Sample Complexity of Separability Testing

MATH 595 Final Presentation

Jacob Beckey, Qizhao Huang, and Nishant Jain

University of Illinois, Urbana-Champaign

What is Separability Testing?

Recall that a bipartite density matrix ρ on $\mathbb{C}^d \otimes \mathbb{C}^d$ is called *separable* if it can be written as

$$\rho_{AB} = \sum_k p_k \, \omega_A^{(k)} \otimes \sigma_B^{(k)}, \tag{1}$$

for some probability distribution $\{p_k\}$.

Problem (Separability Testing)

Given unrestricted measurement access to n *identical copies of an unknown quantum state* ρ_{AB} *on* $\mathbb{C}^d \otimes \mathbb{C}^d$*, determine (with probability at least* $2/3$*) whether* ρ_{AB} *is separable or* ϵ*-far in trace distance from all separable states.*

Longstanding Open Problem

Open Problem

Determine how many samples n are necessary and sufficient to solve this task (up to constant factors) as a function of d and ϵ.

This is stated as Question 8 in [4] and was stated again in an open problem session at FOCS 2024 [1].

Outline

Our talk will be divided into the following three parts:

1. $\Omega(d^2)$ Lower Bound from Quantum Collision Bound
2. $\Omega(d^2)/\epsilon^2$ Lower Bound from Mixedness Testing
3. $O(\frac{d^4}{\epsilon^2})$ Upper Bound from Quantum State Tomography

Outline

Our talk will be divided into the following three parts:

1. $\Omega(d^2)$ Lower Bound from Quantum Collision Bound
2. $\Omega(d^2)/\epsilon^2$ Lower Bound from Mixedness Testing
3. $O(\frac{d^4}{\epsilon^2})$ Upper Bound from Quantum State Tomography

Main takeaways:

- All known lower and upper bounds on separability testing leverage representation theoretic tools
- The sample complexity of separability testing is one of the big open problems in quantum learning theory

$\Omega(d^2)$ Lower Bound from Quantum Collision Bound

Lower Bounds

A general technique for proving testing lower bounds is to construct a difficult *distinguishing* task that a tester could solve, *if one exists*.

Because the tester could solve this task, it is *at least as hard* as the distinguishing task. Thus, a sample complexity lower bound on distinguishing implies a lower bound on testing.

Lower Bounds

A general technique for proving testing lower bounds is to construct a difficult *distinguishing* task that a tester could solve, *if one exists*.

Because the tester could solve this task, it is *at least as hard* as the distinguishing task. Thus, a sample complexity lower bound on distinguishing implies a lower bound on testing.

Upper Bounds

Given a testing algorithm, one "simply" needs to use concentration inequalities to derive an upper bound on the sample complexity.

Lower Bound from Quantum Collision Problem

Theorem (Separability Testing Lower Bound, [4])

Any algorithm used to solve the separability problem must use at least $\Omega(d^2)$ *samples of* $\rho_{AB} \in \mathcal{B}(\mathbb{C}^d \otimes \mathbb{C}^d)$.

Lower Bound from Quantum Collision Problem

Theorem (Separability Testing Lower Bound, [4])

Any algorithm used to solve the separability problem must use at least $\Omega(d^2)$ *samples of* $\rho_{AB} \in \mathcal{B}(\mathbb{C}^d \otimes \mathbb{C}^d)$.

Proof Ingredients

1. Hard instance: distinguishing maximally mixed state on a random subspace from maximally mixed state on the full space (quantum collision sampling problem)
2. Most mixed states are far from separable
3. Continuity of entanglement entanglement of formation

Sample Complexity of Quantum Collision Problem

Theorem (Quantum Collision Problem, [2])

Let d and r be integers such that r strictly divides d. Any algorithm which distinguishes, with probability of success at least $2/3$, between $\rho = I/d$ or ρ is maximally mixed on a uniformly random subspace of dimension r must use $\Omega(r)$ copies of ρ.

Sample Complexity of Quantum Collision Problem

Theorem (Quantum Collision Problem, [2])

Let d and r be integers such that r strictly divides d. Any algorithm which distinguishes, with probability of success at least $2/3$, between $\rho = I/d$ or ρ is maximally mixed on a uniformly random subspace of dimension r must use $\Omega(r)$ copies of ρ.

The proof of this theorem is where the bulk of the representation theoretic tools enter. To motivate you to care about those details, let us explain how it will imply a lower bound on separability testing.

Connection to Separability Testing

Main idea: show that a maximally mixed state on a random subspace of dimension r, denoted σ_{AB}, is far from all separable states (for suitable $r < d$).

Connection to Separability Testing

Main idea: show that a maximally mixed state on a random subspace of dimension r, denoted σ_{AB}, is far from all separable states (for suitable $r < d$).

Let $E(\rho_{AB})$ denote the *entanglement of formation* (a well-defined measure of mixed state entanglement).

We need the following facts from entanglement and information theory:

1. **Fact 1:** If ρ_{AB} is separable, $E(\rho_{AB}) = 0$.
2. **Fact 2:** If $r = \lfloor cd^2 \rfloor$ for some fixed $c \in (0,1)$, there exists a universal constant $C > 0$ such that $E(\sigma_{AB}) \geq C \log d$, with high probability.
3. **Fact 3:** $E(\cdot)$ is continuous: $|E(\rho) - E(\sigma)| \lesssim \log d \cdot d_{\text{tr}}(\rho, \sigma)$.

Connection to Separability Testing

From Fact 2 on the preceding slide, we may write

$$C \log d \leq E(\sigma_{AB}) = |E(\sigma_{AB})| = |0 - E(\sigma_{AB})| = |E(\rho_{AB}) - E(\sigma_{AB})|, \qquad (2)$$

for any separable state ρ_{AB}. In particular, this holds for whatever separable state is closest to σ_{AB}, denote it ρ^*_{AB}.

Connection to Separability Testing

From Fact 2 on the preceding slide, we may write

$$C \log d \leq E(\sigma_{AB}) = |E(\sigma_{AB})| = |0 - E(\sigma_{AB})| = |E(\rho_{AB}) - E(\sigma_{AB})|, \tag{2}$$

for any separable state ρ_{AB}. In particular, this holds for whatever separable state is closest to σ_{AB}, denote it ρ^*_{AB}.

Combining this lower bound with Fact 3 (continuity), we obtain

$$C \log d \leq |E(\rho^*_{AB}) - E(\sigma_{AB})| \leq \log d \cdot d_{\mathrm{tr}}(\rho, \sigma) \implies \Omega(1) \leq d_{\mathrm{tr}}(\rho, \sigma). \tag{3}$$

On the other hand, the maximally mixed state is certainly separable. Thus **a separability tester could solve this distinguishing task**.

Quantum Collision Problem Revisited

Now that we have motivated the study of the quantum collision sampling problem, let us sketch the proof at a very high level.

Theorem (Quantum Collision Problem, [2])

Let d and r be integers such that r strictly divides d. Any algorithm which distinguishes, with probability of success at least $2/3$, between $\rho = I/d$ or ρ is maximally mixed on a uniformly random subspace of dimension r must use $\Omega(r)$ copies of ρ.

