معیارهای فاصله و شباهت در اطلاعات کوانتومی

وحید کریمی پور_ دانشکده فیزیک _ دانشگاه صنعتی شریف

۲۲ آذر ۱۴۰۲

۱ مقدمه

حالت های کوانتومی خالص بردارهایی در یک فضای هیلبرت هستند. حالت های کوانتومی ناخالص نیز عملگرهایی مثبت هستند که روی یک فضای هیلبرت عمل می کنند. بنابراین می توانیم فاصله بین این حالت های خالص یا ناخالص و هم چنین شباهت بین آنها را بر اساس آنچه که در بخش گذشته دیده ایم تعریف کنیم، اما سوال اصلی این است که آیا این فاصله ها معنای عملیاتی [۱] نیز دارند یا نه؟ مقصود از معنای عملیاتی یک تعریف این است که از نظر فیزیکی و آزمایشی آن تعریف می بایست معنای مشخص فیزیکی داشته باشد. ما هیچگاه یک حالت کوانتومی را در آزمایشگاه مشاهده نمی کنیم، بلکه آنچه که مشاهده می کنیم آماری است که از اندازه گیری های مختلف روی یک حالت کوانتومی بدست می آوریم و نه چیزی دیگر. انتظار داریم که اگر دو حالت کوانتومی نزدیک به هم باشند، آماری که در آزمایشگاه تولید می کنند نیز به هم نزدیک باشند. آیا می توان گفت که اگر دو حالت کوانتومی آماری یکسان تولید کنند، آن دو حالت نیز به هم نزدیک اند؟ به نظر می رسد که چنین نیست. به عنوان مثال اگر یک حالت کاملا ناخالص

$$\rho = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|) \qquad (۱)$$

و یک حال خالص مثل

$$|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \qquad (۲)$$

[۱] Operational Meaning

در پایه محاسباتی اندازه گیری کنیم هر دو آماری دقیقا یکسان تولید می کنند و حال آنکه این دو حالت کاملا از هم متمایزند. بنابراین نوع اندازه گیری ای که روی دو حالت انجام می دهیم نیز در آماری که در آزمایشگاه تولید می کنند و در قضاوتی که ما نسبت به فاصله آنها می کنیم موثر است. همه این ها نشان می دهد که قبل از این که تعریفی عملیاتی برای دوری یا نزدیکی حالت های کوانتومی ارائه کنیم، می بایست تعریفی عملیاتی برای دوری یا نزدیکی توابع توزیع احتمال کلاسیک پیدا کنیم. بنابراین سوال اصلی این درس این است که چگونه تعاریفی عملیاتی برای توابع توزیع احتمال و هم چنین برای حالت های کوانتومی پیدا کنیم. پاسخ این سوالات به ما خواهد آموخت که سوال های دیگری نیز بپرسیم و سعی کنیم به آنها پاسخ دهیم، سوالاتی مثل:

- شباهت و فاصله دو حالت آمیخته مثل ρ و σ چگونه تعریف می شود؟

- شباهت و فاصله بین دو کانال کوانتومی مثل Λ_1 و Λ_2 چگونه تعریف می شود؟

- آیا این تعاریف فاصله یکتاست یا می توان فاصله های گوناگون تعریف کرد؟ در این صورت این فاصله ها چه مزیت هایی نسبت به هم دارند؟

در این درس تعاریف مختلفی از مفهوم فاصله و شباهت هم در حوزه کلاسیک (بین توابع توزیع احتمال) و هم در حوزه کوانتومی (بین حالت های کوانتومی) ارائه خواهیم کرد.

---

۲ فاصله بین دو تابع توزیع احتمال

فرض کنید که p و q دو تابع توزیع احتمال روی یک متغیر تصادفی باشند، مثل دو احتمالات روی یک مهره تاس. مثلا می توانید این دو تابع احتمال را به صورت زیر در نظر بگیرید؛

$$p = \{\frac{4}{24},\ \frac{4}{24},\ \frac{3}{24}.\ \frac{3}{24},\ \frac{5}{24},\ \frac{5}{24}\} \qquad q = \{\frac{4}{24},\ \frac{5}{24},\ \frac{4}{24},\ \frac{3}{24},\ \frac{3}{24},\ \frac{5}{24}\}. \tag{۳}$$

از خود می پرسیم که اختلاف این دو تابع توزیع چقدر است؟ در نگاه اول می توانیم مثل هر دو تابع دیگری تفاوت این دو تابع را نیز به صورت زیر تعریف کنیم

$$d(p,q)=\sum_{x=1}^{6}|p_x-q_x|^2. \tag{۴}$$

یا به صورت کلی تر

$$d_a(p,q)=(\sum_{x=1}^{6}|p_x-q_x|^{\alpha})^{\frac{1}{a}}, \qquad a\geq 1. \tag{۵}$$

همه این تعریف ها از نظر ریاضی درست اند اما تنها یکی از آنهاست که معنای عملیاتی دارد و آن فاصله زیر است که برای یک متغیر تصادفی دلخواه چنین است:

$$D(p,q)=\frac{1}{2}\sum_x |p_x-q_x|. \tag{۶}$$

که در آن جمع روی تمام مقادیر متغیر تصادفی است. فضای نمونه متغیر تصادفی یعنی همه مقادیر ممکنی را که متغیر تصادفی می تواند اختیار کند با X نشان می دهیم. هرگاه متغیرتصادفی مقادیر پیوسته به خود بگیرد، رابطه بالا به شکل زیر در می آید:

$$D(p,q)=\frac{1}{2}\int dx|p(x)-q(x)|. \tag{۷}$$

رابطه ای که به این ترتیب تعریف شده است واقعا یک متریک یا فاصله را بین توابع توزیع احتمال تعریف می کند.

■ تمرین: نشان دهید که فاصله $D(p,q)$ دارای خواص زیر است:

$$\begin{aligned}
&0\leq D(p,q)\leq 1,\\
&D(p,q)=0 \leftrightarrow \quad p=q,\\
&D(p,q)=D(q,p),\\
&D(p,q)\leq D(p,r)+D(r,q).
\end{aligned} \tag{۸}$$

■ تمرین: یک مهره تاس را در نظر بگیرید. دو تابع توزیع مثال بزنید که فاصله آنها برابر با یک باشد.

این تعریف از فاصله چه خصوصیتی دارد؟ چه مزیتی بر تعاریف متعدد دیگری که می توانستیم جایگزین آن کنیم دارد؟ مثلا می توانیم فاصله را به صورت $D'(p,q):=\sum_x(P_x-q_x)^2$ تعریف کنیم. یا می توانستیم ضریب $1/2$ را به کار نبریم. در این جا می خواهیم نشان دهیم که این

تعریف یک معنای فیزیکی و شهودی خیلی خوب دارد. یک زیر مجموعه از مقادیر متغیر تصادفی مثل $S \subset X$ را در نظر بگیرید. در نظریه احتمالات اصطلاحا به S یک واقعه یا $Event$ می گوییم. از خود می پرسیم که p احتمال وقوع S را چقدر پیش بینی می کند و q این احتمال وقوع را چقدر پیش بینی می کند و تفاوت پیش بینی های آنها چقدر است؟ به یاد داشته باشیم که

$$p(S) = \sum_{x \in S} p_x, \qquad q(S) = \sum_{x \in S} q_x. \tag{۹}$$

برای بعضی از حوادث (یعنی زیر مجموعه های X) ممکن است پیش بینی آنها یکسان باشد یعنی $p(S) = q(S)$. از روی این نوع حوادث نمی توان در باره تفاوت این دو تابع توزیع احتمال قضاوت کرد. می توان گفت به طور میانگین پیش بینی آنها چقدر تفاوت دارد و یا این که بیشترین تفاوت آنها چقدر است؟ یعنی اینکه آن حادثه ای را در نظر بگیریم که پیش بینی این دو تابع توزیع احتمال برای ان حادثه بیشترین اختلاف را داشته باشد. تعریفی که در بالا ارائه شده است مبتنی بر این معیار دوم است. این امر در قضیه زیر بیان می شود:

- قضیه: فاصله دو تابع توزیع در رابطه زیر صدق می کند:

$$D(p,q) = max_{S \subset X} |p(S) - q(S)|, \tag{۱۰}$$

که در آن

$$|P(S) - q(S)| = |\sum_{x \in S} (p_x - q_x)| \tag{۱۱}$$

تفاوت پیش بینی p و q برای حادثه S است.

در واقع این نامساوی بیان می کند که حداقل یک زیر مجموعه خاص مثل S^* نیز وجود دارد به قسمی که

$$D(p,q) = |p(S^*) - q(S^*)|, \tag{۱۲}$$

و معنای قضیه بالا این است که

$$D(p,q) \geq |p(S) - q(S)| \qquad \forall\, S \subset X. \tag{۱۳}$$

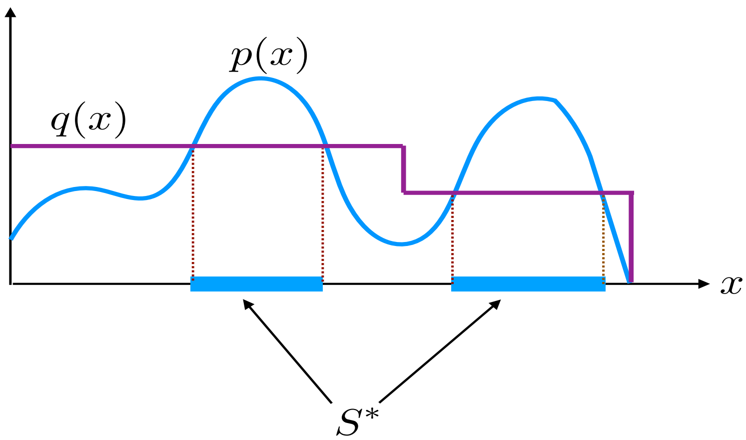
شکل ۱: در این شکل ها رنگ را به معنای مقدار تابع توزیع گرفته ایم. زیرمجموعه یا پیشامدِ S_1 تفاوت دو تابع توزیع را نشان نمی دهد. پیشامد S_2 تفاوت دو تابع توزیع را نشان می دهد. فاصله بین دو تابع توزیع بیشترین اختلاف پیش بینی آنها روی تمام پیشامدهاست.

■ اثبات: برای این که این موضوع را ثابت کنیم، ابتدا باید رابطه ۱۳ را ثابت کنیم و سپس نشان دهیم که یک زیر مجموعه خاص مثل S^* و جود دارد به نحوی که رابطه ۱۲ برقرار می شود. برای اثبات رابطه ۱۳ قرار می دهیم:

$$|p(S)-q(S)| = |\sum_{x\in S}(p_x-q_x)| = \frac{1}{2}|\sum_{x\in S}(p_x-q_x)| + \frac{1}{2}|\sum_{x\in S}(p_x-q_x)|. \tag{۱۴}$$

اما دقت می کنیم که بنا براین که هردو تابع توزیع احتمال بهنجار هستند، یعنی با توجه به اینکه

$$\sum_{x\in S} p_x + \sum_{x\in X-S} p_x = 1, \qquad \sum_{x\in S} q_x + \sum_{x\in X-S} q_x = 1, \tag{۱۵}$$



شکل ۲: قسمت پررنگ محدوده ای را نشان می دهد که در آن یکی از توابع توزیع از دیگری بزرگتر است. این ناحیه همان ناحیه ای است که اندازه گیری احتمال آن به بهترین نحوی تفاوت دو تابع توزیع را نشان می دهد.

نتیجه می گیریم که

$$\sum_{x\in S}(p_x - q_x) = \sum_{x\in X-S}(q_x - p_x). \tag{۱۶}$$

بنابراین خواهیم داشت

$$\begin{aligned} |p(S)-q(S)| &= \frac{1}{2}|\sum_{x\in S}(p_x - q_x)| + \frac{1}{2}|\sum_{x\in X-S}(p_x - q_x)| \\ &\leq \frac{1}{2}\sum_{x\in S}|(p_x - q_x)| + \frac{1}{2}\sum_{x\in X-S}|(p_x - q_x)| = \sum_{x\in X}|p_x - q_x| = D(p,q). \end{aligned} \tag{۱۷}$$

که در آن در خط دوم از نامساوی مثلث برای اعداد، یعنی $|a+b| \leq |a| + |b|$ استفاده کرده ایم. به این ترتیب نامساوی ۱۳ را نشان دادیم. حال باید نشان دهیم که واقعا یک مجموعه خاص یا به اصطلاح یک پیشامد خاص مثل S^* وجود دارد که به ازای آن رابطه ۱۲

برقرار می شود. مجموعه S^* را به صورت زیر تعریف می کنیم:

$$S^* := \{x \in X, \;\; p_x \geq q_x\}. \tag{۱۸}$$

در نتیجه خواهیم داشت:

$$X - S^* = \{x \in X, \;\; p_x < q_x\}. \tag{۱۹}$$

حال با توجه به روابط قبلی خواهیم داشت:

$$\begin{aligned} D(p,q) &= \frac{1}{2}\sum_{x\in X}|p_x - q_x| \\ &= \frac{1}{2}\sum_{x\in S^*}|p_x - q_x| + \frac{1}{2}\sum_{x\in X-S^*}|p_x - q_x| \\ &= \frac{1}{2}\sum_{x\in S^*}(p_x - q_x) + \frac{1}{2}\sum_{x\in X-S^*}(q_x - p_x) \\ &= \frac{1}{2}\sum_{x\in S^*}(p_x - q_x) + \frac{1}{2}\sum_{x\in S^*}(p_x - q_x) \\ &= \sum_{x\in S^*}(p_x - q_x) = |\sum_{x\in S^*}(p_x - q_x)| = |p(S^*) - q(S^*)|. \end{aligned} \tag{۲۰}$$

■ تمرین: دو مهره تاس مختلف با توابع توزیع زیر در نظر بگیرید:

$$\begin{aligned} &\{P_1 = \frac{2}{12}, \quad P_2 = \frac{1}{12}, \quad P_3 = \frac{3}{12}, \quad P_4 = \frac{1}{12}, \quad P_5 = \frac{3}{12}, \quad P_6 = \frac{2}{12}\} \\ &\{Q_1 = \frac{1}{12}, \quad Q_2 = \frac{5}{12}, \quad Q_3 = \frac{3}{2}, \quad Q_4 = \frac{1}{12}, \quad Q_5 = \frac{1}{12}, \quad Q_6 = \frac{1}{12}\} \end{aligned} \tag{۲۱}$$

الف: فاصله این دو تابع توزیع را حساب کنید.

ب: مجموعه S^* را برای این دو تابع توزیع بدست آورید.

پ: برای کدام مجموعه ها این دو تابع توزیع غیرقابل تمیز هستند؟

۱.۲ خواص فاصله بین دو تابع توزیع احتمال

در این بخش خواص فاصله بین دو تابع توزیع را مطالعه می کنیم. مهمترین خاصیتی که این فاصله دارد خاصیت تحدب قوی آن است. از نظر شهودی می دانیم که اگر دو تابع توزیع احتمال را با هم مخلوط کنیم، توزیعی که بدست می آوریم فاصله اش با یک توزیع دلخواه دیگر کم تر می شود زیرا ویژگی های مشخصه هر کدام از دو تابع توزیع در این مخلوط شدن کم رنگ می شود. قضیه زیر بیان کننده این موضوع است:

■ قضیه: فاصله دو تابع توزیع یک تابع مقعر است به این معنا که اگر q ، p_1 و p_2 توابع توزیع باشند آنگاه

$$D(\lambda p_1 + (1-\lambda)p_2, q) \leq \lambda D(p_1, q) + (1-\lambda) D(p_2, q). \tag{۲۲}$$

بهتراست که معنای قضیه را به زبان فارسی بنویسیم. این رابطه می گوید که مخلوط کردن توابع توزیع احتمال فاصله آن ها را از یک تابع توزیع دیگر کم می کند و آن ها را به هم شبیه تر می کند.

■ اثبات: با استفاده از تعریف فاصله می توانیم بنویسیم:

$$\begin{aligned} D(\lambda p_1 + (1-\lambda)p_2, q) &= \frac{1}{2}\sum_x |\lambda p_{1,x} + (1-\lambda)p_{2,x} - q_x| \\ &= \frac{1}{2}\sum_x |\lambda p_{1,x} + (1-\lambda)p_{2,x} - (\lambda + (1-\lambda))q_x| \\ &= \frac{1}{2}\sum_x |\lambda(p_{1,x} - q_x) + (1-\lambda)(p_{2,x} - q_x)|. \end{aligned} \tag{۲۳}$$

حال در طرف راست آخرین رابطه از نامساوی مثلث استفاده می کنیم که بر مبنای آن به ازای هر دو عدد دلخواه

$$|a+b| \leq |a| + |b|. \tag{۲۴}$$

در نتیجه خواهیم داشت:

$$\begin{aligned} D(\lambda p_1 + (1-\lambda)p_2, q) &\leq \lambda\frac{1}{2}\sum_x |p_{1,x} - q_x| + (1-\lambda)\frac{1}{2}\sum_x |p_{2,x} - q_x| \\ &= \lambda D(p_1, q) + (1-\lambda) D(p_2, q). \end{aligned} \tag{۲۵}$$

به این ترتیب اثبات قضیه کامل می شود.

