M525/P622 Quantum Information	Ninnat Dangniam
Homework Assignment 7	
35 points	DUE: 11 Oct 2024

1. Pretty good measurement (15 points).

Consider an ensemble decomposition of a density operator ρ:

$$\rho = \sum_{\alpha=1}^{N} p_\alpha |\psi_\alpha\rangle\langle\psi_\alpha| = \sum_\alpha |\overline{\psi}_\alpha\rangle\langle\overline{\psi}_\alpha|,$$

where the decomposition vectors $|\overline{\psi}_\alpha\rangle = \sqrt{p_\alpha}|\psi_\alpha\rangle$ lie in and span the support of ρ. In this problem, we restrict attention to the support, which we assume to be D-dimensional (with $D \leq N$), and forget about the rest of Hilbert space. This means in particular that ρ is invertible.

We can define a POVM that has POVM elements

$$E_\alpha = \rho^{-1/2}|\overline{\psi}_\alpha\rangle\langle\overline{\psi}_\alpha|\rho^{-1/2} = |\overline{\phi}_\alpha\rangle\langle\overline{\phi}_\alpha|,$$

where $|\overline{\phi}_\alpha\rangle = \rho^{-1/2}|\overline{\psi}_\alpha\rangle = \sqrt{p_\alpha}\rho^{-1/2}|\psi_\alpha\rangle$. The POVM elements are clearly positive operators and the POVM satisfies the completeness relation

$$\sum_\alpha E_\alpha = \rho^{-1/2}\underbrace{\left(\sum_\alpha |\overline{\psi}_\alpha\rangle\langle\overline{\psi}_\alpha|\right)}_{\rho}\rho^{-1/2} = \mathbb{1}.$$

This measurement is called the *pretty good measurement*.

(a) Show that the outcome probabilities $q_\alpha = \operatorname{Tr}(\rho E_\alpha)$, when the state is ρ, are the same as the ensemble probabilities p_α. This is the unique property of the pretty good measurement.

(b) The unique property in part **(a)** means that the preparation information inequality,

$$H(\mathbf{p}) \geq S(\rho),$$

and the POVM inequality,

$$H(\mathbf{q}) + \sum_\alpha q_\alpha \log(\operatorname{Tr} E_\alpha) \geq S(\rho),$$

are constraints on the same probability distribution $\mathbf{p} = \mathbf{q}$. Which of these inequalities provides the tighter constraint on $H(\mathbf{p})$?

(c) Show that

$$H(\mathbf{p}) \geq \sum_\alpha p_\alpha \log\left(\langle\psi_\alpha|\rho^{-1}|\psi_\alpha\rangle\right) \geq S(\rho).$$

(d) Show that if the decomposition vectors $|\overline{\psi}_\alpha\rangle$ are linearly independent, hence $N = D$, then $\langle\overline{\phi}_\alpha|\overline{\phi}_\beta\rangle = \delta_{\alpha\beta}$. For equal probabilities p_α, this is a "democratic" way to generate an ONB from a set of linearly independent vectors since it doesn't favor any vector in the set, unlike the Gram-Schmidt process.

(e) For the case $N = D = 2$ and equal prior probabilities, show that the pretty good measurement is the measurement that minimizes error probability.

---

2. Uniformly random quantum states (20 points).

Suppose that Alice prepares a pure state $|\varphi\rangle$ drawn from the uniform ensemble in a D-dimensional Hilbert space, and Bob performs a projective measurement associated to an ONB $\{|y\rangle\}$. The information Bob gains about Alice's preparation is quantified by the mutual information

$$H(\Phi : Y) = H(Y) - H(Y|\Phi), \tag{1}$$

where

$$H(Y|\Phi) = \sum_{y=1}^{D} \int d\Upsilon_{\varphi}\, p(y|\varphi) \log p(y|\varphi),$$
$$p(y|\varphi) = |\langle y|\varphi\rangle|^2 = \cos^2\theta,$$

and $d\Upsilon_{\varphi}$ is the unitarily invariant measure for rays in the projective Hilbert space.

(a) The measure $d\Upsilon_{\varphi}$ is proportional to $d\Gamma_{\varphi}$ in **Problem 4** of **HW6** for the isotropic measurement, but it is not exactly $d\Gamma_{\varphi}$ because POVM elements and states satisfy different kinds of constraints. Show that

$$d\Upsilon_{\varphi} = -(D-1)(1-\cos^2\theta)^{D-2} d(\cos^2\theta)$$

(b) Show that

$$H(\Phi : Y) = \log D - \frac{1}{\ln 2} \sum_{k=1}^{D-1} \frac{1}{k+1}$$

Hint: To evaluate the integral

$$\int_0^1 dx\, (1-x)^r x \log x,$$

observe that

$$x \ln x = \frac{d}{ds} x^s \big|_{s=1},$$

and integrate by parts.

(c) Show that in the limit of large Hilbert-space dimension D, the information gain becomes

$$H(\Phi : Y) = \frac{1-\gamma}{\ln 2} = 0.60995\ldots,$$

where $\gamma = 0.57721\ldots$ is the Euler's constant.