Quantum Collision Proof Sketch

The authors relate the desired problem to that of distinguishing

$$\left(\frac{\mathbb{I}}{d}\right)^{\otimes n} \quad \text{and} \quad \int_{\mathcal{U}_d} dU \left(\sum_{j=1}^{r} U \left|j\right\rangle\!\left\langle j\right| U^\dagger\right)^{\otimes n}. \tag{4}$$

Quantum Collision Proof Sketch

The authors relate the desired problem to that of distinguishing

$$\left(\frac{\mathbb{I}}{d}\right)^{\otimes n} \quad \text{and} \quad \int_{\mathcal{U}_d} dU \left(\sum_{j=1}^{r} U \left|j\right\rangle\!\left\langle j\right| U^\dagger\right)^{\otimes n}. \tag{4}$$

- Both of these states are invariant under unitary and symmetric group
- Recall from class (spectrum estimation) that weak Schur sampling with classical postprocessing is optimal for any such task
- It suffices to analyze the empirical distributions obtained from weak Schur sampling
- The authors prove an upper bound on the distance between these distributions that is close to zero unless $n = \Omega(r)$.

Final Sample Complexity Lower Bound

Recall from Fact 2 above that we had to take $r = \lfloor cd^2 \rfloor$ to ensure the maximally mixed state on the subspace was far from separable.

Thus, putting these results together, we see that distinguishing these states requires $n = \Omega(r) = \Omega(d^2)$ samples. This, in turn, implies our desired $\Omega(d^2)$ lower bound on separability testing.

$\Omega(d^2/\epsilon^2)$ Lower Bound from Quantum Mixedness Testing

Classical Distribution Testing

Problem (Distribution Property Testing)

Let $\mathcal{P}, \mathcal{Q}$ be a property (subset) of probability distributions over a finite domain $[d] = \{1, 2, \ldots, d\}$. Given sample access to an unknown distribution D over $[n]$ and a distance parameter $\varepsilon > 0$, the goal of distribution property testing is to design a randomized algorithm (a tester) that distinguishes between $D \in \mathcal{P}$ or $D \in \mathcal{Q}$ with success probability at least $2/3$.

Classical Distribution Testing

Problem (Distribution Property Testing)

Let $\mathcal{P}, \mathcal{Q}$ be a property (subset) of probability distributions over a finite domain $[d] = \{1, 2, \ldots, d\}$. Given sample access to an unknown distribution D over $[n]$ and a distance parameter $\varepsilon > 0$, the goal of distribution property testing is to design a randomized algorithm (a tester) that distinguishes between $D \in \mathcal{P}$ or $D \in \mathcal{Q}$ with success probability at least $2/3$.

The **sample complexity** of the tester is the number of independent samples from D it requires.

Uniformity testing

One interesting special case is:

$$\mathcal{P} = \{\mathrm{Unif}_d\}, \quad \mathcal{Q} = \{p : \|p - \mathrm{Unif}_d\|_1 \geq \epsilon\}$$

which is the *Uniformity Testing*.

Uniformity testing

One interesting special case is:

$$\mathcal{P} = \{\mathrm{Unif}_d\}, \quad \mathcal{Q} = \{p : \|p - \mathrm{Unif}_d\|_1 \geq \epsilon\}$$

which is the *Uniformity Testing*.

The upper bound is straightforward: Unif_d minimizes the ℓ_2-norm of the probability distribution, we distinguish two cases by using the unbiased estimator of ℓ_2-norm of the unknown probability distribution:

$$Z := \frac{1}{\binom{n}{2}} \sum_{s,t=1,s\neq t}^{n} \mathbf{1}_{\{w_s = w_t\}}$$

Uniformity testing

One interesting special case is:

$$\mathcal{P} = \{\mathrm{Unif}_d\}, \quad \mathcal{Q} = \{p : \|p - \mathrm{Unif}_d\|_1 \geq \epsilon\}$$

which is the *Uniformity Testing*.

The upper bound is straightforward: Unif_d minimizes the ℓ_2-norm of the probability distribution, we distinguish two cases by using the unbiased estimator of ℓ_2-norm of the unknown probability distribution:

$$Z := \frac{1}{\binom{n}{2}} \sum_{s,t=1, s\neq t}^{n} \mathbf{1}_{\{w_s = w_t\}}$$

The sample complexity is $O(\frac{\sqrt{d}}{\epsilon^2})$.

Toy Model on Lower Bound

Given an unknown distribution D, which would be either onw of two distributions p, q over the finite domain $[d]$, what is the optimal tester $\widetilde{D}(w)$ with only one sample $w \sim D$ to distinguish p, q under two-sided error:

$$\mathcal{E}_{p,q} := \frac{1}{2}\left(\Pr\left[\widetilde{D}(w) = 1 | D = p\right] + \Pr\left[\widetilde{D}(w) = 0 | D = q\right]\right)$$

Toy Model on Lower Bound

Given an unknown distribution D, which would be either onw of two distributions p, q over the finite domain $[d]$, what is the optimal tester $\widetilde{D}(w)$ with only one sample $w \sim D$ to distinguish p, q under two-sided error:

$$\mathcal{E}_{p,q} := \frac{1}{2}\left(\Pr\left[\widetilde{D}(w) = 1 | D = p\right] + \Pr\left[\widetilde{D}(w) = 0 | D = q\right]\right)$$

We would have:

$$\min_{\widetilde{D}} \mathcal{E}_{p,q}(\widetilde{D}) = \frac{1}{2} \min_{S \subseteq [d]} p(\bar{S}) + q(S) = \frac{1}{2}(1 - (\max_{S \subseteq [d]} p(S) - q(S)))$$

Toy Model on Lower Bound

Given an unknown distribution D, which would be either onw of two distributions p, q over the finite domain $[d]$, what is the optimal tester $\widetilde{D}(w)$ with only one sample $w \sim D$ to distinguish p, q under two-sided error:

$$\mathcal{E}_{p,q} := \frac{1}{2}\left(\Pr\left[\widetilde{D}(w) = 1 | D = p\right] + \Pr\left[\widetilde{D}(w) = 0 | D = q\right]\right)$$

We would have:

$$\min_{\widetilde{D}} \mathcal{E}_{p,q}(\widetilde{D}) = \frac{1}{2} \min_{S \subseteq [d]} p(\bar{S}) + q(S) = \frac{1}{2}(1 - (\max_{S \subseteq [d]} p(S) - q(S)))$$

We define the *total variance* as:

$$d_{TV}(p, q) := \max_{S \subseteq [d]} p(S) - q(S) = \frac{1}{2}\|p - q\|_1$$

Toy Model on Lower Bound

Given an unknown distribution D, which would be either onw of two distributions p, q over the finite domain $[d]$, what is the optimal tester $\widetilde{D}(w)$ with only one sample $w \sim D$ to distinguish p, q under two-sided error:

$$\mathcal{E}_{p,q} := \frac{1}{2}\left(\Pr\left[\widetilde{D}(w) = 1 | D = p\right] + \Pr\left[\widetilde{D}(w) = 0 | D = q\right]\right)$$

We would have:

$$\min_{\widetilde{D}} \mathcal{E}_{p,q}(\widetilde{D}) = \frac{1}{2} \min_{S \subseteq [d]} p(\bar{S}) + q(S) = \frac{1}{2}(1 - (\max_{S \subseteq [d]} p(S) - q(S)))$$

We define the *total variance* as:

$$d_{TV}(p, q) := \max_{S \subseteq [d]} p(S) - q(S) = \frac{1}{2}\|p - q\|_1$$

Hence, if a valid distribution tester exists, we need $d_{TV}(p, q) \geq \frac{1}{3}$.

