Noise-tolerant learnability of shallow quantum circuits from statistics and the cost of quantum pseudorandomness

Chirag Wadhwa*,[1] and Mina Doosti[1]

[1] *School of Informatics, University of Edinburgh, United Kingdom*

Abstract

This work studies the learnability of quantum circuits in the near term. We show the natural robustness of quantum statistical queries for learning quantum processes and provide an efficient way to benchmark global depolarizing noise from statistics, which gives us a powerful framework for developing noise-tolerant algorithms. We adapt a learning algorithm for constant-depth quantum circuits to the quantum statistical query setting with a small overhead in the query complexity. We prove average-case lower bounds for learning random quantum circuits of logarithmic and higher depths within diamond distance with statistical queries. Finally, we prove that pseudorandom unitaries (PRUs) cannot be constructed using circuits of constant depth by constructing an efficient distinguisher and proving a new variation of the quantum no-free lunch theorem.

1 Introduction

The problem of learning quantum processes is a fundamental question in quantum information, learning theory, benchmarking and cryptography among several other areas. However, tomography of quantum processes in general is known to require an exponential amount of data and computational time [1, 2]. Even so, many advances have been made to improve the efficiency of learning processes by focusing either on specific classes of quantum processes with special structure [3, 4, 5] or by focusing on learning specific properties of processes rather than fully characterizing them [6]. An interesting property to focus on is the depth of the quantum circuit, especially in the near term, where devices are prone to noise and cannot be used to implement deep circuits. The utility of shallow quantum circuits has been widely studied [7, 8], making the learnability of shallow quantum circuits an interesting area of research, which has already been explored in [4, 5].

Another important property to consider in the near term is the prevalence of noise. Characterizing and correcting the noise of quantum devices are active areas of research that are crucial for the progress of the field of quantum computing [9, 10, 11]. In classical learning theory, the study of learnability under noise was initiated by the work of Kearns [12], who proposed the statistical query model as a naturally robust model. Indeed, an algorithm that can learn a function using only statistical properties can be implemented robustly in a noisy setting. Recently, statistical queries

*chirag.wadhwa@ed.ac.uk

have also found a lot of interest in quantum learning theory, with recent studies of quantum statistical queries for learning many properties of quantum states and processes [13, 14, 15, 16, 17, 18]. In this work, we will focus on quantum statistical queries for learning quantum processes, and show their applicability in the development of robust algorithms. Further, our focus will be on the learnability of shallow quantum circuits, for which we will prove both positive and negative results. Our work signifies the use of quantum statistical queries as a powerful theoretical tool for studying the learnability of quantum processes in the near term.

Another field closely tied to learning theory is cryptography. Even though the techniques used in both fields are quite different, a lot of results in cryptography have been proven using tools from learning theory [19, 20] and vice-versa [4, 21]. An area of research in quantum cryptography that has attracted much attention over the past few years is the notion of quantum pseudorandomness [22], where one hopes to replace truly random quantum objects with deterministic objects that are computationally indistinguishable from truly random ones. Applications of quantum pseudorandomness in cryptography have been studied in [23]. Recently, it has been shown that quantum pseudorandomness may exist independent of classical cryptographic assumptions [24, 25], making it a key focus of the community. In particular, a line of recent work [26, 27] has focused on proving a secure construction for pseudorandom unitaries. While no fully secure construction has been found so far, these constructions do satisfy weaker notions of security. An interesting work is that of [28], where Haug *et al.* showed certain properties pseudorandom unitaries must possess. Another work along these lines was that of [29], where Grewal *et al.* showed that pseudorandom quantum states require a high number of non-Clifford gates, by constructing an efficient distinguisher using a learning algorithm for states prepared with few non-Cliffords. Studying such properties and analyzing the required resources is an important direction to develop efficient and secure constructions for PRUs. Towards this goal, we will use results from learning theory to prove that pseudorandom unitaries cannot be implemented by shallow circuits. In what follows we detail our technical contributions.

1.1 Our contributions

Quantum Statistical Query oracles: Quantum statistical query (QSQ) oracles have previously been studied for the task of learning quantum states [13, 14, 16] and quantum processes [15, 17, 18]. Recently, a new multi-copy QSQ oracle was introduced in [16]. Expanding on these definitions, we propose two new quantum statistical query oracles. First, we define a *multi-copy quantum statistical query oracle for the task of learning a quantum process*. With this oracle, we allow the learner to query a k-register state, to which k-copies of the process are applied in parallel, followed by a simultaneous measurement over the k registers. The learner is allowed to query states that are entangled across the k registers and perform entangled measurements across the registers. This oracle allows us to model learners that have stronger capabilities compared to the single-copy setting for processes but are still restricted compared to general black-box learners. An interesting feature of this oracle is that one can use it to model learners with bounded memory. Next, we introduce a *QSQ oracle for learning an observable*. Here, one queries the unknown observable with a state and receives an approximation of the expectation value. This oracle can be viewed as modelling the behaviour of an unknown quantum apparatus that takes copies of a state, performs some measurements and returns a classical expectation value. In this context, the task of learning the unknown observable is equivalent to identifying the physical quantity the apparatus measures and can be used for benchmarking its behaviour.

Noise-tolerance of QPSQ learners: The main motivation behind statistical query learning in the classical learning theory setting has been to provide a framework to study learnability robust to classification noise. The need for developing robust learners is even more crucial in the quantum setting, particularly in the near-term. In this direction, the robustness of QSQ learners for quantum states to various kinds of noise was shown in [30]. We extend this argument to the QSQ oracle for processes (QPSQ), by showing that learners with access to this oracle are robust to general noise models, as long as the noise is within an acceptable threshold. Further, we show that for global depolarizing noise, there is an efficient method to estimate the deviation of the noisy channel from the noiseless one. This method only uses a single 2-copy QPSQ query. Building upon existing no-go results, we argue that one cannot perform such an estimation using a single-copy oracle, making the use of the 2-copy oracle optimal.

QPSQ learner for shallow quantum circuits: Due to their importance in the near term, we focus on learning shallow quantum circuits. We study the learning algorithm for constant-depth quantum circuits proposed in [5], and show that with a slight overhead and some modifications, this algorithm can be carried out using only access to the QPStat oracle. Along with our framework for developing robust algorithms, this implies a way of ensuring that the algorithm from [5] can be implemented in a robust manner, by using a higher number of samples to simulate the statistical queries.

Lower bounds for learning logarithmic-to-linear depth random quantum circuits: Here, we show an *average-case* query-complexity lower bound for learning quantum circuits with depth in the logarithmic to linear regime. In comparison to the lower bound in [5], which showed worst-case hardness for log-depth circuits, our result shows that *on average*, one might be able to develop efficient learning algorithms at this depth. Our lower bound scales exponentially with the depth, providing a nice characterization of the learnability of shallow quantum circuits with respect to their depth. At higher depths, such random quantum circuits form approximate unitary 2-designs, for which query complexity lower bounds were shown in [18]. Thus, our learning algorithm for constant depth, lower bound for log-to-linear depths, and the previous lower bounds for higher depths together characterize the learnability of random quantum circuits from statistical queries at most depth regimes.

Cost of quantum pseudorandomness: Alongside a learning algorithm for constant-depth circuits, Huang *et al.* [5] show an algorithm for verifying the output of their learning algorithm within an average distance d_{avg}. By extending the quantum no-free lunch theorem from [31], we show the average-case hardness of learning unitaries over the Haar measure with black-box access within bounded d_{avg}. Thus, we use the learner and verifier from [5] in a black-box manner to construct a distinguisher between Haar random unitaries and unitaries implementable by constant-depth circuits. As a result, we prove that *constant-depth circuits cannot form pseudorandom unitaries (PRUs)*.

1.2 Related work

Quantum statistical queries for learning quantum processes were proposed by [18]. [32] also consider a similar access model and show a low-degree learning algorithm in this model, allowing them to learn QAC^0 channels with limited auxiliary qubits. [18] also shows a general average-case query complexity lower bound for learning processes within diamond distance. This lower bound follows from the general bound of [16], which was shown using a technique analogous to the classical one proposed by [33]. [16] also defined multi-copy QSQs for learning quantum states. Lower bounds for learning the output distributions of shallow quantum circuits under a weaker access model were shown in [15]. Here, our lower bound is for learning shallow circuits within diamond distance. The lower bounds in this work and that of [15] use techniques developed in [34, 35]. The learnability of quantum states from quantum statistical queries has been studied in [13, 14, 16]. A model for learning unitaries from statistical queries to the Choi-Jamiolkowski state was proposed in [17].

The problem of learning unknown observables from statistical data was recently considered in [36]. In Section 3, we will show that the access model we consider is stronger than theirs.

The algorithm for learning constant-depth circuits using classical shadows of the circuit was proposed in [5]. A related algorithm for learning the output of arbitrary quantum processes was proposed in [6], and its quantum statistical query analogue was shown in [18].

In [37], the authors show average-case hardness for learning Haar-random unitaries as well as pseudorandom unitaries. While their lower bound for Haar-random unitaries is stronger than ours, it only holds for learning algorithms whose hypothesis is unitary (or close to one). Our weaker bound in Theorem 5 holds even if the hypothesis of the learning algorithm is any CPTP map. Similarly, their hardness for learning PRUs only holds for unitary hypotheses. As the learning algorithm of [5] for constant-depth circuits does not have a unitary hypothesis, the results of [37] do not directly imply that constant-depth circuits cannot form PRUs. Theorems 5 and 6 thus constitute novel contributions.

Note: Since the first version of this work, a new result by [38] showed that PRUs can be constructed optimally in depth $\mathsf{poly}(\log \log n)$, superseding our result in Theorem 6.

1.3 Structure of the paper

In Section 2, we will introduce some preliminary material. Section 3 will introduce QSQ oracles defined previously, as well as the two new oracles defined in this work. In Section 4, we will show the natural robustness of the QPSQ model as well as a method to benchmark various kinds of noise. In Section 5, we will show a learning algorithm for constant-depth circuits. Section 6 covers the lower bounds for learning quantum circuits in logarithmic to linear depth regimes. In Section 7, we will show that shallow circuits cannot be used to construct pseudorandom unitaries.

2 Preliminaries

For basic definitions of quantum computation and information, we refer the reader to [39]. We denote the $N \times N$ identity matrix as I_N and we may omit the index N when the dimension is clear from the context. We will write $\mathcal{M}_{N,N}$ to denote the set of linear operators from $\mathbb{C}^N$ to $\mathbb{C}^N$ and we define the set of quantum states as $\mathcal{S}_N := \{\rho \in \mathcal{M}_{N,N} : \rho \succeq 0, Tr[\rho] = 1\}$. We denote by $U(N)$ the group of N-dimensional unitary operators. We denote by $H_{N,N}$ the set of $N \times N$ Hermitian operators. For a unitary operator U, we may denote the corresponding channel $\mathcal{U} := U(.)U^\dagger$ without explicit definition. We now include important definitions that we will use throughout the paper.

Definition 1 (Pauli operators)**.** The set of Pauli operators is given by

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{1}$$

The set $\mathcal{P}_1 = \{I, X, Y, Z\}$ forms an orthonormal basis for $\mathcal{M}_{2,2}$ with respect to the Hilbert-Schmidt inner product. We will refer to the set of tensor products of Pauli operators and the identity, i.e. the operators of the form $P \in \{I, X, Y, Z\}^{\otimes n} := \mathcal{P}_n$ as *Pauli strings* over n qubits.

Definition 2 (Single-qubit Pauli eigenstates)**.** We define the set of eigenstates of the single-qubit Pauli operators as

$$stab_1 = \{|0\rangle, |1\rangle, |+\rangle, |-\rangle, |+y\rangle, |-y\rangle\}, \tag{2}$$

where $|0\rangle$ & $|1\rangle$ are the eigenstates of Z , $|+\rangle$ & $|-\rangle$ are the eigenstates of X and $|+y\rangle$ & $|-y\rangle$ are the eigenstates of Y.