حال با استفاده از تمرینی که در ضمیمه شماره یک حل کرده اید، می توانیم بنویسیم:

$$D(\sum_\alpha \lambda_\alpha P_\alpha, Q) \leq \sum_\alpha \lambda_\alpha D(P_\alpha, Q). \tag{۲۶}$$

تا کنون خاصیت تقعر فاصله را نسبت به تنها یکی از توابع توزیع نشان داده ایم. ما تابع فاصله دارای یک خاصیت تقعر قوی نیز هست که به صورت زیر بیان می شود: فرض کنید که p_i و q_i دو رشته تابع توزیع دلخواه روی یک متغیر تصادفی باشند. هم چنین فرض کنید که $\{\lambda_i\}$ و $\{\mu_i\}$ نیز دو رشته اعداد مثبت باشند که مجموع هرکدام به تنهایی برابر با ۱ باشد. دقت کنید که هرکدام از رشته های $\lambda := \{\lambda_i\}$ و $\mu := \{\mu_i\}$ را نیز به صورت یک تابع توزیع گسسته احتمال در نظر گرفت. دراین صورت می دانیم که

$$P := \sum_i \lambda_\alpha P_\alpha, \qquad Q := \sum_\alpha \mu_\alpha Q_\alpha, \tag{۲۷}$$

نیز دو تابع توزیع احتمال خواهند بود. دقت کنید که این دو تابع توزیع هر کدام جمع محدب توابع توزیع p_i و q_i هستند. در این صورت قضیه زیر برقرار است:

قضیه:

$$D(\sum_\alpha \lambda_\alpha P_\alpha, \sum_\alpha \mu_\alpha Q_\alpha) \leq D(\lambda, \mu) + \sum_\alpha \lambda_\alpha D(P_\alpha, Q_\alpha). \tag{۲۸}$$

که در آن $D(\lambda, \mu)$ فاصله تابع توزیع λ و μ است.

اثبات:

با اضافه کردن و کم کردن یک جمله مناسب و استفاده از نامساوی مثلث و هم چنین توجه به اینکه μ_α ها و λ_α ها مثبت هستند، به روابط زیر می رسیم.

$$\begin{aligned}
& D(\sum_\alpha \lambda_\alpha P_\alpha, \sum_\alpha \mu_\alpha Q_\alpha) \\
= & \frac{1}{2}\sum_i |\sum_\alpha \lambda_\alpha p_{\alpha,i} - \sum_\alpha \mu_\alpha q_{\alpha,i}| \\
= & \frac{1}{2}\sum_i |\sum_\alpha (\lambda_\alpha - \mu_\alpha) p_{\alpha,i} + \mu_\alpha (p_{\alpha,i} - q_{\alpha,i})| \\
\leq & \frac{1}{2}\sum_i \sum_\alpha |\lambda_\alpha - \mu_\alpha| p_{\alpha,i} + \frac{1}{2}\sum_i \sum_\alpha \mu_\alpha |p_{\alpha,i} - q_{\alpha,i}| \\
= & \frac{1}{2}\sum_\alpha |\lambda_\alpha - \mu_\alpha| + \frac{1}{2}\sum_\alpha \mu_\alpha D(P_\alpha, Q_\alpha).
\end{aligned} \tag{۲۹}$$

در آخرین سطر نیز از این که $\sum_i p_{\alpha,i} = 1$ استفاده کرده ایم.

به چند حالت خاص از این قضیه کلی توجه می کنیم:

- نتیجه یک: در رابطه ۹ فرض کنید که $q_i = q$ و $\lambda_i = \mu_i$. دراین صورت خواهیم داشت:

$$D(\sum_i \lambda_i p_i, q) \leq \sum_i \lambda_i D(p_i, q). \tag{۳۰}$$

که معنایش این است که فاصله نسبت به هرکدام از آرگومان هایش خاصیت تحدب دارد.

- نتیجه دو: در رابطه ۹ فرض کنید که $\lambda_i = \mu_i$. در این صورت خواهیم داشت:

$$D(\sum_i \lambda_i p_i, \sum_i \lambda_i q_i) \leq \sum_i \lambda_i D(p_i, q_i). \tag{۳۱}$$

---

۳ اهمیت روابط تحدب و تقعر:

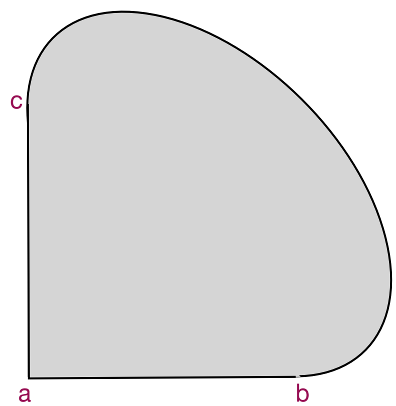
در بسیاری از موارد می خواهیم نقطه می نیمم یا ماکزیمیم یک تابع پیچیده را حساب کنیم. فضای متغیرهای این تابع می تواند دارای یک متغیر یا چند متغیر باشد یا حتی این فضا می تواند مجموعه ای از توابع یا ماتریس ها باشد. در چنین مسایلی اگر تابع مورد نظر یک تابع محدب یا مقعر باشد، بلافاصله می توانیم اطلاعات بسیار مهی در باره این نقاط بدست آوریم.

از نظر شهودی و با نگاه کردن به شکل (۳) معلوم است که یک تابع محدب، حتما می نیمم هایش را در نقاط انتهایی یا حدی مجموعه اختیار می کند. هم چنین یک تابع مقعر، مقدار ماکزیمم اش را در نقاط حدی مجموعه انتخاب می کند. برای اینکه معنای دقیق تر این گزاره را بفهمیم نخست باید مفهوم نقطه حدی [۲] را در یک مجموعه محدب تعریف کنیم.

- **تعریف:**] در یک مجموعه محدب، یک نقطه، مثل نقطه $x \in S$ نقطه حدی نامیده می شود هر گاه نتوانیم آن را به صورت مجموع محدب دو نقطه از داخل همان مجموعه نوشت. به عبارت بهتر x یک نقطه حدی است اگر نتوان آن را به صورت $x = \lambda y + (1-\lambda) z \quad y, z \in S$ نوشت. در شکل (۴) نقاط $a,\ b,\ c$ و تمام نقاط کمان bc نقاط حدی هستند. ولی بقیه نقاط از جمله نقاط داخل پاره خطهای ab و ac نقطه حدی نیستند. دقت کنید که نقطه حدی با نقطه مرزی فرق دارد، همان طور که در این شکل نشان داده می شود.

[۲] extreme point

شکل ۳: یک تابع محدب حتما می نیمم های خود را در نقاط حدی انتخاب می کند. یک تابع مقعر نیز ماکزیمم های خود را در نقاط حدی انتخاب می کند.



شکل ۴: یک مجموعه محدب و نقاط حدی آن. توضیح نقاط حدی در متن داده شده است.

پس از این تعریف می توانیم به قضیه زیر بپردازیم.

■ **قضیه:** یک تابع محدب که روی یک مجموعه محدب تعریف شده باشد، حتما می نیمم های خود را در نقاط حدی اختیار می کند.

■ **اثبات:** از برهان خلف استفاده می کنیم. فرض کنید که تابع $F : S \longrightarrow R$ روی مجموعه محدب S تعریف شده باشد و می نیمم خود را در یک نقطه غیر حدی اختیار کرده باشد. این نقطه غیرحدی را به صورت $x = \sum_i \lambda_i y_i$ در نظر می گیریم که در آن y_i ها نقاط حدی

باشند. در این صورت چون تابع محدب است می نویسیم:

$$F(x) \equiv F(\sum_i \lambda_i y_i) \geq \sum_i \lambda_i F(y_i) \geq (\sum_i \lambda_i) F(y_0) = F(y_0) \tag{۳۲}$$

که در آن y_0 نقطه ای بوده است که در جمع طرف راست بالا کمترین مقدار را داشته است، یعنی $F(y_0) < F(y_i) \quad \forall\, i$. به این ترتیب نقطه ای حدی پیدا کرده ایم که تابع F مقدارش در آنجا کمتر از مقدار تابع در نقطه x است. به این ترتیب حکم ثابت می شود.

علاوه بر این، محدب بودن تابع یک فایده مهم دیگر هم دارد و آن این است که برای یافتن ماکزیمم تابع با روش های عددی می توان از روش شیب گرادیان [۳] استفاده کرد. شبیه این قضیه نیز برای توابع مقعر درست است به این معنا که:

- **قضیه:** یک تابع مقعر که روی یک مجموعه محدب تعریف شده باشد، حتما ماکزیمم های خود را در نقاط حدی اختیار می کند. برای یافتن می نیمم تابع نیز می توان از روش شیب گرادیان استفاده کرد.

۱.۳ تبدیلات استوکاستیک و اثر آنها روی فاصله

یک تبدیل استوکاستیک یا به عبارت بهتر یک تبدیل مارکوف تبدیلی است که روی توزیع های احتمال رخ می دهد و آنها را به توزیع های احتمال دیگر تبدیل می کند. چنین تبدیلی به شکل زیر است:

$$\mathbf{q} = T\mathbf{p} \tag{۳۳}$$

که وقتی آن را به شکل ماتریسی بنویسیم به صورت زیر خواهد بود:

$$\begin{pmatrix} q_1 \\ q_2 \\ \cdot \\ \cdot \\ q_N \end{pmatrix} = \begin{pmatrix} T_{11} & T_{12} & \cdot & \cdot & T_{1N} \\ T_{21} & T_{22} & \cdot & \cdot & T_{2N} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ T_{N1} & T_{N2} & \cdot & \cdot & T_{NN} \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ \cdot \\ \cdot \\ p_N \end{pmatrix} \tag{۳۴}$$

[۳] Gradient Descent

چنین ماتریسی می بایست در دو شرط اصلی زیر صدق کند که اولی برای این است که احتمالات (که اعداد مثبت هستند) به اعداد مثبت نگاشته شوند و دومی هم برای این است که مجموع احتمالات همواره برابر با یک باقی بماند

$$T_{ij} \geq 0 \qquad \sum_i T_{ij} = 1 \ \ \forall\, i. \tag{۳۵}$$

چنین ماتریس های ماتریس های کاتوره ای یا استوکاستیک [۴] نامیده می شوند. این تبدیلات آنالوگ کلاسیک کانال های کوانتومی هستند. از این میان دسته خاصی از ماتریس ها هستند که مجموع درایه های روی هر سطر آنها نیز برابر با یک است، یعنی

$$\sum_j T_{ij} = 1 \qquad \forall\, j. \tag{۳۶}$$

این ماتریس ها که استوکاستیک دوگانه [۵] نامیده می شوند دارای این خاصیت هستند که تابع توزیع یکنواخت را به تابع توزیع یکنواخت می نگارند و آنالوگ کانال های یونیتال هستند. هر تبدیل استوکاستیک فاصله بین توابع توزیع احتمال را کم می کند و این همان خاصیتی است که در مورد کانال های کوانتومی نیز دیدیم. بنابراین می خواهیم نشان دهیم که

- **قضیه:** به ازای هر ماتریس استوکاستیک T رابطه زیر برقرار است:

$$d(T\mathbf{p}, T\mathbf{q}) \leq d(\mathbf{p}, \mathbf{q}). \tag{۳۷}$$

- **اثبات:** از تعریف فاصله استفاده می کنیم و سطر اول را می نویسم و سپس از نامساوی مثلث $|a+b| \leq |a| + |b|$ و مثبت بودن T_{ij} ها در سطر دوم استفاده می کنیم و سرانجام از رابطه (۳۶) در سطر آخر بهره می بریم:

$$\begin{aligned} d(T\mathbf{p}, T\mathbf{q}) &= \frac{1}{2}\sum_i |T(\mathbf{p}-\mathbf{q})_i| = \frac{1}{2}\sum_i |\sum_j T_{ij}(p_j - q_j)| \\ &\leq \frac{1}{2}\sum_i \sum_j |T_{ij}(p_j - q_j)| = \frac{1}{2}\sum_i \sum_j T_{ij}|(p_j - q_j)| \\ &= \frac{1}{2}\sum_j (\sum_i T_{ij})|(p_j - q_j)| = \frac{1}{2}\sum_j |(p_j - q_j)| = d(\mathbf{p}, \mathbf{q}). \end{aligned} \tag{۳۸}$$

---

[۴] Stochastic Matrices

[۵] Bi-Stochastic

■ **تمرین:** سه ماتریس کاتوره ای به شکل

$$T=\frac{1}{N}\sum_{i,j}|i\rangle\langle j| \ \ , \ \ P=|1\rangle\langle 2|+|2\rangle\langle 1|+\sum_{i\neq 1,2}|i\rangle\langle i| \ \ , \ \ I=\sum_i |i\rangle\langle i|$$

را در نظر بگیرید و فاصله هر جفت از آنها را با هم حساب کنید.

۴ شباهت دو تابع توزیع احتمال

حال به معرفی شباهت دو تابع توزیع احتمال می پردازیم که به صورت زیر تعریف می شود:

$$F(p,q):=\sum_x \sqrt{p_x q_x} \tag{۳۹}$$

که برای متغیرهای تصادفی پیوسته به شکل زیر در می آید:

$$F(p,q):=\int dx\sqrt{p(x)q(x)}. \tag{۴۰}$$

ممکن است فکر کنیم چرا شباهت دو تابع توزیع احتمال را به صورت

$$F_2(p,q)=\sum_x p_x q_x \tag{۴۱}$$

تعریف نکرده ایم. یک دلیل اش این است که در این صورت شباهت هر تابع توزیع احتمال با خودش برابر یک نمی شود و حال آنکه یک ملاک شباهت می بایست همواره مقدار این شباهت را با یک برابر کند. برای فهم بهتر و بیشتر، تمرین زیر را حل کنید.

■ تمرین: توابع توزیع زیر را در نظر بگیرید که روی یک مجموعه N تایی تعریف شده اند:

$$\begin{aligned} P_1 &= (1,0,0,0\cdots 0)\\ P_2 &= \frac{1}{2}(1,1,0,0\cdots 0)\\ P_3 &= \frac{1}{3}(1,1,1,0\cdots 0)\end{aligned}$$

$$........$$

$$........ \quad (۴۲)$$

الف: شباهت هر کدام از این توابع را با خودش بدست آورید، یک بار برای تابع شباهت F و یک بار هم برای تابع شباهت F_2.

ب: فاصله های زیر را نیز بدست آورید:

$$d(P_k, P_{k+1}). \quad (۴۳)$$

شباهت دو تابع توزیع احتمال خاصیت های زیر را دارد:

$$\begin{aligned} &F(p,q) = F(q,p), \\ &F(p,q) = 1 \leftrightarrow p = q, \\ &0 \leq F(p,q) \leq 1. \end{aligned} \quad (۴۴)$$

■ **تمرین** : خواص بالا را ثابت کنید. (راهنمایی: از نامساوی کوشی-شوارتز برای دو بردار مناسب استفاده کنید.)

■ **تمرین** : شباهت دو تابع توزیع گاووسی یک متغیره را که هرکدام با یک مقدار میانگین و انحراف معیار مخصوص به خود مشخص می شوند محاسبه کنید.

۱.۴ خواص تحدب برای تابع شباهت

در بخش قبلی ثابت کردیم که تابع فاصله محدب رو به پایین $Convex\ down$ (یعنی یک تابع دارای کمینه) است. در این بخش ثابت می کنیم که تابع شباهت یک تابع محدب رو به بالا $Convex\ up$ (یعنی یک تابع دارای بیشینه) است.

■ از روی شکل تابع $f(x)=\sqrt{x}$ براحتی می توان دید که این تابع یک تابع محدب است یعنی

$$\sqrt{\lambda x+(1-\lambda)y}\geq \lambda\sqrt{x}+(1-\lambda)\sqrt{y}.$$

با استفاده از این تعریف معلوم می شود که

$$F(\lambda p_1+(1-\lambda)p_2,q)\geq \lambda F(p_1,q)+(1-\lambda)F(p_2,q). \tag{۴۵}$$

از نظر شهودی معنای این رابطه این است که مخلوط کردن توزیع های احتمال شباهت آنها را با یک تابع توزیع مشخص زیاد می کند. این خاصیت به شکل کلی تری نیز برقرار است. نخست حالت ساده این خاصیت را بیان و ثابت می کنیم . اثبات کلی آن کاملا مشابه است و ایده جدیدی در بر ندارد.

■ قضیه: فرض کنید که p_1,p_2,q_1 و q_2 توابع توزیع احتمال هستند و پارامترهای λ و μ اعداد مثبت بین صفر و یک هستند. در این صورت نامساوی زیر برقرار است:

$$F(\lambda p_1+(1-\lambda)p_2,\mu q_1+(1-\mu)q_2)\geq \sqrt{\lambda\mu}F(p_1,q_1)+\sqrt{(1-\lambda)(1-\mu)}F(p_2,q_2) \tag{۴۶}$$

هرگاه در این رابطه قرار دهیم $q_1=q_2$ در این صورت و $\lambda=\mu$ این رابطه تبدیل به رابطه (۴۵) می شود.

برای اثبات این قضیه کافی است که تمرین زیر را حل کنید.

■ تمرین: الف: فرض کنید که اعداد a,b,c و d اعداد مثبتی هستند. نامساوی زیر را ثابت کنید:

$$\sqrt{[\lambda a+(1-\lambda)b][\mu c+(1-\mu)d]}\geq \sqrt{\lambda\mu}\sqrt{ac}+\sqrt{(1-\lambda)(1-\mu)}\sqrt{bd}. \tag{۴۷}$$

راهنمایی: مجذور طرفین را باهم مقایسه کنید.

ب: با استفاده از این نامساوی نشان دهید که رابطه (۴۶) برقرار است.

شکل کلی تر این قضیه نیز برقرار است:

■ قضیه:

$$F(\sum_i \lambda_i p_i, \sum_i \mu_i q_i) \geq \sum_i \sqrt{\lambda_i \mu_i} F(p_i, q_i). \qquad (۴۸)$$

■ تمرین : با استفاده از تمرین قبلی نامساوی کلی بالا را ثابت کنید.

بازهم می توان حالت های خاص را بررسی کرد. هرگاه در دو طرف این رابطه قرار دهیم $q_i = q$ و $\mu_i = \lambda_i$ به رابطه تحدب زیر می رسیم:

$$F(\sum_i \lambda_i p_i, q) \geq \sum_i \lambda_i F(p_i, q). \qquad (۴۹)$$

و هرگاه تنها قرار دهیم $\lambda_i = \mu_i$ به رابطه زیر می رسیم:

$$F(\sum_i \lambda_i p_i, \sum_i \lambda_i q_i) \geq \sum_i \lambda_i F(p_i, q_i). \qquad (۵۰)$$

از نظر شهودی معنای تمام این رابطه ها این است که مخلوط کردن توابع توزیع احتمال شباهت آنها را با یکدیگر زیاد و فاصله آنها را باهم کم می کند.

