نظریه اطلاعات کوانتمی ۱، ترم پاییز ۱۳۹۲-۱۳۹۱
مدرسین: ابوالفتح بیگی و امین زاده گوهری

جلسه ۲۲-۲۰

۱ مقدمه

تئوری اطلاعات کلاسیک به بررسی انتقال اطلاعات از یک نقطه به نقطه دیگر و از طریق کانال‌های کلاسیک می‌پردازد. مساله اصلی تئوری اطلاعات کوانتمی همانند مساله اصلی تئوری اطلاعات کلاسیک انتقال پیام از یک نقطه به نقطه دیگر است، با این تفاوت که نحوه انتقال و حتی ماهیت پیام تحت تاثیر آثار کوانتمی است. در نتیجه، نظریه کلاسیک نیازمند بسط و گسترش می‌باشد به گونه‌ای که در حالت خاصی که تمامی سیستم‌ها کلاسیک باشند، نتایج به همان نتایج تئوری اطلاعات کلاسیک منجر شوند. قطعا قدم اول در این راه تعمیم مفهوم آنتروپی به دنیای کوانتمی است که به آن آنتروپی فون نویمان می‌گوییم. ممکن است عجیب به نظر برسد اما آنتروپی فون نویمان (که تعمیمی از آنتروپی شانون است) قبل از آنتروپی شانون معرفی شده است.[۱] البته آنتروپی فون نویمان در مباحث مربوط به ترمودینامیک و جهت اندازه‌گیری بی‌نظمی مطرح شده بود، و در نگاه فون نویمان معنای اطلاعاتی نداشته است. علیرغم مقدم بودن تعریف آنتروپی فون نویمان بر آنتروپی شانون تا قبل از دهه نود میلادی محققین نظریه اطلاعات عموما توجه چندانی به نظریه کوانتمی اطلاعات نداشته‌اند، و تنها دو دهه است که بررسی نظام‌مند نظریه اطلاعات کوانتمی شروع شده است. اگر چه بسیاری از نتایج دنیای کوانتمی شبیه نتایج کلاسیک هستند، اما مواردی از تفاوت‌های مورد توجه نیز به چشم می‌خورد.

نظریه اطلاعات کوانتمی مفاهیمی مانند منبع ، کانال ، کد را تعمیم داده و مفهوم جدید درهم‌تنیدگی را لحاظ می‌کند. در این بخش ابتدا مقدمه‌ای کلی بر این مفاهیم خواهیم داشت. در بخش بعد به مرور مفاهیمی از نظریه اطلاعات کلاسیک می‌پردازیم و سپس وارد تئوری اطلاعات کوانتمی می‌شویم. خلاصه مطالب گفته شده در این بخش در جدول ۲.۱ آمده است.

ابتدا از مفهوم درهم‌تنیدگی شروع می‌کنیم. درهم‌تنیدگی مفهومی کاملا کوانتمی است که معادل کلاسیک ندارد. شاید نزدیک‌ترین تشبیه به مفهوم درهم‌تنیدگی بیت‌های تصادفی به اشتراک گذاشته شده باشد. اما از آزمایش بل دیدیم که درهم‌تنیدگی اکیدا غنی‌تر از بیت‌های تصادفی به اشتراک گذاشته شده است. همچنین کاربرد درهم‌تنیدگی را در پروتکل‌های کدگذاری فوق چگال[۲] و فرابرد[۳] دیدیم. براساس قوانین مکانیک کوانتمی اگر یکی از دو سیستم درهم‌تنیده را اندازه‌گیری کنیم، سیستم دیگر بصورت آنی و لحظه‌ای (سریع‌تر از سرعت نور) تغییر حالت می‌دهد. اما همان طور که

[۱] زمانی که شانون فرمول $\sum_i p_i \log \frac{1}{p_i}$ را کشف کرد، در استفاده از کلمه آنتروپی برای این مفهوم به دلیل بحث‌های فلسفی پیرامون این کلمه در تردید بود. اما فون نیومان شانون را به استفاده از این کلمه ترغیب کرده و به او گفت: "هیچ کس به هر حال نمی‌داند که آنتروپی چیست. پس در هر مناظره‌ای تو یک برتری داری".

[۲] Superdense coding

[۳] Teleportation

شکل ۱: (راست) جان فون نویمان ($1903 - 1957$) مبدع آنتروپی کوانتمی. (چپ) کلود شانون ($1916 - 2001$) مبدع نظریه اطلاعات کلاسیک

نشان داده شد این سقوط حالت باعث انتقال اطلاعات نمی‌شود و همچنان محدودیت سرعت نور برای انتقال اطلاعات برقرار است. یکی از سؤالاتی که در نظریه اطلاعات کوانتمی راجع به آن صحبت می‌شود مفید بودن درهم‌تنیدگی در پروتکل‌های انتقال اطلاعات است. آیا همانطور که درهم‌تنیدگی در بازی CHSH کمک می‌کرد، در انتقال اطلاعات میان دو نفر نیز می‌تواند مفید باشد؟

مفهوم بعدی مفهوم کانال است. در جلسات قبل دیدیم که کانال‌های کلاسیک (متشکل از یک توزیع شرطی $p(y|x)$) جای خود را به کانال‌های کوانتمی می‌دهند. کانال کوانتمی نگاشتی خطی است که یک ماتریس چگالی (ماتریسی مثبت نیمه معین با اثر یک) را به ماتریس‌های چگالی می‌برد (و به آن نگاشت خطی کاملا مثبت و حافظ اثر[۴] نیز گفته می‌شود). کانال میان فرستنده و گیرنده ممکن است کانالی با ورودی کلاسیک و خروجی کلاسیک (اصطلاحا کانال cc یا کلاسیک-کلاسیک)، کانالی با ورودی کلاسیک و خروجی کوانتمی(اصطلاحا کانال cq)، کانالی با ورودی کوانتمی و خروجی کلاسیک (اصطلاحا کانال qc)، و یا کانالی با ورودی کوانتمی و خروجی کوانتمی (اصطلاحا کانال qq) باشد. یک فیبر نوری که انتقال دهنده‌ی فوتون‌هاست می‌تواند مثالی از یک کانال qq باشد.

کانال‌های cc همان کانال‌های کلاسیک هستند که با یک توزیع شرطی $p(y|x)$ قابل نمایش هستند. فرض کنید که ورودی X یک متغیر تصادفی با توزیع $p(x)$ باشد، در نمادگذاری کوانتمی متغیر تصادفی X را می‌توان به عنوان یک هنگرد در فضای هیلبرتی $|\mathcal{X}|$ بعدی در نظر گرفت که حالت $|x\rangle$ را با احتمال $p(X = x)$ اخذ می‌کند. حالات $|x\rangle$ برای x-های مختلف بر هم عمود می‌باشند. بنابراین حالت متغیر تصادفی X با توجه به هنگرد مربوطه عبارت است از

$$\sum_x p(x)|x\rangle\langle x|.$$

بیان حالت مشترک X, Y در دنیای کوانتمی را به دو طریق می‌توان انجام داد: یکی با استفاده از رفتار کانال و دیگری بطور مستقیم. بطور مستقیم می‌بینیم که حالت (x, y) (یا $|x\rangle \otimes |y\rangle$ در شکل کوانتمی آن) با احتمال $p(x, y)$ اتفاق

[۴] Completely Positive Trace-Preserving Map

می‌افتد. پس ماتریس چگالی هنگرد مربوطه از ضرب احتمالات در ماتریس چگالی متناظر با هر حالت به شکل زیر بدست می‌آید:

$$\rho_{XY} = \sum_x p(x,y)(|x\rangle \otimes |y\rangle)(\langle x| \otimes \langle y|) = \sum_x p(x,y)|x,y\rangle\langle x,y| \qquad (۱)$$

اما راه دیگری برای نوشتن عبارت بالا وجود دارد: زمانی که $X = x$ اتفاق بیافتد حالت Y هنگردی خواهد بود که با احتمال $p(y|x)$ بردار $|y\rangle$ را می‌گیرد. در نتیجه ماتریس چگالی هنگرد مربوط به آن $\sum_y p(y|x)|y\rangle\langle y|$ می‌باشد. چون این حالت به مقدار $X = x$ مربوط است می‌توانیم آن را ρ_x بنامیم.

$$\rho_x = \sum_y p(y|x)|y\rangle\langle y|.$$

حال می‌توانیم مساله را این گونه در نظر بگیریم: یک هنگرد داریم که با احتمال $p(x)$ کل سیستم در حالت $|x\rangle\langle x| \otimes \rho_x$ قرار می‌گیرد (یعنی X در حالت x و Y در حالت ρ_x). پس (بنابر تعریف) هنگرد معادل برای کل سیستم با ضرب احتمالات در ماتریس چگالی‌های مربوطه بدست می‌آید:

$$\rho_{XY} = \sum_x p(x)\big(|x\rangle\langle x| \otimes \rho_x\big).$$

با جایگذاری می‌بینیم که این عبارت مساوی عبارت قبلی (۱) از ρ_{XY} است.

ورودی یک کانال cq متغیر تصادفی کلاسیکی مانند X است و خروجی آن یک سیستم A در حالت کوانتمی‌ای است که به مقدار متغیر تصادفی X بستگی دارد. این وابستگی را می‌توانیم با ρ_x نشان دهیم. این نمادگذاری به این معنی است که اگر ورودی کانال $X = x$ باشد، خروجی کانال سیستمی در حالت ρ_x خواهد بود. در صورتی که ورودی X یک متغیر تصادفی با توزیع $p(x)$ باشد، یک هنگرد خواهیم داشت که با احتمال $p(x)$ کل سیستم را در حالت $|x\rangle\langle x| \otimes \rho_x$ قرار می‌دهد. پس حالت مشترک ورودی و خروجی را می‌توان به شکل زیر نوشت:

$$\rho_{AX} = \sum_x p(x)\big(|x\rangle\langle x| \otimes \rho_x\big).$$

برای محاسبه حالت سیستم A می‌توان از دو روش استفاده کرد: یکی روش مستقیم و دیگری با گرفتن اثر جزئی از عبارت بالا. روش مستقیم این است که سیستم A با احتمال $p(x)$ در حالت ρ_x است پس حالت آن با ضرب این احتمالات در حالات مربوطه بدست می آید:

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

اما روش دیگر با گرفتن اثر جزئی است:

$$\begin{aligned}\rho_A &= \mathrm{tr}_X[\rho_{AX}] \\ &= \mathrm{tr}_X\left[\sum_x p(x)\big(|x\rangle\langle x| \otimes \rho_x\big)\right] \\ &= \sum_x p(x)\mathrm{tr}_X\big(|x\rangle\langle x| \otimes \rho_x\big). \\ &= \sum_x p(x)\rho_x.\end{aligned}$$

مفهوم بعدی مفهوم منبع است. در دنیای کلاسیک معمولا منابع را با متغیرهای تصادفی مدل‌سازی می‌کنیم. یک نمونه ساده از آن می‌تواند یک متغیر تصادفی برنولی باشد که دو مقدار صفر و یک - اصطلاحا یک بیت - را می‌گیرد. مفهوم بیت جای خود را به کیوبیت در دنیای کوانتمی می‌دهد که حالات ممکن آن نه تنها صفر و یک، بلکه شامل برهم‌نهی آنها نیز می‌باشد. در حالت کلی یک منبع کوانتمی یک سیستم است که در اختیار ما قرار داده می‌شود. مثلا ممکن است طبیعت با احتمال p_i سیستم را در حالت $|\psi_i\rangle$ تولید کرده باشد؛ یعنی هنگردی از حالات محض $\{p_i, |\psi_i\rangle\}$ را تولید و نمونه‌ای از آن را در اختیار ما قرار داده باشد. حال اطلاعات مورد علاقه ممکن است تشخیص این باشد که سیستم دقیقا در کدام حالت $|\psi_i\rangle$ قرار دارد. در صورتی که بردارهای $|\psi_i\rangle$ بر هم عمود باشند، با اندازه‌گیری در یک پایه‌ی مناسب می‌توانیم این موضوع را تشخیص دهیم. اما در حالت کلی این تشخیص ممکن نیست زیرا دسترسی به این اطلاعات ذخیره شده توسط قوانین مکانیک کوانتمی محدود می‌شود. بر خلاف منابع کلاسیک اندازه‌گیری یک سیستم کوانتمی (در حالت کلی) نتیجه‌ای نامعین و احتمالی در بر داشته و به علاوه باعث تغییر حالت آن می‌شود. این تغییر حالت امکان مطالعه مجدد سیستم اولیه را از ما می‌گیرد. به علاوه یک منبع کوانتمی قابل کپی برداری نیست تا چندین نسخه از آن را ذخیره کرده و به صورت جدا نسخه‌ها را مورد مطالعه قرار دهیم. با توجه به این محدودیت‌ها خواهیم دید که هر کیوبیت حاوی حداکثر یک بیت اطلاعات است. اگر این محدودیت در اندازه‌گیری وجود نداشت می‌توانستیم در هر کیوبیت بینهایت بیت ذخیره کنیم. برای مشخص کردن حالت یک کیوبیت در نمایش بلوخ آن باید دو پارامتر حقیقی را مشخص کنیم که هرکدام بسط اعشاری بینهایت بیتی دارند و در نتیجه می‌توانستیم بینهایت بیت را در حالت یک کیوبیت ذخیره کنیم.

