نظریه اطلاعات کوانتمی ۱ ترم پاییز ۱۳۹۲-۱۳۹۱

مدرسین: ابوالفتح بیگی و امین زاده گوهری
نویسنده: رضا پورمحمدی

جلسه ۲۵

---

۱ نوعی بودن کوانتمی

همانند تئوری اطلاعات کلاسیک، در تئوری اطلاعات کوانتومی نیز مفهوم نوعی بودن از اهمیت زیادی برخوردار است. در قلمرو کلاسیک، از مفهوم دنباله‌های نوعی استفاده می‌کنیم؛ در حالی که در قلمرو کوانتومی صحبت از زیرفضاهای نوعی است. مشابه حالت کلاسیک مفهوم نوعی بودن زمانی معنی می یابد که تکرارهای مستقل زیادی از یک منبع داشته باشیم. همان طور که خواهیم دید قضایای مربوط به نوعی بودن کوانتمی تعمیمی از قضایای کلاسیک آن خواهند بود، با چند تفاوت. اول اینکه مفهوم نوعی بودن شرطی برای دو سیستم کوانتمی قابل تعریف نیست. این مفهوم تنها زمانی که حداقل یکی از سیستم‌ها کلاسیک باشد قابل تعریف است. تفاوت دوم این است که در حالت کلاسیک می‌توانیم اعضای یک دنباله را یکی یکی خوانده و تحقیق کنیم که آیا یک دنباله نوعی هست یا نیست. اما در حالت کوانتمی مشاهده و اندازه‌گیری سیستم‌ها خود سیستم را تغییر می‌دهد. بنابراین در دنیای کوانتمی تنها باید بدنبال اندازه گیری‌های نرمی باشیم که سیستم را «زیاد» تغییر ندهند. خواهیم دید که این کار امکان‌پذیر است.

برای شروع فرض کنید یک منبع اطلاعات کوانتومی، n سیستم مستقل دارای ماتریس چگالی یکسان را انتشار دهد. در این صورت ماتریس چگالی کل سیستم به صورت زیر است:

$$\rho^{S^n} = \rho^{S_1} \otimes \rho^{S_2} \otimes \cdots \otimes \rho^{S_n} = (\rho^S)^{\otimes n},$$

که حاصل ضرب تانسوری n حالت منتشر شده توسط منبع است. حال اگر هر یک از این n حالت را در پایه متعامد یکه تجزیه کنیم، مثلا $(\rho^{S_1} = \sum_{s_1} p(s_1)|s_1\rangle\langle s_1|)$ پس از ضرب و بسط جملات خواهیم داشت:

$$\begin{aligned}
(\rho^S)^{\otimes n} &= \big(\sum_{s_1} p(s_1)|s_1\rangle\langle s_1|\big) \otimes \big(\sum_{s_2} p(s_2)|s_2\rangle\langle s_2|\big) \otimes \cdots \otimes \big(\sum_{s_n} p(s_n)|s_n\rangle\langle s_n|\big) \\
&= \sum_{s^n} p(s_1)p(s_2)\cdots p(s_n) \bigotimes_{i=1}^{n} |s_i\rangle\langle s_i| \\
&= \sum_{s^n} p(s^n)|s^n\rangle\langle s^n|,
\end{aligned}$$

که $|s^n\rangle, p(s^n)$ به صورت زیر تعریف می‌شوند:

$$p(s^n) = \prod_{i=1}^{n} p(s_i) \quad , \quad |s_n\rangle = |s_1\rangle|s_2\rangle\cdots|s_n\rangle.$$

جملات شامل $p(s^n)$-هایی که s^n دنباله نوعی نیست، مجموعشان تقریبا صفر است.[۱] با توجه به این موضوع زیرفضای نوعی را به شکل زیر تعریف کردیم:

$$\mathcal{T}^n_{\rho,\delta} = span\{|s^n\rangle : \text{دنباله } s^n \text{ نوعی باشد}\}.$$

در این نمادگذاری گاهی برای سادگی $\mathcal{T}^n_{\rho,\delta}$ را با $\mathcal{T}_\delta$ نمایش می‌دهیم. از آنجایی که دنباله‌های $|s^n_1\rangle$ و $|s^n_2\rangle$ برای هر $s^n_1 \neq s^n_2$ بر هم عمود هستند بعد زیرفضای نوعی برابر است با

$$\dim(\mathcal{T}_\delta) = \left|\{|s^n\rangle : \text{دنباله } s^n \text{ نوعی باشد}\}\right|$$

تقریبا برابر $2^{nH(p)} = 2^{nH(\rho)}$ می‌باشد. بصورت دقیق‌تر

$$2^{n(H(p)-\epsilon)} = 2^{nH(\rho)-\epsilon} \leq \dim(\mathcal{T}_\delta) \leq 2^{n(H(p)+\epsilon)} = 2^{n(H(\rho)+\epsilon)}.$$

مثال ۱ فرض کنید که

$$\rho = p|0\rangle\langle 0| + \bar{p}|1\rangle\langle 1|$$

که در آن $\bar{p} = 1 - p$. زیرفضاهای نوعی برای n های خیلی بزرگتر از ۲ جالب هستند اما نوشتن دنباله‌ها برای حالت $n = 2$ آسان تر است. داریم:

$$\rho^{\otimes 2} = p^2|00\rangle\langle 00| + p\bar{p}|01\rangle\langle 01| + \bar{p}p|10\rangle\langle 10| + \bar{p}^2|11\rangle\langle 11|.$$

حال اگر $p = \frac{1}{2}$ باشد فرکانس تکرار نصف 0 و نصف 1 را انتظار داریم. پس می‌توان دنباله نوعی را دنباله‌ای با یک 0 و یک 1 بگیریم. در این صورت زیرفضای نوعی برابر خواهد بود با:

$$\mathcal{T} = span\{|01\rangle, |10\rangle\} = \{c_1|01\rangle + c_2|10\rangle : \quad c_1, c_2 \in \mathbb{C}\}$$

