دانشکده: فیزیک

گروه : ذرات بنیادی

عنوان پایان نامه ارشد:

اطلاعات شانون و اطلاعات کوانتمی

نام و نام خانوادگی ارائه دهنده:

حلیمه وحید

استاد راهنما :

دکتر حسین موحدیان

زمستان 1388

تشکر و قدردانی

از جناب آقای دکتر موحدیان که در طی انجام این رساله کمک‌های شایانی به من کردند کمال تشکر و قدردانی را دارم. هرچند نوشتن این چند کلمه هیچ‌گاه زحمات بی‌دریغ ایشان را پاسخ‌گو نبوده و حقی که ایشان بر گردن من دارند ادا نمی‌کند.

اقرارنامه

اینجانب تأیید می نمایم که مطالب مندرج دراین پـایان نامه (رساله) نتیجه تحقیقات خودم می باشد و در صورت استفاده از نتایج دیگران مرجع آن را ذکر نموده است.

کلیه حقوق مادی مترتب از نتایج مطالعات ، آزمایشات و نوآوری ناشی از تحقیق موضوع این پایان نامه (رساله) متعلق به دانشگاه صنعتی شاهرود می باشد.

بهمن 1388

چکیده

توسعه روز افزون علوم بشری و نیازهای انسان، او را ملزم به پیشبرد صنعت و تکنولوژی پا به پای علوم جدید کرده است. کامپیوترها و ابزار محاسباتی یکی از نیازهای انسان می‌باشند که اذهان متخصصین و دانشمندان را همیشه درگیر خود کرده است و دست‌اندرکاران هر روزه سعی بر این دارند که نوع پیشرفته‌تری از آن‌ها را عرضه کنند. ظهور مکانیک کوانتمی و کارایی آن در بسیاری از حوزه‌ها مانند فیزیک، شیمی و ...، بسیاری از دانشمندان را برآن داشت که از خواص منحصربفرد آن در کامپیوترها استفاده کنند. زیرا با گسترش مکانیک کوانتمی در این شاخه از تکنولوژی و ساخت کامپیوترهای کوانتمی می‌توان این سیستم‌ها را به شکل قابل توجهی بهینه کرد که به هیچ عنوان قابل مقایسه با سیستم‌های کلاسیکی نمی‌باشد.

نیازهای اشاره شده لازم می‌دارد که تحقیقات در زمینه مکانیک کوانتمی و کاربرد آن در کامپیوترها روز به روز ادامه یابد. در این اینجا سعی داریم به گوشه‌هایی از مفاهیم فیزیکی کامپیوترهای کوانتمی، هر چند به طور مختصر بپردازیم.

تئوری اطلاعات اهمیت بنیادین در کامپیوترهای کلاسیکی و کوانتمی دارد. در این رساله ابتدا در فصل اول در مورد ارتباط تئوری اطلاعات با مفاهیم فیزیکی به ویژه انرژی و کار صحبت می‌کنیم. در فصل دوم به تئوری اطلاعات کلاسیکی پرداخته و به ویژه اطلاعات شانون و خواص آنرا که نقش مهمی در کامپیوترهای کلاسیکی دارد، از نظر خواهیم گذراند. در فصل سوم در مورد تئوری اطلاعات کوانتمی صحبت کرده و آنتروپی فون نیومن را به عنوان مؤثرترین رابطه این مبحث مورد توجه قرار داده و خواص و کاربردهای آن را هر چند به شکلی مختصر ارائه می‌کنیم. در فصل چهارم به پیشنهادهایی که برای کمی‌کردن اطلاعات در کامپیوترهای کوانتمی شده اشاره کرده و رابطه‌های مفروض اتخاذ شده را با آنتروپی فون نیومن مقایسه می‌کنیم. همچنین در انتهای فصل چهارم به کارایی بهتر سیستم‌های سه‌حالتی یا سیستم‌های سه‌سطحی نسبت به سیستم‌های دو حالتی پرداخته و بعضی از مزیت‌های استفاده از این سیستم‌ها را بیان می‌کنیم.

کلمات کلیدی « آنتروپی آماری، آنتروپی شانون، آنتروپی فون نیومن، اطلاعات کلاسیکی، اطلاعات کوانتمی، بیت، تریت،کیوبیت، کیوتریت ».

