Lecture 7

19 Jun 2014

(1)

Classical Communication over a quantum channel.

Simplest model of a channel for this purpose:

pure-state cq channel

$x \to |\psi_x\rangle$ where $|\psi_x\rangle$ are non-orthogonal

Example:

If $0 \to |0\rangle$

$1 \to |+\rangle$ there is a non-zero error prob. in figuring out whether 0 or 1 was sent.

$|\langle 0|+\rangle|^2 = 1/2$

by repeating, i.e., coding, we get

$00 \to |0\rangle \otimes |0\rangle$

$11 \to |+\rangle \otimes |+\rangle$

$$\left| \left(\langle 0| \otimes \langle 0| \right) \left(|+\rangle \otimes |+\rangle \right) \right|^2$$

$$= |\langle 0|+\rangle|^2 \cdot |\langle 0|+\rangle|^2 = \frac{1}{4}$$

error prob. goes down

More generally, an $(|\mathcal{M}|, \varepsilon)$ code for a channel $x \to |\psi_x\rangle$

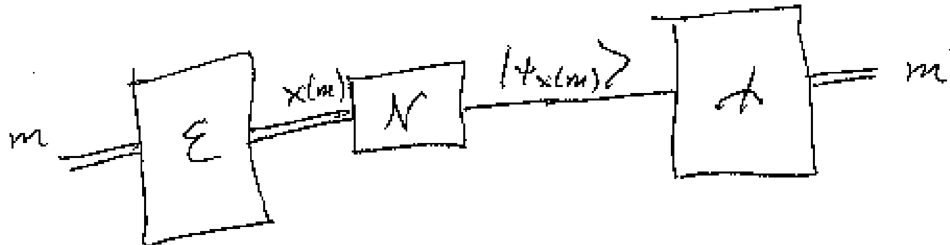
(think that input alphabet is very large, say, 2^{1000})

consists of an encoder $\mathcal{E} : [\mathcal{M}] \to \mathcal{X}$ + a decoding POVM $\{\Lambda_m\}$ such that average success probability $\geq 1 - \varepsilon$,

i.e.,

$$\frac{1}{|\mathcal{M}|} \sum_{m=1}^{|\mathcal{M}|} \mathrm{Tr}\{\Lambda_m \psi_{x(m)}\} \geq 1 - \varepsilon$$

Picture:



- $\mathcal{M}^*(\mathcal{N}, \varepsilon)$ maximum achievable code size for channel ψ_x is the largest $|\mathcal{M}|$ such that $\exists$ an $(|\mathcal{M}|, \varepsilon)$ code
- prove a lower bound on this quantity

Main obstacle for the quantum case:
How to build a decoder?

Simple idea: Consider the measurement

$$\left\{ \underbrace{|\psi_{x(m)}\rangle\langle\psi_{x(m)}|}_{\Pi_m},\ \underbrace{I - |\psi_{x(m)}\rangle\langle\psi_{x(m)}|}_{I - \Pi_m} \right\}$$

If performed on codeword $|\psi_{x(m)}\rangle$ the success probability is

$$\mathrm{Tr}\{\Pi_m |\psi_{x(m)}\rangle\langle\psi_{x(m)}|\} = 1$$

so that this kind of measurement should be helpful in building a decoder.

We could proceed w/ using a sequential decoder, in which the receiver performs these measurements one after another to figure out which codeword was sent, but there is a measurement strategy that outperforms this one, so we will consider this one instead.

Another idea that we will use is borrowed from Shannon. This is the idea of picking an encoder at random. This might seem like a silly strategy at first, but it works well for channels w/ large input alphabets & some more structure, so we will use it.

So, fix some probability distribution $P_X(x)$. To pick a code, for the 1st message, choose a codeword $x(1)$ at random according to $P_X(x)$. For the 2nd msg, independently choose $x(2)$ at random according to $P_X(x)$. Keep doing this for all messages.

Shannon's idea was to analyze the expectation of average error prob. w/ respect to random code, rather than to study the error prob. of any individual code. (easier to analyze the former)

(5)

An important density operator is the expected density operator for the random code, computed as

$$\rho = \sum_x p_X(x)\, |\psi_x\rangle\langle\psi_x|$$

Since $\{|\psi_x\rangle\}$ is not necessarily an orthonormal set, this might not be a spectral decomposition

Let the spectral decomposition be given by $\rho = \sum_z p_Z(z)\, |\phi_z\rangle\langle\phi_z|$

When choosing a code @ random, there are two different kinds of error that are important to consider

1) The probability of detecting the codeword $\psi_{x(m)}$
2) The probability of confusing w/ another codeword $\psi_{x(m')}$

~~Let $\Pi_{x(m)}$ now denote a quantum projector~~

Recall that we were thinking to have

$\Pi_{x(m)} = |\psi_{x(m)}\rangle\langle\psi_{x(m)}|$

The 1st kind of error is characterized by

$$1 - \mathrm{Tr}\{\Pi_{X(m)} \psi_{X(m)}\}$$

As before, we could show that this is zero if we choose $\Pi_{X(m)}$ as above.

