CHAPTER 9: SPECTRUM ESTIMATION

§9.1 Problem setup

Density operators describe the state of a quantum system.

Mathematically: ρ is a quantum state $:\Leftrightarrow \rho \geq 0$ and $\operatorname{tr}\rho = 1$.

Spectral decomposition: $\rho = \sum_{i=1}^{d} \lambda_i |e_i\rangle\langle e_i|$

with: ·) Eigenvalues $(\lambda_i)_{i=1}^{d}$, $\lambda_i \geq 0$, $\sum_i \lambda_i = 1$.

·) Eigenvectors $\{|e_i\rangle\}_{i=1}^{d}$, $\langle e_i|e_j\rangle = \delta_{ij}$.

In this chapter we are interested in the task of estimating the (unknown) density operator ρ of a quantum system.

We focus on estimating the spectrum $\{\lambda_i\}_{i=1}^{d}$ of ρ.

Assumptions: 1) We have access to an experiment that prepares the system (exactly) in the state ρ.

2) We can run this experiment n times and perform joint measurements on all n copies at the same time.

$\rightarrow$ estimate spectrum of ρ by measuring $\rho^{\otimes n}$.

Goal: Devise strategy that gives exact result with probability approaching 1 as $n \rightarrow \infty$.

§ 9.2 Symmetries of spectrum estimation

The state $\rho^{\otimes n}$ is permutation invariant:

$$Q_\pi \, \rho^{\otimes n} \, Q_\pi^\dagger = \rho^{\otimes n} \text{ for all } \pi \in S_n.$$

Hence, w.l.o.g. the desired measurement also has permutation invariance, since for any $P \geq 0$ we have

$$\operatorname{tr}(P \rho^{\otimes n}) = \operatorname{tr}(P \, Q_\pi \rho^{\otimes n} Q_\pi^\dagger) = \operatorname{tr}(Q_\pi^\dagger P Q_\pi \, \rho^{\otimes n}),$$

$$\Rightarrow \quad \operatorname{tr}(P \rho^{\otimes n}) = \operatorname{tr}(\bar{P} \rho^{\otimes n}) \quad \text{with } \bar{P} = \frac{1}{n!} \sum_{\pi \in S_n} Q_\pi P Q_\pi^\dagger.$$

We also know that ρ and $U \rho U^\dagger$ have the same eigenvalues for any unitary $U \in \mathcal{U}_d$.

$\Rightarrow$ Can impose $U^{\otimes n}$ invariance on measurement operators as well!

$S_n + \mathcal{U}_d$ invariance $\rightarrow$ Schur-Weyl duality

$$(\mathbb{C}^d)^{\otimes n} = \bigoplus_{\lambda \vdash_d n} V_\lambda \otimes W_\lambda$$

S_n-irrep $\qquad$ $\mathcal{U}_d$-irrep

For $\lambda \vdash_d n$ let P_λ be the projection onto $V_\lambda \otimes W_\lambda$

$$\Rightarrow \quad P_\lambda \geq 0 \text{ and } \sum_{\lambda \vdash_d n} P_\lambda = \mathbb{1}_{\mathbb{C}^d}^{\otimes n} \qquad \text{(measurement)}$$

Furthermore: $[P_\lambda, Q_\pi] = 0 \quad \forall \pi \in S_n$

$$[P_\lambda, U^{\otimes n}] = 0 \quad \forall U \in \mathcal{U}_d$$

$\rightarrow$ good candidate for spectrum measurement!

What does outcome "$\lambda \vdash_d n$" mean?

Observation: Let $\lambda = (\lambda_1, \dots, \lambda_d) \vdash_d n$, i.e.,

$$\lambda_1 \geq \dots \geq \lambda_d \geq 0 \quad \text{and} \quad \sum^d \lambda_i = n.$$

$\Rightarrow$ $\bar{\lambda} := \lambda/n$ is a valid spectrum of a quantum state!

That is, $\bar{\lambda}_i \geq 0$ and $\sum_{i=1}^{d} \bar{\lambda}_i = 1$.

Idea of spectrum estimation:

Let ρ have spectrum $r = (r_1, \dots, r_d)$ (w.l.o.g. $r_1 \geq r_2 \geq \dots \geq r_d$)

1) Measure $\rho^{\otimes n}$ w.r.t. $\{P_\lambda\}$.

