國立臺灣大學
National Taiwan University

On learning quantum states and measurements

(*Quantum Information and Computation*, **16**(7&8): 0615–0656, 2016, arXiv: 1501.00559)

Hao-Chung Cheng (鄭皓中)

haochung@ntu.edu.tw

Department of Electrical Engineering

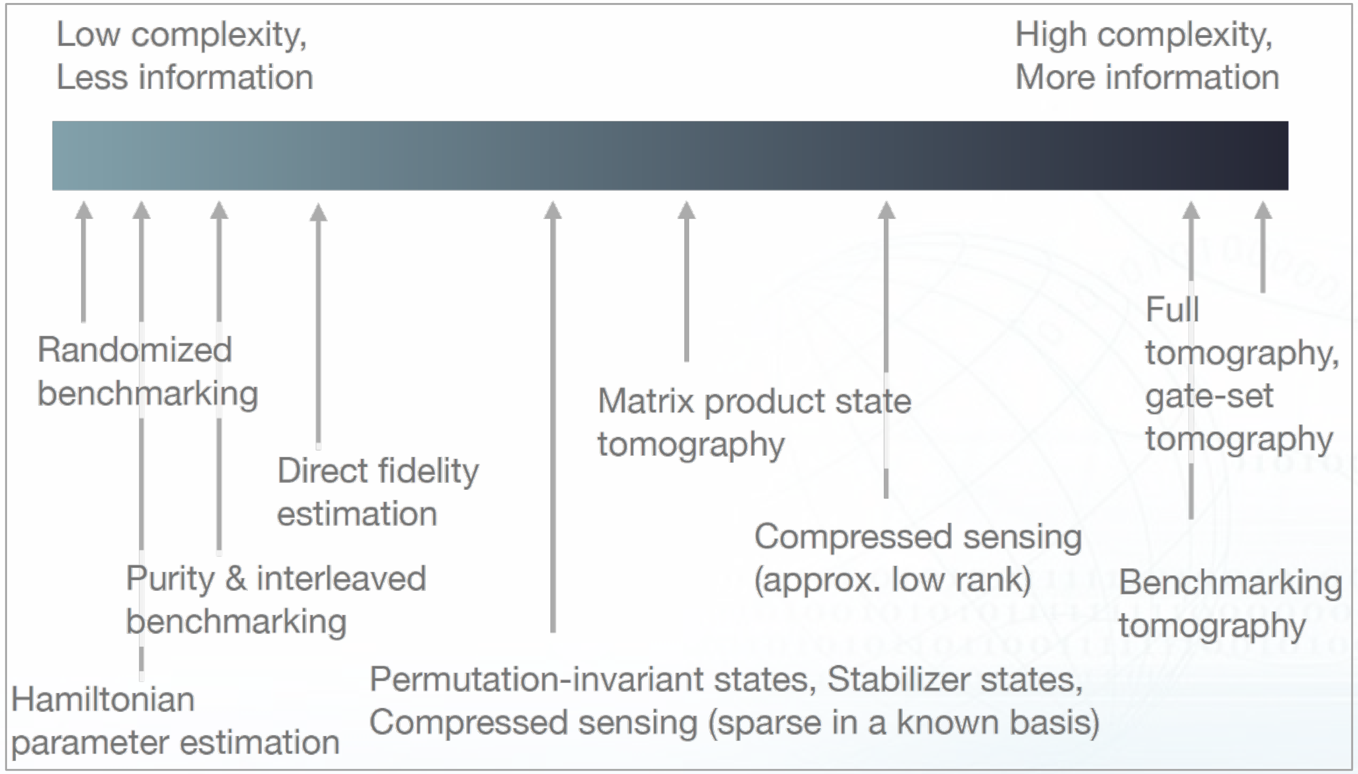
National Taiwan University

Workshop on Quantum Science and Technology (QST)

Online Meeting, August 24–27, 2021

How to learn a quantum device?

- Suppose that you have manufactured a *quantum device*, e.g. a quantum system, circuit, or measurement apparatus. How do you know that it works as expected?



[S. Flammia, QIP 2017]

Problem Formulation – Learning Quantum States

- Given: Many copies of an *unknown* quantum state $\rho \in \mathcal{S}(\mathbb{C}^d), d = 2^n$

$$\mathcal{S}(\mathbb{C}^d) := \{\rho \in \mathbb{C}^{d\times d}: \rho \geq 0, \mathrm{Tr}[\rho] = 1\}$$

Target

- Goal: To propose a *hypothesis state* $\hat{\rho} \approx \rho$.

- Question:
  - **Sample complexity**: How many copies of ρ are necessary and sufficient to produce the hypothesis state $\hat{\rho}$?
  - **Time complexity**: How long it takes to find such a state $\hat{\rho}$?