۱.۱ تصویرگر نوعی

تصویرگر نوعی[۲] عملگری است که تصویر روی زیرفضای نوعی را انجام می‌دهد. این عملگر و عملگر مکمل آن را هم به فرم زیر تعریف میشود:

$$\Pi^n_{\rho,\delta} = \sum_{s^n \in T_\delta} |s^n\rangle\langle s^n|,$$

$$I - \Pi^n_{\rho,\delta} = I - \sum_{s^n \in T_\delta} |s^n\rangle\langle s^n| = \sum_{s^n \notin T_\delta} |s^n\rangle\langle s^n|.$$

در اینجا هم گاهی برای سادگی $\Pi^n_{\rho,\delta}$ را با Π_δ نشان می‌دهیم. تصویرگر نوعی به ما اجازه می دهد تا اندازه‌گیری نوعی $M_1 = \Pi_\delta, M_0 = I - \Pi_\delta$ را تعریف کنیم. این اندازه‌گیری در واقع به ما نشان می‌دهد که آیا حالت ما یک حالت نوعی است یا خیر. در صورتی که حاصل اندازه گیری 1 باشد، حالت نوعی است و اگر 0 باشد، حالت نوعی نیست.

[۱] این نکته با توجه به خواص دنباله‌های نوعی در نظریه‌ی اطلاعات کلاسیک برقرار است.

[۲] Typical subspace projector

مثال ۲ برای مثال ۱ داریم:

$$\Pi_\delta = |01\rangle\langle 01| + |10\rangle\langle 10| = M_1$$

$$I - \Pi_\delta = |00\rangle\langle 00| + |11\rangle\langle 11| = M_0$$

پس از انجام این اندازه‌گیری داریم:

$$p_1 = tr(\Pi_\delta \rho^{\otimes 2} \Pi_\delta) = tr(\Pi \rho^{\otimes 2}), \qquad \rho^{\otimes 2} \longrightarrow \frac{\Pi_\delta \rho^{\otimes 2} \Pi_\delta}{p_1}$$

$$p_0 = tr((I - \Pi_\delta)\rho^{\otimes 2}(I - \Pi_\delta)) = tr((I - \Pi)\rho^{\otimes 2}), \qquad \rho^{\otimes 2} \longrightarrow \frac{(I - \Pi_\delta)\rho^{\otimes 2}(I - \Pi_\delta)}{p_0}.$$

توجه کنید که چون Π_δ و $\rho^{\otimes 2}$ در پایه متعامد یکه یکسانی همزمان قطری می‌شوند داریم:

$$\Pi_\delta \rho^{\otimes 2} \Pi_\delta = \Pi_\delta \Pi_\delta \rho^{\otimes 2} = \Pi_\delta \rho^{\otimes 2}.$$

در حالت کلی نیز چون Π_δ و $\rho^{\otimes 2}$ در پایه متعامد یکه یکسانی همزمان قطری می‌شوند داریم:

$$\Pi_\delta \rho^{\otimes n} \Pi_\delta = \Pi_\delta \Pi_\delta \rho^{\otimes n} = \Pi_\delta \rho^{\otimes n}.$$

همچنین این حاصل ضرب را می‌توان حساب کرد:

$$\Pi_\delta \rho^{\otimes n} \Pi_\delta = \Pi_\delta \rho^{\otimes n} = \rho^{\otimes n} \Pi_\delta = \sum_{s^n \text{ نوعی}} p(s^n) |s^n\rangle\langle s^n|.$$

زیرا

$$\begin{aligned}
\rho^{\otimes n} \Pi_\delta &= \big(\sum_{s^n} p(s^n) |s^n\rangle\langle s^n| \big) \big(\sum_{s'^n \text{ نوعی}} |s'^n\rangle\langle s'^n| \big) \\
&= \sum_{s^n, s'^n \text{ نوعی}} p(s^n) |s^n\rangle\langle s^n | s'^n\rangle\langle s'^n| \\
&= \sum_{s^n, s'^n \text{ نوعی}} p(s^n) \delta[s^n = s'^n] \, |s^n\rangle\langle s'^n| \\
&= \sum_{s^n \text{ نوعی}} p(s^n) |s^n\rangle\langle s^n|.
\end{aligned}$$

بنابراین احتمال اینکه حاصل اندازه گیری برابر یک باشد مساوی است با

$$\begin{aligned} p_1 &= \mathrm{tr}(\Pi_\delta \rho^{\otimes n} \Pi_\delta) \\ &= \mathrm{tr}\big(\sum_{\text{نوعی } s^n} p(s^n) |s^n\rangle\langle s^n| \big) \\ &= \sum_{\text{نوعی } s^n} p(s^n) \mathrm{tr}\big(|s^n\rangle\langle s^n|\big) \\ &= \sum_{\text{نوعی } s^n} p(s^n) \\ &\geq 1-\epsilon \end{aligned}$$

چند خاصیت زیر برای اندازه‌گیری نوعی به شرح زیر است:

۱. در صورتی که n سیستم مستقل که در حالت ρ آماده شده‌اند را توسط اندازه‌گیری نوعی مربوط به آنها اندازه‌گیری کنیم، حاصل با احتمال بالا 1 خواهد بود. به عبارت دیگر:

$$\mathrm{tr}(\Pi_\delta \rho^{\otimes n}) = \sum_{s^n \in T_\delta} p(s^n) \geq 1-\epsilon.$$

اثبات این خاصیت را در بالا دیدیم. دلیل اصلی این خاصیت این است که مجموع جملات $p(s^n)$ که s^n دنباله نوعی نیست، تقریبا صفر است.