لیست مقالات مستخرج از پایان‌نامه:

وحید، حلیمه؛ موحدیان، حسین؛ « *ناکارآمدی آنتروپی شانون در تعیین محتوای اطلاعاتی یک سیستم کوانتمی*»، کنفرانس فیزیک ایران، تابستان 1388، دانشگاه صنعتی اصفهان، اصفهان.

فهرست

فهرست مطالب

فهرست اشکال

فصل اول

رابطه انرژی و اطلاعات

- **مقدمه**
- **اولین گام‌های تعیین کننده رابطه بین انرژی و اطلاعات**
- **استخراج کار از اطلاعات**
- **رابطه آنتروپی شانون و تعریف آنتروپی در مکانیک آماری**

مقدمه

انرژی و اطلاعات موضوعاتی اساسی در حوزه فیزیک و علوم کامپیوتر می‌باشند. اگرچه در دید نخست این مفاهیم کاملاً بی‌ارتباط به نظر می‌رسند ولی جستجوی بیشتر نشان می‌دهد که موضوعاتی کاملاً وابسته بوده و ارتباطی جدا نشدنی دارند. بنابراین در این فصل رابطه بین انرژی و اطلاعات (آنتروپی و اطلاعات) مورد بحث قرار گرفته و تا حدودی مزایای این ارتباط و کمّ و کیف آن روشن می‌شود.

1-1- اولین گام‌های تعیین کننده رابطهٔ بین دو مفهوم انرژی و اطلاعات

1-1-1 پارادوکس شیطانک ماکسول

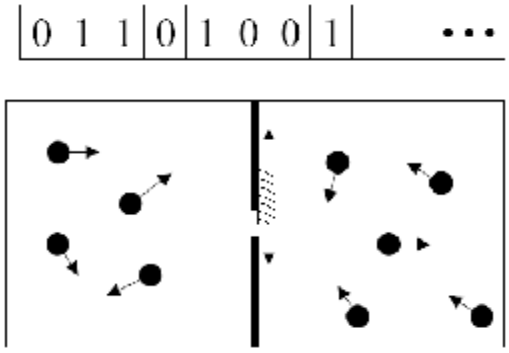
گاهی گفته می‌شود که بحث رابطه بین انرژی و اطلاعات از پارادوکس شیطانک ماکسول شروع شده است. ماکسول نخستین کسی بود که به ارتباط بین انرژی و اطلاعات و یا به عبارتی آنتروپی و اطلاعات پی برد. در واقع او با طرح یک آزمایش فکری، معروف به شیطانک ماکسول متوجه سرشت آماری قانون دوم ترمودینامیک و همچنین وجود رابطه‌ای بین آنتروپی و اطلاعات شد. ماکسول جعبه-ای حاوی گاز را که درون این جعبه به وسیلهٔ دیواری به دو بخش تقسیم می‌شد، در حکم مدل آزمایش خود معرفی کرد. او ابتدا اجازه داد که دما و فشار در هر دو قسمت جعبه برابر شوند شکل (1-1) [3]. سپس بر روی این دیوار در بدون وزن کوچکی در نظر گرفت که به گفتهٔ خودش با یک شیطانک باز و بسته می‌شد. ماکسول بعد از آماده‌سازی سیستم مورد نظر و باز کردن در مشاهده کرد که ذراتی که با سرعت از سمت چپ به دیواره نزدیک می‌شوند می‌توانند از در فرضی عبور کنند اما برای ذراتی که به آهستگی از سمت چپ به دیوار نزدیک می‌شدند، در باز نمی‌شد و نمی‌توانستند عبور کنند. علاوه بر این در مذکور در برابر ذراتی که از راست به دیوار نزدیک می‌شدند رفتار معکوسی از خود نشان می‌داد. یعنی به ذراتی که به آهستگی به دیوار نزدیک می‌شدند اجازهٔ عبور داده ولی برای ذرات سریع‌تر بسته باقی می‌ماند. اگرچه در ابتدا دما در کل جعبه یکسان بوده است اما در پایان تفاوت‌های دمایی ثابت شده و دمای سمت راست بیشتر از دمای سمت چپ می‌باشد. بنابراین آنتروپی

کل سیستم کاهش پیدا کرده است. در واقع ذرات سریعتر در سمت راست جعبه و ذرات کندتر در سمت چپ جمع می‌شوند، پس در سمت راست دما بالاتر از سمت چپ می‌باشد. بنابراین آشکارا شیطانک نظمی را در سیستم ایجاد کرده است. این رفتار ظاهراً قانون دوم ترمودینامیک[1] را نقض می-کند. چون این اینگونه به نظر می‌آید که شیطانک گرمای منبعی را که در دمای یکنواخت T قرار داشته به کار تبدیل کرده است. بعداً نشان داده شد که چنین تناقضی اتفاق نمی‌افتد و جواب مناسبی برای این رفتار شیطانک ارائه شد. البته ماکسول پذیرفت که شیطانک و درِ بدون وزن وجود ندارد، ولی تذکر داد که می‌توان یک شکاف خیلی کوچک روی دیوار تفکیک‌کنندهٔ در نظر گرفت.

