arXiv:2302.07409v5 [cs.LG] 21 Jul 2025

Quantum Learning Theory Beyond Batch Binary Classification

Preetham Mohan[1] and Ambuj Tewari[2,3]

[1]Department of Mathematics, University of Michigan
[2]Department of Statistics, University of Michigan
[3]Department of Electrical Engineering and Computer Science, University of Michigan

Arunachalam and de Wolf [1] showed that the sample complexity of quantum batch learning of boolean functions, in the realizable and agnostic settings, has the *same form and order* as the corresponding classical sample complexities. In this paper, we extend this, ostensibly surprising, message to batch multiclass learning, online boolean learning, and online multiclass learning. For our online learning results, we first consider an adaptive adversary variant of the classical model of Dawid and Tewari [2]. Then, we introduce the first (to the best of our knowledge) model of online learning with quantum examples.

1 Introduction

Bshouty and Jackson [3] provided a quantum extension of the PAC setting by formalizing what it means to learn from quantum examples. This inspired a line of work [4–6], culminating in the work of Arunachalam and de Wolf [1], that has provided *sample complexity* bounds for quantum batch learning of boolean functions. Arunachalam and de Wolf [1] helped crystallize the following message:

1. No new combinatorial dimension is needed to characterize quantum batch learnability of boolean functions–the VC dimension continues to do so.
2. There is *at most a constant* sample complexity advantage for quantum batch learning of boolean functions, in both the realizable and agnostic settings, as compared to the corresponding classical sample complexities.

In this paper, we show that this message continues to hold in three other learning settings: batch learning of multiclass functions, online learning of boolean functions, and online learning of multiclass functions.

Our motivation for considering quantum batch learning of multiclass functions is an open question posed in [1] which asks "what is the quantum sample complexity for learning concepts whose range is $[k]$ rather than $\{0,1\}$, for some $k > 2$?" We resolve this question for $2 < k < \infty$ (see Section 3). In classical multiclass batch learning, an approach to establish the lower and

Preetham Mohan: preetham@umich.edu
Ambuj Tewari: tewaria@umich.edu

upper sample complexity bounds [7] is to proceed via a reduction to the binary case, with an appeal to the definition of Natarajan dimension. While classically straightforward, extending such a proof approach to establish sample complexity bounds for quantum multiclass batch learning involves manipulating quantum examples, which has to be done with utmost care (see Section 3.2.1).

Unlike the batch setting, quantum online learning of classical functions, to the best of our knowledge, has no predefined model. One possible explanation is that we need, as an intermediary, a new classical online learning model (see Sections 4.2, 4.3) where, at each round, the adversary provides a *distribution* over the example (input-label) space instead of a single example. With this new classical model, and the definition of a quantum example, a model for online learning in the quantum setting arises as a natural extension (see Figure 1).

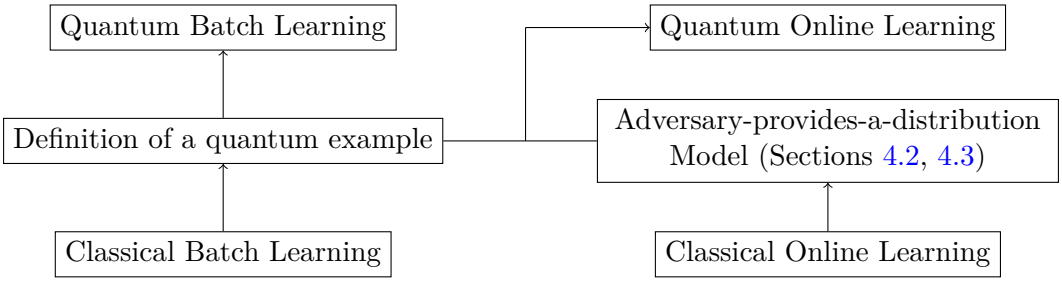

Figure 1: Mapping of the tools necessary for generalizations of learning paradigms from classical to quantum.

1.1 Our Contributions

In Tables 1 and 2, we provide a concise overview of existing results and highlight our contributions in batch and online learning for binary and multiclass classification across realizable and agnostic settings in both classical and quantum paradigms. Our contributions include:

- establishing lower and upper sample complexity bounds for quantum batch multiclass classification in the realizable and agnostic settings,
- proposing a new classical online learning model, which is an adaptive adversary variant of an existing classical online learning model [2],
- proposing a quantum online learning model, as a natural generalization of our proposed classical online learning model,
- establishing tight expected regret bounds for quantum online binary classification in the realizable and agnostic settings,
- establishing a tight expected regret bound for quantum online multiclass classification in the realizable setting, and
- establishing lower and upper expected regret bounds for quantum online multiclass classification in the agnostic setting.

Notes on Tables 1 and 2

- In all cases, we state known results that exhibit the tightest dependence on the combinatorial parameters that characterize learning in the respective settings.

- In the batch multiclass case, we work with Natarajan dimension, instead of DS dimension which was shown to characterize classical batch multiclass learning (including the $k \to \infty$ case) recently [8]. We defer the resolution of the quantum sample complexity in the $k \to \infty$ case to future work.
- In the batch multiclass realizable case, there exists an upper bound with a tighter dependence on ϵ (but looser on $\mathrm{Ndim}(\mathcal{H})$) for both classical and quantum cases (see Section 3.2.2).
- In the (canonical) classical online multiclass agnostic case, the $\tilde{\mathcal{O}}(\sqrt{\mathrm{mcLdim}(\mathcal{H})T})$ bound hides a $\sqrt{\log\left(\frac{T}{\mathrm{mcLdim}(\mathcal{H})}\right)}$ factor [9].
- The definition of loss/regret differs between the canonical classical adversary-provides-an-input model (in Section 4.1) and both the classical adversary-provides-a-distribution model (in Sections 4.2, 4.3) and the quantum online model (in Section 5). Specifically, the former employs the indicator loss (mistake model), whereas the latter two involve probabilistic losses.

		Classical	Quantum
Boolean	Realizable	$\Theta\left(\frac{d+\log(\frac{1}{\delta})}{\epsilon}\right)$ [10, 11]	$\Theta\left(\frac{d+\log(\frac{1}{\delta})}{\epsilon}\right)$ [1]
	Agnostic	$\Theta\left(\frac{d+\log(\frac{1}{\delta})}{\epsilon^2}\right)$ [12, 13]	$\Theta\left(\frac{d+\log(\frac{1}{\delta})}{\epsilon^2}\right)$ [1]
Multiclass	Realizable	$\Omega\left(\frac{d+\log(\frac{1}{\delta})}{\epsilon}\right)$ [14] $\mathcal{O}\left(\frac{d\log(k)\log(\frac{1}{\epsilon})+\log(\frac{1}{\delta})}{\epsilon}\right)$ [7]	$\Omega\left(\frac{d+\log(\frac{1}{\delta})}{\epsilon}\right)$ (Thm. 3.3) $\mathcal{O}\left(\frac{d\log(k)\log(\frac{1}{\epsilon})+\log(\frac{1}{\delta})}{\epsilon}\right)$ (Thm. 3.8)
	Agnostic	$\Omega\left(\frac{d+\log(\frac{1}{\delta})}{\epsilon^2}\right)$ [15] $\mathcal{O}\left(\frac{d\log(k)+\log(\frac{1}{\delta})}{\epsilon^2}\right)$ [15]	$\Omega\left(\frac{d+\log(\frac{1}{\delta})}{\epsilon^2}\right)$ (Thm. 3.3) $\mathcal{O}\left(\frac{d\log(k)+\log(\frac{1}{\delta})}{\epsilon^2}\right)$ (Thm. 3.8)

Table 1: An overview of sample complexity results for **batch** learning in classical and quantum paradigms. Our novel contributions are presented in boxes shaded gray. d denotes the VC dimension for Boolean cases ($d := \mathrm{VCdim}(\mathcal{H})$), and the Natarajan dimension for Multiclass cases ($d := \mathrm{Ndim}(\mathcal{H})$).

		Classical (Input-based)	Classical (Distribution-based)	Quantum
Bool.	Real.	$\Theta(d)$ [16]	$\Theta(d)$ (Thms. 4.6, 4.7)	$\Theta(d)$ (Thms. 5.4, 5.5)
	Agn.	$\Theta(\sqrt{dT})$ [17, 18]	$\Theta(\sqrt{dT})$ (Thms. 4.10, 4.11)	$\Theta(\sqrt{dT})$ (Thms. 5.4, 5.5)
Multi.	Real.	$\Theta(d)$ [7]	$\Theta(d)$ (Thms. 4.13, 4.14)	$\Theta(d)$ (Thms. 5.6, 5.7)
	Agn.	$\Omega(\sqrt{dT})$ [7] $\tilde{\mathcal{O}}(\sqrt{dT})$ [9]	$\Omega(\sqrt{dT})$ (Thm. 4.20) $\mathcal{O}(\sqrt{dT\log(Tk)})$ (Thm. 4.19)	$\Omega(\sqrt{dT})$ (Thm. 5.6) $\mathcal{O}(\sqrt{dT\log(Tk)})$ (Thm. 5.7)

Table 2: An overview of expected regret bounds for **online** learning in the canonical classical (adversary-provides-an-input), classical *adversary-provides-a-distribution*, and quantum paradigms. Our novel contributions are presented in boxes shaded gray. d denotes the Littlestone dimension for Boolean cases ($d := \mathrm{Ldim}(\mathcal{H})$), and the multiclass Littlestone dimension for Multiclass cases ($d := \mathrm{mcLdim}(\mathcal{H})$).

1.2 Related Works

Situating Our Work in the Larger Context. In the survey on the complexity of learning quantum states [19], the authors identify four key avenues of current research organized into a 2×2 framework: (1) learning **all states** versus **some subclasses of states**, and (2) learning with **strong requirements** (e.g., up to small trace distance) versus **weaker requirements** (e.g., PAC learning, statistical query learning). One of these avenues, concerning "learning classical functions encoded as [quantum] states," aligns with the *learning some subclasses of states with weaker learning requirements* category, and it is here that the models introduced in this paper contribute to the ongoing exploration of this research landscape.

Batch Learning. Quantum learning of classical functions in the batch setting has has predominantly focused on learning Boolean functions (see [20] for a survey). In our work, the multiclass extension in the batch setting (Section 3.2) is inspired by an open question of [1]. The lower bound (Section 3.2.1) in this extension is made possible due to the presence of an explicit quantum circuit (see Figure 2) that allows for a *black box*-style reduction of the quantum multiclass case to the quantum binary classification case, enabling the existing quantum batch binary lower bound (and, techniques therewith, see Section 3.1 for a detailed treatment) in [1] to yield the desired multiclass bound. This reduction is analogous to the establishment of the lower bound in the classical batch multiclass setting (as presented in Theorem 5 of [7]). While the batch classification model in this paper assumes the learner has access to copies of quantum states $|\psi_c\rangle$ (i.e., an *example oracle*), with no (i.e., at most a constant) advantage over classical batch models, Grover's algorithm – along with its quadratic quantum sample complexity advantage – has been successfully adapted in [21] to batch *binary* classification in a stronger learning model where the learner has access to a quantum circuit Q_c that generates a quantum sample $|\psi_c\rangle$ (i.e., a *unitary oracle*[1]).

Proper vs. Improper Learning. In this work, we do not explicitly address whether learning is proper or improper in the various settings. However, since most of our proofs involve reductions to the corresponding classical settings, the nature of the algorithms that provide tight bounds in those settings remains unchanged in our cases. The separation between proper and improper quantum learning has garnered significant interest in the community [22, 23]. In particular, this separation has been studied in the context of quantum extensions of the coupon collector problem, which provides a classical example of such a separation. Notably, [23] established a provable separation between proper and improper learning in the quantum setting using a "padded" variant of the quantum coupon collector problem. On the classical side, [24] provides a full characterization of when proper learning achieves optimal sample complexity, based on a combinatorial parameter known as the dual Helly number. To the best of our knowledge, an analogous characterization for quantum proper learning remains an open question.

Online Learning and Adversary-provides-a-distribution. The online *adversary-provides-a-distribution* model (Section 4.2) has analogues in other works. For example, [25] and [26]

[1]The sample complexity here refers to the number of calls to the unitary oracle, providing access to both Q_c and its Hermitian conjugate $Q_c^\dagger$

consider a setting where an adaptive adversary provides a distribution. Their setup is designed for smoothed learning, where at each time step, the adversary provides a sample drawn from their chosen distribution. However, their framework assumes a full-information setting, where the learner observes all losses, whereas in our model, the learner only receives partial feedback. Another related connection can be found in the online game introduced in [27] to analyze batch generalization error. While their setting, like ours, involves partial information on losses (and, therefore, regret), the probabilistic aspect of their model arises from the learner selecting a distribution over a hypothesis class, rather than the adversary providing one.

1.3 Organization

The paper is organized as follows. In Section 2, we present preliminaries, including notation, quantum and probability theory basics, and batch learning frameworks. Section 3 contains one of our main results, addressing the quantum sample complexity in the batch multiclass setting. Moving on to Section 4, we revisit the canonical online model (Section 4.1), followed by the introduction and presentation of results for the classical adversary-provides-a-distribution model in both realizable (Section 4.2) and agnostic (Section 4.3) settings for both binary and multiclass (Section 4.4) classification. This model serves as an intermediary for transitioning from the canonical classical online model to the quantum online model. In Section 5, we introduce the quantum online learning model and summarize results within the established framework. The paper concludes with a discussion and reflection on the obtained results in Sections 5.4 and 6, along with some open questions for future exploration.

2 Preliminaries

In this section, we introduce the key concepts and notation used throughout the paper. First, we briefly review the fundamentals of quantum computing. Next, we present essential probability theory concepts needed for the proofs in Section 4. Finally, we outline the batch learning frameworks in both classical and quantum settings, setting the stage for Section 3.

2.1 Notation

In the bra-ket (Dirac) notation, a ket, $|x\rangle$, denotes a column vector in a complex vector space with an inner product (i.e., a Hilbert space). It is used primarily in the context of describing the state of a quantum system (e.g., see Definition 2.1). A bra $\langle\cdot|$ is the dual of the ket, in that $\langle x| = |x\rangle^\dagger$, where the $\dagger$ operator denotes the conjugate transpose. Typically, the bra notation is used for operators $\langle M|$ (e.g. measurement operators) acting on a ket. This notation lends itself naturally to the notion of inner product $\langle x|x\rangle = \|x\|^2$, and matrix-vector multiplication $\langle M|x\rangle$. Furthermore, note that $|x, y\rangle$ denotes the tensor product $|x\rangle \otimes |y\rangle$, where $\otimes$ denotes the standard tensor product of two vector spaces. The comma may be omitted, and we have numerous equivalent notations for the tensor product: e.g., $|0^2\rangle = |00\rangle = |0, 0\rangle = |0\rangle\,|0\rangle = |0\rangle \otimes |0\rangle = |0\rangle^{\otimes 2}$.

2.2 Quantum Basics

Analogous to how a classical bit (*bit*) is a unit of classical information, a quantum bit (*qubit*) is a unit of quantum information. The difference between the two is best illustrated by

considering how each is realized. A bit is realized via well separated values of a physical property of a system (e.g., voltage across an element in an electric circuit). If the value is higher than a certain threshold, the bit assumes the value 1. Otherwise, it assumes the value 0. Thus, a bit carries the information equivalent of its namesake, a binary digit. A qubit, on the other hand, is realized as a two-level quantum system; e.g., as the spin (up, down) of an electron, the polarization (horizontal, vertical) of a photon, or the discrete energy levels (ground, excited) of an ion. Consequently, it is governed by the postulates of quantum mechanics [28], as detailed in the following paragraphs.

The *state space* of a qubit is a 2-dimensional complex vector space, denoted as $\mathbb{C}^2$. The definition of a qubit as a *state vector* within this space is presented in the following definition.

Definition 2.1 (Qubit)**.** *A single (isolated) qubit is described by a state vector* $|\psi\rangle$*, which is a unit vector in the state space* $\mathbb{C}^2$*. Mathematically,*

$$|\psi\rangle = \alpha_0 |0\rangle + \alpha_1 |1\rangle , \quad \alpha_0, \alpha_1 \in \mathbb{C}, \quad |\alpha_0|^2 + |\alpha_1|^2 = 1,$$

where $|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ *and* $|1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ *are basis vectors for the state space.*

So, although we have two basis states (much as we did for the classical bit), the qubit is allowed to be in a (complex) superposition of the two, whereas a classical bit must deterministically be in one of the basis states. Additionally, as our learning examples (refer to (1), (2)) will involve multiple qubits, it is important to note that the state space of the composite system, comprising many qubits, is the tensor product of the state spaces of its components (i.e., the individual qubits). The joint state of the composite system formed by n qubits, each in state $|\psi_i\rangle$, $i \in \{1, \ldots, n\}$, is given by, $|\psi_1\rangle \otimes |\psi_2\rangle \otimes \cdots \otimes |\psi_n\rangle \in \mathbb{C}^{2^n}$. Now, let us express the state vector for multi-qubit system in terms of the standard basis elements.

Definition 2.2 (Multi-qubit system)**.** *The state vector,* $|\Psi\rangle \in \mathbb{C}^{2^n}$*, describing the system of* n *qubits can be expressed in terms of the standard basis elements* $\{|b\rangle = |b_1\rangle \otimes \cdots \otimes |b_n\rangle \,|\, b = (b_1, \ldots, b_n) \in \{0,1\}^n\}$ *as follows:*

$$|\Psi\rangle = \sum_{b \in \{0,1\}^n} \alpha_b |b\rangle ,$$

where $\alpha_b = \langle \Psi | b \rangle \in \mathbb{C}$*,* $\sum_b |\alpha_b|^2 = 1$*.*

In contrast, the joint state of n classical bits is described by their Cartesian product. This essential distinction between Cartesian and tensor products is precisely the phenomenon of quantum entanglement, namely the existence of (pure) states of a composite system that are not product states of its parts. Quantum entanglement, alongside superposition, lies at the heart of intrinsic advantages of quantum computing.

Any manipulation of a quantum system is confined to unitary evolution. In the context of computation, this implies that *all quantum gates are unitary operators*, restricting their application to reversible computations. An avenue for irreversible computation, and the only way to obtain classical outputs in the quantum realm, is the notion of a *measurement*.