حافظه‌های کلاسیک و حافظه‌های کوانتمی ساختارهای متفاوتی دارند. حافظه های کلاسیک معمولا در سیستمی متشکل از تعدادی بسیار زیادی اتم هستند. معمولا از ولتاژ ۵ ولت برای نمایش یک و از ولتاژ ۰ ولت برای نمایش صفر استفاده می‌شود. جهت حفظ این ولتاژها و مقاومت در برابر تغییر آنها بدلایل محیطی، به طور منظم این ولتاژها دوباره تنظیم می‌شوند. از طرف دیگر واحدهای حافظه‌های کوانتمی از یک اتم یا یک فوتون تشکیل می‌شوند. جهت حفظ این واحدهای حافظه در برابر تغییر آنها بدلایل محیطی باید برهم‌کنش آنها با محیط اطرافشان را کنترل کرد و برای مقابله با از دست رفتن اطلاعات در صورت تغییر از روش‌های بازیابی از طریق کدگذاری استفاده کرد.

دیدیم که حالت یک سیستم کوانتمی را می‌توان با یک هنگرد و یا یک ماتریس چگالی نمایش داد. یک منبع .i.i.d در دنیای کلاسیک معادل با دنباله $X_1, X_2, \cdots, X_n$ است به طوری که

$$p(x_1, x_2, \cdots, x_n) = \prod_{i=1}^{n} p(x_i)$$

که در آن توصیف $p(x_i)$-ها یکسان است. یک منبع i.i.d. در دنیای کوانتمی مجموعه‌ای از n سیستم $A_1, A_2, \cdots, A_n$ در حالت ضرب تانسوری $\rho_{A_1,A_2,\cdots,A_n} = \rho_{A_1} \otimes \rho_{A_2} \otimes \cdots \otimes \rho_{A_n}$ است که توصیف ρ_i-ها یکسان است. گاهی ضرب تانسوری بالا را با $(\rho^{\otimes n})_{A_1,A_2,\cdot,A_n}$ یا به صورت خلاصه‌تر با $\rho^{\otimes n}$ نمایش می‌دهند.[۵]

مشابه حالت کلاسیک که دنباله متغیرهای تصادفی $(X_1, X_2, \cdots, X_n)$ را با $X^{1:n}$ یا X^n نمایش می‌دادیم، دنباله سیستمهای $A_1, A_2, \cdots, A_n$ را می‌توان با $A^{1:n}$ یا A^n نمایش داد.

ارتباط میان منابع کوانتمی و منابع کلاسیک زمانی آشکار می‌شود که حالت سیستم را در یک پایه متعامد یکه بنویسیم. فرض کنید که نمایش قطری ρ به صورت

$$\rho = \sum_{j=1}^{d} p_j |v_j\rangle\langle v_j|$$

باشد، که در آن بردارهای $|v_i\rangle$ یک پایه متعامد یکه تشکیل می‌دهند. در این صورت پس از بسط دادن عبارت ضرب تانسوری خواهیم داشت:

$$\rho^{\otimes n} = \sum_{j_1,j_2,\cdots,j_n} p_{j_1} p_{j_2} \cdots p_{j_n} |v_{j_1}, v_{j_2}, \cdots, v_{j_n}\rangle\langle v_{j_1}, v_{j_2}, \cdots, v_{j_n}|.$$

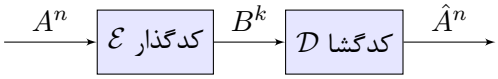
در بالا جمع روی دنباله‌های دلخواه $j_{1:n} = (j_1, j_2, \cdots, j_n)$ به طوری که اندیس‌های j_k متعلق به $\{1, 2, \cdots, d\}$ هستند گرفته می‌شود. توجه کنید که $p_{j_1} p_{j_2} \cdots p_{j_n}$ چیزی جز احتمال وقوع دنباله $j_{1:n}$ به عنوان یک منبع کلاسیکی که اندیس j را با احتمال p_j تولید می‌کند نیست. همچنین بردارهای $\{|v_{j_1}, v_{j_2}, \cdots, v_{j_n}\rangle\}$ یک پایه متعامد یکه برای فضای تانسوری تشکیل می‌دهند زیرا از ضرب تانسوری بردارهای پایه بدست آمده‌اند.

مفهوم بعدی مفهوم کد است که تعریف آن معادل تعریف یک سناریوی مخابراتی است. سناریوهای مخابراتی مختلف بر حسب نوع منبع، کانال و منابع به اشتراک گذاشته شده میان فرستنده و گیرنده تفاوت می‌کنند. منبع ما ممکن است کلاسیک و یا کوانتمی باشد؛ یعنی ممکن است بخواهیم یک رشته بیت‌های کلاسیک یا یک سری کیوبیت را از یک نقطه به نقطه دیگر منتقل کنیم. کانال میان فرستنده و گیرنده می‌تواند cc, cq, qc یا qq باشد. همچنین ممکن است میان فرستنده و گیرنده درهم‌تنیدگی (مثلا جفت EPR) تسهیم (به اشتراک گذاشته) شده باشد. همچنین زمانی که هدف انتقال کیوبیت است ممکن است میان فرستنده و گیرنده یک کانال مخابراتی کلاسیک با ظرفیت بینهایت به صورت رایگان فراهم شده باشد. به وضوح در این حالت میان فرستنده و گیرنده درهم‌تنیدگی به اشتراک گذاشته نشده است چون در این صورت با استفاده از پروتکل فرابرد می‌توان کیوبیت‌ها را از طریق جابجایی اطلاعات کلاسیک منتقل کرد. به حداکثر نرخ ارسال مطمئن کیوبیت‌ها در این حالت ظرفیت کوانتمی یاری شده[۶] می‌گویند.

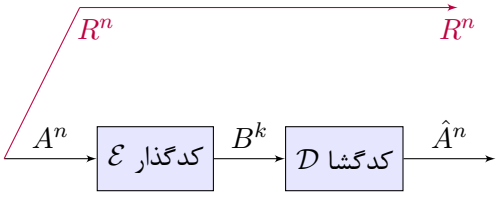
در مساله انتقال اطلاعات روی یک کانال کلاسیک (کدگذاری کانال) در نظریه اطلاعات کلاسیک مفاهیم کدگذار و کدگشا به شکل زیر تعریف می‌شوند. منظور از یک کدگذار یک تابع از مجموعه پیام‌ها به مجموعه سمبل‌های ورودی یک

[۵] در حالت کلاسیک زمانی که می‌نویسیم $p(x)$ منظور تابع وزن احتمال p است که روی مقادیر مختلف $x \in \mathcal{X}$ تعریف شده است. همچنین می‌توان در یک لحظه دو تابع وزن احتمال مختلف (مثلا $p(x)$ و $q(x)$) روی یک مجموعه $\mathcal{X}$ داشت. مشابها در حالت کوانتمی زمانی که صحبت از ρ_A می‌کنیم منظور یک ماتریس چگالی خاص است که روی سیستم A وجود دارد. σ_A یک ماتریس چگالی متفاوت خواهد بود که می‌تواند روی همان سیستم، اما مثلا در لحظه دیگری از زمان، وجود داشته باشد. حال فرض کنید که یک ماتریس چگالی خاص مانند ρ را می‌گیریم، بدون اینکه مشخص کنیم که روی چه سیستمی تعریف شده است. در این صورت $\rho^{\otimes n}$ یک ماتریس چگالی روی n سیستم است. پس $(\rho^{\otimes n})_{A_1,A_2,\cdot,A_n}$ بیانگر این است که این ماتریس چگالی روی ترکیب n سیستم مشخص $A_1, A_2, \cdots, A_n$ تعریف می‌شود.

[۶] Assisted Quantum Capacity



شکل ۲: نمایش شماتیک یک کدگذار منبع کوانتمی



شکل ۳: نمایش شماتیک یک کدگذار منبع کوانتمی به همراه محض کننده منبع

کانال است، و منظور از کدگشا یک تابع از خروجی کانال به سمبل‌های ورودی. یک کدگذار و یک کدگشای کوانتمی چیزی جز فرایندهای کوانتمی نیستند که شکل کلی آن در جلسات قبل بیان شد.

اگر یک کدگذار کلاسیک مانند $f : \{1, 2, \cdots, 2^k\} \to \{1, 2, \cdots, 2^n\}$ داشته باشیم و بخواهیم معادل آن را در فرمول‌بندی کوانتمی بنویسیم، می‌توانیم ورودی را فضای هیلبرتی با بعد 2^k گرفته و یک پایه متعامد یکه برای آن را $|1\rangle, |2\rangle, \cdots, |2^k\rangle$ گرفته و خروجی را فضای هیبلرتی با بعد 2^n گرفته و یک پایه متعامد یکه آن را $|1\rangle, |2\rangle, \cdots, |2^n\rangle$ در نظر بگیریم. فضای هیلبرت ورودی و خروجی این کدگذار توسط فرایند فیزیکی زیر به هم مربوط می‌شوند:

$$\mathcal{E}(|i\rangle\langle i|) = |f(i)\rangle\langle f(i)|, \quad \forall i \in \{1, 2, \cdots, 2^k\}$$

که می‌توان آن را به شکل زیر هم نوشت:

$$\mathcal{E}(\rho) = \sum_{i=1}^{2^k} M_i \rho M_i^\dagger, \quad M_i = |f(i)\rangle\langle i|.$$

درستی رابطه فوق را به عنوان تمرین به خواننده واگذار می‌کنیم. دقت کنید که

$$\sum_{i=1}^{2^k} M_i^\dagger M_i = \sum_{i=1}^{2^k} |i\rangle\langle f(i)| \cdot |f(i)\rangle\langle i| = \sum_{i=1}^{2^k} |i\rangle\langle i| = I.$$

در صورتی که $f(i)$ ها متمایز باشند، تبدیل بالا یک ایزومتری (و در نتیجه وارون پذیر) است چون پایه‌های یک فضا را به پایه‌های متمایز یک فضای دیگر متناظر می‌کند.

فضای هیلبرت با بعد 2^n با فضای هیلبرت متشکل از n کیوبیت یکریخت است. در نتیجه تبدیل بالا را می‌توان فرایند فیزیکی که k کیوبیت را دریافت و n کیوبیت را در خروجی قرار می‌دهد هم قلمداد کنیم.