2) For outcome $\lambda \vdash_d n$, set $\hat{r} = \lambda/n$.

3) $\mathrm{Prob}(\hat{r} \neq r) \rightarrow 0$ as $n \rightarrow \infty$

The measurement in 1) is often called weak Schur sampling.

The main result of this chapter is to show 3).

§ 9.3 Weak Schur sampling

Our goal is to bound the probability of obtaining outcome "λ" (where $\lambda \vdash_d n$ is a Young diagram) in weak Schur sampling. That is, denoting by P_λ the projector onto $V_\lambda \otimes W_\lambda$ in the SW-decomposition, we want to bound

$$\mathrm{tr}\left(P_\lambda \rho^{\otimes n}\right),$$

where ρ is the unknown quantum state whose spectrum we want to estimate.

Since $Q_\pi \rho^{\otimes n} Q_\pi^\dagger = \rho^{\otimes n}$, we can write

$$\rho^{\otimes n} = \bigoplus_{\lambda \vdash_d n} \mathbb{1}_{V_\lambda} \otimes \rho_\lambda$$

for some (PSD) operators $\rho_\lambda \in \mathrm{End}(W_\lambda)$.

Recall that $W_\lambda = e_T (\mathbb{C}^d)^{\otimes n}$, where T is the standard Young tableau of shape $\lambda \vdash_d n$.

First step: characterize W_λ so that we understand the effect of P_λ on $\rho^{\otimes n}$.

Def (Majorization)

Let $x, y \in \mathbb{R}^d$, and denote by $x^\downarrow, y^\downarrow$ the vectors of components of x, y sorted in non-increasing order (e.g., $x_1^\downarrow \geq \ldots \geq x_d^\downarrow$).

Then y is said to majorize x, in symbols $x \prec y$, if

$$\cdot) \quad \sum_{i=1}^{q} x_i^\downarrow \leq \sum_{i=1}^{q} y_i^\downarrow \quad \text{for all } q = 1, \ldots, d-1,$$

$$\cdot) \quad \sum_{i=1}^{d} x_i = \sum_{i=1}^{d} y_i .$$

Now consider a spectral decomposition $\rho = \sum_{i=1}^{d} r_i |e_i\rangle\langle e_i|$,

and form the tensor product basis $B = \left\{ \bigotimes_{j=1}^{n} |e_{i_j}\rangle : i_j \in [d] \right\}$

of $(\mathbb{C}^d)^{\otimes n}$. For $|v\rangle \in B$, let $f = (f_1, \ldots, f_d)$ be the

frequency distribution of $|v\rangle$: f_i is the number of times

$|e_i\rangle$ appears in $|v\rangle$. Note that f is an (ordered) partition of n.

Lem Let $|v\rangle \in B$ with frequency distribution f, and let

T be a standard Young tableau of shape $\lambda \vdash_d n$.

Then $e_T |v\rangle = 0$ unless $f \prec \lambda$.

<u>Proof</u>: First a simple observation: if T has a column with indices j and k such that $|e_{i_j}\rangle = |e_{i_k}\rangle$ in $|v\rangle$, then

$$e_T\,|v\rangle = 0.$$

This is because $e_T \propto r_T c_T$ antisymmetrizes over columns, and $c_T = c_T\,(\mathbb{1} - (jk))$ (exercise!).

Now, w.l.o.g. assume $f_1 \geq f_2 \geq \ldots \geq f_d$.

If $e_T\,|v\rangle \neq 0$, then $f_1 \leq \lambda_1$ (length of first row of λ), because otherwise some column would have two indices j and k with $|e_{i_j}\rangle = |e_{i_k}\rangle$ in $|v\rangle$ (where $i_j = i_k$ has frequency f_1), in which case $e_T\,|v\rangle = 0$ (the basis elements $|e_{i_j}\rangle$ "spill over" into the second row).

Likewise, if $f_1 + f_2 > \lambda_1 + \lambda_2$ then the same thing happens in in row 3 or further down, hence $f_1 + f_2 \leq \lambda_1 + \lambda_2$ if $e_T\,|v\rangle \neq 0$.

