PHYSICS 491: Symmetry and Quantum Information April 20, 2017

Solution of the spectrum estimation problem

Lecture 6 *Michael Walter, Stanford University*

These lecture notes are not proof-read and are offered for your convenience only. They include additional detail and references to supplementary reading material. I would be grateful if you email me about any mistakes and typos that you find.

6.1 Solution of the spectrum estimation problem

Last time we started discussing the *spectrum estimation* problem for qubits. Given $\rho^{\otimes n}$, where ρ had eigenvalues $p \geq 1-p$, we wanted to design a measurement that tells us information about $p \in [\frac{1}{2}, 1]$. For this, we considered the decomposition of $(\mathbb{C}^2)^{\otimes n}$ into irreducible representations for SU(2):

$$(\mathbb{C}^2)^{\otimes n} \cong \bigoplus_j V_j \otimes \mathbb{C}^{m(n,j)} \tag{6.1}$$

and defined P_j as the projector on the spin-j summand. We were led to these projectors because we were looking for a measurement that respected all the symmetries: the unitary invariance of the spectrum of ρ as well the permutation invariance of $\rho^{\otimes n}$. In fact, P_j is the most fine-grained measurement that commutes with $U^{\otimes n}$ and with R_π (problem 3.5). Hoping that $\{P_j\}$ might prove to be a good measurement for solving the spectrum estimation problem, we started to calculate the probability

$$\Pr(\text{outcome } j) = \operatorname{tr}\left[P_j \rho^{\otimes n}\right] = ? \tag{6.2}$$

We will now finish this calculation. Our goal will be to show that this probability is exponentially small in n, unless

$$\hat{p} := \frac{1}{2} + \frac{j}{n} \approx p.$$

Thus we will find that the measurement outcome j will lead to a good estimate $\hat{p} \approx p$ with very high probability.

The key idea to calculating (6.2) was to extend both $(\mathbb{C}^2)^{\otimes n}$ as well as the spin-j representations V_j from SU(2) to SL(2) (see eq. (5.8)). Using that $\rho/\sqrt{\det \rho}$ is an element in SL(2), we found that

$$\rho^{\otimes n} \cong (\det \rho)^{n/2} \bigoplus_j T^{(j)}_{\rho/\sqrt{\det \rho}} \otimes \mathbb{1}_{\mathbb{C}^{m(n,j)}} = \bigoplus_j \underbrace{(\det \rho)^{n/2}\, T^{(j)}_{\rho/\sqrt{\det \rho}}}_{=:T^{(n,j)}_\rho} \otimes \mathbb{1}_{\mathbb{C}^{m(n,j)}} \tag{6.3}$$

(eq. (5.9)) for arbitrary density operators ρ. It followed that:

$$\operatorname{tr}\left[P_j \rho^{\otimes n}\right] = m(n,j) \operatorname{tr}\left[T^{(n,j)}_\rho\right]. \tag{6.4}$$

Last time, we calculated the right-hand side trace but not the multiplicities $m(n,j)$. For this, we will recall one last fact that you learned in your quantum mechanics class when studying the total

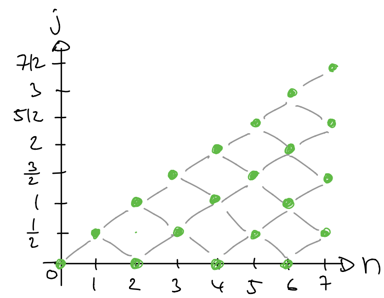

Figure 9: By iterating the Clebsch-Gordan decomposition for $V_{1/2} \otimes V_{1/2} \otimes \ldots$, we obtain a decomposition of $(\mathbb{C}^2)^{\otimes n}$ into irreducible representations of SU(2).

angular momentum. Given two irreducible representations V_{j_1} and V_{j_2}, we can consider their tensor product $V_{j_1} \otimes V_{j_2}$. This is a representation of SU(2), with U acting by $T_U^{(j_1)} \otimes T_U^{(j_2)}$. In general this representation is not irreducible and so it can be decomposed it into irreducibles. The famous *Clebsch-Gordan rule* tells us what that this decomposition look as follows:

$$V_{j_1} \otimes V_{j_2} \cong V_{j_1+j_2} \oplus V_{j_1+j_2-1} \oplus \ldots \oplus V_{|j_1-j_2|}$$

In particular, for $j_2 = \frac{1}{2}$, we have

$$V_j \otimes V_{1/2} = \begin{cases} V_{j+1/2} \oplus V_{j-1/2} & \text{if } j > 0 \\ V_{1/2} & \text{if } j = 0 \end{cases}. \tag{6.5}$$