۱.۱ کدگذاری منبع

یک کدگذار منبع کوانتمی را می‌توان به شکل 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

از این منبع است. فرض کنید که n سیستم $A_1, A_2, \cdots, A_n$ داریم که حالت مشترک آنها $\rho_{A_1,A_2,A_3,\cdots,A_n}$ (یا به صورت خلاصه ρ_{A^n}) برابر با $\rho^{\otimes n}$ است. همان طور که در شکل ۲ 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 $B_1, B_2, \cdots, B_k$ که حالت مشترک آنها را با $\sigma_{B^k} = \mathcal{E}(\rho^{\otimes n})$ نشان داده‌ایم را 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 ($\mu_{\hat{A}^n} = \mathcal{D}(\sigma_{B^k})$). توجه کنید که حالت مشترک σ_{B^k} و حالت مشترک $\mu_{\hat{A}^n}$ لزوما به شکل ضرب تانسوری نیستند همان طور که در حالت کلاسیک پیام فشرده شده و بازیابی شده لزوما i.i.d. نیستند.

چند تفاوت مهم میان حالات کلاسیک و کوانتمی وجود دارد: در کدگذاری کوانتمی دنباله منبع A^n از کدگذار عبور داده می‌شود و نزد فرستنده باقی نمی‌ماند در حالی که در دنیای کلاسیک فرستنده می‌تواند یک نسخه از منبع را برای خود نگاه دارد.[۷] نکته دیگر این که در نمایش شماتیک شکل ۲، سیستم‌های کوانتمی A^n، B^k و $\hat{A}^n$ در یک لحظه مشترک زمانی وجود فیزیکی ندارند، بلکه نابود شدن یکی منجر به تولید شدن دیگر می‌شود.

به هر کد کلاسیک دو پارامتر احتمال خطا و نرخ کد نسبت داده می‌شود. تعریف نرخ کد کوانتمی شبیه تعریف کلاسیک است: $R = \frac{k}{n}$ و واحد آن کیوبیت بر نمونه (سیستم)، و یا کیوبیت بر واحد زمان است (با فرض اینکه منبع یک سیستم A_i در واحد زمان تولید کند).

در دنیای کلاسیک، زمانی که منبع X^n را ارسال و بصورت $\hat{X}^n$ بازیابی کنیم، میزان خطای یک کد با احتمال $p(X^n \neq \hat{X}^n)$ توصیف میشود. اما همانطور که در بالا گفتیم در دنیای کوانتمی سیستم های A^n و $\hat{A}^n$ بصورت همزمان وجود ندارند. در بخش بعد تعمیم مناسب مفهوم خطای یک کد به دنیای کوانتمی را مورد بحث قرار خواهیم داد، اما جهت مقدمه سازی فرض کنید که فرستنده یک کپی از پیامش را پیش از ارسال بسازد $R^n = X^n$ که دارای توزیع مشترک

$$p_{R^n,X^n}(r^n, x^n) = p_{X^n}(x^n)\mathbf{1}[r^n = x^n] = p_{R^n}(r^n)\mathbf{1}[r^n = x^n].$$

میباشد. حال فرض کنید ثابت میکنیم که فاصله مجموع [۸] دو توزیع میان R^n, X^n و $R^n, \hat{X}^n$ همان احتمال خطای کد

[۷] به خاطر قضیه‌ی no-cloning این کار در دنیای کوانتمی امکان‌پذیر نیست.

[۸] Total variation distance

است:

$$\begin{aligned}
\|p_{R^n,X^n} - p_{R^n,\hat{X}^n}\|_1 &= \sum_{x^n,r^n} \left|p_{R^n,X^n}(r^n,x^n) - p_{R^n,\hat{X}^n}(r^n,x^n)\right| \\
&= \sum_{x^n,r^n} \left|p_{R^n}(r^n)\mathbf{1}[r^n = x^n] - p_{R^n}(r^n)p_{\hat{X}^n|R^n}(x^n|r^n)\right| \\
&= \sum_{r^n} p_{R^n}(r^n) \sum_{x^n} \left|\mathbf{1}[r^n = x^n] - p_{\hat{X}^n|R^n}(x^n|r^n)\right| \\
&= \sum_{r^n} p_{R^n}(r^n) P(\hat{X}^n \neq R^n | R^n = r^n) \\
&= P(\hat{X}^n \neq R^n) \\
&= P(\hat{X}^n \neq X^n)
\end{aligned}$$

مفهوم میزان خطای کد و اهمیت محض‌سازی: مفهوم احتمال خطای کد در یک کد کوانتمی جای خود را به نزدیک بودن حالات می‌دهد. منبع A^n ممکن است با محیط خارج درهم‌تنیده باشد. مثلا فرض کنید که A^n بخشی از کیوبیت‌هایی است که میان آلیس و باب در حالت EPR برای انجام پروتکل فرابرد در آینده به اشتراک گذاشته شده است. اما آلیس تصمیم گرفته که این کیوبیت‌ها را برای چارلی بفرستد تا چارلی بجای آلیس با باب پرتوکل فرابرد را انجام دهد. یک راه برای آلیس این است که کیوبیت‌هایش را درون جعبه گذاشته و آنها را برای چارلی بفرستد، اما در این صورت تعداد زیادی کیوبیت باید ارسال کند. سؤالی که وجود دارد این است که آیا با فشرده‌سازی و ارسال تعداد کمتری کیوبیت می‌توان انتقال A^n را انجام داد به طوری که طی فرایند فشرده‌سازی و بازیابی درهم‌تنیدگی آنها با محیط بیرونی حفظ شود؟ به عبارت دیگر برای هر سیستم دلخواه E که با A^n حالت مشترک ρ_{A^nE} را دارند و $\mathrm{tr}_E(\rho_{A^nE}) = (\rho^{\otimes n})_{A^n}$ باید داشته باشیم:

$$\left\|\rho_{A^n,E} - \big(\mathcal{D}\otimes\mathcal{I}_E\big)\circ\big(\mathcal{E}\otimes\mathcal{I}_E\big)\rho_{A^n,E}\right\|_1 \leq \epsilon,$$

که در آن از فاصله اثر [۹] جهت تعیین فاصله میان دو حالت کوانتمی استفاده کرده‌ایم. در اینجا ϵ کران بالایی روی خطای کد است. خطای کد ϵ (که بر حسب فاصله‌ی اثر نوشته شده) به این معنی است که هیچ آزمایشی (یا پروتکل کوانتمی‌ای) وجود ندارد که اختلاف میان دو حالت بالا را با احتمال درستی بیش از $\frac{1}{2} + \frac{1}{4}\epsilon$ حدس بزند؛ یا به طور معادل این تغییر حالت A^n به دلیل انجام فشرده‌سازی احتمال مشاهدات نتایج یک پروتکل دلخواه را بیش از ϵ تغییر دهد.

از آنجا که فرض می‌کنیم رابطه بالا برای هر محیط بیرونی E باید برقرار باشد، به طور معادل می‌توانیم بنویسیم:

$$\max_{E:\mathrm{tr}_E(\rho_{A^nE})=(\rho^{\otimes n})_{A^n}} \left\|\rho_{A^nE} - \big(\mathcal{D}\otimes I_E\big)\circ\big(\mathcal{E}\otimes I_E\big)\rho_{A^n,E}\right\|_1 \leq \epsilon. \qquad (۲)$$

ادعا می‌کنیم که جهت برقراری رابطه بالا برای سیستم دلخواه E کافی است حالت خاصی را در نظر بگیریم که E یک محض‌سازی از $\rho^{\otimes n}$ باشد. به عبارت دیگر برای برقراری (۲) کافی است که برای محض‌سازی‌های دلخواه R از $\rho^{\otimes n}$ داشته باشیم:

$$\left\|\rho_{A^nR} - \big(\mathcal{D}\otimes I_R\big)\circ\big(\mathcal{E}\otimes I_R\big)\rho_{A^nR}\right\|_1 \leq \epsilon.$$

[۹] Trace distance

برای اثبات سیستم دلخواه E و حالت ρ_{A^nE} که $\mathrm{tr}_E(\rho_{A^nE}) = \rho^{\otimes n}$ را در نظر بگیرید. فرض کنید که F یک محض سازی از ρ_{A^nE} باشد. در این صورت $R = EF$ یک محض‌سازی از $\rho^{\otimes n}$ است. پس طبق فرض داریم

$$\left\|\rho_{A^nEF} - \left(\mathcal{D}\otimes\mathcal{I}_E\otimes\mathcal{I}_F\right)\circ\left(\mathcal{E}\otimes\mathcal{I}_E\otimes\mathcal{I}_F\right)\rho_{A^nEF}\right\|_1 \leq \epsilon.$$

اگر قرار دهیم

$$\eta_{\hat{A}^nEF} = \left(\mathcal{D}\otimes\mathcal{I}_E\otimes\mathcal{I}_F\right)\circ\left(\mathcal{E}\otimes\mathcal{I}_E\otimes\mathcal{I}_F\right)\rho_{A^nEF},$$

آنگاه داریم:

$$\left\|\rho_{A^nEF} - \eta_{\hat{A}^nEF}\right\|_1 \leq \epsilon.$$

توجه کنید که دور انداختن بخشی از سیستم فاصله میان دو ماتریس چگالی را کم می‌کند. پس

$$\left\|\rho_{A^nE} - \eta_{\hat{A}^nE}\right\|_1 \leq \epsilon,$$

که همان چیزی است که به دنبال اثبات کردنش بودیم چون

$$\eta_{\hat{A}^n,E} = \left(\mathcal{D}\otimes\mathcal{I}_E\right)\circ\left(\mathcal{E}\otimes\mathcal{I}_E\right)\rho_{A^nE}.$$

این مثال تا حدی کاربرد و اهمیت محض‌سازی را در نظریه اطلاعات کوانتمی نشان می‌دهد.

پس برای اندازه‌گیری احتمال خطا کافی است محض‌سازی‌ها را در نظر بگیریم. حال برای ساده‌سازی بیشتر ادعا می‌کنیم که اگر فقط برای یک محض‌سازی رابطه

$$\left\|\rho_{A^nR} - \left(\mathcal{D}\otimes\mathcal{I}_R\right)\circ\left(\mathcal{E}\otimes\mathcal{I}_R\right)\rho_{A^nR}\right\|_1 \leq \epsilon$$

برقرار باشد، آنگاه برای هر محض‌سازی دیگری هم برقرار خواهد بود. دلیل این موضوع این است که تمام محض‌سازی‌ها توسط ایزومتری‌ها به هم مربوط می‌شوند، و ایزومتری‌ها فاصله‌ی اثر بین دو حالت را تغییر نمی‌دهند. به طور دقیق‌تر اگر $\rho_{A^nR'}$ محض‌سازی دیگر از $\rho^{\otimes n}$ باشد و V یک ایزومتری که $\rho_{A^nR'} = (I_{A^n}\otimes V)\rho_{A^nR}(I_{A^n}\otimes V^\dagger)$ آنگاه داریم

$$\begin{aligned}
&\left\|\rho_{A^nR'} - \left(\mathcal{D}\otimes\mathcal{I}_{R'}\right)\circ\left(\mathcal{E}\otimes\mathcal{I}_{R'}\right)\rho_{A^nR'}\right\| \\
&= \left\|(I_{A^n}\otimes V)\rho_{A^nR}(I_{A^n}\otimes V^\dagger) - \left(\mathcal{D}\otimes\mathcal{I}_{R'}\right)\circ\left(\mathcal{E}\otimes\mathcal{I}_{R'}\right)\left[(I_{A^n}\otimes V)\rho_{A^nR}(I_{A^n}\otimes V^\dagger)\right]\right\| \\
&= \left\|W\rho_{A^nR}W^\dagger - W\left(\mathcal{D}\otimes\mathcal{I}_R\right)\circ\left(\mathcal{E}\otimes\mathcal{I}_R\right)\rho_{A^nR}W^\dagger\right\| \\
&= \left\|\rho_{A^nR} - \left(\mathcal{D}\otimes\mathcal{I}_R\right)\circ\left(\mathcal{E}\otimes\mathcal{I}_R\right)\rho_{A^nR}\right\| \\
&\leq \epsilon,
\end{aligned}$$

که در اینجا قرار داده‌ایم $W = I\otimes V$ و همچنین از رابطه‌ی

$$\mathcal{N}_A\otimes\mathcal{I}_B\big((I_A\otimes T_B)\tau_{AB}(I_A\otimes T_B^\dagger)\big) = (I_A\otimes T_B)(\mathcal{N}_A\otimes\mathcal{I}_B(\tau_{AB}))(I_A\otimes T_B^\dagger)$$

استفاده کرده‌ایم.