Definition 2.3 (Measurement)**.** *Quantum measurements are described by a collection* $\{M_m\}$ *of measurement operators acting on the state space of the system. The index* m *denotes the*

possible classical outcomes of the measurement. If the quantum system is in the state $|\psi\rangle$ before measurement, then the probability that result m occurs is given by $p(m) = \langle\psi| M_m^\dagger M_m |\psi\rangle$, and the state of the system after the measurement, if $m^\star$ is observed, "collapses to" $M_{m^\star}|\psi\rangle / \sqrt{p(m^\star)}$. To ensure the conservation of total probability, $\sum_m M_m^\dagger M_m = I$ is satisfied.

Here are a couple of examples to illustrate the above definition:

- A measurement in the standard basis is implemented by measurement operators $M_0 = |0\rangle\langle 0|$ and $M_1 = |1\rangle\langle 1|$.
- A measurement in the standard basis of the state $|\Psi\rangle = \sum_{b\in\{0,1\}^n} \alpha_b |b\rangle$ yields the classical outcome b with probability $|\alpha_b|^2$.

2.3 Probability Theory Basics

Let $\mathbf{X} := (X_t)_{t\geq 1}$ be a sequence of random variables defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where each X_t represents the outcome of an experiment at time t. Associated with this sequence is a filtration $\mathbb{F} := (\mathcal{F}_t)_{t\geq 1}$, an increasing family of σ-algebras such that $\mathcal{F}_t$ captures the information available up to time t. Formally, $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \cdots \subseteq \mathbb{F}$, where $\mathcal{F}_t$ contains events determined by $\mathbf{X}|_t = (X_1, X_2, \ldots, X_t)$.

Definition 2.4 (Martingale)**.** *A sequence of random variables $\mathbf{M} := (M_t)_{t\geq 1}$ is said to be a martingale with respect to the filtration $\mathbb{F} := (\mathcal{F}_t)_{t\geq 1}$ if it satisfies the following conditions:*

1. *Adaptation: M_t is $\mathcal{F}_t$-measurable for all $t \geq 1$ (i.e., we say $\mathbf{M}$ is adapted to $\mathbb{F}$),*
2. *Integrability: $\mathbb{E}[|M_t|] < \infty$ for all $t \geq 1$,*
3. *Martingale Property: $\mathbb{E}[M_{t+1}|\mathcal{F}_t] = M_t$ for all $t \geq 1$.*

Definition 2.5 (Martingale Difference Sequence)**.** *Given a martingale $\mathbf{M} := \{M_t\}_{t\geq 1}$, the corresponding martingale difference sequence is defined as $\mathbf{D} := \{D_{t+1}\}_{t\geq 1}$, where $D_{t+1} = M_{t+1} - M_t$. By the martingale property, the sequence $\mathbf{D}$ satisfies $\mathbb{E}[D_{t+1}|\mathcal{F}_t] = 0$ for all $t \geq 1$, indicating that each increment D_{t+1} has conditional mean zero given the information up to time t.*

In general, any sequence of random variables $\mathbf{X} := \{X_t\}_{t\geq 1}$ is a martingale difference sequence with respect to the filtration $\mathbb{F} := (\mathcal{F}_t)_{t\geq 1}$ if it satisfies the following conditions:

1. *Adaptation: X_t is $\mathcal{F}_t$-measurable for all $t \geq 1$ (i.e. $\mathbf{X}$ is adapted to $\mathbb{F}$),*
2. *Integrability: $\mathbb{E}[|X_t|] < \infty$ for all $t \geq 1$,*
3. *Conditional Mean-Zero Property: $\mathbb{E}[X_t|\mathcal{F}_{t-1}] = 0$ for all $t \geq 1$*[2]*.*

2.4 PAC Learning Framework

In the classical PAC (Probably Approximately Correct) learning model [29], a learner is provided oracle access to samples (x, y), where x is sampled from some unknown distribution D on $\mathcal{X}$ and $y = h^\star(x)$, for some *target* hypothesis $h^\star : \mathcal{X} \to \mathcal{Y}$. We assume that $h^\star \in \mathcal{H}$, where $\mathcal{H}$ is a predefined hypothesis class, i.e., the learner has prior knowledge of $\mathcal{H}$. The goal of the learning problem is to find[3] $h : \mathcal{X} \to \mathcal{Y}$ such that the generalization error, given by the loss function $\mathcal{L}(h, D, h^\star) = \mathbb{P}_{x\sim D}(h(x) \neq h^\star(x))$, is minimized.

[2] By convention, we treat the "conditioning" on $\mathcal{F}_0$ as an unconditional expectation, i.e., $\mathbb{E}[X_1] = 0$.

[3] Note that h need not necessarily belong to $\mathcal{H}$. If it does, the learner is called *proper*. If not, the learner is *improper*.

Definition 2.6 (PAC learner)**.** *An algorithm $\mathcal{A}$ is an (ϵ, δ)-PAC learner for a hypothesis class $\mathcal{H}$ if, for any unknown distribution D and for all $h^\star \in \mathcal{H}$, $\mathcal{A}$ takes in m pairs of labeled instances, i.e., $\{(x_i, h^\star(x_i))\}_{i=1}^m$, each drawn i.i.d. from D, and outputs a hypothesis h such that $\mathbb{P}[\mathcal{L}(h, D, h^\star) \leq \epsilon] \geq 1 - \delta$, where the outer probability is over the sequence of examples and the learner's internal randomness.*

Indeed, an (ϵ, δ)-PAC learner outputs a hypothesis that is, with high probability ($\geq 1-\delta$), approximately correct ($\mathcal{L} \leq \epsilon$). A hypothesis class $\mathcal{H}$ is *PAC-learnable* if there exists an algorithm $\mathcal{A}$ that is an (ϵ, δ)-PAC learner for $\mathcal{H}$. When $\mathcal{Y} = \{0, 1\}$, we are in the setting of binary classification. To express the sample complexity of learning boolean function classes later on, we define below a key combinatorial parameter known as the VC dimension.

Definition 2.7 (VC dimension)**.** *Given a hypothesis class $\mathcal{H} = \{h : \mathcal{X} \to \{0, 1\}\}$, a set $S = \{s_1, \ldots, s_t\} \subseteq \mathcal{X}$ is said to be shattered by $\mathcal{H}$ if, for every labeling $\ell \in \{0, 1\}^t$, there exists an $h \in \mathcal{H}$ such that $(h(s_1), h(s_2), \ldots, h(s_t)) = \ell$. The VC dimension of $\mathcal{H}$, VCdim($\mathcal{H}$), is the size of the largest set S that is shattered by $\mathcal{H}$.*

2.5 Agnostic Learning Framework

In the PAC learning framework, we worked with the *realizability assumption*, namely that $h^\star \in \mathcal{H}$. If we omit this rather strong assumption, we are able to generalize the PAC learning framework to the agnostic learning framework [12]. Here, a learner is provided with oracle access to samples (x, y), sampled from some unknown distribution D on $\mathcal{X} \times \mathcal{Y}$. The learner has knowledge of a predefined hypothesis class $\mathcal{H}$. The objective of the learning problem is to find[3] $h : \mathcal{X} \to \mathcal{Y}$ such that the *regret*

$$\mathcal{R}(h, D) = \mathbb{P}_{(x,y)\sim D}(h(x) \neq y) - \inf_{h_c \in \mathcal{H}} \mathbb{P}_{(x,y)\sim D}(h_c(x) \neq y),$$

is minimized. One can notice that if the labels happen to satisfy some $h^\star \in \mathcal{H}$, $\mathcal{R} \equiv \mathcal{L}$.

Definition 2.8 (Agnostic learner)**.** *An algorithm $\mathcal{A}$ is an (ϵ, δ)-agnostic learner for a hypothesis class $\mathcal{H}$ if, for any unknown distribution D, $\mathcal{A}$ takes in m pairs of labeled instances, i.e., $(x_i, y_i)_{i=1}^m$, each drawn i.i.d. from D, and outputs a hypothesis h such that $\mathbb{P}[\mathcal{R}(h, D) \leq \epsilon] \geq 1 - \delta$, where the outer probability is over the sequence of examples and the learner's internal randomness.*

2.6 Quantum PAC and Agnostic Learning Frameworks

In the quantum setting, the primary difference from the classical setting lies in how the examples are provided. In particular, in the PAC learning setup, a quantum example [3] takes the form

$$\sum_{x \in \{0,1\}^n} \sqrt{D(x)}\, |x, h^\star(x)\rangle\,, \tag{1}$$

for some $h^\star \in \mathcal{H}$, where $D : \{0, 1\}^n \to [0, 1]$ is a distribution over the instance space[4], as before. This might appear slightly strange, as a single example seemingly contains information

[4]Here, we have taken $\mathcal{X} = \{0, 1\}^n$ for convenience and ease of analysis. However, any finite $\mathcal{X}$ could be mapped to this one, if needed.

about *all* possible classical examples. However, if we view it via the lens of measurement (see Definition 2.3), then it is clear that measuring a quantum example will provide the learner with a *single* classical example $(x, h^\star(x))$ with probability $D(x)$, exactly how it was in the classical PAC learning setup. While we have argued that the quantum example is a natural generalization of the classical example, the question still remains as to whether any sample complexity advantages in the quantum realm arise from the intrinsic description of a quantum example or from the quantum algorithm used or from both.

In the agnostic learning setting, a quantum example takes the form

$$\sum_{(x,y)\in\{0,1\}^{n+1}} \sqrt{D(x,y)}\,|x,y\rangle\,, \tag{2}$$

where, now, $D : \{0,1\}^{n+1} \to [0,1]$. These examples, like in the quantum PAC setting above, are typically *prepared* by acting on the all-zero state $|0^n, 0\rangle$ via an appropriate quantum circuit.

Given quantum examples (instead of classical examples), Definitions 2.6 and 2.8 otherwise stay exactly the same in the quantum setting.

3 Quantum Batch Learning

Under the quantum (batch) learning frameworks outlined in Section 2.6, we investigate the sample complexity of batch learning a hypothesis class $\mathcal{H}$. Specifically, we address the question of how many copies of quantum examples, as given in (1) (resp. (2)), are required to (ϵ, δ)-quantum PAC (resp. quantum agnostic) learn $\mathcal{H}$.

3.1 Binary Classification

In the binary classification setting, this question has been conclusively answered, and we reproduce the corresponding theorem below.

Theorem 3.1 (Sample complexity bounds[5] for quantum batch binary classification; Theorems 23 and 25 in [1])**.** *Let $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$. For every $\delta \in (0, 1/2)$ and $\epsilon \in (0, 1/20)$, the sample complexity of an (ϵ, δ)-quantum PAC learner (and, respectively, an (ϵ, δ)-quantum agnostic learner) for the hypothesis class $\mathcal{H}$ is given by:*

$$m^{PAC} = \Theta\left(\frac{VCdim(\mathcal{H}) + \log(\frac{1}{\delta})}{\epsilon}\right), \quad \textit{and} \quad m^{agnostic} = \Theta\left(\frac{VCdim(\mathcal{H}) + \log(\frac{1}{\delta})}{\epsilon^2}\right).$$

Arunachalam and de Wolf [1] established the *optimal* lower bounds via quantum state identification, employing ideas from Fourier analysis to assess the performance of the Pretty Good Measurement. They also derived *near-optimal* information-theoretic lower bounds, which are off by a factor of $\frac{1}{\log(\mathrm{VCdim}(\mathcal{H})/\epsilon)}$ in the term involving $\mathrm{VCdim}(\mathcal{H})$ in both the PAC and agnostic settings. Their information-theoretic approach has since been refined to provide the *optimal* lower bounds [30].

Having resolved the quantum sample complexity in the (batch) binary classification setting, [1] also presented an open question regarding the quantum sample complexity in the (batch) *multiclass* classification setting, i.e., when $\mathcal{H} \subseteq \mathcal{Y}^{\mathcal{X}}$ with $|\mathcal{Y}| = k > 2$.

[5] The upper bounds are obtained trivially via a measure-and-learn-classically quantum learner, whereas matching lower bounds are provided by [1].

3.2 Multiclass Classification

In this subsection, we provide an answer to the aforementioned question. To express sample complexity results in this setting, we first define the combinatorial parameter, Natarajan dimension ($\mathrm{Ndim}(\cdot)$), which serves as a generalization of the VC dimension to the multiclass setting.

Definition 3.2 (Natarajan dimension)**.** *Given a hypothesis class $\mathcal{H} = \{h : \mathcal{X} \to [k]\}$, a set $S = \{s_1, \ldots, s_t\} \subseteq \mathcal{X}$ is said to be N-shattered by $\mathcal{H}$ if there exist two "witness" functions $f_0, f_1 : S \to [k]$ such that:*

- *For every $x \in S$, $f_0(x) \neq f_1(x)$.*
- *For every $R \subseteq S$, there exists a function $h \in \mathcal{H}$ such that*

$$\forall x \in R, h(x) = f_0(x) \text{ and } \forall x \in S \setminus R, h(x) = f_1(x).$$

The Natarajan dimension of $\mathcal{H}$, $Ndim(\mathcal{H})$, is the size of the largest set S that is N-shattered by $\mathcal{H}$.

3.2.1 Lower Bounds

Theorem 3.3 (Sample complexity lower bounds for quantum batch multiclass classification)**.** *Let $\mathcal{H} \subseteq \mathcal{Y}^{\mathcal{X}}$, with $|\mathcal{Y}| = k > 2$. For every $\delta \in (0, 1/2)$ and $\epsilon \in (0, 1/20)$, the sample complexity of an (ϵ, δ)-quantum PAC learner (and, respectively, an (ϵ, δ)-quantum agnostic learner) for the hypothesis class $\mathcal{H}$ is bounded below as follows:*

$$m^{PAC} = \Omega\left(\frac{Ndim(\mathcal{H}) + \log(\frac{1}{\delta})}{\epsilon}\right), \quad \text{and} \quad m^{agnostic} = \Omega\left(\frac{Ndim(\mathcal{H}) + \log(\frac{1}{\delta})}{\epsilon^2}\right).$$

At its core, the proof involves reducing the problem to the quantum binary case – establishing that a learning algorithm for $\mathcal{H}$ implies a learning algorithm for $\mathcal{H}_d$, where $\mathrm{VCdim}(\mathcal{H}_d) = \mathrm{Ndim}(\mathcal{H}) = d$. This, in turn, enables us to deduce a sample complexity lower bound for learning $\mathcal{H}$ based on the corresponding lower bound for learning $\mathcal{H}_d$. A key step in the reduction involves the following transformation of a quantum binary example[6] into a quantum multiclass example,

$$\sum_{x \in [d]} \sqrt{D(x)}\, |x, y\rangle \to \sum_{x \in [d]} \sqrt{D(x)}\, |x, f_y(x)\rangle\,, \text{ where } y \in \{0, 1\}, \text{ and } f_0, f_1 : [d] \to [k]. \tag{3}$$

While in the corresponding classical reduction proof, converting $(x, y) \to (x, f_y(x))$ is entirely trivial with the knowledge of x, y, f_0, f_1, performing the transformation in (3) using only unitary operations (in a reversible manner) in the quantum realm involves delicate reasoning using an explicit quantum circuit. In particular, it is noteworthy as its existence hinges on the reversibility of the transformation $y \leftrightarrow f_y$, which is guaranteed precisely due to the definition of N-shattering.

As preliminaries for the proof, we first introduce the quantum `X`, `CNOT`, `TOFFOLI` gates, and quantum oracles for computing classical functions. The notation $\oplus$ refers to the classical `XOR` operation (i.e., addition modulo 2).

[6] To maintain consistency with Section 2.6, the input space $[d]$ can be identified with $\{0, 1\}^{\lceil \log_2 d \rceil}$.

Definition 3.4 (`X` gate)**.** *`X` (or the Pauli-`X`) gate is the quantum equivalent of the classical `NOT` gate. It operates on one qubit, mapping* $|0\rangle \to |1\rangle$ *and* $|1\rangle \to |0\rangle$ *(i.e. it "flips" the qubit).*

Definition 3.5 (`CNOT` gate)**.** *`CNOT` is a quantum gate that operates on two qubits, one control and one target. If the control qubit is in the state* $|1\rangle$*, it flips (i.e., applies an `X` gate to) the target qubit.*

Definition 3.6 (`TOFFOLI` gate)**.** *`TOFFOLI` is a quantum gate that operates on three qubits, two control and one target. If the control qubits are both in the state* $|1\rangle$*, it flips (i.e., applies an `X` gate to) the target qubit.*

Definition 3.7 (Quantum oracle U_f)**.** *For classical functions* $f : \{0,1\}^m \to \{0,1\}^n$*, there exists*[7] *a quantum oracle* U_f *that performs the unitary evolution*

$$U_f \, |x, y\rangle = |x, y \oplus f(x)\rangle \,,$$

for $x \in \{0,1\}^m$ *and* $y \in \{0,1\}^n$*.*

Proof. (of Theorem 3.3) Let $\mathcal{H} \subseteq \mathcal{Y}^{\mathcal{X}}$ be a hypothesis class of Natarajan dimension d and let $\mathcal{H}_d = \{0,1\}^{[d]}$. Let $\mathcal{A}$ be a *quantum* PAC (corresp. *quantum* agnostic) learning algorithm for $\mathcal{H}$. We proceed to show that it is possible to construct a quantum PAC (corresp. quantum agnostic) learning algorithm, $\bar{\mathcal{A}}$, for $\mathcal{H}_d$. Therefore, by reduction, we would obtain $m_{\bar{\mathcal{A}},\mathcal{H}_d} \leq m_{\mathcal{A},\mathcal{H}}$, and thus $m^{\text{PAC}}_{\mathcal{H}_d} \leq m^{\text{PAC}}_{\mathcal{H}}$ (corresp. $m^{\text{agnostic}}_{\mathcal{H}_d} \leq m^{\text{agnostic}}_{\mathcal{H}}$). Since, by construction, $\text{VCdim}(\mathcal{H}_d) = d = \text{Ndim}(\mathcal{H})$, the reduction allows us to obtain the sample complexity lower bounds being proven here (for the multiclass case), from the corresponding lower bounds for quantum batch binary classification (Theorem 3.1). Now, for the key step of the proof, given a *quantum* learner $\mathcal{A}$ for $\mathcal{H}$, it is possible to construct a *quantum* learner $\bar{\mathcal{A}}$, for $\mathcal{H}_d$, as follows. We show this for the *quantum PAC* case, and comment here that this reduction in the *quantum agnostic* case will proceed identically.

