Learning Quantum Processes with Quantum Statistical Queries

Chirag Wadhwa and Mina Doosti

School of Informatics, University of Edinburgh, Edinburgh, United Kingdom

In this work, we initiate the study of learning quantum processes from quantum statistical queries. We focus on two fundamental learning tasks in this new access model: shadow tomography of quantum processes and process tomography with respect to diamond distance. For the former, we present an efficient average-case algorithm along with a nearly matching lower bound with respect to the number of observables to be predicted. For the latter, we present average-case query complexity lower bounds for learning classes of unitaries. We obtain an exponential lower bound for learning unitary 2-designs and a doubly exponential lower bound for Haar-random unitaries. Finally, we demonstrate the practical relevance of our access model by applying our learning algorithm to attack an authentication protocol using Classical-Readout Quantum Physically Unclonable Functions, partially addressing an important open question in quantum hardware security.

1 Introduction

Quantum learning theory aims to study the advantages and limitations of quantum machine learning in both classical and quantum problems. Here, the *learner* is a classical, quantum or hybrid algorithm, trying to learn about an unknown object (function, distribution, quantum state, quantum process, etc.) through some limited access to it. Widely studied classical access models such as random examples [1] and statistical queries [2] have been extended to the quantum setting in the form of quantum examples [3, 4] and quantum statistical queries [5–9] respectively. A wide array of results have been shown in this field, ranging from learning functions encoded within quantum states [3, 10, 11], quantum state tomography [12–14], shadow tomography [15, 16], learning diverse classes of probability distributions [10, 17, 18], to learning quantum processes [19–21]. We refer to the following survey for an overview of results [22]. While most of the efforts in quantum learning theory have been focused on quantum states (both as examples and target objects), in this work, we study a new access model for learning about quantum processes.

Learning quantum processes is a fundamental problem that arises in many areas in physics [23, 24] and quantum computing, such as quantum benchmarking [25–28], noise characterisation [29], error mitigation [30, 31], and variational quantum algorithms [32]. Furthermore, with the crucial role of quantum computing in cryptography, another such area is cryptanalysis. In these scenarios, the quantum process of interest can manifest as a quantum oracle, providing a quantum implementation of a classical function [11, 33–36], or as a physical device or hardware component implementing an unknown underlying unitary, which serves as a cryptographic key or fingerprint [37–39]. The primary challenge in learning complex quantum processes lies in the resource-intensive nature of this task, rendering conventional techniques for process tomography [19] impractical, especially for near-term devices. Recent endeavours have explored diverse techniques to devise algorithms and approaches for efficiently tackling specific instances of this challenge [20, 25–27, 40–42].

In this work, we focus on the the *statistical query* access model [2], and naturally extend it to the task of learning quantum processes. In a statistical query access model (quantum or classical)

Chirag Wadhwa: chirag.wadhwa@ed.ac.uk
Mina Doosti: mdoosti@ed.ac.uk

the learner constructs a hypothesis not by accessing a sequence of labelled examples themselves, but instead by adaptively querying an oracle that provides an *estimate* of the statistical properties of the labelled examples. In the quantum case, this estimate is in fact the estimated expectation values of some observable. This extension to the quantum world is quite natural, as this is often the information extracted from a quantum system, by measuring it many times and estimating the expectation value of an observable, which corresponds to a physical quantity. This natural correspondence to the physics of the quantum experiment and the learning tasks designed based on them marks our main motivation for the choice of this model. In our access model, information about a process is accessed by querying it with an input state and observable, and receiving an estimate of the output expectation value. The feasibility of this model in practice, as compared to quantum PAC learning, makes it a good candidate for studying learning algorithms and their limitations in the near-term. Aside from being physically well-motivated, quantum statistical queries have also found applications in the classical verification of quantum learning [43] and quantum error mitigation [6, 31].

There are several noteworthy points regarding our proposed access model. In comparison with conventional quantum process tomography [19], our access model is significantly weaker. In particular, we restrict the learner to measurement statistics on a single copy of the output state, and the learner cannot obtain the results of any entangled measurements with ancillary qubits. In many practical scenarios, due to the limitations of quantum memories, only the statistical results of quantum measurements are relevant and accessible in the near term. This aspect is precisely the focus of our model. Furthermore, our model gains importance when considering scenarios in which direct access to the process is not provided. This becomes especially crucial for cryptanalysis purposes, where certain attack models may not grant direct access to the process, but statistical data can still be accessed by adversaries.

In this access model, we study two fundamental problems in quantum learning theory. First, we look at the task of predicting properties of quantum processes, where given quantum states from a certain distribution and a list of observables, one aims to predict output expectation values after evolution under an unknown quantum process. For this task, we present an efficient learning algorithm and provide a matching lower bound (up to a logarithmic factor) in terms of the number of observables to be predicted. Next, we look at the task of learning unitaries with respect to the diamond distance, for which we provide average-case query complexity lower bounds. Furthermore, we explore the practical applications of our learning algorithm in this access model, by applying it to attack a cryptographic protocol. This result sheds light on identifying the conditions under which a class of quantum physical unclonable functions may be vulnerable, partially addressing an ongoing open question in quantum hardware security.

1.1 Our Contributions

In this section, we summarize our main contributions.

Quantum Statistical Queries to Quantum Processes (QPSQs): We define our access model through the quantum statistical query oracle QPStat for a quantum process $\mathcal{E}$, that produces an estimate of the expectation value $\mathrm{tr}\big(O\mathcal{E}(\rho)\big)$. We discuss how our oracle generalises the previously defined QStat oracles for states. We show that in many situations, efficient quantum statistical query algorithms for learning boolean functions admit equivalent, efficient algorithms in our model. We present our access model in Section 3.

Predicting properties of quantum processes: We present an efficient algorithm that can predict properties of quantum processes from QPSQs. Our algorithm is an adaptation of the classical shadow [15] algorithm of [40] to our access model. In our weaker access model, we obtain a similar query complexity as [40] with a linear overhead in the number of observables to be predicted. We show that this overhead is unavoidable, by presenting a nearly matching lower bound with respect to the number of observables. Our lower bound builds upon the technique used by [6] to show a similar bound for shadow tomography of quantum states from quantum statistical queries. We further demonstrate the performance of our proposed algorithm through numerical simulations. We present these results in Section 4.

Hardness of diamond-distance learning: We provide an exponential lower bound on the query complexity for learning exact and approximate unitary 2-designs from QPStat queries with respect to diamond distance. We also show a doubly exponential lower bound on the hardness of learning unitaries over the Haar-measure from QPStat queries with respect to this distance. We start by proving a general lower bound for learning any class of unitaries from QPSQs, which we obtain through a reduction from a many-vs-one distinguishing task. Our techniques are inspired by standard techniques for proving statistical query lower bounds [44], which have been widely used in the quantum statistical query setting as well [5, 8, 9]. We present our lower bounds for diamond-distance learning in Section 5.

Application to cryptanalysis: We apply our results to cryptanalysis by studying a primitive from quantum hardware security, namely *Classical Readout of Quantum Physically Unclonable Functions* (CR-QPUFs) [38, 39]. The security of this primitive relies on the assumed hardness of predicting statistical properties of an underlying quantum process. However, the existence of a secure realization of CR-QPUFs has not yet been shown, and remains an important open question. We partially address this problem by applying our learning algorithm for an attack against authentication protocols based on CR-QPUFs under practical physical assumptions. Our attack inherits the quasipolynomial complexity of our learning algorithm, preventing us from formally breaking the security of CR-QPUFs. However, our results present an interesting connection between learning theory and cryptography, as any new polynomial-time QPSQ algorithm for predicting expectation values would imply an attack against such protocols.

1.2 Related work

Quantum Statistical Queries: Quantum statistical queries (QSQs) were introduced by [5] as a quantum generalization of the statistical query (SQ) access model [2]. [5] also showed efficient QSQ learning algorithms for various classes of Boolean functions which are provably hard to learn from SQs. QSQs for learning quantum states have been considered in [6, 9] and those for learning output distributions of quantum circuits have been considered in [7, 8]. In concurrent work, [45] consider learning unitaries from QSQs to their Choi states, and present learning algorithms for quantum Boolean functions and quantum k-juntas.

Statistical Query Lower Bounds: Recent works in QSQ learning [6, 8, 9] have obtained lower bounds by adapting the classical SQ lower bound techniques of [44], involving a reduction from a many-vs-one distinguishing task. In [8], the authors showed lower bounds for learning the Born distributions of random circuits at varying depth regimes. [9] provided a unifying framework for a large class of learning models (including many statistical query oracles), presented lower bounds within this framework, and applied them to the task of learning output states of quantum circuits. Compared to learning a process within diamond distance, the tasks of learning output distributions and output states are easier. However, the access model we consider is also stronger than their statistical query models. As such, our bounds cannot be directly compared to those of [8, 9].

In the setting where learning algorithms are restricted to single-copy measurements, [46] showed an exponential sample complexity lower bound for distinguishing between Haar-random unitaries and the depolarizing channel. As QPSQ access is weaker than single-copy measurements, this immediately implies an exponential lower bound in our setting. We improve upon this by showing a doubly exponential lower bound in Theorem 4.

Predicting Properties of Quantum Processes: [26, 40, 47, 48] have studied shadow tomography of quantum processes, i.e. predicting expectation values of observables for quantum states after evolution under an unknown quantum process. [26, 47, 48] focus on this task in the worst-case over input states, i.e., they aim to make predictions that are accurate for every input quantum state. To this end, [26, 47] use classical shadows [15] to make predictions. [48] learn elements of the Pauli transfer matrix of a channel, allowing them to efficiently make predictions for states and observables with a sparse Pauli spectrum decomposition. On the other hand, similar to our work, [40] focus on an average-case version of this task, only requiring the error to be low on average over states sampled randomly. [40] also make use of classical shadows [15] for this task.

Concurrent work: In independent and concurrent work, [49] considered the similar access model of *measurement queries*. Among other results, they presented an efficient algorithm in this access model for learning a class of quantum channels formed by QAC^0 circuits.

1.3 Organization of the paper

We provide the necessary background and notation in Section 2. In Section 3, we define our access model and discuss its relation to other models. In Section 4, we provide our upper and lower bound for predicting properties of an unknown quantum process. In Section 5, we present average-case lower bounds for learning a unitary with respect to diamond distance. Finally, we discuss applications of our algorithm for cryptanalysis in in Section 6.

2 Preliminaries

We start by introducing the notation we use in the paper as well as the essential background.

2.1 Quantum Information

We include some basic definitions of quantum computation and information in this section. For more details, we refer the reader to [50]. We will denote the $d \times d$ identity matrix as I_d and we may omit the index d when the dimension is clear from the context. We use the bra-ket notation, where we denote a vector $v \in \mathbb{C}^N$ using the ket notation $|v\rangle$ and its adjoint using the bra notation $\langle v|$. For $u, v \in \mathbb{C}^n$, we will denote by $\langle u|v\rangle$ the standard Hermitian inner product $u^\dagger v$. A quantum (pure) state is a normalized vector $|v\rangle$, i.e. $|\langle v|v\rangle| = 1$. 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 $\mathcal{S}_N := \{\rho \in \mathcal{M}_{N,N} : \rho \succeq 0, \mathrm{Tr}[\rho] = 1\}$. We denote by $\mathcal{U}_N$ the set of N-dimensional unitary operators,

$$\mathcal{U}_N := \left\{U \in \mathcal{M}_{N,N} : UU^\dagger = U^\dagger U = I\right\}. \tag{1}$$

We will now introduce a useful orthonormal basis for $\mathcal{M}_{N,N}$ which is widely used in quantum information.

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

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2}$$

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. The tensor products of Pauli operators and the identity, i.e. the operators of the form $P \in \{I, X, Y, Z\}^{\otimes n} := \mathcal{P}_n$, are usually referred as *stabilizer operators* or *Pauli strings* over n qubits. Pauli strings over n qubits form an orthonormal basis for $\mathcal{M}_{2^n,2^n}$ with respect to the Hilbert-Schmidt inner product. For a Pauli string $P \in \mathcal{P}_n$, we define its degree, $|P|$, as the number of indices on which it acts non-trivially, i.e., the number of non-identity Paulis. We now look at the eigenstates of the Pauli operators, which are of special interest.

Definition 2 (Single-qubit stabilizer states)**.** *We denote the set of Pauli eigenstates by*

$$\mathrm{stab}_1 = \{|0\rangle, |1\rangle, |+\rangle, |-\rangle, |+y\rangle, |-y\rangle\}, \tag{3}$$

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 (Clifford group)**.** *The Clifford group is the group of unitaries generated by the following 3 gates :*

$$H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, S = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}, CNOT = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{4}$$

We now define quantum processes.

Definition 4 (Quantum process)**.** *A map $\mathcal{E} : \mathcal{M}_{N,N} \to \mathcal{M}_{N,N}$ 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 $\rho, \mathrm{tr}\big(\mathcal{E}(\rho)\big) = \mathrm{tr}(\rho) = 1$. A quantum process $\mathcal{E}$ is defined as a Completely Positive Trace-Preserving (CPTP) map from one quantum state to another. For a unitary U, the associated map is:*

$$\mathcal{E} : \rho \to U\rho U^\dagger. \tag{5}$$

Next, we define the maximally depolarizing channel, which will be particularly useful.

Definition 5 (Maximally depolarizing channel)**.** *The maximally depolarizing channel Φ_{dep} acting on states in S_N is defined as follows:*

$$\Phi_{\mathrm{dep}}(\rho) = \frac{\mathrm{tr}(\rho)}{N} I. \tag{6}$$

We now define some distances between quantum states and channels.

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

$$\mathrm{d}_{\mathrm{tr}}(\rho, \sigma) \triangleq \frac{1}{2}\|\rho - \sigma\|_1 , \tag{7}$$

where $\|.\|_1$ is the Schatten-1 norm. The fidelity *between two quantum states is given by*

$$F(\rho, \sigma) \triangleq \mathrm{tr}\left(\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\right)^2 . \tag{8}$$

When at least one of the states is pure, d_{tr} and F are related as follows:

$$1 - F(|\psi\rangle\langle\psi| , \rho) \leq \mathrm{d}_{\mathrm{tr}}(|\psi\rangle\langle\psi| , \rho). \tag{9}$$

Definition 7 (Diamond norm and diamond distance)**.** *For a map $\mathcal{E} \in \mathcal{M}_{N,N} \to \mathcal{M}_{N,N}$, and $\mathcal{I} \in \mathcal{M}_{N,N} \to \mathcal{M}_{N,N}$ the identity superoperator, we define the diamond norm $\| \ . \ \|_\diamond$*

$$\|\mathcal{E}\|_\diamond = \max_{\rho \in \mathcal{S}_{N^2}} \big\|(\mathcal{E} \otimes \mathcal{I})(\rho)\big\|_1 , \tag{10}$$

where $\|.\|_1$ denotes the Schatten 1-norm. We then define the diamond distance $d_\diamond$:

$$d_\diamond(\mathcal{E}_1, \mathcal{E}_2) = \frac{1}{2}\|\mathcal{E}_1 - \mathcal{E}_2\|_\diamond . \tag{11}$$

Definition 8 (POVM)**.** *A* Positive Operator-Valued Measure *(POVM) is a quantum measurement described by a collection of positive operators $\{E_m\}_m$, such that $\sum_m E_m = I$ and the probability of obtaining measurement outcome m on a state ρ is given by $p(m) = \mathrm{tr}(E_m\rho)$.*

2.2 Unitary Haar Measure and Designs

We now define the unitary Haar measure μ_H, which can be thought as the uniform probability distribution over all quantum states or over all unitary operators in the Hilbert space of dimension N. For a comprehensive introduction to the Haar measure and its properties, we refer to [51].

Definition 9 (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 the set $\mathcal{U}_N$, i.e., for all integrable functions f and for all $V \in \mathcal{U}_N$, we have:*

$$\int_{U(N)} f(U)d\mu_H(U) = \int_{U(N)} f(UV)d\mu_H(U) = \int_{U(N)} f(VU)d\mu_H(U). \tag{12}$$

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

$$\mathbf{E}_{|\psi\rangle\sim\mu_S}\left[|\psi\rangle\langle\psi|^{\otimes k}\right] := \mathbf{E}_{U\sim\mu_H}\left[U^{\otimes k}|\phi\rangle\langle\phi|^{\otimes k}U^{\dagger\otimes k}\right]. \tag{13}$$

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 $|\phi\rangle$.

