مدرسین: ابوالفتح بیگی و امین‌زاده گوهری نویسنده: مرتضی نوشاد

جلسه ۲۶

۱ اثبات بخش مستقیم فشرده‌سازی شوماخر

در جلسه قبل اثبات بخش وارون قضیه فشرده‌سازی شوماخر اثبات شد. در این جلسه به اثبات بخش مستقیم قضیه فشرده‌سازی شوماخر پرداخته و سپس وارد بحث کدگذاری کانال خواهیم شد.

جهت یادآوری ابتدا صورت قضیه شوماخر را دوباره بیان می‌کنیم.

قضیه ۱ فرض کنید که ρ^A تابع چگالی منبع اطلاعات کوانتمی باشد. در این صورت $H(A)_\rho$ کوچکترین نرخ R قابل حصول برای فشرده‌سازی است.

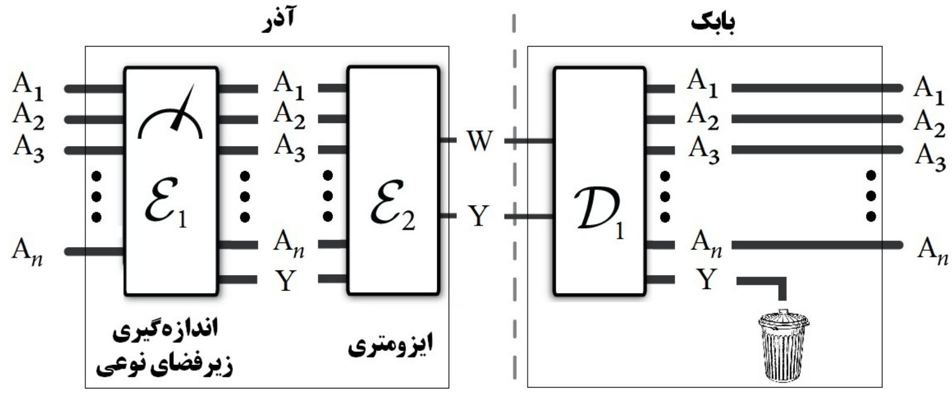
اثبات: نشان می‌دهیم که آنتروپی فون نیومن یک نرخ قابل حصول برای فشرده‌سازی اطلاعات کوانتمی است. پروتکلی که در اینجا ارایه می‌کنیم مشابه پروتکل شانون برای فشرده‌سازی اطلاعات کلاسیک است. اما این دو پروتکل به دلیل ویژگی‌های مکانیک کوانتمی فرق‌هایی نیز دارند. در حالت کلاسیک ابتدا تحقیق می‌کنیم که دنباله منبع (مثلا X^n) نوعی باشد. در صورتی که این دنباله نوعی باشد، شماره آن دنباله (عددی در بازه $\{1,2,3,\cdots,2^{n(H(X)+\epsilon)}\}$) را برای بابک ارسال می‌شود. پیدا کردن شماره مربوط به دنباله نوعی با استفاده از یک تابع $f:\mathcal{T}^n_{p_X,\delta}\to\{1,2,3,\cdots,2^{n(H(X)+\epsilon)}\}$ بدست می‌آید. در پروتکل کوانتمی این دو مرحله، به انجام اندازه‌گیری زیرفضای نوعی و بعد یک ایزومتری تبدیل می‌شوند. اعمال ایزومتری جای پیدا کردن شماره مربوط به دنباله نوعی (اعمال تابع f) را می‌گیرد. بخش کدگشایی عکس فشرده‌سازی است. شکل ۱ مراحل کلی این پروتکل فشرده‌سازی را نشان می‌دهد.

حال به بیان جزئیات این پروتکل می‌پردازیم. فرض کنیم که A^n در حالت $\rho_A^{\otimes n}$ در دست آذر باشد و

$$\rho_A=\sum_z p_Z(z)|z\rangle\langle z|,$$

تجزیه طیفی ρ_A باشد که در آن $p_Z(z)$ یک تابع توزیع احتمال و $\{|z\rangle\}$ یک پایه متعامد یکه است. آذر ابتدا اندازه‌گیری زیرفضای نوعی $\mathcal{E}_1^{A^n\to YA^n}$ متناظر با حالت ρ_A را روی A^n اعمال می‌کند که در آن حاصل اندازه‌گیری در سیستم Y ذخیره می‌شود. توجه کنید که برای ورودی دلخواه σ_{A^n} داریم

$$\begin{aligned}\mathcal{E}_1^{A^n\to YA^n}(\sigma_{A^n})&=p_0\bigg(|0\rangle\langle 0|\otimes\frac{1}{p_0}(I-\Pi^n_\delta)\sigma_{A^n}(I-\Pi^n_\delta)\bigg)+p_1\bigg(|1\rangle\langle 1|\otimes\frac{1}{p_1}\Pi^n_\delta\sigma_{A^n}\Pi^n_\delta\bigg)\\&=|0\rangle\langle 0|\otimes(I-\Pi^n_\delta)\sigma_{A^n}(I-\Pi^n_\delta)+|1\rangle\langle 1|\otimes\Pi^n_\delta\sigma_{A^n}\Pi^n_\delta\end{aligned}$$



شکل ۱: شمای کلی پروتکل بخش قابل حصول فشرده‌سازی شوماخر. در اینجا Y حاصل اندازه‌گیری نوعی است و سیستم‌های کوانتمی (W, Y) از سمت آذر به بابک ارسال می‌شوند. W سیستم متشکل از $n(H(\rho)+\epsilon)$ کیوبیت است.

