Quantum Local Differential Privacy and Quantum Statistical Query Model

Armando Angrisani, Elham Kashefi

Quantum Local Differential Privacy and Quantum Statistical Query Model

Armando Angrisani[1], Elham Kashefi[1,2]

[1]*LIP6, CNRS, Sorbonne Université, 75005 Paris, France*
[2]*School of Informatics, University of Edinburgh, EH8 9AB Edinburgh, United Kingdom*

Abstract

The problem of private learning has been extensively studied in classical computer science. Notably, a striking equivalence between local differentially private learning and statistical query learning has been shown. In addition, the statistical query model has been recently extended to quantum computation. In this work, we give a formal definition of quantum local differential privacy and we extend the aforementioned result to quantum computation.

1 Introduction

Making predictions based on empirical observations is a central task in many scientific fields and it is at the hearth of statistical learning theory. A fundamental tool for the analysis of learning algorithms is undoubtedly the *probably approximately correct* (PAC) model, introduced by [1]. In the classical PAC model of learning, the goal is to learn a collection of Boolean functions $\mathcal{C} \subseteq \{c : \{0,1\}^d \to \{0,1\}\}$. A learner is given in input *labelled examples* $\{x_i, c(x_i)\}$, where x is drawn from a (possibly arbitrary) distribution $\mathcal{X} : \{0,1\}^d \to [0,1]$ and $c \in C$ is a *target concept*. Given two parameters $\varepsilon, \delta \in (0,1)$, the goal of the learner is to output a hypothesis h such that $\Pr_{x\sim\mathcal{X}}[h(x) \neq c(x)] \leq \varepsilon$ with probability at least $1-\delta$, for any choice of c and $\mathcal{X}$. Several extension of the PAC model have been proposed. In particular, [2] introduced the *quantum* PAC model, where the classical labelled examples are replace by the following *quantum example*

$$|\psi_c\rangle = \sum_{x\in\{0,1\}^d} \sqrt{\mathcal{X}(x)}\,|x, c(x)\rangle\,,$$

which is a quantum *superposition* of labelled examples. Note that by measuring the above state in the computational basis we obtain the labelled example $\{x, c(x)\}$ with probability $\mathcal{X}(x)$. A key question is whether quantum learners may be able to learn concepts with less examples than is possible classically. Early results in this direction were both positive and negative, with the distribution from which the examples are sampled being a crucial ingredient. For example, it was shown in [2, 3, 4, 5] that exponential advantages for PAC learners were possible under the uniform distribution. On the other hand, if the distribution is arbitrary, there is only a marginal improvement that quantum samples can hope to provide [6].

Statistical query model. In the (classical) *statistical query* (SQ) model [7], instead of accessing examples directly, the learner can specify some function on the examples, for which he is given an estimate, up an additive perturbation error, of their expectation with respect to the distribution $\mathcal{X}$. PAC learning is strictly stronger than the SQ learning, as shown in [7]. The SQ was extended to quantum computation in [8]. In the *quantum statistical query* (QSQ) model, the learner is still a classical randomized algorithm and can query an oracle to obtain statistics of quantum examples. More concretely, the learner specifies an *observable* M and receives an estimate of the expectation value $\langle\psi_c|M|\psi_c\rangle$. When M is diagonal, this reduces to the classical

SQ model. The QSQ model is considerably more powerful than its classical counterpart. As shown in [8], it is possible to learn parity functions, juntas and DNF formulas under the uniform distribution in polynomial time in the QSQ model, while this is provably hard in the classical SQ model.

Differential privacy. Many of the dataset processed by machine learning algorithms contains sensitive information, namely biometric or financial data. To this end, a mathematical notion of privacy, known as *differential privacy* (DP) [9, 10] has been extensively studied. Differentially private mechanisms ensure that the final output of an algorithm does not depend too heavily on any one input or specific training example. In the standard model, a trusted curator collects the raw data of the individuals and it's responsible of their privacy. On the contrary, in the *local* model the curator is possibly malicious, and hence each individual submits her own privatized data. More formally, consider a statistical database, i.e. a vector $x = (x_1, \ldots, x_n)$ over a domain X, where each entry $x_i \in X$ represents information contributed by one individual. Databases x and x' are neighbors if $x_i \neq x'_i$ for exactly one $i \in [n]$. A randomized algorithm $\mathcal{A}$ is α-differentially private if for any two neighbor databases x, x' and for every subset F of the possible outcomes of $\mathcal{A}$ we have

$$\Pr[\mathcal{A}(x) \in F] \leq e^{\alpha} \Pr[\mathcal{A}(x') \in F].$$

We now turn our attention to the *local model*. Following the notation used in [11], we say that a randomized algorithm over databases is α-local differentially private if it's an α-differentially private algorithm that takes in input a database of size $n = 1$.