Quantum State Tomography

Exponential in the # of qubits

- A standard way – *Quantum state tomography*

  → Haah *et al.* (2017): $\tilde{O}(d^2/\epsilon^2)$ copies are necessary and sufficient such that $\|\hat{\rho} - \rho\|_1 \leq \epsilon$.

  → A general mixed d-dimensional state contains $d^2 - 1$ parameters!

- Sometimes, full tomography on an arbitrary state is overkill.
  - States with certain structures: (1) r-rank ρ: $\tilde{O}(dr/\epsilon^2)$.
    (2) stabilizer states and beyond.
  - To produce the state $\hat{\rho}$ that is ***Probably and Approximately Correct (PAC)***.

G. D'Ariano, M. Paris, M Sacchi, "Quantum Tomography," *Advances in Imaging and Electron Physics*, 128, 205-308, 2003.

Haah *et al.*, "Sample-optimal tomography of quantum states," *IEEE Transactions on Information Theory*, 63(9), 5628-5641, 2017.

The PAC Learning Model

- To learn an unknown target quantum state $\rho \in \mathcal{S}(\mathbb{C}^d)$:

$E_i \in \mathbb{C}^{d \times d}, 0 \le E_i \le I$

Randomly (i.i.d.) draw a set of two-outcome measurements $E_1, \ldots, E_m$

→ *Training set*: $\{(E_1, \mathrm{Tr}[\rho E_1]), \ldots, (E_m, \mathrm{Tr}[\rho E_m])\}$, where $E_i \sim \mu$.

Born's Rule

→ To pick a hypothesis state $\hat{\rho}$ such that

$$\Pr_{\mu}\{|\mathrm{Tr}[\hat{\rho}E] - \mathrm{Tr}[\rho E]| \le \epsilon\} \ge 1 - \delta$$

- Given $0 < \epsilon, \delta < 1$, the **sample complexity** $m_{\mathcal{S}(\mathbb{C}^d)}(\epsilon, \delta)$ is the least integer of m such that the above is satisfied.

L. G. Valiant, "A theory ofthe learnable," *Comm. ACM*, 27(11):1134-31, 1984.

Learnability of Quantum States [Aar07]

- To learn an unknown n-qubit quantum state:

$$m_{\mathcal{S}\left(\mathbb{C}^{2^n}\right)}(\epsilon, \delta) = \Theta\left(n/\epsilon^2\right)$$

Full tomography: $O\left(4^n/\epsilon^2\right)$

Empirical Risk Minimizer

- Protocol: take $m = O\left(n/\epsilon^2\right)$ many samples of training data and find $\hat{\rho}$ that has minimum training error $\frac{1}{m}\sum_{i=1}^{m}|\mathrm{Tr}[\hat{\rho}E_i] - \mathrm{Tr}[\rho E_i]|$.

- Quantum states are *PAC-learnable*, i.e. sample-efficient, but not time-efficient!

- Technique: An entropic inequality in *Quantum Random Access Codes*.

[Aar07] S. Aaronson. "The learnability of quantum states," *Proceedings of the Royal Society A*, 463 (2088), 2007

Quantum Random Access Codes

- Random access code
  - Alice encodes n bits into m (classical or quantum) bits and send them to Bob (n>m)
  - Bob restores any of the n bits with probability greater than p

- $\nexists$ classical $2 \overset{p}{\mapsto} 1$ with $p > 0.5$

  $\exists$ quantum $2 \overset{p}{\mapsto} 1$ with $p = \cos^2(\pi/8) \approx 0.85$

- **Theorem** (Ambainis *et al.*)**.** Let $\frac{1}{2} < p \leq 1$. Any quantum (and hence any classical) $m \overset{p}{\mapsto} n$ encoding satisfies $n \geq (1 - H(p))m$.

A. Ambainis, A. Nayak, A. Ta-Shma, U. Vazirani, "Dense Quantum Coding and Quantum Finite Automata," *J. ACM.* 49 (4), 2002

Why learning quantum states is *sample-efficient*?

Are quantum measurements PAC learnable?

Learning Quantum States vs. Measurements

$E_i \in \mathbb{C}^{d\times d}, 0 \leq E_i \leq I$

- To learn an unknown target quantum state $\rho \in \mathcal{S}(\mathbb{C}^d)$:

  Randomly (i.i.d.) draw a set of two-outcome measurements $E_1, \dots, E_m$

  → *Training set*: $\{(E_1, \mathrm{Tr}[\rho E_1]), \dots, (E_m, \mathrm{Tr}[\rho E_m])\}$.

  → To pick a hypothesis state $\hat{\rho}$ such that $\mathrm{Tr}[\hat{\rho} E] \approx \mathrm{Tr}[\rho E]$.