More on Toy Model

Now, if $\widetilde{D}$ is a valid distribution tester between $\mathcal{P}$ and $\mathcal{Q}$, we have:

$$\mathcal{E}_{p,q}(\widetilde{D}) \leq \frac{1}{3}$$

for any $p \in \mathcal{P}$ and $q \in \mathcal{Q}$.

More on Toy Model

Now, if $\widetilde{D}$ is a valid distribution tester between $\mathcal{P}$ and $\mathcal{Q}$, we have:

$$\mathcal{E}_{p,q}(\widetilde{D}) \leq \frac{1}{3}$$

for any $p \in \mathcal{P}$ and $q \in \mathcal{Q}$.

Note $\mathcal{E}_{p,q}(\widetilde{D})$ is linear with respect to p and q. Then for any distribution Π_0 and Π_1 over $\mathcal{P}$ and $\mathcal{Q}$, we have

$$\mathbb{E}_{p\sim\Pi_0}\mathbb{E}_{q\sim\Pi_1}[\mathcal{E}_{p,q}(\widetilde{D})] = \mathcal{E}_{\mathbb{E}_{p\sim\Pi_0}[p],\mathbb{E}_{q\sim\Pi_1}[q]}(\widetilde{D}) \leq \frac{1}{3}$$

which indicates $d_{TV}(\mathbb{E}_{p\sim\Pi_0}[p], \mathbb{E}_{q\sim\Pi_1}[q]) \leq \frac{1}{3}$

Move to the n-sample tester case, the tester would get a *word* $w = (w_1, \ldots, w_n)$ with $w \sim D^{\otimes n}$. Following the same calculation, we have:

Lemma (Generalized Le Cam Lemma)

If there exists a n-sample (possibly randomized) tester $\widetilde{D}$ that is valid for distribution property testing between $\mathcal{P}$ and $\mathcal{Q}$, then for any distributions Π_0 and Π_1 over probability distributions on $\mathcal{P}$, $\mathcal{Q}$ respectively. We would have

$$d_{TV}(\mathbb{E}_{p\sim\Pi_0}[p^{\otimes n}],\ \mathbb{E}_{q\sim\Pi_1}[q^{\otimes n}]) \geq \frac{1}{3}$$

Working with tensor product

How could we bound the quantity?

$$d_{TV}(\mathbb{E}_{p\sim\Pi_0}[p^{\otimes n}],\ \mathbb{E}_{q\sim\Pi_1}[q^{\otimes n}]) \geq \frac{1}{3}$$

Working with tensor product

How could we bound the quantity?

$$d_{TV}(\mathbb{E}_{p\sim\Pi_0}[p^{\otimes n}],\ \mathbb{E}_{q\sim\Pi_1}[q^{\otimes n}]) \geq \frac{1}{3}$$

Total variation is bad: d_{TV} could only give very loose bound. The best we have is $d_{TV}(p^{\otimes n}, q^{\otimes n}) \leq n \cdot d_{TV}(p, q)$.

Working with tensor product

How could we bound the quantity?

$$d_{TV}(\mathbb{E}_{p\sim\Pi_0}[p^{\otimes n}],\ \mathbb{E}_{q\sim\Pi_1}[q^{\otimes n}]) \geq \frac{1}{3}$$

Total variation is bad: d_{TV} could only give very loose bound. The best we have is $d_{TV}(p^{\otimes n}, q^{\otimes n}) \leq n \cdot d_{TV}(p, q)$.

KL divergence is also not good: KL-divergence only has additive tensorization:

$$D(p^{\otimes n}|q^{\otimes n}) = nD(p||q)$$

Working with tensor product

How could we bound the quantity?

$$d_{TV}(\mathbb{E}_{p\sim\Pi_0}[p^{\otimes n}],\ \mathbb{E}_{q\sim\Pi_1}[q^{\otimes n}]) \geq \frac{1}{3}$$

Total variation is bad: d_{TV} could only give very loose bound. The best we have is $d_{TV}(p^{\otimes n}, q^{\otimes n}) \leq n \cdot d_{TV}(p, q)$.

KL divergence is also not good: KL-divergence only has additive tensorization:

$$D(p^{\otimes n}|q^{\otimes n}) = nD(p||q)$$

Working with tensor product

χ^2**-divergence is good:** We define the χ^2-distance as:

$$d_{\chi^2}(p, q) = \sum_{i=1}^{n} \frac{(p_i - q_i)^2}{q_i}.$$

Lemma (Ingster Method)

$$d_{\chi^2}\left(\mathbb{E}_\theta[p_\theta^{\otimes n}] \,\|\, q^{\otimes n}\right) = \mathbb{E}_{\theta,\theta'}\left[\left(1 + H(\theta, \theta')\right)^n\right] - 1,$$

where θ' is an independent copy of θ, and

$$H_j(\vartheta, \vartheta') := \mathbb{E}_{x\sim q}\left[\frac{(p_\vartheta(x) - q(x))(p_{\vartheta'}(x) - q(x))}{q(x)^2}\right]$$

is the "chi-square inner product" of $p_{j,\vartheta}$ and $p_{j,\vartheta'}$ with respect to p_j.

Uniformity Testing Lower Bound

If the sample complexity of the uniformity testing is n, we would have:

$$d_{TV}(\mathbb{E}_{p\sim\Pi}[p^{\otimes n}],\ \mathrm{Unif}_d^{\otimes n}) \geq \frac{1}{3}$$

for any distribution Π over $\mathcal{Q} = \{p : \|p - \mathrm{Unif}_d\|_1 \geq \epsilon\}$.

Uniformity Testing Lower Bound

If the sample complexity of the uniformity testing is n, we would have:

$$d_{TV}(\mathbb{E}_{p\sim\Pi}[p^{\otimes n}],\ \mathrm{Unif}_d^{\otimes n}) \geq \frac{1}{3}$$

for any distribution Π over $\mathcal{Q} = \{p : \|p - \mathrm{Unif}_d\|_1 \geq \epsilon\}$.

We need "hard"-instance to claim the testing lower bound, one natural guess would a distribution over the robust version of the uniform distribution:

- Partition the domain Ω into $d/2$ disjoint pairs:

$$\Omega = \{1, 2\} \cup \{3, 4\} \cup \cdots \cup \{d-1, d\}.$$

Uniformity Testing Lower Bound

- For each pair $(2i-1, 2i)$, independently choose a sign $s_i \in \{+1, -1\}$ uniformly at random, and define the distribution $\mathcal{D}_S$ as

$$\mathcal{D}_S(2i-1) = \frac{1}{d} + s_i \frac{\varepsilon}{d}, \qquad \mathcal{D}_S(2i) = \frac{1}{d} - s_i \frac{\varepsilon}{d}.$$

Calculation would show that:

$$d_{\chi^2}(\mathbb{E}_{p\sim\Pi}[p^{\otimes n}] || \mathrm{Unif}_d^{\otimes n}) \le \exp\big(81\epsilon^4 n^2/d\big) - 1$$

which gives the sample complexity lower bound $\Omega(\frac{\sqrt{d}}{\epsilon^2})$

Problem (Quantum Mixedness Testing)

Let ρ be an unknown quantum state on an d-dimensional Hilbert space $\mathcal{H}$, which would be either maximally mixed state, or a quantum state that is ϵ-far from maximally mixed state in trace distance. The goal of quantum mixedness testing is to distinguish between these two cases with success probability $\frac{2}{3}$.