As shown in [11], there's a striking equivalence between local differentially private learning and SQ learning. Differential privacy has been extended to quantum computation in [12]. Let σ and ρ be two quantum mixed states on n registers each. We call them *neighbors* if it's possible to reach either σ from ρ, or ρ from σ, by performing a general quantum operation on a single register only. Given a set S of quantum mixed states each on n registers, a measurement M and a parameter $\alpha \geq 0$, we define M to be α-DP on S if for all states $\rho, \sigma \in S$ that are neighbors, and all possible outputs y of M, we have

$$\Pr[M(\rho) = y] \geq e^{\alpha} \Pr[M(\sigma) = y].$$

Interestingly, the authors of [12] showed that the definition above is connected to quantum gentle measurements, and provided an application to quantum shadow tomography. An alternative notion of quantum differential privacy was instead proposed in [13].

Our contribution. In this work, we give a formal definition of quantum local differential privacy and extend the definition of quantum statistical query to mixed states. Thus we prove an equivalence between these two models, extending the classical result of [11] and addressing an open problem posed in [14].

2 Preliminaries

We use $[n]$ to denote the set $\{1, 2, \ldots, n\}$ and $\mathbb{E}[\cdot]$ to denote the expectation of a random variable. Let $\mathbb{I}$ be the d-dimensional identity matrix. We will omit the dimension d when it is clear from the context.

2.1 Quantum computing

We briefly review the basic concepts in quantum computing. We define $|0\rangle := \begin{pmatrix}1 & 0\end{pmatrix}^\intercal$ and $|1\rangle := \begin{pmatrix}0 & 1\end{pmatrix}^\intercal$ as the canonical basis for $\mathbb{C}^2$. A single-qubit pure state $|\psi\rangle$ is a unit vector in $\mathbb{C}^2$, i.e. $\alpha|0\rangle + \beta|1\rangle$ for $\alpha, \beta \in \mathbb{C}$ that satisfy $|\alpha|^2 + |\beta|^2 = 1$. Multi-qubit pure states are obtained by taking tensor products of single-qubit states: an arbitrary n-qubit pure state $|\psi\rangle \in \mathbb{C}^{2n}$ is a unit vector in $\mathbb{C}^{2n}$ and can be expressed as $|\psi\rangle = \sum_{x\in\{0,1\}^n} \alpha_x |x\rangle$ where $\alpha_x \in \mathbb{C}$ and $\sum_x |\alpha_x|^2 = 1$. We denote by $\langle\psi|$ as the conjugate transpose of the quantum state $|\psi\rangle$.

In general, we may also have classical probability distributions over pure states. This scenario is captured by *mixed* states, the most general kind of states in quantum mechanics. Mixed states cannot be expressed as superpositions and are conveniently described by density matrices. Formally, a d-dimensional mixed state ρ is a $d \times d$ positive semidefinite matrix that satisfies $\mathrm{Tr}(\rho) = 1$. Equivalently, ρ is a convex combination of outer products of pure states with themselves:

$$\rho = \sum_{i=1}^{d} p_i |\psi_i\rangle \langle\psi_i|,$$

where $p_i \geq 0$ and $\sum_i p_i = 1$. In the special case where $p_i = 1$ for some i, we obtain a pure state $\rho = |\psi_i\rangle\langle\psi_i|$. A *superoperator* S maps a mixed state ρ to the mixed state $S(\rho) = \sum_{i=1}^{k} B_i \rho B_i^\dagger$, where $B_1, \ldots, B_k$ can be any matrices satisfying $\sum_{i=1}^{k} B_i^\dagger B_i = \mathbb{I}$. This is the most general (norm-preserving) mapping from mixed states to mixed states allowed by quantum mechanics.

