8.S372/18.S996 Quantum Information Science III **Fall 2020**

Lecture 9: Sep 29, 2020

Lecturer: Aram Harrow *Scribe: John Martyn, Thao Dinh*

9.1 Details on the Packing Lemma

Regarding the packing lemma, one might be curious as to why we cannot simply use the POVM $\{\Pi_m\}$ in place of $\{\Lambda_m\}$. We can illustrate that this is not correct with a classical instance of the packing lemma. We also show how to correct this misconception and properly define $\{\Lambda_m\}$.

9.1.1 Classical Case

Look at the simple example in which the codebook is $C = \{1, 2\}$, and the projectors are $\Pi_1 = \text{diag}(1, 1, 1, 1, 0, 0, 0)$ and $\Pi_2 = \text{diag}(0, 0, 0, 1, 1, 1, 1)$. We will take as our signal states $\sigma_1 = \Pi_1/4$ and $\sigma_2 = \Pi_2/4$. As all the operators are diagonal, this is essentially a classical problem.

Anyhow, observe that $\Pi_1 + \Pi_2 = \text{diag}(1, 1, 1, 2, 1, 1, 1) \neq I$, so $\{\Pi_1, \Pi_2\}$ is not a valid POVM. Hence, we cannot simply use $\{\Pi_1, \Pi_2\}$ in place of $\{\Lambda_1, \Lambda_2\}$

However, in this scenario, we can instead use as a POVM $\Lambda_1 = \text{diag}(1, 1, 1, \frac{1}{2}, 0, 0, 0)$ and $\Lambda_2 = \text{diag}(0, 0, 0, \frac{1}{2}, 1, 1, 1)$. In general, when dealing with classcal instances of the packing lemma, we can obtain valid POVM $\{\Lambda_m\}$ as follows:

$$\Pi_{\text{total}} = \sum_m \Pi_m$$
$$\Lambda_m = \Pi_{\text{total}}^{-1} \Pi_m.$$

A quick calculation indices that this yields the above expression for $\{\Lambda_1, \Lambda_2\}$.

9.1.2 Quantum Case

In the quantum case of the packing lemma, we again cannot simply set $\Lambda_m = \Pi_m$. Instead, we can use the following prescription to construct Λ_m, which is analogous to the procedure employed above. Let $P_m = \Pi\Pi_m\Pi$ and $P_{\text{total}} = \sum_m P_m$. Note that

$P_m \geq 0$ because it is constructed as a symmetric product of projection operators. We then define $\Lambda_m = P_{\text{total}}^{-1/2} P_m P_{\text{total}}^{-1/2}$. If it is the case that P_{total} is not full rank, introduce an additinal projector Λ_{fail} such that $\sum_m \Lambda_m = I$. $\{\Lambda_m\}$ is then a valid POVM.

9.2 Packing Lemma Proof

We can use the above prescription to prove the packing lemma. To begin, let's look at the probability of incurring an error on message m, given codebook C:

$$p_{\text{err}}(m|C) = 1 - \text{tr}(\Lambda_m \sigma_m) = \text{tr}\Big((I - \Lambda_m)\sigma_m\Big)$$

To analyze this expression, we will employ the Hayashi-Nagaoka lemma: Given S and T, such that $0 \leq S \leq I$ and $T \geq 0$,

$$I - (S+T)^{-1/2} S (S+T)^{-1/2} \leq 2(I-S) + 4T.$$

Set $T = \sum_{m \neq m'} P_{m'} = P_{\text{total}} - P_m$, and $S = P_m$. These obey $0 \leq S \leq I$ and $T \geq 0$, so we can apply the Hayashi-Nagaoka lemma:

$$I - \Lambda_m = I - P_{\text{total}}^{-1/2} P_m P_{\text{total}}^{-1/2} =$$
$$I - (S+T)^{-1/2} S (S+T)^{-1/2} \leq 2(I - P_m) + 4 \sum_{m' \neq m} P_{m'}.$$

Using this result to evaluate $p_{\text{err}}(m|C)$, we have

$$p_{\text{err}}(m|C) \leq 2(1 - \text{tr}(P_m \sigma_m)) + 4 \sum_{m' \neq m} \text{tr}(P_{m'} \sigma_m)$$

Then, making use of the conditions assumed in the packing lemma, we can establish the bound

$$\text{tr}(P_m c_m) = \text{tr}(\Pi \Pi_m \Pi \sigma_m) = \text{tr}(\Pi_m \Pi \sigma_m \Pi) \geq$$
$$\text{tr}(\Pi_m \sigma_m) - ||\sigma_m - \Pi \sigma_m \Pi||_2 \geq 1 - \epsilon - 2\sqrt{\epsilon},$$

where we obtain the $\sqrt{\epsilon}$ as a result of the bound on gentle measurement proven in problem set 3. Next, we input the above bound into the expression for $p_{\text{err}}(m|C)$ and average this probability over messages m and codebooks C:

$$\mathop{\mathbb{E}}_{C} \mathop{\mathbb{E}}_{m} p_{\text{err}}(m|C) \leq 2(\epsilon + 2\sqrt{\epsilon}) + 4 \mathop{\mathbb{E}}_{m} \frac{1}{M} \sum_{m' \neq m} \text{tr}(P_{m'} \sigma_m).$$