Reduction to Classical Distribution Testing

Theorem

If we want to testing a property $\mathcal{P}$ of ρ that only depends on its spectrum, which is referred as unitary invariant property. Any algorithm for computing/testing the property with success probability at least $1-\epsilon$ in the worst-case, there exists a tester with same worst-case error probability by doing weak Schur sampling followed with classical postprocessing.

Reduction to Classical Distribution Testing

Theorem

If we want to testing a property $\mathcal{P}$ of ρ that only depends on its spectrum, which is referred as unitary invariant property. Any algorithm for computing/testing the property with success probability at least $1-\epsilon$ in the worst-case, there exists a tester with same worst-case error probability by doing weak Schur sampling followed with classical postprocessing.

Proof: Suppose M is the measurement operator corresponding to the n-sample tester output "accept". The acceptance probability is given by $\mathrm{Tr}[M\rho^{\otimes n}]$ and let:

$$1-\epsilon \leq \min_{\rho \in \mathcal{P}} \mathrm{Tr}\left[M\rho^{\otimes n}\right]$$

which is the worst-case probability of acceptance among all possible quantum states.

Reduction to Classical Distribution Testing

By permutation invariance of $\rho^{\otimes n}$, we would have:

$$\mathrm{Tr}\left[M\rho^{\otimes n}\right] = \frac{1}{n!}\sum_{\pi\in S_n}\mathrm{Tr}\left[MS_\pi\rho^{\otimes n}S_\pi^{-1}\right] = \mathrm{Tr}\left[\overline{M}\rho^{\otimes n}\right]$$

with $\overline{M} = \frac{1}{n!}\sum_{\pi\in S_n} S_\pi M S_\pi^{-1}$. Moreover, by unitary invariance assumption, we have:

$$1-\epsilon \le \min_{\rho\in\mathcal{P}}\mathrm{Tr}\left[\overline{M}\rho^{\otimes n}\right] \le \min_{\rho\in\mathcal{P}}\int \mathrm{Tr}\left[\overline{M}(U\rho U^\dagger)^{\otimes n}\right] = \min_{\rho\in\mathcal{P}}\mathrm{Tr}\left[\overline{\overline{M}}\rho^{\otimes n}\right]$$

where $\overline{\overline{M}} := \int U^{\otimes n}\overline{M}\,U^{\dagger\otimes n}$.

Reduction to Classical Distribution Testing

Note that we have:

- $[\overline{\overline{M}}, S_\pi] = 0 \quad \forall \pi \in S_n$.
- $[\overline{\overline{M}}, U^{\otimes n}] = 0 \quad \forall U \in U(d)$.

By Schur-Weyl Duality, we have $\overline{\overline{M}} = \sum_{\lambda \vdash n} \alpha_\lambda \Pi_\lambda$, where Π_λ is the projection operator of isotypical component λ. Similarly, we could have:

$$\epsilon \geq \max_{\rho \notin \mathcal{P}} \mathrm{Tr}\left[\overline{\overline{M}} \rho^{\otimes n}\right]$$

Hence, $\overline{\overline{M}}$ corresponds to a tester with same worst-case error probability. Moreover, it corresponds to the weak schur sampling with classical postprocessing.

Schur-Weyl Distribution

Recall the weak Schur Sampling process, the probability for the measurement outcome $\lambda \vdash n$ is:

$$\Pr[\lambda] = \mathrm{Tr}\left[\Pi_\lambda \rho^{\otimes n}\right]$$

Schur-Weyl Distribution

Recall the weak Schur Sampling process, the probability for the measurement outcome $\lambda \vdash n$ is:

$$\Pr[\lambda] = \mathrm{Tr}\left[\Pi_\lambda \rho^{\otimes n}\right]$$

We denote the corresponding probability distribution over all partition $\lambda \vdash n$ as the *Schur-Weyl Distribution* SW_ρ^n.

More Close look at Schur-Weyl Distribution

By Schur-Weyl Duality, we know $\rho^{\otimes n} \cong \bigoplus_{\lambda \vdash n} \mathbf{1}_\lambda \otimes \rho_\lambda$, where the Schur-Weyl Distribution is determined by the spectrum of ρ_λ.

More Close look at Schur-Weyl Distribution

By Schur-Weyl Duality, we know $\rho^{\otimes n} \cong \bigoplus_{\lambda \vdash n} \mathbf{1}_\lambda \otimes \rho_\lambda$, where the Schur-Weyl Distribution is determined by the spectrum of ρ_λ.

Given a quantum state ρ with spectrum $\text{Spec}(\rho) = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$, one way for characterising the spectrum of ρ_λ is via the SSYT. We have:

$$\text{Spec}(\rho_\lambda) = \{\alpha^T : \text{T is SSYT with shape } \lambda\}$$

where $\alpha^T := \prod_{i=1}^{d} \alpha_i^{\#\text{ of occurrences of letter i in T}}$

More Close look at Schur-Weyl Distribution

By Schur-Weyl Duality, we know $\rho^{\otimes n} \cong \bigoplus_{\lambda \vdash n} \mathbf{1}_\lambda \otimes \rho_\lambda$, where the Schur-Weyl Distribution is determined by the spectrum of ρ_λ.

Given a quantum state ρ with spectrum $\mathrm{Spec}(\rho) = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$, one way for characterising the spectrum of ρ_λ is via the SSYT. We have:

$$\mathrm{Spec}(\rho_\lambda) = \{\alpha^T : \text{T is SSYT with shape } \lambda\}$$

where $\alpha^T := \prod_{i=1}^{d} \alpha_i^{\#\text{ of occurrences of letter i in T}}$

Then $\mathrm{SW}_\rho^n(\lambda) = \dim(Sp_\lambda) \sum_{\text{SSYT T with shape } \lambda} \alpha^T =: \dim(Sp_\lambda) s_\lambda(\alpha_1, \cdots, \alpha_d)$, where s_λ denotes the schur polynomial.

Classical version of Quantum Mixedness Testing

Following the idea of classical uniformity testing, we hope to do the same thing in classical uniformity testing:

- Construct the "hard" instance, which are quantum states with spectrum:

$$P_\epsilon^d := \left(\frac{1+2\epsilon}{d}, \frac{1+2\epsilon}{d}, \ldots, \frac{1-2\epsilon}{d}, \frac{1-2\epsilon}{d}, \ldots, \frac{1-2\epsilon}{d}\right)$$

- Proving the sample lower bound by imposing the condition on:

$$d_{\chi^2}(\mathrm{SW}^n_{P_\epsilon^d}, \mathrm{SW}^n_{\mathrm{Unif}_d}) \geq O(1)$$

for the existence of a valid n-sample quantum mixedness tester.

Combinatorics is hard, but..

The last step would try the best on bounding:

$$d_{\chi^2}(\mathrm{SW}^n_{P^d_\epsilon}, \mathrm{SW}^n_{\mathrm{Unif}_d}) = \mathbb{E}_{\lambda\sim \mathrm{SW}^n_{\mathrm{Unif}_d}}\left[(\frac{\mathrm{SW}^n_{P^d_\epsilon}}{\mathrm{SW}^n_{\mathrm{Unif}_d}})^2 - 1\right]$$

$$= \mathbb{E}_{\lambda\sim \mathrm{SW}^n_{\mathrm{Unif}_d}}\left[(\frac{s_\lambda(\frac{1+2\epsilon}{d}, \frac{1+2\epsilon}{d}, \dots, \frac{1-2\epsilon}{d}, \frac{1-2\epsilon}{d}, \dots, \frac{1-2\epsilon}{d})}{s_\lambda(\frac{1}{d}, \frac{1}{d}, \cdots, \frac{1}{d})})^2 - 1\right]$$

Combinatorics is hard, but..