Definition 3 (Quantum Channels)**.** A map $\mathcal{E} : \mathcal{S}_N \to \mathcal{S}_N$ is said to be completely positive if for any positive operator $A \in \mathcal{M}_{N^2,N^2}, (\mathcal{E} \otimes I)(A)$ is also a positive operator. $\mathcal{E}$ is said to be trace-preserving if for any input density operator $\rho, Tr(\mathcal{E}(\rho)) = 1 = Tr(\rho)$. A quantum process $\mathcal{E}$ is defined as a Completely Positive Trace-Preserving (CPTP) map from one quantum state to another. We may use the terms quantum process and quantum channel interchangeably.

Definition 4 (Maximally Depolarizing Channel)**.** The maximally depolarizing channel Φ_{dep} acting on states in S_N is defined as follows:

$$\Phi_{dep}(\rho) = \frac{Tr(\rho)}{N} I \tag{3}$$

We now define the Haar measure μ_H, which can be thought as the uniform probability distribution over $U(N)$. Similarly, we denote by μ_S the Haar measure over all pure states in the Hilbert space of dimension N. For a comprehensive introduction to the Haar measure and its properties, we refer to [40].

Definition 5 (Haar measure)**.** The Haar measure on the unitary group $U(N)$ is the unique probability measure μ_H that is both left and right invariant over $U(N)$, i.e., for all integrable functions f and for all $V \in U(N)$, we have:

$$\int_{U(N)} f(U) d\mu_H(U) = \int_{U(N)} f(UV) d\mu_H(U) = \int_{U(N)} f(VU) d\mu_H(U). \tag{4}$$

Given a state $|\phi\rangle \in \mathbb{C}^N$, we denote the k-th moment of a Haar random state as

$$\mathbf{E}_{|\psi\rangle \sim \mu_S} \left[|\psi\rangle \langle\psi|^{\otimes k} \right] := \mathbf{E}_{U \sim \mu_H} \left[U^{\otimes k} |\phi\rangle \langle\phi|^{\otimes k} U^{\dagger \otimes k} \right]. \tag{5}$$

Note that the right invariance of the Haar measure implies that the definition of $\mathbf{E}_{|\psi\rangle \sim \mu_S} \left[|\psi\rangle \langle\psi|^{\otimes k} \right]$ does not depend on the choice of $|\phi\rangle$.

Now, we define distance and accuracy measures for quantum states.

Definition 6 (Trace Distance)**.** The trace distance between two quantum states is given by

$$d_{tr}(\rho, \sigma) = \frac{1}{2}\|\rho - \sigma\|_1 \tag{6}$$

Definition 7 (Fidelity)**.** The fidelity between two quantum states ρ and σ is given by

$$F(\rho, \sigma) = Tr\left(\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\right)^2 \tag{7}$$

In particular, when both states are pure, the fidelity is given by

$$F(|\psi\rangle\langle\psi|, |\phi\rangle\langle\phi|) = \left|\langle\psi|\phi\rangle\right|^2 \tag{8}$$

When at least one of the states is pure,

$$F(|\psi\rangle\langle\psi|, \rho) = \langle\psi|\rho|\psi\rangle \tag{9}$$

and we also have the following relation between the fidelity and trace distance.

$$1 - F(|\psi\rangle\langle\psi|, \rho) \leq d_{tr}(|\psi\rangle\langle\psi|, \rho) \tag{10}$$

Next, we define two notions of distance between quantum channels. The first one is the diamond distance, which is a worst-case distance over all states, while the other distance is the average infidelity of the output states of the channels over Haar-random inputs.

Definition 8 (Diamond norm and diamond distance)**.** For a quantum process $\mathcal{E} : \mathcal{S}_N \to \mathcal{S}_N$, and I the identity operator in $\mathcal{M}_{N,N}$, we define the diamond norm $\| \ . \ \|_\diamond$

$$\|\mathcal{E}\|_\diamond = \max_{\rho \in \mathcal{S}_{N^2}} \left\|(\mathcal{E} \otimes I)(\rho)\right\|_1 \tag{11}$$

We then define the diamond distance $d_\diamond$:

$$d_\diamond(\mathcal{E}_1, \mathcal{E}_2) = \frac{1}{2}\|\mathcal{E}_1 - \mathcal{E}_2\|_\diamond \tag{12}$$

Definition 9 (Average distance)**.** We define the average distance between two channels d_{avg} as the infidelity of the output states on average over Haar-random inputs.

$$d_{avg}(\mathcal{E}_1, \mathcal{E}_2) = 1 - \mathop{\mathbf{E}}_{|\psi\rangle \sim \mu_S}\left[F\left(\mathcal{E}_1(|\psi\rangle\langle\psi|), \mathcal{E}_2(|\psi\rangle\langle\psi|)\right)\right] \tag{13}$$

Now, we state a relation between the average and diamond distances, when at least one of the channels is unitary.

Lemma 1 (Average distance and diamond distance)**.** *Consider a unitary $U \in U(N)$ and the associated unitary channel $\mathcal{U}$, as well as a CPTP map $\mathcal{E} : \mathcal{S}_N \to \mathcal{S}_N$. Then,*

$$d_{avg}(\mathcal{U}, \mathcal{E}) \leq d_\diamond(\mathcal{U}, \mathcal{E}) \tag{14}$$

Proof.

$$
\begin{aligned}
d_{avg}(\mathcal{U},\mathcal{E}) &= \mathop{\mathbf{E}}_{|\psi\rangle\sim\mu_S}\Big[1 - F\big(\mathcal{U}(|\psi\rangle\langle\psi|),\mathcal{E}(|\psi\rangle\langle\psi|)\big)\Big] \\
&\leq \mathop{\mathbf{E}}_{|\psi\rangle\sim\mu_S}\Big[d_{tr}\big(\mathcal{U}(|\psi\rangle\langle\psi|),\mathcal{E}(|\psi\rangle\langle\psi|)\big)\Big] \\
&= \frac{1}{2}\mathop{\mathbf{E}}_{|\psi\rangle\sim\mu_S}\big\|\mathcal{U}(|\psi\rangle\langle\psi|) - \mathcal{E}(|\psi\rangle\langle\psi|)\big\|_1 \\
&\leq \frac{1}{2}\max_\rho\big\|\mathcal{U}(\rho) - \mathcal{E}(\rho)\big\|_1 \\
&\leq \frac{1}{2}\max_\rho\big\|(\mathcal{U}\otimes I)(\rho) - (\mathcal{E}\otimes I)(\rho\big\|_1 \\
&= d_\diamond(\mathcal{U},\mathcal{E})
\end{aligned}
\tag{15}
$$

where the first inequality uses the fact that the output of a unitary on a pure state is pure, and that when one of the states is pure, $1 - F \leq d_{tr}$ □

3 Quantum Statistical Query Oracles

In this section, we start by recalling previously defined statistical query oracles for learning quantum states and processes. We then extend these definitions naturally to define a new oracle for learning unknown observables and a new multi-copy statistical query oracle for learning quantum processes. To help keep track of all the oracles, we summarize them in Table 1 at the end of the section.

Definition 10 (QSQs for learning quantum states cf. [13, 14, 16])**.** A quantum statistical query oracle QStat_ρ associated with a state $\rho \in \mathcal{S}^N$ takes as input an observable $O \in H_{N,N}$ with $\|O\|_\infty \leq 1$ and tolerance $\tau \in \mathbb{R}, \tau > 0$, and returns $\alpha \in \mathbb{R}$ satisfying

$$|\alpha - Tr(O\rho)| \leq \tau \tag{16}$$

Definition 11 (Multi-copy QSQs for learning quantum states cf. [16])**.** A multi-copy quantum statistical query oracle MQStat_ρ^k associated with a state $\rho \in \mathcal{S}^N$ takes as input an observable $O \in H_{N^k,N^k}$ with $\|O\|_\infty \leq 1$ and tolerance $\tau \in \mathbb{R}, \tau > 0$, and returns $\alpha \in \mathbb{R}$ satisfying

$$|\alpha - Tr(O\rho^{\otimes k})| \leq \tau \tag{17}$$

For $k = 2$, we denote the oracle by $\mathsf{2QStat}_\rho$ instead.

Definition 12 (QSQs for learning quantum processes (QPSQs) cf. [18])**.** A quantum statistical query oracle $\mathsf{QPStat}_\mathcal{E}$ associated with a quantum process $\mathcal{E} : \mathcal{S}_N \to \mathcal{S}_N$, takes as input an observable $O \in H_{N,N}$ with $\|O\|_\infty \leq 1$, a state $\rho \in \mathcal{S}^N$, and a tolerance $\tau \in \mathbb{R}, \tau > 0$ and outputs a number α satisfying

$$|\alpha - \mathrm{Tr}\big(O\mathcal{E}(\rho)\big)| \leq \tau \tag{18}$$

We will now define two new quantum statistical query oracles. We naturally extend the definition of quantum statistical queries to the multi-copy setting for quantum processes. Here, we allow the learner to query the oracle with a state that may be entangled across k registers as well as with entangled measurements while applying k-copies of the process in parallel.

Definition 13 (Multi-copy QPSQs)**.** A multi-copy quantum statistical query oracle $\mathsf{MQPStat}^k_{\mathcal{E}}$ associated with a quantum process $\mathcal{E} : \mathcal{S}_N \rightarrow \mathcal{S}_N$, takes as input an observable $O \in H_{N^k,N^k}$ with $\|O\|_\infty \leq 1$, a state $\rho \in \mathcal{S}^{N^k}$, and a tolerance $\tau \in \mathbb{R}, \tau > 0$ and outputs a number α satisfying

$$|\alpha - \mathrm{Tr}\Big(O\mathcal{E}^{\otimes k}(\rho)\Big)| \leq \tau \tag{19}$$

Again, for $k = 2$, we denote the oracle by $\mathsf{2QPStat}_{\mathcal{E}}$ instead.

We now present a new definition of quantum statistical queries for learning unknown observables.

Definition 14 (QSQs for learning observables)**.** A quantum statistical query oracle QStat_O associated with an observable $O \in H_{N,N}$ with $\|O\|_\infty \leq 1$, takes as input a state $\rho \in \mathcal{S}^N$ and tolerance $\tau \in \mathbb{R}, \tau > 0$, and returns $\alpha \in \mathbb{R}$ satisfying

$$|\alpha - Tr(O\rho)| \leq \tau \tag{20}$$

The task of learning unknown observables was recently considered by Molteni *et al.* in [36]. There is a distinction between the access model considered in [36] and the oracle QStat_O. Specifically, Molteni *et al.* consider learners with access to random examples of the form $(|\psi_x\rangle, \alpha)$, where $|\psi_x\rangle$ is a classically-described quantum state w.r.t a classical string x, and α is an estimation of the expectation value of the unknown observable on this state. One can see that QStat_O is an approximate evaluation oracle (as defined in [16]) for the function $f_O : \rho \rightarrow Tr(O\rho)$, while Molteni *et al.* consider access to random examples of this function, making QStat_O a stronger access model than that considered in [36].