که در آن p_0 و p_1 احتمال این هستند که حاصل اندازه‌گیری نوعی 0 یا 1 باشد:

$$p_0 = \mathrm{tr}((I - \Pi_\delta^n)\sigma_{A^n}(I - \Pi_\delta^n)) = \mathrm{tr}(\sigma_{A^n}(I - \Pi_\delta^n)(I - \Pi_\delta^n)) = \mathrm{tr}(\sigma_{A^n}(I - \Pi_\delta^n)),$$

$$p_1 = \mathrm{tr}(\sigma_{A^n}\Pi_\delta^n).$$

همچنین $\Pi_\delta^n = \Pi_{\rho,\delta}^n$ عملگر تصویر روی زیرفضای نوعی ρ_A است. در صورتی که حاصل اندازه‌گیری 1 باشد، سیستم A^n در حالتی قرار می‌گیرد که در واقع در یک زیرفضای با بعد $2^{n(H(\rho)+\epsilon)}$ است. لذا این سیستم را می‌توان با یک ایزومتری در یک فضای با بعد $2^{n(H(\rho)+\epsilon)}$ نشاند. یک پایه متعامد یکه برای این زیرفضا بردارهای به شکل $|z^n\rangle$ برای دنباله‌های نوعی z^n است. لذا برای این کار یک فضای هیلبرت $\mathcal{H}_W$ با بعد $2^{n(H(\rho)+\epsilon)}$ در نظر می‌گیریم؛ این فضا، فضای هیلبرت متناظر با $n(H(\rho)+\epsilon)$ کیوبیت است. یک پایه متعامد یکه دلخواه برای $\mathcal{H}_W$ را با $|1\rangle, |2\rangle, \cdots, |2^{nH(\rho)+\epsilon}\rangle$ نشان می‌دهیم. ایزومتری‌ای را در نظر می‌گیریم که پایه‌های متعامد یکه این دو فضا را به هم تصویر کند. برای این کار اگر دنباله‌های نوعی z^n را با اعداد شماره‌گذاری کرده باشیم و تابع $f : \mathcal{T}_{p_Z,\delta}^n \to \{1, 2, \cdots, 2^{nH(\rho)+\epsilon}\}$ شماره دنباله نوعی را به ما بدهد، ایزومتری مورد نظر بردار پایه $|z^n\rangle$ را به بردار پایه $|f(z^n)\rangle$ تصویر می‌کند. پس این ایزومتری به شکل زیر است:

$$U = \sum_{z^n \in \mathcal{T}_{p_Z,\delta}^n} |f(z^n)\rangle_W \langle z^n|_{A^n}.$$

پس قدم بعدی آذر، اعمال این ایزومتری برای فشرده‌سازی با فرض این که حاصل اندازه‌گیری (Y) یک باشد، است. اگر حاصل اندازه‌گیری صفر باشد، می‌توانیم سیستم W را در حالت ثابت دلخواهی بنام $|e\rangle_W$ متناظر با «خطا» قرار دهیم. فرآیند متناظر با این مرحله‌ی آذر را $\mathcal{E}_2^{YA^n\to YW}$ می‌نامیم. جهت نوشتن رابطه $\mathcal{E}_2^{YA^n\to YW}$ به عنوان یک فرایند کوانتمی لازم است که رفتار آن را برای تمامی حالات ورودی تعریف کنیم. فرض کنید که ورودی این فرایند حالتی کلاسیک- کوانتمی به شکل زیر باشد

$$\sigma_{YA^n} = p_0|0\rangle\langle 0|_Y \otimes \sigma_0 + p_1|1\rangle\langle 1|_Y \otimes \sigma_1.$$

برای ورودی‌های از این دست $\mathcal{E}_2^{YA^n\to YW}$ را به صورت زیر تعریف می‌کنیم:

$$\mathcal{E}_2^{YA^n\to YW}(\sigma_{YA^n}) = p_0|0\rangle\langle 0|_Y \otimes |e\rangle\langle e|_W + p_1|1\rangle\langle 1|_Y \otimes U\sigma_1 U^\dagger$$