نتیجه این که کافی است تنها یک محض‌سازی خاص را در نظر بگیریم. از آنجایی که $\rho_{A^n} = \rho^{\otimes n}$ حالت ضرب تانسوری دارد، می‌توانیم هرکدام از زیر سیستم‌های A_i را به صورت مستقل با یک R_i محض‌سازی کنیم. یعنی فرض کنید ρ_{AR} یک محض سازی از ρ_A باشد. در این صورت $\rho_{A_1R_1} \otimes \cdots \otimes \rho_{A_nR_n} = (\rho^{\otimes n})_{A^nR^n}$ یک محض‌سازی از $\rho_{A_1} \otimes \cdots \otimes \rho_{A_n} = (\rho^{\otimes n})_{A^n}$ است. در نظر گرفتن این محض‌سازی در نظریه‌ی اطلاعات کوانتمی معمول است. شکل ۳ نمایش شماتیک این مساله را نشان می‌دهد.

۲.۱ کدگذاری کانال

یک کانال کوانتمی $\mathcal{N}$ با سه مشخصه توصیف می‌شود: فضای هیلبرت سیستم ورودی، فضای هیلبرت سیستم خروجی و توصیف عملکرد کانال که مثلا با عملگرهای $E_1, E_2, \cdots, E_k$ به صورت

$$\mu = \mathcal{N}(\sigma) = \sum_i E_i \sigma E_i^\dagger, \qquad \sum_i E_i^\dagger E_i = I,$$

بیان می‌شود.

در حالت کلاسیک هدف ما انتقال پیام m که عضوی از مجموعه $\{1, 2, \ldots, 2^k\}$ است، با استفاده از کانال است. در دنیای کوانتمی کدگذاری کانال ممکن است کلاسیک یا کوانتمی باشد. یعنی هدف ممکن ارسال پیام کلاسیک یا کوانتمی روی کانال باشد. در اینجا به بررسی ارسال اطلاعات کوانتمی روی یک کانال اکتفا می‌کنیم. پس هدف ما انتقال قابل اطمینان یک سیستم کوانتمی M (با فضای هیلبرت 2^k بعدی) با استفاده از کانال است. مانند مساله‌ی کدگذاری منبع در اینجا هم فرض می‌کنیم که اگر سیستم M با محیط اطراف درهم‌تنیده باشد، این درهم‌تنیدگی پس از انتقال از بین نرود.[۱۰] برای در نظر گرفتن این درهم‌تنیدگی‌های مانند مساله‌ی کدگذاری منبع کافی است فقط یک محض‌سازی از M را در نظر بگیریم.

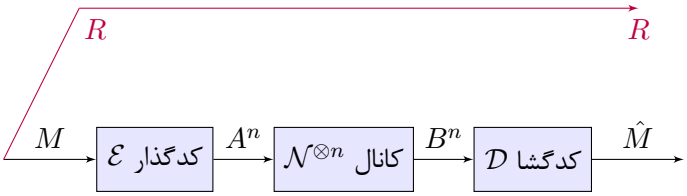
همان طور که در شکل ۴ نشان داده شده است کد کانال سیستم M را که با محض‌سازی R در حالت $|\psi\rangle_{MR}$ قرار گرفته، دریافت و سپس از یک کدگذار $\mathcal{E}^{M\to A^n}$ استفاده کرده تا n سیستم (احتمالا درهم‌تنیده) که فضای هیلبرتشان همان فضای هیلبرت ورودی کانال است را تولید کند ($\sigma_{A^nR} = \mathcal{E} \otimes \mathcal{I}_R(|\psi\rangle\langle\psi|_{MR})$). این سیستم‌های کوانتمی از n نسخه کانال عبور کرده تا n سیستم B^n با حالت

$$\mu_{B^nR} = (\mathcal{N} \otimes \mathcal{N} \otimes \cdots \otimes \mathcal{N} \otimes \mathcal{I}_R)\sigma_{A^nR} = \mathcal{N}^{\otimes n} \otimes \mathcal{I}_R(\sigma_{A^nR})$$

در خروجی کانال بوجود آید. سپس یک کدگشا $\mathcal{D}^{B^n\to\hat{M}}$ در خروجی سیستم $\hat{M}$ را تولید می‌کند. کدگشا یک فرایند کوانتمی است که هدفش بازیابی M است. بنابراین فاصله میان حالت سیستم بوجود آمده در گیرنده با حالت سیستم ارسالی باید محاسبه شود (با در نظر گرفتن محض‌سازی). به عبارت دیگر:

$$\big\||\psi\rangle\langle\psi|_{MR} - \big(\mathcal{D} \otimes \mathcal{I}_R\big) \circ \big(\mathcal{N}^{\otimes n} \otimes \mathcal{I}_R\big) \circ \big(\mathcal{E} \otimes \mathcal{I}_R\big)|\psi\rangle\langle\psi|_{MR}\big\|_1 \leq \epsilon, \quad \forall |\psi\rangle_{MR}.$$

[۱۰] توجه کنید که این فرض به نوعی در دنیای کلاسیک نیز وجود دارد. در دنیای کلاسیک نیز انتظار داریم که وابستگی پیغام با محیط اطرافش پس از رد شدن از کانال از بین نرود. نکته در این است که به طور معمول در دنیای کلاسیک ارسال پیغام با خطای کم خود به خود حفظ این وابستگی را تضمین می‌کند. ولی در دنیای کوانتمی باید وابستگی پیغام و محیط را در بررسی احتمال خطا در نظر بگیریم.



شکل ۴: نمایش شماتیک یک کدگذار کانال کوانتمی

توجه کنید که رابطه بالا باید برای هر بردار حالت ورودی برقرار باشد. با توجه به تعریف «نرم لوزی»[۱۱] این رابطه را می‌توان به طور خلاصه به صورت زیر نوشت:

$$\left\|\mathcal{I}_M - \mathcal{D} \circ \mathcal{N}^{\otimes n} \circ \mathcal{E}\right\|_\diamond \leq \epsilon.$$

با استفاده از بعد فضای هیلبرت M که 2^k است، نرخ کد R را به شکل زیر تعریف می‌کنیم:

$$R = \frac{k}{n}.$$

در کدگذاری کانال هر چه نرخ کد بیشتر باشد کد بهتری داریم. اما در کدگذاری منبع هرچه نرخ کد کمتر باشد، کد بهتر است.

۲ مروری بر نظریه اطلاعات کلاسیک

برای این بخش نوشتار جدایی در نظر گرفته شده است که در وبگاه درس آمده است.

۳ آنتروپی کوانتمی

در این بخش آنتروپی کوانتمی فون نیومان را تعریف می‌کنیم. همان گونه که تئوری اطلاعات کلاسیک مبتنی بر آنتروپی شانون است، تئوری اطلاعات کوانتمی مبتنی بر آنتروپی فون نیومان است و در نتیجه یادگیری تعریف و خواص آن از اهمیت بالایی برخوردار است. آنتروپی فون نیومان به هر ماتریس چگالی دلخواه یک عدد حقیقی نامنفی نسبت می‌دهد. تابع آنتروپی فون نیومان به شکل زیر تعریف می‌شود:

$$H(\rho) = -\text{tr}(\rho \log(\rho))$$

پیش از توجیه این فرمول، ابتدا نحوه محاسبه آن را خواهیم دید. فرض کنید که

$$\rho = \sum_i \lambda_i |v_i\rangle\langle v_i|,$$

قطری‌سازی ρ در یک پایه‌ی متعامد یکه باشد و λ_i ها مقادیر ویژه ρ باشند. بیاد آورید که چون ρ مثبت نیمه معین است و اثر آن یک است داریم:

$$\lambda_i \geq 0, \quad \sum_i \lambda_i = 1.$$

[۱۱] Diamond distance

خاصیت	کلاسیک	کوانتمی
بیان حالت	رشته‌ای از دنباله‌ها $x \in \{0,1\}^n$	رشته‌ای از کیوبیت‌ها $\psi = \sum_{x\in\{0,1\}^n} c_x \lvert x\rangle$
منبع	رشته‌ای از بیت‌ها	رشته‌ای از بیت‌ها رشته‌ای از کیوبیت‌ها
	متغیر تصادفی تابع توزیع	سیستم کوانتمی ماتریس چگالی
کانال	ورودی کلاسیک-خروجی کلاسیک	ورودی کلاسیک-خروجی کلاسیک (cc) ورودی کوانتمی-خروجی کلاسیک (qc) ورودی کلاسیک-خروجی کوانتمی (cq) ورودی کوانتمی-خروجی کوانتمی (qq)
کدگذار/کدگشا	یک تابع	یک فرایند کوانتمی
مخابرات	انتقال بیت‌های کلاسیک	انتقال بیت‌های کلاسیک، انتقال کیوبیت‌ها، تسهیم حالت درهم‌تنیده (جفت EPR)
آنتروپی منبع	آنتروپی شانون: $H(X) = -\sum_x p(x)\log p(x)$	آنتروپی فون نیومان: $H(\rho) = -\mathrm{tr}(\rho\log(\rho))$.
ظرفیت کانال نویزی	ظرفیت کلاسیک یک کانال	ظرفیت کلاسیک، ظرفیت کوانتمی بدون تسهیم درهم تنیدگی، ظرفیت کوانتمی با تسهیم درهم‌تنیدگی، ظرفیت کوانتمی یاری شده (توسط انتقال اطلاعات کلاسیک رایگان)

جدول ۱: مقایسه مفاهیم اصلی تئوری اطلاعات کوانتمی و کلاسیک

حال مشاهده کنید که $\rho\log(\rho)$ تابعی از ρ است و طبق بحثی که در بخش جبرخطی کردیم بردار ویژه‌های آن با بردار ویژه‌های ρ یکسان هستند و مقادیر ویژه‌ی آن برابرند با $\lambda_i \log \lambda_i$:

$$\rho\log(\rho) = \sum_i \lambda_i \log(\lambda_i)|v_i\rangle\langle v_i|.$$

در نتیجه

$$H(\rho) = -\text{tr}(\rho\log(\rho)) = -\sum_i \lambda_i \log(\lambda_i)$$

که همان آنتروپی کلاسیک دنباله $\{\lambda_i\}$، دنباله‌ی مقادیر ویژه ρ است. توجه کنید که $0\log 0$ را برابر 0 قلمداد می‌کنیم. توجه کنید که دنباله‌ی مقادیر ویژه یک توزیع احتمال تشکیل می‌دهند. پس آنتروپی این دنباله قابل تعریف است.

مثال ۱ فرض کنید $\rho = \frac{1}{d}I$ حالت یک سیستم d بعدی باشد. به این حالت کاملا مرکب[۱۲] می‌گویند. آنتروپی این ماتریس چگالی برابر است با $\log(d)$ چون این ماتریس d مقدار ویژه $\lambda_i = \frac{1}{d}$ دارد که متناظر با توزیع یکنواخت است.

مثال ۲ فرض کنید $\rho = |\psi\rangle\langle\psi|$ محض باشد. در این صورت ρ یک مقدار ویژه‌ی 1 دارد و بقیه‌ی مقادیر ویژه‌ی آن صفر هستند. پس داریم

$$H(|\psi\rangle\langle\psi|) = 0.$$

مثال ۳ فرض کنید

$$\rho = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

حالت یک سیستم 2 بعدی باشد. در این صورت مقادیر ویژه 0 و 1 هستند (در واقعا ρ محض است) و در نتیجه آنتروپی این حالت برابر صفر است.