۲. با توجه به بعد زیرفضای نوعی می‌توان اثر عملگر تصویر به این زیر فضا را یافت:

$$\mathrm{tr}(\Pi_\delta) = \dim(\mathcal{T}_\delta) \simeq 2^{nH(\rho)}.$$

بصورت دقیق تر $\mathrm{tr}(\Pi_\delta) \in [2^{n(H(\rho)-\epsilon)}, 2^{n(H(\rho)+\epsilon)}]$. دلیل اینکه $\mathrm{tr}(\Pi_\delta) = \dim(\mathcal{T}_\delta)$ این است که

$$\mathrm{tr}(\Pi_\delta) = \mathrm{tr}(\sum_{\text{نوعی } s^n} |s^n\rangle\langle s^n|) = \sum_{\text{نوعی } s^n} \mathrm{tr}(|s^n\rangle\langle s^n|) = \big|\{|s^n\rangle : \text{دنباله } s^n \text{ نوعی باشد}\}\big| = \dim(\mathcal{T}_\delta).$$

۳. چون برای هر دنباله s^n نوعی داریم

$$(1-\epsilon)2^{-n(H(p)+\epsilon)} \leq p(s^n) \leq 2^{-n(H(p)-\epsilon)},$$

پس

$$(1-\epsilon)2^{-n(H(p)+\epsilon)}\Pi_\delta \leq \Pi_\delta \rho^{\otimes n} \Pi_\delta \leq 2^{-n(H(p)-\epsilon)}\Pi_\delta.$$

این رابطه از اینجا نتیجه می‌شود که

$$\Pi_\delta \rho^{\otimes n} \Pi_\delta = \sum_{\text{نوعی } s^n} p(s^n) |s^n\rangle\langle s^n|.$$

کافی است که از کران بالایی و پایینی روی $p(s^n)$ استفاده کنیم.

۴. چون نتیجه اندازه‌گیری زیرفضای نوعی روی حالت $\rho^{\otimes n}$ با احتمال زیاد برابر 1 می‌باشد، طبق لم اندازه‌گیری نرم[۳]

[۳] Gentle Measurement Lemma

که در تمرین‌ها داشتیم، حالت سیستم قبل و بعد از اندازه‌گیری تفاوت چندانی نمی‌کنند:

$$\|\frac{1}{p_1}\Pi_\delta\rho^{\otimes n}\Pi_\delta - \rho^{\otimes n}\|_1 \leq 2\sqrt{\epsilon}.$$

که در آن $p_1 = \text{tr}(\Pi_\delta\rho^{\otimes n}\Pi_\delta)$ احتمال مشاهده یک است. این قضیه در مورد حالت نرمال نشده هم برقرار است:

$$\|\Pi_\delta\rho^{\otimes n}\Pi_\delta - \rho^{\otimes n}\|_1 \leq 2\sqrt{\epsilon}.$$

۵. اندازه‌گیری زیرفضای نوعی را می‌توان به عنوان یک دینامیک کوانتمی در نظر گرفت که یک حالت دلخواه σ را این گونه تغییر می‌دهد:

$$\sigma \mapsto (I-\Pi_\delta)\sigma(I-\Pi_\delta)\otimes|0\rangle\langle 0|^E + \Pi_\delta\sigma\Pi_\delta\otimes|1\rangle\langle 1|^E$$

که در بالا جواب اندازه‌گیری را در سیستم E ذخیره می‌شود. جهت تحقیق درستی این رابطه توجه کنید که پس از اندازه‌گیری یک هنگرد به شکل زیر خواهیم داشت: با احتمال $p_0 = \text{tr}((I-\Pi_\delta)\sigma(I-\Pi_\delta))$ حاصل اندازه گیری صفر بوده و حالت σ به حالت $\frac{1}{p_0}(I-\Pi_\delta)\sigma(I-\Pi_\delta)$ سقوط میکند. پس با احتمال p_0 در حالت $\frac{1}{p_0}(I-\Pi_\delta)\sigma(I-\Pi_\delta)\otimes|0\rangle\langle 0|^E$ قرار داریم. مشابها با احتمال p_1 در حالت $\frac{1}{p_1}\Pi_\delta\sigma\Pi_\delta\otimes|1\rangle\langle 1|^E$ قرار داریم و هنگرد مربوط به خروجی همان عبارت داده شده خواهد بود.

۲ اندازه‌گیری سیستم‌های ترکیبی

سیستم دوتایی XB را در نظر بگیرید که در آن X کلاسیک و B کوانتومی باشد و هنگرد کلاسیک-کوانتومی زیر را تشکیل دهند:

$$\{p(x), |x\rangle\langle x|\otimes\rho_x^B\}.$$

در این صورت ماتریس چگالی $\rho^{XB} = \sum_x p(x)|x\rangle\langle x|\otimes\rho_x^B$ خواهد بود و ماتریس چگالی برای n حالت کوانتومی که این منبع انتشار می‌دهد، برابر است با:

$$(\rho^{XB})^{\otimes n} = \sum p(x^n)|x^n\rangle\langle x^n|\otimes\rho_{x^n}^{B^n},$$

که در آن :

$$\rho_{x^n}^{B^n} = \rho_{x_1}^{B_1}\otimes\rho_{x_2}^{B_2}\otimes\cdots\otimes\rho_{x_n}^{B_n}, \qquad p(x^n) = \prod_{i=1}^n p(x_i).$$

مجددا جملات شامل $p(x^n)$-هایی که x^n دنباله نوعی نیست، مجموعشان تقریبا صفر است. به عبارت دیگر اگر فرض کنیم که X مقادیر $1,2,...,m$ را با احتمال‌های به ترتیب $p_1,p_2,...,p_m$ اخذ کند، آن گاه دنباله‌هایی که تقریبا np_1 تا یک، np_2 تا دو، و ... دارند دنباله‌های نوعی را تشکیل داده و جمع احتمال آنها تقریبا یک است.