توجه کنید که خروجی این عملگر هم متناظر با یک هنگرد است. در حالت کلی نیز که σ_{YA^n} کلاسیک-کوانتمی نیست نیز می‌توان $\mathcal{E}_2^{YA^n\to YW}$ را به این صورت تعریف کرد که تحت این ورودی ابتدا کیوبیت Y را در پایه‌ی متعامد یکه $\{|0\rangle, |1\rangle\}$ اندازه‌گیری می‌کنیم. اگر حاصل اندازه‌گیری 0 بود سیستم W را در خروجی در حالت $|e\rangle_W$ قرار می‌دهیم و اگر حاصل 1 بود ایزومتری U را روی A^n اعمال می‌کنیم. به عبارت دیگر برای ورودی دلخواه σ_{YA^n} تعریف می‌کنیم

$$\mathcal{E}_2^{YA^n\to YW}(\sigma_{YA^n}) = |0\rangle\langle 0|_Y \mathrm{tr}_{A^n}(\sigma_{YA^n})|0\rangle\langle 0|_Y \otimes |e\rangle\langle e|_W + (|1\rangle\langle 1|_Y\otimes U)\sigma_{YA^n}(|1\rangle\langle 1|_Y\otimes U^\dagger).$$

بنابراین می‌توان کدگذاری آذر را به شکل زیر در نظر گرفت:

$$\mathcal{E}^{A^n\to YW} := \mathcal{E}_2^{YA^n\to YW}\circ \mathcal{E}_1^{A^n\to YW}.$$

پس از اعمال کدگذاری آذر سیستم‌های Y, W را که شامل $n[H(\rho)+\epsilon]$ کیوبیت هستند برای بابک می‌فرستد.

اما تابع کدگشایی $\mathcal{D}^{YW\to A^n}$ بابک عکس ایزومتری آذر عمل خواهد کرد. وقتی که Y برابر یک است وارون ایزومتری U را اعمال می‌کند و وقتی Y برابر صفر است یک بردار دلخواه به نام $|e'\rangle_{A^n}$ در فضای A^n را در نظر گرفته و سیستم A^n را در آن حالت قرار می‌دهد. پس اگر حالت کلاسیک-کوانتمی

$$\tau_{YW} = p_0\ |0\rangle\langle 0|\otimes\tau_0 + p_1\ |1\rangle\langle 1|\otimes \tau_1$$

را به عنوان ورودی داشته باشیم، اثر فرایند کدگشایی روی آن به صورت زیر خواهد بود.

$$\mathcal{D}_1^{YW\to YA^n}(\tau_{YW}) = p_0\ |0\rangle\langle 0|_Y\otimes |e'\rangle\langle e'|_{A^n} + p_1\ |1\rangle\langle 1|_Y\otimes U^\dagger\tau_1 U.$$

مانند $\mathcal{E}_2^{YA^n\to YW}$ اگر ورودی فرایند کلاسیک کوانتمی نبود نیز باز هم می‌توان رابطه‌ای برای $\mathcal{D}^{YW\to YA^n}$ نوشت. حال با صرف نظر کردن از سیستم Y، فرایند کدکشایی بابک به طورت زیر است

$$\mathcal{D}^{YW\to A^n} = \mathrm{tr}_Y\circ\mathcal{D}_1^{YW\to YA^n}.$$

تحلیل خطای پروتکل:

فرض کنید که ρ_{AR} یک محض‌سازی از ρ_A باشد. داریم

$$\begin{aligned}
&\|\rho_{AR}^{\otimes n} - [(\mathcal{D}^{YW\to A^n}\circ\mathcal{E}^{A^n\to YW})\otimes\mathcal{I}_{R^n}](\rho_{AR}^{\otimes n})\|_1\\
&= \|\mathrm{tr}_Y\{|1\rangle\langle 1|_Y\otimes\rho_{AR}^{\otimes n}\} - [(\mathcal{D}^{YW\to A^n}\circ\mathcal{E}^{A^n\to YW})\otimes\mathcal{I}_{R^n}](\rho_{AR}^{\otimes n})\|_1 \qquad (۱)\\
&= \|\mathrm{tr}_Y\{|1\rangle\langle 1|_Y\otimes\rho_{AR}^{\otimes n}\} - \mathrm{tr}_Y\{[(\mathcal{D}_1^{YW\to A^nY}\circ\mathcal{E}^{A^n\to YW})\otimes\mathcal{I}_{R^n}](\rho_{AR}^{\otimes n})\}\|_1 \qquad (۲)\\
&\le \||1\rangle\langle 1|_Y\otimes\rho_{AR}^{\otimes n} - [(\mathcal{D}_1^{YW\to A^nY}\circ\mathcal{E}^{A^n\to YW})\otimes\mathcal{I}_{R^n}](\rho_{AR}^{\otimes n})\|_1 \qquad (۳)\\
&= \||1\rangle\langle 1|_Y\otimes\rho_{AR}^{\otimes n} - p_0\ |0\rangle\langle 0|_Y\otimes|e'\rangle\langle e'|_{A^n}\otimes\rho_R^{\otimes n}\\
&\qquad - p_1\ |1\rangle\langle 1|_Y\otimes\frac{1}{p_1}(\Pi_\delta^n\otimes I_{R^n})\rho_{AR}^{\otimes n}(\Pi_\delta^n\otimes I_{R^n})\|_1. \qquad (۴)
\end{aligned}$$