The most general class of measurements that we can perform on mixed states are the POVMs (Positive Operator Valued Measures). Although they can be represented as superoperators, it's convenient to define them separately. In the POVM formalism, a measurement M is given by a list of $d \times d$ positive semidefinite matrices $E_1, \ldots, E_k$, which satisfy $\sum_i E_i = \mathbb{I}$. The measurement rule is:

$$\Pr[M(\rho) \text{ returns outcome } i] = \mathrm{Tr}(E_i \rho)$$

and hence $\mathbb{E}[M(\rho)] = \sum_i i\mathrm{Tr}(E_i\rho)$.

2.2 Quantum statistical queries.

We present here a definition of QSQ oracle that returns an approximation of the expectation value of any POVM. This definition generalizes the one given in [8], which is stated with respect to projective measurements.

Definition 1 (QSQ oracle)**.** *A quantum statistical query oracle $QSQ_\rho(\cdot,\cdot)$ for some d-dimensional mixed state ρ receives as inputs a tolerance parameter $\tau \geq 0$ and a POVM measurement $M = (E_1, \ldots, E_k)$. Such oracle outputs a number α satisfying*

$$|\alpha - \mathbb{E}[M(\rho)]| \leq \tau.$$

A QSQ algorithm accesses a quantum state ρ via the quantum statistical query oracle QSQ_ρ. QSQ algorithms that prepare all their queries to QSQ_ρ before receiving any answers are called nonadaptive; otherwise, they are called adaptive.

2.3 Quantum differential privacy

We recall the definition of quantum differential privacy given in [12]. For the sake of simplicity, we assume that the input state is a product state.

Definition 2 (DP measurement)**.** *Two product states $\rho = \rho_1 \otimes \ldots \otimes \rho_n$ and $\sigma = \sigma_1 \otimes \cdots \otimes \sigma_n$ are neighbors if there exists exactly one $i \in [n]$ such that $\rho_i \neq \sigma_i$. A POVM measurement M is α-differentially private on some subset S of product states if for all states $\rho, \sigma \in S$ that are neighbors and every possible outcome y of M we have that*

$$\Pr[M(\rho) = y] \leq \exp(\alpha) \Pr[M(\sigma) = y].$$

Definition 3 (Trivial measurement)**.** *A measurement M is α-trivial on some subset S if for all states $\rho, \sigma \in S$ and every possible output y of M we have that*

$$\Pr[M(\rho) = y] \leq \exp(\alpha) \Pr[M(\sigma) = y].$$

We provide here a formal definition of quantum *local* differential privacy (LDP), inspired by its classical counterpart introduced in [11].

Definition 4 (QLDP oracle)**.** *Let $\rho = \rho_1 \otimes \cdots \otimes \rho_n$ be a product state. An α-quantum local differentially private (QLDP) oracle $QL_\rho(\cdot, \cdot)$ gets an index $j \in [n]$ and an α-trivial measurement $M = (E_1, \ldots, E_k)$. Such oracle outputs $i \in [k]$ with probability $Tr(E_i \rho_j)$.*

Given a product state $\rho = \rho_1 \otimes \ldots \otimes \rho_n$, we say that an algorithm is α-QLDP if it accesses the state ρ via the oracle QL_ρ and the following restriction holds: for all $i \in [n]$, if $QL_\rho(i, M_1), \ldots, QL_\rho(i, M_k)$ are the algorithm's invocations of QL_ρ on state ρ_i, where each M_j is an α_j-local differentially private measurement, then $\sum_{j=1}^{k} \alpha_j \leq \alpha$. We stress that, in our model, the invocations above refer to the same state ρ_i, and we ignore post-measurement states. In other words, we replace each measured state with a fresh copy.

QLDP algorithms that prepare all their queries to QL_ρ before receiving any answers are called noninteractive; otherwise, they are interactive.

2.4 Quantum PAC learning

The probably approximately correct (PAC) model of learning gives a formalization of what "learning a function" means [1]. In this learning model, a concept class $\mathcal{C}$ is a collection of Boolean functions $\mathcal{C} \subseteq \{c : \{0,1\}^d \to \{0,1\}\}$. The functions inside $\mathcal{C}$ are referred to as concepts. We recall here the definition of quantum PAC learning, introduced in [2].