۵ آنتروپی نسبی

قبل از پایان بررسی خود بهتر است به یک فاصله مهم دیگر بین توابع توزیع کلاسیک نیز بپردازیم که به دنیای کوانتومی نیز تعمیم پیدا می کند. این فاصله آنتروپی نسبی [۶] نام دارد و به صورت زیر تعریف می شود:

$$S(p||q) := \sum_m p_m \log \frac{p_m}{q_m}. \qquad (۵۱)$$

این فاصله مثبت است و در رابطه $S(p||p) = 0$ نیز صدق می کند، اما به وضوح متقارن نیست. برای اینکه ببینیم این تابع مثبت است کافی است منحنی $\ln(x)$ و $x - 1$ را رسم کنیم تا متوجه شویم که همواره اتحاد زیر برقرار است:

[۶] Relative Entropy

$$\ln(x) \leq x - 1 \tag{۵۲}$$

و از آنجا برای هر دو عدد مثبتِ $p_i,\ q_i$

$$\ln(\frac{p_i}{q_i}) \leq \frac{p_i}{q_i} - 1. \tag{۵۳}$$

حال اگر دو تابع توزیع احتمال داشته باشیم، نتیجه می گیریم:

$$\sum_i q_i \ln(\frac{p_i}{q_i}) \leq p_i - q_i = 0. \tag{۵۴}$$

که مثبت بودن انتروپی نسبی را ثابت می کند. (به عوض شدن جای p_i و q_i دقت کنید.) ممکن است بپرسیم چرا چنین شکل غریبی به عنوان فاصله بین دو تابع توزیع احتمال معرفی شده است. پاسخ اش این است که بازهم یک معنای عملیاتی است که این فاصله از آن برخوردار است. فرض کنید که یک متغیرتصادفی $X = \{x_1,\ x_2,\ \cdots x_k\}$ دارای تابع توزیع احتمال $\{p_1,\ p_2,\ \cdots p_k\}$ است. حال به تعداد N بار از این متغیر تصادفی نمونه گیری می کنیم و مقادیر متغیرهای X به تعداد $\{n_1,\ n_2,\ \cdots n_k\}$ بار ظاهر می شوند. سوال این است: احتمال این که ما از نمونه گیری انجام شده قضاوت کنیم که تابع توزیع $\{q_1,\ q_2, \cdots q_k\}$ است چقدر است؟ پاسخ این سوال این است:

$$\mathcal{P}(q|p) = e^{-NS(q||p)+O(log(N))}. \tag{۵۵}$$

این رابطه را می توان به صورت زیر فهمید. ما وقتی چنین قضاوتی می کنیم که داشته باشیم:

$$n_1 = q_1 N,\ n_2 = q_2 N,\ \cdots\ n_k = q_k N. \tag{۵۶}$$

بنابراین، احتمال اینکه توزیعی را که واقعا p ما به صورت q قضاوت کنیم برابر است با:

$$\mathcal{P}(q|p) = \frac{N!}{\prod_i (Nq_i)!} \prod_i (p_i)^{Nq_i}. \tag{۵۷}$$

با بسط این رابطه و استفاده از تقریب استرلینگ به رابطه (۵۵) می رسیم.

■ تمرین: فاصله آنتروپی نسبی بوضوح متقارن نیست و با توجه به تعریف اش باید هم چنین باشد. یک مهره تاس انتراعی N حالته را با دو تابع توزیع زیر در نظر بگیرید:

$$p = (\frac{1}{N},\ \frac{1}{N}, \frac{1}{N}, \cdots, \frac{1}{N}) \qquad q = (1,\ 0,\ 0,\ 0,\ \cdots,\ 0) \tag{۵۸}$$

آنتروپی های نسبی $S(p||q)$ و $S(q||p)$ را حساب کنید و نتیجه خود را از نظر شهودی توصیف کنید. این دو فاصله را در حد $N \longrightarrow \infty$ با هم مقایسه کنید.

۶ فاصله دو حالت کوانتومی

در این بخش و بخش های بعدی خواهیم دید که با تعاریف مناسبی که از این تعاریف کلاسیک الهام گرفته است می توان فاصله و شباهت را برای حالت های کوانتومی نیز تعریف کرد و این کمیت ها دارای همان خاصیت هایی هستند که تا کنون برای فاصله و شباهت کلاسیکی بیان کرده ایم. نخست از فاصله دو حالت کوانتومی آغاز می کنیم. فرض کنید که ρ و σ دو حالت کوانتومی در یک فضای هیلبرت دلخواه باشند. در این صورت فاصله این دو حالت به شکل زیر تعریف می شود:

$$D(\rho,\sigma) := \frac{1}{2}tr|\rho-\sigma|, \tag{۵۹}$$

که در آن $|A| := \sqrt{A^\dagger A}$. اگر A یک تجزیه طیفی به صورت $A=\sum_i \alpha_i|\phi_i\rangle\langle\phi_i|$ داشته باشد، آنگاه $|A|=\sum_i |\alpha_i||\phi_i\rangle\langle\phi_i|$. این فاصله، فاصله ردّ [۷] خوانده می شود. این تعریف نیز بنا بر خواص فیزیکی و مفهومی آن انتخاب شده است. این فاصله بوضوح خاصیت های زیر را داراست:

$$D(\rho,\sigma)=0, \ \ \leftrightarrow \rho=\sigma,$$
$$D(\rho,\sigma)\leq D(\rho,\mu)+D(\mu,\sigma),$$
$$D(\rho,\sigma)=D(\sigma,\rho),$$

[۷] Trace Distance

$$D(U\rho U^\dagger, U\sigma U^\dagger) = D(\rho, \sigma). \quad (۶۰)$$

از این خاصیت ها تنها نامساوی مثلث بدیهی نیست. این خاصیت را در ادامه درس ثابت می کنیم. قبل از آن بهتر است با خواص فاصله آشنایی بیشتری پیدا کنیم.

■ **تمرین** : اگر ρ ، σ حالت های یک کیوبیت باشند که متناظر با بردارهای $\mathbf{r}$ و $\mathbf{s}$ باشند نشان دهید که

$$D(\rho, \sigma) = \frac{1}{2}|\mathbf{r} - \mathbf{s}|. \quad (۶۱)$$

هرگاه حالت های ρ و σ باهم جابجا شوند می توان پایه ای انتخاب کرد که هردو این ماتریس ها در آن قطری شوند. در این صورت خواهیم داشت:

$$\rho = \sum_i p_i|i\rangle\langle i|, \qquad \sigma = \sum_i q_i|i\rangle\langle i|. \quad (۶۲)$$

■ **تمرین** : نشان دهید که در این صورت خواهیم داشت:

$$D(\rho, \sigma) = D(p, q), \quad (۶۳)$$

که در آن $D(p,q)$ فاصله بین دو تابع توزیع احتمال p و q است.

حال به مهم ترین خاصیت این تعریف از فاصله می رسیم که معنای عملیاتی آن را روشن می کند. هرچه که هست ما دسترسی مستقیمی به توابع موج یا حالت های کوانتومی نداریم، آنچه که از آنها می فهمیم همان آماری است که در اندازه گیری های مختلف در آزمایشگاه تولید می کنند. دو حالت نزدیک به هم همواره آمار نزدیک به هم تولید می کنند ولی برعکس آن درست نیست. به عنوان مثال حالت های $\rho = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|)$ و $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ در اندازه گیری پایه Z آمار یکسان تولید می کنند حال آنکه این دو حالت کاملا با هم متفاوتند. بنابراین مهم است که رابطه بین فاصله بین حالت های کوانتومی و فاصله توابع توزیعی که در آزمایشگاه تولید می کنند را به خوبی بفهمیم. قضیه زیر این رابطه را روشن می کند.

■ **قضیه:** فرض کنید که یک اندازه گیری کلی $E = \{E_m\}$ روی دو حالتِ ρ و σ انجام دهیم. در این صورت حالت ρ توزیع احتمال $p_m = Tr(E_m\rho)$ و حالتِ σ توزیع احتمال $q_m = Tr(E_m\sigma)$ را تولید خواهند کرد. این دو تابع توزیع احتمال که در آزمایشگاه تولید می شوند یک فاصله با هم دارند که به صورت $D(p,q) := \frac{1}{2}\sum_m |p_m - q_m|$ بیان می شود. در این صورت حتما رابطه زیر برقرار است:

$$D(\rho, \sigma) \geq D(p, q). \quad (۶۴)$$

علاوه بر آن حتما یک اندازه گیری خاص مثل E^* وجود دارد که آمار $p_m^* = Tr(E^*\rho)$ و $q_m^* = Tr(E^*\sigma)$ را تولید می کند و برای آن رابطه زیر برقرار است.

$$D(\rho, \sigma) = D(p^*, q^*) \tag{۶۵}$$

دو رابطه اخیر را می توان به شکل زیر خلاصه کرد:

$$D(\rho, \sigma) = Max_{\{E_m\}} D(Tr(E_m\rho), Tr(E_m\sigma)). \tag{۶۶}$$

این رابطه آنالوگ رابطه (۱۰) در مورد فاصله بین توابع توزیع کلاسیک است. ممکن است برای بعضی اندازه گیری ها و نتایج بین ρ و σ تفاوتی حاصل نشود یا برای بعضی اندازه گیری های دیگر تفاوت کمی حاصل شود. آنچه که رابطه بالا بیان می کند که فاصله بین دو حالت برابر است با ماکزیمم فاصله بین نتایج. شکل (۵) این خاصیت را بهتر نشان می دهد.

■ اثبات: ماتریس های ρ و σ هر دو مثبت هستند ولی $\rho - \sigma$ الزاما مثبت نیست. برای محاسبه قدر مطلق آن ماتریس $\rho - \sigma$ را تجزیه طیفی می کنیم. این ماتریس در پایه ویژه مقدارهای خودش قطری است و می توانیم آن را مرتب کرده و به صورت زیر بنویسیم:

$$\rho - \sigma = diagonal(\lambda_1, \lambda_2, \cdots \lambda_k, -\mu_1, -\mu_2, \cdots \mu_l) \tag{۶۷}$$

که در آن ها

$$\lambda_i > 0, \qquad \mu_j \leq 0.$$

به عبارت دیگر داریم:

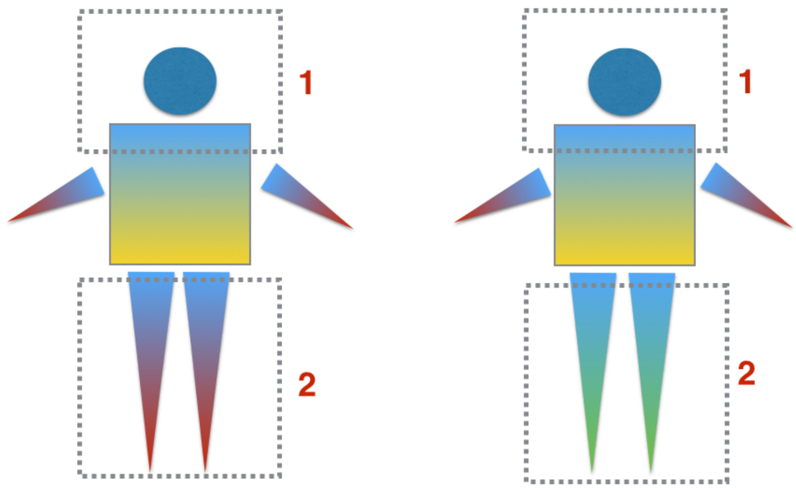
$$\rho - \sigma = \sum_{i=1}^{k} \lambda_i |a_i\rangle\langle a_i| - \sum_{j=1}^{l} \mu_j |b_j\rangle\langle b_j| \tag{۶۸}$$

و

$$\langle a_i|b_j\rangle = 0 \qquad \forall\, i,\ j. \tag{۶۹}$$

بنابراین می توان نوشت

$$\rho - \sigma = Q - S, \tag{۷۰}$$



شکل ۵: تفاوت بین دو حالت با هر نوع اندازه گیری آشکار نمی شود. بعضی از اندازه گیری ها مثل اندازه گیری ۱ این تفاوت را آشکار نمی کنند، بعضی تنها اندکی از این تفاوت را آشکار می کنند و بعضی نیز مثل ۲ تمام این تفاوت را آشکار می کنند.

که در آن

$$Q = diagonal(\lambda_1, \lambda_2, \cdots \lambda_k, 0, 0, \cdots 0) \equiv \sum_{i=1}^{k} \lambda_i |a_i\rangle\langle a_i|$$

و

$$S = diagonal(0, 0, \cdots 0, \mu_1, \mu_2, \cdots \mu_l) \equiv \sum_{j=1}^{l} \mu_j |b_j\rangle\langle b_j|$$

ماتریس های مثبت هستند و پایه های [8] آنها برهم عمود هستند. از آنجا که $tr(\rho - \sigma) = 0$ است نتیجه می گیریم که $tr(S) = tr(Q)$.

حال با توجه به (۶۷) و (۷۰) نتیجه می گیریم که

$$|\rho - \sigma| = S + Q$$

[8] Support

و

$$D(\rho,\sigma)=\frac{1}{2}tr|\rho-\sigma|=\frac{1}{2}tr(Q+S)=tr(Q).$$

حال به ازای هر اندازه گیری $\{E_m\}$ می توانیم بنویسیم:

$$\begin{aligned} D(p,q) &= \frac{1}{2}\sum_m |p_m-q_m| = \frac{1}{2}\sum_m |TrE_m(\rho-\sigma)| \\ &= \frac{1}{2}\sum_m |TrE_m(Q-S)| = \frac{1}{2}\sum_m |TrE_mQ-TrE_mS)| \end{aligned} \quad (۷۱)$$

با استفاده از نامساوی $|a+b|\leq |a|+|b|$ برای اعداد و سپس استفاده از اینکه $Tr(E_mQ)$ و $Tr(E_mS)$ مثبت هستند، به این نتیجه می رسیم که

$$\begin{aligned} D(p,q) &\leq \frac{1}{2}\sum_m |TrE_mQ|+\frac{1}{2}|TrE_mS| \\ &= \frac{1}{2}\sum_m TrE_mQ+\frac{1}{2}TrE_mS=\frac{1}{2}TrQ+\frac{1}{2}TrS=TrQ=D(\rho,\sigma). \end{aligned} \quad (۷۲)$$

در این روابط از این استفاده کرده ایم که اگر چه حاصلضرب دو عملگر مثبت یعنی E_m و Q الزاماً مثبت نیست، اما رد حاصلضرب آنها یعنی $Tr(E_mQ)$ همواره یک عدد مثبت است. هم چنین از این استفاده کرده ایم که $\sum_m E_m=I$.

■ **مثال حل شده:** اگر A و B دو عملگر مثبت باشند، می توانیم یکی از آنها مثلا B را تجزیه طیفی کنیم و بنویسیم: $B=\sum_i \beta_i|u_i\rangle\langle u_i|$ و از آنجا

$$Tr(AB)=\sum_i \beta_i\langle u_i|A|u_i\rangle\geq 0. \quad (۷۳)$$

از آنجا که

■ **تمرین:** برای ماتریس ها می توان یک ترتیب تعریف کرد به این معنا که

$$A\geq B \quad \text{if} \quad A-B\geq 0.$$

معنای مثبت بودن یک ماتریس هم این است که همه ویژه مقادیر آن نا منفی باشند. با یک مثال نشان دهید که رابطه مثلث برای ماتریس ها برقرار نیست، یعنی نمی توان نوشت:

$$|A+B|\leq |A|+|B|. \quad (۷۴)$$

راهنمایی: ماتریس ها ی زیر را امتحان کنید:

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

البته رابطه مثلث برای اندازه ماتریس ها برقرار است، یعنی همواره داریم

$$||A+B||_p \leq ||A||_p + ||B||_p. \tag{۷۵}$$

منظور از این اندازه عبارت زیر است:

$$\|A\|_p = (\sum_i |\lambda_i|^p)^{\frac{1}{p}} \tag{۷۶}$$

است که درآن λ_i ها ویژه مقدارهای A هستند.

به این ترتیب ثابت کردیم که در هر اندازه گیری ای که انجام می دهیم فاصله توزیع های آماری ای که تولید می کنیم کمتر از فاصله دو حالت کوانتومی است.

■ **قضیه:** الف: به ازای هر تصویرگر P داریم:

$$D(\rho,\sigma) \geq Tr(P(\rho-\sigma)) \tag{۷۷}$$

ب: حتما یک تصویرگر خاص مثل P^* وجود دارد به قسمی که

$$D(\rho,\sigma) = tr(P^*(\rho-\sigma)). \tag{۷۸}$$

اثبات الف: یک اندازه گیری تعمیم یافته به صورت زیر در نظر می گیریم:

$$\{E_0 = P,\ E_1 = I - P\} \tag{۷۹}$$

در این صورت توابع توزیعی احتمالی که در این اندازه گیری روی حالت های ρ و σ تولید خواهند شد عبارتند از:

$$(p_0,p_1) = \{TrP\rho, 1-TrP\rho\}, \qquad (q_0,q_1) = \{TrP\sigma, 1-TrP\sigma\} \tag{۸۰}$$

و در نتیجه

$$D(p,q)\equiv\frac{1}{2}(|p_0-q_0|+|p_1-q_1|)=\frac{1}{2}|Tr(P(\rho-\sigma))|+\frac{1}{2}|Tr(I-P)(\rho-\sigma)|=|Tr(P(\rho-\sigma))|. \quad (۸۱)$$

حال از قضیه قبلی استفاده می کنیم و این که به ازای هر عدد حقیقی مثل a همواره می توان نوشت $a\leq|a|$ و با ترکیب این دو به این نتیجه می رسیم که:

$$Tr(P(\rho-\sigma))\leq|Tr(P(\rho-\sigma))|=D(p,q)\leq D(\rho,\sigma). \quad (۸۲)$$

اثبات ب: قبلا دیدیم که همواره می توان نوشت:

$$\rho-\sigma=Q-S$$

که در آن Q و S دو عملگر مثبت با رد یکسان هستند. هم چنین یافتیم که $D(\rho,\sigma)=Tr(Q)$. تصویر گر P^* را چنان انتخاب می کنیم که تصویرگر روی زیرفضای پایه Q باشد. این تصویرگر دارای خاصیت زیر است: $P^*Q=Q,\quad P^*S=0$. می نویسیم:

$$D(\rho,\sigma)=Tr(Q)=Tr(P^*Q)=Tr(P^*Q-P^*S)=Tr(P^*(Q-S))=Tr(P^*(\rho-\sigma)). \quad (۸۳)$$

■ تمرین: ماتریس های زیر را در نظر بگیرید:

$$\rho=\frac{1}{2}(I+\mathbf{r}\cdot\sigma),\qquad \sigma=\frac{1}{2}(I+\mathbf{s}\cdot\sigma). \quad (۸۴)$$

الف: ماتریس های $Q,\ S$ را پیدا کنید و نشان دهید که واقعا $D(\rho,\sigma)=tr(Q)$.