The 2nd kind of error is characterized by

$$\mathrm{Tr}\{\Pi_{X(m')} \psi_{X(m)}\}$$

where $m' \neq m$

This is where Shannon's idea is helpful

We can instead consider

$$\mathbb{E}_{X(m'),X(m)}\{\mathrm{Tr}\{\Pi_{X(m')} \psi_{X(m)}\}\}$$

(expectation of error w.r.t. random code)

$$= \mathbb{E}_{X(m'),X(m)}\{\mathrm{Tr}\{\psi_{X(m')} \psi_{X(m)}\}\}$$

Since $X(m')$ & $X(m)$ are chosen independently, we can distribute the expectations to get

$$= \mathrm{Tr}\{\mathbb{E}_{X(m')}\{\psi_{X(m')}\} \mathbb{E}_{X(m)}\{\psi_{X(m)}\}\}$$

This then becomes

$$\mathrm{Tr}\{\rho \cdot \rho\} = \mathrm{Tr}\{\rho^2\}$$

Now, recall that our goal is to find a code w/ average error probability larger than $1-\varepsilon$

If this is our goal, it is clear that the 1st kind of error does not need to be zero. We can have some "leeway" to allow for a trade-off such that we meet this error constraint.

So for the decoding, we will instead consider $\Pi_\gamma \equiv \{\rho \leq 2^{-\gamma} I\}$
(the projection onto the eigenvectors of ρ w/ eigenvalues no larger than $2^{-\gamma}$)

This parameter γ will allow us to ~~tune~~ the performance in an optimal way.

So instead pick the decoding projectors to be

$$\Pi_{x(m)} = |\psi^{\gamma}_{x(m)}\rangle\langle\psi^{\gamma}_{x(m)}|$$

where $$|\psi^{\gamma}_{x(m)}\rangle = \frac{\Pi_\gamma |\psi_{x(m)}\rangle}{\| \Pi_\gamma |\psi_{x(m)}\rangle \|_2}$$

(projected + renormalize

Now consider the second kind of error

$$\mathbb{E}_{X(m'),X(m)}\left\{ \mathrm{Tr}\left\{ \Pi_{X(m')}\, \psi_{X(m)} \right\}\right\}$$

$$= \mathbb{E}_{X(m'),X(m)}\left\{ \mathrm{Tr}\left\{ \frac{\Pi_\gamma |\psi_{X(m')}\rangle\langle\psi_{X(m')}| \Pi_\gamma}{\langle\psi_{X(m')}|\Pi_\gamma|\psi_{X(m')}\rangle}\, \psi_{X(m)} \right\}\right\}$$

$$= \mathbb{E}_{X(m')}\, \mathrm{Tr}\left\{ \frac{\Pi_\gamma |\psi_{X(m')}\rangle\langle\psi_{X(m')}| \Pi_\gamma}{\langle\psi_{X(m')}|\Pi_\gamma|\psi_{X(m')}\rangle}\, \mathbb{E}_{X(m)}\{\psi_{X(m)}\} \right\}\Big\}$$

$$= \quad '' \quad '' \quad '' \quad '' \quad \rho \Big\}\Big\}$$

$$= \mathbb{E}_{X(m')}\, \mathrm{Tr}\Big\{ \underbrace{|\psi_{X(m')}\rangle\langle\psi_{X(m')}| \Pi_\gamma \rho \Pi_\gamma}_{''} \Big\}$$

$$\leq 2^{-\gamma}$$

So we would like to make this parameter γ as large as it can be while still having the 1st kind of error to be $\lesssim \varepsilon$

this motivates choosing it as

$$\gamma(\rho,\varepsilon) = \sup \{ a : \mathrm{Tr}\{\rho \Pi_a\} \geq 1-\varepsilon \} \quad (*)$$

Indeed if we do so, then one can show that the first kind of error $\leq \varepsilon$

Since the quantity in $(*)$ is important, let us call it the

<u>ε- spectral inf-entropy</u>

$$\underline{H}_s^{\varepsilon}(\rho) = \gamma(\rho,\varepsilon)$$

10

So now let's discuss the design of the decoder. Ideally, we would like to have it be just