A unitary t-design is a measure over unitaries, over which the t-th order moments match those of the Haar-measure exactly. In case the moments are only approximately equal, we call this an approximate unitary t-design. While there are many notions of approximation, here we restrict our attention to additive approximations as defined in [9].

Definition 10 (Exact and Approximate Unitary t-Designs)**.** *The t-th moment superoperator with respect to a distribution ν over $\mathcal{U}_N$ is defined as*

$$\mathcal{M}_\nu^{(t)}(A) = \mathop{\mathbf{E}}_{U\sim\nu}[U^{\otimes t}A(U^\dagger)^{\otimes t}] = \int U^{\otimes t}A(U^\dagger)^{\otimes t}d\nu(U). \tag{14}$$

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

$$\mathcal{M}_\nu^{(t)}(A) = \mathcal{M}_{\mu_H}^{(t)}(A). \tag{15}$$

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

$$\left\|\mathcal{M}_\nu^{(t)}(A) - \mathcal{M}_{\mu_H}^{(t)}(A)\right\|_\diamond \leq \delta. \tag{16}$$

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

We will be particularly interested in the first and second-order moments over the unitary Haar measure. We start by defining the identity and flip permutation operators which will be useful for this purpose.

Definition 11 (Identity and Flip permutation operators, Definition 12 from [51])**.** *The identity operator $\mathbb{I}$ and the flip operator $\mathbb{F}$ act on pure states $|\psi\rangle, |\phi\rangle \in \mathbb{C}^N$ as follows:*

$$\mathbb{I}(|\psi\rangle\otimes|\phi\rangle) = |\psi\rangle\otimes|\phi\rangle. \tag{17}$$

$$\mathbb{F}(|\psi\rangle\otimes|\phi\rangle) = |\phi\rangle\otimes|\psi\rangle. \tag{18}$$

We will use the following property of the flip operator. For all operators $A, B \in \mathcal{M}_{N,N}$:

$$\mathrm{tr}\big(\mathbb{F}(A\otimes B)\big) = \mathrm{tr}(AB). \tag{19}$$

Lemma 1 (First and second-order moments over the Haar-measure, Corollary 13 from [51])**.** *We have, for $O \in \mathcal{M}_{N,N}$,*

$$\mathcal{M}_{\mu_H}^{(1)}(O) = \frac{\mathrm{tr}(O)}{d}I. \tag{20}$$

For $O \in \mathcal{M}_{N^2,N^2}$,

$$\mathcal{M}_{\mu_H}^{(2)}(O) = c_{\mathbb{I},O}\mathbb{I} + c_{\mathbb{F},O}\mathbb{F}, \tag{21}$$

where,

$$c_{\mathbb{I},O} = \frac{\mathrm{tr}(O) - N^{-1}\,\mathrm{tr}(\mathbb{F}O)}{N^2-1} \text{ and } c_{\mathbb{F},O} = \frac{\mathrm{tr}(\mathbb{F}O) - N^{-1}\,\mathrm{tr}(O)}{N^2-1}. \tag{22}$$

We draw attention to the fact that $\mathcal{M}_{\mu_H}^{(1)}$ is the maximally depolarizing channel Φ_{dep}.

2.3 Classical Shadow Tomography

Classical shadow tomography is the technique of using randomized measurements to learn many properties of quantum states [15, 52, 53]. It has recently been shown that classical shadow tomography can be used to predict the outcomes of arbitrary quantum processes [40]. In this section, we include relevant results on classical shadow tomography that will be used in the rest of the work.

Definition 12 (Randomized Pauli Measurement)**.** *Given $n > 0$. A randomized Pauli measurement on an n-qubit state is given by a 6^n-outcome POVM*

$$\mathcal{F}^{Pauli} \triangleq \left\{ \frac{1}{3^n} \bigotimes_{i=1}^{n} |s_i\rangle\langle s_i| \right\}_{s_1,\ldots,s_n \in \mathrm{stab}_1}, \tag{23}$$

which corresponds to measuring every qubit under a random Pauli basis (X, Y, Z). The outcome of $\mathcal{F}^{Pauli}$ is an n-qubit state $|\psi\rangle = \bigotimes_{i=1}^{n} |s_i\rangle$, where $|s_i\rangle \in \mathrm{stab}_1$ is a single-qubit stabilizer state.

Next, we define classical shadows based on randomized Pauli measurements. Other measurements can also be used to define classical shadows.

Definition 13 (Classical shadow of a quantum state [15])**.** *Given $n, N > 0$. Consider an n-qubit state ρ. A size N classical shadow of $S_N(\rho)$ of quantum state ρ is a random set given by*

$$S_N(\rho) \triangleq \left\{ |\psi_l\rangle \right\}_{l=1}^{N}, \tag{24}$$

where $|\psi_l\rangle = \bigotimes_{i=1}^{n} |s_{l,i}\rangle$ is the outcome of the l-th randomized Pauli measurement on a single copy of ρ.

Definition 14 (Classical Shadow Approximation of a quantum state [15])**.** *Given the classical shadow $S_N(\rho)$ of an n-qubit state ρ. We can approximate ρ via*

$$\sigma_N(\rho) = \frac{1}{N} \sum_{l=1}^{N} \bigotimes_{i=1}^{n} (3|s_{l,i}\rangle\langle s_{l,i}| - I), \tag{25}$$

Definition 15 (Classical shadow of a quantum process [40])**.** *Given an n-qubit CPTP map $\mathcal{E}$. A size-N classical shadow $S_N(\mathcal{E})$ of the quantum process $\mathcal{E}$ is a random set given by*

$$S_N(\mathcal{E}) \triangleq \left\{ |\psi_l^{(in)}\rangle, |\psi_l^{(out)}\rangle \right\}_{l=1}^{N}, \tag{26}$$

where $|\psi_l^{(in)}\rangle = \bigotimes_{i=1}^{n} |s_{l,i}^{(in)}\rangle$ is a random input state with $|s_{l,i}^{(in)}\rangle \in stab_1$ sampled uniformly at random, and $|\psi_l^{(out)}\rangle = \bigotimes_{i=1}^{n} |s_{l,i}^{(out)}\rangle$ is the outcome of performing a random Pauli measurement on $\mathcal{E}(|\psi_l^{(in)}\rangle\langle\psi_l^{(in)}|)$.

The authors in [40] recently proposed a machine learning algorithm that is able to learn the average output behaviour of any quantum process, under some restrictions. In the learning phase, the algorithm works with the classical shadow of a generic quantum process $\mathcal{E}$ and a set of observables $\{O_i\}$. In the prediction phase, the algorithm receives an input quantum state ρ sampled from the target distribution D, and aims to predict $\mathrm{tr}\big(O_i\mathcal{E}(\rho)\big)$ for all observables in the set. The algorithm comes with a rigorous performance guarantee on the average prediction error over D, achieved with efficient time and sample complexity with respect to the number of qubits and error parameters. While the guarantee holds for any quantum process, there are certain restrictions on the observables $\{O_i\}$ and the distribution D. We partially state their results in Lemma 3 and refer to [40] for further details.

2.4 Computational Learning Theory

Computational learning theory studies what it means to *learn a function*. One of the most successful formal learning frameworks is undoubtedly the model of *Probably Approximately Correct*

(PAC) learning, which was introduced in [1]. In this model, we consider a class of target Boolean functions $\mathcal{C} \subseteq \{f | f : \{0,1\}^n \to \{0,1\}\}$, usually called the *concept class*. For an arbitrary concept $c \in \mathcal{C}$, a PAC learner receives samples of the form $\{x, c(x)\}$, where, in general, x is sampled from an unknown probability distribution $D : \{0,1\}^n \to [0,1]$. In the setting of noisy PAC learning, the bit $c(x)$ of each sample may independently be incorrect with some probability. The learner aims to output, with high probability, a hypothesis function h with low error on average over the distribution D.

Another widely studied access model is that of learning from Statistical Queries (SQs) [2]. Here, a learner is more restricted in the way it can interact with the data. Rather than learning from individual samples, the algorithm learns using the statistical properties of the data, making it more robust to noise. In particular, an SQ learner receives as input estimates of the expectation values of some chosen functions within specified error tolerance.

Quantum generalizations of both PAC and SQ learning have already been introduced and studied widely. The Quantum PAC (QPAC) model was introduced in [3], where the learner has access to a quantum computer and receives *quantum example states* as input. The quantum example state for a concept c over n input bits with the target distribution D is the $n+1$-qubit state $|\psi_c\rangle = \sum_x \sqrt{D(x)}|x, c(x)\rangle$. It has been shown in [4] that the sample complexity of quantum and classical PAC learning under unknown distributions is the same. However, over a fixed uniform distribution, learning with quantum queries can provide exponential advantage over the classical learner [10, 11]. Efficient learners from quantum queries under product distributions have also been shown [54, 55].

A quantum analogue of statistical queries was introduced in [5]. Here, the statistical query returns an approximation of the expectation value for an input measurement observable on quantum examples of the concept class to be learned. We include the quantum statistical query oracle defined in [5] below:

Definition 16 (QStat, from [5])**.** *Let $\mathcal{C} \subseteq \{c : \{0,1\}^n \to \{0,1\}\}$ be a concept class and $D : \{0,1\}^n \to [0,1]$ be a distribution over n-bit strings. A quantum statistical query oracle $\mathsf{QStat}(O, \tau)$ for some $c^* \in \mathcal{C}$ receives as inputs O, τ, where $\tau \geq 0$ and $O \in (\mathbb{C}^2)^{\otimes n+1} \times (\mathbb{C}^2)^{\otimes n+1}, ||O||_\infty \leq 1$, and returns a number α satisfying*

$$|\alpha - \langle\psi_{c^*}|O|\psi_{c^*}\rangle| \leq \tau,$$

where $|\psi_{c^*}\rangle = \sum_{x\in\{0,1\}^n} \sqrt{D(x)}|x, c^*(x)\rangle$.

Note that in the QSQ model of [5], while the learner can obtain an estimate of any measurement on the quantum examples, it is restricted only to classical computation. Interestingly, several concept classes such as parities, juntas, and DNF formulae are efficiently learnable in the QSQ model, whereas the classical statistical query model necessitates an exponentially larger number of queries. Additionally, the authors of [6] have established an exponential separation between QSQ learning and learning with quantum examples in the presence of classification noise.

One can also define quantum statistical queries for the task of learning states. In this case, the quantum example in Definition 16 would be replaced by the unknown quantum state to be learned.

Quantum statistical queries have also found practical applications in classical verification of quantum learning, as detailed in [43]. Furthermore, they have been employed in the analysis of quantum error mitigation models [6, 31] and quantum neural networks [56]. Alternative variations of quantum statistical queries have also been explored in [7, 8, 57].

3 Quantum Statistical Queries to Quantum Processes

In this section, we propose a framework and definition for learning quantum processes through quantum statistical queries and discuss its importance and relevance to different problems. Previously studied quantum statistical queries (Definition 16) have considered queries to quantum examples associated with some classical function [5]. While this model is the first generalisation of statistical query learning into the quantum setting and this type of query is useful for the problem of learning certain classes of classical functions, they do not encompass the quantum processes, and as such a new framework is needed for studying the learnability of quantum processes through

quantum statistical queries. Learning quantum processes from (often limited amount of) data is a crucial problem in physics and many areas of quantum computing such as error characterisation and error mitigation [19, 29, 30, 58, 59]. In many realistic and near-term scenarios, the only accessible data of the quantum process is through measured outcomes of such quantum channels, which are in fact nothing but statistical queries to such quantum processes. Hence studying the quantum process learnability via statistical queries is well motivated practically from the nature of quantum experiments and measurements. In what follows, we define our access model and then clarify its relationship to the previous definition of QSQs.

Definition 17 (Statistical Query to a Quantum Process (QPSQ))**.** *Let $\mathcal{E} : \mathbb{C}^d \to \mathbb{C}^d$ be a quantum process acting on a d-dimensional Hilbert space. A QPSQ learning algorithm has access to a quantum statistical query oracle* QPStat *of the process $\mathcal{E}$, which receives as input an observable $O \in \mathbb{C}^d \times \mathbb{C}^d$ satisfying $\|O\|_\infty \leq 1$, a quantum state $\rho \in \mathcal{S}_d$, and a tolerance parameter $\tau \geq 0$, and outputs a number α satisfying*

$$|\alpha - \mathrm{tr}\big(O\mathcal{E}(\rho)\big)| \leq \tau. \tag{27}$$

We denote the query as $\alpha \leftarrow \mathsf{QPStat}_{\mathcal{E}}(\rho, O, \tau)$[1]. The output α acts as an estimate of the expectation value of O on the state ρ after evolution under $\mathcal{E}$ within absolute error τ.

Our definition of QPSQs is justified in the setting where a learner has black-box access to a quantum process, with the ability to query the process with any quantum state. We note that our definition does not specify how the input quantum state ρ is provided to the oracle. The algorithm can provide multiple copies of the quantum state or can send the classical description of the quantum state to the oracle where they can be locally prepared, depending on the scenario and application. Obtaining the output of these queries is then achieved by the most natural operation, which is estimating the desired observable after evolution under the process. Again, since most physical properties of a quantum system are extracted through such interactions with its associated quantum channel, the application of this model in physics is straightforward. However, we will show that this model can be used in various scenarios including quantum cryptanalysis (see Section 6) and learning quantumly-encoded classical functions.

We now discuss what it means for a QPSQ learning algorithm to be *efficient*. Any such learning algorithm aiming to learn some property of the process $\mathcal{E}$ must be a quantum polynomial time (QPT) algorithm, making at most polynomially many queries to the oracle QPStat. As the learner itself must provide the input states or their classical description to QPStat, it must be possible to efficiently prepare the required copies of these states. This is an important point for the physical justification of this model. If the learner was able to query QPStat with arbitrary quantum states that require exponential quantum computation for preparation, the learning model would no longer be physically justified. On a similar note, the observables that an efficient learner provides to the oracle must also be efficiently measurable. Then, in the following definition for an *efficient* QPSQ learning algorithm, we only discuss the efficiency of the algorithm, not its correctness, allowing various notions of correctness depending on the desired property of $\mathcal{E}$ to be learned up for consideration.

Definition 18 (Efficient QPSQ learner)**.** *A QPSQ learning algorithm is called an* efficient *QPSQ learner if it makes at most* $\mathsf{poly}(\log(d))$ *queries with tolerance at least* $1/\mathsf{poly}(\log(d))$ *to the* $\mathsf{QPStat}_{\mathcal{E}}$ *oracle with states preparable in* $\mathsf{poly}(\log(d))$ *time, and observables measurable up to precision τ in* $\mathsf{poly}(\log(d), 1/\tau)$ *time, and runs in* $\mathsf{poly}(\log(d))$ *computational time.*

After formally defining our QPSQ model, we now talk about the relationship between our proposed model and other SQ models. It is easy to see that the QPStat oracle in Definition 17 generalizes the QStat oracle from Definition 16. We start by considering the unitary U_c associated with a Boolean function $c : \{0,1\}^n \to \{0,1\}$, $U_c|x, y\rangle = |x, y \oplus c(x)\rangle$, $\forall x \in \{0,1\}^n, y \in \{0,1\}$, and the quantum process $\mathcal{E}_c : \rho \to U_c \rho U_c^\dagger$. Let $|\psi_D\rangle = \sum_{x\in\{0,1\}^n} \sqrt{D(x)}|x, 0\rangle$ be a superposition state associated with a distribution D over $\{0,1\}^n$. Clearly, $\mathcal{E}_c(|\psi_D\rangle\langle\psi_D|) = |\psi_c\rangle\langle\psi_c|$. Thus, for any observable O, we can see that $\mathrm{Tr}\big(O\mathcal{E}_c(|\psi_D\rangle\langle\psi_D|)\big) = \langle\psi_c|O|\psi_c\rangle$, giving us the equivalence

$$\mathsf{QPStat}_{\mathcal{E}_c}(|\psi_D\rangle\langle\psi_D|, O, \tau) \equiv \mathsf{QStat}_{|\psi_c\rangle}(O, \tau). \tag{28}$$

[1]When referring to unitaries, we will sometimes abuse the notation and use QPStat_U to refer to the oracle associated with the unitary channel $\rho \to U\rho U^\dagger$.