The learner $\bar{\mathcal{A}}$ receives m-copies of the quantum example $\sum_{x\in[d]} \sqrt{D(x)}\,|x, y\rangle$, where $(x, y) \in [d] \times \{0,1\}$ and $D : [d] \to [0,1]$ is an *arbitrary* distribution on $[d]$. Now, let $S = \{s_1, \ldots, s_d\} \subseteq \mathcal{X}$ be a set and f_0, f_1 be the functions that witness the N-shattering of S by $\mathcal{H}$. The learner $\bar{\mathcal{A}}$ will now attempt to convert[8] each of its m-copies of $\sum_{x\in[d]} \sqrt{D(x)}\,|x, y\rangle$ to $\sum_{x\in[d]} \sqrt{D(x)}\,|s_x, f_y(s_x)\rangle$. However, as s_x is simply an indexing of the elements of the set S, without loss of generality, we let $\bar{\mathcal{A}}$ convert each of its m-copies of $\sum_{x\in[d]} \sqrt{D(x)}\,|x, y\rangle$ to $\sum_{x\in[d]} \sqrt{D(x)}\,|x, f_y(x)\rangle$ instead. We claim that the transformation

$$|\psi\rangle_{\bar{\mathcal{A}}} = \sum_{x\in[d]} \sqrt{D(x)}\,|x, y\rangle \mapsto \sum_{x\in[d]} \sqrt{D(x)}\,|x, f_y(x)\rangle = |\psi\rangle_{\mathcal{A}} \,, \tag{4}$$

is attainable. Indeed, the quantum circuit shown in Figure 2 (and described subsequently) performs the following augmented transformation,

$$\overline{|\psi\rangle_{\bar{\mathcal{A}}}} = \sum_{x\in[d]} \sqrt{D(x)}\,|x, y, 0^{3\lceil \log_2 k\rceil}\rangle \mapsto \sum_{x\in[d]} \sqrt{D(x)}\,|x, 0, 0^{2\lceil \log_2 k\rceil}, f_y(x)\rangle = \overline{|\psi\rangle_{\mathcal{A}}}. \tag{5}$$

[7] In fact, the quantum oracle U_f can be implemented in a rather straightforward way, by using the truth table of f and generalizations of the `CNOT` gate that use several qubits as controls.

[8] The learner $\bar{\mathcal{A}}$ will then present these transformed examples to $\mathcal{A}$, the quantum PAC learner for $\mathcal{H}$.

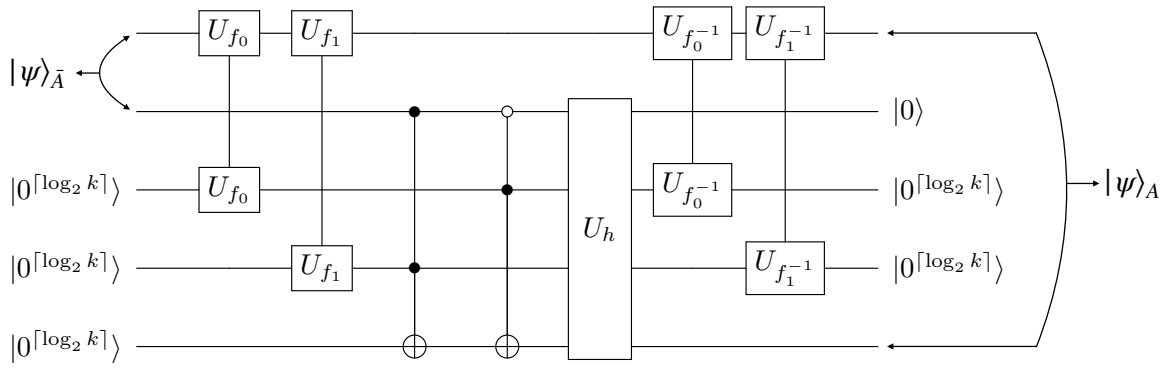

Figure 2: A quantum circuit that performs the transformation in (5). The first row denotes the $\lceil \log_2 d \rceil$ qubits involved in encoding x. The second row denotes the single qubit involved in encoding y. The third and fourth row each denote the $\lceil \log_2 k \rceil$ ancillary qubits designed to hold the intermediate computation of $f_0(x)$ and $f_1(x)$ respectively. The fifth row denotes the $\lceil \log_2 k \rceil$ qubits designed to hold the output $f_y(x)$.

We form the augmented state $\overline{|\psi\rangle_{\bar{\mathcal{A}}}}$ by appending $3\lceil \log_2 k \rceil$ qubits in the state $|0\rangle$ to the input $|\psi\rangle_{\bar{\mathcal{A}}}$. We intend to use one set of $\lceil \log_2 k \rceil$ qubits to encode each of $f_0(x)$, $f_1(x)$ (both ancillary) and $f_y(x)$ (solution). First, we pass the qubits encoding x and one set of $\lceil \log_2 k \rceil$ qubits in the state $|0\rangle$ to the quantum oracle U_{f_0}. From this, we transform

$$\overline{|\psi\rangle_{\bar{\mathcal{A}}}} = \sum_{x\in[d]} \sqrt{D(x)}\,|x, y, 0^{3\lceil \log_2 k \rceil}\rangle \mapsto \sum_{x\in[d]} \sqrt{D(x)}\,|x, y, f_0(x), 0^{2\lceil \log_2 k \rceil}\rangle\,.$$

Next, we pass the qubits encoding x and another set of $\lceil \log_2 k \rceil$ qubits in the state $|0\rangle$ to the quantum oracle U_{f_1}. From this, we transform

$$\sum_{x\in[d]} \sqrt{D(x)}\,|x, y, f_0(x), 0^{2\lceil \log_2 k \rceil}\rangle \mapsto \sum_{x\in[d]} \sqrt{D(x)}\,|x, y, f_0(x), f_1(x), 0^{\lceil \log_2 k \rceil}\rangle\,.$$

Now, we apply $\lceil \log_2 k \rceil$ `TOFFOLI` gates to each set of the qubit encoding y, a qubit involved in encoding $f_1(x)$ and a qubit $|0\rangle$ in the set of $\lceil \log_2 k \rceil$ remaining qubits that have not yet been operated on (that are designed hold the final result $f_y(x)$). From this, we transform

$$\sum_{x\in[d]} \sqrt{D(x)}\,|x, y, f_0(x), f_1(x), 0^{\lceil \log_2 k \rceil}\rangle \mapsto \sum_{x\in[d]} \sqrt{D(x)}\,|x, y, f_0(x), f_1(x), y.f_1(x)\rangle\,.$$

Next, we apply the `X` gate[9] to the qubit encoding y, and then apply $\lceil \log_2 k \rceil$ `TOFFOLI` gates to each set of the qubit (now encoding) $1-y$, a qubit involved in encoding $f_0(x)$ and a qubit that is holding the final result. Finally, we apply the `X` gate again to the qubit encoding $1-y$

[9]The concise description for controlling on $|y\rangle = |0\rangle$ (instead of explicitly on $|1-y\rangle = |1\rangle$) is depicted in Figure 2.

to revert the effect of the first `X` gate. All together, we have transformed

$$\begin{aligned}\sum_{x\in[d]}\sqrt{D(x)}\,|x,y,f_0(x),f_1(x),y.f_1(x)\rangle &\mapsto \sum_{x\in[d]}\sqrt{D(x)}\,|x,1-y,f_0(x),f_1(x),y.f_1(x)\rangle\\ &\mapsto \sum_{x\in[d]}\sqrt{D(x)}\,|x,1-y,f_0(x),f_1(x),y.f_1(x)\oplus(1-y).f_0(x)\rangle\\ &\mapsto \sum_{x\in[d]}\sqrt{D(x)}\,|x,y,f_0(x),f_1(x),y.f_1(x)\oplus(1-y).f_0(x)\rangle\\ &= \sum_{x\in[d]}\sqrt{D(x)}\,|x,y,f_0(x),f_1(x),f_y(x)\rangle\,,\end{aligned}$$

where, in the last line, we recognize that $y.f_1(x)\oplus(1-y).f_0(x)=f_y(x)$.

Now that we have computed our solution, $f_y(x)$, to ensure we are not introducing extraneous outputs to satisfy reversibility of unitary computation, we must *uncompute* all ancillary qubits to their original form. Furthermore, we also do not want the input $|y\rangle$ in our output[10], and would like to "remove" it, by transforming it to the $|0\rangle$ state. We proceed with the *uncomputation* as follows. First, we claim, with the following subproof, that there exists a (genuine) boolean *function* h that takes as input $f_0(x), f_1(x)$, and $f_y(x)$ and outputs y.

> For a short (sub)proof-by-contradiction, consider a particular input $f_0(x), f_1(x)$, and $f_y(x)=y.f_1(x)\oplus(1-y).f_0(x)$ (for some x) that maps to both $y=0$ and $y=1$. This would imply that $f_0(x)=f_1(x)$ for that particular x. However, as f_0 and f_1 are witnesses of the N-shattering of $\mathcal{H}$, they must disagree on all inputs x, providing us with the contradiction.

As h is a genuine (boolean) function, we can apply the oracle U_h to the qubits encoding $f_0(x), f_1(x)$, and $f_y(x)$ and the qubit encoding y. This gives us the transformation[11]

$$\begin{aligned}\sum_{x\in[d]}\sqrt{D(x)}\,|x,y,f_0(x),f_1(x),f_y(x)\rangle &\mapsto \sum_{x\in[d]}\sqrt{D(x)}\,|x,y\oplus y,f_0(x),f_1(x),f_y(x)\rangle\\ &= \sum_{x\in[d]}\sqrt{D(x)}\,|x,0,f_0(x),f_1(x),f_y(x)\rangle\,.\end{aligned}$$

Lastly, we apply $U_{f_0^{-1}}$ to the qubits encoding x and the ancillary qubits encoding $f_0(x)$. And then, we apply $U_{f_1^{-1}}$ to the qubits encoding x and the ancillary qubits encoding $f_1(x)$. All together, we have transformed,

$$\begin{aligned}\sum_{x\in[d]}\sqrt{D(x)}\,|x,0,f_0(x),f_1(x),f_y(x)\rangle &\mapsto \sum_{x\in[d]}\sqrt{D(x)}\,|x,0,0^{\lceil\log_2 k\rceil},f_1(x),f_y(x)\rangle\\ &\mapsto \sum_{x\in[d]}\sqrt{D(x)}\,|x,0,0^{2\lceil\log_2 k\rceil},f_y(x)\rangle = \overline{|\psi\rangle}_{\mathcal{A}}.\end{aligned}$$

[10] We do not want to allow $\mathcal{A}$ to be able to cheat by providing it with this additional knowledge.

[11] In other words, we have been able to "remove" y. It is important to note that this was possible precisely because y was recoverable from f_y (rendering $y\leftrightarrow f_y$ reversible), ensuring its ability to be encoded in a unitary operation.

Now that we have performed the transformation in (5) quantumly, ignoring[12] ancillary (and the "removed" y) qubits in the output, we note that the output of the transformation in (4) is attained.

The learner $\bar{\mathcal{A}}$ now feeds the m-copies of $\sum_{x\in[d]}\sqrt{D(x)}\,|x, f_y(x)\rangle$ as input to the *quantum* learner $\mathcal{A}$ and obtains (from $\mathcal{A}$) a *classical* (black box) function[13] $g : S \to \mathcal{Y}$ (recall $S = \{s_1, \ldots, s_d\} \subseteq \mathcal{X}$ is the subset that is N-shattered by $\mathcal{H}$). Finally, $\bar{\mathcal{A}}$ outputs the hypothesis (as a black box function) $f : [d] \to \{0, 1\}$, given by $f(i) = 1$ if and only if $g(s_i) = f_1(s_i)$ (which by construction learns $\mathcal{H}_d$).

■

3.2.2 Upper Bounds

In general, classical sample complexity upper bounds trivially translate to the corresponding quantum ones, as the quantum learner always has the option of simply performing a measurement on each quantum example, and perform the classical learning algorithm on the resulting m classical examples. We include the theorem statement (Theorem 3.8) and proof below for completeness.

Theorem 3.8 (Sample complexity upper bounds for quantum batch multiclass classification)**.** *Let $\mathcal{H} \subseteq \mathcal{Y}^{\mathcal{X}}$, with $|\mathcal{Y}| = k > 2$. The sample complexity of an (ϵ, δ)-quantum PAC learner (and, respectively, an (ϵ, δ)-quantum agnostic learner) for the hypothesis class $\mathcal{H}$ is bounded above as follows:*

$$m^{PAC} = \mathcal{O}\left(\frac{Ndim(\mathcal{H})\log(k)\log(\frac{1}{\epsilon}) + \log(\frac{1}{\delta})}{\epsilon}\right), \quad and \quad m^{agnostic} = \mathcal{O}\left(\frac{Ndim(\mathcal{H})\log(k) + \log(\frac{1}{\delta})}{\epsilon^2}\right).$$

Proof. The quantum PAC (resp. quantum agnostic) learner performs a measurement on each of the m examples $\sum_x \sqrt{D(x)}\,|x, y\rangle$ (corresp. $\sum_{x,y}\sqrt{D(x,y)}\,|x, y\rangle$) in the standard computational basis. This provides m classical examples, i.e. gives us the training set $\{(x_i, y_i)\}_{i=1}^m$, where a given (x_i, y_i) appears with probability $D(x_i)$ (corresp. $D(x_i, y_i)$). Now, the quantum PAC (resp. quantum agnostic) learners calls upon an (ϵ, δ)-PAC (resp. agnostic) *classical* learner to learn on the training set $\{(x_i, y_i)\}_{i=1}^m$, and outputs the resulting *classically learned* hypothesis. Thus, the classical sample complexity sufficiency requirements [7, 15] continue to hold, and our proof is complete. ■

In [7], the classical upper bound

$$m^{\text{PAC}} = \mathcal{O}\left(\frac{\text{Ndim}(\mathcal{H})(\log(k) + \log(\frac{1}{\epsilon}) + \log(\text{Ndim}(\mathcal{H}))) + \log(\frac{1}{\delta})}{\epsilon}\right),$$

was shown to hold which has a tighter dependence on ϵ, but a looser dependence on $\text{Ndim}(\mathcal{H})$. The proof above naturally extends this bound too to the quantum case.

[12] They are each deterministically $|0\rangle$.

[13] In fact, our identification $s_x \leftrightarrow x$ earlier means that $g : [d] \to \mathcal{Y}$ instead, and $f : [d] \to \{0, 1\}$ would then be given by $f(i) = 1$ if and only if $g(i) = f_1(i)$.

While this paper does not focus on the DS dimension, which was recently shown to characterize multiclass learnability, including the $k \to \infty$ case [8], we note that the upper bound involving the DS dimension can be similarly extended to the quantum setting. The challenges in adapting the proof of the lower bound involving the DS dimension [31] are discussed in the conclusion (Section 6).

4 Classical Online Learning

So far, we have been working with learning in the batch setting, where we are provided with all the examples at once[14]. For several practical applications, it is either impossible to obtain all the examples at once (e.g., recommendation systems), or we simply wish to evolve our learning over time. In these cases, *online learning* [16] – where we iteratively improve our hypothesis using examples we receive over time, and using our current hypothesis to predict for the upcoming example – is the appropriate framework to be placing ourselves in. First, we will introduce known models and results in classical online learning, and a *classical* generalization in Section 4.2 that, in turn, provides us with a *quantum* online learning model (Section 5) as a natural generalization. For ease of exposition, we begin with a treatment of boolean function classes in the realizable setting.

4.1 Adversary provides an input

Let $\mathcal{C} := \{c : \mathcal{X} \to \{0,1\}\}$, and $\mathcal{H} \subseteq \mathcal{C}$ (i.e., $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$). A protocol for online learning is a T-round procedure described as follows: at the t-th round,

1. Adversary provides input point in the domain: $x_t \in \mathcal{X}$.
2. Learner uses a hypothesis[15] $h_t \in \mathcal{C}$, and makes the prediction $\hat{y}_t = h_t(x_t) \in \{0,1\}$.
3. Adversary provides the input point's label, $y_t = h^\star(x_t)$, where $h^\star \in \mathcal{H}$.
4. Learner suffers a loss of 1 (a 'mistake'), if $\hat{y}_t \neq y_t$, i.e. $\mathcal{L}_{\mathbb{I}}(h_t, x_t, h^\star) = \mathbf{1}[h_t(x_t) \neq h^\star(x_t)]$.

Therefore, the learner's total loss is given by,

$$\mathcal{L}_{\mathbb{I}}(\mathbf{h}, \mathbf{x}, h^\star) = \sum_{t=1}^{T} \mathcal{L}_{\mathbb{I}}(h_t, x_t, h^\star) = \sum_{t=1}^{T} \mathbf{1}[h_t(x_t) \neq h^\star(x_t)], \tag{6}$$

where we use $\mathbf{h} := (h_1, \ldots, h_T)$ to denote the sequence of hypotheses that the learner uses, and $\mathbf{x} := (x_1, \ldots, x_T)$ to denote the sequence of instances that the adversary provides. The subscript $\mathbb{I}$ (in $\mathcal{L}_{\mathbb{I}}$) indicates that this is the input-based indicator (0-1) loss function[16].

The learner chooses an algorithm $\mathcal{A}$ that will generate the sequence $\mathbf{h}_{\mathcal{A}}$ following the protocol above. The learner's goal is to minimize $\mathcal{L}_{\mathbb{I}}(\mathbf{h}_{\mathcal{A}}, \mathcal{H}) = \sup_{\mathbf{x}, h^\star \in \mathcal{H}} \mathbb{E}[\mathcal{L}_{\mathbb{I}}(\mathbf{h}_{\mathcal{A}}, \mathbf{x}, h^\star)]$, i.e., make as few mistakes, on average, as possible regardless of the adversary's (potentially worst-case) choices of sequence of instances $\mathbf{x} := (x_1, \ldots, x_T)$ and labeling function $h^\star \in \mathcal{H}$. For

[14] This is typical for most settings where we are trying to learn a hypothesis via inductive reasoning (e.g. learning a function to fit data, etc.)

[15] Note that the learner may choose a hypothesis $h_t \in \mathcal{C} \setminus \mathcal{H}$, i.e., we do not require the learner to be *proper*.

[16] This is distinct from the probabilistic loss function $\mathcal{L}_P$ that we will encounter later in Section 4.2.

the subsequent bound on $\mathcal{L}_{\mathbb{I}}(\mathbf{h}_{\mathcal{A}}, \mathcal{H})$, we first define the combinatorial parameter, Littlestone dimension, $\mathrm{Ldim}(\mathcal{H})$.