- To learn an unknown two-outcome quantum measurement $E \in \mathcal{E}(\mathbb{C}^d)$:

  Randomly (i.i.d.) draw a set of states $\rho_1, \dots, \rho_m$

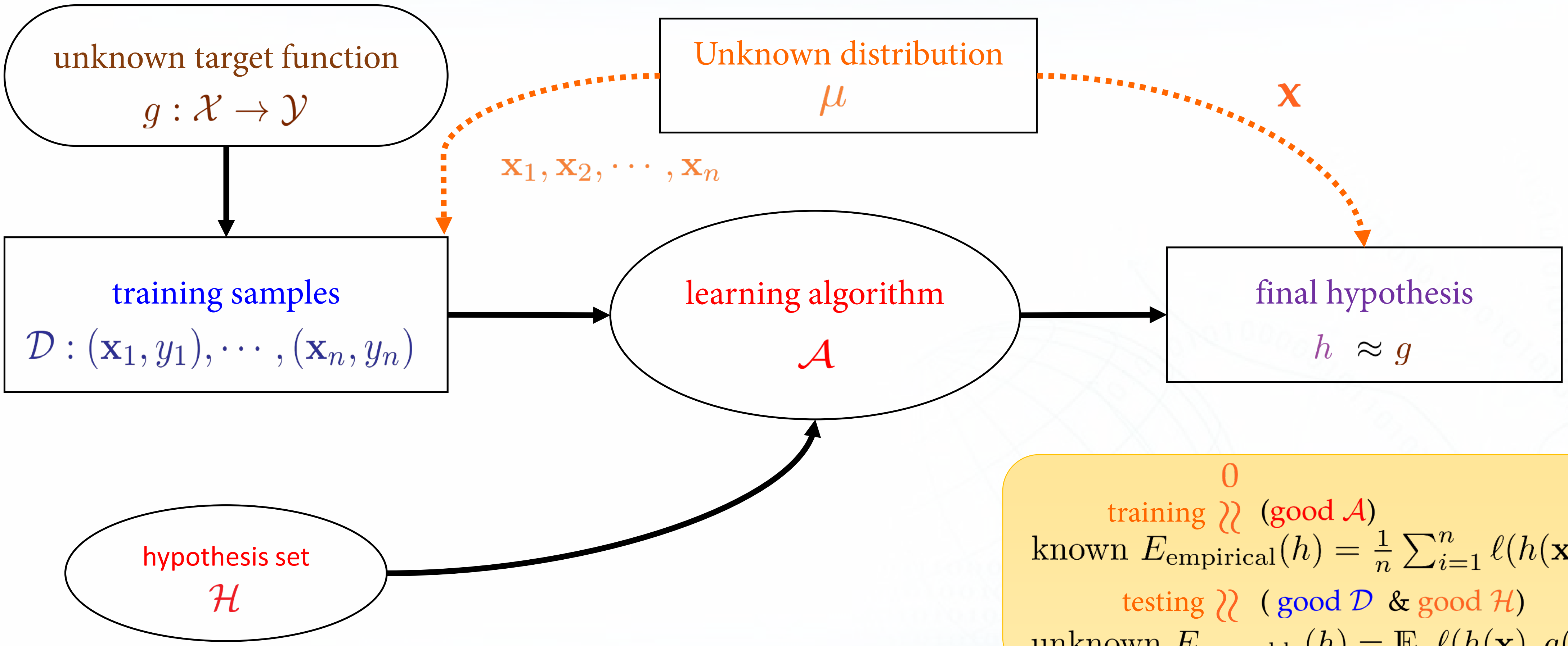
  → *Training set*: $\{(\rho_1, \mathrm{Tr}[\rho_1 E]), \dots, (\rho_m, \mathrm{Tr}[\rho_m E])\}$.

  → To pick a hypothesis operator $\hat{E}$ such that $\mathrm{Tr}[\rho \hat{E}] \approx \mathrm{Tr}[\rho E]$.

Statistical Learning Framework

Different Output Space

- binary classification: $\mathcal{Y} = \{-1, +1\}$
- multiclass classification: $\mathcal{Y} = \{1, 2, \cdots, K\}$
- regression: $\mathcal{Y} = \mathbb{R}$
- unsupervised: $\mathcal{Y} = \emptyset$



0

training $\wr$ (good $\mathcal{A}$)

known $E_{\text{empirical}}(h) = \frac{1}{n}\sum_{i=1}^{n} \ell(h(\mathbf{x}_i), y_i)$