تساوی (۱) با اضافه کردن $|1\rangle_Y$ به $\rho_{RA}^{\otimes n}$ و با اثر جزئی گرفتن نتیجه می‌شود. تساوی (۲) با بیرون کشیدن اثر جزئی گرفتن نسبت به Y از کدگشا حاصل می‌شود؛ در اینجا ضرب تانسوری در $\mathcal{I}_{R^n}$ به داخل عملگر اثر جزئی گرفتن برده شده است. نامساوی (۳) در عبارات بالا، از خاصیت افزایش فاصله اثر با حذف زیر سیستم بدست می‌آید. تساوی (۴) با جایگذاری یا تحقیق مستقیم نتیجه می‌شود که در آن p_0 و p_1 احتمالات مشاهده 0 و 1 در اندازه‌گیری نوعی هستند:

$$p_0 = \mathrm{tr}(\rho_A^{\otimes n}(I - \Pi_\delta^n)) = \mathrm{tr}(\rho_{AR}^{\otimes n}((I - \Pi_\delta^n) \otimes I_{R^n})), \quad p_1 = \mathrm{tr}(\rho_A^{\otimes n}\Pi_\delta^n) = \mathrm{tr}(\rho_{AR}^{\otimes n}(\Pi_\delta^n \otimes I_{R^n})).$$

جهت تحقیق تساوی (۴) توجه کنید که با احتمال p_0 نهایتا سیستم Y, A^n, R در حالت $|0\rangle\langle 0|^Y \otimes |e'\rangle\langle e'|^{A^n} \otimes \rho_R^{\otimes n}$ قرار می‌گیرد و همین طور در مورد جمله دوم. رابطه نوشته شده هنگرد نهایی خروجی است.

درادامه زنجیره نامساوی‌ها می‌توان نوشت:

$$\begin{aligned} &\leq \| |1\rangle\langle 1|_Y \otimes \rho_{AR}^{\otimes n} - |1\rangle\langle 1|_Y \otimes (\Pi_\delta^n \otimes I_{R^n})\rho_{AR}^{\otimes n}(\Pi_\delta^n \otimes I_{R^n})\|_1 \\ &\quad + \|p_0\ |0\rangle\langle 0|_Y \otimes |e'\rangle\langle e'|_{A^n} \otimes \rho_R^{\otimes n}\|_1 \\ &= \|\rho_{AR}^{\otimes n} - (\Pi_\delta^n \otimes I_{R^n})\rho_{AR}^{\otimes n}(\Pi_\delta^n \otimes I_{R^n})\|_1 + p_0 \\ &\leq 2\sqrt{\epsilon} + \epsilon \end{aligned} \tag{۵}$$

اولین نامساوی از نامساوی مثلث برای فاصله اثر و ساده کردن p_1 از صورت و مخرج نتیجه شده است. تساوی (۵) از برابری

$$\|\rho \otimes \sigma - \omega \otimes \sigma\|_1 = \|\rho - \omega\|_1 \|\sigma\|_1 = \|\rho - \omega\|_1$$

و

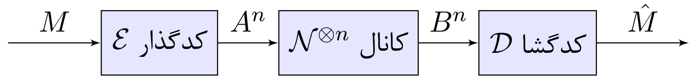
$$\|b\rho\|_1 = |b| \cdot \|\rho\|_1 = |b|$$

به ازای ماتریس چگالی های دلخواه ρ، σ و ω و ثابت b نتیجه می شود. نامساوی آخر نیز از ویژگی اندازه‌گیری زیرفضای نوعی $p_0 \leq \epsilon$ و لم اندازه‌گیری نرم بدست می‌آید. توجه کنید که اگر سیستم A^n, R^n را در نظر بگیریم و روی A^n یک اندازه‌گیری نوعی اعمال کنیم، با احتمال بالا نتیجه برابر 1 خواهد بود. پس میزان اعوجاج کل سیستم در اثر این اندازه‌گیری روی زیر سیستم کم خواهد بود. □

۲ کدگذاری کانال برای انتقال پیام کلاسیک

فرض کنیم که یک کانال کوانتمی $\mathcal{N}^{A\to B}$ داشته باشیم که در هر بار استفاده یک نسخه از سیستم A را دریافت و در خروجی یک نسخه از سیستم B را قرار می‌دهد. در حالت کلی ممکن است علاقه‌مند به انتقال اطلاعات کلاسیک یا کوانتمی باشیم، اما در اینجا تنها مساله انتقال اطلاعات کلاسیک از طریق این کانال را مورد بررسی قرار می‌دهیم.