Definition 4.1 (Littlestone dimension)**.** *Let T be a rooted tree whose internal nodes are labeled by elements from $\mathcal{X}$. Each internal node's left edge and right edge are labeled 0 and 1, respectively. The tree T is L-shattered by $\mathcal{H}$ if, for every path from root to leaf which traverses the nodes $x_1, \ldots, x_d$, there exists a hypothesis $h \in \mathcal{H}$ such that, for all i, $h(x_i)$ is the label of the edge (x_i, x_{i+1}). We define the Littlestone dimension, $Ldim(\mathcal{H})$, to be the maximal depth of a complete binary tree that is L-shattered by $\mathcal{H}$.*

The classical online learning model in this subsection (Section 4.1) has been thoroughly studied, and the following theorem characterizes it in terms of the Littlestone dimension.

Theorem 4.2 (Bounds on $\mathcal{L}_{\mathbb{I}}(\mathbf{h}_{\mathcal{A}}, \mathcal{H})$ for the canonical classical online model; Corollary 21.8 in [32] and Theorem 24 in [7])**.** *Let $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$ be a hypothesis class. The Standard Optimal Algorithm (SOA)*[17] *is a deterministic algorithm that achieves a worst-case total loss of $Ldim(\mathcal{H})$, i.e. $\mathcal{L}_{\mathbb{I}}(\mathbf{h}_{SOA}, \mathcal{H}) = Ldim(\mathcal{H})$. Furthermore, for any algorithm $\mathcal{A}$, the expected total loss on the worst-case sequence is at least*[18] $\frac{1}{2} \cdot Ldim(\mathcal{H})$*, i.e. $\mathcal{L}_{\mathbb{I}}(\mathbf{h}_{\mathcal{A}}, \mathcal{H}) \geq \frac{1}{2} \cdot Ldim(\mathcal{H})$.*

4.1.1 Can we obtain a quantum generalization of the classical online model?

Given the popularity and widespread applications of the classical online model, we explore the feasibility of developing a quantum version of the above online model to ultimately inquire whether such a quantum adaptation would be any more powerful from the perspective of the learner and/or the adversary. In essence, as there exists a well-defined "landscape" for classical and quantum *batch* learning, we seek to delineate the analogous landscape in the *online* learning context.

To this, if one attempts to naïvely generalize the above classical model to the quantum setting, an obvious issue arises: the quantum examples of the form (1) do not split the input-label pair. In particular, an adversary cannot temporally separate its provision of the input point and its label. A first step towards a model that can be generalized to the quantum setting, then, is to reorder the steps at the t-th round to $2, 1\,\&\,3, 4$ (i.e., where the learner provides a prediction $\hat{y}_t$ after which the adversary presents *both* the input and its label (x_t, y_t)).

While this reordering gives an entirely equivalent model that is, once again, characterized by Littlestone dimension (Theorem 4.2), it is not sufficient for a natural quantum generalization. The issue now is that a *classical* adversary only ever presents one (classical) example at each round. How do we go about generalizing a single classical example to a *quantum* adversary's (quantum) example that, in general, sits in superposition? It appears futile to attempt to do so. The missing piece, evidently, is the lack of a notion of a distribution over examples in the classical online model(s) examined so far.

[17] At round t, given input x_t, the SOA predicts $\hat{y}_t \in \{0,1\}$ that maximizes the Littlestone dimension of the version space consistent with $\hat{y}_t$.

[18] The adversary traverses the shattered tree and provides, at every round, the label that the (randomized) algorithm $\mathcal{A}$ is less likely to predict.

4.2 Adversary provides a distribution

Now that we have identified the unfilled gap to transition to the quantum setting, we first state the appropriate *classical* generalization of the canonical classical model (in Section 4.1) by asking the adversary to, at each t, choose a distribution over a set of input-label pairs, from which an explicit input-label pair is then drawn. The protocol for the T-round procedure will be as follows: at the t-th round,

1. Learner provides a hypothesis $h_t \in \mathcal{C}$.
2. Adversary chooses a distribution $D_t : \mathcal{X} \to [0,1]$ on the instance space, draws $x \sim D_t$, and reveals $(x, h^\star(x))$ to the learner, where $h^\star \in \mathcal{H}$.
3. Learner suffers, but does not "see", a loss of $\mathcal{L}_P(h_t, D_t, h^\star) := \mathbb{P}_{x\sim D_t}(h_t(x) \neq h^\star(x))$.

Here, the learner's total loss is given by,

$$\mathcal{L}_P(\mathbf{h}, \mathbf{D}, h^\star) = \sum_{t=1}^{T} \mathcal{L}_P(h_t, D_t, h^\star) = \sum_{t=1}^{T} \mathbb{P}_{x\sim D_t}(h_t(x) \neq h^\star(x)), \tag{7}$$

where we additionally use $\mathbf{D} := (D_1, \ldots, D_T)$ to denote the sequence of distributions that the adversary chooses. Analogously, the learner's objective is to choose $\mathbf{h}$ to minimize $\mathcal{L}_P(\mathbf{h}, \mathcal{H}) = \sup_{\mathbf{D},\, h^\star\in\mathcal{H}} \mathbb{E}[\mathcal{L}_P(\mathbf{h}, \mathbf{D}, h^\star)]$.

We identify this model as the adaptive adversary variant of the online learning model recently considered in Dawid and Tewari [2]. Note that, if we restrict the adversary, allowing it to choose only point masses, we recover the *reordered* model in Section 4.1.

From a learning standpoint, the adversary-provides-a-distribution model differs fundamentally from the canonical model in that the learner, here, does not have full information about its own loss at any given round. Since the learner does not know D_t, it cannot compute $\mathcal{L}_P(h_t, D_t, h^\star)$ for any t. In other words, the learner seeks to minimize a quantity that it cannot even compute. This *partial information* setting here, at least at first, appears to be more challenging for the learner as it not only grapples with the inability to compute its loss but also contends with the larger space available to the adversary for its choices ($D_t \in [0,1]^{\mathcal{X}}$ vs. $x_t \in \{0,1\}^{\mathcal{X}}$).

However, as we will soon illustrate, this perceived challenge proves not to be the case. The key factor influencing this distinction lies in the learner's ability to calculate $\mathcal{L}_{\mathbb{I}}(\mathbf{h}, \mathbf{x}, h^\star)$ for the observed sequence of examples $\mathbf{x}$, providing an unbiased estimator for its total loss $\mathcal{L}_P(\mathbf{h}, \mathbf{D}, h^\star)$. We demonstrate that it is indeed (necessary and) sufficient for a learner in the adversary-provides-a-distribution model to execute SOA on the observed sequence of examples $\mathbf{x}$ to achieve a bound analogous to that in the canonical model (cf. Theorem 4.2). Before delving into the results, we formally define what a learner, an adversary, and learnability entails for the adversary-provides-a-distribution model we have just discussed.

Definition 4.3 (Classical online learner)**.** *An algorithm $\mathcal{A}$ is a classical online learner for a hypothesis class $\mathcal{H} \subseteq \mathcal{C}$ if having received a sequence of examples over the first t rounds, $(x_1, h^\star(x_1)), \ldots, (x_t, h^\star(x_t))$ where $x_i \sim D_i$ with D_i arbitrary (unknown), $\mathcal{A}$ outputs a hypothesis $h_{t+1} \in \mathcal{C}$ at round*[19] *$t+1$.*

[19] Prior to receiving any examples, $\mathcal{A}$ outputs some arbitrary hypothesis $h_1 \in \mathcal{C}$ at round 1.

Definition 4.4 (Classical adversary)**.** *Having received a sequence of hypothesis* $\mathbf{h}|_t = (h_1, \ldots, h_t)$ *from the learner, and a sequence of examples* $\mathbf{x}|_t = (x_t, \ldots, x_t)$ *drawn previously from its own prior choices of distributions* $\mathbf{D}|_t = (D_1, \ldots, D_t)$ *over the first* t *rounds, at round* $t+1$*, a classical (online) adversary chooses a distribution* $D_{t+1} : \mathcal{X} \to [0,1]$ *on the instance space, draws* $x_{t+1} \in D_{t+1}$ *and reveals* $(x_{t+1}, h^\star(x_{t+1}))$ *to the learner, where* $h^\star \in \mathcal{H}$ *is consistent with all preceding labeled examples.*

Definition 4.5 (Classical online learnability)**.** *A hypothesis class* $\mathcal{H}$ *is classical online learnable if there exists a classical online learning algorithm* $\mathcal{A}$ *such that:*
$\mathcal{L}_P(\mathbf{h}_\mathcal{A}, \mathcal{H}) = \sup_{\mathbf{D},\, h^\star \in \mathcal{H}} \mathbb{E}[\mathcal{L}_P(\mathbf{h}_\mathcal{A}, \mathbf{D}, h^\star)] = o(T)$.

With these definitions in place, our next objective is to characterize learnability in the adversary-provides-a-distribution framework. When we later introduce the quantum online learning model (Section 5), these classical insights will serve as a foundation, enabling us to draw direct comparisons and understand the strong links connecting the *classical* adversary-provides-a-distribution model to the *quantum* online learning setup.

Theorem 4.6 (Upper bound on the expected loss for the classical adversary-provides-a-distribution model)**.** *Let* $\mathcal{H} \subseteq \{0,1\}^\mathcal{X}$ *be a hypothesis class, and* $h^\star \in \mathcal{H}$*. For every adversary, there exists a classical online learner for* $\mathcal{H}$ *that satisfies*

$$\mathbb{E}[\mathcal{L}_P(\mathbf{h}, \mathbf{D}, h^\star)] = \mathcal{O}(Ldim(\mathcal{H})).$$

Proof. Let $\mathbf{D}$ and $h^\star \in \mathcal{H}$ be arbitrarily chosen. We proceed by first obtaining a high-probability bound for $\mathcal{L}_P(\mathbf{h}, \mathbf{D}, h^\star)$, and then converting it to an in-expectation one. To obtain the high-probability bound, we begin by establishing that the difference between $\mathcal{L}_P(\mathbf{h}, \mathbf{D}, h^\star)$ and $\mathcal{L}_\mathbb{I}(\mathbf{h}, \mathbf{x}, h^\star)$ (see Section 4.1, (6)) on the revealed stream of examples $\mathbf{x} = (x_t)_{t=1}^T$ (with each $x_t \sim D_t$) is the sum of a martingale difference sequence.

Let $M_t := \underbrace{\mathcal{L}_P(h_t, D_t, h^\star)}_{P_t} - \underbrace{\mathcal{L}_\mathbb{I}(h_t, x_t, h^\star)}_{I_t}$, where $x_t \sim D_t$. With the filtration $\mathbb{F} := (\mathcal{F}_t)_{t=1}^T$, where $\mathcal{F}_t$ corresponds to the information revealed[20] up to (and, including) round t, namely[21] $\mathbf{h}|_t$, $\mathbf{D}|_t$, $\mathbf{x}|_t$ and $\mathbf{y}|_t := (h^\star(x_i))_{i=1}^t$, we note that $\mathbf{M} := (M_t)_{t=1}^T$ is adapted to $\mathbb{F}$ and $\forall t$,

$$\begin{aligned} \mathbb{E}[M_t] &= \mathbb{E}[P_t] - \mathbb{E}[I_t] < \infty \\ \mathbb{E}[M_t|\mathcal{F}_{t-1}] &= \mathbb{E}[P_t|\mathcal{F}_{t-1}] - \mathbb{E}[I_t|\mathcal{F}_{t-1}] = P_t - P_t = 0, \end{aligned}$$

where the first line is due to the boundedness ($0 \le P_t, I_t \le 1,\ \forall t$) of P_t and I_t. And, the second line is due to $\mathbb{E}[I_t|\mathcal{F}_{t-1}] = 1 \cdot \mathbb{P}_{x_t \sim D_t}(h_t(x_t) \neq h^\star(x_t)) + 0 \cdot \mathbb{P}_{x_t \sim D_t}(h_t(x_t) = h^\star(x_t)) = P_t$ and $\mathbb{E}[P_t|\mathcal{F}_{t-1}] = P_t$ (see Remark 1). Therefore, $\mathbf{M}$ is a martingale difference sequence.

[20] Note that this is not alluding to the information revealed to the learner. Instead, we can think of this information as having been revealed to an arbiter until the end of round t, where during each round the arbiter performs the draw $x_t \sim D_t$ on the adversary's communicated choice of D_t and provides $(x_t, h^\star(x_t))$ to the learner.

[21] Recall, $\mathbf{v}|_t = (v_1, \ldots, v_t)$, i.e. $\mathbf{v}$ restricted to the first t rounds.

We first bound the *predictable quadratic variation* $\langle M_T \rangle$ of $\mathbf{M}$, as follows:

$$\begin{aligned}
\langle M_T \rangle &:= \sum_{t=1}^{T} \mathbb{E}[M_t^2 | \mathcal{F}_{t-1}] = \sum_{t=1}^{T} \mathbb{E}[(P_t - I_t)^2 | \mathcal{F}_{t-1}] \\
&= \sum_{t=1}^{T} \Big(\mathbb{E}[P_t^2 | \mathcal{F}_{t-1}] - 2\mathbb{E}[P_t I_t | \mathcal{F}_{t-1}] + \mathbb{E}[I_t^2 | \mathcal{F}_{t-1}] \Big) \\
&= \sum_{t=1}^{T} (P_t^2 - 2P_t^2 + P_t) = \sum_{t=1}^{T} (P_t - P_t^2) \\
&\leq \sum_{t=1}^{T} P_t, \qquad (8)
\end{aligned}$$

where the second line is due to linearity of expectation, the third line uses $\mathbb{E}[P_t I_t | \mathcal{F}_{t-1}] = P_t \mathbb{E}[I_t | \mathcal{F}_{t-1}] = P_t \cdot P_t = P_t^2$, and the fourth line is due to $P_t^2 \geq 0$, $\forall t$. Now, by Theorem 1 of [33], with probability $1 - \delta$ (for any $\delta > 0$), we have

$$\sum_{t=1}^{T} M_t \leq \log\Big(\frac{1}{\delta}\Big) + (e-2)\langle M_T \rangle \iff \sum_{t=1}^{T} P_t \leq \sum_{t=1}^{T} I_t + \log\Big(\frac{1}{\delta}\Big) + (e-2)\langle M_T \rangle.$$

Note here that $\sum_{t=1}^{T} P_t = \mathcal{L}_P(\mathbf{h}, \mathbf{D}, h^\star)$. Therefore, appealing to Theorem 4.2 and the inequality on $\langle M_T \rangle$ in (8), we have that (with probability $1 - \delta$),

$$\begin{aligned}
&\mathcal{L}_P(\mathbf{h}, \mathbf{D}, h^\star) \leq \mathrm{Ldim}(\mathcal{H}) + \log\Big(\frac{1}{\delta}\Big) + (e-2)\mathcal{L}_P(\mathbf{h}, \mathbf{D}, h^\star) \\
\implies &\mathcal{L}_P(\mathbf{h}, \mathbf{D}, h^\star) \leq \frac{1}{1-(e-2)}\mathrm{Ldim}(\mathcal{H}) + \frac{1}{1-(e-2)} \log\Big(\frac{1}{\delta}\Big),
\end{aligned}$$

or equivalently, for any $\delta > 0$,

$$\mathbb{P}\Big[\mathcal{L}_P(\mathbf{h}, \mathbf{D}, h^\star) > \frac{1}{1-(e-2)} \cdot \mathrm{Ldim}(\mathcal{H}) + \frac{\delta}{1-(e-2)}\Big] \leq e^{-\delta}. \qquad (9)$$

Now, we compute the in-expectation bound (i.e. a bound on $\mathbb{E}[\mathcal{L}_P(\mathbf{h}, \mathbf{D}, h^\star)]$) guaranteed by the above tail bound (9). Let $c := \dfrac{1}{1-(e-2)}$.

$$\begin{aligned}
\mathbb{E}[\mathcal{L}_P(\mathbf{h}, \mathbf{D}, h^\star)] &= \int_0^\infty \mathbb{P}[\mathcal{L}_P(\mathbf{h}, \mathbf{D}, h^\star) > \ell]\, d\ell \\
&= \int_0^{c\cdot\mathrm{Ldim}(\mathcal{H})} \mathbb{P}[\mathcal{L}_P(\mathbf{h}, \mathbf{D}, h^\star) > \ell]\, d\ell + \int_{c\cdot\mathrm{Ldim}(\mathcal{H})}^{\infty} \mathbb{P}[\mathcal{L}_P(\mathbf{h}, \mathbf{D}, h^\star) > \ell]\, d\ell \\
&= \int_0^{c\cdot\mathrm{Ldim}(\mathcal{H})} \mathbb{P}[\mathcal{L}_P(\mathbf{h}, \mathbf{D}, h^\star) > \ell]\, d\ell + c\int_0^{\infty} \mathbb{P}\Big[\mathcal{L}_P(\mathbf{h}, \mathbf{D}, h^\star) > c\cdot\mathrm{Ldim}(\mathcal{H}) + c\delta\Big]\, d\delta \\
&\leq \int_0^{c\cdot\mathrm{Ldim}(\mathcal{H})} 1\, d\ell + c\int_0^{\infty} e^{-\delta}\, d\delta \\
&= c\cdot\mathrm{Ldim}(\mathcal{H}) + c \\
&= \mathcal{O}(\mathrm{Ldim}(\mathcal{H})),
\end{aligned}$$

where the first line holds as $\mathcal{L}_P(\mathbf{h}, \mathbf{D}, h^\star) \geq 0$, the third line uses the change of variable $\ell = c \cdot \mathrm{Ldim}(\mathcal{H}) + c\delta$, the fourth line uses a naïve bound of $\mathbb{P}[\mathcal{L}_P(\mathbf{h}, \mathbf{D}, h^\star) > \ell] \leq 1$ (for $0 \leq \ell \leq c \cdot \mathrm{Ldim}(\mathcal{H})$) on the first integral, and our high-probability bound in (9) on the second integral. ■

Remark 1 (Internal Randomness)**.** *In the proof of Theorem 4.6 (and in subsequent proofs throughout the paper), the filtration $\mathbb{F}$ does not include any information on the internal randomness of the learner or the (adaptive) adversary at any round. In other words, the bound in Theorem 4.6 applies to a T-round procedure that is behaving deterministically (in as much as it applies to the learner and the adversary). Crucially, however, it applies to all such procedures. Therefore, the bounds presented in Theorem 4.6 (as well as others throughout the paper) imply identical bounds on a T-round procedure where internal randomness is allowed for both the learner and the adversary.*

To summarize, Theorem 4.6 tells us that $\mathrm{Ldim}(\mathcal{H})$ continues to be a sufficient condition for learnability under the new (adversary-provides-a-distribution) model in Section 4.2. In other words, a learner that performs SOA on the observed sequence of examples $\mathbf{x}$ only suffers a constant overhead under the adversary-provides-a-distribution model as compared to SOA under the canonical online model (Section 4.1). Next, we show (Theorem 4.7) that $\mathrm{Ldim}(\mathcal{H})$ is also a necessary condition for learnability, and thus fully characterizes the learnability of the adversary-provides-a-distribution model.