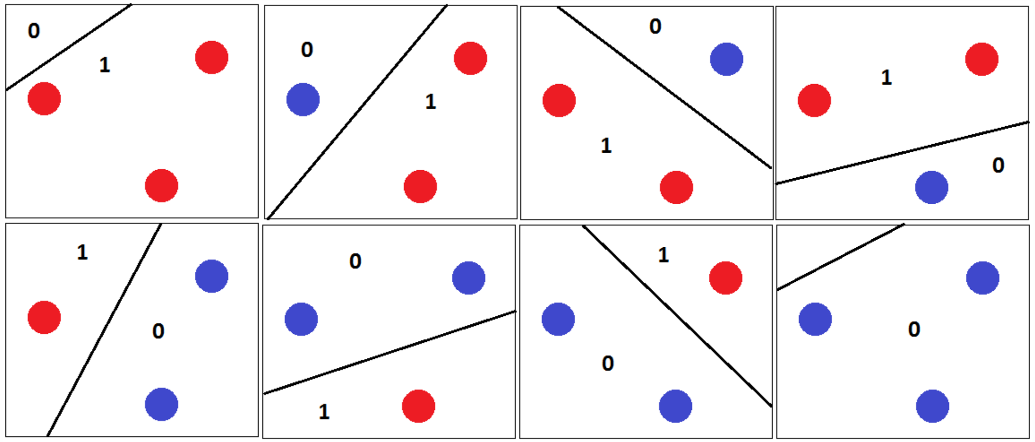
testing $\wr$ (good $\mathcal{D}$ & good $\mathcal{H}$)

unknown $E_{\text{ensemble}}(h) = \mathbb{E}_{\mu}\ell(h(\mathbf{x}), g(\mathbf{x}))$

Vapnik–Chervonenkis Dimension

Definition. Let $\mathcal{F}$ be a set of $\{0,1\}$-valued functions on a domain $\mathcal{X}$.
We say that $\mathcal{F}$ shatters a set $\{x_1,\ldots,x_n\} \subseteq \mathcal{X}$ if for every subset $B \subseteq \{1,\ldots,n\}$ there exists a function $f_B \in \mathcal{F}$ for which $f_B(x_i) = 1$ if $i \in B$, and $f_B(x_i) = 0$ if $i \notin B$. Let

$$\mathrm{VCdim}(\mathcal{F}) = \sup\{|\mathcal{S}| : \mathcal{S} \subseteq \mathcal{X},\ \mathcal{S} \text{ is shattered by } \mathcal{F}\}.$$

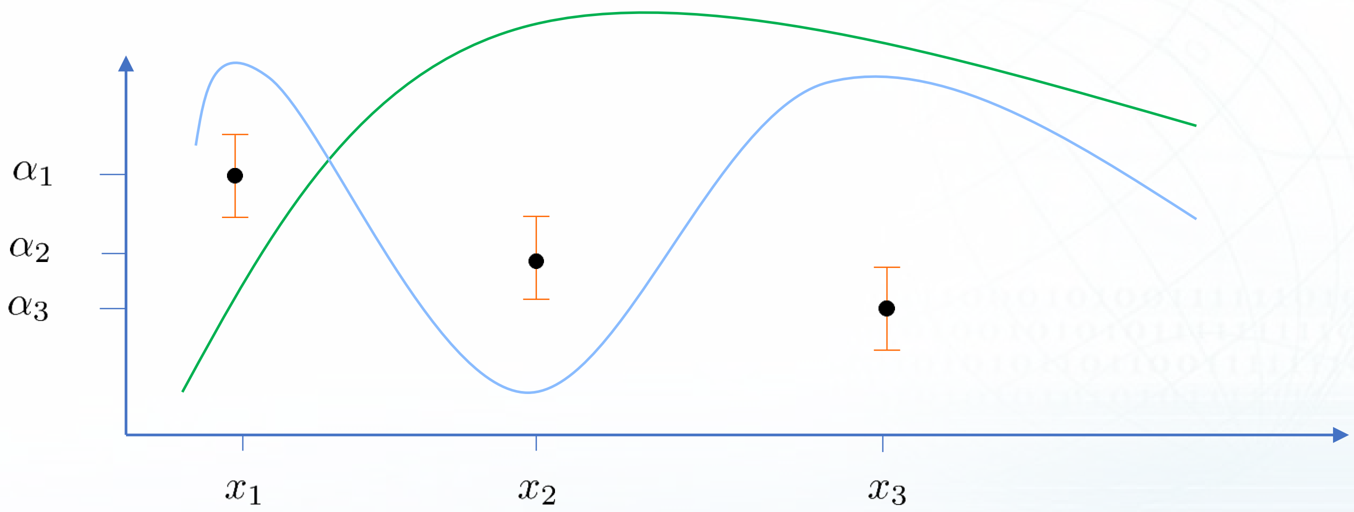


$$m_{\mathcal{F}}(\epsilon,\delta) = \Theta\left(\frac{\mathrm{VCdim}(\mathcal{F}) + \frac{1}{\delta}}{\epsilon}\right)$$