مثال ۴ فرض کنید ماتریس چگالی ρ به شکل بلوکی

$$\rho = \begin{pmatrix} \rho_1 & 0 \\ 0 & 0 \end{pmatrix}$$

برای یک ماتریس چگالی ρ_1 باشد. در این صورت مقادیر ویژه ρ همان مقادیر ویژه ρ_1 به علاوه تعدادی صفر هستند. اما مقادیر ویژه صفر آنتروپی یک توزیع را عوض نمی‌کنند. پس

$$H(\rho) = H(\rho_1).$$

تمرین ۵ آنتروپی حالت

$$\rho = p\cdot|0\rangle + (1-p)\frac{(|0\rangle+|1\rangle)(\langle 0|+\langle 1|)}{2}$$

را یافته و آن را با آنتروپی کلاسیک دو-دویی $H(p, 1-p)$ مقایسه کنید.

[۱۲] Maximally mixed state

مثال ۶ فرض کنید که یک هنگرد از حالات محض $\{p_i, |\psi_i\rangle\}$ ا در اختیار داریم. در صورتی که بردارهای $|\psi_i\rangle$ بر هم عمود باشند، منبع کوانتمی همانند یک منبع کلاسیک خواهد بود و آنتروپی آن برابر خواهد بود با آنتروپی دنباله احتمالات $\{p_i\}$. زیرا در این صورت

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|,$$

که نتیجه می‌دهد که p_i-ها مقادیر ویژه و $|\psi_i\rangle$-ها بردارهای ویژه ρ هستند.

حالت جالب تر زمانی خواهد بود که حالات $|\psi_i\rangle$ بر هم عمود نباشند.

۱.۳ توجیه آنتروپی فون نیومان

حال به توجیه تعریف آنتروپی فون نیومان می‌پردازیم. اصولا در تعریف آنتروپی دو روش متفاوت وجود دارد. یک روش بر اساس اصل موضوع است. در این روش ما فرض می‌کنیم که فرمول مربوط به تابع آنتروپی کوانتمی را نمی‌دانیم. سپس انتظارات خود را از خواص یک تابع آنتروپی کوانتمی به عنوان اصل موضوع می‌نویسیم و نشان می‌دهیم که تابع داده شده تنها تابعی است که در خواص داده شده صدق می‌کند. روش دیگر معنی بخشی عملیاتی [۱۳] به آنتروپی است که آن را بر حسب یک فرایند واقعی که به اندازه‌گیری آن می‌پردازد تعریف می‌کند. مثلا ظرفیت کلاسیک یک کانال تعریفی عملیاتی دارد: ماکزیمم نرخی که به صورت قابل اطمینان می‌توان از یک کانال عبور داد. این تعریف بر حسب یک پروتکل واقعی که دارای نتایج ملموس است تعریف شده است، و نه بر حسب یک فرمول ریاضی که معنی و مفهوم آن مشخص نیست.

ما در این بخش ابتدا از روش اصل موضوعی استفاده می‌کنیم و سپس توجه خود را به معانی عملیاتی آنتروپی فون نیومان جلب می‌کنیم.

فرض کنید که فرمول مربوط به تابع آنتروپی کوانتمی را نمی‌دانیم. انتظارات خود را از خواص یک تابع آنتروپی کوانتمی را به عنوان اصل موضوع می‌نویسیم:

- **خاصیت سازگاری:** اگر ρ حالتی کلاسیک باشد، تابع آنتروپی کوانتمی همان آنتروپی شانون بشود.
- **خاصیت پایایی:** اعمال هر فرایند فیزیکی برگشت‌پذیر (وارون‌پذیر) تابع آنتروپی کوانتمی را تغییر ندهد. یا به عبارت دیگر

$$H(\rho) = H(\Phi(\rho))$$

برای هر فرایند کوانتمی برگشت‌پذیر $\Phi(\cdot)$ و هر حالت دلخواه ρ. منظور از یک فرایند برگشت‌پذیر Φ فرایندی است که در ازای آن فرایند دیگری بنام Φ^{-1} وجود داشته باشد به طوری که برای هر حالت دلخواه ρ

$$\Phi^{-1}(\Phi(\rho)) = \rho.$$

فرایندهای وارون‌پذیر در دنیای کلاسیک معادل توابع وارون‌پذیر هستند.

نکته: خاصیت پایایی از این نظر که اطلاعات ذخیره شده در یک سیستم تحت فرایندهای برگشت‌پذیر قابل بازیابی بوده و تغییر نمی‌کند منطقی است. در حالت کلاسیک هم برای هر تابع وارون پذیر f داریم $H(X) = H(f(X))$.

[۱۳] Operational meaning

قضیه ۷ تنها تابعی که دارای خواص سازگاری و پایایی است، تابع آنتروپی فون نویمان می‌باشد. یا به عبارت دیگر خواص سازگاری و پایایی تابع آنتروپی فون نیومان را بصورت یکتا مشخص می‌کنند.

اثبات: اثبات از دو قسمت تشکیل شده است. ابتدا نشان می‌دهیم که آنتروپی فون نویمان دارای خواص بالا است. فرض کنید که $|0\rangle, |1\rangle, \cdots, |d-1\rangle$ یک پایه متعامد یکه از حالات کلاسیک باشد. همچنین فرض کنید که سیستم با احتمال p_i در وضعیت $|i\rangle$ است. در این صورت ماتریس چگالی سیستم برابر خواهد بود با

$$\rho = \sum_{i=0}^{d-1} p_i |i\rangle\langle i|.$$

به وضوح مقادیر ویژه‌ی این ماتریس چگالی همان دنباله احتمالات p_i است. پس طبق بحثی که قبلا داشتیم آنتروپی فون نیومان آن همان آنتروپی دنباله مقادیر ویژه، یا آنتروپی احتمالات است. پس آنتروپی فون نیومان در این حالت همان آنتروپی شانون می‌شود.

جهت اثبات رابطه دوم توجه کنید که فرایندهای فیزیکی برگشت‌پذیر را می‌توان با یک ایزومتری نشان داد. پس فرض کنید

$$\Phi(\rho) = V\rho V^\dagger.$$

به طوری که

$$V^\dagger V = I.$$

از جبرخطی می‌دانیم که مقادیر ویژه AB و BA یکسان هستند، به استثنای احتمالا تعدادی مقدار ویژه صفر (یعنی مقادیر ویژه ناصفر به همراه تکررهای آنها در این دو ماتریس یکسان است). پس مقادیر ویژه ماتریس $V\rho V^\dagger$ و ماتریس $\rho V^\dagger V = \rho$ یکسان هستند، به استثنای احتمالا تعدادی صفر. اما زیاد کردن تعدادی صفر به مجموعه مقادیر ویژه آنتروپی دنباله مقادیر ویژه را تغییر نمی‌دهد. پس خاصیت پایایی برای آنتروپی فون نیومان برقرار است.

برعکس فرض کنید که تابع دلخواهی دارای خواص سازگاری و پایایی باشد. ثابت می‌کنیم که این تابع همان آنتروپی فون نیومان است. حالت دلخواه ρ را در نظر بگیرید. فرض کنید که قطری‌سازی ρ در یک پایه‌ی متعامد یکه به صورت

$$\rho = \sum_i \lambda_i |v_i\rangle\langle v_i|$$

باشد که در آن λ_i-ها مقادیر ویژه‌ی ρ هستند. تبدیل خطی زیر را در نظر بگیرید:

$$U|v_i\rangle = |i\rangle, \quad i = 0, 1, 2, \cdots d-1.$$

از آنجا که این تبدیل یک پایه متعامد یکه را به یک پایه متعامد یکه می‌برد، پس تبدیلی یکانی است و معادل با یک تحول

زمانی برگشت‌پذیر. اگر سیستم را تحت این تحول زمانی قرار دهیم حالت آن به صورت زیر تغییر می‌کند:

$$\begin{aligned}\rho \to U\rho U^\dagger &= U\big(\sum_i \lambda_i |v_i\rangle\langle v_i|\big)U^\dagger \\ &= \sum_i \lambda_i U|v_i\rangle\langle v_i|U^\dagger \\ &= \sum_i \lambda_i |i\rangle\langle i|\end{aligned}$$

که ماتریس چگالی یک حالت کلاسیک است. طبق خاصیت سازگاری آنتروپی آن برابر است با آنتروپی دنباله $\{\lambda_i\}$ که همان مقادیر ویژه $U\rho U^\dagger$ هستند. پس

$$H(U\rho U^\dagger) = -\sum_i \lambda_i \log(\lambda_i).$$

اما طبق خاصیت پایایی داریم:

$$H(\rho) = H(U\rho U^\dagger).$$

پس

$$H(\rho) = -\sum_i \lambda_i \log(\lambda_i) = -\mathrm{tr}(\rho\log(\rho)).$$

□

حال معانی عملیاتی آنتروپی را به طور خلاصه توضیح می‌دهیم. اگر نسخه‌های مستقل زیادی از یک هنگرد را در اختیار داشته باشیم، آنتروپی فون نیومان برابر میزان اطلاعات غیرقابل فشرده کردن هنگرد می‌باشد (همان طور که آنتروپی شانون طبق قضیه اول شانون برابر میزان اطلاعات غیر قابل فشرده کردن یک منبع کلاسیک است). به عبارت دیگر آنتروپی فون نیومان یک منبع برابر حداقل تعداد کیوبیت‌هایی است که نمونه‌های هنگرد را می‌توان در آنها ذخیره کرد به طوری که روشی برای بازیابی قابل اطمینان آنها وجود داشته باشد.

اما تفسیر عملیاتی دیگری از آنتروپی فون نیومان نیز وجود دارد. خواهیم دید که آنتروپی فون نیومان همچنین برابر ماکزیمم نرخ اطلاعات کلاسیکی که می‌توان توسط بهترین اندازه‌گیری ممکن از نمونه‌های هنگرد بدست آورد نیز هست.

آنتروپی فون نیومان کاربردهای زیادی در ارتباط با مفاهیمی دارد که معادل کلاسیک ندارند. مثلا اگر دو سیستم در یک حالت محض قرار گرفته باشند، از آنتروپی فون نیومان می‌توان برای یافتن میزان درهم‌تنیدگی استفاده کرد. بسیاری از ابزارهایی که در مطالعه آنتروپی فون نیومان مورد استفاده قرار می‌گیرند تعمیم ابزارهایی هستند که در تئوری اطلاعات کلاسیک از آنها بهره می‌بریم. خواهیم دید که همان طور که آنتروپی شانون ارتباط تنگاتنگی با مفهوم دنباله‌های نوعی دارد، آنتروپی فون نیومان نیز ارتباط تنگاتنگی با «زیرفضاهای» نوعی دارد.

۴ خواص آنتروپی فون نیومان

در اینجا برخی خواص مهم آنتروپی را ذکر می‌کنیم. اکثر این خواص شبیه خواص آنتروپی کلاسیک هستند. در مواردی که اثبات کوتاه است، پس از بیان خاصیت اثبات آن آمده است، اما اثبات اکثر این خواص به جلسات بعد موکول می‌شود.

- **حالت محض:** آنتروپی هر حالت محض $\rho = |\psi\rangle\langle\psi|$ صفر است: $H(\rho) = 0$. دلیل این موضوع این است که ماتریس دارای یک مقدار ویژه 1 است و بقیه مقادیر ویژه آن صفر هستند.

- **پایایی:** در صورتی که یک عملگر یکانی روی سیستم اعمال کنیم آنتروپی آن عوض نمی‌شود.

$$H(U\rho U^{\dagger}) = H(\rho).$$

جهت اثبات توجه کنید که یک تغییر پایه یکانی مقادیر ویژه یک ماتریس را تغییر نمی‌دهد.