Our computed value of $H(Y|\Phi)$ may be interpreted in another way: Suppose that we fix an ONB measurement, choose a typical state, and perform the measurement repeatedly on (multiple copies of) that chosen state. Then the measurement outcomes will not be uniformly distributed. Instead the entropy of the outcomes will fall short of maximal by 0.60995 bits, in the limit of large Hilbert-space dimension.

This has an application in benchmarking of sampling-quantum-advantage experiments [1]. Under some plausible complexity-theoretic conjecture, generating samples by making measurements on a typical quantum state in the manner described in the previous paragraph is believed to be hard for classical computers, and this exercise suggests a way to distinguish between samples generated by an actual quantum state and samples made by a classical algorithm to imitate the quantum sampling.

(d) More formally, by repeating a large number m of measurement trials on a chosen typical state $|\varphi\rangle$, with high probability we would observe a string $\mathbf{y} = y_1 y_2 \dots y_m$ with probability $p(\mathbf{y}|\varphi) = p(y_1|\varphi)\dots p(y_m|\varphi)$ satisfying

$$-\frac{1}{m}\log p(\mathbf{y}|\varphi) = H(Y|\Phi = \varphi).$$

A classical algorithm trying to mimic the quantum sampling would have as input a specification of $|\varphi\rangle$ and output a string $\mathbf{y}^{\mathrm{cl}} = y_1^{\mathrm{cl}}\dots y_m^{\mathrm{cl}}$ with some probability $q(\mathbf{y}^{\mathrm{cl}}|\varphi) = q(y_1^{\mathrm{cl}}|\varphi)\dots q(y_m^{\mathrm{cl}}|\varphi)$. Fix this string $\mathbf{y}^{\mathrm{cl}}$ and consider the probability $p(\mathbf{y}^{\mathrm{cl}}|\varphi)$ of observing this string *if it were produced from the quantum state* $|\varphi\rangle$. Show that in the limit of a large number of trials, $m \to \infty$, with high probability,

$$-\frac{1}{m}\log p(\mathbf{y}^{\mathrm{cl}}|\varphi) = -\sum_{y=1}^{D} q(y|\varphi)\log p(y|\varphi) \equiv H(\mathbf{q},\mathbf{p}), \tag{2}$$

where $H(\mathbf{q},\mathbf{p}) = H(\mathbf{q}) + H(\mathbf{q}||\mathbf{p})$ is the *cross-entropy*.

(e) To compare (2) with the conditional entropy $H(Y|\Phi)$, we should take the average over $|\varphi\rangle$. Under the quantum-advantage assumption, the output of the classical algorithm would almost be statistically uncorrelated with $p(y|\varphi)$:

$$\left\langle -\sum_{y=1}^{D} q(y|\varphi)\log p(y|\varphi)\right\rangle_{\varphi} \approx -\sum_{y=1}^{D}\langle q(y|\varphi)\rangle_{\varphi}\langle \log p(y|\varphi)\rangle_{\varphi}. \tag{3}$$

Compute this average by first computing $\langle \log p(y|\varphi)\rangle_{\varphi}$. (This computation should be very similar to what you have done in part **(b)**.) Also show that the RHS of Eq.(3) is the cross-entropy $H(\mathbf{q}_{\mathrm{unif}},\mathbf{p})$ between the uniform distribution $q(y|\varphi) = 1/D$ and $p(y|\varphi)$.

The *cross-entropy benchmarking* (XEB) proposed in [1] compares the *cross-entropy difference* of a given classical algorithm:

$$\Delta H(\mathbf{q}) \equiv H(\mathbf{q}_{\mathrm{unif}},\mathbf{p}) - H(\mathbf{q},\mathbf{p}) = -\sum_{y=1}^{D}\left(\frac{1}{N} - q(y|\varphi)\right)\log p(y|\varphi),$$

to the cross-entropy difference of a realistic quantum experiment that produces samples $\mathbf{y}$ with errors due to noise and imperfection:

$$\Delta H(\mathbf{p}_{\mathrm{exp}}) = -\sum_{y=1}^{D}\left(\frac{1}{N} - p_{\mathrm{exp}}(y|\varphi)\right)\log p(y|\varphi),$$

In particular, the average $\langle \Delta H(\mathbf{p}_{\text{exp}}) \rangle_{\varphi}$ should not only be nonzero, but also larger than the best known classical algorithm for us to claim a demonstration of sampling quantum advantage. A major downside of XEB is that, to calculate the cross-entropy differences, we have to pre-compute $\log p(y|\varphi)$ of an ideal, error-free quantum experiment which simply cannot be done for a sufficiently large Hilbert space.[1]

References

[1] Boixo, S., Isakov, S.V., Smelyanskiy, V.N., Babbush, R., Ding, N., Jiang, Z., Bremner, M.J., Martinis, J.M. and Neven, H., *Characterizing quantum supremacy in near-term devices*, Nature Physics, **14**, 595-600 (2018).

[2] Arute, F, *et al.*, *Quantum supremacy using a programmable superconducting processor*, Nature **574**, 505-510 (2019).

[1]The first Google's sampling quantum advantage experiment [2] used a 53-qubit device, meaning that $D \approx 9.0 \times 10^{15}$.