ب: می دانیم که ρ و σ هر دو ماتریس های مثبت هستند. با این وجود چه نیازی هست که $\rho-\sigma$ را به صورت $Q-S$ بنویسیم که همین خاصیت ها را دارند؟

■ تمرین: حالت های زیر را در نظر بگیرید:

$$\rho=|\phi^+\rangle\langle\phi^+|,\qquad \sigma=|\psi^+\rangle\langle\psi^+|, \quad (۸۵)$$

که در آن $|\phi^+\rangle=\frac{1}{\sqrt{2}}(|00\rangle+|11\rangle)$ و $|\psi^+\rangle=\frac{1}{\sqrt{2}}(|01\rangle+|10\rangle)$. ماتریس های $Q,\ S$ را پیدا کنید و فاصله دو حالت بالا را

پیدا کنید.

■ تمرین: حالت $\rho=\frac{1}{2}I$ و حالت $\sigma=|+\rangle\langle+|$ را در نظر بگیرید. اگر روی این دو حالت اندازه گیری σ_x انجام دهیم هیچ تفاوتی حس نمی کنیم. ولی اگر اندازه گیری σ_x انجام دهیم تفاوت زیادی حس خواهیم کرد.

الف: تفاوت تابع های توزیع ناشی از اندازه گیری σ_x را روی این دو حالت بدست آورید.

ب: تفاوت تابع های توزیع ناشی از اندازه گیری σ_n را روی این دو حالت بدست آورید که در آن n یک حالت دلخواه است. حساب کنید که در کدام جهت می بایست اندازه گیری کنیم تا بیشترین تفاوت بین توابع توزیع ناشی از اندازه گیری بدست آید.

پ: فاصله بین دو حالت ρ و σ را از رابطه (۵۹) بدست آورید و صحت رابطه (۶۴) را تحقیق کنید.

■ **تمرین** : حالت $\rho=|\psi\rangle\langle\psi|$ و حالت $\sigma=|\phi\rangle\langle\phi|$ را در نظر بگیرید که در آن

$$|\psi\rangle=\frac{1}{\sqrt{2}}(|0,0\rangle+|1,1\rangle),\qquad |\phi\rangle=\frac{1}{\sqrt{2}}(|0,0\rangle-|1,1\rangle). \tag{۸۶}$$

نشان دهید که با هیچ نوع اندازه گیری موضعی این دو حالت را نمی توان از هم تشخیص داد. فاصله بین این دو حالت را حساب کنید.

■ **نامساوی مثلث:** با استفاده از رابطه ۶۴ می توان نامساوی مثلث را برای فاصله کوانتومی بین دو حالت ثابت کرد. این نامساوی بیان می کند که برای هر سه حالت کوانتومی رابطه زیر برقرار است:

$$D(\rho,\sigma)\le D(\rho,\nu)+D(\nu,\sigma). \tag{۸۷}$$

برای اثبات این رابطه می نویسیم:

$$D(\rho,\sigma)=tr(P^*(\rho-\sigma)), \tag{۸۸}$$

که در آن P^* تصویرگری است که نامساوی ۶۴ را تبدیل به تساوی می کند. حال می نویسیم:

$$\begin{aligned} D(\rho,\sigma) &= tr(P^*(\rho-\nu)+P^*(\nu-\sigma)) \\ &= tr(P^*(\rho-\nu))+tr(P^*(\nu-\sigma))\le D(\rho,\nu)+D(\nu,\sigma). \end{aligned} \tag{۸۹}$$

که در آن دوباره از نامساوی ۶۴ استفاده کرده ایم. به این ترتیب نامساوی مثلث ثابت می شود.

■ تمرین: ثابت کنید که در هر بعدی همواره رابطه زیر برقرار است:

$$0 \leq D(\rho,\sigma) \leq 1. \tag{۹۰}$$

حل: این که فاصله از صفر بیشتر یا مساوی است واضح است. طرف دوم نامساوی را به شکل زیر ثابت می کنیم:

$$D(\rho,\sigma) = Tr(P^*(\rho-\sigma)) = Tr(P^*\rho) - Tr(P^*\sigma) \leq Tr(P^*\rho) \leq Tr(\rho) = 1. \tag{۹۱}$$

حال به بیان و اثبات خاصیت مهمی از فاصله کوانتومی می پردازیم. این خاصیت بیان می کند که فاصله حالت های کوانتومی در اثر هر عمل کوانتومی (هر نگاشت مثبت و رد نگه دار) کم می شود. به عبارت بهتر تحت اعمال کوانتومی حالت ها به هم شبیه تر شده و فاصله آنها از یکدیگر کمتر می شود. این رابطه به این صورت بیان می شود:

$$D(\Lambda(\rho),\Lambda(\sigma)) \leq D(\rho,\sigma), \tag{۹۲}$$

که در آن Λ یک نگاشت مثبت و رد نگه دار است. برای اثبات آن از آنچه که در اثبات روابط قبلی یادگرفته ایم استفاده می کنیم و P^* را تصویرگری می گیریم که فاصله $D(\Lambda(\rho),\Lambda(\sigma)$ را تعریف می کند. سپس می نویسیم:

$$\begin{aligned} D\left[\Lambda(\rho),\Lambda(\sigma))\right] &= tr\left[P^*(\Lambda(\rho)-\Lambda(\sigma))\right] \\ &= tr\left[P^*\Lambda(\rho-\sigma)\right] = tr\left[P^*\Lambda(Q-S)\right] = tr\left[P^*(\Lambda(Q)-\Lambda(S))\right]. \end{aligned} \tag{۹۳}$$

حال از این موضوع استفاده می کنیم که عبارت $Tr(P^*\Lambda(S))$ به دلیل مثبت بودن P^* و $\Lambda(S)$ مثبت است و می نویسیم:

$$D\left[\Lambda(\rho),\Lambda(\sigma))\right] \leq tr\left[P^*\Lambda(Q)\right] \leq tr\left[\Lambda(Q)\right] = tr(Q) = D(\rho,\sigma). \tag{۹۴}$$

به این ترتیب قضیه ثابت می شود.

■ **مثال:**

به عنوان یک مثال Λ را نگاشتی بگیرید که روی یک قسمت از سیستم دو قسمتی رد می گیرد، یعنی

$$\Lambda(\rho^{AB}) = tr_B(\rho^{AB}) = \rho^A. \tag{۹۵}$$

در این صورت نتیجه این قضیه این است که:

$$D(\rho^A, \sigma^A) \leq D(\rho^{AB}, \sigma^{AB}). \qquad (۹۶)$$

■ تمرین: دو حالت مختلف بل را در نظر بگیرید. فاصله آنها را قبل و بعد از عبور از کانال بالا حساب کنید.

حال به یکی از مهم ترین خاصیت های فاصله یعنی خاصیت تقعر قوی [۹] می پردازیم که مبنای خیلی از خاصیت های دیگر است. این خاصیت درست مشابه حالت کلاسیک است. فرض کنید که $\{\lambda_i\}$ و $\{\mu_i\}$ دو مجموعه اعداد مثبت باشند که مجموع هرکدام از آنها به تنهایی برابر با یک باشد. به بیان دیگر این دو مجموعه را می توان به عنوان دو تابع توزیع گسسته احتمال در نظر گرفت. دراین صورت شرط تحدب قوی بیان می کند که رابطه زیر بین فاصله ها برقرار است:

$$D(\sum_i \lambda_i\rho_i, \sum_i \mu_i\sigma_i) \leq D(\lambda, \mu) + \sum_i \lambda_i D(\rho_i, \sigma_i), \qquad (۹۷)$$

که در آن منظور از $D(\lambda, \mu)$ فاصله بین دو تابع توزیع کلاسیک است که در بخش های قبلی تعریف شده است. برای اثبات این رابطه به ترتیب زیر عمل می کنیم.

$$\begin{aligned}
&D(\sum_i \lambda_i\rho_i, \sum_i \mu_i\sigma_i) = tr\left[P^*\left(\sum_i \lambda_i\rho_i - \sum_i \mu_i\sigma_i\right)\right] \\
&= tr\left[P^*\sum_i(\lambda_i - \mu_i)\rho_i + P^*\sum_i \mu_i(\rho_i - \sigma_i)\right] = tr\left[P^*\sum_i(\lambda_i - \mu_i)\rho_i\right] + \sum_i \mu_i tr\left[P^*(\rho_i - \sigma_i)\right] \\
&\leq tr\left[P^*\sum_i(\lambda_i - \mu_i)\rho_i\right] + \sum_i \mu_i D(\rho_i, \sigma_i). \qquad (۹۸)
\end{aligned}$$

حال باید ثابت کنیم که جمله اول طرف راست از فاصله کلاسیک بین دو تابع توزیع کلاسیک λ و μ کمتر است. برای این کار این جمله را به این صورت می نویسیم

$$tr\left[P^*\sum_i(\lambda_i - \mu_i)\rho_i\right] = tr\left[P^*(\rho - \rho')\right] \qquad (۹۹)$$

که در آن $\rho = \sum_i \lambda_i\rho_i$ و $\rho' = \sum_i \mu_i\rho_i$ دو تابع توزیع جدید هستند. اما بنابرقضایایی که قبلا ثابت کرده ایم داریم

$$tr\left[P^*(\rho - \rho')\right] \leq D(\rho, \rho') = \frac{1}{2}tr|\rho - \rho'| = \frac{1}{2}tr|\sum_i(\lambda_i - \mu_i)\rho_i|. \qquad (۱۰۰)$$

[۹] Strong Concavity

در مرحله آخر از نامساوی مثلث برای رد استفاده می کنیم که بر مبنای آن

$$tr|A+B| \leq tr|A| + tr|B|, \tag{۱۰۱}$$

و در نتیجه خواهیم داشت

$$tr\left[P^*(\rho-\rho')\right] \leq \frac{1}{2}\sum_i tr|(\lambda_i-\mu_i)\rho_i| = \frac{1}{2}\sum_i |\lambda_i-\mu_i| = D(\lambda,\mu). \tag{۱۰۲}$$

با ترکیب رابطه ۹۸ و ۱۰۲ به رابطه ۹۷ می رسیم.

■ **تمرین** : نشان دهید که فاصله دو حالت کوانتومی تحت تبدیلات یکانی ناورداست.

۷ تشابه حالت های کوانتومی

تشابه بین دو حالت کوانتومی نیز می بایست چنان تعریف شود که یک معنای عملیاتی مشخص داشته باشد. در واقع کاری که در آزمایشگاه انجام می دهیم آن است که روی یک حالت کوانتومی آزمایش و اندازه گیری انجام می دهیم و این اندازه گیری یک تابع توزیع احتمال تولید می کند، به این معنا که نتایج مشخص با احتمالات مشخص بدست می آید. مثلا فرض کنید که یک اندازه گیری $POVM$ با المان های $\{E_m\}$ روی حالت های کوانتومی ρ و σ انجام می دهیم. احتمال بدست آمدن نتیجه m از حالت ρ برابر است با $p_m := tr(\rho E_m)$ و احتمال بدست آمدن نتیجه m از حالت σ برابر است با $q_m := tr(\sigma E_m)$. میزان نزدیکی این دو تابع احتمال کلاسیک معیاری است از نزدیکی دو حالت کوانتومی ρ و σ . ولی در این جا یک نکته مهم مطرح است و آن اینکه ممکن است که دو حالت ρ و σ با یک اندازه گیری به خوبی تمیز داده نشوند. به عنوان مثال دو حالت زیر را در نظر بگیرید:

$$\rho = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|), \qquad \sigma = |+\rangle\langle +|, \tag{۱۰۳}$$

که در آن $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. واضح است که اندازه گیری در پایه z روی هر دو حالت یک نتیجه بدست می دهد. بنابراین اگر می خواهیم تفاوت این دو حالت را معین کنیم می بایست یا روی همه اندازه گیری ها متوسط بگیریم یا اینکه آن اندازه گیری را انتخاب کنیم که

کمترین شباهت را یا بیشترین فاصله را بین توابع توزیع احتمال p و q تولید می کند. خواهیم دید که تعریفی که از شباهت برای دو حالت کوانتومی ارائه می کنیم با این معیار دوم مطابق است. برای تعریف شباهت می توانیم از یک تعریف آزمایشی شروع کنیم و سعی کنیم که آن را اصلاح کرده و به یک تعریف نهایی خوب برسیم. در مرحله اول به نظر می رسد که تعریف زیر مناسب باشد:

$$F_0(\rho,\sigma) = tr(\rho\sigma)\ ? \tag{۱۰۴}$$

این تعریف همان ضرب داخلی بین دو ماتریس هرمیتی است و به نظر می رسد که تعریف مناسبی باشد. فقط یک عیب دارد و آن اینکه انتظار داریم که شباهت یک حالت با خودش برابر با یک باشد و حال آنکه این تعریف چنین خاصیتی ندارد زیرا $tr(\rho\rho) \leq 1$ و درنتیجه $F_0(\rho,\rho) \neq 1$. در واقع تنها برای حالت های خالص است که این شرط برآورده می شود. اگر می خواهیم این شرط که شرط مهمی است برآورده شود می توانیم تعریف زیر را امتحان کنیم:

$$F_1(\rho,\sigma) = tr(\sqrt{\rho\sigma})\ ? \tag{۱۰۵}$$

این تعریف اشکال بالا را برطرف می کند ولی عیب اش این است که نسبت به ρ و σ متقارن نیست، یعنی $F_1(\rho,\sigma) \neq F_1(\sigma,\rho)$. برای رفع این اشکال می توانیم تشابه را به شکل زیر تعریف کنیم:

$$F_2(\rho,\sigma) = tr(\sqrt{\rho}\sqrt{\sigma})\ ? \tag{۱۰۶}$$

این تعریف اشکالات هر دو تعریف قبلی را برطرف می کند زیرا هم متقارن است و هم تشابه هر حالت با خودش برابر با یک می شود. اما یک اشکال دیگر دارد زیرا اگر چه $\sqrt{\rho}$ و $\sqrt{\sigma}$ هر دو هرمیتی و مثبت هستند، حاصلضرب آن ها الزاما هرمیتی نیست و ویژه مقدارهای یک ماتریس غیرهرمیتی نیز الزاما حقیقی نیستند. بنابراین تعریف نهایی را به صورت زیر اختیار می کنیم.

$$F(\rho,\sigma) = tr(|\sqrt{\rho}\sqrt{\sigma}|) \tag{۱۰۷}$$

با توجه به اینکه به ازای هر ماتریس A

$$|A| = \sqrt{AA^\dagger} \tag{۱۰۸}$$

بدست می آوریم:

$$F(\rho,\sigma) = tr(\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}). \tag{۱۰۹}$$

این تعریف بوضوح دارای این خاصیت است که $F(\rho,\rho)=1$ است. ولی واضح نیست که نسبت به جابجایی ρ و σ متقارن باشد. اما با توجه به اینکه $A := \sqrt{\rho}\sqrt{\sigma}$ ، و هم چنین این که

$$tr(\sqrt{AA^\dagger}) = tr(|A|) = tr(|A^\dagger|) \tag{۱۱۰}$$

تقارن F نسبت به ρ و σ واضح می شود.

■ **تمرین** : ثابت کنید که $tr(|A|) = tr(|A^\dagger|)$.

■ **تمرین** : دو حالت دلخواه را در کره بلوخ در نظر بگیرید که با بردارهای r و s مشخص می شوند. تشابه این دو حالت را محاسبه کنید.

راهنمایی: از رابطه $(\lambda_1+\lambda_2)^2 = \lambda_1^2+\lambda_2^2+2\lambda_1\lambda_2$ استفاده کنید و رابطه زیر را برای ماتریس های دو بعدی نشان دهید: $TrA^2 = Tr(A^2) + 2\det(A)$ و سپس تشابه دو حالت یاد شده را حساب کنید.

هرگاه که ρ و σ با هم جابجا شوند آنگاه مقدار تشابه دو حالت کوانتومی دقیقا با تشابه کلاسیک بین ویژه مقدارهای آن ها یکی می شود. در واقع تحت این شرایط می توانیم هر دو حالت کوانتومی را در یک پایه قطری کنیم:

$$\rho = \sum_i r_i|i\rangle\langle i|, \qquad \sigma = \sum_i s_i|i\rangle\langle i|. \tag{۱۱۱}$$

در نتیجه خواهیم داشت:

$$\begin{aligned} F(\rho,\sigma) &= tr\sqrt{\sum_i r_i s_i|i\rangle\langle i|} \\ &= tr\left(\sum_i \sqrt{r_i s_i}|i\rangle\langle i|\right) \end{aligned}$$

$$= \sum_i \sqrt{r_i s_i} = F(r,s). \qquad (۱۱۲)$$

که در آن r و s دو تابع توزیع احتمال کلاسیک هستند. یکی دیگر از مواقعی که می توان تشابه دو حالت کوانتومی را به راحتی حساب کرد وقتی است که یکی از حالت ها خالص باشد. در این صورت با توجه به اینکه $|\psi\rangle\langle\psi| = (|\psi\rangle\langle\psi|)^{\frac{1}{2}}$ خواهیم داشت:

$$F(|\psi\rangle,\rho) = tr\sqrt{\langle\psi|\rho|\psi\rangle\langle\psi|\psi\rangle} = \sqrt{\langle\psi|\rho|\psi\rangle}. \qquad (۱۱۳)$$

۱.۷ قضیه اولمان و اثبات آن

یکی از مهمترین قضایایی که در باره تشابه دو حالت کوانتومی وجود دارد قضیه یولمان [۱۰] است. این قضیه در اثبات بسیاری دیگر از خواص تشابه کوانتومی به کار می رود و به شکل زیر بیان می شود.