The last step would try the best on bounding:

$$d_{\chi^2}(\mathrm{SW}^n_{P^d_\epsilon}, \mathrm{SW}^n_{\mathrm{Unif}_d}) = \mathbb{E}_{\lambda\sim \mathrm{SW}^n_{\mathrm{Unif}_d}}\left[(\frac{\mathrm{SW}^n_{P^d_\epsilon}}{\mathrm{SW}^n_{\mathrm{Unif}_d}})^2 - 1\right]$$

$$= \mathbb{E}_{\lambda\sim \mathrm{SW}^n_{\mathrm{Unif}_d}}\left[(\frac{s_\lambda(\frac{1+2\epsilon}{d}, \frac{1+2\epsilon}{d}, \dots, \frac{1-2\epsilon}{d}, \frac{1-2\epsilon}{d}, \dots, \frac{1-2\epsilon}{d})}{s_\lambda(\frac{1}{d}, \frac{1}{d}, \cdots, \frac{1}{d})})^2 - 1\right]$$

Clever Combinatoric researcher tells us:

$$d_{\chi^2}(\mathrm{SW}^n_{P^d_\epsilon}, \mathrm{SW}^n_{\mathrm{Unif}_d}) \leq \exp\big((4n\epsilon^2/d)^2\big) - 1$$

which implies a $\Omega(d/\epsilon^2)$ lower bound on quantum mixedness testing.

Now, we could build a quantum separability tester from a quantum mixedness testing tester.

Quantum Separability Testing

Now, we could build a quantum separability tester from a quantum mixedness testing tester.

A similar argument we could show is: For an unknown quantum state ρ on $\mathbb{C}^d \otimes \mathbb{C}^d$, which is either maximally mixed state or state in the conjugacy class $\mathcal{Q} := \{UD_\epsilon U^\dagger : U \in U(d)\}$, where $D_\epsilon := \text{diag}(\frac{1+2\epsilon}{d^2}, \ldots, \frac{1+2\epsilon}{d^2}, \frac{1-2\epsilon}{d^2}, \ldots, \frac{1-2\epsilon}{d^2})$. The sample complexity is at least $\Omega(d^2/\epsilon^2)$.

Now, we could build a quantum separability tester from a quantum mixedness testing tester.

A similar argument we could show is: For an unknown quantum state ρ on $\mathbb{C}^d \otimes \mathbb{C}^d$, which is either maximally mixed state or state in the conjugacy class $\mathcal{Q} := \{UD_\epsilon U^\dagger : U \in U(d)\}$, where $D_\epsilon := \mathrm{diag}(\frac{1+2\epsilon}{d^2}, \ldots, \frac{1+2\epsilon}{d^2}, \frac{1-2\epsilon}{d^2}, \ldots, \frac{1-2\epsilon}{d^2})$. The sample complexity is at least $\Omega(d^2/\epsilon^2)$.

Useful Lemma

Lemma

There is a universal constant $C_0 > 0$ *such that for all* ϵ *with* $C_0/\sqrt{d} \leq \epsilon \leq 1/2$, *the following holds when* $\varrho = U D_\epsilon U^\dagger$, $U \sim$ *Haar*$(U(d))$ *is a uniformly random state in the ensemble* $\mathcal{Q}$:

$$\Pr\left[\forall \sigma \in \mathrm{Sep}, \|\varrho - \sigma\|_1 \geq 2\epsilon\right] \geq \frac{2}{3}.$$

Quantum Mixedness Testing via Quantum Separability Testing

Given the unknown quantum state ρ, which is either maximally mixed state or state in the conjugacy class $\mathcal{Q} := \{UD_\epsilon U^\dagger : U \in U(d)\}$. Consider the following tester:

- Sample some $U \in U(d^2)$ from Haar random unitaries.
- Apply quantum separability testing on $U\rho U^\dagger$.

Quantum Mixedness Testing via Quantum Separability Testing

Given the unknown quantum state ρ, which is either maximally mixed state or state in the conjugacy class $\mathcal{Q} := \{UD_\epsilon U^\dagger : U \in U(d)\}$. Consider the following tester:

- Sample some $U \in U(d^2)$ from Haar random unitaries.
- Apply quantum separability testing on $U\rho U^\dagger$.

If $\rho = \frac{\mathbb{I}}{d^2}$: Then $U\rho U^\dagger = \frac{\mathbb{I}}{d^2}$ and the success probability is at least $2/3$.

Quantum Mixedness Testing via Quantum Separability Testing

Given the unknown quantum state ρ, which is either maximally mixed state or state in the conjugacy class $\mathcal{Q} := \{UD_\epsilon U^\dagger : U \in U(d)\}$. Consider the following tester:

- Sample some $U \in U(d^2)$ from Haar random unitaries.
- Apply quantum separability testing on $U\rho U^\dagger$.

If $\rho = \frac{\mathbb{I}}{d^2}$: Then $U\rho U^\dagger = \frac{\mathbb{I}}{d^2}$ and the success probability is at least $2/3$.

If $\rho \in \mathcal{Q}$: According to previous lemma, $\Pr\left[\forall\sigma \in \text{Sep}, \|\varrho - \sigma\|_1 \geq 2\epsilon\right] \geq \frac{2}{3}$, then

Quantum Mixedness Testing via Quantum Separability Testing

Given the unknown quantum state ρ, which is either maximally mixed state or state in the conjugacy class $\mathcal{Q} := \{UD_\epsilon U^\dagger : U \in U(d)\}$. Consider the following tester:

- Sample some $U \in U(d^2)$ from Haar random unitaries.
- Apply quantum separability testing on $U\rho U^\dagger$.

If $\rho = \frac{\mathbb{I}}{d^2}$: Then $U\rho U^\dagger = \frac{\mathbb{I}}{d^2}$ and the success probability is at least $2/3$.

If $\rho \in \mathcal{Q}$: According to previous lemma, $\Pr\left[\forall\sigma \in \text{Sep}, \|\varrho - \sigma\|_1 \geq 2\epsilon\right] \geq \frac{2}{3}$, then

$$\begin{aligned}\Pr[\text{test success}] &= \Pr\left[\text{test success}\middle| U\rho U^\dagger \text{ is } 2\epsilon\text{-far from Sep}\right] \\ &\quad + \Pr\left[\text{test success}\middle| U\rho U^\dagger \text{ is } 2\epsilon\text{-close to Sep}\right] \\ &\geq 2/3 \cdot 2/3 + 0 = 4/9\end{aligned}$$

Hence, we could build the tester from a quantum separability tester with the same order of sample complexity, which gives the following result:

Theorem (Lower bound for separability testing)

Testing whether a state is separable or ϵ-far from separable requires at least $\Omega(d^2/\epsilon^2)$ copies.

$O(d^4/\epsilon^2)$ Upper Bound from Tomography

For this, we can just do the following:

- Quantum Tomography to recover the density matrix ρ
- Classical algorithm to solve the set membership problem for the recovered density matrix.