Theorem 4.7 (Lower bound on the expected loss for the classical adversary-provides-a-distribution model)**.** *Let $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$ be a hypothesis class, and $h^\star \in \mathcal{H}$. For every classical online learner of $\mathcal{H}$, there exists an adversary such that*

$$\mathbb{E}[\mathcal{L}_P(\mathbf{h}, \mathbf{D}, h^\star)] = \Omega(\mathit{Ldim}(\mathcal{H})).$$

Proof. Consider an adversary which chooses each D_t to be a point mass on the instance space, i.e. the adversary simply chooses an instance x_t at each t. Since each $x_t \sim D_t$ is deterministic, we have

$$\mathcal{L}_P(\mathbf{h}, \mathbf{D}, h^\star) = \mathcal{L}_{\mathbb{I}}(\mathbf{h}, \mathbf{x}, h^\star) = \Omega(\mathrm{Ldim}(\mathcal{H})),$$

where the second equality is due to (the lower bound part of) Theorem 4.2. By taking expectations, we conclude our proof. ■

Now that we have characterized the learnability of the adversary-provides-a-distribution model in Section 4.2 (via Theorems 4.6 and 4.7), we proceed to introduce and present results for related classical online models which arise from successively relaxing, first, the *realizability* assumption and then, the *boolean function class* assumption. These serve to define our scope and lay the groundwork for a comprehensive understanding before introducing the anticipated quantum generalization.

4.3 Adversary provides a distribution in the agnostic setting

In the agnostic framework, we dispense with the realizability assumption that $h^\star \in \mathcal{H}$ (i.e., the labels need not be consistent with any hypothesis in the hypothesis class). In fact, the labels

need not arise from a labeling function at all, i.e., the examples may be inconsistent[22]. The agnostic generalization of the adversary-provides-a-distribution model in Section 4.2 is given by the following protocol for the T-round procedure: at the t-th round,

1. Learner provides a hypothesis $h_t \in \mathcal{C}$.
2. Adversary chooses a distribution $D_t : \{0,1\}^{n+1} \to [0,1]$ on the instance space, draws and reveals $z_t = (x_t, y_t) \sim D_t$ to the learner.
3. Learner suffers, but does not "see", a loss of $\mathcal{L}_P(h_t, D_t) := \mathbb{P}_{(x,y)\sim D_t}(h_t(x) \neq y)$.

As there may be no hypothesis that provides the true label on every instance over the T rounds, we resort to comparing the learner to the best fixed hypothesis in $\mathcal{H}$ in hindsight. In other words, the learner's total *regret* is given by,

$$\mathcal{R}_P(\mathbf{h}, \mathbf{D}, \mathcal{H}) = \sum_{t=1}^{T} \mathcal{L}_P(h_t, D_t) - \inf_{h\in\mathcal{H}} \sum_{t=1}^{T} \mathcal{L}_P(h, D_t). \tag{10}$$

Definitions 4.3, 4.4, and 4.5 are adapted analogously to give us the notions of a classical online learner, adversary and learnability in the agnostic setting. Additionally, we introduce further notation and definitions to facilitate the subsequent theorem.

An **$\mathcal{X}$-valued tree $\underline{X}$ of depth** T is a rooted complete binary tree with nodes labeled by elements of $\mathcal{X}$. We identify the tree $\underline{X}$ with the sequence $x^{(1)}, \ldots, x^{(T)}$ of labeling functions $x^{(i)} : \{\pm 1\}^{i-1} \to \mathcal{X}$ which provide the labels for each node. Here, $x^{(1)}$ labels the root of the tree, and $x^{(i)}$ for $i > 1$ labels the node obtained by following the path of length $i-1$ from the root, with $+1$ indicating 'right' and -1 indicating 'left'. A path of length T is given by the sequence $\epsilon = (\epsilon_1, \ldots, \epsilon_T) \in \{\pm 1\}^T$. We denote the label at round t along this path as $x^{(t)}(\epsilon)$, understanding that $x^{(t)}$ depends only on the prefix $(\epsilon_1, \ldots, \epsilon_{t-1})$ of ϵ. With this notion of a tree, we define the *sequential Rademacher complexity* of a hypothesis class, $\mathcal{H} \subseteq \mathcal{Y}^{\mathcal{X}}$ [34].

Definition 4.8 (Sequential Rademacher complexity of $\mathcal{H}$)**.** *The sequential Rademacher complexity of a function class $\mathcal{H} \subseteq \boldsymbol{R}^{\mathcal{X}}$ on an $\mathcal{X}$-valued tree $\underline{X}$ is defined as*

$$\mathfrak{R}_T(\mathcal{H}, \underline{X}) = \mathbb{E}\Big[\sup_{h\in\mathcal{H}} \frac{1}{T} \sum_{t=1}^{T} \epsilon_t h(x^{(t)}(\epsilon))\Big],$$

and

$$\mathfrak{R}_T(\mathcal{H}) = \sup_{\underline{X}} \mathfrak{R}_T(\mathcal{H}, \underline{X}),$$

where the outer supremum is taken over all $\mathcal{X}$-valued trees of depth T; $\epsilon = (\epsilon_1, \ldots, \epsilon_T) \in \{\pm 1\}^T$ is a sequence of i.i.d. Rademacher random variables.

Definition 4.9 (The loss class $\mathcal{L}_{\mathbb{I}} \circ \mathcal{H}$)**.** *The loss class, $\mathcal{L}_{\mathbb{I}} \circ \mathcal{H}$, is a boolean hypothesis class given by*

$$\mathcal{L}_{\mathbb{I}} \circ \mathcal{H} = \{l_h : (x, y) \mapsto \mathbf{1}[h(x) \neq y] \,|\, h \in \mathcal{H}\}.$$

[22] That is, it is entirely possible to encounter both $(x, 0)$ and $(x, 1)$ in the sequence of examples.

Note that $\mathcal{L}_{\mathbb{I}} \circ \mathcal{H} \subseteq \{0,1\}^{\mathcal{X}\times\mathcal{Y}}$, and that its sequential Rademacher complexity is defined analogously to Definition 4.8. With these definitions in place, we can proceed to present the theorems for the bounds on expected regret for the adversary-provides-a-distribution model in the agnostic setting.

Theorem 4.10 (Upper bound on the expected regret for the classical adversary-provides-a-distribution model in the agnostic setting)**.** *Let $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$ be a hypothesis class. For every adversary, there exists a classical online learner for $\mathcal{H}$ that satisfies*

$$\mathbb{E}[\mathcal{R}_P(\mathbf{h}, \mathbf{D}, \mathcal{H})] = \mathcal{O}(\sqrt{Ldim(\mathcal{H}) \cdot T}).$$

Proof. Let $\mathbf{D}$ be an arbitrary sequence of distributions, and let $\mathbf{z} = (z_1, \ldots, z_t)$ be a sequence of instances such that $z_i \sim D_i$. Then, defining[23] $\mathcal{R}_{\mathbb{I}}(\mathbf{h}, \mathbf{z}, \mathcal{H}) := \sum_{t=1}^{T} \mathcal{L}_{\mathbb{I}}(h_t, z_t) - \inf_{h\in\mathcal{H}} \sum_{t=1}^{T} \mathcal{L}_{\mathbb{I}}(h, z_t)$, we have

$$\begin{aligned}\mathcal{R}_P(\mathbf{h}, \mathbf{D}, \mathcal{H}) &= \mathcal{R}_P(\mathbf{h}, \mathbf{D}, \mathcal{H}) - \mathcal{R}_{\mathbb{I}}(\mathbf{h}, \mathbf{z}, \mathcal{H}) + \mathcal{R}_{\mathbb{I}}(\mathbf{h}, \mathbf{z}, \mathcal{H}) \\ &= \underbrace{\sum_{t=1}^{T} \mathcal{L}_P(h_t, D_t) - \sum_{t=1}^{T} \mathcal{L}_{\mathbb{I}}(h_t, z_t)}_{\Delta_1} + \underbrace{\inf_{h\in\mathcal{H}} \sum_{t=1}^{T} \mathcal{L}_{\mathbb{I}}(h, z_t) - \inf_{h\in\mathcal{H}} \sum_{t=1}^{T} \mathcal{L}_P(h, D_t)}_{\Delta_2} + \mathcal{R}_{\mathbb{I}}(\mathbf{h}, \mathbf{z}, \mathcal{H}).\end{aligned}$$

We proceed by bounding the expected value of Δ_1, Δ_2 and $\mathcal{R}_{\mathbb{I}}(\mathbf{h}, \mathbf{z}, \mathcal{H})$ separately.

Working with Δ_1, we let $M_t := \mathcal{L}_P(h_t, D_t) - \mathcal{L}_{\mathbb{I}}(h_t, z_t)$. With the filtration $\mathbb{F} := (\mathcal{F}_t)_{t=1}^{T}$, where $\mathcal{F}_t$ corresponds to the information revealed up to (and, including) round t, namely $\mathbf{h}|_t$, $\mathbf{D}|_t$ and $\mathbf{z}|_t$, we note that $\mathbf{M} := (M_t)_{t=1}^{T}$ is adapted to $\mathbb{F}$ and $\forall t$,

$$\begin{aligned}\mathbb{E}[M_t] &= \mathbb{E}[\mathcal{L}_P(h_t, D_t)] - \mathbb{E}[\mathcal{L}_{\mathbb{I}}(h_t, z_t)] < \infty \\ \mathbb{E}[M_t|\mathcal{F}_{t-1}] &= \mathbb{E}[\mathcal{L}_P(h_t, D_t)|\mathcal{F}_{t-1}] - \mathbb{E}[\mathcal{L}_{\mathbb{I}}(h_t, z_t)|\mathcal{F}_{t-1}] = \mathcal{L}_P(h_t, D_t) - \mathcal{L}_P(h_t, D_t) = 0,\end{aligned}$$

where the first line is due to the boundedness ($0 \leq \mathcal{L}_P(h_t, D_t), \mathcal{L}_{\mathbb{I}}(h_t, z_t) \leq 1,\ \forall t$) of $\mathcal{L}_P(h_t, D_t)$ and $\mathcal{L}_{\mathbb{I}}(h_t, z_t)$. And, the second line is due to $\mathbb{E}[\mathcal{L}_{\mathbb{I}}(h_t, z_t)|\mathcal{F}_{t-1}] = 1 \cdot \mathbb{P}_{z_t\sim D_t}(h_t(x_t) \neq y_t) + 0 \cdot \mathbb{P}_{z_t\sim D_t}(h_t(x_t) = y_t) = \mathcal{L}_P(h_t, D_t)$. Therefore, $\mathbf{M}$ is a martingale difference sequence. Now, as $\Delta_1 = \sum_{t=1}^{T} M_t$, by Azuma-Hoeffding's inequality, since $|M_t| < 1$ for all t, we have that

$$\mathbb{P}[|\Delta_1| \geq \delta] \leq 2\exp\Big(-\frac{\delta^2}{2T}\Big), \quad \text{for all } \ \delta \in \mathbf{R}^+, T \in \mathbb{Z}^+. \tag{11}$$

This allows us to compute the following bound on $\mathbb{E}[\Delta_1]$:

$$\begin{aligned}\mathbb{E}[\Delta_1] \leq \mathbb{E}[|\Delta_1|] &\leq \int_0^\infty \mathbb{P}[|\Delta_1| \geq \delta]\, dB \\ &\leq \int_0^\infty 2\exp\Big(-\frac{\delta^2}{2T}\Big)\, dB = \sqrt{2\pi T} = \mathcal{O}(\sqrt{T}),\end{aligned}$$

where the first line holds as $\Delta_1 < |\Delta_1|$ (for the first inequality) and $|\Delta_1| \geq 0$ (for the second inequality), and the second line uses the bound in (11).

[23] $\mathcal{R}_{\mathbb{I}}(\mathbf{h}, \mathbf{z}, \mathcal{H})$ is precisely the regret of an algorithm in the agnostic generalization of the canonical (adversary-provides-an-input) classical online model in Section 4.1.

Next, working with Δ_2, we obtain the following chain of (in)equalities:

$$\begin{aligned}
\Delta_2 &= \inf_{h\in\mathcal{H}} \sum_{t=1}^{T} \mathcal{L}_{\mathbb{I}}(h, z_t) - \inf_{h\in\mathcal{H}} \sum_{t=1}^{T} \mathcal{L}_P(h, D_t) \\
&= -\sup_{h\in\mathcal{H}} \sum_{t=1}^{T} -\mathcal{L}_{\mathbb{I}}(h, z_t) + \sup_{h\in\mathcal{H}} \sum_{t=1}^{T} -\mathcal{L}_P(h, D_t) \\
&= -\sup_{h\in\mathcal{H}} \sum_{t=1}^{T} (-\mathcal{L}_{\mathbb{I}}(h, z_t) + \mathcal{L}_P(h, D_t) - \mathcal{L}_P(h, D_t)) + \sup_{h\in\mathcal{H}} \sum_{t=1}^{T} -\mathcal{L}_P(h, D_t) \\
&\leq \sup_{h\in\mathcal{H}} \sum_{t=1}^{T} (\mathcal{L}_{\mathbb{I}}(h, z_t) - \mathcal{L}_P(h, D_t)),
\end{aligned}$$

where the final inequality is a result of the subadditivity of the supremum; to elaborate, we derive $\sup_h f \leq \sup_h (f-g) + \sup_h g \iff \sup_h f - \sup_h g \leq \sup_h (f-g)$, to which we substitute $f = -\mathcal{L}_P(h, D_t)$ and $g = -\mathcal{L}_{\mathbb{I}}(h, z_t) + \mathcal{L}_P(h, D_t) - \mathcal{L}_P(h, D_t)$. So far, taking expectations both sides, we have $\mathbb{E}[\Delta_2] \leq \mathbb{E}[\sup_{h\in\mathcal{H}} \sum_{t=1}^{T} (\mathcal{L}_{\mathbb{I}}(h, z_t) - \mathcal{L}_P(h, D_t))]$. Next, we bound the in-expectation quantity on the right-hand side to obtain,

$$\begin{aligned}
\mathbb{E}\Big[\sup_{h\in\mathcal{H}} \sum_{t=1}^{T} (\mathcal{L}_{\mathbb{I}}(h, z_t) - \mathcal{L}_P(h, D_t))\Big] &= \mathbb{E}\left[\sup_{h\in\mathcal{H}} \sum_{t=1}^{T} \Big(\mathbf{1}[h(x_t) \neq y_t] - \mathbb{P}_{(x_t,y_t)\sim D_t}[h(x_t) \neq y_t]\Big)\right] \\
&= \mathbb{E}\left[\sup_{h\in\mathcal{H}} \sum_{t=1}^{T} \Big(\mathbf{1}[h(x_t) \neq y_t] - \mathbb{E}_{(x_t,y_t)\sim D_t}\Big[\mathbf{1}[h(x_t) \neq y_t]|\mathcal{F}_{t-1}\Big]\Big)\right] \\
&\leq 2T \cdot \mathfrak{R}_T(\mathcal{L}_{\mathbb{I}} \circ \mathcal{H}) \\
&\leq 2T \cdot \mathfrak{R}_T(\mathcal{H}) \\
&= \mathcal{O}(\sqrt{\mathrm{Ldim}(\mathcal{H}) \cdot T}),
\end{aligned}$$

where the second line uses a conditioning on the filtration $\mathcal{F}_{t-1}$, which corresponds to the information revealed up to (and, including) round $t-1$ (namely $\mathbf{h}|_{t-1} := (h_1, \ldots, h_{t-1})$, $\mathbf{D}|_{t-1} := (D_1, \ldots, D_{t-1})$ and $\mathbf{z}|_{t-1} := ((x_1, y_1), \ldots, (x_{t-1}, y_{t-1}))$), the third line is due to Theorem 2 in [34][24], the fourth line is due to Theorem 16 of [2] (with $\mathcal{Y} = \{0, 1\}$) and the last line is from the proof of Theorem 12.1 in [18].

Finally, also from Theorem 12.1 of [18], we have $\mathcal{R}_{\mathbb{I}}(\mathbf{h}, \mathbf{z}, \mathcal{H}) = \mathcal{O}(\sqrt{\mathrm{Ldim}(\mathcal{H}) \cdot T})$. Putting

[24] Theorem 2 in [34], as stated, provides a bound on $\mathbb{E}\Big[\sup_{h\in\mathcal{H}} \sum_{t=1}^{T} \Big(\mathbb{E}_{(x_t,y_t)\sim D_t}\Big[\mathbf{1}[h(x_t) \neq y_t]|\mathcal{F}_{t-1}\Big] - \mathbf{1}[h(x_t) \neq y_t]\Big)\Big]$. However, the proof of Lemma 9 in [34] notes its validity even with absolute values around the sum, which subsequently ensures that Theorem 2 in [34] also holds in the same generality. This justifies its use here in the following sense: $\mathbb{E}\Big[\sup_{h\in\mathcal{H}} \sum_{t=1}^{T} \Big(\mathbf{1}[h(x_t) \neq y_t] - \mathbb{E}_{(x_t,y_t)\sim D_t}\Big[\mathbf{1}[h(x_t) \neq y_t]|\mathcal{F}_{t-1}\Big]\Big)\Big] \leq \mathbb{E}\Big[\sup_{h\in\mathcal{H}} \Big|\sum_{t=1}^{T} \Big(\mathbb{E}_{(x_t,y_t)\sim D_t}\Big[\mathbf{1}[h(x_t) \neq y_t]|\mathcal{F}_{t-1}\Big] - \mathbf{1}[h(x_t) \neq y_t]\Big)\Big|\Big] \leq 2T \cdot \mathfrak{R}_T(\mathcal{L}_{\mathbb{I}} \circ \mathcal{H})$.

everything together, using the linearity of expectation, we have,

$$\begin{aligned}\mathbb{E}[\mathcal{R}_P(\mathbf{h},\mathbf{D},\mathcal{H})] &= \mathbb{E}[\Delta_1+\Delta_2+\mathcal{R}_{\mathbb{I}}(\mathbf{h},\mathbf{z},\mathcal{H})]\\ &= \mathbb{E}[\Delta_1]+\mathbb{E}[\Delta_2]+\mathbb{E}[\mathcal{R}_{\mathbb{I}}(\mathbf{h},\mathbf{z},\mathcal{H})]\\ &= \mathcal{O}(\sqrt{T})+\mathcal{O}(\sqrt{\mathrm{Ldim}(\mathcal{H})\cdot T})+\mathcal{O}(\sqrt{\mathrm{Ldim}(\mathcal{H})\cdot T})\\ &= \mathcal{O}(\sqrt{\mathrm{Ldim}(\mathcal{H})\cdot T}).\end{aligned}$$

■

In summary, our theorem reveals that the optimal classical learner for the canonical (adversary-provides-an-input) agnostic model, when provided with the observed sequence of instances $\mathbf{z}$ in the new protocol, experiences, at most, a constant overhead when assessed under the new (adversary-provides-a-distribution) framework. Upon closer examination, it was critical for the bound on $\mathbb{E}[\Delta_2]$ in our proof not to be worse than the bound on $\mathbb{E}[\mathcal{R}_{\mathbb{I}}(\mathbf{h},\mathbf{z},\mathcal{H})]$. It is noteworthy that the bound on $\mathbb{E}[\Delta_2]$ is guaranteed by the rate of online uniform convergence (in the frameworks of the sequential Rademacher complexity [34] and the adversarial (uniform) laws of large numbers [18]), whereas the bound on $\mathbb{E}[\mathcal{R}_{\mathbb{I}}(\mathbf{h},\mathbf{z},\mathcal{H})]$ is guaranteed by the rate of canonical (agnostic) online learnability. The equivalence between these two rates, for the boolean function class case, played a pivotal role in establishing our result.