Continuing in this manner, we get

$$\sum_{i=1}^{q} f_i \leq \sum_{i=1}^{q} \lambda_i \quad \text{for } 1 \leq q \leq d-1 \quad \text{and} \quad \sum_{i=1}^{d} f_i = n = \sum_{i=1}^{d} \lambda_i,$$

if $e_T\,|v\rangle \neq 0$. $\square$

$\boxed{\text{Prop}}$ Let ρ be a density operator with spectrum $r = (r_1, \dots, r_d)$ where $r_1 \geq \dots \geq r_d \geq 0$, and let $\lambda = (\lambda_1, \dots, \lambda_d) \vdash_d n$ and $\bar{\lambda} = \lambda/n$. Then,

$$\operatorname{tr}(P_\lambda \rho^{\otimes n}) \leq (n+1)^{d(d-1)/2} \exp(-n D(\bar{\lambda} \| r)),$$

with the Kullback-Leibler divergence $D(p\|q) = \sum_i p_i \log \frac{p_i}{q_i}$,

defined for probability distributions p and q with $\operatorname{supp} p := \{i : p_i \neq 0\} \subseteq \operatorname{supp} q$, and satisfying

$$D(p\|q) \geq 0 \;\; \forall p, q, \quad \text{and} \quad D(p\|q) = 0 \text{ iff } p = q.$$

<u>Proof:</u> We first recall the following fact:

For $\lambda \vdash_d n$ we denote by $SYT(\lambda)$ the set of standard Young tableau of shape λ. Then,

$$P_\lambda = \sum_{T \in SYT(\lambda)} e_T,$$

with the Young projector e_T associated to $T \in SYT(\lambda)$.

Note that $|SYT(\lambda)| = \dim V_\lambda = \frac{n!}{\prod_{(i,j)\in\lambda} h(i,j)} \leq \frac{n!}{\prod_{i=1}^d \lambda_i!}$,

where the last bound is a simple exercise.

Hence, for $\lambda \vdash_d n$ we have

$$\operatorname{tr}(P_\lambda \rho^{\otimes n}) = \sum_{T \in SYT(\lambda)} \operatorname{tr}(e_T \rho^{\otimes n}).$$

Fix some $T \in SYT(\lambda)$, and recall that $\rho^{\otimes n}$ has eigenvectors $|v\rangle \in B$ (with B the tensor product basis of eigenvectors of ρ defined before) with eigenvalues $\prod_i r_i^{f_i}$, where $f = (f_1, \dots, f_d)$ is the frequency distribution of $|v\rangle$. We can thus write

$$\rho^{\otimes n} = \sum_{|v\rangle \in B} \prod_i r_i^{f_i} |v\rangle\langle v|,$$

and using the preceding lemma,

$$\operatorname{tr}(e_T \rho^{\otimes n}) = \sum_{|v\rangle \in B} \prod_i r_i^{f_i} \operatorname{tr}(e_T |v\rangle\langle v|)$$

$$= \sum_{\substack{|v\rangle \in B: \\ f \prec \lambda}} \prod_i r_i^{f_i} \operatorname{tr}(e_T |v\rangle\langle v|)$$

To bound this expression further, we use the following simple fact from majorization theory (see exercises):

If $x \prec y$ and $u \in \mathbb{R}^d$ is arbitrary, $\langle x^\downarrow, u^\downarrow \rangle \leq \langle y^\downarrow, u^\downarrow \rangle$

Choosing $x = f$, $y = \lambda$ and $u = (\log r_1, \dots, \log r_d)$, we get

$$\langle f, u \rangle = \sum_{i=1}^{d} f_i \log r_i \le \sum_{i=1}^{d} \lambda_i \log r_i = \langle \lambda, u \rangle .$$