یک دنباله نوعی x^n را در نظر بگیرید و برای راحتی فرض کنید که مقادیر یکسان در دنباله پشت سر هم قرار گرفته باشند: در ابتدا np_1 تا یک، سپس np_2 تا دو، ... ظاهر شوند:

$$x^n = \underbrace{11\cdots 1}_{np_1}\underbrace{22\cdots 2}_{np_2}\cdots\cdots\underbrace{mm\cdots m}_{np_m}. \tag{۱}$$

در این صورت داریم:

$$\rho_{x_n}^{B_n} = \left(\rho_1^{\otimes np_1}\right)^{B_1:B_{np_1}} \otimes \left(\rho_2^{\otimes np_2}\right)^{B_{np_1+1}:B_{np_1+np_2}} \otimes \cdots \left(\rho_m^{\otimes np_m}\right)^{B_{n-np_m+1}:B_n}. \tag{۲}$$

تعریف ۳ برای دنباله نوعی خاص x^n که در بالا تعریف شد عملگر تصویر نوعی شرطی را با استفاده از عملگرهای تصویر نوعی برای اجزا به فرم زیر تعریف می‌کنیم:

$$\Pi_\delta^{B^n|x^n} = (\Pi_{\rho_1,\delta}^{np_1})^{B_1:B_{np_1}} \otimes (\Pi_{\rho_2,\delta}^{np_2})^{B_{np_1+1}:B_{np_1+np_2}} \otimes \cdots (\Pi_{\rho_m,\delta}^{np_m})^{B_{n-np_m+1}:B_n} = \bigotimes_x \Pi_{\rho_x,\delta}^{np_x}.$$

توجه کنید که منظور از $\Pi_{\rho_i,\delta}^{np_i}$ عملگر تصویر نوعی متناظر با ρ_i است و مثلا داریم $tr(\Pi_{\rho_i,\delta}^{np_i}\rho_i^{np_i}) \geq 1-\epsilon$. همچنین توجه کنید که ضرب تانسوری عملگرهای تصویری، عملگری تصویری است. پس $\Pi_\delta^{B^n|x^n}$ نیز یک عملگر تصویر است و به آن تصویر نوعی شرطی (با شرط $X^n = x^n$) می‌گویند.

عملگر تصویر نوعی شرطی در حالت کلی، و نه فقط وقتی x^n فرم خاص (۱) را داشته باشد، نیز تعریف می‌شود. هر دنباله‌ی نوعی دلخواه x^n تقریبا np_i تا i دارد. پس تحت یک جایگشت κ (از اندیس‌های 1 تا n) همان فرم (۱) را دارد. در این صورت برای تعریف $\Pi_\delta^{B^n|x^n}$ کافی است همان عملگری که در بالا تعریف شد را در نظر گرفته و بعد جایگشت κ را روی اندیس‌های $B_1,\ldots,B_n$ اعمال کنیم.

۱.۲ اندازه‌گیری با تصویرگر نوعی شرطی

در صورتی که با عملگر تصویر نوعی شرطی $\{M_1 = \Pi_\delta^{B^n|x^n}, M_0 = I - \Pi_\delta^{B^n|x^n}\}$ حالت $\rho_{x^n}^{B^n}$ را اندازه‌گیری کنیم، خواهیم داشت:

$$\mathrm{tr}(\Pi_\delta^{B^n|x^n}\rho_{x^n}^{B^n}) = \mathrm{tr}(\Pi_{\rho_1,\delta}^{np_1}\rho_1^{\otimes np_1})\mathrm{tr}(\Pi_{\rho_2,\delta}^{np_2}\rho_2^{\otimes np_2})\cdots\mathrm{tr}(\Pi_{\rho_m,\delta}^{np_m}\rho_m^{\otimes np_m}),$$

که رابطه بالا را با توجه به این که اثر ضرب تانسوری دو عملگر برابر ضرب اثرهای آنها است نوشته‌ایم. توجه کنید که این رابطه حتی اگر x^n فرم (۱) را نداشته باشد نیز برقرار است زیرا جایگشتی که روی x^n و جایگشتی که روی اندیس‌های B_i اعمال می‌شوند یکسانند.

براساس خاصیت اول اندازه‌گیری زیرفضای نوعی حاصل هرکدام از اثرهای فوق $\mathrm{tr}(\Pi_{\rho_i,\delta}^{np_i}\rho_i^{\otimes np_i})$ نزدیک 1 است، پس داریم:

$$\mathrm{tr}(\Pi_\delta^{B^n|x^n}\rho_{x^n}^{B^n}) \geq (1-\epsilon)^m.$$

خواص اندازه گیری تصویرگر نوعی شرطی:

۱. نزدیک یک بودن احتمال حاصل اندازه‌گیری

$$\mathrm{tr}(\Pi_\delta^{B^n|x^n}\rho_{x^n}^{B^n}) \geq (1-\epsilon).$$

۲. بعد زیرفضایی که بر روی آن تصویر می‌شود

$$\begin{aligned}\mathrm{tr}(\Pi_\delta^{B^n|x^n}) &\simeq 2^{np_1H(\rho_1)} \times 2^{np_2H(\rho_2)} \times \cdots \times 2^{np_mH(\rho_m)}\\ &= 2^{n\sum_i p_iH(\rho_i)}\\ &= 2^{nH(B|X)}.\end{aligned}$$

۳. $\Pi_\delta^{B^n|x^n}$ و $\rho_{x^n}^{\otimes n}$ با هم جابجا می‌شوند چون ضرب تانسوری عملگرهایی هستند که خود با هم جابجا می‌شوند.