کلی‌ترین پروتکلی که می‌توان برای این کار در نظر گرفت در شکل ۲ نشان داده شده است. در این پروتکل، آذر پیام کلاسیک m را بین پیام‌های $\{1, \ldots, |\mathcal{M}|\}$ انتخاب می‌کند و با توجه به آن حالت سیستم ورودی کانال A^n را مشخص می‌کند. به این حالت سیستم ورودی A^n کلمات کد کوانتمی گفته می‌شود و آن را با $\rho_m^{A^n}$ نشان می‌دهیم. آذر این حالت



شکل ۲: نمایش شماتیک یک کدگذار کانال کوانتمی

را با n بار استفاده مستقل از کانال $\mathcal{N}$ به بابک انتقال می‌دهد. بنابراین حالتی که بابک در اختیار می‌گیرد به صورت زیر خواهد بود.

$$\sigma_m^{B^n} = \mathcal{N}^{\otimes n}(\rho_m^{A^n})$$

سپس بابک برای یافتن پیام آذر یک اندازه‌گیری POVM با اعضای $\{\Lambda_m\}$ روی سیستم خروجی کانال انجام می‌دهد. در این صورت احتمال این که بابک پیام m را درست کدگشائی کند برابر است با

$$Pr\{\hat{M} = m | M = m\} = \text{tr}\{\Lambda_m \mathcal{N}^{\otimes n}(\rho_m^{A^n})\}.$$

همچنین احتمال رخداد خطا در کدگشائی پیام m نیز برابر است با:

$$\begin{aligned} P_e(m) &= 1 - Pr\{\hat{M} = m | M = m\} \\ &= \text{tr}\{(I - \Lambda_m)\mathcal{N}^{\otimes n}(\rho_m^{A^n})\}. \end{aligned}$$

یک نرخ قابل حصول

مساله ظرفیت انتقال پیام کلاسیک روی یک کانال کوانتمی در حالت کلی حل نشده است. اما یک کران پایین (نرخ قابل حصول) برای آن وجود دارد که به اطلاعات هولوو[1] معروف است. در حالتی که کانال دارای ورودی و خروجی کلاسیک باشد فرمول اطلاعات هولوو به فرمول ظرفیت کانال شانون تبدیل می‌شود. از این جهت اطلاعات هولوو را می‌توان تعمیمی از فرمول ظرفیت شانون دانست. اطلاعات هولو به صورت زیر تعریف می‌شود:

$$\chi(\mathcal{N}) := \max_{\rho_{XA}} I(X;B),$$

که در آن ماکزیمم‌گیری روی تمامی حالت‌های کلاسیک-کوانتمی ρ^{XA} به فرم

$$\rho^{XA} = \sum_x p(x) |x\rangle\langle x|^X \otimes \rho_x^A,$$

انجام می‌شود. توجه کنید که پس از انتخاب ρ^{XA} برای محاسبه اطلاعات متقابل $I(X;B)$ نیاز به حالت مشترک X, B داریم که به صورت زیر محاسبه می‌شود:

$$\rho^{XB} = \sum_x p(x) |x\rangle\langle x|^X \otimes \mathcal{N}^{A\to B}(\rho_x^A).$$

[1] Holevo Information

به عبارت دیگر داریم

$$\begin{aligned}\chi(\mathcal{N}) &= \max_{p(x),\rho_x} I(X;B) \\ &= \max_{p(x),\rho_x} H(B) - H(B|X) \\ &= \max_{p(x),\rho_x} H\left(\sum_x p(x)\mathcal{N}(\rho_x)\right) - \sum_x p(x) H(\mathcal{N}(\rho_x)).\end{aligned}$$

تمرین ۲ نشان دهید که در ماکزیمم‌گیری فوق می‌توان فرض کرد که ρ_x-ها محض هستند.

نکته ۳ اگر ورودی کانال یعنی A کلاسیک باشد، آنگاه سیستم A حتی بعد از اعمال کانال $\mathcal{N}$ نیز وجود دارد و لذا عبارت $I(A;B)$ با معنی است. حال با استفاده از نامساوی پردازش داده داریم $I(X;B) \leq I(A;B)$. نتیجه این که اگر ورودی کانال کلاسیک باشد انتخاب $X = A$ ممکن خواهد بود و بدلیل نامساوی فوق بهترین انتخاب است. به عبارت دیگر اگر کانال به صورتی باشد که حالت کلاسیک a را به σ_a ببرد آنگاه داریم

$$\chi(\mathcal{N}) = \max_{p(a)} H\left(\sum_a p(a)\sigma_a\right) - \sum_a p(a) H(\sigma_a).$$

به عنوان تمرین این نکته را با استفاده از تمرین ۲ نیز ثابت کنید.