Oracle	Object to learn	Inputs	Output($\pm\tau$)
QStat_ρ	$\rho \in \mathcal{S}_N$	$O \in H_{N,N}, \|O\|_\infty \leq 1$	$Tr(O\rho)$
MQStat^k_ρ	$\rho \in \mathcal{S}_N$	$O \in H_{N^k,N^k}, \|O\|_\infty \leq 1$	$Tr(O\rho^{\otimes k})$
$\mathsf{QPStat}_{\mathcal{E}}$	$\mathcal{E} : \mathcal{S}_N \rightarrow \mathcal{S}_N$	$\rho \in \mathcal{S}_N, O \in H_{N,N}, \|O\|_\infty \leq 1$	$Tr(O\mathcal{E}(\rho))$
$\mathsf{MQPStat}^k_{\mathcal{E}}$ [*This work*]	$\mathcal{E} : \mathcal{S}_N \rightarrow \mathcal{S}_N$	$\rho \in \mathcal{S}_{N^k}, O \in H_{N^k,N^k}, \|O\|_\infty \leq 1$	$Tr(O\mathcal{E}^{\otimes k}(\rho))$
QStat_O [*This work*]	$O \in H_{N,N}, \|O\|_\infty \leq 1$	$\rho \in \mathcal{S}_N$	$Tr(O\rho)$

Table 1: Summary of quantum statistical query oracles

4 Noise-tolerance of QPSQ

Statistical query algorithms, in classical learning theory, are known to be robust to classification noise [12]. A similar result was shown for quantum statistical queries in [30] for classification noise, depolarizing noise and any bounded noise. We now demonstrate the noise-robustness of learners with access to the QPStat oracle. Informally, we show that a QPSQ learning algorithm for a class of quantum channels in a noiseless setting can successfully learn the same class even when given noisy data; more specifically, with QPStat access to the noisy version of the channel (within a noise threshold that is not too high). In particular, the number of queries needed and the algorithm itself remain unchanged, while the tolerance of the queries needs to be lowered. The statement is formalised as follows:

Theorem 1 (Noise tolerance of QPSQ learner)**.** *Suppose there exists a learning algorithm that learns a class of channels* $\mathcal{C} = \{\mathcal{E}_i\}_i$ *using* q *queries to the oracle* $\mathsf{QPStat}_{\mathcal{E}}$ *for an unknown channel* $\mathcal{E} \in \mathcal{C}$*, with tolerance at least* τ*, for observables* $\{O_j\}_{j\in[q]}$*, with* $\left\|O_j\right\|_\infty \leq 1$*,* $\forall j \in [q]$*. Let* Λ *be any unknown noise channel with the guarantee that for some* $\eta, 0 < \eta < \tau$*, we have*

$$\left\|\Lambda(\mathcal{E}_i) - \mathcal{E}_i\right\|_\diamond \leq \eta \quad \forall i \tag{21}$$

Then, there exists an algorithm for learning $\mathcal{C}$ *given access to* $\mathsf{QPStat}_{\Lambda(\mathcal{E}_i)}$*, using* q *queries of tolerance at least* $\tau - \eta$*.*

Proof. Let the algorithm in the noisy scenario make the exact same queries as in the noiseless case while changing the tolerance from τ to $\tau - \eta$. Then, on input O, ρ, the oracle responds with α such that

$$|\alpha - Tr(O\Lambda(\mathcal{E}_i)(\rho))| \leq \tau - \eta \tag{22}$$

By the definition of $\|.\|_\diamond$ and the matrix Hölder inequality, we have

$$|Tr(O\mathcal{E}(\rho))| \leq \|O\|_\infty \|\rho\|_1 \|\mathcal{E}\|_\diamond \leq \|\mathcal{E}\|_\diamond \tag{23}$$

Therefore,

$$|Tr(O\mathcal{E}_i(\rho)) - Tr(O\Lambda(\mathcal{E}_i)(\rho))| \leq \left\|\mathcal{E}_i - \Lambda(\mathcal{E}_i)\right\|_\diamond \leq \eta \tag{24}$$

From triangle inequality, we have

$$|\alpha - Tr(O\mathcal{E}_i(\rho))| \leq \tau \tag{25}$$

Thus, the noisy learner now receives data with identical guarantees to the noiseless case and can proceed in the same way. □

4.1 Estimating noise with a single query to 2QPStat

As Theorem 1 requires us to reduce the query tolerance by the value of the diamond norm between the noisy and noiseless channels, it is important to identify this quantity. In many situations, such an upper bound may already be known from previous benchmarking. However, it would be interesting to see if such a bound can be obtained directly using statistical queries. We show a method to perform such an estimation when the noise is global depolarizing in Theorem 2.

An immediate challenge to this problem is that the strength of the depolarizing noise is closely related to the purity of the output state, which has been shown to be hard to estimate from single-copy quantum statistical queries [14]. Nevertheless, it was shown in [16] that a single query to 2QStat suffices to estimate the purity of a state. We use this single-query purity estimation method to characterize the depolarizing strength from a single query to 2QPStat. For our estimation to succeed, we require prior knowledge of an upper bound on the depolarizing noise. However, this bound need not be tight. For example, $\gamma \leq 1/2$ is a good enough bound for our purposes and would hold in most practical situations.

Theorem 2 (Estimating noise with a single 2QPStat query)**.** *Consider a noiseless unitary channel* $\mathcal{U} = U(.)U^\dagger$*, and the corresponding noisy channel* $\Lambda(\mathcal{U}) = \Lambda(\gamma) \circ \mathcal{U}$*, where* $\Lambda(\gamma)$ *is the depolarizing channel with strength* γ

$$\Lambda(\gamma) : \rho \rightarrow (1-\gamma)\rho + \gamma Tr(\rho)\frac{\mathbb{I}}{d} \tag{26}$$

Then, given an initial (loose) upper bound $\gamma \leq \gamma_u$, there exists a method to characterize the noise

$$\left\|\mathcal{U} - \Lambda(\mathcal{U})\right\|_\diamond \in [l, u] \tag{27}$$

such that $u - l \leq \epsilon$, for $0 < \epsilon \leq 1 - \gamma_u$, using a single query to $\mathsf{2QPStat}_{\Lambda(\mathcal{U})}$ with tolerance $\tau = \Theta((1-\gamma_u)\epsilon)$.

Proof sketch. We will use a single query to $\mathsf{2QPStat}$ to estimate the purity of the output state.First, observe that

$$Tr(\mathbb{F}(\Lambda(\mathcal{U})(|\psi\rangle\langle\psi|))^{\otimes 2}) = Tr((\Lambda(\mathcal{U})(|\psi\rangle\langle\psi|))^2) \tag{28}$$

gives us the purity of the output state, where $\mathbb{F}$ is the flip operator. A τ-accurate estimate of this quantity can be obtained by making a query of the form $\mathsf{2QPStat}_{\Lambda(\mathcal{U})}(|\psi\rangle\langle\psi|^{\otimes 2}, \mathbb{F}, \tau)$. Note that the purity of the output state will be the same for any pure input $|\psi\rangle$. For simplicity, we choose $|\psi\rangle = |0\rangle$. We define the output state

$$\rho^{out} = (1-\gamma)U|0\rangle\langle 0|U^\dagger + \gamma\frac{\mathbb{I}}{2^n} \tag{29}$$

We have

$$(\rho^{out})^2 = (1-\gamma)^2 U|0\rangle\langle 0|U^\dagger + \frac{2\gamma(1-\gamma)}{2^n}U|0\rangle\langle 0|U^\dagger + \gamma^2\frac{\mathbb{I}}{2^n} \tag{30}$$

This gives us the purity of the output state

$$\begin{aligned} Tr((\rho^{out})^2) &= (1-\gamma)^2 + \frac{2\gamma(1-\gamma)}{2^n} + \frac{\gamma^2}{2^n} \\ &= 1 - (2\gamma - \gamma^2)(1 - \frac{1}{2^n}) \end{aligned} \tag{31}$$

Suppose the query to $\mathsf{2QPStat}$ gives us a quantity α

$$\alpha \leftarrow \mathsf{2QPStat}_{\Lambda(\mathcal{U})}(|0\rangle\langle 0|^{\otimes 2}, \mathbb{F}, \tau) \tag{32}$$

Then, by definition of the oracle,

$$\alpha \in \Big[1 - (2\gamma - \gamma^2)(1 - 2^{-n}) - \tau, 1 - (2\gamma - \gamma^2)(1 - 2^{-n}) + \tau\Big] \tag{33}$$

We will use this range of the estimated purity to find a range for γ, and then extend it to characterize the diamond distance between the noiseless and noisy channels. We defer the rest of the analysis to Appendix A. □

In practice, we are interested in the regime when $\tau \geq 1/\mathsf{poly}(n)$, and thus $\epsilon \geq 1/\mathsf{poly}(n)$ and $1 - \gamma_u \geq 1/\mathsf{poly}(n)$. For known upper bounds on the depolarizing strength that are tighter than being arbitrarily close to 1, one can thus obtain an estimate of the diamond norm between the noisy and noiseless channels up to inverse-polynomial precision using a single query to $\mathsf{2QPStat}$ of inverse-polynomial tolerance.

As stated earlier, this method only applies to global depolarizing noise. While this is a physically relevant noise model, it would be interesting to consider the estimation of more general kinds of noise from statistical queries.

5 QPSQ learner for shallow circuits

In this section, we show an efficient algorithm for learning constant-depth circuits within diamond distance. Huang *et al.* [5] proposed a learning algorithm for this problem using classical shadows. The algorithm proceeds by learning all $3n$ single-qubit Pauli observables after Heisenberg-evolution under the unknown circuit, using classical shadows of the circuit, similar to the algorithm of [6]. Then, the algorithm combines these learned observables using an innovative *circuit sewing* process. The algorithm only makes random queries to the circuit to learn the observables, and has a sample complexity of $\mathcal{O}\left(\frac{n^2 \log(n/\delta)}{\epsilon^2}\right)$. We show that it is possible to learn all the Heisenberg-evolved observables using $\mathcal{O}\left(\frac{n^3 \log(n/\delta)}{\epsilon^2}\right)$ queries to $\mathsf{QPStat}_{\mathcal{U}}$. This allows us to efficiently learn arbitrary constant-depth circuits using statistical queries within bounded diamond distance. Thus, with a linear overhead in the query complexity, we can gain the robustness guarantees of Theorem 1 for learning any constant-depth circuit.

First, we state the quantum statistical query algorithm for learning a few-body observable with unknown support. We denote the support of an observable O, i.e. the set of qubits it acts on, by $\mathsf{supp}(O)$.

Lemma 2 (Learning a few-body observable with unknown support from QSQs)**.** *There exists a QSQ algorithm for learning an unknown n-qubit observable O with $\|O\|_\infty \leq 1$ that acts on an unknown set of k qubits , such that with probability at least $1-\delta$, the learned observable $\hat{O}$ satisfies*

$$\left\|\hat{O} - O\right\| \leq \epsilon \quad \text{and} \quad \mathsf{supp}(\hat{O}) \subseteq \mathsf{supp}(O) \tag{34}$$

using

$$N = \frac{2^{\mathcal{O}(k)} \log(n/\delta)}{\epsilon^2} \tag{35}$$

queries to QStat_O *of tolerance*

$$\tau = \frac{\epsilon}{4\left(6\sqrt{2}\right)^k} \tag{36}$$

running in computational time $\mathcal{O}(n^k \log(n/\delta)/\epsilon^2)$

Proof. Consider $O = \sum_{P \in \mathcal{P}_n, |P| \leq k} \alpha_P P$. The Pauli coefficients α_P can be represented as

$$\alpha_P = 3^{|P|} \mathop{\mathbf{E}}_{|\psi\rangle \sim \mathrm{stab}_1^{\otimes n}} \langle\psi|O|\psi\rangle\langle\psi|P|\psi\rangle \tag{37}$$

The algorithm makes random queries and uses the output to estimate each coefficient. Specifically, it makes N queries $\mathsf{QStat}_O(|\psi_l\rangle\langle\psi_l|, \tau)$ for $l \in [N]$, with random input states $|\psi_l\rangle \sim \mathrm{stab}_1^{\otimes n}$. Denote the output of the l^{th} query by y_l. Denote by $\hat{\alpha}_P$ the estimate

$$\hat{\alpha}_P = \frac{3^{|P|}}{N} \sum_{l \in [N]} y_l \langle\psi_l|P|\psi_l\rangle \tag{38}$$

Define the coefficients of the learned observables

$$\hat{\beta}_P = \begin{cases} \hat{\alpha}_P & \hat{\alpha}_P \geq 0.5\epsilon/(2\sqrt{2})^k, \\ 0 & \hat{\alpha}_P < 0.5\epsilon/(2\sqrt{2})^k \end{cases} \tag{39}$$

Then, the algorithm outputs the observable

$$\hat{O} = \sum_{P \in \mathcal{P}_n, |P| \leq k} \hat{\beta}_P P \tag{40}$$

Now, we show the correctness of the algorithm. Denote the intermediate quantity $\bar{\alpha}_P$

$$\bar{\alpha}_P = \frac{3^{|P|}}{N} \sum_{l \in [N]} \langle \psi_l | O | \psi_l \rangle \langle \psi_l | P | \psi_l \rangle \tag{41}$$

By the definition of QStat_O,

$$|\hat{\alpha}_P - \bar{\alpha}_P| \leq 3^{|P|}\tau \tag{42}$$

Using N queries as specified in equation 35, with the tolerance τ given in equation 36, from Hoeffding's inequality, we see that with probability at least $1-\delta$, it holds for all $P \in \mathcal{P}_n, |P| \leq k$,

$$|\bar{\alpha}_P - \alpha_P| \leq \frac{\epsilon}{4(2\sqrt{2})^k} \leq \frac{0.5\epsilon}{(2\sqrt{2})^k} - 3^k\tau \leq \frac{0.5\epsilon}{(2\sqrt{2})^k} - 3^{|P|}\tau \tag{43}$$

Thus, N queries of tolerance τ suffice to obtain

$$|\hat{\alpha}_P - \alpha_P| \leq \frac{0.5\epsilon}{(2\sqrt{2})^k} \tag{44}$$

From this point, we follow the presentation of [5]. First, we show that $\mathsf{supp}(\hat{O}) \subseteq \mathsf{supp}(O)$. For all $P \in \mathcal{P}_n$ with α_P, equation 44 tells us $|\hat{\alpha}_P| < \frac{0.5\epsilon}{(2\sqrt{2})^k}$. Thus, $\hat{\beta}_P = 0$, showing that $\mathsf{supp}(\hat{O}) \subseteq \mathsf{supp}(O)$.