Fat-Shattering Dimension

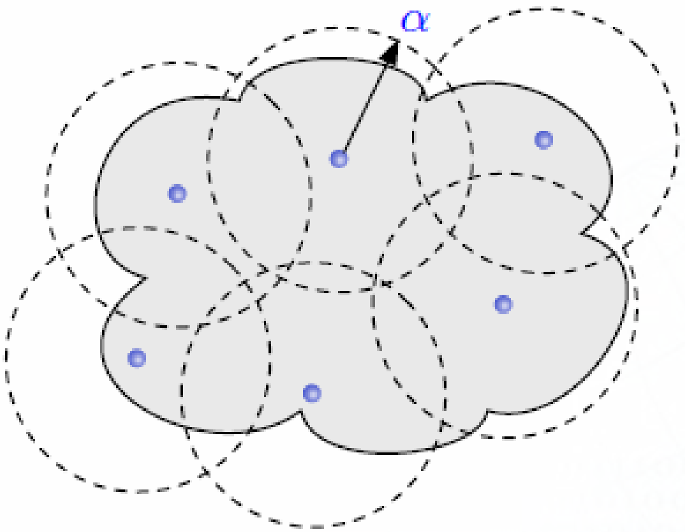
Definition. Let $\mathcal{F}$ be a set of real-valued functions on a domain $\mathcal{X}$. For every $\epsilon > 0$, a set $\mathcal{S} = \{x_1, \ldots, x_n\} \subseteq \mathcal{X}$ is said to be ϵ-shattered by the $\mathcal{F}$ if there exists a set $\{\alpha_i\}_{i=1}^n \subset \mathbb{R}$ such that for every $B \subseteq \{1, \ldots, n\}$ there is some function $f_B \in \mathcal{F}$ for which $f_B(x_i) \geq \alpha_i + \epsilon$ if $i \in B$, and $f_B(x_i) < \alpha_i - \epsilon$ if $i \notin B$. Define the *fat-shattering dimension* of $\mathcal{F}$ on the domain $\mathcal{X}$ as

$$\mathrm{fat}_{\mathcal{F}}(\epsilon, \mathcal{X}) = \sup \{|\mathcal{S}| : \mathcal{S} \subseteq \mathcal{X}, \mathcal{S} \text{ is } \epsilon\text{-shattered by } \mathcal{F}\}.$$



Covering Number

Definition. Let (Y, τ) be a metric space and let $\mathcal{F} \subset Y$. For every $\epsilon > 0$, the set $\{y_1, \ldots, y_n\}$ is called an ϵ-cover of $\mathcal{F}$ if every $f \in \mathcal{F}$ has some y_i such that $\tau(f, y_i) < \epsilon$. The covering number $\mathcal{N}(\epsilon, \mathcal{F}, \tau)$ is the minimum cardinality of a ϵ-covering set for $\mathcal{F}$ with respect to the metric τ.



Rademacher Complexity

Definition. Let μ be a probability measure on $\mathcal{X}$ and $\mathcal{F}$ be a set of uniformly bounded functions on $\mathcal{X}$. For every positive integer n, define

$$\mathcal{R}_n(\mathcal{F}) = \mathbb{E} \sup_{f \in \mathcal{F}} \frac{1}{\sqrt{n}} \left| \sum_{i=1}^{n} \gamma_i f(x_i) \right|,$$

where $\{x_i\}_{i=1}^n$ are independent random variables distributed according to μ and $\{\gamma_i\}_{i=1}^n$ independently takes values in $\{-1, +1\}$ with equal probability (which are also independent of $\{x_i\}_{i=1}^n$). The quantity $\mathcal{R}_n(\mathcal{F})$ is called the Rademacher complexity associated with the class $\mathcal{F}$.

Proof Roadmap

1. Formulate learning *operators* into learning **real-valued functions**.

2. Invoke a bound in Banach space theory that relates the number of samples to the expected norm of the input space.

3. Apply some **matrix concentration inequalities** to calculate the expected norm so as to obtain an upper bound to the sample complexity.

Take-Home Message –Real-Valued Functions

- **Definition** (Schatten p-ball). $S_p^d \triangleq \{M \in \mathbb{C}^{d\times d} : M = M^\dagger, \|M\|_p \leq 1\}$

$$\|M\|_p := \left(\mathrm{Tr}\left[|M|^p\right]\right)^{1/p}$$

- Learning quantum state ↔ learning linear functionals parameterized by S_1^d:

$$\mathcal{F} = \{X \mapsto \langle X, W\rangle_{\mathrm{HS}} = \mathrm{Tr}[XW] : \|W\|_1 \leq 1\}, \quad X : \|X\|_\infty \leq 1$$

- Learning measurements ↔ learning linear functionals parameterized by S_∞^d:

$$\mathcal{F} = \{X \mapsto \langle X, W\rangle_{\mathrm{HS}} = \mathrm{Tr}[XW] : \|W\|_\infty \leq 1\}, \quad X : \|X\|_1 \leq 1$$