Definition 5 (Quantum PAC learning)**.** *A concept class $\mathcal{C} \subseteq \{c : \{0,1\}^d \to \{0,1\}\}$ is quantum PAC learnable if there exist a quantum algorithm $\mathcal{A}$ and a polynomial $poly(\cdot, \cdot, \cdot)$ such that for all concepts $c \in \mathcal{C}$, all distributions $\mathcal{X}$ on $\{0,1\}^d$, and all $\alpha, \beta \in (0,1)$, given in input α, β and n copies of $|\psi_c\rangle = \sum_{x \in \{0,1\}^d} \sqrt{\mathcal{X}(x)} \, |x, c(x)\rangle$, where $n = poly(d, 1/\alpha, \log(1/\beta))$, algorithm $\mathcal{A}$ outputs a hypothesis $h : \{0,1\}^d \to \{0,1\}$ satisfying*

$$\Pr_{x \sim \mathcal{X}}[h(x) \neq c(x)] \leq \alpha,$$

with probability at least $1 - \beta$. The probability is taken over $\mathcal{A}$. Class $\mathcal{C}$ is (inefficiently) PAC learnable if there exists a PAC learner $\mathcal{A}$ such that $\mathcal{A}$ PAC learns $\mathcal{C}$. Class $\mathcal{C}$ is efficiently PAC learnable if $\mathcal{A}$ runs in time polynomial in $d, 1/\alpha$, and $\log(1/\beta)$.

We recall as well the definition of QSQ learning, introduced in [8].

Definition 6 (QSQ learning)**.** *QSQ learnable is defined identically to PAC learnable (Definition 5), except that instead of having access to n copies of $|\psi_c\rangle = \sum_{x\in\{0,1\}^d} \sqrt{\mathcal{X}(x)}\,|x, c(x)\rangle$, a QSQ learner $\mathcal{A}$ can make $poly(d, 1/\alpha, \log(1/\beta))$ queries to oracle $QSQ_{|\psi\rangle_c}$ with tolerance $\tau \geq 1/poly(d, 1/\alpha, \log(1/\beta))$. Class $\mathcal{C}$ is efficiently QSQ learnable if both: (i) the running time of $\mathcal{A}$ and (ii) the time to evaluate each query that $\mathcal{A}$ makes are bounded by some polynomial in d, $1/\alpha$, and $\log(1/\beta)$.*

Here we give a definition of QLDP learning, which naturally extend the notion of local learning given in [11].

Definition 7 (QLDP learning)**.** *Let α, β be as in Definition 5 and $\varepsilon > 0$. Concept class $\mathcal{C} \subseteq \{c : \{0,1\}^d \to \{0,1\}\}$ is (inefficiently) QLDP learnable if there exists a quantum algorithm $\mathcal{A}$ that takes inputs $\varepsilon, \alpha, \beta$ and n copies of $|\psi_c\rangle = \sum_{x\in\{0,1\}^d} \sqrt{\mathcal{X}(x)}\,|x, c(x)\rangle$, where n, the number of labeled examples in z, is polynomial in $1/\varepsilon, d, 1/\alpha, \log(1/\beta)$, and satisfies*

1. *(Local differential privacy) For all $\varepsilon > 0$, algorithm $\mathcal{A}(\varepsilon, \cdot, \cdot, \cdot)$ is ε-QLDP.*
2. *(Utility) Algorithm $\mathcal{A}$ PAC learns $\mathcal{C}$ (Definition 5).*

$\mathcal{C}$ is efficiently privately PAC learnable if $\mathcal{A}$ runs in time polynomial in $d, 1/\varepsilon, 1/\alpha$, and $\log(1/\beta)$.

3 The equivalence

In this section, we relate the QSQ model and the QLDP model. Specifically, we show that a QSQ algorithm that queries the oracle QSQ_ρ can be simulated by a QLDP algorithm that queries the oracle $QL_{\rho^{\otimes n}}$. Moreover, the expected query complexity is preserved up to polynomial factors. For the sake of simplicity, we restrict our analysis to noninteractive QLDP algorithms and nonadaptive QSQ algorithms.