Exponentiating this yields $\prod_{i=1}^{d} r_i^{f_i} \le \prod_{i=1}^{d} r_i^{\lambda_i}$, so that

$$\begin{aligned}
\operatorname{tr}(e_T \rho^{\otimes n}) &= \sum_{\substack{|v\rangle \in B: \\ f \prec \lambda}} \prod_i r_i^{f_i} \operatorname{tr}(e_T |v\rangle\langle v|) \\
&\le \prod_i r_i^{\lambda_i} \operatorname{tr}\Big[e_T \overbrace{\sum_{\substack{|v\rangle \in B: \\ f \prec \lambda}} |v\rangle\langle v|}^{\le \mathbb{1}} \Big] \\
&\le \prod_i r_i^{\lambda_i} \operatorname{tr} e_T \\
&= \prod_i r_i^{\lambda_i} \dim W_\lambda \\
&\le \prod_i r_i^{\lambda_i} (n+1)^{d(d-1)/2},
\end{aligned}$$

where we used the dimension bound $\dim W_\lambda \le (n+1)^{d(d-1)/2}$ (see, e.g., Christandl's PhD thesis).

Taking everything together:

$$\operatorname{tr}(P_\lambda \rho^{\otimes n}) = \sum_{T \in SYT(\lambda)} \operatorname{tr}(e_T \rho^{\otimes n})$$

$$\leq (n+1)^{d(d-1)/2} \prod_i r_i^{\lambda_i} \sum_{T \in SYT(\lambda)} 1$$

$$= (n+1)^{d(d-1)/2} \frac{n!}{\prod_i \lambda_i!} \prod_i r_i^{\lambda_i}$$

The result now follows from a well-known bound on the multinomial coefficient $\binom{n}{\lambda} = \frac{n!}{\lambda_1! \cdots \lambda_d!}$:

$$\binom{n}{\lambda} \leq \prod_{i=1}^{d} \left(\frac{n}{\lambda_i}\right)^{\lambda_i},$$

together with the observation that

$$-n D(\bar{\lambda} \| r) = -n \sum_i \frac{\lambda_i}{n} \log \frac{\lambda_i / n}{r_i}$$

$$= \sum_i -\lambda_i \log \frac{\lambda_i}{r_i n}$$

$$= \sum_i \log \left(\frac{r_i n}{\lambda_i}\right)^{\lambda_i},$$

so that $\exp(-n D(\bar{\lambda} \| r)) = \prod_i r_i^{\lambda_i} \left(\frac{n}{\lambda_i}\right)^{\lambda_i}$. $\square$

§ 9.4 Asymptotics of spectrum estimation

We have proved that for a quantum state ρ with spectrum $r = (r_1, \dots, r_d)$, $r_i \geq r_{i+1}$, and $\lambda \vdash_d n$,

$$\operatorname{tr}(P_\lambda \rho^{\otimes n}) \leq (n+1)^{d(d-1)/2} \exp(-n D(\bar{\lambda} \| r)),$$

where $\bar{\lambda} = \lambda/n$ and $D(\cdot \| \cdot)$ is the so-called relative entropy.

We can extend this bound to a set S of possible spectra:

Set $P_S = \sum_{\substack{\lambda \vdash n: \\ \bar{\lambda} \in S}} P_\lambda$, then

$$\operatorname{tr}(P_S \rho^{\otimes n}) \leq (n+1)^{d(d+1)/2} \exp\Big(-n \min_{\substack{\lambda \vdash n: \\ \bar{\lambda} \in S}} D(\bar{\lambda} \| r)\Big),$$

which follows from picking the λ with slowest convergence (equiv., $\min D(\bar{\lambda} \| r)$), and using

$$|S| \leq |\{\lambda \vdash_d n\}| \leq (n+1)^d.$$

(which heavily overestimates the number of Young diagrams with n boxes in d rows, but is still ok.)

Finally, we consider the ε-ball

$$B_\varepsilon(r) = \left\{ r' : \sum_i |r_i - r_i'| < \varepsilon \right\}$$

around the true spectrum r. Choosing $S = \overline{B_\varepsilon(r)}$ (complement), we then obtain:

$\boxed{\text{Prop}}$ Let ρ be a quantum state with (ordered) spectrum $r = (r_1, \dots, r_d)$, and for given $\delta > 0$ let $P_\lambda = \sum_{\substack{\lambda \vdash n: \\ \bar{\lambda} \in B_\delta(r)}} P_\lambda$.

Then for any $\varepsilon > 0$ there exists n_0 s.t. for all $n \geq n_0$,

$$\operatorname{tr}\left(P_\lambda \rho^{\otimes n}\right) \geq 1 - \varepsilon.$$