Learning Quantum States and Measurements

Learning Setup	Learning Measurements	Learning States
Space $\mathcal{Z} = \mathcal{X} \times \mathcal{Y} \sim \mathcal{D}$	Input space $\mathcal{X} = \mathcal{S}(\mathbb{C}^d) \ni \rho \sim \mathbb{P}_{\mathcal{X}}$	Input space $\mathcal{X} = \mathcal{E}(\mathbb{C}^d) \ni E \sim \mathbb{P}_{\mathcal{X}}$
Target concept $t : \mathcal{X} \to \mathcal{Y}$	Unknown E $\rho \mapsto \mathrm{Tr}[\rho E]$	Unknown ρ $E \mapsto \mathrm{Tr}[\rho E]$
Hypothesis set $f \in \mathcal{G} : \mathcal{X} \to \mathbb{R}$	$\{\rho \mapsto \mathrm{Tr}(E\rho)\colon \forall E \in \mathcal{E}(\mathbb{C}^d)\}$	$\{E \mapsto \mathrm{Tr}(E\rho)\colon \forall \rho \in \mathcal{S}(\mathbb{C}^d)\}$
Loss function $\ell \in \mathcal{F} : \mathcal{G} \times \mathcal{Z} \to \mathbb{R}$	Absolute or Square Error	
Risk $R(g) = \mathbb{E}[f(g, \mathcal{Z})]$	$\mathop{\mathbb{E}}_{\mathcal{Z}\sim\mathcal{D}}\left[(g(\rho) - y)^2\right]$	
Learnability	Fat-shattering dimension, etc.	

Output space $\mathcal{Y} = [0, 1]$

Dual problem

Key Ingredient: A Banach Space Theory

- **Theorem** (Mendelson and Schechtman '04)
  The set $\mathcal{S} = \{x_1, \ldots, x_m\} \subset B_X$ is ϵ-shattered by B_{X^*} if and only if $\{x_i\}_{i=1}^m$ are linearly independent and for every $a_1, \ldots, a_m \in \mathbb{R}$,

$$\epsilon \sum_{i=1}^{m} |a_i| \leq \left\| \sum_{i=1}^{m} a_i x_i \right\|_X ,$$

  where B_X is the unit ball of some Banach space $\mathcal{X}$ and B_{X^*} is its dual unit ball.

- Denote $\{\gamma_i\}$ as the Rademacher variables (symmetric $-1, 1$-valued random variables). By selecting $a_i = \gamma_i$, we have

$$\epsilon m \leq \left\| \sum_{i=1}^{m} \gamma_i x_i \right\|_X .$$

[MS04] S. Mendelson and G. Schechtman, "The shattering dimension of sets of linear functionals," Annals of Probability, vol. 32, pp. 1746-1770, 2004

Reduction

- Learning quantum states : $\mathcal{X} = S_\infty^d, \mathcal{F} = \{X \mapsto \langle X, W \rangle_{\text{HS}} : W \in S_1^d\}$

$\Rightarrow$ To find $\mathbb{E}\left\|\sum_{i=1}^m \gamma_i x_i\right\|_\infty$

- Learning measurements: $\mathcal{X} = S_1^d, \mathcal{F} = \{X \mapsto \langle X, W \rangle_{\text{HS}} : W \in S_\infty^d\}$

$\Rightarrow$ To find $\mathbb{E}\left\|\sum_{i=1}^m \gamma_i x_i\right\|_1$

Matrix Concentration Inequality

- **Theorem** (Rademacher Series [Tro12])
  Consider a finite sequence $\{x_i\}$ of deterministic Hermitian matrices with dimension d, and let $\{\gamma_i\}$ be independent Rademacher variables. Set $\|\cdot\|_\infty$ be the operator norm. Form the matrix Rademacher series

$$Y = \sum_i \gamma_i x_i.$$

  Compute the variance parameter $\sigma^2 = \sigma^2(Y) = \|\mathbb{E}\left(Y^2\right)\|_\infty$. Then

$$\mathbb{E}\|Y\|_\infty \leq \sqrt{2\sigma^2 \log d},$$

  Furthermore, for all $t \geq 0$,

$$\Pr\{\|Y\|_\infty \geq t\} \leq d e^{-t^2/w\sigma^2}.$$

- $\epsilon m \leq \mathbb{E}\left\|\sum_{i=1}^{m} \gamma_i x_i\right\|_\infty \leq \sqrt{2m \log d}$

$$\Rightarrow m \leq \frac{2\log d}{\epsilon^2}$$

J. A. Tropp, "User-friendly tail bounds for sums of random matrices," *Found. Comput. Math.*, 12(4), pp. 389-434, 2012

Noncommutative Khintchine Inequalities

- Let x_i be deterministic $d \times d$ matrices, γ_i be independent Rademacher random variables. Then

$$\mathbb{E}\left\|\sum_{i=1}^{m} \gamma_i x_i\right\|_{S_p} \approx_p \begin{cases} \left(\|(\sum_{i=1}^m x_i x_i^*)^{1/2}\|_{S_p}^p + \|(\sum_{i=1}^m x_i^* x_i)^{1/2}\|_{S_p}^p\right)^{1/p} & \text{if } 2 \le p \le \infty \\ \inf_{x_i = a_i + b_i} \left(\|(\sum_{i=1}^m a_i a_i^*)^{1/2}\|_{S_p}^p + \|(\sum_{i=1}^m b_i^* b_i)^{1/2}\|_{S_p}^p\right)^{1/p} & \text{if } 1 \le p \le 2 \end{cases}$$

$$\Rightarrow \mathbb{E}\,\|\textstyle\sum_{i=1}^m \gamma_i x_i\|_1 \le \left\|\left(\sum_{i=1}^m x_i^2\right)^{\frac{1}{2}}\right\|_1 \text{ for } x_i \text{ being Hermitian matrices.}$$