- **قضیه یولمان :** فرض کنید که ρ و σ دو حالت ازیک سیستم کوانتومی B باشند. نمونه دیگری از سیستم کوانتومی A که آن را با B نمایش می دهیم درنظرمی گیریم. دراین صورت هرگاه $|\psi\rangle$ و ϕ دو خالص سازی از ρ و σ روی AB باشند آنگاه

$$F(\rho,\sigma) = max_{|\psi\rangle,|\phi\rangle} \mid \langle\psi|\phi\rangle \mid, \qquad (۱۱۴)$$

که درآن ماکزیمم گیری روی تمام خالص سازی ها صورت می گیرد.

قبل از اینکه به اثبات این قضیه در حالت کلی بپردازیم بهتر است که یک مثال ساده را حل کنیم.

- تمرین: دو حالت $\rho_A = \frac{1}{2}(1 + \mathbf{r}\cdot\sigma)$ و $\sigma_A = \frac{1}{2}(1 + \mathbf{s}\cdot\sigma)$ را در نظر بگیرید.

الف: یک بار فاصله آنها را مستقیما از تعریف بدست آورید.

[۱۰] Uhlmann

ب: حال دو خالص سازی زیر را در نظر بگیرید:

$$\rho_A = Tr_B(|\psi\rangle_{AB}\langle\psi|) \quad , \quad \sigma_A = Tr_B(|\phi\rangle_{AB}\langle\phi|) \tag{۱۱۵}$$

که در آن

$$\begin{aligned} |\psi\rangle_{AB} &= \sqrt{\frac{1+r}{2}}|\hat{\mathbf{r}}\rangle|0\rangle + \sqrt{\frac{1-r}{2}}|-\hat{\mathbf{r}}\rangle|1\rangle \\ |\phi\rangle_{AB} &= \sqrt{\frac{1+s}{2}}|\hat{\mathbf{s}}\rangle|\alpha\rangle + \sqrt{\frac{1-s}{2}}|-\hat{\mathbf{s}}\rangle|\alpha^\perp\rangle \end{aligned} \tag{۱۱۶}$$

که در آن $|\alpha\rangle = a|0\rangle + b|1\rangle$ یک حالت دلخواه است. نخست خود را قانع کنید که این دو حالت خالص شده ماتریس های چگالی اولیه هستند. سپس نشان دهید که اگر بیشترین مقدار همپوشانی این دو حالت را حساب کنید به شباهت دو ماتریس چگالی می رسید.

پس از این مثال به اثبات قضیه اولمان می پردازیم. برای اثبات این قضیه نخست به یک لم احتیاج داریم:

- **لم یک:** هرگاه A یک ماتریس دلخواه باشد و U یک ماتریس یکانی باشد، آنگاه

$$|tr(AU)| \leq tr|A|.$$

- **اثبات:** نخست توجه می کنیم که

$$|AU| = |A|,$$

(این تساوی ناشی از تعریف قدر مطلق یک ماتریس است: $|A| := \sqrt{A^\dagger A}$.) سپس توجه می کنیم که

$$|TrA| \leq Tr|A|,$$

(این نامساوی همان نامساوی مثلث است که در پایه ای که A در آن قطری شده باشد، بدست می آید.)

با ترکیب این دو رابطه می نویسیم:

$$|TrAU| \leq Tr|AU| = Tr|A|. \tag{۱۱۷}$$

- به طور کلی نامساوی زیر را داریم:

$$\begin{aligned} &|Tr(A)| \leq Tr|A| = Tr|AU| \\ &|Tr(AU)| \leq Tr|A|. \end{aligned} \tag{۱۱۸}$$

■ **لم دو:** در این لم یک مسئله کلی را طرح می کنیم که در جاهای دیگر هم کاربرد دارد. می خواهیم یک طرح کلی از خالص سازی یک ماتریس چگالی ارایه دهیم و آن را به طور ساده تری بفهمیم. هرگاه یک حالت مثل ρ_A از سیستم A داشته باشیم، حالت $|\psi\rangle_{AB}$ از یک سیستم بزرگ تر را یک خالص سازی از ρ_A می خوانیم هرگاه داشته باشیم $tr_B(|\psi\rangle_{AB}\langle\psi|) = \rho_A$. ساده ترین روش برای خالص سازی آن است که نخست ρ_A را به صورت زیر تجزیه طیفی کنیم

$$\rho_A = \sum_i \lambda_i |a_i\rangle\langle a_i| \qquad (۱۱۹)$$

و سپس با تعریف یک مجموعه از بردارهای متعامد و یکه $|b_i\rangle$ از سیستم B قرار دهیم

$$|\psi_\rho\rangle_{AB} := \sum_i \sqrt{\lambda_i}|a_i\rangle_A|b_i\rangle_B. \qquad (۱۲۰)$$

اما تمام خالص سازی های ممکن با عمل ماتریس های یکانی روی سیستم B به هم مرتبط هستند، دلیل آن هم این است که برای سیستم B مجموعه بردارهای متعامدِ متفاوتی می توانیم اختیار کنیم. حال یک حالت مرجع مثل حالت زیر را در نظر بگیرید:

$$|\phi^+\rangle_{AB} = \sum_i |i\rangle_A \otimes |i\rangle_B. \qquad (۱۲۱)$$

که در آن $\{|i\rangle\}$ ها پایه ای برای هر کدام از فضاها هستند. در واقع این حالت یک حالت بل تعمیم یافته است. (این حالت را نزمالیزه نگرفته ایم.). می توان پایه های فوق را با ماتریس های یکانی $U\otimes V$ طوری چرخاند که حالت زیر درست شود:

$$(U\otimes V)|\phi^+\rangle = \sum_i |a_i\rangle \otimes |b_i\rangle \qquad (۱۲۲)$$

سپس می توان با اثر عملگر $\sqrt{\rho}\otimes I$ روی دو طرف به حالت خالص شده رسید یعنی

$$|\psi_\rho\rangle = (\sqrt{\rho}U\otimes V)|\phi^+\rangle. \qquad (۱۲۳)$$

در این خالص سازی با تغییر ماتریس U حالت آمیخته اولیه عوض نمی شود. بنابراین با تغییر V می توان به خالص سازی های متفاوت از یک حالت آمیخته رسید. در واقع هر حالت آمیخته ای با یک ماتریس V متناظر می شود.

■ حال می توانیم به اثبات قضیه یولمان بپردازیم. با خالص سازی دو حالت ِ ρ و σ می توانیم بنویسیم:

$$|\psi_\rho\rangle \quad = \quad (\sqrt{\rho}U\otimes V)|\phi^+\rangle$$

$$|\psi_\sigma\rangle \quad = \quad (\sqrt{\sigma}U' \otimes V')|\phi^+\rangle \tag{۱۲۴}$$

و در نتیجه

$$\langle\psi_\rho|\psi_\sigma\rangle = \langle\phi^+|(U^\dagger\sqrt{\rho}\sqrt{\sigma}U') \otimes (V^\dagger V')|\phi^+\rangle \tag{۱۲۵}$$

و یا با نمادهای خلاصه تر به شکل :

$$\langle\psi_\rho|\psi_\sigma\rangle = \langle\phi^+|X \otimes Y|\phi^+\rangle \tag{۱۲۶}$$

که در آن

$$X = U^\dagger\sqrt{\rho}\sqrt{\sigma}U' \quad , \quad Y = V^\dagger V'.$$

حال می توانیم طرف راست را به صورت ساده تری بنویسیم. برای این کار از تمرین زیر استفاده می کنیم:

■ تمرین: درستی اتحاد زیر را نشان دهید

$$\langle\phi^+|X \otimes Y|\phi^+\rangle = tr(XY^T). \tag{۱۲۷}$$

حال با استفاده از این تمرین می نویسیم

$$|\langle\psi_\rho|\psi_\sigma\rangle| = |tr(XY^T)| = |tr(U^\dagger\sqrt{\rho}\sqrt{\sigma}U')V'^TV^*)| = |tr(\sqrt{\rho}\sqrt{\sigma}W)| \tag{۱۲۸}$$

که در آن $W = U'V'^TV^*U^\dagger$ ماتریس یکانی است. با استفاده از لمی که در ابتدا ثابت کردیم خواهیم داشت:

$$|\langle\psi_\rho|\psi_\sigma\rangle| = |tr(\sqrt{\rho}\sqrt{\sigma}W)| \leq tr(|\sqrt{\rho}\sqrt{\sigma}|) = F(\rho, \sigma). \tag{۱۲۹}$$

حال باید نشان دهیم که تساوی نیز برقرار می شود، یعنی یک خالص سازی وجود دارد که تساوی را برقرار می کند. برای یافتن این خالص سازی به ترتیب زیر عمل می کنیم. فرض کنید که $\sqrt{\rho}\sqrt{\sigma}$ دارای تجزیه قطبی زیر است:

$$\sqrt{\rho}\sqrt{\sigma} = |\sqrt{\rho}\sqrt{\sigma}|\Omega.$$

در این صورت کافی است که ماتریس یکانی بالا را به ترتیبی اختیار کنیم که داشته باشیم:

$$\sqrt{\rho}\sqrt{\sigma}W = |\sqrt{\rho}\sqrt{\sigma}|.$$

معنای این حرف این است که قرار دهیم $W = \Omega^\dagger$ و با توجه به فرم W که در انتخاب ماتریس های U و U' آزادی داریم همواره می توان این کار را انجام داد. اثبات قضیه در این جا کامل می شود.

۲.۷ تشابه حالت های کوانتومی و اثر اندازه گیری

دو حالت کوانتومی مثل

$$\rho = |+\rangle\langle+|, \qquad \sigma = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|) \tag{۱۳۰}$$

را در نظر بگیرید. هرگاه روی این دو حالت در پایه Z یعنی $|0\rangle\langle 0|, |1\rangle\langle 1|$ اندازه گیری کنیم به هیچ وجه متوجه اختلاف این دو حالت نمی شویم زیرا در هر دو حالت احتمال بدست آوردن 0 و 1 برابر با $1/2$ است. بنابراین با این نوع اندازه گیری ممکن است که به غلط ادعا کنیم که این دو حالت یکسان هستند یا شباهت بسیار زیادی به هم دارند و حال انکه این دو حالت با هم یکی نیستند. برای آنکه شباهت واقعی این دو حالت را بدست آوریم می بایست اندازه گیری خود را تغییر دهیم. در وقاع اگر این دو حالت را در پایه X یعنی $|+\rangle\langle+|, |-\rangle\langle-|$ اندازه گیری کنیم آنگاه نتایج زیر بدست می آیند:

$$P_\rho(+) = 1, \qquad P_\rho(-) = 0, \qquad P_\sigma(+) = P_\sigma(-) = \frac{1}{2}. \tag{۱۳۱}$$

بنابراین شباهت نتایج اندازه گیری در این جا برابر است با:

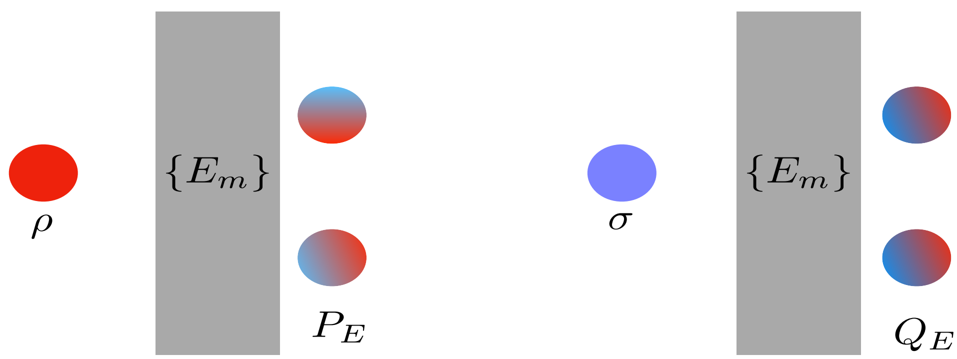
$$F(P_\rho, P_\sigma) = \sum_{i=+,-} P_\rho(i) P_\sigma(i) = \frac{1}{\sqrt{2}}. \tag{۱۳۲}$$

می توانیم اندازه گیری های دیگری را امتحان کنیم. منطقی است که شباهت این دو حالت را بر اساس مقدار کمینه ای که در تمام اندازه گیری ها بدست می آوریم تعریف کنیم و این درست همان چیزی است که رابطه تشابه کوانتومی به ما می گوید. در حقیقت قضیه زیر را می توانیم ثابت کنیم:

■ **قضیه:** هرگاه $E = \{E_m\}$ یک اندازه گیری $POVM$ باشد آنگاه

$$F(\rho, \sigma) \leq F(P_E, Q_E) \tag{۱۳۳}$$

که در آن $P_E = \{q_m := tr(\rho E_m)\}$ و $Q_E = \{q_m := tr(\sigma E_m)\}$ دو تابع توزیع کلاسیک ناشی از اندازه گیری روی این دو حالت هستند.



شکل ۶: آمار کلاسیکی که از دو حالت مختلف کوانتومی ایجاد می شود همواره تشابه بیشتری دارند تا خود آن دو حالت. به عنوان مثال آماری که از اندازه گیری S_z روی دو حالتِ $|+\rangle$ و $|-\rangle$ ایجاد می شود، کاملا یکسان هستند و حال آنکه حالت های کوانتومی مربوطه هیچ شباهتی به هم ندارند.

دلیل این نامساوی هم معلوم است چرا که آماری که یک اندازه گیری کوانتومی تولید می کند همواره همراه با یک عدم قطعیت ذاتی ناشی از مکانیک کوانتومی است که باعث می شود نتوانیم حالت ها را به درستی از هم تشخیص دهیم. شکل (۶) این موضوع را به صورت شماتیک نشان می دهد.

■ **اثبات:** نخست توجه می کنیم که $F(\rho,\sigma)=tr|A|$ که در آن $A=\sqrt{\rho}\sqrt{\sigma}$. سپس از این استفاده می کنیم که هر ماتریسی مثل A یک تجزیه قطبی دارد به این معنا که

$$A=|A|U^\dagger, \longrightarrow |A|=AU, \tag{۱۳۴}$$

که در آن U یک ماتریس یکانی است. بنابراین می توانیم بنویسیم:

$$F(\rho,\sigma)=tr(\sqrt{\rho}\sqrt{\sigma}U)=\sum_m tr(\sqrt{\rho}\sqrt{E_m}\sqrt{E_m}\sqrt{\sigma}U). \tag{۱۳۵}$$

در اینجا از این استفاده کرده ایم که $\sum_m E_m = I$. حال از قضیه کوشی_شوارتز استفاده می کنیم که بر مبنای آن برای دو ماتریس دلخواه

$$tr(A^\dagger B) \leq \sqrt{tr(AA^\dagger}\sqrt{tr(BB^\dagger}. \qquad (۱۳۶)$$

هرگاه در این رابطه قرار دهیم $A = \sqrt{\rho}\sqrt{E_m}$ و $B = \sqrt{E_m}\sqrt{\sigma}U$ ، و آن را با رابطه قبلی ترکیب کنیم بدست می آوریم

$$F(\rho,\sigma) \leq \sum_m \sqrt{tr(E_m\rho)tr(E_m\sigma)}. \qquad (۱۳۷)$$

به این ترتیب ثابت کرده ایم که تشابه دو حالت کوانتومی به ترتیبی که تعریف کرده ایم از تشابه احتمالات کلاسیکی که ناشی از هر نوع اندازه گیری است کمتر است. این که آیا واقعا اندازه گیری ای وجود دارد که تشابه احتمالات کلاسیکی اش دقیقا برابر با تشابه کوانتومی بین حالات باشد سوالی است که پاسخ آن مثبت است اگرچه اثبات آن را ما بیان نخواهیم کرد. بازهم برای اثبات این قضیه خواننده می تواند به کتاب نیلسون و چوانگ مراجعه کند. این قضیه بیان می کند که حتما یک اندازه گیری $POVM$ مثل $\{E_m^*\}$ وجود دارد به قسمی که

$$F(\rho,\sigma) = \sum_m \sqrt{tr(E_m^*\rho)tr(E_m^*\sigma)}. \qquad (۱۳۸)$$

از این موضوع در آینده استفاده های زیادی خواهیم کرد.

۳.۷ تاثیر کانال های کوانتومی روی تشابه حالت ها

به همانگونه که ثابت کردیم فاصله حالت های کوانتومی تحت عبور از یک کانال کم می شود، اکنون می خواهیم ثابت کنیم که شباهت حالت های کوانتومی در عبور از یک کانال زیاد می شود یعنی حالت ها به هم شبیه تر می شوند. این مسئله در قضیه زیر بیان می شود.

- **قضیه: یکنوایی فیدلیته** در عبور از هر کانالی فیدلیته یا تشابه حالت های کوانتومی زیاد می شود یعنی به ازای هر نگاشت مثبت Λ :

$$F(\Lambda(\rho),\Lambda(\sigma)) \geq F(\rho,\sigma) \qquad (۱۳۹)$$

دقت کنید که Λ یک نگاشت مثبت است و نه الزاما نه یک نگاشت کاملا مثبت.