قضیه ۴ اطلاعات هولوو یک نرخ قابل حصول برای ظرفیت کلاسیک یک کانال کوانتمی است.

کلیات اثبات

اثبات این قضیه شبیه اثبات قسمت قابل حصول قضیه شانون می‌باشد که مبتنی بر تولید کتاب کد تصادفی است. فرض کنید که

$$\rho^{XA} = \sum_x p(x)\, |x\rangle\langle x|^X \otimes \rho_x^A,$$

داده شده باشد. در ابتدا به ازای هر پیام، یک دنباله n بیتی x^n به صورت i.i.d. از توزیع حاشیه‌ای $p(x)$ تولید می‌شود. این کار دقیقا مشابه کاری است که در حالت کلاسیک برای تولید کتاب کد انجام می‌دادیم. در نتیجه جدول دنباله‌های x^n را می‌توان به شکل زیر تشکیل داد که در آن سطرهای این جدول دنباله‌های x^n هستند. در این جدول فرض شده که X دودویی است، $\mathcal{X} = \{0,1\}$. به همین دلیل خانه‌های جدول با 0 و 1 پر شده‌اند.

	n	$\cdots$	3	2	1
1	1	$\cdots\ \cdots$	0	1	1
2	0	$\cdots\ \cdots$	0	0	1
3	1	$\cdots\ \cdots$	0	1	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
2^{nR}	0	$\cdots\ \cdots$	0	1	0

حال برای انتقال پیام $m \in \{1,2,3,\cdots,2^{nR}\}$ نیاز به یک کدگذار داریم. وظیفه کدگذار تبدیل پیام m به حالتی از ورودی کانال یعنی A^n است. برای این کار به سطر m-ام جدول بالا نگاه کرده و دنباله

$$x_m^n = (x_{m1}, x_{m2}, \cdots, x_{mn})$$

را پیدا می‌کنیم. سپس ورودی کانال را در حالت

$$\rho_{x_{m1}}^{A_1} \otimes \rho_{x_{m2}}^{A_2} \otimes \cdots \otimes \rho_{x_{mn}}^{A_n},$$

قرار می‌دهیم.

در گیرنده بابک سیستم B^n را در حالت

$$\mathcal{N}(\rho_{x_{m1}}^{A_1}) \otimes \mathcal{N}(\rho_{x_{m2}}^{A_2}) \otimes \cdots \otimes \mathcal{N}(\rho_{x_{mn}}^{A_n})$$

دریافت می‌کند. برای سادگی نمادگذاری فرض کنید $\sigma_x^B = \mathcal{N}(\rho_x^A)$. در نتیجه با این نمادگذاری سیستم B^n در حالت

$$\sigma_{x_m^n}^{B^n} := \bigotimes_{i=1}^{n} \sigma_{x_{mi}}^{B_i},$$

قرار خواهد داشت.

حال هدف گیرنده اعمال اندازه‌گیری مناسب بر روی سیستم B^n است تا بتواند m را پیدا کند. در حالت کلاسیک کدگشایی را این گونه انجام می‌دادیم که ابتدا بررسی می‌کردیم که آیا دنباله خروجی نوعی هست یا نه؟ اگر نوعی نبود خطا اعلام می‌کردیم. مشابه این کار را در اینجا نیز می‌توانیم انجام دهیم: با فرض اینکه دنباله x_m^n با توجه به $p(x)$ نوعی باشد، می‌دانیم که اگر $\sigma_{x_m^n}^{B^n}$ را با توجه به اندازه‌گیری نوعی مربوط به حالت متوسط سیستم B اندازه بگیریم با احتمال زیاد جواب 1 خواهد بود. به عبارت دیگر

$$\text{tr}(\Pi_{\sigma,\delta}^n \sigma_{x_m^n}^{B^n}) \geq 1 - \epsilon,$$

که در آن

$$\sigma = \sum_x p(x)\sigma_x.$$

پس این مرحله شبیه حالت کلاسیک است. سپس در حالت کلاسیک دنباله خروجی را با تک تک کلمات کد مقایسه می‌کردیم تا بتوانیم دنباله مشترکا نوعی را پیدا کنیم. اما در اینجا امکان این کار وجود ندارد زیرا هر اندازه‌گیری‌ای که بر روی سیستم B^n انجام بدهیم، این سیستم را تغییر می‌دهد. تعداد کلمات کد به صورت نمایی زیاد است و اگر مقایسه با هر کدام بخواهد کمی حالت B^n را تغییر دهد، پس از مدتی حالت B^n تغییرات عمده‌ای خواهد کرد.