Noting that $\mathop{\mathbb{E}}_{C} \sigma_m = \mathop{\mathbb{E}}_{C} \sigma_{C_m} = \sum_x p(x) \sigma_x = \sigma$, we can write the second term as

$$\frac{1}{M} \sum_{m \neq m'} \text{tr}(\Pi \Pi_{m'} \Pi) \mathop{\mathbb{E}}_{C} \sigma_m = \sum_{m \neq 1} \text{tr}(\Pi_{m'} \Pi \sigma \Pi) \leq (M-1) \text{tr}\Big(\Pi_{m'} \frac{I}{D}\Big) \leq M \frac{d}{D},$$

where we have employed the inequalities assumed in the packing lemma. Combining all the terms and inequalities above, we have

$$p_{\text{err}}(m|C) = 1 - \text{tr}(\Lambda_m \sigma_m) \leq 2\epsilon + 4\sqrt{\epsilon} + 4M\frac{d}{D} \quad \Rightarrow$$
$$\text{tr}(\Lambda_m \sigma_m) \geq 1 - 2\epsilon - 4\sqrt{\epsilon} - 4M\frac{d}{D}.$$

This is the claim of the packing lemma, which is now proven.

9.3 Aside: Pretty Good Measurement

Imagine that given a state $\sigma = \sum_x p(x)\sigma_x$, we wish to distinguish between the states σ_x. We can do this decently well with the "pretty good measurement" which is defined by the POVM $M_x = \sigma^{-1/2}p(x)\sigma_x\sigma^{-1/2}$. The Barnum-Knill theorem proves that the pretty good measurement can distinguish between the states σ_x with error probability

$$p_{\text{err}}(\text{Pretty Good Measurement}) \leq 2p_{\text{err}}(\text{Optimal Measurement}).$$

So in general, the "pretty good measurement" achieves an error probability that is comparable to the optimal error probability.

The pretty good measurement can be thought of as reversing the action of the channel $\mathcal{N} : x \to \sigma_x$, and applying this reversal to the state $\rho = \sum_x p(x)|x\rangle\langle x|$. In particular, if $\mathcal{N}$ has Kraus operators $\{E_k\}$, then the reversal of this channel, which we call the recovery channel, has Kraus operators $F_k = \rho^{1/2}E_k^\dagger\rho^{-1/2}$. This generalizes the "pretty good measurement" to a more general construction known as the "Petz recovery map".

9.4 Sequential Coding

In sequential decoding, one decodes a message by enumerating through the set of all possible message sequences. Specifically, we are given a state σ_x, and perform on it the set of measurements $\{\Pi, I - \Pi\}$, $\{\Pi_{c_1}, I - \Pi_{c_1}\}$, ..., $\{\Pi_{c_m}, I - \Pi_{c_m}\}$. These measurements dictate whether we fail or continue in the sequential coding procedure as follows

$$\Pi \to \text{continue}, \quad I - \Pi \to \text{fail}$$
$$\Pi_{c_m} \to \text{stop, output m}, \quad I - \Pi_{c_m} =: \hat{\Pi}_{c_m} \to \text{continue}$$

The probability that this procedure fails to output m is

$$p_{\text{err}}(m) = 1 - p_{\text{success}} = 1 - \text{tr}\Big(\Pi_{C_m}\hat{\Pi}_{c_{m-1}}...\hat{\Pi}_{c_1}\Pi\sigma_{c_m}\Pi\hat{\Pi}_{c_1}\hat{\Pi}_{c_{m-1}}\Pi_{C_m}\Big).$$

To analyze this expression, which we will do in the future, we will make use of the non-commutative union bound:

$$\omega \geq 0, \quad \mathrm{tr}(\omega) \leq 1, \quad P_1, ..., P_L = \text{set of projectors} \quad \Rightarrow$$
$$\mathrm{tr}(\omega) - \mathrm{tr}(P_L ... P_1 \omega P_1 ... P_L) \leq \sqrt{2 \sum_i \mathrm{tr}(\hat{P}_i \omega)}, \quad \hat{P}_i = I - P_i.$$

We will prove this relation next class. For now, we can observe that it is not at all obvious. Imagining that ω is a density matrix, the above quantity on the LHS will measure the difference between the density matrix, and the density matrix after a set of L projective measurements are applied to it. In general, applying a set of projective measurements can change the state drastically. For instance, imagine states

$$|\phi_j\rangle = \cos\left(\frac{\pi}{2}\frac{j}{L}\right)|0\rangle + \sin\left(\frac{\pi}{2}\frac{j}{L}\right)|1\rangle, \quad j = 1, ..., L,$$

to which we apply projectors

$$P_j = |\phi_j\rangle\langle\phi_j|.$$

With this setup, we have $\langle\phi_j|P_{j+1}|\phi_j\rangle = |\langle\phi_j|\phi_{j+1}\rangle|^2 = \cos^2(\frac{\pi}{2}\frac{1}{L}) = 1 - O(L^{-2})$, which indicates that applying the measurement P_{j+1} to $|\phi_j\rangle$, transitions one to state $|\phi_{j+1}\rangle$ with high probability. Therefore, one can begin in the state $|\phi_0\rangle \approx |0\rangle$ and end in $|\phi_L\rangle \approx |1\rangle$ with probability $1 - O(\frac{1}{L})$. These states are very different from each other (nearly orthogonal!), so it can be tricky to place a bound on $\mathrm{tr}(\omega) - \mathrm{tr}(P_L ... P_1 \omega P_1 ... P_L)$. We will discuss this further in the next class.