■ نخست ازیک لم استفاده می کنیم: فرض کنید که $|\psi\rangle_{AB}$ خالص شده یک ماتریس چگالی ρ_A باشد. خالص شده ی $\Lambda(\rho_A)$ چیست؟

برای پاسخ به این سوال توجه می کنیم که $E(\rho_A)$ را همواره می توان به این شکل نوشت:

$$E(\rho_A) = Tr_C\Big(U_{AC}(\rho_A\otimes|0\rangle_C\langle 0|)U^\dagger_{AC}\Big) \tag{۱۴۰}$$

اما می دانیم که

$$\rho_A = Tr_B(|\psi\rangle_{AB}\langle\psi|). \tag{۱۴۱}$$

با ترکیب این دو رابطه و استفاده از خاصیت های رد جزیی بدست می آوریم:

$$\begin{aligned}
E(\rho_A) &= Tr_C\Big(U_{AC}\big[\rho_A\otimes|0\rangle_C\langle 0|\big]U^\dagger_{AC}\Big)\\
&= Tr_C\Big(U_{AC}\big[Tr_B(|\psi\rangle_{AB}\langle\psi|)\otimes|0\rangle_C\langle 0|\big]U^\dagger_{AC}\Big)\\
&= Tr_{BC}\Big(|\Psi\rangle\langle\Psi|\Big)_{ABC},
\end{aligned} \tag{۱۴۲}$$

که در آن

$$|\Psi\rangle_{ABC} = U_{AC}\big(|\psi\rangle_{AB}\otimes|0\rangle_C\big). \tag{۱۴۳}$$

حال از این لم برای اثبات قضیه اصلی استفاده می کنیم. فرض کنید که $|\psi^*\rangle$ و $|\phi^*\rangle$ حالت های خالص شده ρ و σ هستند که بنابر قضیه یولمان در رابطه زیر صدق می کنند: یعنی

$$F(\rho,\sigma) = |\langle\psi^*|\phi^*\rangle|. \tag{۱۴۴}$$

بنابر لمی که ثابت کردیم خالص شده های $E(\rho)$ و $E(\sigma)$ به ترتیب عبارتند از:

$$|\Psi\rangle = U|\psi*\rangle|0\rangle \quad , \quad |\Phi\rangle = U|\phi^*\rangle|0\rangle. \tag{۱۴۵}$$

از قضیه اولمان استفاده می کنیم و قرار می دهیم

$$F(\Lambda(\rho),\Lambda(\sigma)) \geq |\langle\Psi|\Phi\rangle| = |\langle\psi^*\langle 0|U^\dagger U|\phi^*\rangle|0\rangle| = |\langle\psi^*|\phi^*\rangle| = F(\rho,\sigma). \tag{۱۴۶}$$

در قسمت اول این تساوی بازهم از قضیه یولمان استفاده کرده ایم .

■ تمرین: کانال واقطبش با پارامتر p را در نظر بگیرید. تشابه دو حالت را که با بردارهای بلوخ r و s مشخص می شوند قبل از ورود به کانال و بعد از خروج از کانال با هم مقایسه کنید.

■ **تمرین:** دو حالت خالص d بعدی $|\psi\rangle$ و $|\phi\rangle$ را در نظر بگیرید و آنها را از کانال واقطبش d بعدی عبور دهید. کانال واقطبش d بعدی به صورت زیر تعریف می شود:

$$\Lambda(\rho) = (1-p)\rho + p\frac{I}{d}. \qquad (۱۴۷)$$

الف: تشابه دو حالت بالا را قبل از ورود به کانال و بعد از خرو ج از آن با هم مقایسه کنید.

ب: فاصله دو حالت بالا را قبل از ورود به کانال و بعد از خرو ج از آن با هم مقایسه کنید.

۴.۷ خاصیت تحدب تشابه حالت ها

قبلا دیدیم که وقتی حالت های کوانتومی را با هم مخلوط می کنیم فاصله آنها از هم کمتر می شود. این خاصیت همان خاصیت تقعر قوی فاصله بود. در مورد تشابه نیز خاصیتی شبیه به این برقرار است به این معنا که وقتی که حالت های کوانتومی را باهم مخلوط می کنیم تشابه این حالت ها زیاد تر می شود. این خاصیت را اصطلاحا خاصیت تحدب قوی [۱۱] می گوییم که به شکل زیر بیان می شود:

■ **قضیه: خاصیت تحدب قوی برای فیدلیته**

$$F(\sum_i p_i\rho_i, \sum_i q_i\sigma_i) \geq \sum_i \sqrt{p_iq_i}F(\rho_i,\sigma_i). \qquad (۱۴۸)$$

■ **اثبات:** برای اثبات این خاصیت از قضیه یولمان استفاده می کنیم. فرض کنید که $|\psi_i^*\rangle$ و $|\phi_i^*\rangle$ به ترتیب آن نوع ویژه از حالت های خالصی باشند که مطابق با قضیه اولمان تساوی زیر را برقرار می کنند:

$$F(\rho_i,\sigma_i) = |\langle\psi_i^*|\phi_i^*\rangle|. \qquad (۱۴۹)$$

[۱۱] Fidelity of Convexity trongS

دقت کنید که همواره می توانیم $|\psi_i^*\rangle$ و $|\phi_i^*\rangle$ را طوری انتخاب کنیم که $\langle\psi_i^*|\phi_i^*\rangle$ یک عدد مثبت باشد. این کار را می توانیم با ضرب کردن این حالت های در فازهای مناسب انجام دهیم.

حال حالت های بزرگتری به شکل زیر در نظر می گیریم:

$$|\Psi\rangle := \sum_i \sqrt{p_i}|\psi_i^*\rangle_{AB}|i\rangle_C, \qquad |\Phi\rangle := \sum_i \sqrt{p_i}|\phi_i^*\rangle_{AB}|i\rangle_C. \tag{۱۵۰}$$

این حالت های خالص، حالت هایی هستند که ρ و σ را خالص می کنند یعنی اینکه

$$\rho_A \equiv \sum_i p_i\rho_i = tr_{BC}|\Psi\rangle\langle\Psi|, \qquad \sigma_A \equiv \sum_i q_i\sigma_i = tr_{BC}|\Phi\rangle\langle\Phi|. \tag{۱۵۱}$$

دقت کنید که در اینجا ρ_i و σ_i ماتریس چگالی مربوط به سیستم A هستند. حالت های $|\psi_i\rangle_{AB}$ و $|\phi_i\rangle_{AB}$ حالت های خالصی هستند که این ماتریس های چگالی را خالص می کنند. با افزودن یک سیستم C با پایه های متعامد $\{|i\rangle\}$ و رد گرفتن روی B و C به ماتریس های حالت ρ و σ می رسیم. متعامد بودن بردارهای $\{|i\rangle\}$ برای این مرحله ضروری است. حال از قضیه اولمان می دانیم که

$$\begin{aligned} F(\rho_A,\sigma_A) &\equiv F(\sum_i p_i\rho_i, \sum_i q_i\sigma_i) \geq |\langle\Psi|\Phi\rangle| \\ &= |\sum_i \sqrt{p_iq_i}\langle\psi *_i|\phi *_i\rangle| = \sum_i \sqrt{p_iq_i}\langle\psi *_i|\phi *_i\rangle = \sum_i \sqrt{p_iq_i}F(\rho_i,\sigma_i), \end{aligned} \tag{۱۵۲}$$

که در آن از از مثبت بودن $\langle\psi_i^*|\phi *_i\rangle$ استفاده کرده ایم و اینکه $F(\rho_i,\sigma_i) = |\langle\psi_i^*|\phi_i^*\rangle|$

در اینجا توجه به یک نکته مهم است و آن اینکه بردارهای حالتِ $|\Psi\rangle$ و $|\Phi\rangle$ آن خالص سازی های ویژه برای حالت های ρ و σ نیستند که ضرب داخلی شان تشابه حالت های ρ و σ را بدست بدهد.

۸ کیفیت یک کانال کوانتومی

یک کانال کوانتومی خوب آن است که حالت های کوانتومی را بدون هیچ گونه تغییری به مقصد ارسال کند. این کانال کوانتومی خوب می تواند یک فیبر نوری باشد که از آن برای ارسال حالت های فوتون ها استفاده می شود یا اینکه یک حافظه کوانتومی که حالت های کوانتومی را در خود

نگاه می دارد. بنابراین معیاری برای کیفیت یک کانال کوانتومی آن است که ببینیم فیدلیته حالت ورودی و حالت خروجی آن چقدراست. بنابراین نخست می بایست تشابه یک حالت ورودی و خروجی را به صورت زیر تعریف کنیم:

$$Q(\Lambda, |\psi\rangle) := F(|\psi\rangle, \Lambda(|\psi\rangle)). \quad (۱۵۳)$$

اما طبیعی است که این عبارت هم بستگی به حالت ورودی و هم بستگی به نوع کانال دارد. ممکن است که یک کانال بعضی از حالت های کوانتومی را سالم منتقل کند ولی بعضی دیگر را بکلی خراب کند. بنابراین معیاری برای کیفیت کانال کوانتومی آن است که یا روی ورودی های مختلف متوسط بگیریم یا اینکه بدترین حالت را در نظر بگیریم یعنی حالت هایی که کمترین تشابه بین ورودی و خروجی را تولید می کنند. این تعریف دوم است که بعنوان معیاری برای کیفیت یک کانال به کار می رود:

$$Q(\Lambda) := min_{|\psi\rangle} F(|\psi\rangle, \Lambda(|\psi\rangle)). \quad (۱۵۴)$$

- **مثال: کیفیت کانال واقطبش:** برای یک حالت دلخواه داریم:

$$\begin{aligned} F(|\psi\rangle, \Lambda_{dep}(|\psi\rangle\langle\psi|)) &= \sqrt{\langle\psi|\left(p\frac{I}{2} + (1-p)|\psi\rangle\langle\psi|\right)|\psi\rangle} \\ &= \sqrt{1 - \frac{p}{2}}. \end{aligned} \quad (۱۵۵)$$

این عبارت آخری مستقل از حالت ورودی است بنابراین همین مقدار کیفیت کانال واقطبش است. وقتی که $p = 1$ است، یعنی همه اطلاعات ورودی از بین برود، کیفیت کانال تا حد $\frac{1}{2}$ پایین می آید.

- **مثال: کیفیت کانال فازبرگردان:** با استفاده از تعریف (۱۵۴) داریم

$$Q(\Lambda) := min_{|\psi\rangle} F(|\psi\rangle, \Lambda(|\psi\rangle\langle\psi|)), \quad (۱۵۶)$$

که در آن

$$\Lambda(\rho) = p\rho + (1-p)Z\rho Z. \quad (۱۵۷)$$

با یک محاسبه ساده معلوم می شود

$$F(|\psi\rangle, \Lambda(|\psi\rangle\langle\psi|)) = \sqrt{\langle\psi|\left(p|\psi\rangle\langle\psi| + (1-p)Z|\psi\rangle\langle\psi|Z\right)|\psi\rangle}$$

$$= \quad \sqrt{p + (1-p)\langle\psi|Z|\psi\rangle^2}. \qquad (۱۵۸)$$

حالتی که کمترین مقدار فیدلیته را تولید می کند عبارت است از $|\psi\rangle = |+\rangle$ که برای آن $\langle +|Z|+\rangle$ برابر با صفر است. بنابراین برای کانال فازبرگردان بدست می آوریم:

$$Q(\Lambda_{\mathrm{Phase-Flip}}) = \sqrt{p}. \qquad (۱۵۹)$$

آیا می توان رابطه ای کلی برای کیفیت یک کانال کوانتومی یک کیوبیتی بدست آورد؟ پاسخ این سوال مثبت است. برای این کار حالت خالص ورودی را به شکل زیر می نویسیم:

$$|\psi\rangle\langle\psi| = \frac{1}{2}(I + \mathbf{n}\cdot\boldsymbol{\sigma}) \qquad (۱۶۰)$$

که در آن $\mathbf{n}$ یک بردار یکه روی سطح کره بلوخ است. می دانیم که اثر یک کانال کوانتومی روی چنین حالتی آن است که بردار بلوخ با یک تبدیل آفین تغییر می کند، یعنی

$$\mathbf{n} \longrightarrow \mathbf{r} = M\mathbf{n} + \mathbf{t}. \qquad (۱۶۱)$$

در این رابطه $\mathbf{n}$ و $\mathbf{t}$ به صورت بردارهای ستونی و M به صورت یک ماتریس مربعی است. بنابراین داریم:

$$Q(\Lambda) = Tr(|\psi\rangle\langle\psi|\Lambda(|\psi\rangle\langle\psi|)) = Tr(\frac{1}{2}(I + \mathbf{n}\cdot\boldsymbol{\sigma})\frac{1}{2}(I + \mathbf{r}\cdot\boldsymbol{\sigma})) \qquad (۱۶۲)$$

و با توجه به خواص ماتریس های پاوولی

$$Q(\Lambda) = \frac{1}{2}(1 + \mathbf{n}\cdot\mathbf{r}) = \frac{1}{2}(1 + \mathbf{n}^T(M\mathbf{n} + \mathbf{t})) \qquad (۱۶۳)$$

اگر کانال یکانی (یونیتال) باشد، می دانیم که $\mathbf{t}$ برابر با صفر است. در این صورت بردار ستونی $\mathbf{n}$ را می بایست ویژه بردار M گرفت تا کمترین مقدار تشابه بدست بیاید. در نتیجه پاسخ نهایی برابر است با:

$$Q(\Lambda) = \frac{1}{2}(1 + \lambda_{min}) \qquad (۱۶۴)$$

که در آن λ_{min} کوچکترین ویژه مقدار ماتریس M است.

■ **تمرین:** اگر کانال یونیتال نباشد، عبارت (۱۶۴) با چه عبارتی جایگزین خواهد شد؟ با یک محاسبه دقیق این عبارت را پیدا کنید.

ممکن است بپرسیم چرا ورودی کانال را یک حالت خالص گرفته ایم، زیرا در عمل ورودی یک کانال کوانتومی الزاما یک حالت خالص نیست. پاسخ در خاصیت تحدب تابع تشابه نهفته است. دیدیم که برای چنین تابعی همیشه مخلوط کردن باعث افزایش مقدار آن می شود بنابراین اگر می خواهیم کمینه تابع را پیدا کنیم باید در جستجوی یک حالت خالص باشیم. به عبارت صریح تر فرض کنید که حالت ورودی یک حالت آمیخته باشد. یک تجزیه آنزامبلی آن را در نظر می گیریم. داریم:

$$F(\rho,\Lambda(\rho)) = F(\sum_{i=1}^n \lambda_i|i\rangle\langle i|, \sum_i \lambda_i\Lambda(|i\rangle\langle i|)) \quad (۱۶۵)$$

از رابطه (۱۴۸) بدست می آوریم:

$$\begin{aligned} F(\rho,\Lambda(\rho)) &= F(\sum_{i=1}^n \lambda_i|i\rangle\langle i|, \sum_i \lambda_i\Lambda(|i\rangle\langle i|)) \\ &\geq (\sum_i \lambda_i)F(|i_0\rangle\langle i_0|, \Lambda(|i_0\rangle\langle i_0|)) = F(|i_0\rangle\langle i_0|, \Lambda(|i_0\rangle\langle i_0|)), \end{aligned} \quad (۱۶۶)$$

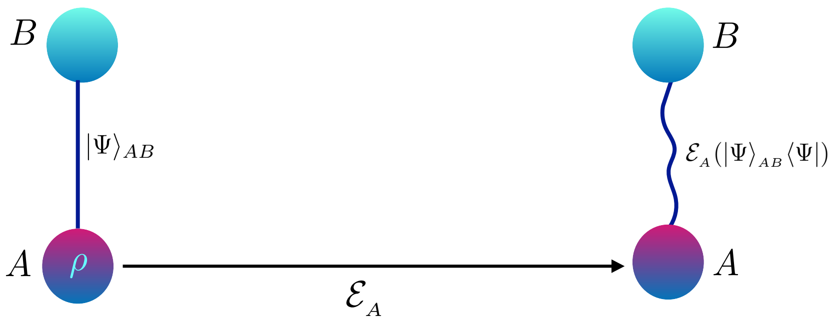
که در آن $|i_0\rangle$ حالت خالصی است که در این تجزیه آنزامبلی کمترین مقدار فیدلیته را تولید کرده است. بنابراین به ازای هر حالت آمیخته حتما یک حالت خالص وجود دارد که فیدلیته کمتری را تولید می کند.

■ **تمرین:** کانال میراکننده دامنه را برای کیوبیت ها در نظر بگیرید و کیفیت آن را حساب کنید.

■ **تمرین:** کیفیت کانال میراکننده فاز را برای کیوبیت ها محاسبه کنید.

۹ تشابه درهم تنیدگی

شکل (۷) یک سیستم دو بخشی را که از دو بخش A و B تشکیل شده نشان می دهد. این دو بخش در یک حالت درهم تنیده مثل $|\Psi\rangle_{AB}$ قرار دارند. سیستم A در حالتِ ρ قرار دارد. حال سیستمِ A را تحت تاثیر یک کانال کوانتومی Λ قرار می دهیم. هدف ما این است که بفهمیم این



شکل ۷: اثر یک کانال کوانتومی روی یک سیستم میزان درهم تنیدگی آن سیستم را با محیط تغییر می دهد.