Theorem 1 (Simulation of QSQ algorithms by QLDP algorithms)**.** *Let $\mathcal{A}$ be a QSQ algorithm that makes at most t queries to a QSQ oracle QSQ_ρ, each with tolerance at least τ, with respect to some POVM with outcomes in $\{1, \ldots, k\}$. Then for any $\alpha \in (0, 1)$, there exists a $\ln\left(\frac{1+\alpha k}{1-\alpha}\right)$-QLDP algorithm that makes $n = t\frac{k^2 \ln(2t/\beta)}{2\tau^2\alpha^2}$ queries to $QL_{\rho^{\otimes n}}$ and simulates $\mathcal{A}$ correctly with probability at least $1 - \beta$.*

Proof. Let ρ a quantum state and $M = (E_1, \ldots, E_k)$ a POVM measurement. For $\tau \geq 0$, we show that any query to $QSQ_\rho(\tau, M)$ can be simulated efficiently by a QLDP algorithm or, equivalently, with nearly trivial measurements on single copies of ρ.

Define the measurement $M' = (E'_1, \ldots, E'_k)$ such that, for $i \in \{1, \ldots, k\}$,

$$E'_i = \alpha E_i + \left(\frac{1-\alpha}{k}\right)\mathbb{I}.$$

Informally, we perform M with probability α and we output an index i sampled u.a.r. with the remaining probability. Recall that, for any ρ, $\Pr[M'(\rho) = i] = \mathrm{Tr}(E'_i\rho)$. We observe that,

$$\frac{1-\alpha}{k} \leq \mathrm{Tr}(E'_i\rho) \leq \alpha + \frac{1-\alpha}{k} \leq \frac{\alpha k + 1}{k}$$

and hence the measurement M' is $\ln\left(\frac{1+\alpha k}{1-\alpha}\right)$-trivial. Apply M' on n fresh copies of ρ and let $x_1, \ldots, x_n$ be the random variables corresponding to the measurements. Define $z_i := \frac{x_i}{\alpha} + \frac{\alpha-1}{\alpha}$.

Observe that

$$\mathbb{E}[z_i] = \frac{\mathbb{E}[x_i]}{\alpha} + \frac{\alpha-1}{\alpha}$$
$$= \frac{1}{\alpha}\left(\sum_i i\mathrm{Tr}(E_i'\rho)\right) + \frac{\alpha-1}{\alpha} = \frac{1}{\alpha}\left(\sum_i i\mathrm{Tr}\left[\left(\alpha E_i + \frac{1-\alpha}{k}\mathbb{I}\right)\rho\right)\right) + \frac{\alpha-1}{\alpha}$$
$$= \sum_i i\mathrm{Tr}(E_i\rho) = \mathbb{E}[M(\rho)].$$

Observe that $x_i \in [0,k]$ and hence $z_i \in \left[\frac{\alpha-1}{\alpha}, \frac{\alpha-1}{\alpha} + \frac{k}{\alpha}\right]$. By Chernoff-Hoeffding bound,

$$\Pr\left[\left|\frac{1}{n}\sum_i z_i - \mathbb{E}[M(\rho)]\right| \geq \tau\right] \leq 2\exp\left(-\frac{2\tau^2\alpha^2 n}{k^2}\right).$$

Thus we can approximate $\mathbb{E}[M(\rho)]$ up to an additive error τ with probability $1-\delta$ by performing the measurement M' on $n = \frac{k^2\ln(2/\delta)}{2\tau^2\alpha^2}$ fresh copies of ρ.

Recall that the algorithm $\mathcal{A}$ performs t queries to the QSQ oracle. By union bound, the probability of any of the queries not being approximated within additive error τ is bounded by $\beta := t\delta$. Thus the simulation succeeds with probability at least $1-\beta$.