Since a single qubit is nothing but a spin-1/2 representation, this allows us to decompose $(\mathbb{C}^2)^{\otimes n}$ by successively applying the Clebsch-Gordan rule (6.5):

$$\begin{aligned}
(\mathbb{C}^2)^{\otimes 1} &\cong V_{1/2} \\
(\mathbb{C}^2)^{\otimes 2} &\cong V_{1/2} \otimes V_{1/2} = V_1 \oplus V_0 \\
(\mathbb{C}^2)^{\otimes 3} &\cong (V_1 \oplus V_0) \otimes V_{1/2} = V_{3/2} \oplus \big(V_{1/2} \oplus V_{1/2}\big) \oplus V_0 \\
&\vdots
\end{aligned}$$

This process is visualized in fig. 9 and the general result is as follows: The multiplicity $m(n,j)$ of V_j in $(\mathbb{C}^2)^{\otimes n}$ is precisely equal to the number of paths from $(0,0)$ to (n,j) in fig. 9.

How can we estimate the number of paths? Any path can be specified by a number of n "ups" and "downs". The number of "ups" u must satisfy $(u-(n-u))/2 = u - n/2 = j$ in order to end up at (n,j). Thus there are at most $\binom{n}{\frac{n}{2}+j}$ such paths. (This is only an upper bound because paths that go below zero are invalid.) As a consequence of eq. (5.1), this means that

$$m(n,j) \leq \binom{n}{\frac{n}{2}+j} \leq 2^{nh(\hat{p})}, \tag{6.6}$$

where we recall the binary Shannon entropy

$$h(\hat{p}) = -\hat{p}\log\hat{p} - (1-\hat{p})\log(1-\hat{p})$$

from the compression of coin flips in section 5.1. Thus the multiplicites $m(n,j)$ grow at most exponentially, with exponent is given by precisely by the binary entropy. Note that, as a consequence

$$\operatorname{rk} P_j = (\dim V_j)\, m(n,j) \leq (2j+1) 2^{n h(\hat{p})} \leq (n+1) 2^{n h(\hat{p})}. \tag{6.7}$$

This fact will prove important later for information theoretic applications.

Remark. *More generally, given two representations $\mathcal{H}$ and $\mathcal{H}'$ of some group G, we can always consider their tensor product $\mathcal{H} \otimes \mathcal{H}'$ as a representation of the group G, with representation operators $T_g \otimes T_{g'}$. Note that this is precisely the same notation as used in eq. (6.1) if we think of $\mathbb{C}^{m(n,j)}$ as an $m(n,j)$-dimensional trivial representation of* $\mathrm{SU}(2)$.

The other ingredient in eq. (6.4) is the trace of the operator $T_\rho^{(n,j)}$. Last time, we computed the following upper bound (eq. (5.10)):

$$\operatorname{tr}\left[T_\rho^{(n,j)}\right] \leq (2j+1) p^{\frac{n}{2}+j} (1-p)^{\frac{n}{2}-j}$$

We can rewrite this as follows,

$$\begin{aligned}
\operatorname{tr}\left[T_\rho^{(n,j)}\right] &\leq (2j+1) 2^{n\left[\hat{p}\log p + (1-\hat{p})\log(1-p)\right]} \leq (2j+1) 2^{-n\left[\hat{p}\log\frac{1}{p} + (1-\hat{p})\log\frac{1}{1-p}\right]} \\
&= (2j+1) 2^{-n\left[-\hat{p}\log\hat{p} - (1-\hat{p})\log(1-\hat{p}) + \hat{p}\log\frac{\hat{p}}{p} + (1-\hat{p})\log\frac{1-\hat{p}}{1-p}\right]} \\
&\leq (2j+1) 2^{-n\left[h(\hat{p}) + \delta(\hat{p}\|p)\right]},
\end{aligned} \tag{6.8}$$

where we have introduced the *binary relative entropy*

$$\delta(\hat{p}\|p) = \hat{p}\log\frac{\hat{p}}{p} + (1-\hat{p})\log\frac{1-\hat{p}}{1-p}.$$

Remark. *The relative entropy is an important quantity in information theory and statistics. Note that it is* not *symmetric under exchanging $p \leftrightarrow \hat{p}$.*

What is the purpose of this rewriting? If we plug eqs. (6.6) and (6.8) into eq. (6.4) we obtain the following result:

$$\Pr(\text{outcome } j) = \operatorname{tr}\left[P_j \rho^{\otimes n}\right] \leq (2j+1) 2^{-n\delta(\hat{p}\|p)} \tag{6.9}$$