کانال که روی تنها قسمتی از حالت $|\Psi\rangle_{AB}$ اثر می کند چه مقدار روی درهم تنیدگی دو قسمتِ A و B اثر می گذارد. هدف ما این است که میزان شباهت درهم تنیدگی سیستم A را با B قبل از اثر کانال با بعد از آن مقایسه کنیم. حالت $|\Psi\rangle$ را تجزیه اشمیت می کنیم:

$$|\Psi\rangle = \sum_i \sqrt{p_i}|i, a_i\rangle.$$

آنچه که می خواهیم محاسبه کنیم کمیت زیر است:

$$\mathcal{E}_f(\rho, \Lambda) := \langle\Psi|\Bigg((\Lambda \otimes I)|\Psi\rangle\langle\Psi|\Bigg)|\Psi\rangle. \tag{۱۶۷}$$

نمادی که برای عبارت سمت چپ پیدا کرده ایم در واقع نتیجه محاسبه ای است که اکنون انجام می دهیم که نشان می دهد این تشابه در هم تنیدگی در واقع به حالت زیرسیستم ρ بستگی دارد و به حالت کل سیستم AB بستگی ندارد. حال با کمی محاسبه سرراست می نویسیم:

$$(\Lambda \otimes I)|\Psi\rangle\langle\Psi| = \sum_{i,j} \sqrt{p_i p_j}\Lambda(|i\rangle\langle j|) \otimes |a_i\rangle\langle a_j|. \tag{۱۶۸}$$

در نتیجه با استفاده از تعامد حالت های $|a_i\rangle$ بر هم به رابطه زیر می رسیم:

$$\mathcal{E}_f(\rho, \Lambda) = \sum_{i,j} p_i p_j \langle i|\Lambda(|i\rangle\langle j|)|j\rangle \tag{۱۶۹}$$

اما نکته مهم این است که این کمیت را می توان کاملا بر حسب حالت فقط سیستمِ A نوشت، زیرا با توجه به این که $\rho=\sum_i p_i|i\rangle\langle i|$ و با توجه به نمایش کراوس برای کانال $\Lambda(\rho)=\sum_m K_m\rho K_m^\dagger$ می توانیم بنویسیم:

$$\begin{aligned}\mathcal{E}_f(\rho,\Lambda) &= \sum_{i,j} p_ip_j\langle i|\Lambda|i\rangle\langle j|j\rangle=\sum_m K_m\langle i|K_m|i\rangle\langle j|K_m^\dagger|j\rangle \\ &= \sum_m tr(K_m\rho)tr(K_m^\dagger\rho)=\sum_m|tr(K_m\rho)|^2. \end{aligned}\tag{۱۷۰}$$

به این ترتیب می بینیم که تشابه درهم تنیدگی تنها به حالت زیر بخش A بستگی دارد و نه به حالت کل سیستم AB. به همین دلیل می توانیم کمیت F را با $\mathcal{E}_f(\rho,\Lambda)$ نشان دهیم. بنابراین نشان داده ایم:

$$\mathcal{E}_f(\rho,\Lambda)=\Big(|\Psi\rangle_{AB},\Lambda_A(|\Psi\rangle_{AB}\langle\Psi|)\Big)=\sum_m tr(K_m\rho)tr(K_m^\dagger\rho)=\sum_m|tr(K_m\rho)|^2=\mathcal{E}_f(\rho).\tag{۱۷۱}$$

■ **تمرین:** نشان دهید که تشابه درهم تنیدگی بستگی به نوع نمایش کراوسی که برای کانال انتخاب می کنیم ندارد.

■ **تمرین:** تشابه درهم تنیدگی را برای یک حالت دلخواه و کانال های بیت برگردان، فاز برگردان و واقطبش پیدا کنید و باهم مقایسه کنید.

حال از خود می پرسیم تشابه درهم تنیدگی چه فایده دیگری دارد؟ برای اینکه پاسخ این سوال و سوال های مشابه را بفهمیم، نیاز به چند تعریف و قضیه کوچک داریم.

■ **لم:** تابع $\mathcal{E}_f(\rho,\Lambda)$ یک تابع مقعر است. یعنی

$$\mathcal{E}_f(\sum_i\lambda_i\rho_i,\Lambda)\le\sum_i\lambda_i\mathcal{E}_f(\rho_i,\Lambda).\tag{۱۷۲}$$

■ **اثبات:** برای اثبات این خاصیت کافی است که یک تابع جدید به صورت زیر تعریف کنیم.

$$f(x):=\mathcal{E}_f(x\rho_1+(1-x)\rho_2)=\sum_m|tr(K_m(x\rho_1+(1-x)\rho_2))|^2$$

مشتق دوم این تابع به راحتی محاسبه می شود: (این محاسبه برعهده خواننده است.)

$$f"(x)=\sum_m tr(|(\rho_1-\rho_2)K_m|^2)\ge 0.\tag{۱۷۳}$$

بنابراین، این تابع یک تابع مقعر است. در نتیجه داریم

$$f(\lambda) \leq (1-\lambda)f(0) + \lambda f(1).$$

با توجه به تعریف تابع $f(x)$ بدست می آوریم:

$$\mathcal{E}_f(\lambda\rho_1 + (1-\lambda)\rho_2, \Lambda) \leq \lambda\mathcal{E}_f(\rho_1, \Lambda) + (1-\lambda)\mathcal{E}_f(\rho_2, \Lambda). \tag{۱۷۴}$$

با تعمیم این رابطه به (۱۷۲) می رسیم.

این خاصیت نشان می دهد که مخلوط ترین حالت ها تشابه درهم تنیدگی شان از همه حالت های دیگر کمتر است. اما مخلوط ترین حالت، حالت بیشسینه مخلوط یعنی $\frac{I}{d}$ است. این حالت برای ρ_A به معنای این است که حالت اولیه ψ_{AB} یک حالت درهم تنیده ماکزیمال بوده است. بنابراین، این نتیجه از نظر شهودی می گوید که حالت های با بیشترین درهم تنیدگی، وقتی که یک بخش از آنها تحت تاثیر یک کانال قرار می گیرند، بیشترین آسیب را می بینند و شباهت آنها با حالت قبل از اثر کانال مقدار کمتری است. به عبارت دیگر حالت های درهم تنیده ماکزیمال بیشترین میزان شکنندگی را دارند.

حال می پرسیم که اگر خود حالت ρ را مستقیما به کانال Λ بفرستیم، تشابه بدست آمده بین ورودی و خروجی چه رابطه ای با آنچه که بدست آوردیم دارد؟ در واقع یک بار حالت ρ را مستقیما به کانال می فرستیم و تشابه ورودی و خروجی را اندازه می گیریم که آن را با $F(\rho, \Lambda(\rho))$ نشان می دهیم و یک بار دیگر هم تشابه حالت های درهم تنیده ی $|\psi_{AB}\rangle$ را قبل و بعد از ورود به کانالِ $\Lambda \otimes I_B$ محاسبه می کنیم. می خواهیم این دو را باهم مقایسه کنیم. نکته این است که محاسبه مستقیمِ $F(\rho, \Lambda(\rho))$ به دلیل آنکه هر دو حالت مخلوط هستند آسان نیست. اما محاسبه $\mathcal{E}_f(\rho, \Lambda)$ آنچنان که دیدیم آسان است. این رابطه در قضیه بعدی بیان شده است:

- **قضیه:** به ازای هر حالت دلخواه ρ تشابه حالت ورودی و خروجی کانال Λ از $\mathcal{E}_f(\rho, \Lambda)$ بیشتر است. به عبارت دیگر رابطه زیر را داریم:

$$\mathcal{E}_f(\rho, \Lambda) \leq F(\rho, \Lambda(\rho)). \tag{۱۷۵}$$

- **اثبات:** از این خاصیت که تشابه حالت ها تحت یک کانال همواره زیاد می شود استفاده می کنیم . در این مورد کانالی که از آن استفاده می کنیم، کانالِ رد جزیی یعنی کانال Tr_B است:

$$Tr_B : |\Psi\rangle\langle\Psi| \longrightarrow \rho. \tag{۱۷۶}$$

حالت $|\Psi\rangle$ را به عنوان خالص سازی شده ی ρ در نظر می گیریم. یعنی

$$\rho = Tr_B|\Psi\rangle\langle\Psi|. \tag{۱۷۷}$$

بنابراین می نویسیم:

$$F(\rho,\Lambda(\rho)) = F\Big(Tr_B|\Psi\rangle\langle\Psi|, \Lambda(Tr_B|\Psi\rangle\langle\Psi|)\Big). \tag{۱۷۸}$$

حال از این نکته استفاده می کنیم که

$$\Lambda(Tr_B(|\Psi\rangle\langle\Psi|)) = Tr_B(\Lambda\otimes I)(|\Psi\rangle\langle\Psi|) \tag{۱۷۹}$$

و در نتیجه رابطه بالا به شکل زیر در می آید:

$$F(\rho,\Lambda(\rho)) = F\Big(Tr_B|\Psi\rangle\langle\Psi|, Tr_B(\Lambda\otimes I)|\Psi\rangle\langle\Psi|\Big). \tag{۱۸۰}$$

حال از این موضوع استفاده می کنیم که Tr_B مثل یک کانال عمل می کند و تحت تاثیر این کانال، مثل هر کانال دیگری، تشابه افزایش می یابد. بنابراین یک نامساوی به سمت راست رابطه بالا اضافه می شود:

$$F\Big(|\Psi\rangle\langle\Psi|, \Lambda|\Psi\rangle\langle\Psi|\Big) \le F\Big(Tr_B|\Psi\rangle\langle\Psi|, Tr_B(\Lambda\otimes I)|\Psi\rangle\langle\Psi|\Big) = F(\rho,\Lambda(\rho)). \tag{۱۸۱}$$

اما آخرین عبارت سمت راست این تساوی چیزی نیست جز $\mathcal{E}_f(\rho,\Lambda)$. بنابراین ثابت کردیم:

$$\mathcal{E}_f(\rho,\Lambda) \le F(\rho,\Lambda(\rho)). \tag{۱۸۲}$$

این نامساوی چه اهمیتی دارد؟ بسیاری از اوقات می خواهیم تشابه یک حالت کلی را قبل از اثر کانال با همان حالت وقتی که از کانال خارج می شود، یعنی $F(\rho,\Lambda(\rho))$ را حساب کنیم. از آنجا که هر دو حالت آمیخته هستند، محاسبه این تشابه بسیار مشکل است و در ابعاد بالا به صورت تحلیلی غیرممکن است. رابطه بالا به ما این امکان را می دهد که یک حد پایین برای این تشابه پیدا کنیم، زیرا محاسبه تشابه درهم تنیدگی یعنی $\mathcal{E}_f(\rho,\Lambda)$ با توجه به رابطه (۱۷۰) آسان است. این که نشان دهیم تشابه یک حالت با وقتی که از کانال عبور می کند از یک مقدار معینی نمی تواند کمتر شود، کار خیلی مهمی است.

این قضیه را در باره یک حالت ورودی ثابت کرده ایم. اما اغلب اوقات پیش می آید که به متوسطِ تشابهِ ورودی و خروجی یک کانال برای یک مجموعه از حالت ها نیاز داریم. فرض کنید که یک آنزامبل از حالت ها مثل $S=\{p_i,\rho_i\}$ داشته باشیم، که در آن ρ_i حالت ورودی و p_i احتمال انتخاب آن حالت ورودی برای ورود به کانال است. بنابراین آنچه که می خواهیم محاسبه کمیت زیر است:

$$\sum_i p_i F(\rho_i,\Lambda(\rho_i)). \tag{۱۸۳}$$

حال از این تعریف و نامساوی (۱۸۲) استفاده می کنیم و می نویسیم:

$$\sum_i p_i \mathcal{E}_f(\rho_i,\Lambda) \leq \sum_i p_i F(\rho_i,\Lambda(\rho_i)) \tag{۱۸۴}$$

و در قدم بعد نیز از خاصیت تقعر تشابه درهم تنیدگی استفاده می کنیم و به رابطه زیر می رسیم:

$$\mathcal{E}_f(\sum_i p_i\rho_i,\Lambda) \leq \sum_i p_i \mathcal{E}_f(\rho_i,\Lambda) \leq \sum_i p_i F(\rho_i,\Lambda(\rho_i)) \tag{۱۸۵}$$

نتیجه نهایی عبارت است از:

$$\mathcal{E}_f(\sum_i p_i\rho_i,\Lambda) \leq \sum_i p_i F(\rho_i,\Lambda(\rho_i)). \tag{۱۸۶}$$

این رابطه آخر می گوید که متوسط تشابه حالت های ورودی و خروجی از تشابه درهم تنیدگی برای حالت متوسط یعنی $\overline{\rho}=\sum_i p_i\rho_i$ بیشتر است.

۱.۹ رابطه فاصله و شباهت

اگر شباهت نشان می دهد که دو حالت کوانتومی چقدر به هم نزدیک اند و فاصله نشان می دهد که این دو حالت چقدر از هم دورند، خیلی طبیعی است که بین این دو یک رابطه ساده وجود داشته باشد. مثلا ممکن است سوال کنیم چرا فاصله را خیلی ساده به شکل $D(\rho,\sigma)=1-F(\rho,\sigma)$ تعریف نمی کنیم. در واقع چنین فاصله هایی تعریف می شوند که البته تعریف آنها هم چندان مثل رابطه اخیر ساده و سرراست نیست چرا که چیزی را که تعریف می کنیم می بایست در خواص فاصله و بخصوص نامساوی مثلث صدق کند. هرکدام خواص ویژه خود را دارند. در زیر دو تا از این فاصله ها را معرفی می کنیم. هر دوی این فاصله ها، فاصله های بورِس [۱۲] نامیده می شوند و چنین تعریف می شوند:

[۱۲] Bures Distance

■

$$D_b(\rho,\sigma) := \sqrt{1-F^2(\rho,\sigma)} \tag{۱۸۷}$$

و

■

$$D(\rho,\sigma) := \cos^{-1}(F(\rho,\sigma)). \tag{۱۸۸}$$

در ادامه خواص فاصله دوم را می کاویم. روابط اول و دوم زیر ساده و سرراست هستند:

$$\begin{aligned} & D(\rho,\rho)=0 \ \ , \ \ if\ D(\rho,\sigma)=0 \ \ \longrightarrow \ \ F(\rho,\sigma)=1, \ \ \longrightarrow \ \ \rho=\sigma, \\ & 0\leq D(\rho,\sigma)\leq 1 \\ & D(\rho,\gamma)\leq D(\rho,\sigma)+D(\sigma,\gamma). \end{aligned} \tag{۱۸۹}$$

رابطه سوم را با استفاده از قضیه یولمان می توان ثابت کرد. فرض کنید که $|\phi\rangle$ یک خالص سازی از حالت ρ باشد. بنابر قضیه یولمان می دانیم که

$$F(\rho,\sigma) = Max_{\psi}|\langle\phi|\psi\rangle| \tag{۱۹۰}$$

که در آن $|\psi\rangle$ یک خالص سازی از σ است. دقت کنید که در این قضیه یک خالص سازی دلخواه از یکی از حالت ها در نظر گرفته می شود و سپس بیشینه روی همه خالص سازی های حالت دیگر محاسبه می شود. حال با استفاده از این قضیه یک خالص سازی دلخواه از حالت σ در نظر می گیریم. این حالت را با $|\phi\rangle$ نشان می دهیم. حالت های $|\psi\rangle$ و $|\chi\rangle$ حالت هایی هستند که تساوی های زیر را بنا بر قضیه یولمان برقرار می کنند:

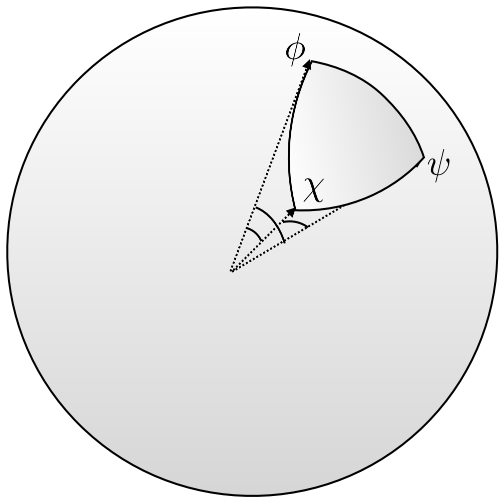
$$\begin{aligned} F(\rho,\sigma) &= \langle\phi|\psi\rangle \\ F(\sigma,\gamma) &= \langle\psi|\chi\rangle. \end{aligned} \tag{۱۹۱}$$

حال به دو نکته توجه می کنیم. نکته اول این که با توجه به شکل (۸)

$$\cos^{-1}\langle\phi,\chi\rangle \leq \cos^{-1}\langle\phi,\psi\rangle + \cos^{-1}\langle\psi,\chi\rangle = \cos^{-1}F(\rho,\sigma) + \cos^{-1}F(\sigma,\gamma), \tag{۱۹۲}$$

و دوم اینکه با توجه به قضیه یولمان

$$F(\rho,\gamma) \geq \langle\phi|\chi\rangle.$$



شکل ۸: از آنجا که زاویه های بین بردارها متناسب با طول های روی کره هستند، رابطه $\cos^{-1}\langle\phi,\chi\rangle \leq \cos^{-1}\langle\phi,\psi\rangle + \cos^{-1}\langle\psi,\chi\rangle$ برقرار است.

اما با توجه به نزولی بودن تابع کسینوس از این رابطه آخر نتیجه می گیریم

$$\cos^{-1} F(\rho,\gamma) \leq \cos^{-1}\langle\phi|\chi\rangle. \tag{۱۹۳}$$

با ترکیب این دو رابطه آخر به نتیجه نهایی می رسیم:

$$\cos^{-1} F(\rho,\gamma) \leq F(\rho,\sigma) + \cos^{-1} F(\sigma,\gamma),$$

یعنی نامساوی مثلث برای این فاصله برقرار است.