- $\epsilon m \le \mathbb{E}\,\|\sum_{i=1}^m \gamma_i x_i\|_1 \le \sqrt{md}$

$$\Rightarrow m \le \frac{d}{\epsilon^2}$$

[LP91] F. Lust-Piquard and G. Pisier, "Noncommutative Khintchine and Paley inequalities," *Ark. Mat.* 29, 2 (1991), 241-260

Our Main Results (1/2)

- Learning quantum state $\rho \in \mathcal{S}(\mathbb{C}^d)$:

→ Sufficiency: $m_{\mathcal{S}(\mathbb{C}^d)}(\epsilon, \delta) \leq O(\log d / \epsilon^2)$.

- Learning quantum measurement $E \in \mathcal{E}(\mathbb{C}^d)$:

→ Sufficiency: $m_{\mathcal{E}(\mathbb{C}^d)}(\epsilon, \delta) \leq O(d / \epsilon^2)$.

→ Necessity: $m_{\mathcal{E}(\mathbb{C}^d)}(\epsilon, \delta) \geq d / \epsilon^2$.

(1)∃ examples of d many states that can be shattered by $\mathcal{E}(\mathbb{C}^d)$.

(2) By Kai-Min: Boolean functions are embedded in $\mathcal{E}(\mathbb{C}^d)$.

Our Main Results (2/2)

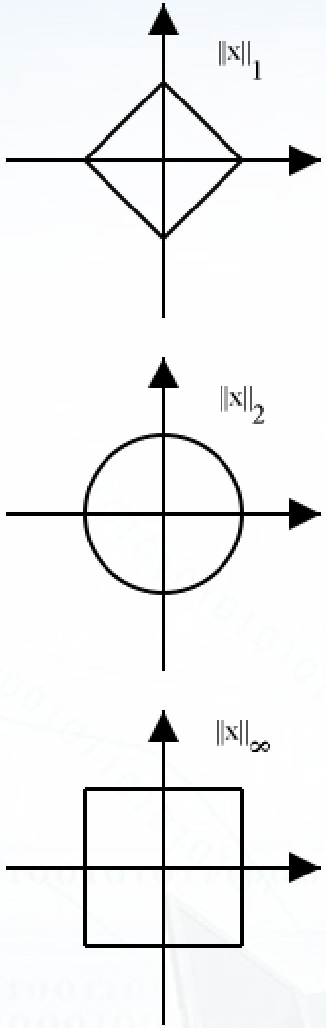
	Learning Measurements	Learning States
Pseudo-Dimension	d^2	$d^2 - 1$ ← Full tomography
Fat-Shattering Dimension	$\frac{d}{\epsilon^2}$	$\frac{\log d}{\epsilon^2}$
Uniform Entropy Number	$\frac{d}{\epsilon^2}$	$\frac{\log d}{\epsilon^2}$
Rademacher/Gaussian Complexity	$\sqrt{d}$	$\sqrt{\log d}$
Sample Complexity $m_{\mathcal{F}}(\epsilon, \delta)$	$\frac{\max\left\{d, \log\frac{1}{\delta}\right\}}{\epsilon^2}$	$\frac{\max\left\{\log d, \log\frac{1}{\delta}\right\}}{\epsilon^2}$ ← Exponential

↑ Quadratic

Intuitions

- Schatten ∞-ball is much larger than the Schatten 1-ball .

$$\frac{|\mathcal{E}(\mathbb{C}^d)|^{1/d^2}}{|\mathcal{S}(\mathbb{C}^d)|^{1/(d^2-1)}} \simeq \left(\frac{|S_\infty^d|}{|S_1^d|}\right)^{1/d^2} \simeq d$$



Related Works

Other Learning Models (1/2)

- Adversarial online learning (for states):
  - Learning happens in rounds; no more distributions on measurements.
  - In each round, adversary sends E_i; the learner replies $\mathrm{Tr}[\hat{\rho}_i E_i]$.
  - If $|\mathrm{Tr}[\hat{\rho}_i E_i] - \mathrm{Tr}[\rho E_i]| \geq \epsilon$, the adversary says 'mistake'; as fewer mistakes as possible.
  - Aaronson *et al.* (2019): $O(\log d)$ mistakes are sufficient.

- Relations between various learning models (for states):
  - Arunachalam, Quek, Smolin (2021):
    Information-theoretic implications for online learning, differential private PAC learning, etc.