Along with the definition of QStat, Arunachalam et. al. [5] presented algorithms for learning various concept classes in their QSQ model. The generalization of QStat by QPStat implies that these algorithms also hold in our QPSQ learning model. Given QPStat oracle access to the process $\mathcal{E}_c$ associated with the target concept, and sufficient copies of the state $|\psi_D\rangle$, the required output from QStat queries can be obtained using QPStat queries, and the remainder of the learning algorithms proceed identically.

Remark 1. Any concept class efficiently learnable in the QSQ model under a *known distribution D* can be learned efficiently given QPSQ access to the unitary encoding of the target function instead if $|\psi_D\rangle$ *can be prepared efficiently.* Thus, by extending the algorithms from [5], one can show that *there exist efficient QPSQ algorithms for learning parities, juntas and DNFs under the uniform distribution.* The generalization also holds in the *unknown distribution* setting assuming access to copies of $|\psi_D\rangle$.

4 Predicting Properties of Quantum Processes

In this section, we consider the problem of predicting properties of quantum processes, i.e., shadow process tomography, previously studied by [26, 47, 48]. We define the problem as follows.

Problem 1 (Shadow Tomography of a Quantum Process)**.** *Let* $0 < \epsilon, \delta < 1$ *and* $M \geq 1$*. Given access to an unknown quantum process* $\mathcal{E}$*, and a list of quantum states and observables* $\{(\rho_i, O_i)\}_{i\in[M]}$*, produce predictions* $\{h(\rho_i, O_i)\}_{i\in[M]}$ *such that, with probability at least* $1-\delta$*,*

$$|h(\rho_i, O_i) - \mathrm{tr}\big(O_i\mathcal{E}(\rho_i)\big)| \leq \epsilon, \forall i \in [M]. \tag{29}$$

We will focus primarily on the average-case relaxation of this problem [40], where the error only needs to be low on average over random input states, rather than on all input states.

Problem 2 (Average-case Shadow Process Tomography)**.** *Let* $0 < \epsilon, \delta < 1$ *and* $M \geq 1$*. Given access to an unknown quantum process* $\mathcal{E}$*, and a list of observables* $\{O_i\}_{i\in[M]}$ *as well as quantum states drawn randomly from some distribution* $\mathcal{D}$*, produce predictions* $h(\rho, O_i)$ *such that, with probability at least* $1-\delta$*,*

$$\mathbf{E}_{\rho\sim\mathcal{D}}\left[|h(\rho, O_i) - \mathrm{tr}\big(O\mathcal{E}(\rho)\big)|^2\right] \leq \epsilon^2, \forall i \in [M]. \tag{30}$$

We state our main result on Problem 2 from access to QPStat queries in the following theorem.

Theorem 1 (Average-case Shadow Process Tomography from QPSQs)**.** *There exists an algorithm for solving Problem 2 for* M *observables* $\{O_i\}_{i\in[M]}$ *and a distribution* $\mathcal{D}$*, where each* O_i *is an* n*-qubit observable with* $\|O_i\|_\infty \leq 1$*, given as a sum of few-body* ($\leq \kappa = \mathcal{O}(1)$) *observables, where each qubit is acted on by* $\mathcal{O}(1)$ *of the few-body observables, and* $\mathcal{D}$ *is a distribution over quantum states that is invariant under single-qubit Clifford operations, using*

$$N = M\log\big(Mn/\delta\big)2^{\mathcal{O}(\log(n)\log\left(1/\epsilon\right))} \tag{31}$$

queries of tolerance $1/2^{\mathcal{O}(\log(n)\log\left(1/\epsilon\right))}$ *to* $\mathsf{QPStat}_{\mathcal{E}}$ *and computational time* $N\cdot 2^{\mathcal{O}(\log(n)\log\left(1/\epsilon\right))}$*. Moreover, any algorithm solving Problem 2 with high probability for* M *Pauli observables and any distribution over target quantum states must make*

$$\Omega\left(\frac{M\tau^2}{\epsilon^2}\right) \tag{32}$$

QPStat *queries of tolerance* τ *to the unknown channel.*

We prove our upper bound by adapting the learning algorithm from [40] to the QPSQ setting, resulting in a factor-M overhead in the query complexity compared to their result. Our lower bound complements this by showing that this overhead is necessary - any QPSQ algorithm for average-case shadow tomography of any quantum process must make $\Omega(M)$ queries. Further, as

the lower bound holds for any M Paulis and any distribution over states, it also applies to the kinds of local observables and distributions considered in the upper bound. Thus, we see that with respect to the number of observables, the upper bound is tight up to a logarithmic factor. Moreover, for $\epsilon = \Theta(1)$, the query and time complexities of the upper bound scale polynomially with n.

The proof of our lower bound uses a similar construction to the lower bound for shadow tomography of quantum states from QSQs shown in [6]. For our upper bound, we adapt the learning algorithm of [40] to the QPSQ setting. We present the proof of our lower bound in Section 4.1 and that of our upper bound in Section 4.2. We support the theoretical guarantees of our learning algorithm with numerical simulations in Section 4.3.

4.1 Proof of Lower Bound

In this section, we prove the lower bound of Theorem 1 on average-case shadow process tomography (Problem 2) from QPStat queries. We start by presenting a lower bound for Problem 1 when the observables are Paulis. First, we recall a result from [6] on Pauli Shadow Tomography of quantum states from QSQs.

Lemma 2 (Adapted from Theorem 27 of [6])**.** *Let $C \subseteq \mathcal{P}_n$, with $|C| = M$, be a set of n-qubit Pauli operators. Let S be the associated class of states*

$$\mathcal{S}_\epsilon = \left\{ \rho_{\epsilon,P} = \frac{\mathbb{I} + 3\epsilon P}{2^n}, P \in \mathcal{C} \right\}. \tag{33}$$

Then, any algorithm that distinguishes between $\rho \in \mathcal{S}_\epsilon$ and $\rho = \frac{\mathbb{I}}{2^n}$ must make $\Omega(M\tau^2/\epsilon^2)$ queries of tolerance τ to QStat_ρ.

While [6] showed the above result for the specific case of $\mathcal{C} = \mathcal{P}_n$, their proof can easily be adapted to show this for any subset of $\mathcal{P}_n$. Further, [6] noted that this distinguishing task reduces to shadow tomography of the set of Pauli operators, resulting in a QSQ lower bound for the latter. We construct a similar distinguishing task for quantum channels, giving us a lower bound for Problem 1.

Proposition 1 (Worst-Case Hardness of Shadow Process Tomography from QPSQs)**.** *Any algorithm solving Problem 1 for M Pauli observables and arbitrary input states within precision ϵ must make $\Omega(M\tau^2/\epsilon^2)$ queries of tolerance τ to $\mathsf{QPStat}_\mathcal{E}$.*

Proof. Let $C \subseteq \mathcal{P}_n$, with $|C| = M$, be the set of M Pauli observables. Consider the class of channels $\mathcal{C}_\epsilon$ that prepare states from $\mathcal{S}_\epsilon$ (Lemma 2), i.e.,

$$\mathcal{C}_\epsilon = \left\{ \Phi_{\epsilon,P} : \rho_{\text{in}} \to \text{tr}(\rho_{\text{in}}) \frac{\mathbb{I} + 3\epsilon P}{2^n} \right\}. \tag{34}$$

Given access to $\mathsf{QPStat}_\mathcal{E}$ queries of tolerance τ, and promised that $\mathcal{E} \in \mathcal{C}_\epsilon$ or $\mathcal{E} = \Phi_{\text{dep}}$, we consider the task of determining which is the case. Clearly, for any input state ρ, and Paulis $P, Q \in \mathcal{C}$,

$$\text{tr}\big(P\Phi_{\epsilon,Q}(\rho)\big) = 3\epsilon\delta_{P,Q}. \tag{35}$$

Meanwhile, for any Pauli $P \in \mathcal{C}$,

$$\text{tr}\big(P\Phi_{\text{dep}}(\rho)\big) = 0. \tag{36}$$

Thus, any algorithm for solving Problem 1 for all observables in $\mathcal{C}$ and arbitrary input states, implies an algorithm for distinguishing between $\mathcal{C}_\epsilon$ and Φ_{dep}.

Now, to lower bound the query complexity of this distinguishing task, we note that the channels in this problem simply discard the input state and prepare a fixed output state. Suppose the output state of a channel $\mathcal{E} \in \mathcal{C}_\epsilon \cup \{\Phi_{\text{dep}}\}$ is ρ_{out}. Clearly, the responses of $\mathsf{QPStat}_\mathcal{E}(\rho_{\text{in}}, O, \tau)$ are indistinguishable from those of $\mathsf{QStat}_{\rho_{\text{out}}}(O, \tau)$. Recall that the class of output states of $\mathcal{C}_\epsilon$ is exactly $\mathcal{S}_\epsilon$ and that the output of Φ_{dep} is always $\frac{\mathbb{I}}{2^n}$. Thus, any algorithm for distinguishing $\mathcal{C}_\epsilon$ versus Φ_{dep} using QPStat queries can be used to distinguish between $\mathcal{S}_\epsilon$ and $\frac{\mathbb{I}}{2^n}$ using the same number of QStat queries. Together with Lemma 2, we obtain the desired lower bound. □

While we have shown a lower bound for worst-case shadow tomography of quantum processes (Problem 1), Problem 2 may be easier as it only requires the error to be low on average over states, rather than on any input quantum state. We prove the lower bound by showing that the same distinguishing task from Proposition 1 can be solved by an algorithm for Problem 2 without additional queries.

Proof of Theorem 1 (Lower Bound). Suppose there exists an algorithm $\mathcal{A}$ for Problem 2 from QPStat queries. Let $C \subseteq \mathcal{P}_n$, with $|C| = M$, be the set of M Pauli observables. We apply $\mathcal{A}$ to a channel $\mathcal{E} \in \mathcal{C}_\epsilon \cup \{\Phi_{\text{dep}}\}$. Then, from the correctness of $\mathcal{A}$, we obtain a hypothesis h, such that for the target distribution $\mathcal{D}$ over input states,

$$\mathbf{E}_{\rho\sim D}\left[|h(\rho, P) - \operatorname{tr}\big(P\mathcal{E}(\rho)\big)|^2\right] \leq \epsilon^2, \forall P \in \mathcal{C}. \tag{37}$$

For $\mathcal{E} \in \mathcal{C}_\epsilon \cup \{\Phi_{\text{dep}}\}$, the output state $\mathcal{E}(\rho)$ is independent of ρ. Let $\alpha_P = \operatorname{tr}\big(P\mathcal{E}(\rho)\big)$ for any input ρ. Then, we have

$$\mathbf{E}_{\rho\sim D}\left[|h(\rho, P) - \alpha_P|^2\right] \leq \epsilon^2, \forall P \in \mathcal{C}. \tag{38}$$

Now, we compute the mean of the hypothesis function, $\mu_P = \mathbf{E}_{\rho\sim\mathcal{D}}[h(\rho, P)]$. While this step may be computationally expensive, it does not require any additional QPStat queries. Our distinguisher will use μ_P as an estimate for α_P. We bound the error as follows.

$$\begin{aligned} |\mu_P - \alpha_P| &= \left|\mathbf{E}_{\rho\sim D}\left[h(\rho, P) - \alpha_P\right]\right| && (39)\\ &\leq \sqrt{\mathbf{E}_{\rho\sim D}\left[|h(\rho, P) - \alpha_P|^2\right]} && (40)\\ &\leq \epsilon && (41) \end{aligned}$$

where the first inequality follows from Jensen's inequality, and the second inequality follows from Equation (38). Thus, using $\mathcal{A}$, we can estimate α_P within error ϵ for all $P \in \mathcal{C}$, without any additional QPStat queries. From the proof of Proposition 1, we see that this is enough to distinguish between $\mathcal{C}_\epsilon$ and Φ_{dep}. Together with Proposition 1, we now obtain the stated lower bound. □

4.2 Proof of Upper Bound

In this section, we present an algorithm achieving the upper bound of Theorem 1. The algorithm follows from a straightforward adaptation of the classical shadow algorithm for this problem by [40]. The key result of [40] we use is that to succeed at Problem 2, it suffices to estimate (in a certain sense) the low-degree Pauli coefficients of the Heisenberg-evolved observables. We state this result in the following lemma.

Lemma 3 (Adapted from [40])**.** *Given $\epsilon, \delta > 0$, a distribution $\mathcal{D}$ over quantum states that is invariant under single-qubit Clifford operations, and an n-qubit observable O given as a sum of few-body observables with degree $\mathcal{O}(1)$, where each qubit is acted on by $\mathcal{O}(1)$ of the few-body observables and $\|O\|_\infty \leq 1$. Define hyperparameters $k, \tilde{\epsilon}$, where*

$$k = \lceil log_{1.5}(2/\epsilon^2)\rceil, \quad \tilde{\epsilon} = \Theta\left(\frac{\epsilon^2}{(2n)^k}\right). \tag{42}$$

Let $\mathcal{E}$ be the unknown process and consider the Heisenberg-evolved observable $\mathcal{E}^\dagger(O)$ with Pauli coefficients $\alpha_P(O)$, i.e.,

$$\mathcal{E}^\dagger(O) = \sum_{P\in\mathcal{P}_n} \alpha_P(O)P. \tag{43}$$

Then, given the description of a random state $\rho \sim \mathcal{D}$, in order to achieve low average-case error over $\mathcal{D}$,

$$\mathbf{E}_{\rho\sim\mathcal{D}}\left[|h(\rho, O) - \operatorname{tr}\big(O\mathcal{E}(\rho)\big)|^2\right] \leq \epsilon^2, \tag{44}$$

it suffices to obtain $\tilde{\epsilon}$-accurate estimates of $\frac{1}{3^{|P|}}\alpha_P(O)$ for all Paulis P with degree $|P| \leq k$ along with $\mathcal{O}(k(3n)^k)$ computational time for each prediction.

We note that [40] additionally consider another setting for the hyperparameters $k, \tilde{\epsilon}$ when an additional error parameter is included. For simplicity, we only consider the hyperparameter setting stated in Equation (42), but our analysis can be extended to the other setting in [40].

Now, from Lemma 3, we see that we need to learn the low-degree Pauli coefficients of $\mathcal{E}^\dagger(O)$ from queries to QPStat. Apart from this step, the rest of the algorithm proceeds identically to that of [40]. For completeness, we detail the full algorithm in Algorithm 1, where we only discuss the case for a single observable. This can be extended to M observables by repeating the algorithm for each observable and increasing the number of queries by readjusting the failure probability of each repetition to δ/M. In Algorithm 1, we make use of $\|O\|_{\text{Pauli},1}$, which denotes the l_1 norm of the Pauli coefficients of O. Specifically, for $O = \sum_{P\in\mathcal{P}_n} a_P P$, we have $\|O\|_{\text{Pauli},1} = \sum_{P\in\mathcal{P}_n} |a_P|$. While

Algorithm 1 Learning to predict properties of a quantum process from QPSQs

$k \leftarrow \lceil log_{1.5}(2/\epsilon^2)\rceil, \tilde{\epsilon} \leftarrow \Theta\left(\frac{\epsilon^2}{(2n)^k}\right)$

Gather Data:
for $l = 1$ to N **do**
  $|\psi_l^{(in)}\rangle \leftarrow \bigotimes_{i=1}^n |s_{l,i}^{(in)}\rangle, |s_{l,i}^{(in)}\rangle \in \text{stab}_1$, chosen uniformly at random
  $y_l \leftarrow \mathsf{QPStat}_{\mathcal{E}}(|\psi_l^{(in)}\rangle\langle\psi_l^{(in)}|, O, \tau)$
end for
return $S_N(\mathcal{E}, O) = \{|\psi_l^{(in)}\rangle, y_l\}_{l=1}^N$

Learning:
for all $P \in \mathcal{P}_n, |P| \leq k$ **do**
  $\hat{x}_P(O) \leftarrow \frac{1}{N}\sum_{l=1}^N y_l \operatorname{tr}\big(P|\psi_l^{in}\rangle\langle\psi_l^{in}|\big)$
  if $(\frac{1}{3})^{|P|} > 2\tilde{\epsilon}$ and $|\hat{x}_P(O)| > 2.3^{|P|/2}\sqrt{\tilde{\epsilon}}\|O\|_{\text{Pauli},1}$ **then**
    $\hat{\alpha}_P(O) \leftarrow 3^{|P|}\hat{x}_P(O)$
  else
    $\hat{\alpha}_P(O) \leftarrow 0$
  end if
end for

Prediction for target state ρ:
$h(\rho) \leftarrow \sum_{P:|P|\leq k} \hat{\alpha}_P(O)\operatorname{tr}(P\rho)$
return $h(\rho)$

Lemma 3 requires the classical description of a state ρ to make predictions on it, from Algorithm 1, we see that it suffices to have estimates of $\operatorname{tr}(P\rho)$ for all low-degree Paulis $P \in \mathcal{P}_n, |P| \leq k$. Instead of a complete classical description, this could also be estimated using classical shadows [15] or from QStat_ρ queries, depending on the access available.