$\{\Pi_{x(m)}\}$ these are all ≥ 0 but it might not be the case that

$$\sum_m \Pi_{x(m)} = I$$

(they might overshoot the identity)

so to make them form a legitimate POVM, we normalize them. So we make

$$\Lambda_m = \left[\sum_{m'} \Pi_{x(m')}\right]^{-1/2} \Pi_{x(m)} \left[\;''\;\right]^{-1/2}$$

So now we can analyze the error prob.

$$\mathrm{Tr}\{(I-\Lambda_m)\rho_{x(m)}\}$$

We can invoke the following operator ineq. due to Hayashi + Nagaoka (Lemma 3 of 0206186)

$$I-\Lambda_m \leq (1+c)(I-\Pi_{x(m)}) + (2+c+c^{-1})\sum_{m'\neq m} \Pi_{x(m')}$$

where $c > 0$

& we find that the error is bounded from above by

$$(1+c)\, Tr\{(I-\Pi_{x(m)})\, \rho_{x(m)}\} +$$
$$(2+c+c^{-1}) \sum_{m' \neq m} Tr\{\Pi_{x(m')}\, \rho_{x(m)}\}$$

using our bounds from before & invoking the expectation over a random choice of code we find that this is bounded from above by

$$(1+c)\, Tr\{(I-\Pi_\gamma)\,\rho\} +$$
$$(2+c+c^{-1})\, 2^{-\gamma} \cdot M$$

If we pick $\log M = \underline{H}_s^{\varepsilon-\eta}(\rho) - \log\left(\frac{4\varepsilon}{\eta^2}\right)$
where $\eta \in (0,\varepsilon)$

& $c = \dfrac{\eta}{2\varepsilon-\eta}$

we find that the error probability is bounded from above by the desired ε. derandomize code

So we have

$$\log M^{\#}(N,\varepsilon) \geq \underline{H}_s^{\varepsilon-\eta}(\rho) - \log\left(\frac{4\varepsilon}{\eta^2}\right)$$

A very important kind of channel model is memoryless channel, i.e., ~~there~~ small input alphabet but we can use the channel many times:

$$x_1 x_2 \cdots x_n \to |\psi_{x_1}\rangle \otimes |\psi_{x_2}\rangle \otimes \cdots \otimes |\psi_{x_n}\rangle$$

In such a case we can choose the distribution to be ~~[illegible]~~ IID

$$\cancel{P_{X^n}(x^n)} \prod_{i=1}^{n} P_X(x_i)$$

In this case, the average density operator is $\rho^{\otimes n}$, so that our theorem says that

$$\log(M^*(N^{\otimes n}, \varepsilon)) \geq \underline{H}_s^{\varepsilon - \eta}(\rho^{\otimes n}) - \log\left(\frac{4\varepsilon}{\eta^2}\right)$$

From definition of $\underline{H}_s^{\varepsilon-\eta}(\rho^{\otimes n})$, can show that this is equal to

$$\sup\left\{ a : \Pr\left\{ -\sum_{i=1}^{n} \log P_Z(Z_i) \leq a \right\} \leq \varepsilon - \eta \right\}$$

where Z_i is an R.V. distributed according to $P_Z(z)$ where $\rho = \sum_z P_Z(z) |\psi_z\rangle\langle\psi_z|$

(13)

Apply central limit theorem of Berry-Esseen, i.e.,

For $\{Y_i\}_{i=1}^n$ IID w/ mean μ & ~~standard~~ deviation s

letting $\bar{Y} = \frac{1}{n}\sum_i Y_i$

$$\left| \Pr\left\{ \sqrt{n}\,\frac{\bar{Y}-\mu}{s} \le y \right\} - \Phi(y) \right| \le \frac{C t^3}{s^3 \sqrt{n}}$$

can turn this around to get the expansion for $\eta = \frac{1}{\sqrt{n}}$ & n large enough

$$\underline{H}_s^{\varepsilon-\eta}(\rho^{\otimes n}) = n H(\rho) + \sqrt{n V(\rho)}\, \Phi^{-1}(\varepsilon) + O(1)$$

where
$$H(\rho) = -\mathrm{Tr}\{\rho \log \rho\}$$
$$V(\rho) = \mathrm{Tr}\{\rho\, [\log \rho - H(\rho)]^2\}$$

$$\Rightarrow \log M^*(N,\varepsilon) \ge n H(\rho) + \sqrt{n V(\rho)}\, \Phi^{-1}(\varepsilon) + O(\log n)$$