۲.۹ رابطه بین فاصله های مختلف

برای حالت های خالص فاصله ردّ و فیدلیته معادلند. برای فهم این موضوع فرض کنید که یکی از این حالات $|a\rangle$ و دیگری $|b\rangle$ باشد. بافرایند گرام‌ـاشمیت دو حالت متعامد پیدامی کنیم به نحوی که $|a\rangle = |0\rangle$ و $|b\rangle = \cos\theta|0\rangle + \sin\theta|1\rangle$. دراین صورت

$$F(|a\rangle, |b\rangle) = |\cos\theta|. \tag{۱۹۴}$$

$$\begin{aligned} D(|a\rangle, |b\rangle) &= \frac{1}{2} tr \left| \begin{bmatrix} 1-\cos^2\theta & -\cos\theta\sin\theta \\ -\cos\theta\sin\theta & -\sin^2\theta \end{bmatrix} \right| \\ &= |\sin\theta| = \sqrt{1 - F(|a\rangle, |b\rangle)^2}. \end{aligned} \tag{۱۹۵}$$

بنابراین، برای حالت های خالص فاصله ردّ تابعی از فیدلیته است. حال از خاصیت های تحدب فاصله استفاده می کنیم و برای حالت های آمیخته یک نامساوی بدست می آوریم. به این ترتیب که قرار می دهیم:

$$D(\rho, \sigma) \leq D(|\psi\rangle, |\phi\rangle) = \sqrt{1 - F(\rho, \sigma)^2}. \tag{۱۹۶}$$

بنابراین یک حد بالا برای فاصله دو حالت دلخواه بدست آوردیم. حال یک حد پایین بدست می آوریم. برای اثبات به نامساوی زیر توجه می کنیم. فرض کنید که $\{p_m\}$ و $\{q_m\}$ دو تابع توزیع احتمال کلاسیک هستند. در این صورت می توانیم بنویسیم:

$$\sum_m (\sqrt{p_m} - \sqrt{q_m})^2 = \sum_m p_m + q_m - 2\sum_m \sqrt{p_m q_m} = 2(1 - F). \tag{۱۹۷}$$

از طرف دیگر داریم

$$\begin{aligned} \sum_m (\sqrt{p_m} - \sqrt{q_m})^2 &\leq \sum_m |\sqrt{p_m} - \sqrt{q_m}||\sqrt{p_m} + \sqrt{q_m}| = \sum_m |p_m - q_m| \\ &= \sum_m |tr(\rho E_m) - tr(\sigma E_m)| \leq D(\rho, \sigma). \end{aligned} \tag{۱۹۸}$$

هرگاه نامساوی هایی را که بدست آورده ایم ترکیب کنیم به رابطه زیر می رسیم که یک حد بالا و پایین برای فاصله برحسب فیدلیته بدست می دهد:

$$2(1 - F(\rho, \sigma)) \leq D(\rho, \sigma) \leq \sqrt{1 - F^2(\rho, \sigma)}. \tag{۱۹۹}$$

۱۰ مسئله ها

- این مسئله بیشتر به اندازه گیری ها مربوط است و ربط مستقیمی به درس کنونی ندارد.

دو حالتِ

$$\rho = \frac{1}{3}|0\rangle\langle 0| + \frac{2}{3}|1\rangle\langle 1| \tag{۲۰۰}$$

و

$$\psi = \frac{1}{\sqrt{3}}|0\rangle + \frac{\sqrt{2}}{\sqrt{3}}|1\rangle \tag{۲۰۱}$$

را در نظر بگیرید. یک اندازه گیری POVM با المان های

$$E_0 = a|0\rangle\langle 0| + b|0\rangle\langle 1| + c|1\rangle\langle 0| + d|1\rangle\langle 1|, \qquad E_1 = I - E_0 \tag{۲۰۲}$$

چنان طراحی کنید که امکان تشخیص خطا را بین این دو حالت به حداقل برساند.

- دو حالت دو کیوبیتی زیر را در نظر بگیرید:

$$\hat{\rho} = \frac{1}{4}(I + \sum_{i=1}^{3} r_i\sigma_i \otimes \sigma_i) \qquad \hat{\sigma} = \frac{1}{4}(I + \sum_{i=1}^{3} s_i\sigma_i \otimes \sigma_i). \tag{۲۰۳}$$

فاصله این دو حالت و هم چنین شباهت آنها را با هم حساب کنید.

- نشان دهید که در قضیه مربوط به فاصله بین دو توزیع احتمال می توان قدر مطلق را برداشت و نوشت:

$$D(p,q) = max_S(p(S) - q(S)) = max_S\left(\sum_{x\in S} p_x - \sum_{x\in S} q_x\right). \tag{۲۰۴}$$

- فاصله رد [۱۳] بین دو حالت زیر را محاسبه کنید:

$$\rho = \frac{3}{4}|0\rangle\langle 0| + \frac{1}{4}|1\rangle\langle 1|, \qquad \sigma = \frac{2}{3}|+\rangle\langle +| + \frac{1}{3}|-\rangle\langle -|. \tag{۲۰۵}$$

[۱۳] distance trace

■ نشان دهید که فاصله رد دارای خاصیت تقعر زیر است:

$$D(\sum_i p_i\rho_i,\sigma) \leq \sum_i p_i D(\rho_i,\sigma). \quad (۲۰۶)$$

■ قضیه نقطه ثابت براوور [۱۴] قضیه ای در آنالیز است که ثابت می کند هر تابع پیوسته ای مثل $f : H \longrightarrow H$ که روی یک زیرمجموعه محدب از یک فضای هیلبرت تعریف شده باشد، دارای یک نقطه ثابت است، یعنی نقطه ای مثل x در آن زیر مجموعه وجود دارد به قسمی که $f(x)=x$. از این قضیه استفاده کنید و نشان دهید که هر کانال کوانتومی ردّ ـ نگهدار یک حالت ناوردا دارد یعنی

$$\exists\, \rho \ \mid\ \Lambda(\rho)=\rho.$$

■ فرض کنید که Λ یک کانال کوانتومی ردّ ـ نگهدار است که فاصله ها را اکیدا کم می کند، یعنی $D(\Lambda(\rho),\Lambda(\sigma)) < D(\rho,\sigma)$. این نوع کانال ها، اکیداً انقباضی [۱۵] نامیده می شوند. نشان دهید که این کانال حتما یک نقطه ثابت دارد.

■ فرض کنید که Λ یک کانال کوانتومی ردّ ـ نگهدار است که به صورت زیر عمل می کند:

$$\Lambda(\rho) = p\rho_0 + (1-p)\Lambda'(\rho) \qquad 0 < p \leq 1. \quad (۲۰۷)$$

به عبارت دیگر این کانال حالت ρ را با یک احتمال غیر صفر با یک حالت ثابت ρ_0 جایگزین می کند و با احتمال $(1-p)$ کانال Λ' را روی آن اثر می دهد. نشان دهید که کانال Λ یک کانال اکیدا انقباضی است و در نتیجه یک نقطه ثابت دارد.

■ نشان دهید که

$$F(\rho,\sigma) = Max_{|\phi\rangle}(|\langle\psi|\phi\rangle|)$$

که در آن $|\psi\rangle$ یک خالص سازی مشخص از ρ است و ماکزیمم روی همه خالص سازی های σ است.

■ نامساوی زیر را ثابت کنید:

$$1 - F^2(\psi,\sigma) \leq D(|\psi\rangle,\sigma). \quad (۲۰۸)$$

■ یک راه برای اندازه گیری اثر یک کانال نیز محاسبه تشابه متوسط خروجی و ورودی یک کانال است که به صورت زیر تعریف می شود:

$$\overline{F}(\Lambda) := \int d\psi F\Big(|\psi\rangle, \Lambda(|\psi\rangle\langle\psi|)\Big) \quad (۲۰۹)$$

[۱۴] Brower's Fixed Point Theorem

[۱۵] Strictly Contractive

- خلوص یک حالت با رابطه $F(\rho) := Tr(\rho^2)$ تعریف می شود. نشان دهید که خلوص یک حالت می تواند با مخلوط کردن آن حالت با یک حالت دیگر افزایش یابد. مثال مشخصی ارایه کنید.

- فرض کنید که ρ و ρ' دو حالت متفاوت باشند. نشان دهید که حتما یک عملگر مثبت E وجود دارد به نحوی که $Tr(\rho E) \neq Tr(\rho' E)$ باشد. این نتیجه به این معناست که حتما یک اندازه گیری ای وجود دارد که به کمک آن بتوان حالت های متفاوت را از هم تمیز داد.

- نخست نشان دهید که

$$|x-y| \geq (\sqrt{x}-\sqrt{y})^2, \tag{۲۱۰}$$

سپس نشان دهید که

$$d_1(p,q) \geq d_2^2(p,q). \tag{۲۱۱}$$

در اینجا d_1 و d_2 دو نوع تعریف برای فاصله بین دو تابع توزیع احتمال هستند:

$$d_1(p,q) := \frac{1}{2}\sum_i |p_i - q_i|, \tag{۲۱۲}$$

و

$$d_2(p,q) := \sqrt{1-\sum_i \sqrt{p_i q_i}}. \tag{۲۱۳}$$

اولی فاصله وردشی [۱۶] و دومی [۱۷] بین دو تابع توزیع احتمال نام دارد.

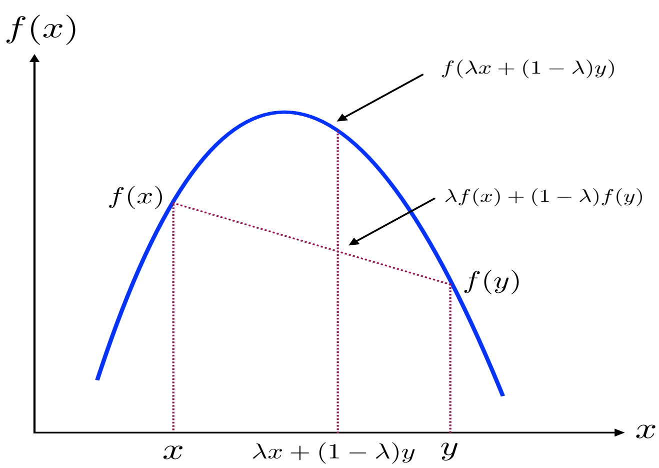
[۱۶] Variational Distance
[۱۷] Helinger Distance

۱۱ ضمیمه یک

- تعریف: تابع حقیقی و یک متغیره f محدب [۱۸] نامیده می شود اگر به ازای هر دو مقدار x_1 و x_2 در دامنه تعریف خود، دارای خاصیت زیر باشد:

$$f(x_1 + (1-\lambda)x_2) \geq \lambda f(x_1) + (1-\lambda) f(x_2). \tag{۲۱۴}$$

یک تابع محدب در شکل ۹ نشان داده شده است. چنین تابعی در تمامی حوزه تعریف خود در رابطه $\frac{d^2}{dx^2} f(x) \geq 0$ صدق می کند.



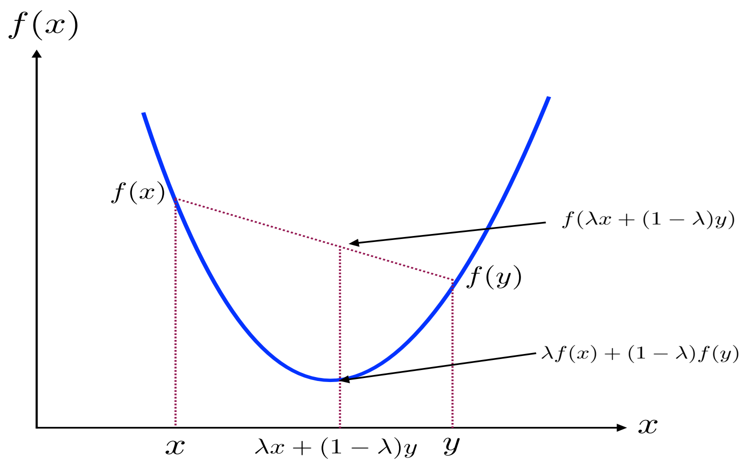
شکل ۹: خاصیت یک تابع محدب این است که اگر متغیرهایش را با هم مخلوط کنیم به سمت بیشینه تابع می رویم. بنابراین اگرتابعی از ماتریس های چگالی داشته باشیم که محدب باشد، مطمئن هستیم که بیشینه این تابع یک حالت کاملا آمیخته و کمینه آن یک حالت کاملا خالص است.

- تابع حقیقی و یک متغیره f مقعر [۱۹] نامیده می شود اگر به ازای هر دو مقدار x_1 و x_2 در دامنه تعریف خود، دارای خاصیت زیر باشد:

$$f(\langle x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2). \tag{۲۱۵}$$

یک تابع مقعر در شکل ۱۰ نشان داده شده است. چنین تابعی در تمامی حوزه تعریف خود در رابطه $\frac{d^2}{dx^2} f(x) \leq 0$ صدق می کند.

---
[۱۸] Convex
[۱۹] Concave



شکل ۱۰: خاصیت یک تابع مقعر ین است که اگر متغیرهایش را با هم مخلوط کنیم به سمت کمینه تابع می رویم. بنابراین اگرتابعی از ماتریس های چگالی داشته باشیم که مقعر باشد باشد، مطمئن هستیم که بیشینه این تابع یک حالت کاملا خالص و کمینه آن یک حالت کاملا آمیخته است.

■ تمرین: نشان دهید که یک تابع مقعر به ازای هر مجموعه از مقادیر x_1 تا x_n در رابطه زیر صدق می کند:

$$f(\sum_i \lambda_i x_i) \leq \sum_i \lambda_i f(x_i), \tag{۲۱۶}$$

که در آن λ_i ها اعداد مثبتی هستند که مجموع آنها برابر با یک است. رابطه مشابهی برای توابع محدب نیز برقرار است.

۱۲ ضمیمه دو

در این ضمیمه با یک معیار جدید از اندازه در فضاهای برداری آشنا می شویم که در نظریه اطلاعات کوانتومی زیاد به آن بر می خوریم. این معیار جدید اندازه شاتن [۲۰] است.

[۲۰] Schatten Norm (p-norm)

■ **اندازه شاتن:** یا اندازه ی p. به ازای هر عدد حقیقی $p \in (1, \infty)$ این اندازه برای یک بردار مختلط به صورت زیر تعریف می شود:

$$||\mathbf{v}||_p := (\sum_{i=1}^{n} |v_i|^p)^{\frac{1}{p}} \tag{۲۱۷}$$

به ازای $p = 2$ این اندازه همان اندازه اقلیدسی یا اندازه هیلبرت-اشمیت [۲۱] نامیده می شود. به ازای $p \longrightarrow \infty$ داریم:

$$\lim_{p\longrightarrow\infty} ||\mathbf{v}||_p = |v_{max}|, \tag{۲۱۸}$$

که در آن v_{max} مولفه ای از بردار است که بیشترین اندازه را دارد.

اگر $p = 1$ باشد، این نرم را به دلایل واضح نرم راننده تاکسی [۲۲] یا $l_1 - norm$ هم می گویند.

■ **تمرین:** یک دیسک به شعاع واحد به صورت زیر تعریف می شود:

$$D_p(1) := \{\mathbf{v} \in R^n \quad | \quad ||v||_p \leq 1.\} \tag{۲۱۹}$$

در فضای دو بعدی یعنی به ازای $n = 2$ شکل دیسک های به شعاع واحد را برای $p = 1, p = 2$ و $p = \infty$ رسم کنید.

■ **لم: نامساوی هولدر** که تعمیمی است از نامساوی کوشی شوارتز به شکل زیر است. (برای اثبات آن می توانید به هر کتاب جبر خطی مراجعه کنید.)

$$\sum_{i=1}^{n} |x_i y_i| \leq ||x||_p ||y||_q, \qquad \frac{1}{p} + \frac{1}{q} = 1. \tag{۲۲۰}$$

اندازه شاتن برای عملگرها نیز به طریق مشابه تعریف می شود.

■ **تعریف:** فرض کنید که

$$T : H \longrightarrow H$$

یک عملگر خطی متناهی بین روی یک فضای هیلبرت باشد. در این صورت نرم شاتن آن به صورت زیر تعریف می شود:

$$||T||_p := \big(Tr(|T|^p)\big)^{\frac{1}{p}} = (\sum_{i=1}^{n} \lambda_i^p)^{\frac{1}{p}} \tag{۲۲۱}$$

---

[۲۱] Hilbert-Schmidt

[۲۲] Taxicab norm

که در آن λ_i ها ویژه مقادیر مثبت عملگر $\sqrt{T^\dagger T}$ هستند.

■ **مثال:** اندازه عملگر

$$T = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

برابر است با:

$$||T||_p = (1+3^p)^{\frac{1}{p}}$$

و اندازه عملگر

$$T = \begin{pmatrix} 1 & 2 \\ 3 & \sqrt{6} \end{pmatrix}$$

برابر است با:

$$||T||_p = \left(\sqrt{12+3\sqrt{6}}^{\,p} + \sqrt{8+3\sqrt{6}}^{\,p}\right)^{\frac{1}{p}}.$$

در حالت های خاص داریم:

$$\begin{aligned} &||T||_1 = \text{Trace norm or } l_1\text{norm} \\ &||T||_2 = \text{Hilbert Schmidt Norm} \\ &||T||_\infty = \text{Sup Norm} = \sup_{\mathbf{x}} \frac{||T\mathbf{x}||}{||\mathbf{x}||}, \quad ||\mathbf{x}|| = \text{Euclidean or Hilbert Schmidt Norm of } \mathbf{x}. \end{aligned} \qquad (۲۲۲)$$

اندازه شاتن برای عملگرها چند خاصیت مهم دارد که به بعضی از آنها اشاره می کنیم. خواننده می تواند بعضی از این خاصیت ها را برای خود اثبات کند.

یک: اندازه شاتن ناوردای یکانی است. p نزولی است. یعنی

$$||UTV||_p = ||T||_p, \qquad \forall\, U\,,\, V.$$

دو: اندازه شاتن بر حسب p نزولی است. یعنی

$$||T||_1 \ge ||T||_p \ge ||T||_{p'} \ge ||T||_\infty, \qquad 1 \ge p \ge p' \ge \infty.$$

سه: به ازای هر دو عملگر در نامساوی زیر صدق می کند

$$||ST||_p \leq ||S||_p ||T||_p$$

چهار: با توجه به نامساوی هولدر اندازه شاتن در نامساوی زیر نیز صدق می کند:

$$||ST||_1 \leq ||S||_p ||T||_q \qquad \frac{1}{p} + \frac{1}{q} = 1.$$

پنج:

$$||T||_p = ||T^*||_p = ||T^t||_p = ||T^\dagger||_p.$$