Proof of Theorem 1 (Upper Bound). From Lemma 3, it suffices to estimate the Pauli coefficients $\frac{1}{3^{|P|}}\alpha_P(O)$ within precision $\tilde{\epsilon}$. Let $\text{stab}_1^{\otimes n}$ be the uniform distribution over the tensor product of n single-qubit stabilizer states. Then, [40] showed that for all Paulis $P \in \mathcal{P}_n$,

$$\frac{1}{3^{|P|}}\alpha_P(O) = \mathop{\mathbf{E}}_{\rho\sim\text{stab}_1^{\otimes n}} \operatorname{tr}(P\rho)\operatorname{tr}\big(O\mathcal{E}(\rho)\big). \tag{45}$$

Denote $x_P(O) \triangleq \frac{1}{3^{|P|}}\alpha_P(O)$. We construct estimators for $x_P(O)$ by performing queries to QPStat with random states $\rho_l \sim \text{stab}_1^{\otimes n}, l \in [N]$. Let $y_l \leftarrow \mathsf{QPStat}_{\mathcal{E}}(\rho_l, O, \tau)$. Then, for all low-degree Paulis $P \in \mathcal{P}_n, |P| \leq k$, we construct estimators

$$\hat{x}_P(O) = \frac{1}{N}\sum_{l\in[N]} \operatorname{tr}(P\rho_l)y_l. \tag{46}$$

Before bounding the error between $\hat{x}_P(O)$ and $x_P(O)$, we first define the intermediate random variable $x'_P(O)$

$$x'_P(O) \triangleq \frac{1}{N}\sum_{l=1}^{N} \operatorname{tr}(P\rho_l)\operatorname{tr}\big(O\mathcal{E}(\rho_l)\big). \tag{47}$$

Now, we bound the error between $x_P(O)$ and $x'_P(O)$. Using $N = \mathcal{O}\left(\frac{\log\left((3n)^k/\delta\right)}{(\tilde{\epsilon}-\tau)^2}\right)$, from Hoeffding's inequality and assuming $\tau < \tilde{\epsilon}$, we see that for any $P \in \mathcal{P}_n, |P| \leq k$, with probability at least $1 - \delta/(3n)^k$,

$$|x_P(O) - x'_P(O)| \leq \tilde{\epsilon} - \tau. \tag{48}$$

Using a union bound over failure probabilities, and noting that there are at most $(3n)^k$ Pauli operators of degree k, we see that Equation (48) holds for all $P \in \mathcal{P}_n, |P| \leq k$ with probability at least $1 - \delta$. For the following arguments, we condition on this event. Next, we bound the error between $x'_P(O)$ and $\hat{x}_P(O)$.

$$|x'_P(O) - \hat{x}_P(O)| = \left|\frac{1}{N}\sum_{l\in[N]} \operatorname{tr}(P\rho_l)\left(\operatorname{tr}\big(O\mathcal{E}(\rho_l)\big) - y_l\right)\right| \tag{49}$$

$$\leq \frac{1}{N}\sum_{l\in[N]} \left|\left(\operatorname{tr}\big(O\mathcal{E}(\rho_l)\big) - y_l\right|\right| \tag{50}$$

$$\leq \tau, \tag{51}$$

where the first inequality uses the triangle inequality and that $\|P\|_\infty = 1$. Now, applying the triangle inequality to Equations (48) and (51), we have

$$|x_P(O) - \hat{x}_P(O)| \leq \tilde{\epsilon}, \tag{52}$$

which is the desired bound on the error. The stated query complexity for $M = 1$ can be obtained by setting $\tau = \tilde{\epsilon}/2$ and selecting k and $\tilde{\epsilon}$ as stated in Equation (42). To obtain the query complexity for general M, we can repeat the entire algorithm for each observable with failure probability at most δ/M.

The computational complexity is dominated by estimating the Pauli coefficients, each of which requires time $\mathcal{O}(kN)$. The stated complexity can be obtained by noting that we estimate at most $(3n)^k$ coefficients. □

In certain practical cases, the following additional assumption about the output α of the oracle $\mathsf{QPStat}_{\mathcal{E}}$ may hold.

Assumption 1. *Consider any input state ρ, observable O and tolerance $\tau > 0$. Let $\alpha \leftarrow \mathsf{QPStat}_{\mathcal{E}}(\rho, O, \tau)$ be the output of a quantum statistical query. Then, we assume that α satisfies*

$$\mathbf{E}[\alpha] = \operatorname{tr}\big(O\mathcal{E}(\rho)\big), \tag{53}$$

where the expectation is over the oracle's internal randomness.

In such cases, we can obtain a tighter query complexity upper bound to solve Problem 2, and without the restriction that $\tau < \tilde{\epsilon}$. We present this result in the following corollary.

Corollary 1. *Under Assumption 1 on the output of QPStat, Algorithm 1 succeeds at Problem 2 using*

$$N = \tau^2 M \log\big(Mn/\delta\big) 2^{\mathcal{O}(\log(n)\log\left(1/\epsilon\right))} \tag{54}$$

queries of tolerance τ to $\mathsf{QPStat}_{\mathcal{E}}$.

Proof. Under this assumption, we can eliminate the intermediate quantity $x'_P(Q)$ in the previous proof, and directly apply Hoeffding's inequality to obtain the bound on $|\hat{x}_P(Q) - x_P(Q)|$. Note that for all $l \in [N]$,

$$|\operatorname{tr}(P\rho_l)y_l - \operatorname{tr}(P\rho_l)\operatorname{tr}(O\mathcal{E}\rho_l)| \leq \tau, \tag{55}$$

and

$$\mathbf{E}_{\rho_l}[\mathrm{tr}(P\rho_l)y_l] = x_P(O). \tag{56}$$

Thus, using $N = \mathcal{O}\left(\frac{\tau^2 \log\left((3n)^k/\delta\right)}{\tilde{\epsilon}^2}\right)$, from Hoeffding's inequality, we see that for any $P \in \mathcal{P}_n, |P| \leq k$, with probability at least $1 - \delta/(3n)^k$,

$$|x_P(O) - \hat{x}_P(O)| \leq \tilde{\epsilon}. \tag{57}$$

Using a union bound over failure probabilities, and noting that there are at most $(3n)^k$ Pauli operators of degree k, we see that Equation (57) holds for all $P \in \mathcal{P}_n, |P| \leq k$. The stated query complexity can be obtained using Equation (42). □

4.3 Numerical Simulations

In this section, we demonstrate the performance of Algorithm 1 through numerical simulations. The code for our simulations is available in a public Github repository [2]. Before presenting our simulations, we remark on our approach towards simulating the QPStat oracle. In order to construct the output of QPStat, we assume the oracle uses a method to estimate the expectation value of an observable, such as the ones shown in [15, 60–62]. In our simulation, in order to emulate the behaviour of these methods, we compute the true expectation value and output the result after adding a normally distributed error to it. The error is sampled from a normal distribution such that it is within the specified tolerance with high probability. We note that our learning model puts no assumptions on the error in the output of the queries, and in theory, this error can come from any arbitrary distribution as long as it is within the tolerance with high probability. However, we will argue that this simple method of generating the sample data already captures these scenarios well enough for the purpose of our simulations. We also compare this method with the use of classical shadows for evaluating the outcome of an observable [15] in Figure 1. We see that for the same target tolerance and success probabilities, the classical shadow method produces less error than that generated using a normal distribution. In fact, the normally distributed error achieves the exact value for the success probability, while any practical method would produce the same error or less, as it would come with a potentially looser bound. Thus, our emulated oracle would produce greater deviations than in a practical implementation, implying that the real-life performance of the algorithm would only be similar to or better than the simulations. We use these emulated oracles for the simulation of the learning algorithms.

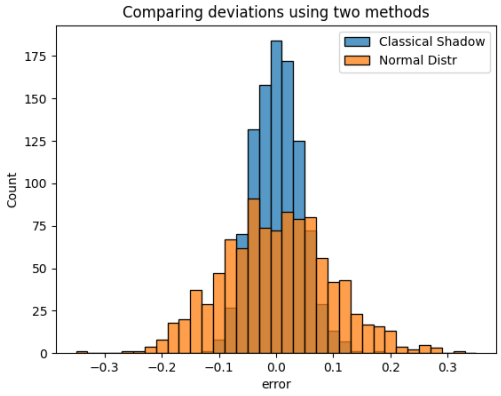
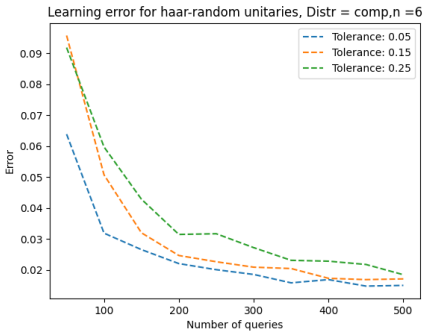
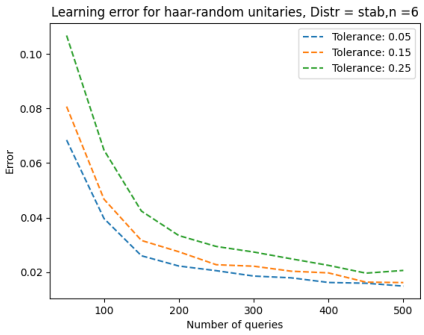


Figure 1: Comparison between simulated errors generated from a normal distribution and those generated using classical shadow tomography to evaluate the EV of the Pauli-Z observable on random single-qubit stabilizer states after evolution under a fixed haar-random unitary. We fix a tolerance value $\tau = 0.2$, and the probability of the deviation lying outside the tolerance, $\delta = 0.0455$

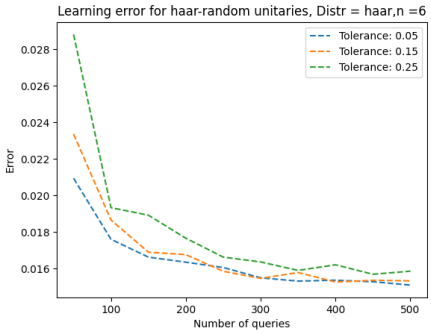
[2] https://github.com/chirag-w/qpsq-learning



(a) Computational basis states



(b) Stabilizer states



(c) Haar-random states

Figure 2: Average performance of the learning algorithm on 10 haar-random 6-qubit unitaries, in predicting the outcome of Z_1 on three target distributions

In Figure 2, we show the simulated performance of Algorithm 1 in predicting properties of 10 haar-random unitaries over 6 qubits for a range of tolerances. We consider $O = Z \otimes I ... \otimes I$, the Pauli-Z observable on the first qubit. We consider three distributions of target states, namely the uniform distributions over the computational basis states, the stabilizer product states and haar-random states. We can see from Figure 2 that a lower QPSQ tolerance results in a lower prediction error for the same number of queries. We also see that the algorithm achieves similar performance when predicting the outcome on computational basis and stabilizer product states, even though the uniform distribution over the computational basis states is not locally flat and thus outside the performance guarantee. On the other hand, the distribution over haar-random states is within the guarantee, and the algorithm performs best on this distribution.

5 Lower Bounds for Diamond Distance Learning

In this section, rather than predicting properties of quantum processes, we consider the harder problem of learning them within diamond distance from QPStat queries. Here, we provide average-case query complexity lower bounds for learning unitary 2-designs and Haar-random unitaries. We start by stating our lower bound for average-case learning exact 2-designs (Definition 10).

Theorem 2 (Lower bound for learning exact unitary 2-designs)**.** *Let $0 < \tau \leq \epsilon, \epsilon + 2\tau < 1 - 1/2^n$, and $\mu_H^{(2)}$ be an exact unitary 2-design over n-qubits. Let there be an algorithm that, given access to QPStat_U for some $U \sim \mu_H^{(2)}$, outputs a quantum channel Φ with $d_\diamond(\mathcal{U}, \Phi) \leq \epsilon$ with probability α over its internal randomness and probability β over $U \sim \mu_H^{(2)}$. Then, such an algorithm must make q queries of tolerance τ to QPStat_U, where*

$$q + 1 \geq (2\alpha - 1)\beta\tau^2(2^n + 1). \tag{58}$$

Remark 2. As random Clifford circuits are known to form unitary 3-designs [63, 64], Theorem 2 implies an exponential lower bound for learning random Cliffords from QPStat queries. Mean-

while, Clifford circuits can be learned efficiently using black-box access [65, 66]. This gives us an exponential separation between QPSQ learners and those in the standard setting.

We now state our lower bound for learning additive approximate unitary 2-designs (Definition 10).

Theorem 3 (Lower bound for learning approximate unitary 2-designs)**.** *Let* $0 < \tau \leq \epsilon < 1, \epsilon + 2\tau < 1 - 1/2^n, 0 \leq \delta$*, and* $\mu_H^{(2,\delta)}$ *be a* δ*-approximate unitary* 2*-design and exact unitary* 1*-design over* n*-qubits. Let there be an algorithm that, given access to* QPStat_U *for some* $U \sim \mu_H^{(2,\delta)}$*, outputs a quantum channel* Φ *with* $d_\diamond(\mathcal{U}, \Phi) \leq \epsilon$ *with probability* α *over its internal randomness and probability* β *over* $U \sim \mu_H^{(2,\delta)}$*. Then, such an algorithm must make* q *queries of tolerance* τ *to* QPStat_U*, where*

$$q + 1 \geq (2\alpha - 1)\beta\tau^2\Omega(\min\{2^n, 1/\delta\}). \tag{59}$$

Remark 3. For our lower bound on approximate 2-designs, Theorem 3 additionally assumes that the measure is an exact 1-design. This assumption is satisfied in practice by brickwork random quantum circuits, which form exact 1-designs at any circuit depth and approximate 2-designs at depths $\mathcal{O}(n + log(1/\delta))$ [67]. This restriction can also be lifted in some cases. One can see that the proof of Theorem 3 also holds when the measure is a δ'-approximate 1-design with $\delta' < \tau/c$ for some absolute constant $c > 1$, but we do not explicitly discuss this case here.

Theorem 2 already implies an exponential lower bound for learning Haar-random unitaries. However, using stronger concentration properties of the Haar measure, we show a doubly exponential lower bound in the following theorem.