□

Theorem 2 (Simulation of QLDP algorithms by QSQ algorithms)**.** *Let $\mathcal{A}$ be a noninteractive ε-QLDP algorithm that makes at most t queries to a QLDP oracle $QL_{\rho^{\otimes n}}$. There exists a nonadaptive QSQ algorithm $\mathcal{B}$ that makes in expectation at most $O(te^{\varepsilon})$ queries to QSQ_ρ with tolerance $\tau := \beta/(3t)$ and the statistical difference between $\mathcal{B}$'s and $\mathcal{A}$'s output distributions is at most β.*

Proof. Since the QLDP algorithm is noninteractive, we can assume without loss of generality that it accesses each copy of ρ just once. Thus, we want to simulate a ε-trivial measurement $M = (E_1, \ldots, E_k)$ on input ρ with a QSQ oracle QSQ_ρ. In other words, we seek for an approximation of the following probability distribution $p(\cdot)$ over the set $[k]$:

$$p(i) := \mathrm{Tr}(M_i\rho).$$

To this end, we adopt the following "rejection-sampling" strategy:

1. Apply M to $|0\ldots0\rangle\langle 0\ldots0|$.

   Let w be the output and define $q(w) := \mathrm{Tr}(E_w|0\ldots0\rangle\langle 0\ldots0|)$.

2. Define $M' = (E_0', E_1')$, where $E_1' = E_w$ and $E_0' = \mathbb{I} - E_w$.

3. Set $\tau := \frac{\beta}{3t}$ and query $QSQ_\rho(M', \tau)$. Denote the output with $\widetilde{p}(w)$.

4. Accept w with probability
$$\frac{\widetilde{p}(w)}{\exp(\varepsilon)(1+\tau)q(w)}.$$
   Otherwise, reject w and repeat the procedure from step (1).

Observe that $\mathbb{E}[M'(\rho)] = \mathrm{Tr}(E'_1\rho)$ and then $\widetilde{p}(w) \in p(w) \pm \tau$.
In a given iteration, any particular element w gets output with probability

$$q(w) \times \frac{\widetilde{p}(w)}{\exp(\varepsilon)(1+\tau)q(w)} = \frac{\widetilde{p}(w)}{\exp(\varepsilon)(1+\tau)}.$$

Then the probability that a given iteration terminates (i.e., outputs some w) is

$$p_{terminate} = \sum_w \frac{\widetilde{p}(w)}{\exp(\varepsilon)(1+\tau)}.$$

Since $\widetilde{p}(w) \in p(w) \pm \tau$, we have that $\sum_w \widetilde{p}(w) \in 1 \pm \tau$. Hence, $p_{terminate} \in \frac{1}{\exp(\varepsilon)(1+\tau)}(1 \pm \tau)$. Thus, conditioned on the iteration terminating, element w is output with probability

$$\frac{1-\tau}{1+\tau}p(w) \leq \frac{\widetilde{p}(w)}{\exp(\varepsilon)(1+\tau)p_{terminate}} \leq \frac{1+\tau}{1-\tau}p(w),$$

Since $\tau \leq 1/3$,

$$\Pr[w \text{ output in a given iteration} \mid \text{iteration produces output}] \in (1 \pm 3\tau)p(w).$$

This implies that no matter which iteration produces output, the statistical difference between the distribution of w and $p(\cdot)$ will be at most $3\tau = \beta/t$. Because the algorithm $\mathcal{A}$ makes t queries, the overall statistical distance between the output distribution of $\mathcal{A}$ and the distribution resulting from the simulation is at most β, as desired.

We proved the correctness of the simulation. It remains to bound the number of queries to QSQ_ρ. Each iteration of the sampling procedure terminates with probability at least $\frac{1-\tau}{(1+\tau)\exp(\varepsilon)}$, thus the expected number of iteration is at most $\frac{\exp(\varepsilon)(1+\tau)}{1-\tau} = O(e^\varepsilon)$, and the total number of queries to the QSQ oracle is $O(te^\varepsilon)$.

□

4 Implications for quantum local learning

We defined learning in the QLDP and QSQ model in Section 2.4. An immediate but important corollary of the previous section is that QLDP learning and QSQ learning are equivalent.

Combining Theorems 1 and 2, we can state the following result.

Theorem 3. *Let $\mathcal{C} \subseteq \{c : \{0,1\}^d \to \{0,1\}\}$ be a concept class. Let $\mathcal{X}$ be a distribution over $\{0,1\}^d$. Let $|\psi_c\rangle := \sum_{x\in\{0,1\}^d} \sqrt{\mathcal{X}(x)}\,|x, c(x)\rangle$. Concept class $\mathcal{C}$ is quantum locally learnable by an adaptive QSQ learner with inputs α, β, and access to $QL_{|\psi_c\rangle^{\otimes n}}$ if and only if $\mathcal{C}$ is QSQ learnable by an adaptive QSQ learner with inputs α, β, and access to $QSQ_{|\psi_c\rangle}$.*

Furthermore, the simulations guarantee the following additional properties: (i) an efficient QSQ learner is simulatable by an efficient QLDP learner; (ii) an efficient QLDP learner is simulatable by an efficient QSQ learner; (iii) a nonadaptive QSQ (resp. noninteractive QLDP) learner is simulatable by a noninteractive QLDP (resp. nonadaptive QSQ) learner.