پیش از آنکه به حل این مشکل بپردازیم برای یک لحظه فرض کنید که در گیرنده m درست را می‌دانیم. با توجه به خواص اندازه‌گیری نوعی مشخص است که اگر با استفاده از اندازه گیری نوعی شرطی $\Pi^{B^n|x_m^n}$ حالت $\sigma_{x_m^n}^{B^n}$ را اندازه بگیریم جواب با احتمال زیاد 1 خواهد بود:

$$\text{tr}(\Pi_\delta^{B^n|x_m^n} \sigma_{x_m^n}^{B^n}) \geq 1 - \epsilon.$$

در اینجا ثابت خواهیم کرد که اگر $m' \neq m$ را درگیرنده انتخاب کنیم و سپس اندازه‌گیری نوعی متناظر با $x^n_{m'}$ را روی $\sigma^{B^n}_{x^n_m}$ انجام دهیم حاصل با احتمال خیلی کمی (تقریبا بطور متوسط $2^{-nI(X;B)}$) برابر 1 خواهد بود که با شهود ما از دنیای کلاسیک سازگار است. برای این به دلیلی که در ادامه مشخص می‌شود بجای عملگر $\Pi_\delta^{B^n|x^n_{m'}}$ عملگر $\Pi^{B^n}_{\sigma,\delta}\Pi_\delta^{B^n|x^n_{m'}}\Pi^{B^n}_{\sigma,\delta}$ را در نظر می‌گیریم. پس هدف اثبات این است که

$$\mathbb{E}_{X^n_m,X^n_{m'}}\left[\mathrm{tr}(\Pi^{B^n}_{\sigma,\delta}\Pi_\delta^{B^n|x^n_{m'}}\Pi^{B^n}_{\sigma,\delta}\sigma^{B^n}_{x^n_m})\right],$$

نزدیک به صفر است.

توجه کنید که

$$\mathbb{E}_{X^n_m,X^n_{m'}}\left[\mathrm{tr}(\Pi^{B^n}_{\sigma}\Pi^{B^n|X^n_{m'}}\Pi^{B^n}_{\sigma}\sigma^{B^n}_{X^n_m})\right] = \mathbb{E}_{X^n_{m'}}\left[\mathrm{tr}(\Pi^{B^n}_{\sigma}\Pi^{B^n|X^n_{m'}}\Pi^{B^n}_{\sigma}\sum_{x^n_m}p(x^n_m)\sigma^{B^n}_{x^n_m})\right].$$

اما می‌دانیم

$$\sum_{x^n_m}p(x^n_m)\sigma^{B^n}_{x^n_m} = (\sum_x p(x)\sigma_x)^{\otimes n} = \sigma^{\otimes n}.$$

در نتیجه

$$\mathbb{E}_{X^n_m,X^n_{m'}}\left[\mathrm{tr}(\Pi^{B^n}_{\sigma}\Pi^{B^n|X^n_{m'}}\Pi^{B^n}_{\sigma}\sigma^{B^n}_{X^n_m})\right] = \mathbb{E}_{X^n_{m'}}\left[\mathrm{tr}(\Pi^{B^n}_{\sigma}\Pi^{B^n|X^n_{m'}}\Pi^{B^n}_{\sigma}\sigma^{\otimes n})\right]$$

امید ریاضی باقیمانده روی دنباله‌های $X^n_{m'}$ است. ثابت می‌کنیم که برای هر «دنباله نوعی» $x^n_{m'}$ داریم:

$$\mathrm{tr}(\Pi^{B^n}_{\sigma}\Pi^{B^n|x^n_{m'}}\Pi^{B^n}_{\sigma}\sigma^{\otimes n}) = \mathrm{tr}(\Pi^{B^n|x^n_{m'}}\Pi^{B^n}_{\sigma}\sigma^{\otimes n}\Pi^{B^n}_{\sigma}) \leq 2^{-n(I(X;B)-\epsilon)}$$

از قضایای زیرفضاهای نوعی می‌دانیم که

$$\Pi^{B^n}_{\sigma}\sigma^{\otimes n}\Pi^{B^n}_{\sigma} \leq 2^{-n(H(B)-\epsilon)}\Pi^{B^n}_{\sigma}.$$

پس داریم[۲]

$$\mathrm{tr}(\Pi^{B^n|x^n_{m'}}\Pi^{B^n}_{\sigma}\sigma^{\otimes n}\Pi^{B^n}_{\sigma}) \leq 2^{-n(H(B)-\epsilon)}\mathrm{tr}(\Pi^{B^n|x^n_{m'}}\Pi^{B^n}_{\sigma}).$$

از طرف دیگر با استفاده از $\Pi^{B^n}_{\sigma} \leq I$ داریم

$$\mathrm{tr}(\Pi^{B^n|x^n_{m'}}\Pi^{B^n}_{\sigma}) \leq \mathrm{tr}(\Pi^{B^n|x^n_{m'}}) \leq 2^{n(H(B|X)+\epsilon)}.$$

نتیجه این که برای هر دنباله نوعی $x^n_{m'}$ داریم:

$$\mathrm{tr}(\Pi^{B^n}_{\sigma}\Pi^{B^n|x^n_{m'}}\Pi^{B^n}_{\sigma}\sigma^{\otimes n}) \leq 2^{-n(H(B)-\epsilon)}2^{n(H(B|X)+\epsilon)} = 2^{-n(I(X;B)-2\epsilon)}$$

[۲] در اینجا از این رابطه استفاده کردیم که اگر $A \leq B$ و $C \geq 0$ آنوقت $\mathrm{tr}(AC) \leq \mathrm{tr}(BC)$ که برقرار است زیرا $\mathrm{tr}((B-A)C) \geq 0$. اثر ضرب دو ماتریس نامنفی، نامنفی است.