Theorem 4 (Lower bound for learning Haar-random unitaries)**.** *Let* $0 < \tau \leq \epsilon < 1, \epsilon + 2\tau < 1 - 1/2^n$*, and* μ_H *be the unitary Haar-measure over* n*-qubits. Let there be an algorithm that, given access to* QPStat_U *for some* $U \sim \mu_H$*, outputs a quantum channel* Φ *with* $d_\diamond(\mathcal{U}, \Phi) \leq \epsilon$ *with probability* α *over its internal randomness and probability* β *over* $U \sim \mu_H$*. Then, such an algorithm must make* q *queries of tolerance* τ *to* QPStat_U*, where*

$$q + 1 \geq (\alpha - 1/2)\beta \exp\left(\frac{2^n\tau^2}{48}\right). \tag{60}$$

To prove these lower bounds, we first describe a many-vs-one distinguishing task and lower bound its QPStat query complexity in Section 5.1. Then, in Section 5.2, we show that for any measure over unitaries, average-case learning is at least as hard as a certain instance of the distinguishing task. Finally, we conclude the proofs of our lower bounds by computing the specific query complexities for this distinguishing task in Sections 5.3 and 5.4. Our techniques are similar to those used in other quantum statistical query lower bounds [6, 8, 9], which are in turn inspired by lower bounds in the classical SQ setting [44].

5.1 Many-vs-One Query Complexity

In this section, we will lower bound the QPSQ complexity of a many-vs-one distinguishing task for quantum channels. Specifically, consider a class of quantum channels $\mathcal{C}$ and another reference quantum channel $\Phi \notin C$. Then, given access to an unknown channel $\mathcal{E}$ through $\mathsf{QPStat}_\mathcal{E}$, one must determine whether $\mathcal{E} \in \mathcal{C}$, or if $\mathcal{E} = \Phi$, promised that one of these is the case. We provide the query complexity lower bound for this task in the following lemma.

Lemma 4 (Many-vs-one distinguishing query complexity)**.** *Let* $\mathcal{C}$ *be a class of quantum channels from* $\mathcal{M}_{N,N}$ *to* $\mathcal{M}_{N,N}$*,* $\Phi \notin \mathcal{C}$ *be another quantum channel from* $\mathcal{M}_{N,N}$ *to* $\mathcal{M}_{N,N}$*. Then, given access to* $\mathsf{QPStat}_\mathcal{E}$*, any algorithm that can distinguish, with probability* α *over its internal randomness, between* $\mathcal{E} \in \mathcal{C}$ *and* $\mathcal{E} = \Phi$*, for any measure* μ *over* $\mathcal{C}$*, must make at least*

$$q \geq \frac{2\alpha - 1}{\max_{\rho,O} \mathbf{Pr}_{\mathcal{E}\sim\mu}\left[\left|\operatorname{tr}\big(O\mathcal{E}(\rho)\big) - \operatorname{tr}\big(O\Phi(\rho)\big)\right| > \tau\right]} \tag{61}$$

queries of tolerance τ *to* $\mathsf{QPStat}_\mathcal{E}$*.*

Proof. Let $\mathcal{A}$ be the algorithm that distinguishes between $\mathcal{C}$ and Φ. Suppose $\mathcal{A}$ makes q QPStat queries $\{(\rho_i, O_i)\}_{i\in[q]}$ based on its internal randomness and the previous responses. Let p_d be the probability that $\mathcal{A}$ makes a distinguishing query, i.e.

$$p_d = \mathbf{Pr}_{\mathcal{A},\mathcal{E}\sim\mu}\left[\exists i \in [q] : |\operatorname{tr}\big(O_i\mathcal{E}(\rho_i)\big) - \operatorname{tr}\big(O_i\Phi(\rho_i)\big)| > \tau\right]. \tag{62}$$

Now, suppose $\mathcal{A}$ receives $\operatorname{tr}\big(O_i\Phi(\rho_i)\big)$ as responses for all q queries. When $\mathcal{E} \in \mathcal{C}$, this happens with probability at most $1 - p_d$. In this case, as the queries are consistent with $\mathcal{E} = \Phi$, by the correctness of $\mathcal{A}$, we obtain "$\mathcal{E} \in \mathcal{C}$" with probability at most $1 - \alpha$. Thus, by the correctness of $\mathcal{A}$, we have

$$\alpha \leq p_d \cdot 1 + (1 - p_d)(1 - \alpha) \leq 1 - \alpha + p_d, \tag{63}$$

where the second inequality uses that $\alpha \leq 1$. Rearranging the terms, we obtain

$$p_d \geq 2\alpha - 1. \tag{64}$$

We can now upper bound p_d using the union bound.

$$\begin{aligned} p_d &= \mathbf{Pr}_{\mathcal{A},\mathcal{E}\sim\mu}\left[\exists i \in [q] : |\operatorname{tr}\big(O_i\mathcal{E}(\rho_i)\big) - \operatorname{tr}\big(O_i\Phi(\rho_i)\big)| > \tau\right] && (65)\\ &\leq \sum_{i\in[q]} \mathbf{Pr}_{\mathcal{A},\mathcal{E}\sim\mu}\left[|\operatorname{tr}\big(O_i\mathcal{E}(\rho_i)\big) - \operatorname{tr}\big(O_i\Phi(\rho_i)\big)| > \tau\right] && (66)\\ &\leq q \max_{\rho,O} \mathbf{Pr}_{\mathcal{E}\sim\mu}\left[|\operatorname{tr}\big(O\mathcal{E}(\rho)\big) - \operatorname{tr}\big(O\Phi(\rho)\big)| > \tau\right]. && (67) \end{aligned}$$

Together with Equation (64), we obtain the desired result. □

5.2 Average-Case Learning Query Complexity

In this section, we provide an average-case query complexity lower bound for learning unitaries from some measure. In Lemma 5, we show that learning a class of unitaries is as hard as a certain instance of the many-vs-one distinguishing task of Section 5.1. We then combine Lemmas 4 and 5 to obtain our average-case learning lower bound in Lemma 6.

Given an algorithm for learning classes of quantum states and distributions from QSQs, [6, 8, 9] show that a single additional query suffices to distinguish this class from any reference object that is sufficiently far. However, such a result does not hold in general when using QPStat queries for distinguishing channels far in diamond distance. This is due to the fact that the optimal distinguishing state may need to be prepared over an additional ancillary register, and such a query cannot be made with our chosen definition of the QPStat oracle.

While we are unable to show a general reduction from distinguishing to learning with a single query, we will look at the specific case when the class of channels consists only of unitary channels and the fixed reference object is the depolarizing channel. In particular, we show that learning a class of unitaries from QPSQs is as hard as distinguishing it from the depolarizing channel.

Lemma 5 (Learning unitaries is as hard as distinguishing them from Φ_{dep}.)**.** *Let $0 < \tau \leq \epsilon < 1$, $\epsilon + 2\tau < 1 - \frac{1}{2^n}$. Let $\mathcal{C} \subseteq \mathcal{U}_{2^n}$ be a class of unitary channels. Let $\mathcal{A}$ be a learning algorithm that, given access to QPStat_U for some $U \in \mathcal{C}$, and with probability α over its internal randomness, outputs a channel Φ such that $d_\diamond(\mathcal{U}, \Phi) \leq \epsilon$, using q_L queries of tolerance τ to QPStat_U. Then,*

$$q_L + 1 \geq q_D, \tag{68}$$

where q_D is the number of QPStat queries of tolerance τ necessary to distinguish $\mathcal{C}$ from the maximally depolarizing channel Φ_{dep} with probability α.

Proof. We will use $\mathcal{A}$ as a subroutine and construct a distinguisher making at most one additional query. We first run $\mathcal{A}$ on the unknown channel $\mathcal{E} \in \mathcal{C} \cup \{\Phi_{\text{dep}}\}$. If the output of the algorithm is not a valid quantum channel, we output "$\mathcal{E} = \Phi_{\text{dep}}$". Otherwise, let the output be a quantum channel Φ. Now, we check if $\min_{\mathcal{E}\in\mathcal{C}} d_\diamond(\mathcal{E}, \Phi) \leq \epsilon$, and output "$\mathcal{E} = \Phi_{\text{dep}}$" if not. While this step may be computationally expensive, it does not incur any additional queries.

Now, the output of $\mathcal{A}$ will be some quantum channel Φ. Then, we query $\mathsf{QPStat}_{\mathcal{E}}$ with $\rho = |0\rangle\langle 0|$ and $O = \Phi(|0\rangle\langle 0|)$. Let $v \leftarrow \mathsf{QPStat}_{\mathcal{E}}(\rho, O, \tau)$. If $v \geq 1 - \epsilon - \tau$, output "$\mathcal{E} \in \mathcal{C}$". Otherwise, output "$\mathcal{E} = \Phi_{\text{dep}}$".

To prove the correctness of this algorithm in distinguishing $\mathcal{C}$ from Φ_{dep}, we will condition on the success of $\mathcal{A}$, which occurs with probability α. Now, when $\mathcal{E} \in \mathcal{C}$, let U be the unitary corresponding to $\mathcal{E}$. By the correctness of $\mathcal{A}$, we obtain a quantum channel Φ with $d_\diamond(\mathcal{E}, \Phi) \leq \epsilon$, so we proceed to the next step and make the QPStat query. We can now lower bound v as follows:

$$\begin{aligned} v &\geq \operatorname{tr}\left(OU\rho U^\dagger\right) - \tau && (69)\\ &= \operatorname{tr}\left(\Phi(|0\rangle\langle 0|)U\,|0\rangle\langle 0|\,U^\dagger\right) - \tau && (70)\\ &= F(\Phi(|0\rangle\langle 0|), U\,|0\rangle\langle 0|\,U^\dagger) - \tau && (71)\\ &\geq 1 - d_{\text{tr}}(\Phi(|0\rangle\langle 0|), U\,|0\rangle\langle 0|\,U^\dagger) - \tau && (72)\\ &\geq 1 - d_\diamond(\Phi, \mathcal{E}) - \tau && (73)\\ &\geq 1 - \epsilon - \tau, && (74) \end{aligned}$$

where the fourth line makes use of Equation (9). Thus, our distinguisher will correctly output "$\mathcal{E} \in \mathcal{C}$".

In the second case, when $\mathcal{E} = \Phi_{\text{dep}}$, the action of $\mathcal{A}$ may not be well defined or $\mathcal{A}$'s output may be far from all channels $\mathcal{C}$. In this case, our distinguisher correctly outputs "$\mathcal{E} = \Phi_{\text{dep}}$". On the other hand, if $\mathcal{A}$'s output is some channel Φ close to some channel in $\mathcal{C}$, we make the QPStat query. Then, we can upper bound v as follows:

$$\begin{aligned} v &\leq \operatorname{tr}\left(O\Phi_{\text{dep}}(\rho)\right) + \tau && (75)\\ &= \operatorname{tr}\left(\Phi(|0\rangle\langle 0|)\frac{I}{2^n}\right) + \tau && (76)\\ &= \frac{1}{2^n} + \tau. && (77) \end{aligned}$$

By assumption, $\epsilon + 2\tau < 1 - \frac{1}{2^n}$. Thus, in this case, $v < 1 - \epsilon - \tau$ and our distinguisher correctly outputs "$\mathcal{E} = \Phi_{\text{dep}}$".

As the distinguisher succeeds conditioned on the success of $\mathcal{A}$, it has an overall success probability of at least α. Further, the distinguisher makes at most $q_L + 1$ queries. This concludes the proof. □

Now, we extend the above result to an average-case learning lower bound for learning unitaries.

Lemma 6 (Average-case lower bound)**.** *Let $0 < \tau \leq \epsilon, \epsilon + 2\tau < 1 - \frac{1}{2^n}$, $\mathcal{C} \subseteq \mathcal{U}_{2^n}$ be a class of unitaries, and μ some measure over $\mathcal{C}$. Let there be a learning algorithm that, given access to QPStat_U for some $U \in \mathcal{C}$, outputs a quantum channel Φ such that $d_\diamond(\mathcal{U}, \Phi) \leq \epsilon$ with probability α over its internal randomness and probability β over the random unitary $U \sim \mu$. Then, such an algorithm must make q queries of tolerance τ to QPStat_U, where*

$$q + 1 \geq \frac{(2\alpha - 1)\beta}{\max_{\rho,O} \mathbf{Pr}_{\mathcal{E}\sim\mu}\left[\left|\operatorname{tr}\left(O\mathcal{E}(\rho)\right) - \operatorname{tr}\left(O\Phi_{\text{dep}}(\rho)\right)\right| > \tau\right]}. \tag{78}$$

Proof. Consider the subset $\mathcal{C}' \subseteq \mathcal{C}$ of measure $\mu(\mathcal{C}') = \beta$ on which the learning algorithm succeeds. Let $\tilde{\mu}$ be the measure conditioned on $\mathcal{C}'$, i.e. $\tilde{\mu}(U) = \mu(U|U \in \mathcal{C}')$. Then, for any ρ, O,

$$\begin{aligned} \mathbf{Pr}_{U\sim\tilde{\mu}}\left[\left|\operatorname{tr}\left(OU\rho U^\dagger\right) - \operatorname{tr}\left(O\Phi_{\text{dep}}(\rho)\right)\right| > \tau\right] &= \mathbf{Pr}_{U\sim\mu}\left[\left|\operatorname{tr}\left(OU\rho U^\dagger\right) - \operatorname{tr}\left(O\Phi_{\text{dep}}(\rho)\right)\right| > \tau | U \in \mathcal{C}'\right] && (79)\\ &\leq \frac{\mathbf{Pr}_{U\sim\mu}\left[\left|\operatorname{tr}\left(OU\rho U^\dagger\right) - \operatorname{tr}\left(O\Phi_{\text{dep}}(\rho)\right)\right| > \tau\right]}{\mathbf{Pr}_{U\sim\mu}\left[U \in \mathcal{C}'\right]} && (80)\\ &= \frac{\mathbf{Pr}_{U\sim\mu}\left[\left|\operatorname{tr}\left(OU\rho U^\dagger\right) - \operatorname{tr}\left(O\Phi_{\text{dep}}(\rho)\right)\right| > \tau\right]}{\beta}. && (81) \end{aligned}$$

The average-case learning algorithm with respect to μ implies a worst-case learner for the class $\mathcal{C}'$. We can then invoke Lemma 5 to lower bound the complexity for this task. The corresponding distinguishing task is C' against Φ_{dep} with respect to the measure $\tilde{\mu}$ over C', and its query complexity can be lower bounded using Lemma 4 and Equation (81). This concludes the proof. □

5.3 Proofs of Theorems 2 and 3

In this section, we detail the proofs of Theorems 2 and 3. Using Lemma 6, it suffices to bound the probability in the denominator of Equation (78) for exact and approximate unitary 2-designs. To this end, we first bound the variance of $\operatorname{tr}\left(OU\rho U^\dagger\right)$ for exact 2-designs.

Lemma 7 (Variance over 2-designs is exponentially small)**.** *Let $\rho \in \mathcal{S}_N, O \in \mathcal{M}_{N,N}, \|O\|_\infty \leq 1$ and $\mu_H^{(2)}$ be a unitary 2-design. Then,*

$$\operatorname*{\mathbf{Var}}_{U\sim\mu_H^{(2)}} \operatorname{tr}\left(OU\rho U^\dagger\right) \leq \frac{1}{N+1}. \tag{82}$$

Proof. As the variance is a second-order moment, we can equivalently bound the variance over the unitary Haar-measure μ_H. From Lemma 1, we have

$$\operatorname*{\mathbf{E}}_{U\sim\mu_H} \operatorname{tr}\left(OU\rho U^\dagger\right) = \frac{\operatorname{tr}(O)}{N}. \tag{83}$$

Again, from Lemma 1,

$$\begin{aligned}
\operatorname*{\mathbf{E}}_{U\sim\mu_H}\left[\operatorname{tr}\left(OU\rho U^\dagger\right)^2\right] &= \operatorname*{\mathbf{E}}_{U\sim\mu_H}\left[\operatorname{tr}\left(O^{\otimes 2}U^{\otimes 2}\rho^{\otimes 2}U^{\dagger\otimes 2}\right)\right] && (84)\\
&= \operatorname{tr}\left(O^{\otimes 2}\operatorname*{\mathbf{E}}_{U\sim\mu_H}[U^{\otimes 2}\rho^{\otimes 2}U^{\dagger\otimes 2}]\right) && (85)\\
&= \operatorname{tr}\left(O^{\otimes 2}(c_{\mathbb{I},\rho^{\otimes 2}}\mathbb{I} + c_{\mathbb{F},\rho^{\otimes 2}}\mathbb{F})\right) && (86)\\
&= c_{\mathbb{I},\rho^{\otimes 2}}\operatorname{tr}(O)^2 + c_{\mathbb{F},\rho^{\otimes 2}}\operatorname{tr}\left(O^2\right). && (87)
\end{aligned}$$

Now,

$$\begin{aligned}
c_{\mathbb{I},\rho^{\otimes 2}} &= \frac{N\operatorname{tr}\left(\rho^{\otimes 2}\right) - \operatorname{tr}\left(\rho^{\otimes 2}\mathbb{F}\right)}{N(N^2-1)} && (88)\\
&= \frac{N\operatorname{tr}(\rho)^2 - \operatorname{tr}\left(\rho^2\right)}{N(N^2-1)} && (89)\\
&= \frac{N - \operatorname{tr}\left(\rho^2\right)}{N(N^2-1)}, && (90)
\end{aligned}$$

and,

$$\begin{aligned}
c_{\mathbb{F},\rho^{\otimes 2}} &= \frac{N\operatorname{tr}\left(\mathbb{F}\rho^{\otimes 2}\right) - \operatorname{tr}\left(\rho^{\otimes 2}\right)}{N(N^2-1)} && (91)\\
&= \frac{N\operatorname{tr}\left(\rho^2\right) - 1}{N(N^2-1)} && (92)\\
&\leq \frac{N-1}{N(N^2-1)} && (93)\\
&= \frac{1}{N(N+1)}, && (94)
\end{aligned}$$

where the inequality follows from the fact that $\operatorname{tr}\left(\rho^2\right) \leq 1$.

