitaries. *CoRR*, abs/2108.07192, 2021.

[42] Tony Metger and Henry Yuen. stateqip = statepspace, 2023.

[43] Vadim Lyubashevsky, Chris Peikert, and Oded Regev. On ideal lattices and learning with errors over rings. In *Advances in Cryptology–EUROCRYPT 2010: 29th Annual International Conference on the Theory and Applications of Cryptographic Techniques, French Riviera, May 30–June 3, 2010. Proceedings 29*, pages 1–23. Springer, 2010.

[44] Nengkun Yu and Tzu-Chieh Wei. Learning marginals suffices! *arXiv preprint arXiv:2303.08938*, 2023.

[45] Lisa Yang and Netta Engelhardt. The complexity of learning (pseudo) random dynamics of black holes and other chaotic systems. *arXiv preprint arXiv:2302.11013*, 2023.

[46] Amira Abbas, Robbie King, Hsin-Yuan Huang, William J Huggins, Ramis Movassagh, Dar Gilboa, and Jarrod R McClean. On quantum backpropagation, information reuse, and cheating measurement collapse. *arXiv preprint arXiv:2305.13362*, 2023.

[47] Kyle Poland, Kerstin Beer, and Tobias J Osborne. No free lunch for quantum machine learning. *arXiv preprint arXiv:2003.14103*, 2020.

[48] Kunal Sharma, Marco Cerezo, Zoë Holmes, Lukasz Cincio, Andrew Sornborger, and Patrick J Coles. Reformulation of the no-free-lunch theorem for entangled datasets. *Physical Review Letters*, 128(7):070501, 2022.

[49] Lukas Gonon and Antoine Jacquier. Universal approximation theorem and error bounds for quantum neural networks and quantum reservoirs. *arXiv preprint arXiv:2307.12904*, 2023.

[50] Adrián Pérez-Salinas, David López-Núñez, Artur García-Sáez, Pol Forn-Díaz, and José I Latorre. One qubit as a universal approximant. *Physical Review A*, 104(1):012405, 2021.

[51] Maria Schuld, Ryan Sweke, and Johannes Jakob Meyer. Effect of data encoding on the expressive power of variational quantum-machine-learning models. *Physical Review A*, 103(3):032430, 2021.

[52] Alberto Manzano, David Dechant, Jordi Tura, and Vedran Dunjko. Parametrized quantum circuits and their approximation capacities in the context of quantum machine learning. *arXiv preprint arXiv:2307.14792*, 2023.

[53] Costin Bădescu and Ryan O'Donnell. Improved quantum data analysis. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, pages 1398–1411, 2021.

[54] Hsin-Yuan Huang, Richard Kueng, and John Preskill. Predicting many properties of a quantum system from very few measurements. *Nature Physics*, 16(10):1050–1057, 2020.

[55] Thomas Chen, Shivam Nadimpalli, and Henry Yuen. Testing and learning quantum juntas nearly optimally. In *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1163–1185. SIAM, 2023.

[56] Vivek V Shende, Stephen S Bullock, and Igor L Markov. Synthesis of quantum logic circuits. In *Proceedings of the 2005 Asia and South Pacific Design Automation Conference*, pages 272–275, 2005.

[57] A. S. Holevo. Some estimates of the information transmitted by quantum communication channels. *Probl. Inf. Transm.*, 9(3):177–183, 1973.

[58] Jeongwan Haah, Aram W Harrow, Zhengfeng Ji, Xiaodi Wu, and Nengkun Yu. Sample-optimal tomography of quantum states. In *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, pages 913–925, 2016.

[59] Juha J Vartiainen, Mikko Möttönen, and Martti M Salomaa. Efficient decomposition of quantum gates. *Physical review letters*, 92(17):177902, 2004.

[60] Jessica Bavaresco, Mio Murao, and Marco Túlio Quintino. Unitary channel discrimination beyond group structures: Advantages of sequential and indefinite-causal-order strategies. *Journal of Mathematical Physics*, 63(4), 2022.

[61] Dominic W Berry, Andrew M Childs, and Robin Kothari. Hamiltonian simulation with nearly optimal dependence on all parameters. In *2015 IEEE 56th Annual Symposium on Foundations of Computer Science*, pages 792–809. IEEE, 2015.

[62] Richard Cleve, Daniel Gottesman, Michele Mosca, Rolando D Somma, and David Yonge-Mallo. Efficient discrete-time simulations of continuous-time quantum query algorithms. In *Proceedings of the forty-first annual ACM symposium on Theory of computing*, pages 409–416, 2009.

[63] Zhengfeng Ji, Yi-Kai Liu, and Fang Song. Pseudorandom quantum states. In *Advances in Cryptology–CRYPTO 2018: 38th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 19–23, 2018, Proceedings, Part III 38*, pages 126–152. Springer, 2018.

[64] Oded Goldreich, Shafi Goldwasser, and Silvio Micali. How to construct random functions. *Journal of the ACM (JACM)*, 33(4):792–807, 1986.

[65] Srinivasan Arunachalam, Alex Bredariol Grilo, and Aarthi Sundaram. Quantum hardness of learning shallow classical circuits. *SIAM Journal on Computing*, 50(3):972–1013, 2021.

[66] Zvika Brakerski and Omri Shmueli. (pseudo) random quantum states with binary phase. In *Theory of Cryptography Conference*, pages 229–250. Springer, 2019.