S. Aaronson, X. Chen, E. Hazan, S. Kale, A. Naya, "Online learning of quantum states," J. Stat. Mech. (2019) 124019.

S. Arunachalam, Y. Quek, J. Smolin, "Private learning implies quantum stability," arXiv:2102.07171

Other Learning Models (2/2)

- PAC Learning quantum circuits:
  - Chung and Lin (2018): Sample-efficient for finite sets of quantum circuits.
  - Caro and Datta (2020): Sample-efficient (pseudo-dimension) for certain classes of circuits.

K.-M. Chung, H.-H. Lin, "Sample Efficient Algorithms for Learning Quantum Channels in PAC Model and the Approximate State Discrimination Problem," arXiv:1810.10938.

M. Caro, I. Datta, "Pseudo-dimension of quantum circuits," *Quantum Machine Intelligence*, 2(2), 2020.

Shadow Tomography

- Instead of learning quantum states on some random measurements, but on a *fixed set of measurements*.
  - Given $\{E_1, \ldots, E_k\}$, how many copies of ρ are sufficient to estimate $\mathrm{Tr}[\rho E_1], \ldots, \mathrm{Tr}[\rho E_k]$?
  - Aaronson *et al.* (2018): $\mathrm{poly}(\log k, \log d)$ copies are sufficient (exponentially better).

- Related works: classical shadows
  - Huang and Kueng (2019):
    $O(k)$ many measurement statistics are sufficient to predict k linear functions of ρ.

S. Aaronson, "Shadow tomography of quantum states," *STOC*, 2018.

H.-Y. Huang, Richard Kueng, "Predicting features of quantum systems using classical shadows", arXiv:1908.08909.

Learning States under Certain Structures (1/3)

- Hamiltonian learning:

Given copies of the Gibbs state $\rho_{\vec{\mu}} = \frac{1}{Z_\beta}\mathrm{e}^{-\beta H}$ and basis $\{E_i\}_i$, where $H = \sum_i \mu_i E_i$ is an κ-local Hamiltonian, output an appion of $\vec{\mu}$.

→ Anshu *et al.* (2020): $\tilde{\Theta}\left(\mathrm{poly}(\mathrm{e}^{\beta+\kappa}, \beta^{-1}, \epsilon^{-1}, n^3)\right)$ copies are necessary and sufficient.

→ Learning a generic Hamiltonian is NOT time-efficient.

A. Anshu, S. Arunachalam, T. Kuwahara, M. Soleimanifar, "Sample efficient learning of quantum many-body systems," FOCS 2020.

Learning States under Certain Structures (2/3)

- So far…
  - Full state tomography is not sample-efficient (even for pure states).
  - PAC learning measurements is not sample-efficient (but better than full tomography)
  - PAC learning, shadow tomography, online learning, and learning Hamiltonian are sample-efficient but time-expensive in general.

- Is it possible to time-efficiently learn certain interesting classes of states?

→ Yes!
  - Exact learning (with high probability producing the hypothesis = target)
  - PAC learning

Learning States under Certain Structures (3/3)

- Montanaro (2017): Exact learn **stabilizer states** via Bell sampling in $O(n)$.
- Low (2009): Sample-efficient for some **Clifford hierarchies**.
- Rocchetto (2018): Efficiently PAC learn **stabilizer states**.
- Lai and Cheng (2021): Exact learn the following
  - **Clifford circuits** using $O(n^2)$ queries in time $O(n^3)$.
  - Output states of an ***T*-depth one circuit** using $O(3^k n)$ queries in time $O(n^3 + 3^k n)$.

A. Montanaro, "Learning stabilizer states by Bell sampling," arXiv:1707.04012.

R. Low, "Learning and testing algorithms for the Clifford group," *Phys. Rev. A*, 80(5) 052314, 2009.

A. Rocchetto, "Stabiliser states are efficiently PAC-learnable," *Quantum Information and Computation*, 18(7&8), 2018.

Rocchetto1 *et al.*, "Experimental learning of quantum states," *Science Advances*, 5(3), 2019.

C.-Y. Lai, H.-C. Cheng, "Learning quantum circuits of some T gates," arXiv:2106.12524.

- An excellent overview talk by Srinivasan Arunachalam (TQC 2021): https://www.youtube.com/VqQTIjS8bDQ?start=32721

Discussions

Open Problems

- Learning global quantum states using only *local operations and classical communication (LOCC).*
- Learning **separable** (i.e. not entangled) measurements.
- Learning output states of the **IQP circuit**: $|\psi\rangle = \sum_x (-1)^{f(x)} |x\rangle$ where f is a degree-3 polynomials.
- Learning quantum circuits beyond T-depth one.
- Learning certain parameterized quantum circuits.
- Relation between *quantum circuit simulation* and *learnability*.
- Learning with noisy samples.

Hao-Chung Cheng (鄭皓中)
haochung@ntu.edu.tw