Finally,

$$\underset{U\sim\mu_H}{\mathbf{Var}}\left[\mathrm{tr}\left(OU\rho U^\dagger\right)\right] = c_{\mathbb{I},\rho^{\otimes 2}}\,\mathrm{tr}(O)^2 + c_{\mathbb{F},\rho^{\otimes 2}}\,\mathrm{tr}\left(O^2\right) - \frac{\mathrm{tr}(O)^2}{N^2} \tag{95}$$

$$\leq \left(\frac{N-\mathrm{tr}\left(\rho^2\right)}{N(N^2-1)} - \frac{1}{N^2}\right)\mathrm{tr}(O)^2 + \frac{\mathrm{tr}\left(O^2\right)}{N(N+1)} \tag{96}$$

$$= \frac{1-N\,\mathrm{tr}\left(\rho^2\right)}{N^2(N^2-1)}\,\mathrm{tr}(O)^2 + \frac{Tr(O^2)}{N(N+1)} \tag{97}$$

$$\leq \frac{\mathrm{tr}\left(O^2\right)}{N(N+1)} \tag{98}$$

$$\leq \frac{1}{N+1}, \tag{99}$$

where the second inequality uses that $\mathrm{tr}\left(\rho^2\right) \geq 1/N$ and the last inequality is due to $\|O\|_\infty \leq 1 \Rightarrow \mathrm{tr}\left(O^2\right) \leq N$. □

We are now in a position to prove Theorem 2.

Proof of Theorem 2. We note that Φ_{dep} is the expected channel over the unitary Haar measure. As this is a first-order moment, Φ_{dep} is also the expected channel over unitary 2-designs. We can thus use Chebyshev's inequality to obtain,

$$\max_{\rho,O}\underset{U\sim\mu_H^{(2)}}{\mathbf{Pr}}\left(\left|\mathrm{tr}\left(OU\rho U^\dagger\right) - \left[\mathrm{tr}\left(O\Phi_{\mathrm{dep}}(\rho)\right)\right]\right| > \tau\right) \leq \max_{\rho,O}\frac{1}{\tau^2}\underset{U\sim\mu_H^{(2)}}{\mathbf{Var}}\left[\mathrm{tr}\left(OU\rho U^\dagger\right)\right]. \tag{100}$$

The desired result then follows by combining Lemmas 6 and 7. □

Before we prove Theorem 3, we show an upper bound on the variance of $\mathrm{tr}\left(OU\rho U^\dagger\right)$ over approximate unitary 2-designs.

Lemma 8 (Variance over approximate 2-designs)**.** *Let $\rho \in \mathcal{S}_N, O \in \mathcal{M}_{N,N}, \|O\|_\infty \leq 1$ and $\mu_H^{(2,\delta)}$ be a δ-approximate unitary 2-design. Then,*

$$\underset{U\sim\mu_H^{2,\delta}}{\mathbf{Var}}\left(\mathrm{tr}\left(OU\rho U^\dagger\right)\right) \leq \frac{1}{N+1} + 3\delta. \tag{101}$$

Proof. We will bound the difference in variances over the Haar measure and approximate 2-designs. For concise notation, let $f_U(\rho, O) = \mathrm{tr}\left(OU\rho U^\dagger\right)$. Then, the difference in variances can be bounded as follows.

$$\left|\underset{\mu_H^{2,\delta}}{\mathbf{Var}}\left(f_U(\rho,O)\right) - \underset{\mu_H}{\mathbf{Var}}\left(f_U(\rho,O)\right)\right| \leq \left|\mathbf{E}\mu_H^{2,\delta} f_U(\rho,O)^2 - \mathbf{E}_{\mu_H} f_U(\rho,O)^2\right| + \left|\left(\mathbf{E}_{\mu_H^{2,\delta}} f_U(\rho,O)\right)^2 - \left(\mathbf{E}_{\mu_H} f_U(\rho,O)\right)^2\right|, \tag{102}$$

where the inequality follows from the triangle inequality. We bound the first term as follows:

$$\left|\mathbf{E}_{\mu_H^{2,\delta}} f_U(\rho,O)^2 - \mathbf{E}_{\mu_H} f_U(\rho,O)^2\right| = \left|\mathrm{tr}\left(O^{\otimes 2}\left(\mathcal{M}^{(2)}_{\mu_H^{2,\delta}} - \mathcal{M}^{(2)}_{\mu_H}\right)(\rho^{\otimes 2})\right)\right| \leq \delta, \tag{103}$$

where the inequality holds since $\mu_H^{2,\delta}$ is an additive δ-approximate unitary 2-design. We now bound

the second term in Equation (102).

$$\left|(\mathbf{E}_{\mu_H^{2,\delta}} f_U(\rho,O))^2 - (\mathbf{E}_{\mu_H} f_U(\rho,O))^2\right| = |\mathbf{E}_{\mu_H^{2,\delta}} f_U(\rho,O) + \mathbf{E}_{\mu_H} f_U(\rho,O)| \cdot |\mathbf{E}_{\mu_H^{2,\delta}} f_U(\rho,O) - \mathbf{E}_{\mu_H} f_U(\rho,O)| \tag{104}$$

$$\leq 2 \cdot \left|\operatorname{tr}\left(O\left(\mathcal{M}^{(1)}_{\mu_H^{2,\delta}} - \mathcal{M}^{(1)}_{\mu_H}\right)(\rho)\right)\right| \tag{105}$$

$$\leq 2\delta. \tag{106}$$

In the first inequality, we use that $|f_U(\rho,O)| \leq 1$ when $\|O\|_\infty \leq 1$, and the last inequality holds since $\mu_H^{2,\delta}$ is an additive δ-approximate unitary 2-design. Lemma 7 and the bounds on the two terms in Equation (102) suffice to prove the lemma. □

Now, we are in a position to prove Theorem 3.

Proof of Theorem 3. Here, we have the additional assumption that the approximate 2-design is an exact 1-design. Thus, Φ_{dep} is also the expected channel over $\mu_H^{(2,\delta)}$. We can thus use Chebyshev's inequality to obtain,

$$\max_{\rho,O} \mathbf{Pr}_{U\sim\mu_H^{(2,\delta)}}\left(\left|\operatorname{tr}\left(OU\rho U^\dagger\right) - \left[\operatorname{tr}\left(O\Phi_{\text{dep}}(\rho)\right)\right]\right| > \tau\right) \leq \max_{\rho,O}\frac{1}{\tau^2}\mathbf{Var}_{U\sim\mu_H^{(2,\delta)}}\left[\operatorname{tr}\left(OU\rho U^\dagger\right)\right]. \tag{107}$$

From Lemma 8, we see that when $\delta = \mathcal{O}(1/2^n)$, the variance is $\mathcal{O}(1/2^n)$. On the other hand, when $\delta = \omega(1/2^n)$, the variance is $\mathcal{O}(\delta)$. Together with Lemma 6, we obtain the desired result. □

5.4 Proof of Theorem 4

In this section, we prove Theorem 4 on the hardness of learning Haar-random unitaries. Using Lemma 6, it again suffices to bound the probability in the denominator of Equation (78). To do this, we will use the following concentration result from [68].

Lemma 9 (Concentration of Measure for Haar-random unitaries, Corollary 17 of [68])**.** *Let $f : \mathcal{M}_{N,N} \to \mathbb{R}$ be a function from $N \times N$-dimensional matrices to the real numbers that is Lipschitz with Lipschitz constant L with respect to the Schatten-2 norm. Let μ_H be the unitary Haar-measure. For any $\tau > 0$,*

$$\mathbf{Pr}_{U\sim\mu_H}\left(f(U) - \mathop{\mathbf{E}}_{U\sim\mu_H}[f(U)] > \tau\right) \leq \exp\left(\frac{-N\tau^2}{12L^2}\right) \tag{108}$$

We are now in a position to prove Theorem 4.

Proof of Theorem 4. We will start by considering $f_{\rho,O}(U) = \operatorname{tr}\left(OU\rho U^\dagger\right)$. For any two unitaries U, V,

$$|f_{\rho,O}(U) - f_{\rho,O}(V)| = |\operatorname{tr}\left(OU\rho U^\dagger\right) - \operatorname{tr}\left(OV\rho V^\dagger\right)| \tag{109}$$

$$= \left|\operatorname{tr}\left(\rho U^\dagger OU\right) - \operatorname{tr}\left(\rho V^\dagger OV\right)\right| \tag{110}$$

$$= \left|\operatorname{tr}\left(\rho(U^\dagger OU - V^\dagger OV)\right)\right| \tag{111}$$

$$\leq \|\rho\|_2 \|U^\dagger OU - V^\dagger OV\|_2 \tag{112}$$

$$\leq \|U^\dagger OU - V^\dagger OU\|_2 + \|V^\dagger OU - V^\dagger OV\|_2 \tag{113}$$

$$\leq \|U^\dagger - V^\dagger\|_2 + \|U - V\|_2 \tag{114}$$

$$= 2\|U - V\|_2. \tag{115}$$

Here, $\|\cdot\|_2$ denotes the Schatten 2-norm. In the second and third lines we use cyclicity and linearity of the trace respectively, in the fourth line we use $\operatorname{tr}\left(A^\dagger B\right) \leq \|A\|_2\|B\|_2$ (see [69]), in the fifth line

we use the triangle inequality and the fact that $\|\rho\|_2 \leq 1$ for all states. Finally, in the second to last line, we use that $\|AB\|_2 \leq \|A\|_2 \|B\|_\infty$. Thus, the Lipschitz constant of $f_{\rho,O}$ is at most 2.

Then, from Lemma 9, and the fact that Φ_{dep} is the expected channel over μ_H, we obtain

$$\Pr_{U\sim\mu_H}\left[\operatorname{tr}\left(OU\rho U^\dagger\right) - \operatorname{tr}\left(O\Phi_{\text{dep}}(\rho)\right) > \tau\right] \leq \exp\left(\frac{-2^n\tau^2}{48}\right). \tag{116}$$

By analyzing $g_{\rho,O} = -f_{\rho,O}$, we can obtain a similar upper bound on the probability of $f_{\rho,O}(U)$ being significantly lower than its expectation. By taking a union bound, we obtain

$$\Pr_{U\sim\mu_H}\left[\left|\operatorname{tr}\left(OU\rho U^\dagger\right) - \operatorname{tr}\left(O\Phi_{\text{dep}}(\rho)\right)\right| > \tau\right] \leq 2\exp\left(\frac{-2^n\tau^2}{48}\right). \tag{117}$$

Together with Lemma 6, we obtain the desired result. □

6 Application to CR-QPUF Security

In this section, we present applications of our QPSQ model in the realm of cryptography. Specifically, we focus on a particular class of hardware security primitives used for authentication known as Classical Readout Quantum Physically Unclonable Functions (CR-QPUFs). We demonstrate an attack against CR-QPUF-based authentication protocols using Algorithm 1. As our attack has a quasi-polynomial time complexity, this does not break the formal security definition of CR-QPUFs. However, our approach highlights that any new, polynomial-time algorithm for shadow tomography of quantum processes from QPSQs would imply vulnerability for an appropriate class of CR-QPUFs.

Physical Unclonable Functions (PUFs) are hardware devices designed to resist cloning or replication, making them suitable for cryptographic tasks like authentication, identification, and fingerprinting [70–73]. Classically, PUFs have been realized using specific electrical circuits or optical materials [74–77]. However, many of these implementations remain susceptible to various attacks, such as side-channel and machine-learning attacks [78–81]. To overcome vulnerabilities due to ML-based attacks, Quantum PUFs were introduced and analyzed in [37]. However, the secure realization of Quantum PUFs shown in [37] requires implementing Haar-random unitaries, quantum communication and quantum memory, making them impractical in the near term.

As a result, a range of PUF variations necessitating different levels of quantum capability have been proposed and explored. Among these are the Classical Readout Quantum Physically Unclonable Functions (CR-QPUFs), examined in [38, 39]. CR-QPUFs represent a middle ground between fully quantum PUFs and classical PUFs, aiming to achieve the security of Quantum PUFs while relying only on classical communication and storage. However, [39] introduces a classical machine-learning attack directed at the initial proposal for CR-QPUFs and a slightly more sophisticated quantum circuit of the same type, demonstrating the vulnerability of such constructions. Nevertheless, because this construction is based on a relatively simple quantum circuit, it raises a significant open question in this domain: *Does the susceptibility of CR-QPUFs to learning attacks stem from the simplicity of the quantum process itself, or does it arise fundamentally from the way CR-QPUFs have been defined?*

We partially address this question by showing that regardless of the underlying quantum process, CR-QPUFs relying on local measurements for authentication are vulnerable to a learning attack based on Algorithm 1. First, we discuss the desired security properties and adversarial model for CR-QPUFs in Section 6.1. Then, we outline an authentication protocol for CR-QPUFs in Section 6.2, and discuss the relation between CR-QPUFs and the QPStat oracle in Section 6.3. We present our main result for this section in the form of a learning attack in Section 6.4.

6.1 Security property and attack model

As discussed in [37], the main cryptographic notion associated with PUFs in general, and particularly for the use-case of authentication and identification, is *unforgeability*. Essentially, unforgeability means that an adversary having access to an efficient-size database of input-output pairs

(also called challenge-response pairs (CRPs)) of a specific function of interest, should not be able to reproduce the output of the function on a *new* input. This database is often obtained via the adversary querying an oracle that realises the black-box access to the function. Unforgeability is also extended to the quantum world generally in two directions: it can be defined for quantum processes (instead of classical functions) [82] and also for quantum adversaries having access to quantum oracles of a classical function [33, 82, 83]. As there are many notions of unforgeability involving quantum adversaries, and the details are outside the scope of this paper, we only present the vulnerability of the protocol through the lens of learnability.