As shown in [8], parities, juntas and DNFs are QSQ learnable (with nonadaptive algorithms). By Theorem 3 they are QLDP learnable with noninteractive algorithms.

5 Discussion and future work

This paper provided a notion of quantum local differential privacy and established a connection with quantum statistical queries. But one could also consider a more general definition for quantum local differential privacy, allowing multiple trivial measurements on the same state. In this case, the proof of Theorem 2 would still hold for noninteractive QLDP algorithms. In the interactive case, we would need to take into account the effect of sequential adaptive measurements on the same state, possibly using the connection between quantum DP and gentleness established in [12]. Indeed, a possible separation between the interactive QLDP model and the adaptive QSQ model would shed light on the power of quantum adaptive measurements, which are widely employed in experiments [15, 16, 17].

Building upon the result of [8], we showed that parities, juntas and DNFs are efficiently learnable in the QLDP model. However, it's still unclear whether the quantum PAC model is more powerful than the QSQ and the QLDP models. So far, all the concepts learnable in the quantum PAC model seem to be learnable in the QSQ and in the QLDP models. Moreover, we can also consider the learnability of *quantum states* in the QSQ and QLDP models. Many algorithms for learning quantum states can be expressed in the QSQ model [18, 19, 20, 21], while in [12] they provide a quantum DP algorithm for shadow tomography. Thus, it would interesting to design QLDP algorithms for shadow tomography. A QLDP algorithm would measure each copy of the state separately, and hence bring shadow tomography much closer to experimental feasibility.

Acknowledgements We thank Alex B. Grilo for discussions about quantum statistical query model and Mina Doosti for discussions about differential privacy

References

[1] L. G. Valiant. A theory of the learnable. *Commun. ACM*, 27(11):1134–1142, nov 1984. ISSN 0001-0782. doi: 10.1145/1968.1972. URL https://doi.org/10.1145/1968.1972.

[2] Nader H. Bshouty and Jeffrey C. Jackson. Learning dnf over the uniform distribution using a quantum example oracle. In *Proceedings of the Eighth Annual Conference on Computational Learning Theory*, COLT '95, page 118–127, New York, NY, USA, 1995. Association for Computing Machinery. ISBN 0897917235. doi: 10.1145/225298.225312. URL https://doi.org/10.1145/225298.225312.

[3] Srinivasan Arunachalam, Sourav Chakraborty, Troy Lee, Manaswi Paraashar, and Ronald de Wolf. Two new results about quantum exact learning. *Quantum*, 5: 587, November 2021. ISSN 2521-327X. doi: 10.22331/q-2021-11-24-587. URL https://doi.org/10.22331/q-2021-11-24-587.

[4] Alex B. Grilo, Iordanis Kerenidis, and Timo Zijlstra. Learning-with-errors problem is easy with quantum samples. *Phys. Rev. A*, 99:032314, Mar 2019. doi: 10.1103/PhysRevA.99.032314. URL https://link.aps.org/doi/10.1103/PhysRevA.99.032314.

[5] Alp Atici and Rocco A. Servedio. Improved bounds on quantum learning algorithms. *Quantum Information Processing*, 4(5):355–386, Oct 2005. ISSN 1573-1332. doi: 10.1007/s11128-005-0001-2. URL http://dx.doi.org/10.1007/s11128-005-0001-2.

[6] Srinivasan Arunachalam and Ronald de Wolf. Optimal quantum sample complexity of learning algorithms. In *Proceedings of the 32nd Computational Complexity Conference*, CCC '17, Dagstuhl, DEU, 2017. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. ISBN 9783959770408.

[7] Michael Kearns. Efficient noise-tolerant learning from statistical queries. *J. ACM*, 45(6):983–1006, nov 1998. ISSN 0004-5411. doi: 10.1145/293347.293351. URL https://doi.org/10.1145/293347.293351.