The point now is that the relative entropy is a distance measure between probability distributions: It is nonnegative and $\delta(\hat{p}\|p) = 0$ if and only if $p = \hat{p}$. More quantitatively, it satisfies the following inequality, a special case of the so-called *Pinsker's inequality* (problem 3.4):

$$\delta(\hat{p}\|p) \geq \frac{2}{\ln 2}(\hat{p}-p)^2 \tag{6.10}$$

Thus, unless $\hat{p} \approx$, the probability in eq. (6.9) is exponentially small. This at last allows us to solve the spectrum estimation problem for qubits:

Given $\rho^{\otimes n}$, perform a total spin measurement in the state $\rho^{\otimes n}$ using the projective measurement $\{P_j\}$. Upon outcome j, estimate that the maximal eigenvalue of the state ρ is $\hat{p} = \frac{1}{2} + \frac{j}{n}$. Then,

$$\Pr(|\hat{p}-p| > \varepsilon) = \sum_{j \text{ s.th. } |\hat{p}-p|>\varepsilon} \operatorname{tr}\left[P_j \rho^{\otimes n}\right] \leq (n+1)^2 2^{-\frac{2}{\ln 2} n\varepsilon^2}, \tag{6.11}$$

where we have used that eqs. (6.9) and (6.10), that $2j+1 \leq n+1$, and that the sum runs certainly over no more than $n+1$ values of j. This means that $\hat{p} \approx p$ with very high probability.

In lecture 10, we will discuss how to implement the spectrum estimation measurement concretely by a quantum circuit. Spectrum estimation has been realized experimentally in Beverland et al. (2016).

6.2 Towards quantum data compression

There is another interpretation of what we have just achieved. For fixed $\varepsilon > 0$, consider the projection operator

$$\tilde{P}_n = \sum_{j \text{ s.th. } |\hat{p}-p|<\varepsilon} P_j$$

on all sectors j in eq. (6.1) for which $|\hat{p} - p| < \varepsilon$. Equation (6.11) asserts precisely that

$$\operatorname{tr}\left[\tilde{P}_n \rho^{\otimes n}\right] \to 1 \tag{6.12}$$

as $n \to \infty$, which in turn implies that

$$\tilde{P}_n \rho^{\otimes n} \tilde{P}_n \approx \rho^{\otimes n}.$$

This means that if we perform a measurement $\{\tilde{P}_n, \mathbb{1} - \tilde{P}_n\}$ on $\rho^{\otimes n}$ then, for large n, this measurement will proceed with very high probability and leave the state $\rho^{\otimes n}$ almost unchanged. We will call the subspace $\tilde{\mathcal{H}}_n$ that $\tilde{P}_n$ projects on a *typical subspace* for $\rho^{\otimes n}$ (although we caution that the traditional definition is somewhat different).

Since the binary entropy is continuous,

$$|\hat{p} - p| < \varepsilon \Rightarrow |h(\hat{p}) - h(p)| < \delta(\varepsilon)$$

for some function δ such that $\delta(\varepsilon) \to 0$ as $\varepsilon \to 0$. (To obtain a more quantitive bound, you could use Fannes' inequality that your derive in problem 3.4.) In view of eq. (6.7), this implies that the subspace that $\tilde{P}_n$ projects on has dimension no larger than

$$\dim \tilde{\mathcal{H}}_n \leq (n+1)^2 2^{n(h(p)+\delta(\varepsilon))}. \tag{6.13}$$

Thus, the post-measurement state is supported on a possibly much smaller subspace of roughly $n(h(p) + \delta)$ qubits.

Let us end with a word of caution: In the coin flip example in section 5.1, the purpose of the compression scheme was to communicate Alice' actual sequence of coin flips to Bob – *not* for Bob to flip its own biased coin. The latter would only reproduce the probability distribution of the biased coin, but not the actual sequence of coin flips observed by Alice! In the same way, the purpose of a quantum compression scheme is *not* simply to produce the quantum state $\rho^{\otimes n}$ at Bob's side.

In fact, compression protocols are usually designed for known information sources. In the coin flip example, this means that Bob already knows the parameter p of the coin and could flip his own biased coin with no communication required at all. (Since its quantum analogue is the eigenvalue spectrum of ρ, you might in fact be concerned that spectrum estimation solves a problem that is completely irrelevant to compression.)

Next time, we will carefully define what it means to compress quantum information and see that the properties in eqs. (6.12) and (6.13) above are nevertheless precisely the properties required to solve the problem.

Bibliography

Michael E Beverland, Jeongwan Haah, Gorjan Alagic, Gretchen K Campbell, Ana Maria Rey, and Alexey V Gorshkov. Spectrum estimation of density operators with alkaline-earth atoms. 2016. arXiv:1608.02045.