We consider two honest computationally bounded parties, a verifier $\mathcal{V}$ and a prover $\mathcal{P}$, communicating through a quantum or classical channel (depending on the challenge type). The purpose of the protocol is for the honest prover to prove their identity to the verifier with the promise that no quantum adversary can falsely identify themselves as the honest prover. A quantum adversary here is a quantum polynomial time (QPT) algorithm, that sits on the communication channel and can run arbitrary efficient quantum processes. The prover possesses a CR-QPUF device denoted as $\mathcal{C}$, which they will use for identification, and which is associated with a CPTP $\mathcal{E}$. The verifier on the other hand has a database of CRPs of $\mathcal{C}$, which is obtained by having direct access to $\mathcal{C}$ in the setup phase and recording the queries and their respective outputs. In this scenario, it is often considered that the device is later sent to the prover physically and from that point is possessed by the prover. To consider a stronger attack model, in addition to having access to the communication channel, we also assume the adversary has access to a polynomial-size database of CRPs. Here, the adversarial models are often categorized into different classes depending on the level of access assumed for the adversary to obtain such data. A *weak adversary* only has access to a randomly selected set of CRPs, often obtained by recording the communications on the channel over a certain period of time. We will consider a stronger security model, where the adversary is *adaptive* and can have oracle access to the device, i.e. issuing their desired queries to the CR-QPUFs. Although the adaptive adversarial model seems very strong, it is still realistic and often the desired security level when it comes to identification protocols and PUF-based schemes. A reason for this is that the device needs to be physically transferred at least once, which gives an adversary the chance to directly query and interact with it. In the standard adversarial model, a QPT adversary $\mathcal{A}$ can collect polynomially-many CRPs, by issuing any desired state as input, and then use this dataset to learn $\mathcal{C}$. Then, in the challenge phase of the protocol, the adversary can provide a verifiable outcome for a new challenge state, and hence break the security. However, we note that the specific attack we demonstrate has a quasi-polynomial query and time complexity for providing such an outcome.

6.2 The general structure of a CR-QPUF-based authentication protocol

We now define a general authentication protocol for CR-QPUFs and discuss how their security is defined. For the CR-QPUFs, we work with a definition similar to [39] based on quantum statistical queries. While our definition differs slightly from that of [39], it is a natural extension of secure Quantum PUFs [37] to the classical readout setting. We will discuss the differences between these frameworks in Section 6.3. For now, we denote a CR-QPUF as $\mathcal{C}$ and abstractly define it as a completely positive trace preserving (CPTP) map $\mathcal{E}$ over the n-qubit state space, able to produce statistical queries for any challenge in the challenge set, given an observable O and a tolerance parameter τ. We define the protocol between a verifier $\mathcal{V}$ and a prover $\mathcal{P}$ in Protocol 2.

We also note that for a physical device such as a CR-QPUF, the statistical query oracle QPStat abstractly models a natural and physical interaction with the device, which is querying it with the given challenge and measuring the output quantum states on a desired observable and over several copies to estimate the expectation value of the observable. In other words, the oracle is the physical device itself and not a separate entity or implementation.

In the context of these authentication protocols utilizing CR-QPUFs, we encounter two main factors governing the complexity of the underlying components. Firstly, there is the complexity of the channel representing CR-QPUF. Secondly, there is the choice of the observable. In our work, we assume the protocol is carried out using observables that are efficiently estimable, taking into account practical scenarios where both the verifier and prover can effectively measure and estimate the CR-QPUF's outcome, regardless of the underlying circuit's complexity. As such, and

Protocol 2 The general structure of the CR-QPUF-based Authentication Protocol

1. **Setup phase:**
   (a) The Verifier $\mathcal{V}$ possesses a CR-QPUF $\mathcal{C}$ associated with the quantum process $\mathcal{E}$.
   (b) The Verifier $\mathcal{V}$ and the Prover $\mathcal{P}$ agree on an observable O and a threshold τ.
   (c) $\mathcal{V}$ builds a database D of CRPs of $\mathcal{C}$ by querying $\mathsf{QPStat}_{\mathcal{E}}$ on O, with threshold τ
   For $i = 1$ to N: For challenge x_i, a quantum state ρ_i (or $\rho(x_i)$), is sampled from a selected distribution $\mathcal{D}$ and prepared as multiple copies, issued to $\mathsf{QPStat}_{\mathcal{E}}$, and the respective statistical query response y_i is recorded. Thus the database $D = \{(x_i, y_i)\}_{i=1}^{N}$ is then constructed.
   (d) $\mathcal{V}$ physically sends $\mathcal{C}$ to the Prover $\mathcal{P}$.
   At this point $\mathcal{V}$ possesses database D and $\mathcal{P}$ possesses the device $\mathcal{C}$.
2. **Authentication phase:**
   (a) $\mathcal{V}$ selects a CRP (x_i, y_i) uniformly at random from D and issues the challenge x_i^t to $\mathcal{P}$.
   if $t = c$ the challenge is classical; if $t = q$ the challenge has been sent as multiple copies of the associate quantum state ρ_i.
   (b) $\mathcal{P}$ receives the challenge x_i and proceeds as follows:
   if t=c: $\mathcal{P}$ creates $\mathsf{poly}(1/\tau)$ copies of the state $\rho_i = \rho(x_i)$.
   else if t=q: $\mathcal{P}$ proceeds to next step.
   (c) $\mathcal{P}$ obtains the output of the statistical query y_i' by issuing ρ_i to $\mathcal{C}$. This step is similar to the setup phase.
   (d) $\mathcal{P}$ sends y_i' to $\mathcal{V}$.
   (e) $\mathcal{V}$ receives y_i' and verifies it. If $|y_i - y_i'| \leq 2\tau$, the authentication is successfully passed. Otherwise, $\mathcal{V}$ aborts.

given that we aim to provide attacks with provable guarantees, we assume that the observable O selected during the setup phase is an efficient observable, i.e. we assume that O has a polynomially bounded number of terms in its Pauli representation as well as a bounded number of qubits each term acts non-trivially on. This is a physically well-motivated assumption, as demonstrated in current state-of-the-art research on estimating the expectation value of an observable [62]. In [62], the authors provided a framework which unifies a number of the most advanced and commonly studied methods, such as those in [15, 61]. While this assumption covers a wide class of observables, there are some non-local observables that can be measured efficiently using more specific techniques, such as the one in [84]. Nevertheless, under this assumption, we are able to formally demonstrate the vulnerability of a very large class of CR-QPUF authentication protocols. However, that does not imply the security of the cases that might be excluded due to our assumption on the observable, and heuristic attacks might still be applicable to scenarios in which the protocol uses a complicated and highly non-local observable.

The *correctness* or *completeness* of the protocol, which is defined as the success probability of an honest prover in the absence of any adversary or noise over the channel, should be 1. The *soundness* or *security* of the protocol, ensures that the success probability of any adversary (depending on the adversarial model) in passing the authentication should be negligible in the security parameter. For protocols defined as above, the completeness is straightforward, hence we only define and discuss the soundness.

Definition 19 (Soundness (security) of the CR-QPUF-based Authentication Protocol)**.** *We say the CR-QPUF-based authentication protocol 2 is secure if the success probability of any QPT adaptive adversary $\mathcal{A}$ in producing an output $\tilde{y}$ for any x sampled at random from a database D, over a distribution $\mathcal{D}$, that passes the verification by satisfying the condition $|y - \tilde{y}| \leq 2\tau$, is negligible in the security parameter.*

6.3 CR-QPUFs in QPSQ framework

CR-QPUF Model: While we have been discussing the QPStat oracles abstractly so far, the CR-QPUF device must, in practice, be able to take a quantum state, an observable and a tolerance

parameter τ, and output with high-probability a τ-estimate of the expectation value. The device can thus be modeled by the oracle $\mathsf{QPStat}_{\mathcal{E}}$ for the fixed underlying channel $\mathcal{E}$. On receiving a query, the device would apply $\mathcal{E}$ to multiple copies of the state, estimate the expectation value of the observable, and respond with that value as output. There are multiple methods for this estimation, such as those shown in [15, 60–62], usually requiring measurements on $\mathsf{poly}(1/\tau)$ copies of the state for statistical estimation. In a real implementation of the protocol, the required copies of the input state would either be received through a quantum communication channel or prepared by the device given a classical description as input.

Our definition of CR-QPUFs is similar to the one considered in prior work [39], where the device was also modeled by a quantum statistical query oracle. While the two models are similar in spirit, we note that there are some differences between the two definitions. The main difference is that in [39], the challenges have been defined in the form of descriptions of unitaries instead of quantum states. Starting with a fixed state initialized as $|0\rangle\langle 0|$, the input unitary U_{in} is applied on the noisy hardware, followed by repeated measurements in the computational basis. Finally, the mean of some statistical function is computed over the measurement results. The idea behind this kind of construction is that the unique noise fingerprint of the device may result in unforgeability. On the other hand, our model considers a device repeatedly implementing a fixed quantum process, that takes input states as challenges. Here, the implemented quantum process acts as a unique fingerprint of the device instead. Considering the setting of a fixed channel is also a natural extension of prior work on Quantum PUFs [37].

6.4 Vulnerability from learning results

We are now ready to present our attack on Protocol 2. To show the extreme case of our result, we can assume the quantum process $\mathcal{E}$ corresponds to a Haar-random unitary. We can also consider any arbitrary, fixed noise model on top of it to model the hardware-specific imperfections of the CR-QPUF. We use Algorithm 1 for our specific attack strategy. We construct the attack in Algorithm 3.

Algorithm 3 QPSQ attack on Protocol 2 with observable O, tolerance τ

Setting hyperparameters:
$\epsilon \leftarrow \tau$
Set N according to Corollary 1

for $i = 1$ to N **do**
  $\rho_i \sim \text{stab}_1^{\otimes n}$
  Issue challenge ρ_i to $\mathcal{C}$
  Receive response y_i
end for
$S_N \leftarrow \{(\rho_i, y_i)\}_{i=1}^N$
Learn h according to **Learning** phase of Algorithm 1

Forgery
Given challenge x from $\mathcal{V}$, respond with $h(\rho(x))$

Using Corollary 1, we present the performance guarantees of this algorithm in Theorem 5. Our result is valid for any challenge distribution $\mathcal{D}$ which satisfies the assumptions of our proposed algorithm (i.e. invariance under local Clifford operations). Two specific examples of such distributions are $\mathcal{D}_{Haar}$, consisting of Haar-random states over the n-qubit Hilbert space and $\mathcal{D}_{stab}$, uniformly random states from $\text{stab}_1^{\otimes n}$. In the first case the challenge states $\rho(x)$ are Haar-random states indexed by x and in the second case, the challenge states are in the form of $\rho(x) = \bigotimes_{i=1}^n |\psi_{x^i}\rangle\langle\psi_{x^i}|$, where $x \in \{0,1,2,3,4,5\}^n$, and we have $\{|\psi_0\rangle = |0\rangle, |\psi_1\rangle = |1\rangle, |\psi_2\rangle = |+\rangle, |\psi_3\rangle = |-\rangle, |\psi_4\rangle = |+i\rangle, |\psi_5\rangle = |-i\rangle\}$. These two specific selections of distributions for challenge states give rise to two very distinct instances of the authentication protocol. In the first case where the challenges are selected from a Haar-random distribution, the challenge state is communicated through a quantum channel, in the form of multiple copies of the state $\rho(x)$, for the prover to be able to produce the

response y_x. Intuitively we expect this to increase the security of the protocol since the adversary is unlikely to gain any information about the challenge state itself. However, this extra hiding comes with the price of generating and communicating n-qubit Haar-random states, which is often very resource-extensive. On the other hand, studying this case (especially when also considering the underlying process of $\mathcal{C}$ to be a Haar-random unitary) would be interesting because it would allow the comparison between a QPUF and a CR-QPUF with the same level of underlying resources. We note that Haar-random unitaries have been shown to satisfy the requirements of secure QPUFs [37]. This highlights the importance of the type of challenge and the verification process in the security of these hardware-based protocols. Our result is formalized in the following theorem.

Theorem 5. *Under Assumption 1 on the output of the CR-QPUF $\mathcal{C}$, for any underlying quantum process $\mathcal{E}$, any choice of observable O given as the sum of few-body observables, and under any choice of challenge state sampled from a distribution $\mathcal{D}$ invariant under single-qubit Cliffords, there exists an attack against Protocol 2 which successfully passes the verification with non-negligible probability using $\tau^2 n^{\mathcal{O}(\log(1/\tau))}$ queries and running in time $\mathsf{poly}(n, 1/\tau) n^{\mathcal{O}(\log(1/\tau))}$.*

Proof. Consider the attack described in Algorithm 3. The stated complexities can be obtained from Corollary 1 and the hyperparameter settings stated in Algorithm 3. We now focus on the correctness of the attack. As we learn within average squared error τ^2, and using Jensen's inequality, we see that on average over $\rho(x)$ drawn from $\mathcal{D}$,

$$|h(\rho(x)) - \mathrm{tr}\big(O\mathcal{E}(\rho(x))\big)| \leq \tau. \tag{118}$$

For any query x issued by $\mathcal{V}$, the associated y stored in D was received as the output of a statistical query, implying that

$$|y - \mathrm{tr}\big(O\mathcal{E}(\rho(x))\big)| \leq \tau. \tag{119}$$

By triangle inequality, the error between the adversary's prediction $h(\rho(x))$ and the stored y is bounded by 2τ. Thus, using Algorithm 3, an adversary is able to efficiently pass Protocol 2 with non-negligible probability over the challenge distribution. □

Remark 4. In Protocol 2, the parties agree on a query tolerance beforehand. As a consequence, the adversary does not have control of the query tolerance during the attack, necessitating the use of the guarantee of Corollary 1, which places no restriction on the tolerance, as opposed to Theorem 1. However, when using Corollary 1, the output of the CR-QPUF must satisfy Assumption 1, i.e., the output should be unbiased. This assumption is satisfied by standard methods for expectation value estimation that could be used to implement the oracle. However, one could add a small amount of biased noise to the output of the CR-QPUF, resulting in our attack failing, and the resulting protocol may be secure. We do not investigate this setting here and leave it for future work. Alternatively, one could also consider a stronger adversarial model, where the adversary can make queries for any tolerance of their choice, in which case the guarantee of Theorem 1 can be used and Assumption 1 would not be necessary.

Remark 5. Naively, one might anticipate that our hardness of learning results would imply a positive security result for CR-QPUFs. However, due to the nature of the verification procedure of the protocol, an adversary only needs to learn to predict approximately correct expectation values, rather than learn the underlying process up to accuracy in diamond distance. This is a significantly easier task, and as such, the hardness results cannot directly be used to prove the security of this protocol. In fact, contrary to this anticipation, we observe the opposite outcome. When the underlying process of the CR-QPUF is chosen from a hard-to-learn ensemble, such as a (approximate) unitary 2-design, then the output expectation values are highly concentrated regardless of the choice of observable and input state, enabling an easy attack with high success probability for multiple rounds. This highlights the importance of the verification method when designing such protocols.

7 Conclusion and Future Work

We have presented a physically well-motivated access model for learning quantum processes. Within this access model, we have studied two important tasks in quantum learning theory, namely

shadow tomography of a quantum process as well as learning unitaries with respect to the diamond distance. We have also demonstrated the practical relevance of this access model and our learning algorithm by partially addressing an open question regarding the security of Classical Readout Quantum PUFs.

For shadow tomography of quantum processes, Algorithm 1 succeeds for arbitrary quantum processes, but requires the observables to be local and the distribution over states to be invariant under local Clifford operations. An exciting direction for future work is to consider natural restrictions on the process instead, potentially resulting in efficient algorithms for more general states and observables.

For the case of diamond distance learning, we have shown a lower bound for learning a general class of unitaries, through a reduction from distinguishing it from the depolarizing channel. While this bound has allowed us to show the hardness of learning unitary 2-designs and the Haar measure, our bound does not hold for non-unitary channels. We believe this bound can be generalized to the non-unitary setting by defining *ancilla-assisted* QPSQs, and leave this for future work.

Furthermore, it is interesting to compare learnability in this framework to the usual quantum statistical query framework for learning states. When considering the task of learning classical functions encoded either as a unitary or as a quantum example, it is interesting to see whether the additional choice of input state available to a learner in our model can provide any advantage, i.e. *are there any separations between QPSQ and QSQ learners when looking at classical functions?*. Another compelling question in this context is whether we can show a meaningful, formal separation between QPSQ and classical learners (beyond what has already been shown through generalizing QSQs).

In terms of applications, there is much to explore in cryptography. We have demonstrated a connection between learning algorithms for shadow process tomography and attacks against CR-QPUFs. As our algorithm has a quasipolynomial query and time complexity, we are unable to conclusively show the vulnerability of CR-QPUFs. However, our approach shows that any new algorithm for this problem with a polynomial complexity would result in an efficient attack against such protocols, motivating further exploration of efficient algorithms for shadow process tomography from QPSQs. Another interesting direction of research would be to instead identify cryptographic schemes whose security can be proven from lower bounds for learning in the QPSQ access model.