۴. داریم:

$$(1-\epsilon)\frac{1}{2^{n(H(B|X)+\epsilon)}}\Pi_\delta^{B^n|x^n} \leq \Pi_\delta^{B^n|x^n}\rho_{x^n}^{\otimes n}\Pi_\delta^{B^n|x^n} \leq \frac{1}{2^{n(H(B|X)-\epsilon)}}\Pi_\delta^{B^n|x^n}.$$

اثبات: مانند قبل بدون از دست رفتن کلیت مساله، فرض می‌کنیم که x^n به صورت (۱) و در نتیجه $\rho_{x^n}^{B^n}$ به فرم (۲) است. با استفاده از خاصیت اندازه‌گیری زیرفضای نوعی داریم:

$$(1-\epsilon)\frac{1}{2^{np_i(H(\rho_i)+\epsilon)}}\Pi_{\rho_i,\delta}^{np_i} \leq \Pi_{\rho_i,\delta}\rho_i^{\otimes np_i}\Pi_{\rho_i,\delta} \leq \frac{1}{2^{np_i(H(\rho_i)-\epsilon)}}\Pi_{\rho_i,\delta}^{np_i}.$$

حال با استفاده از این که اگر $A \leq B, C \leq D$، آنگاه $A\otimes C \leq B\otimes D$ می‌توان اثبات را کامل کرد. □

اندازه‌گیری یک حالت شرطی با استفاده از عملگر تصویر نوعی غیر مشروط

فرض کنید که n نسخه از حالت کلاسیک-کوانتمی XB تولید شده و دنباله X^n نزد آذر و سیستم B^n نزد بابک است. اگر آذر دنباله x^n را مشاهده کند، از نظر او حالت سیستم بابک $\rho_{x^n}^{B^n}$ خواهد بود. اما از نقطه نظر بابک که به دنباله x^n دسترسی ندارد، حالت سیستمش $\rho^{\otimes n}$ می‌باشد که در آن ρ حالت متوسط سیستم B است:

$$\rho = \sum_x p(x)\rho_x.$$

از نقطه نظر بابک اندازه‌گیری با استفاده از عملگر تصویر نوعی $\Pi_{\rho,\delta}^n$ با احتمال زیاد جواب 1 خواهد داد. اما اگر بابک این اندازه گیری را انجام دهد، از نقطه نظر آذر چه اتفاقی می‌افتد؟ ثابت می‌کنیم که اگر دنباله x^n مشاهده شده توسط آذر نوعی باشد، آن گاه با احتمال زیاد از نقطه نظر آذر نیز جواب اندازه‌گیری بابک برابر 1 است. به عبارت دیگر برای هر دنباله نوعی x^n

$$\mathrm{tr}(\Pi_{\rho,\delta}^n\rho_{x^n}^{B^n}) \approx 1$$

توجه کنید که این گزاره تعمیم این گزاره در حالت کلاسیک است که "اگر x^n نوعی باشد و y^n با عبور x^n از کانال $p(y|x)$ تولید شده باشد ، آنگاه با احتمال زیاد y^n نیز نوعی است."

اثبات: طبق تعریف عملگر تصویر نوعی $\Pi^n_{\rho,\delta} = \sum_{s^n:\text{نوعی}} |s^n\rangle\langle s^n|$ داریم:

$$\begin{aligned}
\mathrm{tr}(\Pi^n_{\rho,\delta}\rho^{B^n}_{x^n}) &= \mathrm{tr}(\sum_{s^n\text{نوعی}} |s^n\rangle\langle s^n|\rho^{B^n}_{x^n}) \\
&= \sum_{s^n\text{نوعی}} \langle s^n|\rho^{B^n}_{x^n}|s^n\rangle \\
&= \sum_{s^n\text{نوعی}} \prod_{i=1}^{n} \langle s_i|\rho_{x_i}|s_i\rangle.
\end{aligned}$$

حال یک توزیع شرطی به شکل زیر تعریف می‌کنیم:

$$p_{S|X}(s|x) := \langle s|\rho_x|s\rangle.$$

این یک توزیع شرطی مشروع است زیرا اولا

$$\langle s_i|\rho_{x_i}|s_i\rangle \geq 0,$$

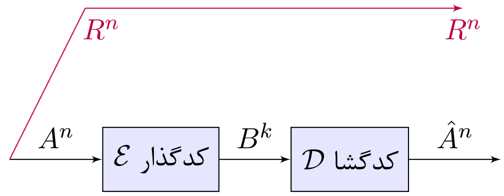
و ثانیا برای هر x داریم

$$\begin{aligned}
\sum_s \langle s|\rho_x|s\rangle &= \sum_s \mathrm{tr}(\langle s|\rho_x|s\rangle) \\
&= \sum_s \mathrm{tr}(\rho_x|s\rangle\langle s|) \\
&= \mathrm{tr}(\rho_x \sum_s |s\rangle\langle s|) \\
&= \mathrm{tr}(\rho_x I) \\
&= 1.
\end{aligned}$$

حال داریم:

$$\begin{aligned}
\mathrm{tr}(\Pi^n_{\rho,\delta}\rho^{B^n}_{x^n}) &= \sum_{s^n:\text{نوعی}} \prod_{i=1}^{n} \langle s_i|\rho_{x_i}|s_i\rangle \\
&= \sum_{s^n:\text{نوعی}} \prod_{i=1}^{n} p_{S|X}(s_i|x_i) \\
&= \sum_{s^n:\text{نوعی}} p_{S^n|X^n}(s^n|x^n) \\
&= p(\mathcal{T}_{S\ \text{نوعی}}|x^n) \\
&\approx 1,
\end{aligned}$$

که در اینجا از گزاره‌ی کلاسیکی که در بالا به آن اشاره شد استفاده کردیم. □



شکل ۱: نمایش شماتیک یک کدگذار منبع کوانتمی به همراه محض کننده منبع

۳ کدگذاری کانال(فشرده‌سازی شوماخر)

همان طور که قبلا به صورت مفصل بحث شد یک کدگذار منبع کوانتمی را می‌توان به صورت زیر تعریف کرد. فرض کنید که تکرارهای .i.i.d منبع دلخواهی مانند ρ را در اختیار داریم. ρ یک ماتریس چگالی روی فضای هیلبرت دلخواهی مانند $\mathcal{H}_A$ است. هدف فشرده‌سازی n نسخه مستقل A^n از این منبع است که حالت مشترک $\rho^{\otimes n}$ دارند. همان طور که در شکل ۱ نشان داده شده است یک کد کوانتمی منبع از یک کدگذار $\mathcal{E}^{A^n \to B^k}$ و یک کدگشا $\mathcal{D}^{B^k \to \hat{A}^n}$ تشکیل شده است که هر دو فرایندهای کوانتمی هستند. فرایند کوانتمی $\mathcal{E}$ سیستم‌های $A_1, A_2, \cdots, A_n$ را به عنوان ورودی گرفته و در خروجی k کیوبیت $B_1, B_2, \cdots, B_k$ را که حالت مشترک آنها را با $\sigma_{B^k} = \mathcal{E}(\rho^{\otimes n})$ نشان داده‌ایم، تولید می‌کند. پس فضای هیلبرت σ_{B^k} یک فضای 2^k بعدی است. در انتها کدگشا نیز یک فرایند کوانتمی است که هدفش بازیابی منبع است ($\mu_{\hat{A}^n} = \mathcal{D}(\sigma_{B^k})$). نرخ این کدگذاری $\frac{k}{n}$ است.