Textbook Tomography Algorithm

Lets consider a density matrix $\rho \in \mathbb{C}^{d \times d}$,

- Consider the *Pauli basis*: set of d^2 $d \times d$ matrices.
- Learn the coefficients of each of the matrix in this basis: $\mathcal{O}\left(\frac{d^2}{\epsilon^2}\right)$.

This leads to the overall complexity of $\mathcal{O}\left(\frac{d^4}{\epsilon^2}\right)$, however we can do better: $\mathcal{O}\left(\frac{d^2}{\epsilon^2}\right)$.

Schur Weyl Duality For Tomography

Definition

The polynomial irreducible representations of $U(d)$ are indexed by partitions λ of height at most d. We will write them as $(\pi_\lambda, V_\lambda^d)$, where V_λ^d is a vector space known as the Weyl module.

By Schur-Weyl Duality, we know $\rho^{\otimes n} \cong \bigoplus_{\lambda \vdash n} \mathbf{1}_\lambda \otimes \rho_\lambda$, where the Schur-Weyl Distribution is determined by the spectrum of ρ_λ.

Schur Weyl Duality For Tomography

Recall: For any diaginalizable matrix D with eigenvalues $\alpha_1, ..., \alpha_d$:

$$\chi_{\pi_\lambda}(D) = s_\lambda(\alpha_1, \dots, \alpha_d).$$

The Schur polynomial $s_\lambda(x_1, \dots, x_d)$ is the degree-n homogeneous polynomial defined by

$$s_\lambda(x_1, \dots, x_d) = \sum_{\text{SSYT T with shape } \lambda} x^T, \qquad x^T := \prod_{i=1}^{d} x_i^{\#\text{ of occurrences of letter } i \text{ in } T}$$

$$s_\lambda(1^d) = \dim\left(V_\lambda^d\right).$$

Also,

$$\Pr[\lambda] = \mathrm{Tr}\left(\Pi_\lambda \, \rho^{\otimes n}\right) = \mathrm{tr}(I_{\dim \lambda} \otimes \pi_\lambda(\rho)) = \dim(\lambda) \cdot s_\lambda(\alpha).$$

Pretty Good Measurement

Definition (PGM tomography algorithm [3])

Given n copies of ρ, the procedure is as follows:

1. Perform weak Schur sampling on ρ, resulting in a random partition λ. Discard the permutation irrep register. Then ρ collapses to $\frac{\pi_\lambda(\rho)}{s_\lambda(\alpha)}$.
2. Measure within the space V_λ^d using the POVM with elements

$$\frac{\dim(V_\lambda^d)}{s_\lambda(\lambda)}\,\pi_\lambda(U\,\mathrm{diag}(\lambda)\,U^\dagger)\,dU$$

   for each $U \in U(d)$.
3. Output $\mathbf{U}\,\mathrm{diag}(\lambda)\,\mathbf{U}^\dagger$.

PGM Tomography Algorithm

Theorem

Let $\rho \in \mathbb{C}^{d\times d}$ be a mixed state, and suppose we are given n copies of ρ. Let $\hat{\rho}$ denote the random output of the PGM tomography algorithm. Then

$$\mathbb{E}\left\|\hat{\rho}-\rho\right\|_F^2 = \mathop{\mathbb{E}}_{\substack{\lambda\sim \mathrm{SW}^n(\alpha)\\ U\sim \mathrm{PGM}_\lambda(\rho)}} \left\| U \operatorname{diag}(\lambda)\, U^\dagger - \rho\right\|_F^2 \leq \frac{4d-3}{n}.$$

Corollary

We have $\mathbb{E}\left\|\hat{\rho}-\rho\right\|_1 \leq \sqrt{\frac{4d-3}{n}}$ and thus, $n = \mathcal{O}\left(\frac{d^2}{\epsilon^2}\right)$ copies suffice for $\epsilon-$ accurate estimation in trace distance.

Proof Sketch

The POVM is valid due to Schur's Lemma and the weight it gives to a particular unitary $U \in U(d)$ is:

$$\frac{\dim(V_\lambda^d)}{s_\lambda(\lambda)\, s_\lambda(\alpha)} \operatorname{tr}\Big(\pi_\lambda(\rho)\, \pi_\lambda(U \operatorname{diag}(\lambda)\, U^\dagger)\Big)\, dU = \frac{\dim(V_\lambda^d)}{s_\lambda(\lambda)\, s_\lambda(\alpha)} \operatorname{tr}\Big(\pi_\lambda(\rho\, U \operatorname{diag}(\lambda)\, U^\dagger)\Big)\, dU$$

Proof Sketch

The POVM is valid due to Schur's Lemma and the weight it gives to a particular unitary $U \in U(d)$ is:

$$\frac{\dim(V_\lambda^d)}{s_\lambda(\lambda)\, s_\lambda(\alpha)} \operatorname{tr}\Big(\pi_\lambda(\rho)\, \pi_\lambda(U \operatorname{diag}(\lambda)\, U^\dagger)\Big)\; dU = \frac{\dim(V_\lambda^d)}{s_\lambda(\lambda)\, s_\lambda(\alpha)} \operatorname{tr}\Big(\pi_\lambda(\rho\, U \operatorname{diag}(\lambda)\, U^\dagger)\Big)\; dU$$

$$= \frac{\dim(V_\lambda^d)}{s_\lambda(\lambda)\, s_\lambda(\alpha)}\, s_\lambda(\rho\, U \operatorname{diag}(\lambda)\, U^\dagger)\, dU.$$

Proof Sketch

Just integrating the last result, we can derive:

$$\int_{U(d)} s_\lambda\left(AUBU^\dagger\right)\, dU = \frac{s_\lambda(A)\, s_\lambda(B)}{\dim\left(V_\lambda^d\right)}\,.$$

For Schur Polynomials:

$$\Pr\left[\lambda^{(n+1)} = \lambda + e_i \,\middle|\, \lambda^{(n)} = \lambda\right] = \frac{s_{\lambda+e_i}(\alpha)}{s_\lambda(\alpha)}.$$

Pieri's Rule:

$$(x_1 + \cdots + x_d)\, s_\lambda(x_1, \ldots, x_d) = \sum_{i=1}^{d} s_{\lambda+e_i}(x_1, \ldots, x_d).$$

Proof Sketch

We also have:

$$\left(\frac{s_{\lambda+e_1}(\alpha)}{s_\lambda(\alpha)}, \ldots, \frac{s_{\lambda+e_d}(\alpha)}{s_\lambda(\alpha)}\right) \succ (\alpha_1, \ldots, \alpha_d). \qquad \text{(majorization)}$$

and also:

$$\frac{s_\lambda(\alpha)}{s_\lambda(1,..,1)} \leq \frac{s_\mu(\alpha)}{s_\mu(1,..,1)} \iff \lambda \prec \mu \qquad \text{(Sra)}$$

and therefore $\frac{s_{\lambda+e_i}(\alpha)}{\dim\left(V^d_{\lambda+e_i}\right)} = \frac{s_{\lambda+e_i}(\alpha)}{s_{\lambda+e_i}(1,..,1)}$ forms a decreasing sequence.