- **مقدار ماکزیمم:** اگر ρ دارای D مقدار ویژه ناصفر باشد، یا به عبارتی اگر $\text{rank}\rho = D$ آنگاه آنتروپی آن در نامساوی زیر صدق می‌کند:

$$H(\rho) \leq \log D.$$

تساوی تنها وقتی برقرار می‌شود که تمامی مقادیر ویژه ناصفر با هم برابر باشند. آنتروپی ρ همان آنتروپی شانون دنباله‌ی مقادیر ویژه است و این خاصیت نتیجه‌ای از این موضوع است که توزیع یکنواخت آنتروپی شانون را بیشینه می‌کند و این مقدار بیشینه لگاریتم تعداد الفبای منبع است.

- **تحدب:** آنتروپی فون نیو مان محدب است. یعنی برای هر مجموعه از حالات $\rho_1, \rho_2, \cdots, \rho_k$ و احتمالات $p_1, p_2, \cdots, p_k$ (به طوری که $\sum_i p_i = 1$) داریم:

$$H\big(\sum_i p_i\rho_i\big) \geq \sum_i p_i H(\rho_i).$$

به علاوه اگر ماتریس‌های چگالی ρ_i دو به دو بر هم عمود باشند[۱۴] آنگاه اختلاف سمت راست و چپ نامساوی بالا را می‌توان به شکل آنتروپی کلاسیک دنباله احتمالات $\{p_i\}$ نوشت:

$$H\big(\sum_i p_i\rho_i\big) = \sum_i p_i H(\rho_i) + H(\{p_i\}). \qquad (۳)$$

ملاحظات: محدب بودن آنتروپی کوانتمی شبیه محدب بودن آنتروپی شانون است. بیاد آورید که محدب بودن آنتروپی شانون نتیجه می‌داد که اطلاعات مشترک دو متغیر کلاسیک نامنفی است زیرا عبارت

$$I(X;Y) = H(X) - H(X|Y) = H(X) - \sum_y p(y)H(X|Y=y)$$

[۱۴] ماتریس های چگالی ρ_i بر هم عمودند اگر $\rho_i\rho_j = 0$ برای هر $i \neq j$. اگر ماتریس های چگالی ρ_i بر هم عمود باشند فضای پشتیبان (Support) آنها دو زیرفضای برداری عمود بر هم تشکیل خواهند داد. فضای پشتیبان یک ماتریس چگالی با برد آن به عنوان یک عملگر یکسان است:

$$Supp(\rho) = \{\rho|v\rangle : |v\rangle \in \mathcal{H}_d\}.$$

اختلاف میان آنتروپی در نقطه $p(x)$ و میانگین وزن‌دار آن در نقاط $p(x|y)$ می‌باشد، و این وزن‌ها به گونه‌ای هستند که مرکز ثقل $\sum_y p(y)p(x|y)$ همان نقطه $p(x)$ است. مشابها محدب بودن آنتروپی کوانتمی نتیجه می‌دهد که اطلاعات مشترک میان یک متغیر تصادفی (سیستم کلاسیک) و یک سیستم کوانتمی نامنفی است.[۱۵]

تساوی (۳) تعمیم تساوی زیر برای آنتروپی کلاسیک است:

$$\begin{aligned}
&H(\{p_1, p_2, \cdots, p_{m-1}, p_m, p_{m+1}, \cdots, p_r\}) = \\
&\quad H(\{p_1 + \cdots + p_m,\ \ p_{m+1} + \cdots + p_r\}) \\
&\quad + (p_1 + \cdots + p_m) H(\frac{p_1}{p_1 + \cdots + p_m}, \frac{p_2}{p_1 + \cdots + p_m}, \cdots, \frac{p_m}{p_1 + \cdots + p_m}) \\
&\quad + (p_{m+1} + \cdots + p_r) H(\frac{p_{m+1}}{p_{m+1} + \cdots + p_r}, \frac{p_{m+2}}{p_{m+1} + \cdots + p_r}, \cdots, \frac{p_r}{p_{m+1} + \cdots + p_r}).
\end{aligned}$$

- **آنتروپی اندازه‌گیری:** در صورتی که حالت دلخواه ρ را در یک پایه متعامد یکه اندازه‌گیری کنیم، آنتروپی حاصل اندازه‌گیری (متغیر تصادفی کلاسیک Y) همواره بزرگتر یا مساوی آنتروپی فون نیومان ρ است:

$$H(Y) \geq H(\rho).$$

به علاوه تساوی برقرار است اگر و فقط اگر ρ در پایه انتخاب شده قطری باشد.

ملاحظات: رابطه‌ی بالا به این معنی است که اگر ρ را در پایه نامناسبی اندازه‌گیری کنیم، حاصل اندازه‌گیری «نویزی‌تر» خواهد بود و در نتیجه ابهام آن بیشتر است.

مثال ۸ فرض کنید که یک کیوبیت را در حالت $|0\rangle$ آماده کرده باشیم. در این صورت اگر این کیوبیت را در پایه

$$|v_0\rangle = \sin(\theta)|0\rangle + \cos(\theta)|1\rangle, \qquad |v_1\rangle = -\cos(\theta)|0\rangle + \sin(\theta)|1\rangle$$

اندازه‌گیری کنیم حاصل اندازه‌گیری با احتمال $\sin^2(\theta)$ برابر 0 و با احتمال $\cos^2(\theta)$ برابر 1 خواهد بود. در نتیجه آنتروپی حاصل اندازه‌گیری برابر $H(\{\sin^2(\theta), \cos^2(\theta)\})$ خواهد بود که به حداقل خود در $\theta = 0$ می‌رسد.

تمرین ۹ با یک مثال نشان دهید که اگر بجای اندازه‌گیری در یک پایه‌ی متعامد یکه، اندازه‌گیری دلخواهی را در نظر بگیریم آنگاه خاصیت «آنتروپی اندازه‌گیری» برقرار نیست.

- **آنتروپی تولید:** اگر یک هنگرد از حالات محض $\{p_i, |\psi_i\rangle\}$ تولید کنیم، آنگاه آنتروپی ماتریس چگالی حاصل حداکثر برابر آنتروپی دنباله احتمالات استفاده شده در تولید هنگرد خواهد بود. به عبارت دیگر

$$H\big(\sum_i p_i |\psi_i\rangle\langle\psi_i|\big) \leq H(\{p_i\}).$$

[۱۵] اطلاعات مشترک در حالت کوانتمی در ادامه تعریف خواهد شد.

تساوی فقط زمانی اتفاق خواهد افتاد که بردارهای $|\psi_i\rangle$ دو به دو بر هم عمود باشند.

ملاحظات: فرض کنید که طبیعت با احتمال p_i سیستم را در حالت $|\psi_i\rangle$ تولید کرده باشد؛ یعنی هنگردی از حالات محض $\{p_i, |\psi_i\rangle\}$ را تولید و نمونه ای از آن را در اختیار ما قرار داده باشد. در صورتی که بردارهای $|\psi_i\rangle$ بر هم عمود نباشند، قابلیت بازیابی اطلاعات مربوط به اینکه هنگرد را در کدام حالت میباشد را از دست خواهیم داد. زیرا با ادغام این خاصیت و خاصیت قبل میبینیم که هر اندازه گیری روی سیستم حداکثر $H(\rho)$ بیت به ما خواهد داد که از آنتروپی $H(\{p_i\})$ کمتر است. اگر بازیابی کامل ممکن بود نیز به کسب $H(\{p_i\})$ بیت اطلاعات داشتیم.

تمرین ۱۰ با یک مثال نشان دهید که اگر بجای یک هنگرد از حالات محض هنگرد دلخواهی را در نظر بگیریم، آنگاه خاصیت «آنتروپی تولید» برقرار نیست.

۱.۴ عبارات آنتروپیک

در صورتی که سیستم A در حالت ρ_A باشد، آنتروپی آن را با نمادهای $H(\rho_A)$ و یا $H(A)$ نشان می‌دهند. همچنین در مورد سیستم ترکیبی AB بجای $H(\rho_{AB})$ از $H(AB)$ استفاده می‌شود؛ گاهی جهت تاکید و مشخص کردن حالت سیستم از نمادگذاری $H(AB)_\rho$ نیز استفاده می‌شود. در دنیای کوانتمی همانند دنیای کلاسیک اطلاعات متقابل با

$$I(A;B) = H(\rho_A) + H(\rho_B) - H(\rho_{AB})$$

تعریف می‌شود که در آن

$$\rho_A = \mathrm{tr}_B(\rho_{AB}), \quad \rho_B = \mathrm{tr}_A(\rho_{AB})$$

جای توزیع‌های حاشیه‌ای را می‌گیرد. عبارات دیگری مانند $I(A;B|C)$ یا $H(A|B)$ به صورت مشابه با استفاده از بسط فرمول‌های آنتروپیک آنها تعریف می‌شوند. بنابراین

$$H(A|B) = H(AB) - H(B),$$

و

$$I(A;B|C) = H(A|C) + H(B|C) - H(AB|C).$$

نکته ۱۱ رابطه

$$H(X|Y) = \sum_y p(y) H(X|Y=y)$$

که در حالت کلاسیک داشتیم در مورد سیستم‌های کوانتمی $H(A|B)$ برقرار نیست. زیرا اصولا نوشتن عبارت $B=b$ در مورد سیستم کوانتمی B بی‌معنی است. زیرا تا سیستم را مشاهده نکرده‌ایم، مقدار خاصی ندارد و به علاوه حالت آن پس از انجام اندازه‌گیری تغییر می‌کند. بنابراین رابطه اصلی که برای تعریف $H(A|B)$ داریم همان $H(A,B) - H(B)$ است.

مفهوم	آنتروپی کلاسیک گسسته	آنتروپی کلاسیک تفاضلی	آنتروپی کوانتمی
آنتروپی $H(A)$	همواره نامنفی	مثبت یا منفی	همواره نامنفی
آنتروپی شرطی $H(A\|B)$	همواره نامنفی	مثبت یا منفی	مثبت یا منفی
اطلاعات متقابل $I(A;B)$	همواره نامنفی	همواره نامنفی	همواره نامنفی
اطلاعات متقابل شرطی $I(A;B\|C)$	همواره نامنفی	همواره نامنفی	همواره نامنفی

جدول ۲: مقایسه آنتروپی‌های کلاسیک و کوانتمی

با این حال برای یک سیستم کوانتمی A و یک سیستم کلاسیک Y رابطه‌ی فوق برقرار است:

$$H(A|Y) = H(A,Y) - H(Y) = \sum_y p(Y=y)H(A|Y=y) = \sum_y p(Y=y)H(\rho_y^A), \quad (۴)$$

که منظور از ρ_y^A حالت A به شرط $Y=y$ است. در این صورت حالت مشترک AY به صورت زیر خواهد بود:

$$\rho_{AY} = \sum_y p(y)|y\rangle\langle y| \otimes \rho_y^A.$$

تمرین ۱۲ تساوی (۴) را ثابت کنید.

نکته ۱۳ اطلاعات متقابل بین دو سیستم کوانتمی $I(A;B)$ تنها وقتی تعریف می‌شود که آن دو سیستم در یک لحظه از زمان با هم وجود داشته باشند. برای مثلا اگر A سیستمی باشد و پس از یک تحول زمانی تغییر حالت داده و به B تبدیل شود، آن وقت نمی‌توان صحبت از اطلاعات متقابل $I(A;B)$ کرد. همچنین اگر متغیر تصادفی حاصل از اندازه‌گیری A را X بنامیم، نمی‌توان صحبت از عباراتی مانند $I(A;X)$ یا $H(A|X)$ کرد.