Now, we prove the error bound on the learned observable. From the fact that $\alpha_P = 0$ implies $\hat{\beta}_P = 0$, we have

$$\begin{aligned} \hat{O} - O &= \sum_{P \in \mathcal{P}_n : \mathsf{supp}(P) \subseteq \mathsf{supp}(O)} \left(\hat{\beta}_P - \alpha_P \right) P \\ &= \sum_{Q \in \mathcal{P}_k} \left(\hat{\beta}_{P(Q)} - \alpha_{P(Q)} \right) P(Q) \end{aligned} \tag{45}$$

where $P(Q)$ denotes $Q \otimes I_{[n] \setminus \mathsf{supp}(O)}$, where $I_{[n] \setminus \mathsf{supp}(O)}$ is the identity on all qubits outside the support of O. Now,

$$\begin{aligned}
\left\|\hat{O}-O\right\|_\infty &= \left\|\sum_{Q\in\mathcal{P}_k}\left(\hat{\beta}_{P(Q)}-\alpha_{P(Q)}\right)P(Q)\right\|_\infty \\
&= \left\|\sum_{Q\in\mathcal{P}_k}\left(\hat{\beta}_{P(Q)}-\alpha_{P(Q)}\right)Q\right\|_\infty \\
&\leq \sqrt{\sum_{Q\in\mathcal{P}_k}\left(\hat{\beta}_{P(Q)}-\alpha_{P(Q)}\right)^2 Tr(Q^2)} \\
&\leq (2\sqrt{2})^k \max_{|P|\leq k}\left|\hat{\beta}_P-\alpha_P\right| \\
&\leq (2\sqrt{2})^k \max_{|P|\leq k}\left(\left|\hat{\beta}_P-\hat{\alpha}_P\right|+|\hat{\alpha}_P-\alpha_P|\right) \\
&\leq \epsilon
\end{aligned} \tag{46}$$

where the first inequality follows from the fact that $\|A\|_\infty \leq \sqrt{Tr(A^2)}$ for any Hermitian matrix A, the second inequality follows from the facts that Q is a k-qubit Pauli and that there are 4^k terms in the summation, and the third inequality follows from triangle inequality. □

We now state the following lemma from [5] on the size of the support of single-qubit Paulis after Heisenberg evolution under constant depth circuits.

Lemma 3 (Support of Heisenberg-evolved observables is bounded cf. [5], Lemma 14)**.** *Given an n-qubit unitary U generated by a constant-depth circuit. For each qubit $i \in [n]$ and Pauli $P \in \{X, Y, Z\}$, we have*

$$|\mathsf{supp}(U^\dagger P_i U)| = \mathcal{O}(1) \tag{47}$$

Theorem 3 (Learning arbitrary shallow circuits from QPStat queries)**.** *There exists a QPSQ algorithm for learning an unknown n-qubit unitary U generated by a constant-depth circuit over any two-qubit gates in* SU(4)*, and with an arbitrary number of ancilla qubits, such that the algorithm outputs an n-qubit quantum channel $\hat{\mathcal{E}}$ that can be implemented by a constant-depth quantum circuit over $2n$ qubits, which satisfies*

$$\left\|\hat{\mathcal{E}}-\mathcal{U}\right\|_\diamond \leq \epsilon \tag{48}$$

with probability at least $1-\delta$. The algorithm uses

$$N = \mathcal{O}\left(\frac{n^3\log(n/\delta)}{\epsilon^2}\right) \tag{49}$$

queries to $\mathsf{QPStat}_\mathcal{U}$ with tolerance $\tau = \Omega(\epsilon/n)$ and runs in computational time $\mathcal{O}(\mathsf{poly}(n)\log(1/\delta)/\epsilon^2)$

Proof. The algorithm proceeds in two steps. First, using $\mathsf{QPStat}_\mathcal{U}$ queries, we learn the $3n$ Heisenberg-evolved observables $O_{i,P} = U^\dagger P_i U, P \in \{X, Y, Z\}, i \in [n]$. From Lemma 3, we know that for each $O_{i,P}$,

$$\mathsf{supp}(O_{i,P}) = \mathcal{O}(1) \tag{50}$$

Observe that

$$\mathsf{QPStat}_{\mathcal{U}}(\rho, P_i, \tau) \equiv \mathsf{QStat}_{O_{i,P}}(\rho, \tau) \tag{51}$$

Thus, one can run the algorithm of Lemma 2 for learning few-body observables by using queries to QPStat[1]. Unlike the algorithm of [5], we require a new set of queries to be made for each observable to be learned. To learn the unitary successfully, we require the algorithm of Lemma 2 to learn each $O_{i,P}$ independently up to error $\epsilon/6n$ with probability at least $1-\delta/3n$. From equations 35 and 36 we then obtain the desired query complexity $N = \mathcal{O}\left(\frac{n^3 \log(n/\delta)}{\epsilon^2}\right)$ and tolerance $\tau = \Omega(\epsilon/n)$.

After learning all $O_{i,P}$, using the circuit sewing procedure of [5], one can construct a circuit implementing a channel $\mathcal{E}$ over $2n$ qubits such that

$$\left\| \mathcal{E} - \mathcal{U} \otimes \mathcal{U}^\dagger \right\| \leq \epsilon \tag{52}$$

Thus, by tracing out the channel over the ancillary qubits, one can obtain the desired result. Having learned the $3n$ observables, the circuit sewing procedure takes $\mathcal{O}(n)$ computational time. We omit this procedure here and refer to [5] for a detailed analysis. Thus, the overall computational time is $\mathcal{O}(\mathsf{poly}(n)\log(1/\delta)/\epsilon^2)$, and is dominated by learning the Heisenberg-evolved Pauli observables. ☐

Remark 1. We observe that with a slight overhead in the query complexity, we are able to adapt the algorithm from [5] into the QPSQ access model. In practice, one can simulate the outcome of a single QPStat query using $\mathcal{O}(1/\tau^2) = \mathcal{O}(n^2/\epsilon^2)$ separate measurements. Thus, with a sample complexity overhead that is cubic in the number of qubits, and quadratic in the error, this algorithm can be run with the natural robustness guarantees of a statistical query model using the framework of Theorems 1 and 2. It would be interesting to study the robustness of the original algorithm as it stands or to see if it can be made robust with a lower overhead. Extending the analysis of works such as [41, 42] on the robustness of classical shadows for learning properties of states to the setting of learning circuits and general processes may prove useful in this regard.

6 Lower bounds for shallow random quantum circuits

In this section, we will show average-case query complexity lower bounds learning brickwork random quantum circuits with depth at least logarithmic and at most linear in the number of qubits. Before stating the main result of this section, we state some important definitions.

Definition 15 (Unitary t-Designs)**.** We denote the haar measure over $U(N)$ by μ_H. The t-th moment superoperator with respect to a distribution ν over $U(N)$ is defined as

$$\mathcal{M}_\nu^{(t)}(A) = \mathop{\mathbf{E}}_{U\sim\nu}[U^{\otimes t} A (U^\dagger)^{\otimes t}] = \int U^{\otimes t} A (U^\dagger)^{\otimes t} d\nu(U) \tag{53}$$

ν is said to be an exact unitary t-design if and only if

$$\mathcal{M}_\nu^{(t)}(A) = \mathcal{M}_{\mu_H}^{(t)}(A) \tag{54}$$

[1]Notice now the similarity of this algorithm to the algorithms of [6, 18], with the difference being that we are now certain that the support of the Heisenberg-evolved observables is bounded and thus there is no penalty for truncation.

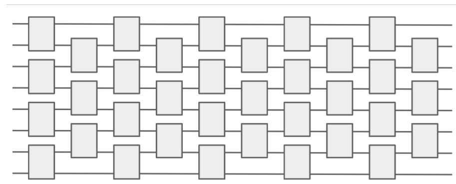

Figure 1: Structure of brickwork random quantum circuits. Each box is a 2-qubit Haar-random gate.

Similarly, ν is said to be an additive δ-approximate unitary t-design if and only if

$$\left\|\mathcal{M}_{\nu}^{(t)}(A)-\mathcal{M}_{\mu_H}^{(t)}(A)\right\|_{\diamond}\leq\delta \tag{55}$$

We denote an exact unitary t-design by $\mu_H^{(t)}$ and an additive δ-approximate unitary t-design by $\mu_H^{(t,\delta)}$.

Definition 16 (Brickwork random quantum circuits)**.** Denote by $\mathsf{RQC}(n,d)$ the measure over brickwork random quantum circuits of n qubits with depth d. $\mathsf{RQC}(n,d)$ consists of unitaries of the form

$$U=(I_2\otimes U_{2,3}^{(d)}\otimes U_{4,5}^{(d)}\otimes\dots)(U_{1,2}^{(d-1)}\otimes U_{3,4}^{(d-1)}\otimes\dots)\dots(U_{1,2}^{1}\otimes U_{3,4}^{1}\otimes\dots) \tag{56}$$

where $U_{i,j}^{(l)}$ are 2-qubit unitaries distributed according to the Haar-measure over $\mathcal{U}(4)$ and I_2 is the identity on a single qubit (Figure 1). $\mathsf{RQC}(n,d)$ for any $d>0$ forms an exact unitary 1-design. At infinite depth, the distribution over brickwork random quantum circuits converges to the Haar measure.

Now, we will state our result on the query-complexity lower bound for learning shallow brickwork random quantum circuits.

Theorem 4 (Average-case lower bound for shallow BRQCs)**.** *Let $N=2^n, 0<\tau\leq\epsilon\leq\frac{1}{3}\left(1-\frac{1}{N}\right), \mathsf{RQC}(n,d)$ be an ensemble of n-qubit brickwork-random quantum circuits of depth d, with*

$$\frac{\log(n)}{\log(5/4)}\leq d\leq\frac{n+\log(n)}{\log(5/4)} \tag{57}$$

Assume there exists an algorithm that with probability β over $U\sim\mathsf{RQC}(n,d)$ and probability α over its internal randomness produces a hypothesis $\hat{\mathcal{E}}$ such that $d_\diamond(\hat{\mathcal{E}},U(.)U^\dagger)\leq\epsilon$, using q queries with tolerance τ to QPStat_U. Then, it holds

$$q+1\geq\Omega\left(\frac{(2\alpha-1)\beta}{n}\tau^2\left(\frac{5}{4}\right)^d\right) \tag{58}$$

Before proving this theorem, we will state some useful lemmas. We start by stating the general lower bound for learning processes up to diamond distance using QPSQs from [18].