Next, we establish (Theorem 4.11) a matching lower bound for expected regret within the agnostic adversary-provides-a-distribution framework. This fully characterizes agnostic learnability under the adversary-provides-a-distribution framework. As with all lower bound proofs within this framework, we efficiently conclude the proof statement by considering an adversary that exclusively plays point masses.

Theorem 4.11 (Lower bound on the expected regret for the classical adversary-provides-a-distribution model in the agnostic setting)**.** *Let $\mathcal{H}\subseteq\{0,1\}^{\mathcal{X}}$ be a hypothesis class. For every classical online learner of $\mathcal{H}$, there exists an adversary such that*

$$\mathbb{E}[\mathcal{R}_P(\mathbf{h},\mathbf{D},\mathcal{H})] = \Omega(\sqrt{Ldim(\mathcal{H})\cdot T}).$$

Proof. As in Theorem 4.7, we again consider an adversary which chooses each D_t to be a point mass on the instance space, i.e. the adversary simply chooses an instance z_t at each t. Since each $z_t\sim D_t$ is deterministic, we have

$$\mathbb{E}[\mathcal{R}_P(\mathbf{h},\mathbf{D},\mathcal{H})] = \mathbb{E}[\mathcal{R}_{\mathbb{I}}(\mathbf{h},\mathbf{z},\mathcal{H})] = \Omega(\sqrt{\mathrm{Ldim}(\mathcal{H})\cdot T}),$$

where the second equality is due to (the lower bound part of) Theorem 21.10 in [32]. ■

4.4 Adversary provides a distribution in the multiclass setting

In Sections 4.1, 4.2 and 4.3, we considered boolean hypothesis classes, i.e. $\mathcal{H}\subseteq\{0,1\}^{\mathcal{X}}$. Here, we consider the adversary-provides-a-distribution models in Sections 4.2 and 4.3 extended to the setting of multiclass learning, i.e. $\mathcal{H}\subseteq\mathcal{C}:=\{c:\mathcal{X}\to\mathcal{Y}\}$, with $|\mathcal{Y}|>2$. As stated earlier, the objective here is to lay the groundwork, with a clear understanding of these models in the classical paradigm, before delving into their quantum generalizations.

4.4.1 Realizable setting

The online learning protocol in the realizable setting is identical to that specified in Section 4.2, with the added specification of a multiclass hypothesis (and, concept) class. To express our results in this setting, we first define the combinatorial parameter, multiclass Littlestone dimension ($\mathrm{mcLdim}(\mathcal{H})$), which is a generalization of the Littlestone dimension to the multiclass setting.

Definition 4.12 (Multiclass Littlestone dimension)**.** *Let T be a rooted tree whose internal nodes are labeled by elements from $\mathcal{X}$ and whose edges are labeled by elements from $\mathcal{Y}$, such that the edges from a single parent to its child nodes are each labeled with a different label*[25]*. The tree T is mcL-shattered by $\mathcal{H}$ if, for every path from root to leaf which traverses the nodes $x_1, \ldots, x_d$, there exists a hypothesis $h \in \mathcal{H}$ such that, for all i, $h(x_i)$ is the label of the edge (x_i, x_{i+1}). We define the multiclass Littlestone dimension, $mcLdim(\mathcal{H})$, to be the maximal depth of a complete **binary** tree that is mcL-shattered by $\mathcal{H}$.*

Theorem 4.13 (Upper bound on the expected loss for the classical adversary-provides-a-distribution model in the multiclass setting)**.** *Let $\mathcal{H} \subseteq \mathcal{Y}^{\mathcal{X}}$, with $|\mathcal{Y}| = k > 2$, be a hypothesis class, and $h^\star \in \mathcal{H}$. For every adversary, there exists a classical online learner for $\mathcal{H}$ that satisfies*

$$\mathbb{E}[\mathcal{L}_P(\mathbf{h}, \mathbf{D}, h^\star)] = \mathcal{O}(mcLdim(\mathcal{H})).$$

Proof. We follow the steps in the proof of Theorem 4.6, which continue to hold in the multiclass setting, with a minor difference: the upper bound on $\sum_{t=1}^{T} \mathcal{L}_{\mathbb{I}}(h_t, x_t, h^\star)$ is now $\mathrm{mcLdim}(\mathcal{H})$ instead of $\mathrm{Ldim}(\mathcal{H})$. As a result, we arrive at the following high-probability bound: with probability $1-\delta$ (for any $\delta > 0$), $\mathcal{L}_P(\mathbf{h}, \mathbf{D}, h^\star) \leq \frac{1}{1-(e-2)}\mathrm{mcLdim}(\mathcal{H}) + \frac{1}{1-(e-2)}\log(\frac{1}{\delta})$, or equivalently, the following tail bound: for any $\delta > 0$,

$$\mathbb{P}\Big[\mathcal{L}_P(\mathbf{h}, \mathbf{D}, h^\star) > \frac{1}{1-(e-2)} \cdot \mathrm{mcLdim}(\mathcal{H}) + \frac{\delta}{1-(e-2)}\Big] \leq e^{-\delta},$$

from which we obtain the desired in-expectation result: $\mathbb{E}[\mathcal{L}_P(\mathbf{h}, \mathbf{D}, h^\star)] = \mathcal{O}(\mathrm{mcLdim}(\mathcal{H}))$. ■

Theorem 4.14 (Lower bound on the expected loss for the classical adversary-provides-a-distribution model in the multiclass setting)**.** *Let $\mathcal{H} \subseteq \mathcal{Y}^{\mathcal{X}}$, with $|\mathcal{Y}| = k > 2$, be a hypothesis class, and $h^\star \in \mathcal{H}$. For every classical online learner of $\mathcal{H}$, there exists an adversary such that*

$$\mathbb{E}[\mathcal{L}_P(\mathbf{h}, \mathbf{D}, h^\star)] = \Omega(mcLdim(\mathcal{H})).$$

Proof. As in the online lower bound proofs thus far, we consider an adversary which chooses each D_t to be a point mass on the instance space, i.e. the adversary simply chooses an instance x_t at each t. Since each $x_t \sim D_t$ is deterministic, we have

$$\mathbb{E}[\mathcal{L}_P(\mathbf{h}, \mathbf{D}, h^\star)] = \mathbb{E}[\mathcal{L}_{\mathbb{I}}(\mathbf{h}, \mathbf{x}, h^\star)] = \Omega(\mathrm{mcLdim}(\mathcal{H})),$$

where the second equality is due to (the lower bound part of) Theorem 24 in [7]. ■

[25] In the binary case (where the only "different labels" are 0 and 1), it is not hard to see that the definition reduces to that of the Littlestone dimension (Definition 4.1).

Theorems 4.13 and 4.14 together imply that $\mathrm{mcLdim}(\mathcal{H})$ continues to characterize multiclass learnability in the realizable setting under the adversary-provides-a-distribution framework. Given that the rate $\Theta(\mathrm{mcLdim}(\mathcal{H}))$ is independent of k, $\mathrm{mcLdim}(\mathcal{H})$ characterizes realizable learnability even in the unbounded label space case, mirroring the scenario in the adversary-provides-an-input framework (Theorem 24, [7]).

4.4.2 Agnostic setting

Here, the online learning protocol is identical to that specified in Section 4.3, with the added specification of a multiclass hypothesis (and, concept) class. We introduce specific definitions and lemmas to facilitate the subsequent theorem, which provides an upper bound on the expected regret of an online learner in this multiclass agnostic adversary-provides-a-distribution setting. For notation related to trees, please refer to Section 4.3 (paragraph preceding Definition 4.8).

Definition 4.15 (0-cover of a hypothesis class $\mathcal{H}$ on a tree $\underline{X}$)**.** *A set $\mathcal{V}$ of $\mathcal{Y}$-valued trees is a 0-cover of $\mathcal{H} \subseteq \mathcal{Y}^{\mathcal{X}}$ on an $\mathcal{X}$-valued tree $\underline{X}$ of depth T if*

$$\forall h \in \mathcal{H}, \forall \epsilon \in \{\pm 1\}^T, \exists \underline{V} \in \mathcal{V},\ s.t.,\ v^{(t)}(\epsilon) = h(x^{(t)}(\epsilon)),$$

for all $t \in [T]$.

Definition 4.16 (Covering number of a hypothesis class $\mathcal{H}$ on a tree $\underline{X}$)**.** *The covering number of a hypothesis class $\mathcal{H} \subseteq \mathcal{Y}^{\mathcal{X}}$ on an $\mathcal{X}$-valued tree $\underline{X}$, $\mathcal{N}(\mathcal{H}, \underline{X})$, is defined as follows:*

$$\mathcal{N}(\mathcal{H}, \underline{X}) := \min\{|\mathcal{V}| : \mathcal{V} \text{ is a 0-cover of } \mathcal{H} \text{ on } \underline{X}\},$$

i.e. the size of the smallest set $\mathcal{V}$ (of trees) that 0-covers $\underline{X}$.

Lemma 4.17 ($\mathcal{N}(\mathcal{L}_{\mathbb{I}} \circ \mathcal{H}, \underline{Z}) \leq \mathcal{N}(\mathcal{H}, \underline{X})$)**.** *Let $\mathcal{H} \subseteq \mathcal{Y}^{\mathcal{X}}$. Let $\underline{Z}$ be a $(\mathcal{X} \times \mathcal{Y})$-valued tree,*[26] *and let $\underline{X}$ be the $\mathcal{X}$-valued tree obtained by extracting the $\mathcal{X}$-component from each node of $\underline{Z}$. Then,*

$$\mathcal{N}(\mathcal{L}_{\mathbb{I}} \circ \mathcal{H}, \underline{Z}) \leq \mathcal{N}(\mathcal{H}, \underline{X}).$$

Proof. Let the set $\mathcal{V}$ be the smallest set of trees that form a 0-cover of $\mathcal{H}$ on $\underline{X}$. For each tree $\underline{V} \in \mathcal{V}$, we construct (and add to $\mathcal{W}$) a tree $\underline{W}$, given by

$$w^{(t)}(\epsilon) = \begin{cases} 0, & \text{if } v^{(t)}(\epsilon) = y^{(t)}(\epsilon) \\ 1, & \text{otherwise} \end{cases},$$

where $y^{(t)}(\epsilon)$ is the y-label of the node of $\underline{Z}$ encountered after having traversed the path $(\epsilon_1, \ldots, \epsilon_{t-1})$. We claim that $\mathcal{W}$ provides a 0-cover of $\mathcal{L}_{\mathbb{I}} \circ \mathcal{H}$ on $\underline{Z}$.

As the set $\mathcal{V}$ of trees forms a 0-cover of $\mathcal{H}$ on $\underline{X}$, we know that

$$\forall h \in \mathcal{H}, \forall \epsilon \in \{\pm 1\}^T, \exists \underline{V} \in \mathcal{V},\ \text{s.t.},\ v^{(t)}(\epsilon) = h(x^{(t)}(\epsilon)),$$

for all $t \in [T]$. Therefore, by construction,

$$\forall h \in \mathcal{H}, \forall \epsilon \in \{\pm 1\}^T, \exists \underline{W} \in \mathcal{W},\ \text{s.t.},\ w^{(t)}(\epsilon) = \mathbf{1}[h(x^{(t)}(\epsilon)) \neq y^{(t)}(\epsilon)],$$

[26] The nodes of $\underline{Z}$ are of the form $z = (x, y)$, where $x \in \mathcal{X}$ and $y \in \mathcal{Y}$.

for all $t \in [T]$. So, we obtain

$$\mathcal{N}(\mathcal{L}_{\mathbb{I}} \circ \mathcal{H}, \underline{Z}) \leq |\mathcal{W}| = |\mathcal{V}| = \mathcal{N}(\mathcal{H}, \underline{X}),$$

completing our proof. ■

Lemma 4.18 ($\mathcal{N}(\mathcal{H}, \underline{X}) \leq (Tk)^{\mathrm{mcLdim}(\mathcal{H})}$)**.** *Let $\mathcal{H} \subseteq \mathcal{Y}^{\mathcal{X}}$ with $|\mathcal{Y}| = k > 2$, and $\underline{X}$ be an $\mathcal{X}$-valued tree. Then,*

$$\mathcal{N}(\mathcal{H}, \underline{X}) \leq (Tk)^{mcLdim(\mathcal{H})}.$$

Proof. Our main idea is that the set of experts, as defined in the proof of Theorem 25 in [7] (reproduced below), can be used to construct a 0-cover of $\mathcal{H}$ on $\underline{X}$.

> Given time horizon T, let $A_T = \{A \subset [T] \mid |A| \leq \mathrm{mcLdim}(\mathcal{H})\}$. For every $A \in A_T$ and $\phi : A \to \mathcal{Y}$, we define an expert $E_{A,\phi}$. The expert $E_{A,\phi}$ imitates the SOA algorithm when it errs exactly on the examples $\{x_t \mid t \in A\}$ and the true labels of these examples are determined by ϕ. Formally, the expert $E_{A,\phi}$ proceeds as follows:
>
> Set $V_1 = \mathcal{H}$.
> For $t = 1, 2, \ldots, T$:
> Receive x_t.
> If $t \in A$, set $\hat{y}_t = \phi(t)$.
> If $t \notin A$, set $\hat{y}_t = \mathrm{argmax}_{y \in \mathcal{Y}} \mathrm{mcLdim}(\{h \in V_t : h(x_t) = y\})$.
> Predict $\hat{y}_t$ and update $V_{t+1} = \{h \in V_t : h(x_t) = \hat{y}_t\}$.

Throughout this proof, we use the notation $E_{A,\phi}(x_1, \ldots, x_t)$ to denote the label $\hat{y}_t$ predicted at round t by the expert $E_{A,\phi}$ after processing the sequence of instances $(x_1, \ldots, x_t)$. Similarly, given any true labeling function $h^\star \in \mathcal{H}$, we use the notation $\mathrm{SOA}^{h^\star}(x_1, \ldots, x_t)$ to denote the label predicted at round t by the *standard* SOA algorithm, after processing the sequence of instances $(x_1, \ldots, x_t)$ and updating its version space based on the labels $h^\star(x_1), \ldots, h^\star(x_{t-1})$.

Now, the set of experts $\mathcal{E} = \{E_{A,\phi}\}$ has size $|\mathcal{E}| = \sum_{j=0}^{\mathrm{mcLdim}(\mathcal{H})} \binom{T}{j} k^j \leq (Tk)^{\mathrm{mcLdim}(\mathcal{H})}$ with the following *expert guarantee*:

> For any sequence $(x_1, \ldots, x_T)$ of instances, and any $h \in \mathcal{H}$, there exists an expert $E^\star \in \mathcal{E}$ such that $E^\star(x_1, \ldots, x_t) = h(x_t)$, $\forall\, t \in [T]$.

Indeed, given an arbitrary sequence of instances $(x_1, \ldots, x_T)$ and an arbitrary $h \in \mathcal{H}$, the expert $E^\star = E_{A^\star,\phi^\star}$ that satisfies the guarantee corresponds to $A^\star := \{t \in [T] : \mathrm{SOA}^h(x_1, \ldots, x_t) \neq h(x_t)\}$ and $\phi^\star : A^\star \to \mathcal{Y}$ defined by $\phi^\star(t) := h(x_t)$. Since the SOA algorithm makes at most $\mathrm{mcLdim}(\mathcal{H})$ mistakes, we have $|A^\star| \leq \mathrm{mcLdim}(\mathcal{H})$, and therefore $E^\star \in \mathcal{E}$ by construction.

Next, for each expert $E \in \mathcal{E}$, we add a tree $\underline{V_E}$ to $\mathcal{V}$, given by

$$\underline{V_E}^{(t)}(\epsilon) = E(x^{(1)}(\epsilon), \ldots, x^{(t)}(\epsilon)), \quad \forall\, t \in [T],$$

where $\epsilon = (\epsilon_1, \ldots, \epsilon_T) \in \{\pm 1\}^T$ is a sequence of i.i.d. Rademacher random variables. We now verify that the set of trees $\mathcal{V}$ forms a 0-cover of $\mathcal{H}$ on $\underline{X}$.