توجه کنید که در دنیای کلاسیک مثلا اگر X یک متغیر تصادفی بوده و $Y=f(X)$ تابعی از X باشد، آنگاه $I(X;Y)$ معنی‌دار است به این دلیل که X حتی بعد از اعمال تابع f همچنان مقداری مشخص دارد. انگاری قبل از اعمال تابع یک «کپی» از X برداشته می‌شود با نام X'. در این صورت X' و Y در یک لحظه هم‌زمان وجود دارند و لذا می‌توانیم قرار دهیم $I(X;Y) = I(X';Y)$. به دلیل قضیه‌ی عدم کپی برداری[۱۶] این کار در دنیای کوانتمی امکان‌پذیر نیست.

نکته ۱۴ اگر سیستم ترکیبی AB تحت یک تبدیل یکانی (تحول زمانی) قرار بگیرند، آنگاه $H(AB)$ تحت این تحول زمانی تغییری نمی‌کند ($H(\rho_{AB}) = H(U_{AB}\rho_{AB}U_{AB}^\dagger)$)، اما آنتروپی‌های $H(A)$ و $H(B)$ ممکن است تغییر کنند. در نتیجه در حالت کلی ممکن است $I(A;B)$ و $H(A|B)$ تحت این تحول زمانی تغییر کنند.

۲.۴ نامساوی‌های آنتروپیک

نامساوی‌های آنتروپیک، نامساوی‌هایی هستند که در مورد آنتروپی و اطلاعات متقابل برقرار هستند. از آنجایی که اطلاعات متقابل را می‌توان بر حسب آنتروپی بیان کرد، تمامی این نامساوی‌ها را می‌توان بر حسب آنتروپی نوشت. در دنیای

[۱۶] No-cloning

کلاسیک نامساوی‌های آنتروپیک به دو دسته نامساوی‌های از نوع شانون[۱۷] و نامساوی‌های از نوع غیر شانون تقسیم می‌شوند. نامساوی‌های از نوع شانون نامساوی‌هایی هستند که نامنفی بودن آنتروپی و اطلاعات متقابل را تضمین می‌کنند. این نامساوی‌ها (در دنیای کلاسیک) عبارتند از

$$H(X) \geq 0, \quad H(X|Y) \geq 0, \quad I(X;Y) \geq 0, \quad I(X;Y|Z) \geq 0.$$

به نمودارهای ون شکل های ۵ و ۶ توجه کنید. نامساوی‌های فوق در واقع بیان می‌کنند که در حالت کلاسیک تنها ناحیه وسط

$$I(X;Y;Z) := I(X;Y) - I(X;Y|Z) = I(X;Z) - I(X;Z|Y) = I(Y;Z) - I(Y;Z|X)$$

می‌تواند منفی باشد و تمامی نواحی دیگر همواره نامنفی بودند. اما در حالت کوانتمی برخی از نواحی کناری (آنتروپی‌های شرطی) می‌توانند منفی باشند. با این حال تمامی اطلاعات متقابل همچنان نامنفی هستند. از این نظر آنتروپی فون نیومان شبیه آنتروپی تفاضلی پیوسته است که می‌تواند منفی یا مثبت باشد اما اطلاعات متقابل همواره مثبت می‌شود. اما باید توجه کرد که این تشابه اصلا دقیق نیست و ارتباط مفهومی خاصی میان این دو وجود ندارد. جدول ۲ روابط موجود میان این مفاهیم آنتروپی را خلاصه می‌کند.

اما چرا آنتروپی شرطی کوانتمی می‌تواند منفی باشد؟ یک جفت EPR را در نظر بگیرید:

$$|\psi\rangle_{AB} = \frac{1}{\sqrt{2}}\big(|00\rangle + |11\rangle\big).$$

داریم $\rho_A = \rho_B = \frac{1}{2}I$. از آنجایی که ρ_{AB} یک حالت محض است آنتروپی آن صفر است.

$$H(AB) = 0.$$

از طرف دیگر

$$H(A) = H(B) = H(\{1/2, 1/2\}) = 1.$$

پس

$$H(AB) < H(B) \quad \Rightarrow \quad H(A|B) < 0.$$

یعنی آنتروپی کل یک سیستم از آنتروپی اجزای آن می‌تواند کمتر باشد. در حالت کلی‌تر لم زیر را داریم:

لم ۱۵ اگر سیستم ترکیبی AB در حالت محض درهم‌تنیده باشند، یعنی

$$\rho_{AB} = |\varphi\rangle\langle\varphi|_{AB},$$

به طوری که $|\varphi\rangle$ به شکل ضرب تانسوری دو بردار نیست، در این صورت

$$H(AB) = 0, \quad \text{و} \quad H(A) = H(B) > 0$$

[۱۷] Shannon type

که نتیجه می‌دهد

$$H(A|B) = H(B|A) < 0, \quad \text{و} \quad I(A;B) = 2H(A) = 2H(B).$$

اثبات: تجزیه اشمیت $|\varphi\rangle_{AB}$ را به صورت

$$|\varphi\rangle_{AB} = \sum_{j=1}^{l} \mu_j |u_j\rangle_A |z_j\rangle_B$$

در نظر بگیرید که در آن $\{|u_j\rangle : j = 1, \ldots, l\}$ و $\{|z_j\rangle : j = 1, \ldots, l\}$ متعامد یکه هستند و $\mu_j > 0, j = 1, \ldots, l$ داریم.

$$\begin{aligned}
\rho_A &= \mathrm{tr}_B \left(|\varphi\rangle\langle\varphi|_{AB} \right) \\
&= \mathrm{tr}_B \left(\sum_{i,j=1}^{l} \mu_i \mu_j \left(|u_i\rangle_A |z_i\rangle_B \right) \left(\langle u_j|_A \langle z_j|_B \right) \right) \\
&= \sum_{i,j=1}^{l} \mu_i \mu_j |u_i\rangle\langle u_j|_A \delta_{ij} \\
&= \sum_{i=1}^{l} \mu_i^2 |u_i\rangle\langle u_i|_A
\end{aligned}$$

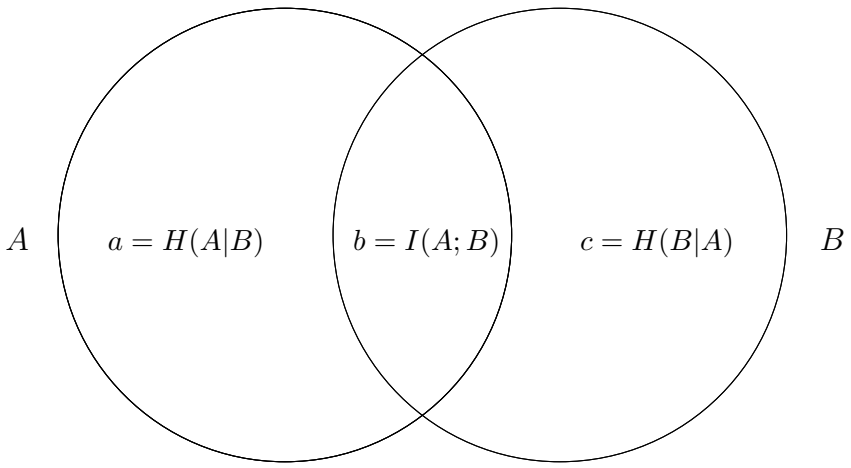
بنابراین مقادیر ویژه ρ_A برابر با $\{\mu_i^2\}$ هستند. به طور مشابه می‌توان نشان داد که مقادیر ویژه ρ_B هم برابر با $\{\mu_i^2\}$ هستند. پس $H(A) = H(B)$. اما چون حالت ما درهم‌تنیده بود، در تجزیه اشمیت حداقل دو جمله یا بیشتر داریم. پس $H(A) = H(B) > 0$. □

مثال ۱۶ فرض کنید که دو سیستم دلخواه A, B داریم، و R یک محض سازی از AB باشد. در این صورت ABR در حالتی محض قرار دارد و تساوی‌های زیر برقرارند.

$$H(ABR) = 0, \quad H(R) = H(AB), \quad H(RA) = H(B), \quad H(RB) = H(A)$$

۳.۴ نامساوی های آنتروپیک در نمودار ون

استفاده از نمودار ون در نوشتن، فهم و تحقیق نامساوی های آنتروپیک میتواند مفید باشد، زیرا هر نامساوی داده شده را میتوان بر حسب اجزای مختلف نمودار ون نوشت و با استفاده از نامساوی هایی که روی بخش های مختلف این شکل داریم به تحقیق درستی نامساوی مورد نظر پرداخت. در شکل ۵ نمودار ون برای دو متغیر آمده است. در حالت کلاسیک تمامی سه قسمتی که نشان داده شده نامنفی هستند، اما در حالت کوانتمی نامساوی $a \geq 0$ جای خود را به $a + \frac{b}{2} \geq 0$ می‌دهد. به طور مشابه $c \geq 0$ جای خود را به $c + \frac{b}{2} \geq 0$ می‌دهد. اگر این دو نامساوی را جمع بزنیم به نتیجه منطقی $a + b + c \geq 0$ می‌رسیم چون آنتروپی کل نامنفی است. در حالت کلاسیک $I(X;Y)$ را می‌توان با $H(X)$ از بالا کران زد، اما در حالت کوانتمی $I(A;B)$ را می‌توان با $2H(A)$ از بالا کران زد. دلیل این موضوع این است که در دنیای کوانتمی



شکل ۵: نمودار ون برای دو متغیر کوانتمی. در حالت کلاسیک تمامی سه قسمتی که نشان داده شده نامنفی هستند، اما در حالت کوانتمی نامساوی $a \geq 0$ جای خود را به $a + \frac{b}{2} \geq 0$ می‌دهد. به طور مشابه $c \geq 0$ جای خود را به $c + \frac{b}{2} \geq 0$ می‌دهد.

آنتروپی شرطی می‌تواند منفی شود، اما این میزان منفی شدن حداکثر به اندازه منهای آنتروپی خود سیستم می‌تواند باشد. نامساوی $I(A;B) \leq 2H(A)$ معادل با نامساوی $b \leq 2(a+b)$ می‌باشد که برقرار است.

در زیرنویس شکل ۶ خواص مربوط به آنتروپی را برای سه متغیر خلاصه کرده‌ایم.

تمرین ۱۷ در شکل ۵ نمودار ون برای دو متغیر و نامساوی $a + \frac{b}{2} \geq 0$ آمده است. این نامساوی را برای نمودار ون شکل ۶ برای دو متغیر A, C نوشته و تحقیق کنید که شرط جدیدی به جز آن چه در زیرنویس آن شکل آمده بدست نمی‌دهد.