Finally, it would be intriguing to observe the implementation of our learner on actual hardware or its application to data acquired from real physical experiments.

Acknowledgements

The authors thank Armando Angrisani for his valuable inputs and discussions, especially the discussions towards establishing the definition of QPSQ, and for sharing related results from [45] with us during the project. We thank Alexander Nietner for pointing out an error in a security result regarding CR-QPUFs in a previous version of this work and other valuable discussions. We also thank Elham Kashefi, Dominik Leichtle, Laura Lewis, Yao Ma, Eliott Mamon and Sean Thrasher for interesting discussions and comments at different stages of this work. We are incredibly grateful to anonymous reviewers for valuable feedback on this manuscript. 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.

References

[1] Leslie G Valiant. "A theory of the learnable". Communications of the ACM **27**, 1134–1142 (1984).

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

[3] Nader H. Bshouty and Jeffrey C. Jackson. "Learning dnf over the uniform distribution using a quantum example oracle". In Proceedings of the Eighth Annual Conference on Computational Learning Theory. Page 118–127. COLT'95. Association for Computing Machinery (1995).

[4] Srinivasan Arunachalam and Ronald De Wolf. "Optimal quantum sample complexity of learning algorithms". The Journal of Machine Learning Research **19**, 2879–2878 (2018).
[5] Srinivasan Arunachalam, Alex B. Grilo, and Henry Yuen. "Quantum statistical query learning" (2020). arxiv:2002.08240.
[6] Srinivasan Arunachalam, Vojtech Havlicek, and Louis Schatzki. "On the role of entanglement and statistics in learning". Advances in Neural Information Processing Systems **36**, 55064–55076 (2024).
[7] 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". Physical Review Letters **130**, 240602 (2023).
[8] 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" (2023). arxiv:2305.05765.
[9] Alexander Nietner. "Unifying (quantum) statistical and parametrized (quantum) algorithms" (2023). arxiv:2310.17716.
[10] Alp Atici and Rocco A Servedio. "Improved bounds on quantum learning algorithms". Quantum Information Processing **4**, 355–386 (2005).
[11] Alex B. Grilo, Iordanis Kerenidis, and Timo Zijlstra. "Learning-with-errors problem is easy with quantum samples". Physical Review A **99**, 032314 (2019).
[12] 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**, 1646–1660 (2009).
[13] 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).
[14] Qian Xu and Shuqi Xu. "Neural network state estimation for full quantum state tomography" (2018). arxiv:1811.06654.
[15] Hsin-Yuan Huang, Richard Kueng, and John Preskill. "Predicting many properties of a quantum system from very few measurements". Nature Physics **16**, 1050–1057 (2020).
[16] Scott Aaronson. "Shadow tomography of quantum states". In Proceedings of the 50th annual ACM SIGACT symposium on theory of computing. Pages 325–338. (2018).
[17] Andrew M Childs, Tongyang Li, Jin-Peng Liu, Chunhao Wang, and Ruizhe Zhang. "Quantum algorithms for sampling log-concave distributions and estimating normalizing constants". Advances in Neural Information Processing Systems **35**, 23205–23217 (2022).
[18] Ashley Montanaro. "Learning stabilizer states by bell sampling" (2017). arxiv:1707.04012.
[19] Masoud Mohseni, Ali T Rezakhani, and Daniel A Lidar. "Quantum-process tomography: Resource analysis of different strategies". Physical Review A **77**, 032322 (2008).
[20] Kai-Min Chung and Han-Hsuan Lin. "Sample efficient algorithms for learning quantum channels in pac model and the approximate state discrimination problem". In 16th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2021). Pages 3:1–3:22. Schloss Dagstuhl–Leibniz-Zentrum für Informatik (2021).
[21] 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).
[22] Srinivasan Arunachalam and Ronald de Wolf. "Guest column: A survey of quantum learning theory". ACM Sigact News **48**, 41–67 (2017).
[23] Nathan Wiebe, Christopher Granade, Christopher Ferrie, and David G Cory. "Hamiltonian learning and certification using quantum resources". Physical review letters **112**, 190501 (2014).
[24] Harper R Grimsley, Sophia E Economou, Edwin Barnes, and Nicholas J Mayhall. "An adaptive variational algorithm for exact molecular simulations on a quantum computer". Nature communications **10**, 3007 (2019).
[25] Andrew James Scott. "Optimizing quantum process tomography with unitary 2-designs". Journal of Physics A: Mathematical and Theoretical **41**, 055308 (2008).
[26] Ryan Levy, Di Luo, and Bryan K Clark. "Classical shadows for quantum process tomography on near-term quantum computers". Physical Review Research **6**, 013029 (2024).

[27] Hsin-Yuan Huang, Steven T Flammia, and John Preskill. "Foundations for learning from noisy quantum experiments" (2022). arxiv:2204.13691.
[28] 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**, 14485 (2017).
[29] Robin Harper, Steven T Flammia, and Joel J Wallman. "Efficient learning of quantum noise". Nature Physics **16**, 1184–1188 (2020).
[30] Armands Strikis, Dayue Qin, Yanzhu Chen, Simon C Benjamin, and Ying Li. "Learning-based quantum error mitigation". PRX Quantum **2**, 040330 (2021).
[31] Yihui Quek, Daniel Stilck França, Sumeet Khatri, Johannes Jakob Meyer, and Jens Eisert. "Exponentially tighter bounds on limitations of quantum error mitigation". Nature Physics **20**, 1648–1658 (2024).
[32] Maria Schuld, Ilya Sinayskiy, and Francesco Petruccione. "An introduction to quantum machine learning". Contemporary Physics **56**, 172–185 (2015).
[33] Dan Boneh and Mark Zhandry. "Quantum-secure message authentication codes". In Thomas Johansson and Phong Q. Nguyen, editors, Advances in Cryptology – EUROCRYPT 2013. Pages 592–608. Berlin, Heidelberg (2013). Springer Berlin Heidelberg.
[34] Marc Kaplan, Gaëtan Leurent, Anthony Leverrier, and María Naya-Plasencia. "Breaking symmetric cryptosystems using quantum period finding". In Matthew Robshaw and Jonathan Katz, editors, Advances in Cryptology – CRYPTO 2016. Pages 207–237. Berlin, Heidelberg (2016). Springer Berlin Heidelberg.
[35] Thomas Santoli and Christian Schaffner. "Using simon's algorithm to attack symmetric-key cryptographic primitives". Quantum Information & Computation **17**, 65–78 (2017).
[36] Céline Chevalier, Ehsan Ebrahimi, and Quoc-Huy Vu. "On security notions for encryption in a quantum world". In International Conference on Cryptology in India. Pages 592–613. Springer (2022).
[37] Myrto Arapinis, Mahshid Delavar, Mina Doosti, and Elham Kashefi. "Quantum physical unclonable functions: Possibilities and impossibilities". Quantum **5**, 475 (2021).
[38] Koustubh Phalak, Abdullah Ash-Saki, Mahabubul Alam, Rasit Onur Topaloglu, and Swaroop Ghosh. "Quantum PUF for security and trust in quantum computing". IEEE Journal on Emerging and Selected Topics in Circuits and Systems **11**, 333–342 (2021).
[39] Niklas Pirnay, Anna Pappa, and Jean-Pierre Seifert. "Learning classical readout quantum PUFs based on single-qubit gates". Quantum Machine Intelligence **4**, 14 (2022).
[40] Hsin-Yuan Huang, Sitan Chen, and John Preskill. "Learning to predict arbitrary quantum processes". PRX Quantum **4**, 040337 (2023).
[41] Ashley Montanaro and Tobias J Osborne. "Quantum boolean functions" (2008). arxiv:0810.2435.
[42] Marco Fanizza, Yihui Quek, and Matteo Rosati. "Learning quantum processes without input control". PRX Quantum **5**, 020367 (2024).
[43] Matthias C Caro, Marcel Hinsche, Marios Ioannou, Alexander Nietner, and Ryan Sweke. "Classical verification of quantum learning". In 15th Innovations in Theoretical Computer Science Conference (ITCS 2024). Pages 24:1–24:23. Schloss Dagstuhl–Leibniz-Zentrum für Informatik (2024).
[44] Vitaly Feldman. "A general characterization of the statistical query complexity". In Conference on learning theory. Pages 785–830. PMLR (2017).
[45] Armando Angrisani. "Learning unitaries with quantum statistical queries" (2023). arxiv:2310.02254.
[46] Sitan Chen, Jordan Cotler, Hsin-Yuan Huang, and Jerry Li. "Exponential separations between learning with and without quantum memory". In 2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS). Pages 574–585. IEEE (2022).
[47] Jonathan Kunjummen, Minh C Tran, Daniel Carney, and Jacob M Taylor. "Shadow process tomography of quantum channels". Physical Review A **107**, 042403 (2023).
[48] Matthias C Caro. "Learning quantum processes and hamiltonians via the pauli transfer matrix". ACM Transactions on Quantum Computing **5**, 1–53 (2024).

[49] 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).
[50] Michael A Nielsen and Isaac L Chuang. "Quantum computation and quantum information". Cambridge university press. (2010).
[51] Antonio Anna Mele. "Introduction to haar measure tools in quantum information: A beginner's tutorial". Quantum **8**, 1340 (2024).
[52] Andreas Elben, Steven T. Flammia, Hsin-Yuan Huang, Richard Kueng, John Preskill, Benoît Vermersch, and Peter Zoller. "The randomized measurement toolbox". Nature Reviews Physics **5**, 9–24 (2022).
[53] Hsin-Yuan Huang, Richard Kueng, Giacomo Torlai, Victor V Albert, and John Preskill. "Provably efficient machine learning for quantum many-body problems". Science **377**, eabk3333 (2022).
[54] V Kanade, A Rocchetto, and S Severini. "Learning dnfs under product distributions via μ-biased quantum fourier sampling". Quantum Information and Computation **19**, 1261–1278 (2019).
[55] Matthias C. Caro. "Quantum learning boolean linear functions w.r.t. product distributions". Quantum Information Processing **19**, 172 (2020).
[56] Yuxuan Du, Min-Hsiu Hsieh, Tongliang Liu, Dacheng Tao, and Nana Liu. "Quantum noise protects quantum classifiers against adversaries". Physical Review Research **3**, 023153 (2021).
[57] Aravind Gollakota and Daniel Liang. "On the hardness of pac-learning stabilizer states with noise". Quantum **6**, 640 (2022).
[58] Changjun Kim, Kyungdeock Daniel Park, and June-Koo Rhee. "Quantum error mitigation with artificial neural network". IEEE Access **8**, 188853–188860 (2020).
[59] Hsin-Yuan Huang, Richard Kueng, and John Preskill. "Information-theoretic bounds on quantum advantage in machine learning". Physical Review Letters **126**, 190505 (2021).
[60] Charles Hadfield, Sergey Bravyi, Rudy Raymond, and Antonio Mezzacapo. "Measurements of quantum hamiltonians with locally-biased classical shadows". Communications in Mathematical Physics **391**, 951–967 (2022).
[61] Hsin-Yuan Huang, Richard Kueng, and John Preskill. "Efficient estimation of pauli observables by derandomization". Physical Review Letters **127**, 030503 (2021).
[62] Bujiao Wu, Jinzhao Sun, Qi Huang, and Xiao Yuan. "Overlapped grouping measurement: A unified framework for measuring quantum states". Quantum **7**, 896 (2023).
[63] Zak Webb. "The clifford group forms a unitary 3-design" (2015). arxiv:1510.02769.
[64] Huangjun Zhu. "Multiqubit clifford groups are unitary 3-designs". Physical Review A **96**, 062336 (2017).
[65] Richard A Low. "Learning and testing algorithms for the clifford group". Physical Review A—Atomic, Molecular, and Optical Physics **80**, 052314 (2009).
[66] Ching-Yi Lai and Hao-Chung Cheng. "Learning quantum circuits of some t gates". IEEE Transactions on Information Theory **68**, 3951–3964 (2022).
[67] Jonas Haferkamp and Nicholas Hunter-Jones. "Improved spectral gaps for random quantum circuits: Large local dimensions and all-to-all interactions". Physical Review A **104**, 022417 (2021).
[68] Elizabeth Meckes and Mark Meckes. "Spectral measures of powers of random matrices". Electronic Communications in Probability **18**, 1 – 13 (2013).
[69] Bernhard Baumgartner. "An inequality for the trace of matrix products, using absolute values" (2011). arxiv:1106.6189.
[70] Ulrich Rührmair and Daniel E Holcomb. "Pufs at a glance". In the conference on Design, Automation & Test in Europe. Page 347. European Design and Automation Association (2014).
[71] Chip-Hong Chang, Yue Zheng, and Le Zhang. "A retrospective and a look forward: Fifteen years of physical unclonable function advancement". IEEE Circuits and Systems Magazine **17**, 32–62 (2017).
[72] Mina Doosti, Niraj Kumar, Mahshid Delavar, and Elham Kashefi. "Client-server identification protocols with quantum PUF". ACM Transactions on Quantum Computing **2**, 1–40 (2021).

[73] Mahshid Delavar, Sattar Mirzakuchaki, Mohammad Hassan Ameri, and Javad Mohajeri. "Puf-based solutions for secure communications in advanced metering infrastructure (ami)". International Journal of Communication Systems **30**, e3195 (2017).
[74] Ravikanth Pappu, Ben Recht, Jason Taylor, and Neil Gershenfeld. "Physical one-way functions". Science **297**, 2026–2030 (2002).
[75] Jorge Guajardo, Sandeep S Kumar, Geert-Jan Schrijen, and Pim Tuyls. "Fpga intrinsic pufs and their use for ip protection". In International workshop on cryptographic hardware and embedded systems. Pages 63–80. Springer (2007).
[76] Blaise Gassend, Dwaine Clarke, Marten Van Dijk, and Srinivas Devadas. "Silicon physical random functions". In 9th ACM conference on Computer and communications security. Pages 148–160. ACM (2002).
[77] G Edward Suh and Srinivas Devadas. "Physical unclonable functions for device authentication and secret key generation". In 44th ACM/IEEE Design Automation Conference. Pages 9–14. IEEE (2007).
[78] Ulrich Rührmair, Frank Sehnke, Jan Sölter, Gideon Dror, Srinivas Devadas, and Jürgen Schmidhuber. "Modeling attacks on physical unclonable functions". In Proceedings of the 17th ACM conference on Computer and communications security. Pages 237–249. (2010).
[79] Fatemeh Ganji, Shahin Tajik, Fabian Fäßler, and Jean-Pierre Seifert. "Strong machine learning attack against PUFs with no mathematical model". In International Conference on Cryptographic Hardware and Embedded Systems. Pages 391–411. Springer (2016).
[80] Lars Tebelmann, Michael Pehl, and Vincent Immler. "Side-channel analysis of the tero PUF". In International Workshop on Constructive Side-Channel Analysis and Secure Design. Pages 43–60. Springer (2019).
[81] Mahmoud Khalafalla and Catherine Gebotys. "PUFs deep attacks: Enhanced modeling attacks using deep learning techniques to break the security of double arbiter PUFs". In 2019 Design, Automation & Test in Europe Conference & Exhibition (DATE). Pages 204–209. IEEE (2019).
[82] Mina Doosti, Mahshid Delavar, Elham Kashefi, and Myrto Arapinis. "A unified framework for quantum unforgeability" (2021). arxiv:2103.13994.
[83] Gorjan Alagic, Tommaso Gagliardoni, and Christian Majenz. "Unforgeable quantum encryption". In Jesper Buus Nielsen and Vincent Rijmen, editors, Advances in Cryptology – EUROCRYPT 2018. Pages 489–519. Cham (2018). Springer International Publishing.
[84] Ophelia Crawford, Barnaby van Straaten, Daochen Wang, Thomas Parks, Earl Campbell, and Stephen Brierley. "Efficient quantum measurement of pauli operators in the presence of finite sampling error". Quantum **5**, 385 (2021).