Fix an arbitrary $h \in \mathcal{H}$ and an arbitrary $\epsilon = (\epsilon_1, \ldots, \epsilon_T) \in \{\pm 1\}^T$. For the sequence of examples on the path $(x^{(1)}(\epsilon), \ldots, x^{(T)}(\epsilon))$, by the expert guarantee and our construction of

$\mathcal{V}$ above, we have that there exists $E^\star \in \mathcal{E}$ such that $\underline{V_{E^\star}}^{(t)}(\epsilon) = E^\star(x^{(1)}(\epsilon), \ldots, x^{(t)}(\epsilon)) = h(x^{(t)}(\epsilon))$, $\forall t \in [T]$. Therefore, by Definition 4.15, we see that $\mathcal{V}$ forms a 0-cover of $\mathcal{H}$ on $\underline{X}$. Hence,

$$\mathcal{N}(\mathcal{H}, \underline{X}) = \min\{|\mathcal{V}| : \mathcal{V} \text{ is a 0-cover of } \mathcal{H} \text{ on } \underline{X}\} \leq |\mathcal{V}| = |\mathcal{E}| \leq (Tk)^{\text{mcLdim}(\mathcal{H})},$$

as desired. ■

Theorem 4.19 (Upper bound on the expected regret for the classical adversary-provides-a-distribution model in the multiclass agnostic setting)**.** *Let $\mathcal{H} \subseteq \mathcal{Y}^{\mathcal{X}}$, with $|\mathcal{Y}| = k > 2$, be a hypothesis class, and $h^\star \in \mathcal{H}$. For every adversary, there exists a classical online learner for $\mathcal{H}$ that satisfies*

$$\mathbb{E}[\mathcal{R}_P(\mathbf{h}, \mathbf{D}, \mathcal{H}))] = \mathcal{O}(\sqrt{mcLdim(\mathcal{H}) \cdot T \log(Tk)}).$$

Proof. We follow the steps in the proof of Theorem 4.10, which, in a general sense, are applicable in the multiclass setting. However, a key bound used in the proof of Theorem 4.10 ($\mathfrak{R}_T(\mathcal{L}_\mathbb{I} \circ \mathcal{H}) \leq \mathfrak{R}_T(\mathcal{H})$) is not known to continue to hold in the multiclass setting, forcing us to handle the rest of the proof differently. Some interesting insights follow from this deviation, which is elaborated in the proof below, as well as the discussion that follows.

We recall the preliminaries. Let $\mathbf{D}$ be an arbitrary sequence of distributions, and let $\mathbf{z} = (z_1, \ldots, z_t)$ be a sequence of instances such that $z_i \sim D_i$. Then, defining[27] $\mathcal{R}_\mathbb{I}(\mathbf{h}, \mathbf{z}, \mathcal{H}) = \sum_{t=1}^T \mathcal{L}_\mathbb{I}(h_t, z_t) - \inf_{h \in \mathcal{H}} \sum_{t=1}^T \mathcal{L}_\mathbb{I}(h, z_t)$, we have

$$\begin{aligned}\mathcal{R}_P(\mathbf{h}, \mathbf{D}, \mathcal{H}) &= \mathcal{R}_P(\mathbf{h}, \mathbf{D}, \mathcal{H}) - \mathcal{R}_\mathbb{I}(\mathbf{h}, \mathbf{z}, \mathcal{H}) + \mathcal{R}_\mathbb{I}(\mathbf{h}, \mathbf{z}, \mathcal{H}) \\ &= \underbrace{\sum_{t=1}^T \mathcal{L}_P(h_t, D_t) - \sum_{t=1}^T \mathcal{L}_\mathbb{I}(h_t, z_t)}_{\Delta_1} + \underbrace{\inf_{h \in \mathcal{H}} \sum_{t=1}^T \mathcal{L}_\mathbb{I}(h, z_t) - \inf_{h \in \mathcal{H}} \sum_{t=1}^T \mathcal{L}_P(h, D_t)}_{\Delta_2} + \mathcal{R}_\mathbb{I}(\mathbf{h}, \mathbf{z}, \mathcal{H}).\end{aligned}$$

We proceed by bounding the expected value of Δ_1, Δ_2 and $\mathcal{R}_\mathbb{I}(\mathbf{h}, \mathbf{z}, \mathcal{H})$ separately. First, our bound $\mathbb{E}[\Delta_1] \leq \mathcal{O}(\sqrt{T})$ using Azuma-Hoeffding's inequality in Theorem 4.10 is independent of the form of $\mathcal{H}$ and, in particular, continues to hold for a multiclass $\mathcal{H}$.

Next, with regard to $\mathbb{E}[\Delta_2]$, we recover the chain of inequalities in Theorem 4.10 leading up to

$$\mathbb{E}[\Delta_2] \leq 2T \cdot \mathfrak{R}_T(\mathcal{L}_\mathbb{I} \circ \mathcal{H}). \tag{12}$$

However, the result in Theorem 16 of [2] ($\mathfrak{R}_T(\mathcal{L}_\mathbb{I} \circ \mathcal{H}) \leq \mathfrak{R}_T(\mathcal{H})$) only applies when $\mathcal{H}$ is a boolean hypothesis class. Therefore, we proceed with an explicit bound on $\mathfrak{R}_T(\mathcal{L}_\mathbb{I} \circ \mathcal{H})$ using a covering number argument. Let $\underline{Z}$ be an $(\mathcal{X} \times \mathcal{Y})$-valued tree, and let $\underline{X}$ be the $\mathcal{X}$-valued tree obtained by extracting the $\mathcal{X}$-component from each node of $\underline{Z}$. We provide a chain of

[27] $\mathcal{R}_\mathbb{I}(\mathbf{h}, \mathbf{z}, \mathcal{H})$ is the regret of an algorithm in the *multiclass agnostic* generalization of the canonical (adversary-provides-an-input) classical online model in Section 4.1.

inequalities starting from (12):

$$\begin{aligned}
\mathbb{E}[\Delta_2] &\leq 2T \cdot \mathfrak{R}_T(\mathcal{L}_{\mathbb{I}} \circ \mathcal{H}) \\
&\leq \frac{24T\sqrt{\log \mathcal{N}(\mathcal{L}_{\mathbb{I}} \circ \mathcal{H}, \underline{Z})}}{\sqrt{T}} \\
&\leq \frac{24T\sqrt{\log \mathcal{N}(\mathcal{H}, \underline{X})}}{\sqrt{T}} \\
&\leq \frac{24T\sqrt{\mathrm{mcLdim}(\mathcal{H}) \cdot \log(Tk)}}{\sqrt{T}} \\
&= \mathcal{O}(\sqrt{\mathrm{mcLdim}(\mathcal{H}) \cdot T \log(Tk)}),
\end{aligned}$$

where the second line is from Theorem 3 and Definition 5 of [34], the third line is from Lemma 4.17, and the fourth line is from Lemma 4.18. Finally, from Theorem 4 of [9], we have $\mathcal{R}_{\mathbb{I}}(\mathbf{h}, \mathbf{z}, \mathcal{H}) = \mathcal{O}\Big(\sqrt{\mathrm{mcLdim}(\mathcal{H}) \cdot T \log\left(\frac{T}{\mathrm{mcLdim}(\mathcal{H})}\right)}\Big)$, when $T \geq 2 \cdot \mathrm{mcLdim}(\mathcal{H})$. Putting everything together, using the linearity of expectation, we have,

$$\begin{aligned}
\mathbb{E}[\mathcal{R}_P(\mathbf{h}, \mathbf{D}, \mathcal{H})] &= \mathbb{E}[\Delta_1 + \Delta_2 + \mathcal{R}_{\mathbb{I}}(\mathbf{h}, \mathbf{z}, \mathcal{H})] \\
&= \mathbb{E}[\Delta_1] + \mathbb{E}[\Delta_2] + \mathbb{E}[\mathcal{R}_{\mathbb{I}}(\mathbf{h}, \mathbf{z}, \mathcal{H})] \\
&= \mathcal{O}(\sqrt{T}) + \mathcal{O}(\sqrt{\mathrm{mcLdim}(\mathcal{H}) \cdot T \log(Tk)}) \\
&\qquad + \mathcal{O}\left(\sqrt{\mathrm{mcLdim}(\mathcal{H}) \cdot T \log\left(\frac{T}{\mathrm{mcLdim}(\mathcal{H})}\right)}\right) \\
&= \mathcal{O}(\sqrt{\mathrm{mcLdim}(\mathcal{H}) \cdot T \log(Tk)}),
\end{aligned}$$

where the third line holds as point-wise bounds imply bounds in-expectation[28], and the last line holds as $k > 2$ and $\mathrm{mcLdim}(\mathcal{H}) \geq 1 \implies k \geq \frac{1}{\mathrm{mcLdim}(\mathcal{H})}$. ■

In summary, our theorem provides a k-dependent upper bound on $\mathbb{E}[\mathcal{R}_P(\mathbf{h}, \mathbf{D}, \mathcal{H})]$, while the corresponding lower bound (presented next, Theorem 4.20), is k-independent. Although the optimal classical learner for the canonical multiclass agnostic model bridges this gap, as indicated by the k-independent bound used on $\mathbb{E}[\mathcal{R}_{\mathbb{I}}(\mathbf{h}, \mathbf{z}, \mathcal{H})]$, it remains unclear whether we can establish a *k-independent* upper bound on $\mathbb{E}[\mathcal{R}_P(\mathbf{h}, \mathbf{D}, \mathcal{H})]$. Our current proof strategy[29] faces challenges in achieving this goal due to the following observation. The bound on $\mathbb{E}[\Delta_2]$ is guaranteed by the rate of online uniform convergence of the loss $(\mathcal{L}_{\mathbb{I}} \circ \mathcal{H})$ class (in the frameworks of sequential Rademacher complexity [34] and adversarial (uniform) laws of large numbers [18]). Meanwhile, the bound on $\mathbb{E}[\mathcal{R}_{\mathbb{I}}(\mathbf{h}, \mathbf{z}, \mathcal{H})]$ is guaranteed by the rate of canonical (agnostic) online learnability of $\mathcal{H}$. However, in the multiclass function class case, as demonstrated by Theorem 7 and Example 1 in [9], these two rates are *not* equivalent.

[28] Since the point-wise bound holds for $T \geq 2 \cdot \mathrm{mcLdim}(\mathcal{H})$, it is clear that $\mathbb{E}[\mathcal{R}_{\mathbb{I}}(\mathbf{h}, \mathbf{z}, \mathcal{H})] \leq 2 \cdot \mathrm{mcLdim}(\mathcal{H}) + \mathcal{O}\Big(\sqrt{\mathrm{mcLdim}(\mathcal{H}) \cdot T \log\left(\frac{T}{\mathrm{mcLdim}(\mathcal{H})}\right)}\Big)$, where the latter term dominates when $T \geq 4 \cdot \mathrm{mcLdim}(\mathcal{H})$.

[29] Currently, we employ the optimal classical learner for the canonical multiclass agnostic model by presenting it with the observed sequence of instances $\mathbf{z}$ in the adversary-provides-a-distribution protocol and evaluate its performance under the adversary-provides-a-distribution framework.

Theorem 4.20 (Lower bound on the expected regret for the classical adversary-provides-a-distribution model in the multiclass agnostic setting)**.** *Let $\mathcal{H} \subseteq \mathcal{Y}^{\mathcal{X}}$, with $|\mathcal{Y}| = k > 2$, be a hypothesis class. For every classical online learner of $\mathcal{H}$, there exists an adversary such that*

$$\mathbb{E}[\mathcal{R}_P(\mathbf{h}, \mathbf{D}, \mathcal{H})] = \Omega(\sqrt{mcLdim(\mathcal{H}) \cdot T}).$$

Proof. As in the online lower bound proofs thus far, we consider an adversary which chooses each D_t to be a point mass on the instance space, i.e., the adversary simply chooses an instance z_t at each t. Since each $z_t \sim D_t$ is deterministic, we have

$$\mathbb{E}[\mathcal{R}_P(\mathbf{h}, \mathbf{D}, \mathcal{H})] = \mathbb{E}[\mathcal{R}_{\mathbb{I}}(\mathbf{h}, \mathbf{z}, \mathcal{H})] = \Omega(\sqrt{\text{mcLdim}(\mathcal{H}) \cdot T}),$$

where the second equality is due to Theorem 26 in [7]. ■

Due to Theorems 4.19 and 4.20, we have characterized multiclass agnostic learnability in the adversary-provides-a-distribution setting for the *bounded* label space ($k < \infty$) case.

5 Quantum Online Learning

Equipped with our models in Sections 4.2 and 4.3, we are finally ready to introduce our quantum online learning model. However, prior to the model description, we clarify our scope. In its nascent existence, quantum online learning has primarily focused on the online learning of *quantum states* [19, 35, 36]. In contrast, our focus in this paper is on the online learning of *classical functions* via quantum examples. Our scope is motivated by the abundance of classical online learning literature [7, 16, 17, 32], as well as our results in Sections 4.2, 4.3 and 4.4, that presents us with an at-the-ready comparison.

5.1 Model Description

Let $\mathcal{H} \subseteq \mathcal{Y}^{\mathcal{X}}$ be a hypothesis class. Identifying the T-round protocol in Section 4.2 (corresp. Section 4.3) with the definition of a quantum example in (1) (corresp. (2)), we obtain the following "natural" model for quantum online learning. The T-round protocol proceeds as follows: at the t-th round,

1. Learner provides a hypothesis $h_t : \mathcal{X} \to \mathcal{Y}$.
2. Adversary reveals an example $|\psi_t\rangle$ where
   (a) $|\psi_t\rangle = \sum_{x \in \mathcal{X}} \sqrt{D_t(x)}\, |x, h^\star(x)\rangle$ for some $D_t : \mathcal{X} \to [0,1]$ and $h^\star \in \mathcal{H}$ (realizable),
   (b) $|\psi_t\rangle = \sum_{x \in \mathcal{X},\, y \in \mathcal{Y}} \sqrt{D_t(x,y)}\, |x, y\rangle$ for some $D_t : \mathcal{X} \times \mathcal{Y} \to [0,1]$ (agnostic[30]).
3. Learner incurs loss[31]
   (a) $\mathcal{L}_P(h_t, D_t, h^\star) := \mathbb{P}_{x \sim D_t}(h_t(x) \neq h^\star(x))$ (realizable),

[30] As in the classical case, the adversary need not be consistent: i.e., they could reveal, for e.g., both $|x, 0\rangle$ and $|x, 1\rangle$ during the T-round protocol.

[31] As an aside, for those who favor a mistake model, it is possible to define it by specifying a threshold ϵ. In this case, a mistake occurs in a round iff $\mathcal{L}_P > \epsilon$, i.e. $\mathcal{L}^{\epsilon}_{\mathbb{I}} = \mathbf{1}[\mathcal{L}_P > \epsilon]$. We do not investigate this mistake model.

(b) $\mathcal{L}_P(h_t, D_t) := \mathbb{P}_{(x,y)\sim D_t}(h_t(x) \neq y)$ (agnostic).

As in the classical adversary-provides-a-distribution models, the learner's total loss in the realizable case continues to be given by $\mathcal{L}_P(\mathbf{h}, \mathbf{D}, h^\star) = \sum_{t=1}^{T} \mathcal{L}_P(h_t, D_t, h^\star)$ (ref. (7)), while in the agnostic case, the learner's total regret continues to be expressed as $\mathcal{R}_P(\mathbf{h}, \mathbf{D}, \mathcal{H}) = \sum_{t=1}^{T} \mathcal{L}_P(h_t, D_t) - \inf_{h\in\mathcal{H}} \sum_{t=1}^{T} \mathcal{L}_P(h, D_t)$ (ref. (10)).

For this model, we formally define what a quantum online learner, a quantum adversary, and quantum online learnability entails. While these definitions are analogous to Definitions 4.3, 4.4, and 4.5, we present them here for the sake of completeness.

Definition 5.1 (Quantum online learner)**.** *An algorithm $\mathcal{A}$ is a quantum online learner for a hypothesis class $\mathcal{H}$ if having received a sequence of quantum examples $(|\psi_i\rangle)_{i=1}^{t}$ (of the form in 2. (a) or 2. (b) of Section 5.1) over the first t rounds, $\mathcal{A}$ outputs a hypothesis $h_{t+1} : \mathcal{X} \to \mathcal{Y}$ at round*[32] *$t+1$.*

Definition 5.2 (Quantum adversary)**.** *Having received a sequence of hypothesis $\mathbf{h}|_t = (h_1, \ldots, h_t)$ from the learner, and with knowledge of its own prior choices of quantum examples, $(|\psi_i\rangle)_{i=1}^{t}$, over the first t rounds, at round $t+1$, a quantum (online) adversary chooses a distribution D_{t+1} (on $\mathcal{X}$ (realizable) or on $\mathcal{X} \times \mathcal{Y}$ (agnostic)) and discloses the corresponding quantum example $|\psi_{t+1}\rangle$ (with consistent labeling throughout the protocol in the realizable case) to the learner.*

Definition 5.3 (Quantum online learnability)**.** *A hypothesis class $\mathcal{H}$ is quantum online learnable if there exists a quantum online learning algorithm $\mathcal{A}$ such that*

- $\mathcal{L}_P(\mathbf{h}_\mathcal{A}, \mathcal{H}) = \sup_{\mathbf{D},\, h^\star\in\mathcal{H}} \mathbb{E}[\mathcal{L}_P(\mathbf{h}_\mathcal{A}, \mathbf{D}, h^\star)] = o(T)$ *(realizable),*
- $\mathcal{R}_P(\mathbf{h}_\mathcal{A}, \mathcal{H}) = \sup_{\mathbf{D}} \mathbb{E}[\mathcal{R}_P(\mathbf{h}_\mathcal{A}, \mathbf{D}, \mathcal{H})] = o(T)$ *(agnostic).*

5.2 Binary Classification

Under the quantum online learning model described above in Section 5.1, we bound the expected regret (in both the realizable and the agnostic cases) of a quantum online learner for a boolean hypothesis class.

Theorem 5.4 (Lower bounds on expected loss/regret for quantum online binary classification)**.** *Let $\mathcal{H} \subseteq \{0,1\}^\mathcal{X}$, be a hypothesis class, and $h^\star \in \mathcal{H}$. For every quantum online learner of $\mathcal{H}$, there exists a quantum adversary such that*

$$\mathbb{E}[\mathcal{L}_P(\mathbf{h}, \mathbf{D}, h^\star)] = \Omega(Ldim(\mathcal{H})) \quad \textit{(realizable), and}$$
$$\mathbb{E}[\mathcal{R}_P(\mathbf{h}, \mathbf{D}, \mathcal{H})] = \Omega(\sqrt{Ldim(\mathcal{H}) \cdot T}) \quad \textit{(agnostic).}$$

Proof. Let $\mathcal{A}_Q$ be an arbitrary, but fixed, *quantum* online learning algorithm for $\mathcal{H}$. We proceed using a reduction argument. To do this, we examine the scenario where a classical adversary chooses D_t to be a point mass for each t (i.e. the adversary simply chooses an instance x_t (realizable) or $z_t = (x_t, y_t)$ (agnostic) at each t), and analyze the loss/regret bound for the following classical learner $\mathcal{A}_C$ that accesses $\mathcal{A}_Q$ as a "black box". At the t-th round,

[32] Prior to receiving any examples, $\mathcal{A}$ outputs some arbitrary hypothesis $h_1 : \mathcal{X} \to \mathcal{Y}$ at round 1.