یک کد خطای ϵ دارد اگر برای محض‌سازی دلخواهی از $\rho^{\otimes n}$ که می‌توانیم فرض کنیم به فرم $(\rho^{AR})^{\otimes n} = \rho^{A^n R^n}$ است داشته باشیم:

$$\|\rho_{A^n R^n} - (\mathcal{D} \otimes \mathcal{I}_{R^n})(\mathcal{E} \otimes \mathcal{I}_{R^n})\rho_{A^n R^n}\|_1 \leq \epsilon.$$

قضیه زیر شرط لازم و کافی برای کدگذاری منبع را بیان می‌کند. این قضیه تعمیم قضیه کدگذاری منبع شانون است.

قضیه ۴ (قضیه فشرده سازی شوماخر) فشرده سازی با خطای به اندازه دلخواه کوچک قابل انجام است اگر و فقط اگر

$$R = \frac{k}{n} > H(\rho).$$

ابتدا اثبات وارون قضیه را بیان می‌کنیم. این اثبات بسیار شبیه اثبات کلاسیک است، پس ابتدا اثبات حالت کلاسیک را یاد آوری می‌کنیم. برای حالت کلاسیک اگر خروجی کدگذار را با B^k نمایش دهیم، داریم:

$$H(A^n) = nH(A) \cong I(A^n; \hat{A}^n) \leq I(A^n; B^k) \leq H(B^k) \leq k$$

$$\Rightarrow \frac{k}{n} \geq H(A).$$

در اینجا از این نکته که $\hat{A}^n$ تقریبا یک کپی از A^n استفاده کردیم و همچنین از نامساوی پردازش داده‌ها.

در حالت کوانتومی تساوی $H(A^n) = nH(A)$ همچنان برقرار است. اما عبارت $I(A^n; \hat{A}^n)$ بی‌معنی است زیرا A^n و $\hat{A}^n$ در یک زمان واحد وجود ندارند. اگر بخواهیم اثبات حالت کلاسیک را بگونه‌ای بازنویسی کنیم که A^n و $\hat{A}^n$ هم‌زمان

ظاهر نشوند، آنگاه می‌توانیم در ابتدا یک کپی از A^n بشکل $A'^n = A^n$ برداشته و سپس بجای $I(A^n;\hat{A}^n)$ بنویسیم $I(A'^n;\hat{A}^n)$. اما در دنیای کوانتمی نمی‌توان از A^n نسخه برداری کرد. به همین دلیل از R^n که یک محض‌سازی A^n است استفاده می‌کنیم. داریم:

$$I(A^n;R^n) = 2H(A^n) = 2nH(A).$$

پس می‌بینیم که $I(A^n;R^n)$ برای ما راه گشا است زیرا این اطلاعات متقابل متناسب با همان $nH(A)$ است، منتها با یک ضریب 2 اضافه که مهم نیست. همچنین انتظار داریم که با توجه به کم بودن خطا $I(A^n;R^n) \simeq I(\hat{A}^n;R^n)$ برقرار باشد. برای اثبات این نامساوی به تعمیمی از قضیه فانو نیاز داریم که در ادامه خواهد آمد. انجام مرحله مربوط به نامساوی پردازش داده مشکل نیست:

$$I(R^n;\hat{A}^n) \leq I(R^n;B^k).$$

همچنین ارتباط دادن $I(R^n;B^k)$ با $H(B^k)$ نیز قابل انجام است

$$I(R^n;B^k) = H(B^k) - H(B^k|R^n) \leq 2H(B^k),$$

که ظاهر شدن این ضریب 2 اضافه نهایتا مشکلی را ایجاد نخواهد کرد. با کنار هم گذاشتن این نامساوی‌ها به اثبات زیر خواهیم رسید:

$$\begin{aligned}
2nH(A) &= I(A^n;R^n) \\
&\simeq I(\hat{A}^n;R^n) \\
&\leq I(B^k;R^n) = H(B^k) - H(B^k \mid R^n) \\
&\leq H(B^k) + H(B^k) \\
&\leq 2H(B^k) \\
&\leq 2k
\end{aligned}$$

در نتیجه $H(A) \leq \frac{k}{n}$.

۱.۳ نامساوی فانو و تعمیم کوانتمی آن

یکی از مراحل اصلی اثبات فوق نامساوی فانو بود. فرض کنید می‌خواهیم مقدار متغیر تصادفی X را بر اساس اطلاعاتی که از متغیر تصادفی Y داریم بیابیم. فرض کنید $\hat{X} = f(Y)$ تخمین ما از X بر حسب Y باشد. اگر ϵ را احتمال نادرست بودن تخمین در نظر بگیریم براساس نامساوی فانو داریم:

$$h(\epsilon) + \epsilon \cdot \log(|\mathcal{X}| - 1) \geq H(X|Y),$$

که در آن

$$h(\epsilon) = \epsilon \log \frac{1}{\epsilon} + \bar{\epsilon} \log \frac{1}{\bar{\epsilon}},$$

تابع آنتروپی دودویی است و منظور از $|\mathcal{X}|$ اندازه مجموعه الفبای X است. اولین مشکل در تعمیم کوانتمی این نامساوی مفهوم احتمال خطا است. زمانی که دو سیستم کوانتمی دلخواه داریم، نمی‌توانیم صحبت از احتمال مساوی بودن آنها بکنیم زیرا این مفهوم مساوی بودن به مفهوم اندازه‌گیری وابسته است و نتیجه اندازه‌گیری در حالت کوانتمی می‌تواند تصادفی باشد. به همین دلیل بجای مفهوم احتمال خطا از فاصله اثر استفاده می‌کنیم. پیش از بحث بیشتر در مورد ارتباط فاصله اثر و احتمال خطا تعمیم نامساوی فانو را بیان می‌کنیم.