Lemma 4 (General lower bound for QPSQ-learning within $d_\diamond$, cf. [18])**.** *Let $0 < \tau \leq \epsilon$, $\mathcal{C} \subseteq \mathcal{S}_N \to \mathcal{S}_N$ be a set of quantum processes , and μ some measure over $\mathcal{C}$. Assume there exists an algorithm that with probability β over $\mathcal{E} \sim \mu$ and probability α over its internal randomness produces a hypothesis $\hat{\mathcal{E}}$ such that $d_\diamond(\hat{\mathcal{E}}, \mathcal{E}) \leq \epsilon$, using q queries with tolerance τ to $\mathsf{QPStat}_{\mathcal{E}}$. Then, for every $\Phi : \mathcal{S}_N \to \mathcal{S}_N$, it holds*

$$q + 1 \geq (2\alpha - 1) \frac{\beta - \mathbf{Pr}_{\mathcal{E}\sim\mu}(d_\diamond(\mathcal{E}, \Phi) < 2\epsilon + \tau)}{\max_{\rho,O} \mathbf{Pr}_{\mathcal{E}\sim\mu}(|Tr(O\mathcal{E}(\rho)) - Tr(O\Phi(\rho))| > \tau)} \tag{59}$$

We will refer to the term in the denominator,

$$\max_{\rho,O} \mathbf{Pr}_{\mathcal{E}\sim\mu}(|Tr(O\mathcal{E}(\rho)) - Tr(O\Phi(\rho))| > \tau)$$

*as the **maximally distinguishable fraction**. We refer to the term in the numerator,*

$$\mathbf{Pr}_{\mathcal{E}\sim\mu}(d_\diamond(\mathcal{E}, \Phi) < 2\epsilon + \tau)$$

*as the **trivial fraction**.*

To prove theorem 4, we will bound both the maximally distinguishable fraction and the trivial fraction with respect to the maximally depolarizing channel as the reference channel. In order to bound the maximally distinguishable fraction, we will first need results on the first and second-order moments over the Haar measure. We state without proof the necessary results from [18, 40] in the following lemma.

Lemma 5 (Moments over the Haar-measure, cf. [18, 40])**.** *The first moment superoperator of the Haar measure μ_H over $U(N)$ is the maximally depolarizing channel.*

$$\mathcal{M}^{(1)}_{\mu_H} = \Phi_{dep} \tag{60}$$

Further, for all $O \in H_{N,N}$ and $\rho \in \mathcal{S}_N$,

$$\mathop{\mathbf{E}}_{U\sim\mu_H}[Tr(OU\rho U^\dagger)] = \frac{Tr(O)}{N} \tag{61}$$

The second moment is given by

$$\mathop{\mathbf{E}}_{U\sim\mu_H}[Tr(OU\rho U^\dagger)^2] = \left(\frac{N - Tr(\rho^2)}{N(N^2-1)}\right) Tr(O)^2 + \left(\frac{NTr(\rho^2) - 1}{N(N^2-1)}\right) Tr(O^2) \tag{62}$$

and the variance is bounded by

$$\mathop{\mathbf{Var}}_{U\sim\mu_H}[Tr(OU\rho U^\dagger)] \leq \frac{1}{N+1} \tag{63}$$

Now, we will prove a bound on the variance of expectation values over $\mathsf{RQC}(n, d)$

Lemma 6 (Bounded variance for $\mathsf{RQC}(n, d)$)**.** *For depth*

$$\frac{\log(n)}{\log(5/4)} \leq d \leq \frac{n + \log(n)}{\log(5/4)} \tag{64}$$

the variance over $\mathsf{RQC}(n,d)$ *is bounded by*

$$\mathop{\mathbf{Var}}_{U\sim\mathsf{RQC(n,d)}}[Tr(OU\rho U^\dagger)] = \mathcal{O}\left(n\left(\frac{4}{5}\right)^d\right) \tag{65}$$

$\forall O \in H_{2^n,2^n}$ *with* $\|O\|_\infty \leq 1$ *and* $\forall \rho \in \mathcal{S}_{2^n}$*.*

Proof sketch. Bounds on low-depth moments for BRQCs have been shown in [15, 35] by counting partitions over a statistical mechanical model, a technique developed originally in [34]. We use the same technique to show an analogous bound for our problem:

$$\mathop{\mathbf{E}}_{U\sim\mathsf{RQC}(n,d)}[Tr(OU\rho U^\dagger)^2] \leq \left(1+\left(\frac{4}{5}\right)^d\right)^{n/2} \mathop{\mathbf{E}}_{U\sim\mu_H}[Tr(OU\rho U^\dagger)^2] \tag{66}$$

Then, using the fact that brickwork random quantum circuits form exact 1-designs at any depth, and substituting in the moments from Lemma 5, the desired result is obtained from direct computation. We show the detailed proof in Appendix B. □

In order to bound the trivial fraction, we will use the following lemma from [18]

Lemma 7 (Unitaries are far from the maximally depolarizing channel, cf. [18])**.** *For any unitary* $U \in U(N)$*, we have :*

$$d_\diamond(U(.)U^\dagger, \Phi_{dep}) \geq 1 - \frac{1}{N} \tag{67}$$

Finally, we are in a position to prove Theorem 4

Proof of Theorem 4. As stated earlier, we will use the maximally depolarizing channel as our reference channel Φ in Lemma 4. For $0 < \tau \leq \epsilon \leq \frac{1}{3}\left(1-\frac{1}{N}\right)$, Lemma 7 tells us that $d_\diamond(U(.)U^\dagger, \Phi_{dep}) > 2\epsilon + \tau$ for all $U \sim \mathsf{RQC}(n,d)$. Thus, the trivial fraction is 0.

As $\mathsf{RQC}(n,d)$ is a 1-design, the expected channel over this ensemble is the maximally depolarizing channel. Thus, we use Chebyshev's inequality to bound the maximally distinguishable fraction.

$$\begin{aligned}\mathop{\mathbf{Pr}}_{U\sim\mathsf{RQC}(n,d)}(|Tr(OU(\rho)U^\dagger) - Tr(O\Phi_{dep}(\rho))| > \tau) &\leq \frac{\mathbf{Var}_{U\sim\mathsf{RQC}(n,d)}[Tr(OU\rho U^\dagger)]}{\tau^2} \\ &= \mathcal{O}\left(\frac{n}{\tau^2}\left(\frac{4}{5}\right)^d\right)\end{aligned} \tag{68}$$

Now, using the bounds on the maximally distinguishable fraction and the trivial fraction, we obtain the desired result

$$q + 1 \geq \Omega\left(\frac{(2\alpha-1)\beta}{n}\tau^2\left(\frac{5}{4}\right)^d\right) \tag{69}$$

□

7 PRUs cannot be shallow

In this section, we will show that constant-depth circuits cannot be used to construct pseudorandom unitaries (PRUs). We start by recalling the definition of PRUs.

Definition 17 (Pseudorandom unitaries (PRUs) cf. [22, 26])**.** Let $n \in \mathbb{N}$ be the security parameter. Let $\mathcal{K}$ denote the key space. An infinite sequence $\mathcal{U} = \{\mathcal{U}_n\}_{n\in\mathbf{N}}$ of n-qubit unitary ensembles $\mathcal{U}_n = \{U_k\}_{k\in\mathcal{K}}$ is said to be pseudorandom if

- **(Efficient computation)** There exists a polynomial-time quantum algorithm $\mathcal{Q}$ such that for all keys $k \in \mathcal{K}$, and any n-qubit pure state $|\psi\rangle$, $\mathcal{Q}(k, |\psi\rangle) = U_k|\psi\rangle$
- **Pseudorandomness** The unitary U_k, for a random key $k \sim \mathcal{K}$, is computationally indistinguishable from a Haar-random unitary $U \sim \mu_H$. In other words, for any quantum polynomial-time algorithm $\mathcal{A}$, it holds that

$$\left| \mathop{\mathbf{Pr}}_{k\sim\mathcal{K}} \left[\mathcal{A}^{U_k}(1^n) = 1\right] - \mathop{\mathbf{Pr}}_{U\sim\mu_H} \left[\mathcal{A}^{U}(1^n) = 1\right] \right| \leq \mathrm{negl}(n) \tag{70}$$

To show that constant-depth unitaries cannot form PRUs, we will construct an efficient distinguisher that achieves non-negligible advantage in distinguishing any constant-depth unitary from Haar-random unitaries. As an overview, our distinguisher will consist of the learning and verification algorithms for constant-depth circuits of [5]. We will use these algorithms in a black-box manner. Given as input any constant-depth circuit, the learning algorithm correctly learns it with high probability. Then, the verification algorithm also passes with high probability. On the other hand, given a Haar-random unitary, the algorithm is not likely to perform well. Rather than analyzing the performance of the algorithm on Haar-random unitaries, we formalize this notion by extending the quantum no-free lunch theorem [31], proving that any algorithm that learns a Haar-random unitary with high probability requires an exponential number of queries. Then, since the learned unitary is not close to the actual one, the verification algorithm fails with high probability. We state the results of the learning and verification algorithms from [5] in the following lemma.

Lemma 8 (Learning and Verification Algorithms for Constant-Depth Circuits, cf. [5])**.** *There exists a learning algorithm $\mathcal{A}_L(n, \epsilon, \delta)$ and a verification algorithm $\mathcal{A}_V(n, \epsilon, \delta)$ such that*

- *(Learning) $\mathcal{A}_L(n, \epsilon, \delta)$ makes $\mathcal{O}(n^2 \log(n/\delta)/\epsilon^2)$ queries to an n-qubit unitary U implemented by a constant-depth circuit composed of two-qubit gates, runs in time $\mathsf{poly}(n)/\epsilon^2$, and outputs a channel $\mathcal{E}$ such that $d_\diamond(U(.)U^\dagger, \mathcal{E}) \leq \epsilon$ with probability at least $1 - \delta$*
- *(Verification) Given a learned implementation of a n-qubit CPTP map $\hat{\mathcal{E}}$ and query access to an unknown CPTP map $\mathcal{C}$, $\mathcal{A}_V(n, \epsilon, \delta)$ makes $\mathcal{O}((n^2 \log(n/\delta)/\epsilon^2)$ queries to $\mathcal{C}$, runs in computational time $\mathcal{O}((n^3 \log(n/\delta)/\epsilon^2)$, and*
  1. *If $d_{avg}(\hat{\mathcal{E}}, \mathcal{C}) > \epsilon$, $\mathcal{A}_V$ outputs **FAIL** with probability at least $1 - \delta$*
  2. *If $d_{avg}(\hat{\mathcal{E}}, \mathcal{C}) \leq \epsilon/12n$ and $d_\diamond(\mathcal{C}^\dagger\mathcal{C} - \mathcal{I}) \leq \epsilon/24n$, $\mathcal{A}_V$ outputs **PASS** with probability at least $1 - \delta$*

Moreover, the queries made by both $\mathcal{A}_L$ and $\mathcal{A}_V$ are pure states.

In other words, $\mathcal{A}_L$ efficiently learns any constant depth quantum circuit, and $\mathcal{A}_V$ ensures that the unknown channel is unitary and that it is learned correctly. Next, we state a theorem on the average-case hardness of learning Haar-random unitaries from black-box queries.

Theorem 5 (Average-case hardness of learning Haar-random unitaries)**.** *Let $\epsilon, \delta \in (0,1)$. Suppose there exists a learning algorithm that queries a unitary U distributed according to the Haar-measure over $U(2^n)$ with q pure states, and outputs a channel $\mathcal{E}$, such that with probability at least $1-\delta$, the channel approximates the Haar-random unitary within average distance ϵ, i.e*

$$\Pr_{U\sim\mu_H}[d_{avg}(U(.)U^\dagger, \mathcal{E}) \leq \epsilon] \geq 1-\delta \tag{71}$$

Then, the algorithm must make at least

$$q \geq 2^n(1-\delta)(1-\epsilon) - 1 \tag{72}$$

queries.