1. $\mathcal{A}_C$ provides hypothesis h_t^Q (received from $\mathcal{A}_Q$ in the previous round).
2. Adversary reveals (x_t, y_t) to $\mathcal{A}_C$ (in the realizable case, $y_t = h^\star(x_t)$ for some $h^\star \in \mathcal{H}$).
3. $\mathcal{A}_C$ state prepares $|\psi_t\rangle = |x_t, y_t\rangle$ and passes it as input to $\mathcal{A}_Q$.
4. $\mathcal{A}_Q$ outputs hypothesis h_{t+1}^Q to $\mathcal{A}_C$.

We provide a bound first for the realizable case. Since $\mathcal{A}_C$ plays h_t^Q at each t, it is clear, for our setup, that $\mathcal{L}_P(\mathbf{h}_{\mathcal{A}_C}, \mathbf{D}, h^\star) = \mathcal{L}_{\mathbb{I}}(\mathbf{h}_{\mathcal{A}_C}, \mathbf{x}, h^\star) \leq \mathcal{L}_{\mathbb{I}}(\mathbf{h}_{\mathcal{A}_Q}, \mathbf{x}, h^\star) = \mathcal{L}_P(\mathbf{h}_{\mathcal{A}_Q}, \mathbf{D}, h^\star)$. Taking expectations, and noting that $\mathbb{E}[\mathcal{L}_P(\mathbf{h}_{\mathcal{A}_C}, \mathbf{D}, h^\star)] = \Omega(\text{Ldim}(\mathcal{H}))$ (from Theorem 4.7), we have shown $\mathbb{E}[\mathcal{L}_P(\mathbf{h}_{\mathcal{A}_Q}, \mathbf{D}, h^\star)] = \Omega(\text{Ldim}(\mathcal{H}))$. Since, $\mathcal{A}_Q$ was chosen arbitrarily, we deduce that $\mathbb{E}[\mathcal{L}_P(\mathbf{h}, \mathbf{D}, h^\star)] = \Omega(\text{Ldim}(\mathcal{H}))$.

The agnostic case follows an identical argument; we obtain the following chain of (in)equalities, $\mathcal{R}_P(\mathbf{h}_{\mathcal{A}_C}, \mathbf{D}, \mathcal{H}) = \mathcal{R}_{\mathbb{I}}(\mathbf{h}_{\mathcal{A}_C}, \mathbf{z}, \mathcal{H}) \leq \mathcal{R}_{\mathbb{I}}(\mathbf{h}_{\mathcal{A}_Q}, \mathbf{z}, \mathcal{H}) = \mathcal{R}_P(\mathbf{h}_{\mathcal{A}_Q}, \mathbf{D}, \mathcal{H})$. Taking expectations, and noting $\mathbb{E}[\mathcal{R}_P(\mathbf{h}_{\mathcal{A}_C}, \mathbf{D}, \mathcal{H})] = \Omega(\sqrt{\text{Ldim}(\mathcal{H}) \cdot T})$ (from Theorem 4.11) and that $\mathcal{A}_Q$ was chosen arbitrarily, we deduce $\mathbb{E}[\mathcal{R}_P(\mathbf{h}, \mathbf{D}, \mathcal{H})] = \Omega(\sqrt{\text{Ldim}(\mathcal{H}) \cdot T})$. ■

Theorem 5.5 (Upper bounds on expected loss/regret for quantum online binary classification)**.** *Let $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$, be a hypothesis class, and $h^\star \in \mathcal{H}$. For every quantum adversary, there exists a quantum online learner for $\mathcal{H}$ that satisfies*

$$\mathbb{E}[\mathcal{L}_P(\mathbf{h}, \mathbf{D}, h^\star)] = \mathcal{O}(Ldim(\mathcal{H})) \quad \textit{(realizable), and}$$
$$\mathbb{E}[\mathcal{R}_P(\mathbf{h}, \mathbf{D}, \mathcal{H})] = \mathcal{O}(\sqrt{Ldim(\mathcal{H}) \cdot T}) \quad \textit{(agnostic).}$$

Proof. For a naïve algorithm that, at each round t, measures $|\psi_t\rangle$ in the standard basis and employs a classical learner to learn from the observed classical outputs, the desired upper bounds are guaranteed by Theorems 4.6 and 4.10. ■

5.3 Multiclass Classification

Here, we present bounds on the expected regret (in both the realizable and the agnostic cases) of a quantum online learner for a multiclass hypothesis class.

Theorem 5.6 (Lower bounds on expected loss/regret for quantum online multiclass classification)**.** *Let $\mathcal{H} \subseteq \mathcal{Y}^{\mathcal{X}}$, with $|\mathcal{Y}| = k > 2$, be a hypothesis class, and $h^\star \in \mathcal{H}$. For every quantum online learner of $\mathcal{H}$, there exists a quantum adversary such that*

$$\mathbb{E}[\mathcal{L}_P(\mathbf{h}, \mathbf{D}, h^\star)] = \Omega(mcLdim(\mathcal{H})) \quad \textit{(realizable), and}$$
$$\mathbb{E}[\mathcal{R}_P(\mathbf{h}, \mathbf{D}, \mathcal{H})] = \Omega(\sqrt{mcLdim(\mathcal{H}) \cdot T}) \quad \textit{(agnostic).}$$

Proof. The proof is identical to that of Theorem 5.4, where now, for the corresponding classical learners, $\mathbb{E}[\mathcal{L}_P(\mathbf{h}_{\mathcal{A}_C}, \mathbf{D}, h^\star)] = \Omega(\text{mcLdim}(\mathcal{H}))$ (from Theorem 4.14), and $\mathbb{E}[\mathcal{R}_P(\mathbf{h}_{\mathcal{A}_C}, \mathbf{D}, \mathcal{H})] = \Omega(\sqrt{\text{mcLdim}(\mathcal{H}) \cdot T})$ (from Theorem 4.20). ■

Theorem 5.7 (Upper bounds on expected loss/regret for quantum online multiclass classification)**.** *Let $\mathcal{H} \subseteq \mathcal{Y}^{\mathcal{X}}$, with $|\mathcal{Y}| = k > 2$, be a hypothesis class, and $h^\star \in \mathcal{H}$. For every quantum adversary, there exists a quantum online learner for $\mathcal{H}$ that satisfies*

$$\mathbb{E}[\mathcal{L}_P(\mathbf{h}, \mathbf{D}, h^\star)] = \mathcal{O}(mcLdim(\mathcal{H})) \quad \textit{(realizable), and}$$
$$\mathbb{E}[\mathcal{R}_P(\mathbf{h}, \mathbf{D}, \mathcal{H})] = \mathcal{O}(\sqrt{mcLdim(\mathcal{H}) \cdot T \log(Tk)}) \quad \textit{(agnostic).}$$

Proof. Once again, we consider the measure-and-learn-classically quantum learner for which the desired upper bounds are now guaranteed by Theorems 4.13 and 4.19. ■

5.4 Takeaways

Before we end this section on online learning with quantum examples, we note that the proofs for the expected regret upper bounds were established by a quantum online learner that performs a measurement and subsequently learns classically. The fact that the upper bounds thus obtained are *identical* to the lower bounds, in all but one setting[33], shows that the performance of this measure-and-learn-classically learner is as good as the best "genuine" quantum online learner in these settings. We feel that this is consistent with the overall message of this paper, viz., that there is limited power in quantum examples to speed up learning especially when the adversary is allowed to play arbitrary distributions (including very degenerate ones like point masses).

Recently, [9] improved the classical upper bound for the online multiclass agnostic case to $\tilde{\mathcal{O}}(\sqrt{\mathrm{mcLdim}(\mathcal{H})T})$[34], which removes all k dependence (cf. the $\sqrt{\log k}$ factor that appears in our corresponding *quantum* upper bound in Theorem 5.7). Meanwhile, we believe our analysis in the proof of Theorem 4.19 (which establishes the classical bound for the measure-and-learn-classically quantum learner in Theorem 5.7) is tight, and so we suspect that any removal of the k-dependence in this setting would involve investigating into a "genuine" quantum online learning algorithm, which may involve a quantum-specific combinatorial parameter that characterizes learning. We identify this as an open question for future work.

- What is the *tight* expected regret bound for quantum online multiclass agnostic learning when the label space is unbounded (i.e., when the number of classes $k \to \infty$)?

6 Conclusion

In this work, we partially resolved an open question of [1] by characterizing the sample complexity of multiclass learning (for $2 < k < \infty$). With recent work [8] fully characterizing *classical* multiclass learnability (including the case when $k \to \infty$) via the DS dimension, we ask whether *quantum* multiclass learnability is also fully characterized by the DS dimension. We know that the upper bound in [8] also holds in the quantum case by *measure-and-learn-classically*. However, since the classical lower bound involving the DS dimension (Theorem 2 of [31]) uses *transductive learning* which has no clear analog for a quantum example (ref. (1) and (2)), providing a quantum lower bound involving the DS dimension has proved to be non-trivial. We identify this as an open question for future work.

- What is the *tight* quantum sample complexity bound for batch multiclass learning, in both the realizable and agnostic settings, when the label space is unbounded (i.e., when $k \to \infty$)?

[33] The exception is the online multiclass agnostic case, where the quantum upper and lower bounds differ by a factor of $\sqrt{\log(Tk)}$.

[34] Here, $\tilde{\mathcal{O}}(\cdot)$ hides $\sqrt{\log\left(\frac{T}{\mathrm{mcLdim}(\mathcal{H})}\right)}$ factors.

In the batch setting, the sample complexity upper bounds were trivial to establish due to the quantum learner's ability to measure quantum examples and learn classically on the resulting output. In the online setting, the expected regret lower bounds, in turn, were trivial due to the adversary's ability to provide point masses D_t at each t, rendering each quantum example equivalent to a classical example. This prompts us to ask the following question.

- What happens when we impose restrictions on D_t to force it away from a point mass? Would the expected regret bounds for the canonical classical online model (Section 4.1), classical adversary-provides-a-distribution model (Sections 4.2 and 4.3), and the quantum online model in Section 5.1 all diverge from one another?

References

[1] Srinivasan Arunachalam and Ronald de Wolf. "Optimal Quantum Sample Complexity of Learning Algorithms". Journal of Machine Learning Research **19**, 1–36 (2018). url: jmlr.org/papers/v19/18-195.html.

[2] Philip Dawid and Ambuj Tewari. "On Learnability under General Stochastic Processes". Harvard Data Science Review **4** (2022).

[3] Nader H. Bshouty and Jeffrey C. Jackson. "Learning DNF over the uniform distribution using a quantum example oracle". In Proceedings of the 8th Conference on Computational Learning Theory (COLT). Pages 118–127. ACM (1995).

[4] Rocco A. Servedio and Steven J. Gortler. "Equivalences and Separations Between Quantum and Classical Learnability". SIAM Journal on Computing **33**, 1067–1092 (2004).

[5] Alp Atıcı and Rocco A. Servedio. "Improved Bounds on Quantum Learning Algorithms". Quantum Information Processing **4**, 355–386 (2005).

[6] Chi Zhang. "An improved lower bound on query complexity for quantum PAC learning". Information Processing Letters **111**, 40–45 (2010).

[7] Amit Daniely, Sivan Sabato, Shai Ben-David, and Shai Shalev-Shwartz. "Multiclass Learnability and the ERM Principle". Journal of Machine Learning Research **16**, 2377–2404 (2015). url: jmlr.org/papers/v16/daniely15a.html.

[8] Nataly Brukhim, Daniel Carmon, Irit Dinur, Shay Moran, and Amir Yehudayoff. "A Characterization of Multiclass Learnability". In Proceedings of the 63rd Symposium on Foundations of Computer Science (FOCS). Pages 943–955. IEEE (2022). arXiv:2203.01550.

[9] Steve Hanneke, Shay Moran, Vinod Raman, Unique Subedi, and Ambuj Tewari. "Multiclass Online Learning and Uniform Convergence". In Proceedings of 36th Conference on Learning Theory (COLT). Pages 5682–5696. PMLR (2023). url: proceedings.mlr.press/v195/hanneke23b.html.

[10] Anselm Blumer, Andrzej Ehrenfeucht, David Haussler, and Manfred K. Warmuth. "Learnability and the Vapnik-Chervonenkis dimension". Journal of the ACM **36**, 929–965 (1989).

[11] Steve Hanneke. "The Optimal Sample Complexity of PAC Learning". Journal of Machine Learning Research **17**, 1–15 (2016). url: jmlr.org/papers/v17/15-389.html.

[12] Michael J. Kearns, Robert E. Schapire, and Linda M. Sellie. "Toward efficient agnostic learning". In Proceedings of the 5th Workshop on Computational Learning Theory (COLT). Pages 341–352. ACM (1992).

[13] Michel Talagrand. "Sharper Bounds for Gaussian and Empirical Processes". The Annals of Probability **22**, 28–76 (1994).

[14] Balas K. Natarajan. "On Learning Sets and Functions". Machine Learning **4**, 67–97 (1989).

[15] Shai Ben-David, Nicolo Cesa-Bianchi, David Haussler, and Philip M. Long. "Characterizations of Learnability for Classes of $\{0, ..., n\}$-Valued Functions". Journal of Computer and System Sciences **50**, 74–86 (1995).

[16] Nick Littlestone. "Learning Quickly When Irrelevant Attributes Abound: A New Linear-Threshold Algorithm". Machine Learning **2**, 285–318 (1988).

[17] Shai Ben-David, Dávid Pál, and Shai Shalev-Shwartz. "Agnostic Online Learning". In Proceedings of the 22nd Conference on Learning Theory (COLT). (2009). url: api.semanticscholar.org/bendavid09agnostic.

[18] Noga Alon, Omri Ben-Eliezer, Yuval Dagan, Shay Moran, Moni Naor, and Eylon Yogev. "Adversarial Laws of Large Numbers and Optimal Regret in Online Classification". In Proceedings of the 53rd ACM SIGACT Symposium on Theory of Computing (STOC). Pages 447–455. ACM (2021). arXiv:2101.09054.

[19] Anurag Anshu and Srinivasan Arunachalam. "A survey on the complexity of learning quantum states". Nature Reviews Physics **6**, 59–69 (2024).

[20] Srinivasan Arunachalam and Ronald de Wolf. "Guest Column: A Survey of Quantum Learning Theory". ACM SIGACT News **48**, 41–67 (2017).

[21] Wilfred Salmon, Sergii Strelchuk, and Tom Gur. "Provable Advantage in Quantum PAC Learning". In Proceedings of the 37th Conference on Learning Theory (COLT). Pages 4487–4510. PMLR (2024). url: proceedings.mlr.press/v247/salmon24a.html.

[22] Srinivasan Arunachalam, Aleksandrs Belovs, Andrew M. Childs, Robin Kothari, Ansis Rosmanis, and Ronald de Wolf. "Quantum Coupon Collector". In Proceedings of the 15th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC). Pages 10:1–10:17. Schloss Dagstuhl – Leibniz-Zentrum für Informatik (2020).

[23] Ashwin Nayak and Pulkit Sinha. "Proper vs Improper Quantum PAC Learning" (2024). arXiv:2403.03295.

[24] Olivier Bousquet, Steve Hanneke, Shay Moran, and Nikita Zhivotovskiy. "Proper Learning, Helly Number, and an Optimal SVM Bound". In Proceedings of the 33rd Conference on Learning Theory (COLT). Pages 582–609. PMLR (2020). url: proceedings.mlr.press/v125/bousquet20a.html.

[25] Nika Haghtalab, Tim Roughgarden, and Abhishek Shetty. "Smoothed Analysis of Online and Differentially Private Learning". In Proceedings of the 34th Conference on Neural Information Processing Systems (NeurIPS). Pages 9203–9215. Curran Associates, Inc. (2020). url: proceedings.neurips.cc/haghtalab20smoothed.

[26] Adam Block, Yuval Dagan, Noah Golowich, and Alexander Rakhlin. "Smoothed Online Learning is as Easy as Statistical Learning". In Proceedings of the 35th Conference on Learning Theory (COLT). Pages 1716–1786. PMLR (2022). url: proceedings.mlr.press/v178/block22a.html.

[27] Gábor Lugosi and Gergely Neu. "Online-to-PAC Conversions: Generalization Bounds via Regret Analysis" (2024). arXiv:2305.19674.

[28] Michael A. Nielsen and Isaac L. Chuang. "Quantum Computation and Quantum Information: 10th Anniversary Edition". Cambridge University Press. (2010).

[29] Leslie G. Valiant. "A Theory of the Learnable". Communications of the ACM **27**, 1134–1142 (1984).

[30] Shima Bab Hadiashar, Ashwin Nayak, and Pulkit Sinha. "Optimal Lower Bounds for Quantum Learning via Information Theory". IEEE Transactions on Information Theory **70**, 1876–1896 (2024).

[31] Amit Daniely and Shai Shalev-Shwartz. "Optimal Learners for Multiclass Problems". In Proceedings of the 27th Conference on Learning Theory (COLT). Pages 287–316. PMLR (2014). url: proceedings.mlr.press/v35/daniely14b.html.

[32] Shai Shalev-Shwartz and Shai Ben-David. "Understanding Machine Learning: From Theory to Algorithms". Cambridge University Press. (2014).

[33] Alina Beygelzimer, John Langford, Lihong Li, Lev Reyzin, and Robert Schapire. "Contextual Bandit Algorithms with Supervised Learning Guarantees". In Proceedings of the 14th International Conference on Artificial Intelligence and Statistics (AISTATS). Pages 19–26. PMLR (2011). url: proceedings.mlr.press/v15/beygelzimer11a.html.

[34] Alexander Rakhlin, Karthik Sridharan, and Ambuj Tewari. "Sequential complexities and uniform martingale laws of large numbers". Probability Theory and Related Fields **161**, 111–153 (2015).

[35] Scott Aaronson, Xinyi Chen, Elad Hazan, Satyen Kale, and Ashwin Nayak. "Online learning of quantum states". Journal of Statistical Mechanics: Theory and Experiment **2019**, 124019 (2019).

[36] Yihui Quek, Srinivasan Arunachalam, and John A Smolin. "Private learning implies quantum stability". In Proceedings of the 35th International Conference on Neural Information Processing Systems (NeurIPS). Pages 20503–20515. (2021). url: openreview.net/quek21private.