قضیه ۵ نامساوی فینز-آدنرت[۴]: برای هر دو حالت ρ و σ روی یک فضای هیلبرت با بعد d داریم:

$$|H(\rho) - H(\sigma)| \leq \epsilon \log(d-1) + h(\epsilon)$$

که در آن $\epsilon = \frac{1}{2}\|\rho - \sigma\|_1$.

نشان خواهیم داد که نامساوی بالا در حالت کلاسیک از نامساوی فانو نتیجه می‌شود. نامساوی بالا همچنین پیوستگی تابع آنتروپی را نسبت به فاصله اثر را بیان می‌کند.

۲.۳ تکمیل وارون قضیه شوماخر

حال بخش مربوط به اثبات وارون قضیه شوماخر را به صورت دقیق ثابت می‌کنیم. توجه کنید که

$$I(A^n; R^n) = H(A^n) + H(R^n) - H(A^n, R^n)$$

$$I(\hat{A}^n; R^n) = H(\hat{A}^n) + H(R^n) - H(\hat{A}^n, R^n)$$

مقدار $H(R^n)$ در ابتدای فرایند و انتهای آن یکی است زیرا تمام اندازه گیری‌ها و فرایندهای کوانتمی روی بخش‌های دیگر سیستم زده شده و طبق قضیه عدم علامت دهی ماتریس چگالی کاهش یافته R^n نباید تغییری کند. طبق شرط مربوط به خطا

$$\|\rho_{A^nR^n} - \mu_{\hat{A}^nR^n}\|_1 \leq \epsilon.$$

پس

$$\begin{aligned}|H(A^n, R^n) - H(\hat{A}^n, R^n)| &\leq \epsilon \log(d_A^n - 1) + h(\epsilon)\\ &\leq n\epsilon \log(d_A) + h(\epsilon)\end{aligned}$$

کوچک است. همچنین

$$\|\rho_{A^n} - \mu_{\hat{A}^n}\|_1 \leq \|\rho_{A^nR^n} - \mu_{\hat{A}^nR^n}\|_1 \leq \epsilon$$

پس $|H(A^n) - H(\hat{A}^n)|$ نیز کوچک است. نتیجه این که $I(A^n; R^n)$ نزدیک به $I(\hat{A}^n; R^n)$ است.

[۴] Fannes-Audenaert Inequality

۴ اثبات نامساوی فینز

اثبات نامساوی فینز در دو بخش انجام می‌شود. ابتدا آن را در حالت کلاسیک با استفاده از فانو اثبات می‌کنیم. سپس حالت کوانتمی آن را با استفاده از حالت کلاسیک آن اثبات می‌کنیم.

۱.۴ اثبات نامساوی فینز با استفاده از نامساوی فانو در حالت کلاسیک

برای اثبات نیاز به تکنیک جفت کردن[۵] داریم که ابتدا آن را بیان می‌کنیم.

فاصله اثر و تکنیک جفت کردن: مفهوم فاصله اثر در حالت کلاسیک همان فاصله مجموع میان دو توزیع است. فرض کنید که $p(x)$ و $q(x)$ دو توزیع روی یک مجموعه $\mathcal{X}$ باشند. در این صورت

$$\frac{1}{2}\|p-q\|_1 = \frac{1}{2}\sum_x |p(x)-q(x)|.$$

حال فرض کنید که متغیر تصادفی X_1 دارای توزیع $p(x)$ و متغیر تصادفی X_2 دارای توزیع $q(x)$ باشند:

$$p(X_1=x)=p(x), \quad p(X_2=x)=q(x).$$

تا اینجا توزیع حاشیه‌ای X_1 و X_2 را تعریف کردیم، و هنوز صحبتی از توزیع مشترک X_1 و X_2 نکرده‌ایم. این توزیع مشترک با توجه به اطلاعات داده شده از توزیع‌های حاشیه‌ای به صورت یکتا مشخص نمی‌شود. فرض کنید از میان تمامی توزیع‌های مشترک ممکن روی X_1X_2 با توزیع‌های حاشیه‌ای داده شده، آن توزیع مشترکی را انتخاب کنیم که احتمال $p(X_1 \neq X_2)$ را کمینه کند. به این کار جفت کردن[۶] دو متغیر تصادفی گفته می‌شود و کاربردهای زیادی دارد. واضح است که اگر $p(x)=q(x), \forall x$ آن وقت می‌توان $X_1=X_2$ قرار داد و احتمال $p(X_1 \neq X_2)$ را به صفر رساند. در حالت کلی‌تر این مقدار کمینه نصف فاصله مجموع دو توزیع می‌شود.

تمرین ۶ ثابت کنید

$$\min p(X_1 \neq X_2) = \frac{1}{2}\|p-q\|_1,$$

که در آن مینیمم روی همه توزیع‌های X_1X_2 با توزیع‌های حاشیه‌ای p و q گرفته می‌شود.

اثبات نامساوی فینز در حالت کلاسیک: حال اگر فاصله اثر $\epsilon = \frac{1}{2}\|p(x)-q(x)\|_1$ را داشته باشیم می‌توانیم دو متغیر X_1 و X_2 بسازیم که توزیع حاشیه‌ای آنها همان $p(x)$ و $q(x)$ باشند و به علاوه $p(X_1 \neq X_2)=\epsilon$. طبق نامساوی فانو برای حالات کلاسیک باید داشته باشیم

$$H(X_1|X_2) \leq \epsilon \log(d-1) + h(\epsilon),$$

که در آن $d=|\mathcal{X}|$. اما $H(X_1)-H(X_2) = H(X_1|X_2) - H(X_2|X_1) \leq H(X_1|X_2)$ پس

$$|H(X_1)-H(X_2)| \leq \epsilon \log(d-1) + h(\epsilon),$$

و اثبات کامل است.