We defer the proof of Theorem 5 to Appendix C. While our proof strategy results in a very loose dependence on ϵ and δ, this bound is sufficient for our purposes. The original quantum no-free-lunch theorem (QNFLT) from [31] gave a similar result, where they lower bounded the error of any learning algorithm on average over both μ_H as well as a randomly sampled dataset, and when the learned channel is a unitary. On the other hand, our result shows a relation between the number of queries and the accuracy as well as the probability, while giving the learner the ability to make adaptive queries. Our result has a couple of notable differences compared to the QNFLT reformulation from [4]. The result of [4] is worst-case over all unitaries and considers access to classical descriptions of output states under bounded noise. [37] also show a similar result to Theorem 5, with a stronger lower bound. However, their result only holds when the hypothesis of the learning algorithm is unitary. QNFLT variations when considering states entangled with ancillary registers, as well as mixed states, have been shown in [4, 43], but queries of this kind are outside the scope of our proof.

Now, we will state and prove our main theorem on the no-go result for PRUs.

Theorem 6 (Constant-depth unitaries cannot form PRUs)**.** *For sufficiently large n, a unitary sampled from an ensemble $\mathcal{C}_n$ over n qubits constructed using circuits composed of 2-qubit gates with depth $\mathcal{O}(1)$ can be distinguished from a random n-qubit unitary from the Haar measure with non-negligible advantage using $\mathcal{O}\left(n^4\log(n)\right)$ queries to the unknown unitary in time $\mathrm{poly}(n)$. Thus, $\mathcal{C}_n$ is not an ensemble of pseudorandom unitaries.*

Proof. We consider our distinguisher $\mathcal{A}$ to be the composition of the two algorithms $\mathcal{A}_L(n, \frac{1}{48n}, 1/6)$ and $\mathcal{A}_V(n, 1/4, 1/6)$. $\mathcal{A}_L$ makes $\mathcal{O}(n^4\log(n))$ queries and $\mathcal{A}_V$ makes $\mathcal{O}(n^2\log(n))$ queries. When acting on a unitary from $\mathcal{C}_n$, $\mathcal{A}_L$ produces a channel $\mathcal{E}$ within $\frac{1}{48n}$ diamond distance of the unitary with probability at least 5/6. Using Lemma 1, we see that the learned channel satisfies $d_{avg}(\mathcal{E}, U(.)U^\dagger) \leq d_\diamond(\mathcal{E}, U(.)U^\dagger) \leq \frac{1}{48n}$. Conditioned on correctly learning, and due to the fact that the queried channel is a unitary, an application of $\mathcal{A}_V$ outputs **PASS** with probability at least 5/6. Overall, a union bound tells us that $\mathcal{A}$ outputs **PASS** with probability at least 2/3.

On the other hand, Theorem 5, with $\epsilon = 1/4$ and $\delta = 1/3$, tells us that when acting on a Haar-random unitary, the channel $\mathcal{E}'$ produced by $\mathcal{A}_L$, which queries the unitary polynomially many times with just pure states, satisfies $d_{avg}(\mathcal{E}', U(.)U^\dagger) \leq 1/4$ with probability less than 1/3. Conditioned on $\mathcal{A}_L$ not learning the unknown unitary, the probability that an application of $\mathcal{A}_V$ outputs **PASS** is at most 1/6. Thus, the probability that $\mathcal{A}$ outputs **PASS** over the Haar-measure is at most $1/3 + 1/6 = 1/2$.

We obtain the advantage of $\mathcal{A}$ in distinguishing $\mathcal{C}_n$ from μ_H is at least $2/3 - 1/2 = 1/6$, which is clearly non-negligible. Moreover, as both $\mathcal{A}_L$ and $\mathcal{A}_V$ run in $\mathsf{poly}(n)$ time, $\mathcal{A}$ is a computationally efficient distinguisher. □

8 Outlook

We have used the natural noise-tolerance of the QPSQ model to demonstrate the significance of statistical data for studying the learnability of quantum circuits in the near term. Our method for characterizing the noise from the entangled statistics applies to global depolarizing noise, and opens up an important line of research - *Can all kinds of noise be efficiently benchmarked using statistical queries?*.

Together, the natural noise tolerance of our access model and the benchmarking method provide a useful framework for developing robust algorithms. By developing algorithms that only require statistical data, one can efficiently make them robust. We have shown how this can be done with the learning algorithm for shallow quantum circuits from [5], by adapting it to our statistical query setting with only a linear overhead in query complexity. Developing other learning algorithms in this access model is a promising line of research towards robust learning.

In the statistical query setting, we have improved upon the worst-case lower bound of [5] for logarithmic depth, by showing an average-case query-complexity lower bound for random quantum circuits of the same depth. In fact, our bound shows that even at logarithmic depth, one might be able to develop efficient learning algorithms that succeed *on average*. A nice property of our lower bound is that up to linear depth, we have shown an exponential scaling of the query complexity with the depth. At greater depths, random quantum circuits converge to approximate 2-designs. For such circuits, exponential hardness for statistical query learning has already been shown in [18]. Our learning algorithm and lower bound thus provide a characterization of the statistical query learnability of random quantum circuits at all depths.

In regards to our quantum statistical query oracle for learning unknown observables, we have only used it as an abstract tool, by simulating it as a part of our learning algorithm. We believe this oracle can have much wider applicability, especially for learning physical experiments. Benchmarking the behaviour of unknown physical apparatus from statistical data is a critical problem in quantum information, and we believe this oracle can prove quite useful in the theoretical study of such problems. The same goes for our multi-copy oracle for learning processes. By allowing multi-copy queries, we are able to model a much wider range of learning scenarios, and it would be quite interesting to observe new separations between multi-copy and single-copy statistical query oracles, similar to the separation for purity testing shown in [16].

Our lower bounds for shallow random quantum circuits only hold when the depth is at least logarithmic. As we do not provide a lower bound at constant depths, one may wonder *if the* QPStat

algorithm of Theorem 3 is optimal. As the technique we have used requires some level of indistinguishability between the outputs of QPStat queries, we believe constant-depth circuits may not be scrambled enough for this technique to provide meaningful lower bounds. Using other techniques to prove QPStat query complexity lower bounds is thus an interesting direction of research.

Finally, we have shown an important limitation in constructing pseudorandom unitaries. Our result on the depth requirement for PRUs is not surprising. Most candidate constructions [26, 27] consist of circuits with a polynomial depth. While a lot of work is actively being done on proving the security of these constructions, proving stronger bounds on the depth necessary to achieve pseudorandomness would be a crucial result. Our method of combining the building blocks of learning and verification algorithms indicates a potential technique for achieving such a result.

Acknowledgements: The authors acknowledge the support of the Quantum Advantage Pathfinder (QAP), with grant reference EP/X026167/1 and the UK Engineering and Physical Sciences Research Council. The authors thank anonymous reviewers of the conferences YQIS and QTML for helpful comments.

References

[1] Alessandro Bisio, Giulio Chiribella, Giacomo Mauro D'Ariano, Stefano Facchini, and Paolo Perinotti. Optimal quantum tomography. *IEEE Journal of Selected Topics in Quantum Electronics*, 15(6):1646–1660, 2009.

[2] Jeongwan Haah, Robin Kothari, Ryan O'Donnell, and Ewin Tang. Query-optimal estimation of unitary channels in diamond distance. In *2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 363–390. IEEE, 2023.

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

[4] Haimeng Zhao, Laura Lewis, Ishaan Kannan, Yihui Quek, Hsin-Yuan Huang, and Matthias C Caro. Learning quantum states and unitaries of bounded gate complexity. *arXiv preprint arXiv:2310.19882*, 2023.

[5] Hsin-Yuan Huang, Yunchao Liu, Michael Broughton, Isaac Kim, Anurag Anshu, Zeph Landau, and Jarrod R McClean. Learning shallow quantum circuits. *arXiv preprint arXiv:2401.10095*, 2024.

[6] Hsin-Yuan Huang, Sitan Chen, and John Preskill. Learning to predict arbitrary quantum processes. *PRX Quantum*, 4(4):040337, 2023.

[7] Sergey Bravyi, David Gosset, and Robert König. Quantum advantage with shallow circuits. *Science*, 362(6412):308–311, 2018.

[8] Sergey Bravyi, David Gosset, Robert König, and Marco Tomamichel. Quantum advantage with noisy shallow circuits. *Nature Physics*, 16(10):1040–1045, 2020.

[9] Jens Eisert, Dominik Hangleiter, Nathan Walk, Ingo Roth, Damian Markham, Rhea Parekh, Ulysse Chabaud, and Elham Kashefi. Quantum certification and benchmarking. *Nature Reviews Physics*, 2(7):382–390, 2020.

[10] Robin Harper, Steven T Flammia, and Joel J Wallman. Efficient learning of quantum noise. *Nature Physics*, 16(12):1184–1188, 2020.

[11] 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. *Reviews of Modern Physics*, 95(4):045005, 2023.

[12] Michael Kearns. Efficient noise-tolerant learning from statistical queries. *Journal of the ACM (JACM)*, 45(6):983–1006, 1998.

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

[14] Srinivasan Arunachalam, Vojtech Havlicek, and Louis Schatzki. On the role of entanglement and statistics in learning. *Advances in Neural Information Processing Systems*, 36, 2024.

[15] Alexander Nietner, Marios Ioannou, Ryan Sweke, Richard Kueng, Jens Eisert, Marcel Hinsche, and Jonas Haferkamp. On the average-case complexity of learning output distributions of quantum circuits. *arXiv preprint arXiv:2305.05765*, 2023.

[16] Alexander Nietner. Unifying (quantum) statistical and parametrized (quantum) algorithms. *arXiv preprint arXiv:2310.17716*, 2023.

[17] Armando Angrisani. Learning unitaries with quantum statistical queries. *arXiv preprint arXiv:2310.02254*, 2023.

[18] Chirag Wadhwa and Mina Doosti. Learning quantum processes with quantum statistical queries. *arXiv preprint arXiv:2310.02075*, 2023.

[19] Krzysztof Pietrzak. Cryptography from learning parity with noise. In *International Conference on Current Trends in Theory and Practice of Computer Science*, pages 99–114. Springer, 2012.

[20] Mohammed M Alani. Applications of machine learning in cryptography: a survey. In *Proceedings of the 3rd International Conference on cryptography, security and privacy*, pages 23–27, 2019.

[21] Sofiene Jerbi, Casper Gyurik, Simon C Marshall, Riccardo Molteni, and Vedran Dunjko. Shadows of quantum machine learning. *arXiv preprint arXiv:2306.00061*, 2023.

[22] 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.

[23] Prabhanjan Ananth, Luowen Qian, and Henry Yuen. Cryptography from pseudorandom quantum states. In *Annual International Cryptology Conference*, pages 208–236. Springer, 2022.

[24] William Kretschmer. Quantum pseudorandomness and classical complexity. *arXiv preprint arXiv:2103.09320*, 2021.

[25] William Kretschmer, Luowen Qian, Makrand Sinha, and Avishay Tal. Quantum cryptography in algorithmica. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, pages 1589–1602, 2023.

[26] Tony Metger, Alexander Poremba, Makrand Sinha, and Henry Yuen. Simple constructions of linear-depth t-designs and pseudorandom unitaries. *arXiv preprint arXiv:2404.12647*, 2024.

[27] Chi-Fang Chen, Adam Bouland, Fernando GSL Brandão, Jordan Docter, Patrick Hayden, and Michelle Xu. Efficient unitary designs and pseudorandom unitaries from permutations. *arXiv preprint arXiv:2404.16751*, 2024.

[28] Tobias Haug, Kishor Bharti, and Dax Enshan Koh. Pseudorandom unitaries are neither real nor sparse nor noise-robust. *arXiv preprint arXiv:2306.11677*, 2023.