[8] Srinivasan Arunachalam, Alex B. Grilo, and Henry Yuen. Quantum statistical query learning, 2020.

[9] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating noise to sensitivity in private data analysis. In *Proceedings of the Third Conference on Theory of Cryptography*, TCC'06, page 265–284, Berlin, Heidelberg, 2006. Springer-Verlag. ISBN 3540327312. doi: 10.1007/11681878_14. URL https://doi.org/10.1007/11681878_14.

[10] Cynthia Dwork and Aaron Roth. The algorithmic foundations of differential privacy. 9(3–4):211–407, August 2014. ISSN 1551-305X. doi: 10.1561/0400000042. URL https://doi.org/10.1561/0400000042.

[11] Shiva Prasad Kasiviswanathan, Homin K. Lee, Kobbi Nissim, Sofya Raskhodnikova, and Adam Smith. What can we learn privately? *SIAM J. Comput.*, 40(3):793–826, June 2011. ISSN 0097-5397. doi: 10.1137/090756090. URL https://doi.org/10.1137/090756090.

[12] Scott Aaronson and Guy N. Rothblum. Gentle measurement of quantum states and differential privacy. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2019, page 322–333, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450367059. doi: 10.1145/3313276.3316378. URL https://doi.org/10.1145/3313276.3316378.

[13] Li Zhou and Mingsheng Ying. Differential privacy in quantum computation. In *2017 IEEE 30th Computer Security Foundations Symposium (CSF)*, pages 249–262, 2017. doi: 10.1109/CSF.2017.23.

[14] Srinivasan Arunachalam, Yihui Quek, and John Smolin. Private learning implies quantum stability, 2021.

[15] H.M. Wiseman, D.W. Berry, S.D. Bartlett, B.L. Higgins, and G.J. Pryde. Adaptive measurements in the optical quantum information laboratory. *IEEE Journal of Selected Topics in Quantum Electronics*, 15(6):1661–1672, 2009. ISSN 1077-260X. doi: 10.1109/jstqe.2009.2020810. URL http://dx.doi.org/10.1109/JSTQE.2009.2020810.

[16] Guillermo García-Pérez, Matteo A.C. Rossi, Boris Sokolov, Francesco Tacchino, Panagiotis Kl. Barkoutsos, Guglielmo Mazzola, Ivano Tavernelli, and Sabrina Maniscalco. Learning to measure: Adaptive informationally complete generalized measurements for quantum algorithms. *PRX Quantum*, 2(4), Nov 2021. ISSN 2691-3399. doi: 10.1103/prxquantum.2.040342. URL http://dx.doi.org/10.1103/PRXQuantum.2.040342.

[17] Marco A. Rodríguez-García, Isaac Pérez Castillo, and P. Barberis-Blostein. Efficient qubit phase estimation using adaptive measurements. *Quantum*, 5:467, June 2021. ISSN 2521-327X. doi: 10.22331/q-2021-06-04-467. URL https://doi.org/10.22331/q-2021-06-04-467.

[18] Anurag Anshu, Srinivasan Arunachalam, Tomotaka Kuwahara, and Mehdi Soleimanifar. Sample-efficient learning of interacting quantum systems. *Nature Physics*, 17 (8):931–935, May 2021. ISSN 1745-2481. doi: 10.1038/s41567-021-01232-0. URL http://dx.doi.org/10.1038/s41567-021-01232-0.

[19] Kai-Min Chung and Han-Hsuan Lin. Sample efficient algorithms for learning quantum channels in pac model and the approximate state discrimination problem, 2021.

[20] Scott Aaronson, Xinyi Chen, Elad Hazan, Satyen Kale, and Ashwin Nayak. Online learning of quantum states. *Journal of Statistical Mechanics: Theory and Experiment*, 2019(12):124019, Dec 2019. ISSN 1742-5468. doi: 10.1088/1742-5468/ab3988. URL http://dx.doi.org/10.1088/1742-5468/ab3988.

[21] Andrea Rocchetto. Stabiliser states are efficiently pac-learnable. *Quantum Info. Comput.*, 18(7–8):541–552, jun 2018. ISSN 1533-7146.