Using $\lambda \sim \mathrm{SW}^n(\alpha)$ and $U \sim \mathrm{PGM}_\lambda(\rho)$:

$$n^2 \mathbb{E}_{\lambda, U} \left\| U \operatorname{diag}(\lambda)\, U^\dagger - \rho \right\|_F^2 = \mathbb{E}_{\lambda, U} \left[\sum_{i=1}^{d} (\alpha_i n)^2 + \sum_{i=1}^{d} \lambda_i^2 - 2n^2 \operatorname{Tr}\left(\rho\, U \operatorname{diag}(\lambda)\, U^\dagger \right) \right].$$

Proof Sketch

Considering the cross term, for a fixed λ we have:

$$\mathbb{E}_U \operatorname{tr}\left(\rho\, U \operatorname{diag}(\lambda)\, U^\dagger\right) = \frac{\dim(V_\lambda^d)}{s_\lambda(\lambda)\, s_\lambda(\alpha)} \int_{U(d)} \operatorname{tr}\left(\rho\, U \operatorname{diag}(\lambda)\, U^\dagger\right)\, s_\lambda\left(\rho\, U \operatorname{diag}(\lambda)\, U^\dagger\right)\, dU$$

Proof Sketch

Considering the cross term, for a fixed λ we have:

$$\mathbb{E}_U \operatorname{tr}\Big(\rho\, U \operatorname{diag}(\lambda)\, U^\dagger\Big) = \frac{\dim(V_\lambda^d)}{s_\lambda(\lambda)\, s_\lambda(\alpha)} \int_{U(d)} \operatorname{tr}\Big(\rho\, U \operatorname{diag}(\lambda)\, U^\dagger\Big)\; s_\lambda\Big(\rho\, U \operatorname{diag}(\lambda)\, U^\dagger\Big)\; dU$$

$$= \frac{\dim(V_\lambda^d)}{s_\lambda(\lambda)\, s_\lambda(\alpha)} \int_{U(d)} \sum_{i=1}^{d} s_{\lambda+e_i}\Big(\rho\, U \operatorname{diag}(\lambda)\, U^\dagger\Big)\; dU \qquad \text{(Pieri)}$$

Proof Sketch

Considering the cross term, for a fixed λ we have:

$$\mathbb{E}_U \operatorname{tr}\left(\rho\, U \operatorname{diag}(\lambda)\, U^\dagger\right) = \frac{\dim(V_\lambda^d)}{s_\lambda(\lambda)\, s_\lambda(\alpha)} \int_{U(d)} \operatorname{tr}\left(\rho\, U \operatorname{diag}(\lambda)\, U^\dagger\right)\, s_\lambda\left(\rho\, U \operatorname{diag}(\lambda)\, U^\dagger\right)\, dU$$

$$= \frac{\dim(V_\lambda^d)}{s_\lambda(\lambda)\, s_\lambda(\alpha)} \int_{U(d)} \sum_{i=1}^{d} s_{\lambda+e_i}\left(\rho\, U \operatorname{diag}(\lambda)\, U^\dagger\right)\, dU \qquad \text{(Pieri)}$$

$$= \frac{\dim(V_\lambda^d)}{s_\lambda(\lambda)\, s_\lambda(\alpha)} \sum_{i=1}^{d} \frac{s_{\lambda+e_i}(\alpha)\, s_{\lambda+e_i}(\lambda)}{\left(\dim(V_{\lambda+e_i}^d)\right)}$$

Proof Sketch

Considering the cross term, for a fixed λ we have:

$$\mathbb{E}_U \operatorname{tr}\left(\rho\, U \operatorname{diag}(\lambda)\, U^\dagger\right) = \frac{\dim(V_\lambda^d)}{s_\lambda(\lambda)\, s_\lambda(\alpha)} \int_{U(d)} \operatorname{tr}\left(\rho\, U \operatorname{diag}(\lambda)\, U^\dagger\right)\, s_\lambda\left(\rho\, U \operatorname{diag}(\lambda)\, U^\dagger\right)\, dU$$

$$= \frac{\dim(V_\lambda^d)}{s_\lambda(\lambda)\, s_\lambda(\alpha)} \int_{U(d)} \sum_{i=1}^{d} s_{\lambda+e_i}\left(\rho\, U \operatorname{diag}(\lambda)\, U^\dagger\right)\, dU \qquad \text{(Pieri)}$$

$$= \frac{\dim(V_\lambda^d)}{s_\lambda(\lambda)\, s_\lambda(\alpha)} \sum_{i=1}^{d} \frac{s_{\lambda+e_i}(\alpha)\, s_{\lambda+e_i}(\lambda)}{\left(\dim(V_{\lambda+e_i}^d)\right)}$$

$$\geq \sum_{i=1}^{d} \frac{\dim(V_\lambda^d) s_{\lambda+e_i}(\alpha)}{\left(s_{\lambda+e_i}(\alpha) \dim(V_{\lambda+e_i}^d)\right)} \cdot \frac{\lambda_i}{n} \qquad \text{(sra, majorization)}$$

Proof Sketch

This leads to:

$$
\begin{aligned}
\mathbb{E}_\lambda \sum_{i=1}^{d} \frac{s_{\lambda+e_i}(\alpha)\dim(V_\lambda^d)}{\left(s_{\lambda+e_i}(\alpha)\dim(V_{\lambda+e_i}^d)\right)} \cdot \frac{\lambda_i}{n} &\geq \mathbb{E}_\lambda \sum_{i=1}^{d} \frac{s_{\lambda+e_i}(\alpha)}{s_{\lambda+e_i}(\alpha)} \cdot \frac{\lambda_i}{n} \left(2 - \frac{\dim(V_{\lambda+e_i}^d)}{\dim(V_\lambda^d)}\right) \\
&\geq 2\sum_i \alpha_i^2 - \mathbb{E}_\lambda \sum_{i=1}^{d} \frac{s_{\lambda+e_i}(\alpha)}{s_{\lambda+e_i}(\alpha)} \cdot \frac{\lambda_i}{n} \frac{\dim(V_{\lambda+e_i}^d)}{\dim(V_\lambda^d)} \\
&\geq 2\sum_i \alpha_i^2 - \sum_i \alpha_i^2 + \frac{3d}{2n} - \frac{3}{2n}
\end{aligned}
$$

where the last step requires some algebraic manipulation.

Testing Separability of the estimated matrix

Weak Membership Problem: Given the classical description of the quantum state ρ over $\mathcal{A}_1 \otimes \mathcal{A}_2$, decide whether this state is inside or $\epsilon-$trace distance from $\mathrm{SepD}(\mathcal{A}_1 \otimes \mathcal{A}_2)$

Problem (Reframed Optimization Problem [5])

Given a Hermitian matrix Q over $A_1 \otimes A_2$ (of dimension $d \times d$), compute the optimum value, denoted by $\mathrm{OptSep}(Q)$, of the optimization problem [5]

$$\max \langle Q, X \rangle \qquad \textit{subject to } X \in \mathrm{Sep}_D(A_1 \otimes A_2).$$

Testing Separability of the estimated matrix

Theorem

Given any Hermitian matrix Q and its decomposition

$$Q = \sum_{i=1}^{M} Q_i^1 \otimes \cdots \otimes Q_i^k,$$

the quantity $\mathrm{OptSep}(Q)$ *can be approximated with additive error δ in quasi-polynomial time in d and $1/\delta$, provided that kM is bounded by poly-logarithms of d.*

Testing Separability of the estimated matrix

Consider the following two spaces:

$$\mathrm{SP}(Q) = \{\, (\langle Q_1, \rho\rangle, \langle Q_2, \rho\rangle, \ldots, \langle Q_M, \rho\rangle) : \rho \in \mathcal{D}(\mathcal{H}) \,\} \subseteq \mathbb{C}^M.$$

$$\text{Raw-}(M, w) = \big\{(q_1, q_2, \ldots, q_M) : \forall i,\ q_i \in \mathbb{C},\ \|q_i\| \le w\big\}.$$