[29] Sabee Grewal, Vishnu Iyer, William Kretschmer, and Daniel Liang. Low-stabilizer-complexity quantum states are not pseudorandom. *arXiv preprint arXiv:2209.14530*, 2022.

[30] Aravind Gollakota and Daniel Liang. On the hardness of pac-learning stabilizer states with noise. *Quantum*, 6:640, 2022.

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

[32] Shivam Nadimpalli, Natalie Parham, Francisca Vasconcelos, and Henry Yuen. On the pauli spectrum of qac0. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, pages 1498–1506, 2024.

[33] Vitaly Feldman. A general characterization of the statistical query complexity. In *Conference on learning theory*, pages 785–830. PMLR, 2017.

[34] Nicholas Hunter-Jones. Unitary designs from statistical mechanics in random quantum circuits. *arXiv preprint arXiv:1905.12053*, 2019.

[35] Boaz Barak, Chi-Ning Chou, and Xun Gao. Spoofing linear cross-entropy benchmarking in shallow quantum circuits. *arXiv preprint arXiv:2005.02421*, 2020.

[36] Riccardo Molteni, Casper Gyurik, and Vedran Dunjko. Exponential quantum advantages in learning quantum observables from classical data. *arXiv preprint arXiv:2405.02027*, 2024.

[37] 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.

[38] Thomas Schuster, Jonas Haferkamp, and Hsin-Yuan Huang. Random unitaries in extremely low depth. *arXiv preprint arXiv:2407.07754*, 2024.

[39] Michael A Nielsen and Isaac L Chuang. *Quantum computation and quantum information*. Cambridge university press, 2010.

[40] Antonio Anna Mele. Introduction to haar measure tools in quantum information: A beginner's tutorial. *Quantum*, 8:1340, 2024.

[41] Dax Enshan Koh and Sabee Grewal. Classical shadows with noise. *Quantum*, 6:776, 2022.

[42] Senrui Chen, Wenjun Yu, Pei Zeng, and Steven T Flammia. Robust shadow estimation. *PRX Quantum*, 2(3):030348, 2021.

[43] 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.

[44] Myrto Arapinis, Mahshid Delavar, Mina Doosti, and Elham Kashefi. Quantum physical unclonable functions: Possibilities and impossibilities. *Quantum*, 5:475, 2021.

A Estimating the noise with a single query

Proof of Theorem 2. Recall the statistical query in the proof sketch

$$\alpha \leftarrow \mathsf{2QPStat}_{\Lambda(\mathcal{U})}(|0\rangle\langle 0|^{\otimes 2}, \mathbb{F}, \tau) \tag{73}$$

and the purity of the output state

$$Tr((\rho^{out})^2) = 1 - (2\gamma - \gamma^2)(1 - \frac{1}{2^n}) \tag{74}$$

Denote by $f : [0,1] \to [0,1], f(\gamma) = 2\gamma - \gamma^2$. Note this function is strictly increasing for $\gamma \in [0,1)$. We have $f^{-1}(y) = 1 - \sqrt{1-y}$. Then, we have

$$\alpha \in \left[1 - f(\gamma)(1 - 2^{-n}) - \tau, 1 - f(\gamma)(1 - 2^{-n}) + \tau\right] \tag{75}$$

Equivalently,

$$f(\gamma) \in \left[\frac{1-\alpha-\tau}{1-2^{-n}}, \frac{1-\alpha+\tau}{1-2^{-n}}\right] \tag{76}$$

Therefore, we obtain a range for the depolarizing strength γ

$$\gamma \in \left[f^{-1}\left(\frac{1-\alpha-\tau}{1-2^{-n}}\right), f^{-1}\left(\frac{1-\alpha+\tau}{1-2^{-n}}\right)\right] \tag{77}$$

While we have an estimate of the noise strength, we need to estimate the diamond distance between the noiseless and noisy channels.

$$\begin{aligned}
\left\|\mathcal{U} - \Lambda(\mathcal{U})\right\|_\diamond &= \left\|\mathcal{U} - \Lambda(\gamma) \circ \mathcal{U}\right\|_\diamond \\
&= \left\|\mathcal{I} - \Lambda(\gamma)\right\|_\diamond \\
&= \max_\rho \left\|\mathcal{I} \otimes \mathcal{I}(\rho) - \Lambda(\gamma) \otimes \mathcal{I}(\rho)\right\|_1 \\
&= \max_\rho \left\|\rho - (1-\gamma)\rho - \gamma(\Phi \otimes \mathcal{I})(\rho)\right\|_1 \\
&= \gamma \max_\rho \left\|\rho - \Phi \otimes \mathcal{I}(\rho)\right\|_1 \\
&= \gamma \|\mathcal{I} - \Phi\|_\diamond
\end{aligned} \tag{78}$$

where the second equality follows from the unitary invariance of the diamond norm and Φ is the maximally depolarizing channel. From Lemma 3 of [18] , we know that for any unitary channel $\mathcal{U}$,

$$\|\mathcal{U}-\Phi\|_\diamond \geq 2-\frac{2}{2^n} \tag{79}$$

As this lower bound is close to the maximum possible for the diamond norm, we use the upper bound

$$\|\mathcal{I}-\Phi\|_\diamond \leq 2 \tag{80}$$

Combining equations 77, 78, 79 and 80, we obtain

$$\|\mathcal{U}-\Lambda(\mathcal{U})\|_\diamond \in \left[2(1-2^{-n})f^{-1}\left(\frac{1-\alpha-\tau}{1-2^{-n}}\right), 2f^{-1}\left(\frac{1-\alpha+\tau}{1-2^{-n}}\right)\right] \tag{81}$$

Denote

$$[l,u] = \left[2(1-2^{-n})f^{-1}\left(\frac{1-\alpha-\tau}{1-2^{-n}}\right), 2f^{-1}\left(\frac{1-\alpha+\tau}{1-2^{-n}}\right)\right] \tag{82}$$

Now, to show that this is a good estimate for the diamond norm between the noisy and noiseless unitaries, we show the choice of τ for which the difference between these bounds is small, i.e.

$$u-l \leq \epsilon \tag{83}$$

$$\begin{aligned} u-l &= 2\left(f^{-1}\left(\frac{1-\alpha+\tau}{1-2^{-n}}\right) - f^{-1}\left(\frac{1-\alpha-\tau}{1-2^{-n}}\right)\right) + \frac{2}{2^n}f^{-1}\left(\frac{1-\alpha-\tau}{1-2^{-n}}\right) \\ &\leq 2\left(\frac{\sqrt{\alpha-2^{-n}+\tau}-\sqrt{\alpha-2^{-n}-\tau}}{\sqrt{1-2^{-n}}}\right) + \frac{2}{2^n} \end{aligned} \tag{84}$$

Going forward, we will focus on the numerator of the term on the left.

$$\begin{aligned} \sqrt{\alpha-2^{-n}+\tau}-\sqrt{\alpha-2^{-n}-\tau} &= \frac{(\alpha-2^{-n}+\tau)-(\alpha-2^{-n}-\tau)}{\sqrt{\alpha-2^{-n}+\tau}+\sqrt{\alpha-2^{-n}-\tau}} \\ &\leq \frac{2\tau}{2\sqrt{\alpha-2^{-n}-\tau}} \\ &\leq \frac{\tau}{\sqrt{Tr(\rho)^2-2^{-n}-2\tau}} \\ &= \frac{\tau}{\sqrt{(1-\gamma)^2(1-2^{-n})-2\tau}} \end{aligned} \tag{85}$$

where the second inequality uses $\alpha \in [Tr(\rho^2)-\tau, Tr(\rho^2)+\tau]$. Suppose we start with some initial, loose upper bound on $\gamma \leq \gamma_u$. For instance, $\gamma_u = 0.5$. Then, choose

$$\tau = \frac{(1-\gamma_u)\epsilon}{4} \tag{86}$$

Denote by C

$$C = \frac{\epsilon}{2(1-\gamma_u)} \tag{87}$$

For $\epsilon \leq (1-\gamma_u)$, we have $C \leq 0.5$. Further,

$$\tau = \frac{C(1-\gamma_u)^2}{2} \leq \frac{C(1-\gamma)^2}{2} \tag{88}$$

Now, for this value of τ,

$$\begin{aligned}\sqrt{\alpha - 2^{-n} + \tau} - \sqrt{\alpha - 2^{-n} - \tau} &\leq \frac{C(1-\gamma_u)^2}{2(1-\gamma)\sqrt{1-2^{-n}-C}} \\ &\leq C(1-\gamma_u)\end{aligned} \tag{89}$$

where we use $1-\gamma_u \leq 1-\gamma$ and $\sqrt{1-2^{-n}-C} \geq 1/2, \forall n \geq 2, C \leq 1/2$ in the last inequality. Thus,

$$\begin{aligned}u - l &\leq 2\frac{C(1-\gamma_u)}{\sqrt{1-2^{-n}}} - \frac{2}{2^n} \\ &= \frac{\epsilon}{\sqrt{1-2^{-n}}} - \frac{2}{2^n} \\ &= \frac{\epsilon(\sqrt{1+2^{-n}})}{1-2^{-n}} - \frac{2}{2^n} \\ &\leq \frac{\epsilon(1+2^{-n-1})}{1-2^{-n}} - \frac{2}{2^n} \\ &= \frac{\epsilon(2^n+1/2) - 2(1-2^{-n})}{2^n - 1}\end{aligned} \tag{90}$$

where the second to last line uses $\sqrt{1+x} \leq 1 + x/2 \ \forall x \geq 0$ .For all $n \geq 2, \epsilon \leq 1$, the above simplifies to

$$u - l \leq \epsilon \tag{91}$$

□

B Bounding variance for $\mathsf{RQC}(n,d)$

Proof of Lemma 6. Recall that our goal is to show a bound on the variance of $Tr(OU\rho U^\dagger)$ where U is a brickwork random quantum circuit. First, we must show that in the depth range of Lemma 6,

$$\mathop{\mathbf{E}}_{U\sim\mathsf{RQC}(n,d)}[Tr(OU\rho U^\dagger)^2] \leq \left(1 + \left(\frac{4}{5}\right)^d\right)^{n/2} \mathop{\mathbf{E}}_{U\sim\mu_H}[Tr(OU\rho U^\dagger)^2] \tag{92}$$

To do this, we will look at the technique used in [15, 34, 35]. The proof technique starts by representing the second-order moment in the form of a tensor network diagram, mapping the diagram to a statistical mechanics model, and then counting domain walls over this model. While we refer to [34, 35] for the diagrams and the detailed mapping, we will follow along the argument of [35], where they find a similar upper bound for $\sum_{x\in\{0,1\}^n}(\langle 0|U|x\rangle\langle x|U^\dagger|0\rangle)^2$. Using their arguments, we obtain

$$\mathop{\mathbf{E}}_{U\sim\mathsf{RQC}(n,d)}[Tr(OU\rho U^\dagger)^2] = f_1(\rho)f_2(n,d)f_3(O) \tag{93}$$

where f_1 and f_3 are functions that map states and observables to reals and f_2 is a real-valued function of the lattice constructed by averaging over $\mathsf{RQC(n,d)}$. Intuitively, this expression is obtained by contracting the first layer of the tensor network to get f_1, the last layer to get f_3, and then averaging over all random circuits to get f_2. However, as we will show later, there is no need to explicitly compute f_1 and f_3, as the final bound is independent of these quantities. Instead, we will focus on f_2. The key result of [35] that we use is the relation between $f_2(n,d)$ and $f_2(n,d^*)$ as $d^* \to \infty$

$$f_2(n,d) \leq \left(1 + \left(\frac{4}{5}\right)^d\right)^{n/2} f_2(n,d^*) \tag{94}$$

Thus, we obtain,

$$\begin{aligned}\underset{U\sim\mathsf{RQC}(n,d)}{\mathbf{E}}[Tr(OU\rho U^\dagger)^2] &\leq \left(1 + \left(\frac{4}{5}\right)^d\right)^{n/2} f_1(\rho)f_2(n,d^*)f_3(O) \\ &= \left(1 + \left(\frac{4}{5}\right)^d\right)^{n/2} \underset{U\sim\mathsf{RQC}(n,d^*)}{\mathbf{E}}[Tr(OU\rho U^\dagger)^2]\end{aligned} \tag{95}$$