۴.۴ بیانی مشروح تر از نامساوی های آنتروپیک

برخی از خواص آنتروپی که در بالا بصورت خلاصه به آنها اشاره کردیم به شرح زیر هستند:

- **زیرجمع پذیری:**[۱۸] اگر سیستم ترکیبی AB در حالت ρ_{AB} باشد، داریم:

$$I(A;B) = H(A) + H(B) - H(AB) \geq 0,$$

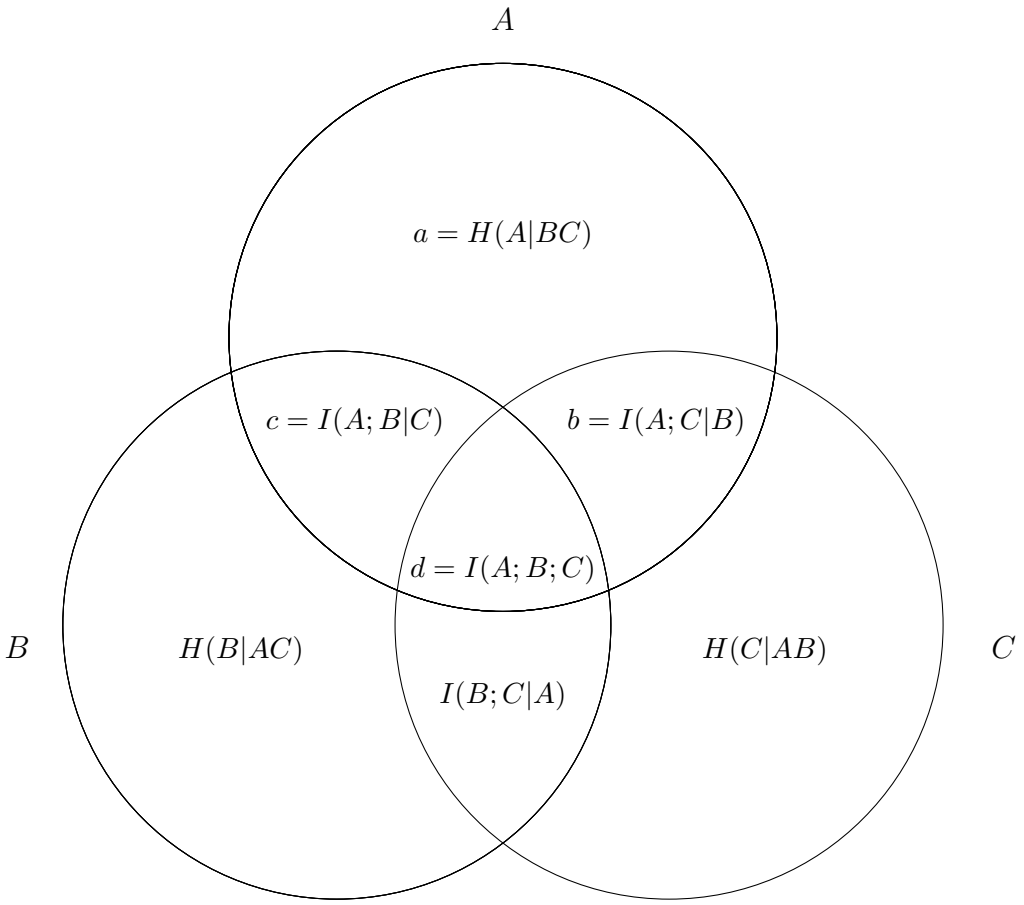
تساوی زمانی برقرار خواهد بود که $\rho_{AB} = \rho_A \otimes \rho_B$.

ملاحظات: این رابطه بیان می‌دارد که ابهام کل یک سیستم از جمع ابهام اجزای آن بیشتر نیست. همان‌طور که می‌دانیم این رابطه در دنیای کلاسیک نیز برقرار است.

مثال ۱۸ (شکلی از قانون دوم ترمودینامیک) فرض کنید که یک سیستم A و محیط خارج از آن، E، در حالتی مستقل از هم قرار گرفته باشند، یعنی $\rho_{AE} = \rho_A \otimes \rho_E$. حال اگر میان محیط و سیستم یک برهم‌کنش صورت بگیرد حالت سیستم به σ_{AE} تغییر می‌کند:

$$\sigma_{AE} = U \rho_{AE} U^\dagger.$$

[۱۸] Subadditivity



شکل ۶: نمودار ون برای متغیرهای کوانتمی. مثبت و منفی بودن بخش‌های کناری در حالت کلاسیک و کوانتمی با هم تفاوت می‌کنند. در حالت کلاسیک تنها ناحیه وسط $I(A;B;C)$ است که می‌تواند منفی شود. در حالت کوانتمی ناحیه وسط و نواحی کناری $I(A;B;C), H(A|BC), H(B|AC), H(C|AB)$ می‌توانند منفی شوند اما تمامی اطلاعات متقابل (شرطی یا غیرشرطی) نامنفی هستند. در مورد نواحی کناری شرط $a \geq 0$ از حالت کلاسیک جای خود را به $a + \frac{b+c}{2} \geq 0$ می‌دهد. همین رابطه در مورد ناحیه‌های گوشه‌ای دیگر نیز برقرار است. همچنین $a + \frac{b+c+d}{2} \geq 0$ که از نامساوی مربوط به دو سیستم نتیجه میشود، اگر A را به عنوان یک سیستم و (B,C) را به عنوان یک سیستم دیگر بگیریم. همین رابطه در مورد ناحیه‌های گوشه‌ای دیگر نیز برقرار است. این تنها تغییرات از حالت کلاسیک به کوانتمی است.

حال داریم:

$$H(A)_\rho + H(E)_\rho = H(AE)_\rho = H(AE)_\sigma \leq H(A)_\sigma + H(E)_\sigma.$$

یعنی جمع آنتروپی سیستم و آنتروپی محیط در اثر تحول زمانی زیاد شده است.

- **نامساوی مثلث (نامساوی آراکی-لیب[۱۹]):** برای هر سیستم ترکیبی AB داریم

$$H(AB) \geq \big|H(A) - H(B)\big|.$$

ملاحظات: این نامساوی را می‌توان به شکل دو نامساوی

$$H(AB) + H(B) \geq H(A), \quad \text{و} \quad H(AB) + H(A) \geq H(B),$$

نوشت. توجه کنید که در دنیای کلاسیک نامساوی قوی‌تر $H(XY) \geq H(X)$ برقرار است (یعنی نیازی به جمله اضافی $H(Y)$ نداریم).

- **زیرجمع پذیری قوی:**[۲۰] آنتروپی شرطی کوانتمی نامنفی است، یعنی برای هر سیستم ترکیبی ABC داریم:

$$I(A;B|C) \geq 0.$$

با بسط اطلاعات متقابل بر حسب آنتروپی به طور معادل داریم:

$$H(AC) + H(BC) \geq H(ABC) + H(C).$$

ملاحظات: برای بخاطر سپردن رابطه‌ی بالا به این نکته توجه کنید که ABC اجتماع AC و BC است، و C اشتراک آنها. رابطه بالا می‌گوید که آنتروپی دو سیستم بیشتر مساوی آنتروپی اجتماع به علاوه آنتروپی اشتراک آنهاست. اگر C را یک بعدی در نظر بگیریم، زیرجمع‌پذیری قوی به زیرجمع‌پذیری منجر می‌شود. بر خلاف حالت کلاسیک، اثبات این نامساوی برای متغیرهای کوانتمی آسان نیست.

۵.۴ کاربردهای زیرجمع پذیری قوی

زیرجمع پذیری قوی نتایج نابدیهی و جالبی دارد که در زیر به برخی از آنها اشاره می‌کنیم:

- **مشروط کردن باعث کاهش آنتروپی می‌شود:** برای هر سه سیستم دلخواه A, B, C داریم

$$H(A|B) \geq H(A|BC). \tag{۵}$$

این نامساوی معادل زیر جمع پذیری قوی است.

[۱۹] Araki-Lieb inequality
[۲۰] Strong subadditivity

- **دور ریختن سیستم‌ها باعث افزایش اطلاعات متقابل نمی‌شود:** برای هر سه سیستم دلخواه A, B, C داریم

$$I(A; BC) \geq I(A; B).$$

این نامساوی معادل زیر جمع پذیری قوی است.

- **قضیه پردازش داده:**[۲۱] اگر یک سیستم ترکیبی AB در حالت ρ_{AB} داشته باشیم و روی زیرسیستم B یک فرایند کوانتمی اعمال کنیم تا سیستم ترکیبی AB' در حالت $\sigma_{AB'} = \mathcal{I}_A \otimes \Phi_{B\to B'}(\rho_{AB})$ بدست آید، آنگاه

$$I(A; B)_\rho \geq I(A; B')_\sigma.$$

اثبات: توجه کنید که هر فرایند کوانتمی دلخواه روی سیستم B معادل است با اعمال ایزومتری $V_{B\to B'E}$ و بعد دور انداختن زیر سیستم E:

$$\sigma_{AB'} = \mathrm{tr}_E(\sigma_{AB'E}), \qquad \sigma_{AB'E} = (I_A \otimes V)\rho_{AB}(I_A \otimes V^\dagger).$$

حال با توجه به این که ایزومتری‌ها آنتروپی را تغییر نمی‌دهند و $\sigma_A = \rho_A$ داریم:

$$H(AB'E)_\sigma = H(AB)_\rho, \quad H(B'E)_\sigma = H(B)_\rho, \quad H(A)_\sigma = H(A)_\rho$$

و در نتیجه

$$I(A; B'E)_\sigma = I(A; B)_\rho.$$

حال از آنجا که دور ریختن زیرسیستم‌ها باعث افزایش اطلاعات متقابل (نامساوی قبل) نمی‌شود داریم

$$I(A; B')_\sigma \leq I(A; B'E)_\sigma = I(A; B)_\rho.$$

- **شکلی معادل از زیرجمع پذیری قوی:** برای هر سه سیستم دلخواه A, B, C داریم

$$H(A) + H(B) \leq H(CA) + H(CB)$$

یا

$$H(C|A) + H(C|B) \geq 0.$$

یعنی با اینکه $H(C|B)$ و $H(C|A)$ ممکن است به تنهایی منفی شوند، اما جمع آنها همواره نامنفی است. نامساوی بالا را به طور معادل به شکل زیر هم می‌توان نوشت:

$$I(A; B) + I(A; C) \leq 2H(A).$$

توجه کنید که این نامساوی در حالت کلاسیک برقرار است زیرا $I(X; Y) \leq H(X)$ و $I(X; Z) \leq H(X)$.

[۲۱] Data processing

اثبات: فرض کنید که R یک محض‌سازی از ABC باشد. در این صورت

$$H(R) = H(ABC), \quad H(RC) = H(AB).$$

در نتیجه

$$H(RC) - H(R) = H(AB) - H(ABC)$$

پس نامساوی زیرجمع پذیری قوی به صورت

$$H(RC) + H(BC) \geq H(R) + H(B)$$

در می‌آید که همان نامساوی مورد نظر است.

- **مقعر بودن آنتروپی شرطی:** همانند حالت کلاسیک آنتروپی شرطی تابعی مقعر است. یعنی اگر

$$\rho_{AB} = p\sigma_{AB} + (1-p)\mu_{AB}$$

آنگاه

$$H(A|B)_\rho \geq pH(A|B)_\sigma + (1-p)H(A|B)_\mu.$$

اثبات: X را یک سیستم کلاسیک گرفته و تعریف کنید

$$\rho_{XAB} = p|0\rangle\langle 0| \otimes \sigma_{AB} + (1-p)|1\rangle\langle 1| \otimes \mu_{AB}.$$

در این صورت $\mathrm{tr}_X(\rho_{XAB})$ همان ماتریس چگالی تعریف شده در بالاست. حال کافی است توجه کنیم که سمت راست نامساوی مورد نظر برابر است با $H(A|XB)$. اثبات با توجه به (۵) تمام است.

مثال ۱۹ A را یک سیستم کوانتمی و X را یک سیستم کلاسیک بگیرید:

$$\rho_{XA} = \sum_x p_x |x\rangle\langle x| \otimes \rho_x.$$

با توجه به بسط

$$H(A|X) = \sum_x p_x H(\rho_x),$$

می‌دانیم که $H(A|X)$ نامنفی است. ادعا می‌کنیم $H(X|A)$ نیز نامنفی است. برای این کار حالت

$$\rho_{XX'A} = \sum_x p_x |x\rangle\langle x|_X \otimes |x\rangle\langle x|_{X'} \otimes \rho_x$$

را در نظر بگیرید. X' در واقع یک کپی از X است و $\rho_{XX'A}$ یک توسعه از ρ_{XA} است. حال با استفاده از زیرجمع پذیری قوی (نامنفی بودن اطلاعات متقابل شرطی) داریم

$$0 \leq I(X;X'|A) = H(X|A) + H(X'|A) - H(XX'|A) = H(X|A).$$

نتیجه این که گرچه $H(A|B)$ می‌تواند منفی باشد، اگر هر یک از A یا B کلاسیک باشند حاصل نامنفی است.

تمرین ۲۰ با استفاده از زیر جمع پذیری قوی برای هر چهار سیستم دلخواه A, B, C, D ثابت کنید:

$$H(AB|CD) \leq H(A|C) + H(B|D).$$