Now, lets see whether we can calculate the distance of some point p to $\mathrm{SP}(Q)$:

$$\mathrm{dis}(\tilde{p}) = \min_{\tilde{q} \in \mathrm{SP}(Q)} \|\tilde{p} - \tilde{q}\|_1 = \min_{\rho \in \mathcal{D}(\mathcal{H})} \max_{\tilde{z} \in B(\mathbb{C}^M, \|\cdot\|_\infty)} \mathrm{Re}\langle \tilde{p} - \tilde{q}(\rho),\, \tilde{z}\rangle$$

where

$$\tilde{q}(\rho) = \big(\langle Q_1, \rho\rangle,\, \langle Q_2, \rho\rangle,\, \ldots,\, \langle Q_M, \rho\rangle\big) \in \mathbb{C}^M.$$

Testing Separability of the estimated matrix

The Min–Max theorem implies

$$\min_{\rho\in\mathcal{D}(\mathcal{H})}\ \max_{\tilde{z}\in B(\mathbb{C}^M,\|\cdot\|_\infty)} \operatorname{Re}\langle \tilde{p}-\tilde{q}(\rho),\tilde{z}\rangle = \max_{\tilde{z}\in B(\mathbb{C}^M,\|\cdot\|_\infty)}\ \min_{\rho\in\mathcal{D}(\mathcal{H})} \operatorname{Re}\langle \tilde{p}-\tilde{q}(\rho),\tilde{z}\rangle.$$

This can be solved in $poly(d, M, w, 1/\epsilon)$.

Testing Separability of the estimated matrix

Definition (ϵ-net). Let (X, d) be a metric space and let $\epsilon > 0$. A subset N_ϵ is called an ϵ-net of X if for each $x \in X$ there exists $y \in N_\epsilon$ such that $d(x, y) \leq \epsilon$.

Construction of the ϵ-net of $(\mathrm{SP}(Q), \ell_1)$ [**5**]:

1. Construct an ϵ-net of the set Raw-(M, w) with respect to the metric induced by the ℓ_1 norm. Denote such an ϵ-net by R_ϵ.

Testing Separability of the estimated matrix

Definition (ϵ-net). Let (X, d) be a metric space and let $\epsilon > 0$. A subset N_ϵ is called an ϵ-net of X if for each $x \in X$ there exists $y \in N_\epsilon$ such that $d(x, y) \leq \epsilon$.

Construction of the ϵ-net of $(\mathrm{SP}(Q), \ell_1)$ [**5**]:

1. Construct an ϵ-net of the set Raw-(M, w) with respect to the metric induced by the ℓ_1 norm. Denote such an ϵ-net by R_ϵ.
2. For each point $\tilde{p} \in R_\epsilon$, compute $\mathrm{dis}(\tilde{p})$, and include $\tilde{p}$ in N_ϵ if $\mathrm{dis}(\tilde{p}) \leq \epsilon$.

The final algorithm

1. Let $Q^t(M, w_t) = (Q_1^t, Q_2^t, \ldots, Q_M^t)$ for $t = 1, \ldots, k-1$, where $\max_i \|Q_i^t\|_{op} \leq w_t$. Let $W = \prod_{i=1}^k w_i$. Generate the ϵ_t-net of $(\mathrm{SP}(Q^t), \ell_1)$ for each $t = 1, \ldots, k-1$, where $\epsilon_t = \frac{w_t \delta}{(k-1)W}$, and denote such a set by $N_{\epsilon_t}^t$.

The final algorithm

1. Let $Q^t(M, w_t) = (Q_1^t, Q_2^t, \ldots, Q_M^t)$ for $t = 1, \ldots, k-1$, where $\max_i \|Q_i^t\|_{op} \leq w_t$. Let $W = \prod_{i=1}^k w_i$. Generate the ϵ_t-net of $(\mathrm{SP}(Q^t), \ell_1)$ for each $t = 1, \ldots, k-1$, where $\epsilon_t = \frac{w_t \delta}{(k-1)W}$, and denote such a set by $N_{\epsilon_t}^t$.
2. For each point

$$q = (q^1, q^2, \ldots, q^{k-1}) \in N_{\epsilon_1}^1 \times N_{\epsilon_2}^2 \times \cdots \times N_{\epsilon_{k-1}}^{k-1},$$

define $Q_k = \sum_{i=1}^M q_i^1 q_i^2 \cdots q_i^{k-1} Q_i^k$. Then compute $\widetilde{Q}_k = \frac{1}{2}\left(Q_k + Q_k^*\right)$ Next, compute the maximum eigenvalue of $\widetilde{Q}_k$, denoted by $\lambda_{\max}(\tilde{q})$. Update

$$\mathsf{OPT} = \max\{\mathsf{OPT}, \lambda_{\max}(\tilde{q})\}.$$

The final algorithm

1. Let $Q^t(M, w_t) = (Q_1^t, Q_2^t, \ldots, Q_M^t)$ for $t = 1, \ldots, k-1$, where $\max_i \|Q_i^t\|_{op} \leq w_t$. Let $W = \prod_{i=1}^k w_i$. Generate the ϵ_t-net of $(\mathrm{SP}(Q^t), \ell_1)$ for each $t = 1, \ldots, k-1$, where $\epsilon_t = \frac{w_t \delta}{(k-1)W}$, and denote such a set by $N_{\epsilon_t}^t$.
2. For each point

$$q = (q^1, q^2, \ldots, q^{k-1}) \in N_{\epsilon_1}^1 \times N_{\epsilon_2}^2 \times \cdots \times N_{\epsilon_{k-1}}^{k-1},$$

   define $Q_k = \sum_{i=1}^M q_i^1 \, q_i^2 \cdots q_i^{k-1} \, Q_i^k$. Then compute $\widetilde{Q}_k = \frac{1}{2}\left(Q_k + Q_k^*\right)$ Next, compute the maximum eigenvalue of $\widetilde{Q}_k$, denoted by $\lambda_{\max}(\tilde{q})$. Update

$$\mathsf{OPT} = \max\{\mathsf{OPT}, \lambda_{\max}(\tilde{q})\}.$$

3. Return OPT.

References

Sitan Chen, Jordan Cotler, Hsin-Yuan Huang, Jerry Li, and Ewin Tang.
Some Open Problems in Quantum Learning.

Andrew M. Childs, Aram W. Harrow, and Paweł Wocjan.
Weak Fourier-Schur Sampling, the Hidden Subgroup Problem, and the Quantum Collision Problem.
In Wolfgang Thomas and Pascal Weil, editors, *STACS 2007*, pages 598–609, Berlin, Heidelberg, 2007. Springer.

- Jeongwan Haah, Aram W. Harrow, Zhengfeng Ji, Xiaodi Wu, and Nengkun Yu.
  Sample-Optimal Tomography of Quantum States.
  IEEE Transactions on Information Theory, 63(9):5628–5641, September 2017.
- Ashley Montanaro and Ronald De Wolf.
  A Survey of Quantum Property Testing.
  Theory of Computing, 1(1):1–81, 2016.
- Yaoyun Shi and Xiaodi Wu.
  Epsilon-net method for optimizations over separable states, 2011.