We obtain equation 92 by observing that BRQCs converge to the Haar measure at infinite depth. Now, we use Lemma 5 and the fact that BRQCs form exact 1-designs to compute the variance. We will denote $N = 2^n$.

$$\begin{aligned}\underset{U\sim\mathsf{RQC}(n,d)}{\mathbf{Var}}[Tr(OU\rho U^\dagger)] &= \underset{U\sim\mathsf{RQC}(n,d)}{\mathbf{E}}[Tr(OU\rho U^\dagger)^2] - \left(\underset{U\sim\mathsf{RQC}(n,d)}{\mathbf{E}}[Tr(OU\rho U^\dagger)]\right)^2 \\ &\leq \left(1 + \left(\frac{4}{5}\right)^d\right)^{n/2} \underset{U\sim\mu_H}{\mathbf{E}}[Tr(OU\rho U^\dagger)^2] - \left(\underset{U\sim\mu_H}{\mathbf{E}}[Tr(OU\rho U^\dagger)]\right)^2\end{aligned} \tag{96}$$

Note that

$$\left(1 + \left(\frac{4}{5}\right)^d\right)^{n/2} \leq \exp\left(\frac{n\,(4/5)^d}{2}\right) \tag{97}$$

For $d \geq \frac{\log n}{\log 5/4}$, we have $\frac{n(4/5)^d}{2} \leq 1/2$. We then use the fact that $e^x \leq 1 + 2x$,$\forall\ 0 \leq x \leq 1$ to show

$$\left(1 + \left(\frac{4}{5}\right)^d\right)^{n/2} \leq \left(1 + n\left(\frac{4}{5}\right)^d\right) \tag{98}$$

Now, we can bound the variance as:

$$
\begin{aligned}
\underset{U\sim\mathsf{RQC}(n,d)}{\mathbf{Var}}[Tr(OU\rho U^\dagger)] &\le \left(1+n\left(\frac{4}{5}\right)^d\right)\underset{U\sim\mu_H}{\mathbf{E}}[Tr(OU\rho U^\dagger)^2] - \left(\underset{U\sim\mu_H}{\mathbf{E}}[Tr(OU\rho U^\dagger)]\right)^2 \\
&= \underset{U\sim\mu_H}{\mathbf{Var}}[Tr(OU\rho U^\dagger)] + n\left(\frac{4}{5}\right)^d \underset{U\sim\mu_H}{\mathbf{E}}[Tr(OU\rho U^\dagger)^2] \\
&\le \frac{1}{N+1} + n\left(\frac{4}{5}\right)^d \left(\left(\frac{N-Tr(\rho^2)}{N(N^2-1)}\right)Tr(O)^2 + \left(\frac{NTr(\rho^2)-1}{N(N^2-1)}\right)Tr(O^2)\right) \\
&\le \frac{1}{N+1} + n\left(\frac{4}{5}\right)^d \left(\frac{Tr(O)^2}{N^2} + \frac{Tr(O^2)}{N(N+1)}\right) \\
&\le \frac{1}{N+1} + n\left(\frac{4}{5}\right)^d \left(1+\frac{1}{N+1}\right) \\
&= n\left(\frac{4}{5}\right)^d + \frac{1}{2^n+1}\left(1+n\left(\frac{4}{5}\right)^d\right)
\end{aligned} \tag{99}
$$

where the second inequality follows from substitution from Lemma 5, the third inequality follows from the fact that $1/N \le Tr(\rho^2) \le 1$. To obtain the fourth inequality, observe that for $\|O\|_\infty \le 1, |Tr(O)| \le N, Tr(O^2) \le N$. The final equation is obtained by rearranging the terms and substituting $N = 2^n$. Now, to obtain the desired bound on the variance, we need

$$
n\left(\frac{4}{5}\right)^d = \Omega(2^{-n}) \tag{100}
$$

which is satisfied when

$$
d \le \frac{n+\log_2(n)}{\log_2(5/4)} \tag{101}
$$

Thus, the desired bound on the variance is obtained for the outlined depth range, concluding the proof. □

C Average-case hardness for Haar-random unitaries

Our proof will build upon the technique of [44], where the authors proved a bound on the ability of an adversary to predict, with high fidelity, the output of a Haar-random unitary on a Haar-random state. We will also use the following lemma on the average infidelity of any CPTP map with Haar-random unitaries on average.

Lemma 9 (Quantum no-free lunch theorem with no samples)**.** *Any CPTP map $\mathcal{E} : \mathcal{S}_N \to \mathcal{S}_N$ has high d_{avg} with Haar-random unitaries on average.*

$$
\underset{U\sim\mu_H}{\mathbf{E}} d_{avg}(\mathcal{E}, U(.)U^\dagger) = 1 - \frac{1}{N} \tag{102}
$$

Proof.

$$\begin{aligned}
\mathop{\mathbf{E}}_{U\sim\mu_H} d_{avg}(\mathcal{E}, U(.)U^\dagger) &= 1 - \mathop{\mathbf{E}}_{U\sim\mu_H}\left[\mathop{\mathbf{E}}_{|\psi\rangle\sim\mu_S} F\left(\mathcal{E}(|\psi\rangle\langle\psi|), U|\psi\rangle\langle\psi|U^\dagger\right)\right] \\
&= 1 - \mathop{\mathbf{E}}_{|\psi\rangle\sim\mu_S}\left[\mathop{\mathbf{E}}_{U\sim\mu_H} F\left(\mathcal{E}(|\psi\rangle\langle\psi|), U|\psi\rangle\langle\psi|U^\dagger\right)\right] \\
&= 1 - \mathop{\mathbf{E}}_{|\psi\rangle\sim\mu_S}\left[\mathop{\mathbf{E}}_{U\sim\mu_H} Tr(\mathcal{E}(|\psi\rangle\langle\psi|)U|\psi\rangle\langle\psi|U^\dagger)\right] \\
&= 1 - \mathop{\mathbf{E}}_{|\psi\rangle\sim\mu_S}\left[Tr\left(\mathcal{E}(|\psi\rangle\langle\psi|)\mathop{\mathbf{E}}_{U\sim\mu_H}[U|\psi\rangle\langle\psi|U^\dagger]\right)\right] \\
&= 1 - \mathop{\mathbf{E}}_{|\psi\rangle\sim\mu_S}\left[Tr\left(\mathcal{E}(|\psi\rangle\langle\psi|)\frac{I}{N}\right)\right] \\
&= 1 - \mathop{\mathbf{E}}_{|\psi\rangle\sim\mu_S}\left[\frac{1}{N}\right] \\
&= 1 - \frac{1}{N}
\end{aligned} \tag{103}$$

where the third equality uses the property of the fidelity when one of the states is pure, the fourth equality uses the linearity of expectation and trace, the fifth equality follows from Lemma 5, and the second to last equality follows from the fact that $\mathcal{E}$ is trace-preserving. □

The lemma can be interpreted as follows: any learning algorithm for a Haar-random unitary that makes *no queries*, can only guess either a fixed channel or a channel from some fixed distribution. Lemma 9 then tells us that such an algorithm will have high error on average over all Haar-random unitaries. This interpretation is similar to the original QNFLT from [31]. In fact, by setting the number of queries to 0 in the original theorem, we obtain the same bound as the original QNFLT. Now, we will prove Theorem 5

Proof of Theorem 5. Suppose the learner makes queries $\{|\psi^i_{in}\rangle\}_{i=1}^q$ and receives states $\{|\psi^i_{out}\rangle\}_{i=1}^q$, with $|\psi^i_{out}\rangle = U|\psi_{in}\rangle$. We strengthen the access of the learner by assuming the learner also has access to the classical descriptions of the states. We will show the lower bound for this strengthened learner, which will then hold for the original setting as well.

Given the classical description, the learner has full knowledge of the action of U over the subspace spanned by the input states. Denote the Hilbert space over n-qubit states as $\mathcal{H}^N$, where $N = 2^n$. Denote the space of states with non-zero overlap with the input states as $\mathcal{H}^q$, and the space of states orthogonal to all input states as $\mathcal{H}^{q^\perp}$. Denote by μ'_H (μ'_S) the Haar measure over unitaries (states) on the Hilbert space $\mathcal{H}^{q^\perp}$. Similar to [44], we strengthen the learner beyond the assumption of [31], by giving it perfect fidelity on every state in $\mathcal{H}^q$. Thus, as long as the overlap of a state with any $|\psi^i_{in}\rangle$ is non-zero, even if it's arbitrarily small, we assume the learner succeeds perfectly. Thus, the learner has no error on states in $\mathcal{H}^q$. On the other hand, for states in $\mathcal{H}^{q^\perp}$, the action of the unitary is completely random. Denote by d_μ the average infidelity between two channels over states sampled from some measure μ. In particular, when $\mu = \mu_S$, d_μ is d_{avg}. Now, for a learner in the original setting, with learned channel $\mathcal{E}$, we bound the expected error as

$$\mathop{\mathbf{E}}_{U\sim\mu_H}[d_{avg}(U,\mathcal{E})] \geq \mathop{\mathbf{Pr}}_{|\psi\rangle\sim\mu_S}\left[|\psi\rangle \notin \mathcal{H}^{q^\perp}\right]0 + \mathop{\mathbf{Pr}}_{|\psi\rangle\sim\mu_S}\left[|\psi\rangle \in \mathcal{H}^{q^\perp}\right]\mathop{\mathbf{E}}_{U'\sim\mu'_H}\left[d_{\mu'_S}(U',\mathcal{E}|_{\mathcal{H}^{q^\perp}})\right] \tag{104}$$

where $\mathcal{E}|_{\mathcal{H}^{q^\perp}}$ denotes $\mathcal{E}$ restricted to input states from $\mathcal{H}^{q^\perp}$. The expected error over μ'_H is now precisely given by Lemma 9 for dimension $N-q$. Next, we compute the probabilities in the equation. As shown in [44], the probability of $|\psi\rangle$ belonging to the subspace $\mathcal{H}^{q^\perp}$ is given by the ratio of the dimensions, i.e.

$$\mathop{\mathbf{Pr}}_{|\psi\rangle\sim\mu_S}\left[|\psi\rangle \in \mathcal{H}^{q^\perp}\right] = \frac{N-q}{N} \tag{105}$$

Thus, we obtain

$$\mathop{\mathbf{E}}_{U\sim\mu_H}[d_{avg}(U,\mathcal{E})] \geq \frac{N-q}{N}\left(1-\frac{1}{N-q}\right) = 1-\frac{q+1}{N} \tag{106}$$

Now, we will find a trivial upper bound on the average error.

$$\begin{aligned}\mathop{\mathbf{E}}_{U\sim\mu_H}[d_{avg}(U,\mathcal{E})] &\leq (\epsilon)\mathop{\mathbf{Pr}}_{U\sim\mu_H}[d_{avg}(U,\mathcal{E})\leq\epsilon] + (1)\mathop{\mathbf{Pr}}_{U\sim\mu_H}[d_{avg}(U,\mathcal{E})>\epsilon] \\ &= 1-(1-\epsilon)\left(\mathop{\mathbf{Pr}}_{U\sim\mu_H}[d_{avg}(U,\mathcal{E})\leq\epsilon]\right)\end{aligned} \tag{107}$$

where the first inequality uses the fact that the average error is upper bound by 1. Now, for success probability at least $1-\delta$, we have the upper bound

$$\mathop{\mathbf{E}}_{U\sim\mu_H}[d_{avg}(U,\mathcal{E})] \leq 1-(1-\epsilon)(1-\delta) \tag{108}$$

Combining the upper bound (108) and lower bound (106), we obtain

$$q \geq N(1-\epsilon)(1-\delta)-1 \tag{109}$$

This concludes the proof. □