[۵] Coupling
[۶] Coupling

۲.۴ اثبات نامساوی فینز با استفاده از نامساوی فانو

اگر $p_1 \geq p_2 \geq p_3 \cdots \geq p_d$ مقادیر ویژه ρ و $q_1 \geq q_2 \geq q_3 \cdots \geq q_d$ مقادیر ویژه σ باشند، نشان خواهیم داد که

$$\|\rho - \sigma\|_1 \geq \sum_i |p_i - q_i|. \tag{۳}$$

در این صورت چون $(p_1, p_2, p_3, \cdots, p_d)$ و $(q_1, q_2, q_3, \cdots, q_d)$ دو توزیع احتمال هستند و

$$H(\sigma) = H(\{q_1, q_2, q_3, \cdots, q_d\}),$$

$$H(\rho) = H(\{p_1, p_2, p_3, \cdots, p_d\}),$$

با استفاده از اثبات قضیه در حالت کلاسیک می‌توانیم اثبات در حالت کوانتمی را کامل کنیم. پس کافی است (۳) را ثابت کنیم.

بیاد آورید که برای ماتریس‌های چگالی ρ و σ، عملگرهای مثبت Q و R با فضای پشتیبان عمود بر هم وجود دارند به طوری که $\rho - \sigma = Q - R$ و $\|\rho - \sigma\|_1 = 2\text{tr}(Q) = 2\text{tr}(R)$. تعریف کنید $T = \rho + R = \sigma + Q$ پس

$$\text{tr}(T) = 1 + \text{tr}(R) = 1 + \text{tr}(Q) = 1 + \frac{1}{2}\|\rho - \sigma\|_1$$

پس بجای اینکه ثابت کنیم

$$\|\rho - \sigma\|_1 \geq \sum_i |p_i - q_i|.$$

کافی است ثابت کنیم

$$\begin{aligned}\text{tr}(T) &\geq 1 + \frac{1}{2}\sum_i |p_i - q_i| \\ &= \sum_i \frac{1}{2}(p_i + q_i + |p_i - q_i|) \\ &= \sum_i \max(p_i, q_i).\end{aligned}$$

نتیجه این که اگر مقادیر ویژه T را با $t_1 \geq t_2 \geq \cdots \geq t_d$ نشان دهیم کافی است که ثابت کنیم

$$t_i \geq \max(p_i, q_i).$$

با توجه به اینکه $T = \rho + R = \sigma + Q$ این نامساوی از دو لم زیر بدست می‌آید.

لم ۷ فرض کنید A یک ماتریس هرمیتی باشد و مقادیر ویژه‌ی آن را با $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d$ نمایش دهید. در این صورت برای هر k داریم

$$\lambda_k = \max_{W \in Gr(k)} \min_{|v\rangle \in W : \||v\rangle\| = 1} \langle v|A|v\rangle,$$

که در آن منظور از $Gr(k)$ مجموعه‌ی زیرفضاهای با بعد k است.

اثبات: این لم برای $k=1$ معادل است با

$$\lambda_1 = \max_{|v\rangle:\||v\rangle\|=1} \langle v|A|v\rangle,$$

و برای $k=d$ خواهیم داشت

$$\lambda_d = \min_{|v\rangle:\||v\rangle\|=1} \langle v|A|v\rangle,$$

که اثبات آنها را قبلا دیده‌ایم. این لم را برای هر k ثابت می‌کنیم. فرض کنید $|w_i\rangle$ بردار ویژه با طول واحد متناظر با λ_i باشد. اگر V_k زیرفضای k بعدی تشکیل شده توسط بردارهای $\{|w_1\rangle, |w_2\rangle, \ldots, |w_k\rangle\}$ بگیریم آنگاه

$$\min_{|v\rangle\in V_k:\||v\rangle\|=1} \langle v|A|v\rangle = \lambda_k.$$

پس کافی است نشان دهیم برای هر زیرفضای k بعدی دلخواه W وجود دارد $|v\rangle \in W$ با طول واحد به طوری که $\langle v|A|v\rangle \leq \lambda_k$.

از آنجا که W زیرفضایی k بعدی است و زیرفضای $span\{|v_k\rangle, |v_{k+1}\rangle, \ldots, |v_d\rangle\}$ بعد $d-k+1$ دارد، این دو زیرفضا اشتراکی نابدیهی دارند. یعنی $|v\rangle = \alpha_k|v_k\rangle + \cdots + \alpha_d|v_d\rangle \in W$ با طول واحد وجود دارد. حال داریم

$$\langle v|A|v\rangle = \sum_{i=k}^{d} \lambda_i|\alpha_i|^2 \leq \lambda_k \sum_{i=k}^{d} |\alpha_i|^2 = \lambda_k\||v\rangle\|^2 = \lambda_k.$$

اثبات تمام است. $\square$

لم ۸ فرض کنید A, B دو ماتریس هرمیتی باشند و $A \geq B$. مقادیر ویژه‌ی A را با $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d$ و مقادیر ویژه‌ی B را با $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_d$ نشان دهید. در این صورت برای هر k داریم $\lambda_k \geq \mu_k$.

اثبات این لم با استفاده از لم قبل واضح است.

تمرین ۹ فرض کنید A یک ماتریس هرمیتی باشد و مقادیر ویژه‌ی آن را با $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d$ نمایش دهید. نشان دهید که برای هر k داریم

$$\lambda_k = \min_{W\in Gr(d-k+1)} \max_{|v\rangle\in W:\||v\rangle\|=1} \langle v|A|v\rangle.$$