توجه کنید در صورتی که دنباله $x^n_{m'}$ نوعی نباشد، رابطه بالا برقرار نیست. جهت حل این مشکل بعدا راه حلی ارائه خواهیم کرد.

ادغام چند اندازه گیری در یک اندازه گیری: دیدیم که جهت بررسی اینکه پیام مثلا $m = 1$ هست یا نه، روش مشخصی وجود دارد. کافی است که اندازه‌گیری نوعی مشروط به $X^n = x^n_1$ را اعمال کنیم. اما این کار حالت سیستم B^n را کمی تغییر می‌دهد و اگر این تحقیق را برای $m = 2$، $m = 3$، و الی آخر نیز انجام دهیم، پس از مدتی حالت B^n ممکن است به شکل قابل ملاحظه‌ای تغییر کند. پس می‌خواهیم بجای تعداد زیادی اندازه‌گیری تنها یک اندازه‌گیری داشته باشیم.

اولین حدس می‌تواند انتخاب عملگرهای $E_m = \Pi^{B^n}_\sigma \Pi^{B^n|x^n_{m'}} \Pi^{B^n}_\sigma$ باشد برای اندازه‌گیری POVM ما باشد. طبق شهود ما این عملگرها «تقریبا» بر هم عمود هستند و نه دقیقا و متاسفانه شرط تمامیت را ارضا نمی‌کنند. پس باید این انتخاب را تصحیح کرد.

برای دو عملگر مثبت معین S, T توجه کنید که

$$E_1 = (S+T)^{-\frac{1}{2}} S (S+T)^{-\frac{1}{2}}, \qquad E_2 = (S+T)^{-\frac{1}{2}} T (S+T)^{-\frac{1}{2}}$$

شرط تمامیت را برقرار می‌کنند:

$$E_1 + E_2 = (S+T)^{-\frac{1}{2}} (S+T) (S+T)^{-\frac{1}{2}} = I.$$

به این روش تصحیح اندازه‌گیری تقریبا خوب[۳] می‌گویند.

حال احتمال خطای این اندازه‌گیری تصحیح شده را با استفاده از لم زیر بدست می‌آوریم:

لم ۵ نامساوی اپراتور *Hayashi-Nagaoka*

فرض کنید که $0 \le S \le I$ و $0 \le T$ را داشته باشیم. در این صورت نامساوی زیر همواره برقرار است:

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

دقت کنید که این نامساوی نتیجه می‌دهد که

$$I - E_1 = E_2 \le 2(I - S) + 4T$$

پس

$$\mathrm{tr}(E_2\sigma) \le 2\mathrm{tr}((I - S)\sigma) + 4\mathrm{tr}(T\sigma).$$

در جلسه بعد از این لم استفاده خواهیم کرد و اثبات را به طور دقیق کامل می‌کنیم.

[۳] Pretty Good Measurement

اثبات: اثبات لم ۵: به ازای هر دو ماتریس A و B داریم:

$$(A-B)^{\dagger}(A-B) \geq 0,$$

و در نتیجه

$$A^{\dagger}B + B^{\dagger}A \leq A^{\dagger}A + B^{\dagger}B.$$

با قرار دادن $A = \sqrt{T}$ و $B = \sqrt{T}((S+T)^{-1/2} - I)$ بدست می آید،

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

حال با استفاده از عملگر یکنوا بودن تابع $f(x) = \sqrt{x}$ و شرط $0 \leq S \leq I$ داریم $\sqrt{S+T} \geq \sqrt{S} \geq S$. در نتیجه

$$\begin{aligned}
I - (S+T)^{-1/2}S(S+T)^{-1/2} &= (S+T)^{-1/2}T(S+T)^{-1/2} \\
&= T + T((S+T)^{-1/2} - I) + ((S+T)^{-1/2} - I)T \\
&\quad + ((S+T)^{-1/2} - I)T((S+T)^{-1/2} - I) \\
&\leq 2T + 2((S+T)^{-1/2} - I)T((S+T)^{-1/2} - I) \\
&\leq 2T + 2((S+T)^{-1/2} - I)(S+T)((S+T)^{-1/2} - I) \\
&= 2T + 2(I + S + T - 2\sqrt{S+T}) \\
&\leq 2T + 2(I + S + T - 2S) \\
&= 2(I-S) + 4T.
\end{aligned}$$

$\square$