یک جنبهٔ خیلی جالب آزمایش فکری ماکسول این است که شیطانکش برای آنکه بتواند ذرات را به این شکل مرتب کند از اطلاعاتی استفاده کرده است. شیطانک اطلاعاتی دربارهٔ سرعت ذرات نزدیک شونده کسب کرده و از این طریق تشخیص می‌دهد که در را برای عبور باز کند یا نه[1]. یعنی کسب اطلاعات راجع به سیستم آنتروپی سیستم را کاهش می‌دهد. سرانجام بعد از اینکه ذرات مرتب شدند و ذرات سریعتر در سمت راست و ذرات کندتر در سمت چپ جمع شدند می‌توان دیوار تفکیک کننده را به کناری کشید و اجازه داد که گازهای گرم و سرد مخلوط شوند. در طی این مخلوط‌سازی آنتروپی سیستم افزایش می‌یابد و اطلاعات راجع به سیستم کم می‌شود[1]. مقصود از آنتروپی سیستم بی نظمی سیستم است. پس می‌توان این گونه گفت که هرچه بی‌نظمی سیستم بیشتر باشد اطلاعات لازم برای شناخت کامل سیستم بیشتر بوده و ناآگاهی راجع به سیستم بیشتر برای کم کردن ناآگاهی راجع به سیستم و افزایش دانش دربارهٔ آن، لازم است که آنتروپی سیستم کاهش داده شود، یا به عبارتی سیستم به سمت حالتی منظم پیش برده شود. به کمک اصل لانداوو[2] می‌توان توضیحات مناسب و کامل‌تری برای این عمل شیطانک ارائه داد و رابطهٔ بین آنتروپی و است[2].

[1] این قانون بیان می‌کند که آنتروپی سیستم همواره در حال افزایش است.

[2] Landauer



شکل (1-1). طرح نمادین پارادوکس شیطانک ماکسول. حافظه شیطانک به‌وسیله یک رشته دودویی نمایش داده شده است که نتیجه اندازه‌گیری‌اش از موقعیت و سرعت مولکول‌ها را ذخیره می‌کند.

برای کم کردن ناآگاهی راجع به سیستم و افزایش دانش دربارهٔ آن، لازم است که آنتروپی سیستم کاهش داده شود، یا به عبارتی سیستم به سمت حالتی منظم پیش برده شود. به کمک اصل لانداوو[3] می‌توان توضیحات مناسب و کامل‌تری برای این عمل شیطانک ارائه داد و رابطهٔ بین آنتروپی و اطلاعات را به خوبی بررسی کرد. در زیر دربارهٔ این اصل و کاربرد آن در حل پارادوکس شیطانک ماکسول به طور مختصر صحبت می‌شود.

1-1-2 اصل لانداوو؛

شیطانک ماکسول بحث‌های زیادی پیش آورد و روش‌های متفاوتی برای حل این پارادوکس ارائه گردید. در آغاز عده‌ای معتقد بودند که شیطانک مقداری انرژی برای اندازه‌گیری‌های انجام گرفته، صرف کرده است. اگرچه بنت[4] و لانداوو قادر بودند نشان دهند که روند اندازه‌گیری را می‌توان بدون صرف انرژی نیز انجام داد و سرانجام توانستند پاسخی برای پارادوکس پیدا کنند. آن‌ها به این نتیجه رسیدند که نتایج اندازه‌گیری باید در حافظه شیطانک ذخیره شود و چون حافظه‌اش محدود است درنهایت باید آنرا پاک کند تا بتواند اندازه‌گیری‌های جدید انجام دهد.

[3] Landauer

[4] Bennet

اصل لانداوو: هرگاه یک بیت از اطلاعات پاک شود مقداری انرژی وارد محیط شده و هدر می‌رود که حداقل برابر $k_B T \ln 2$ می‌باشد. k_B ثابت بولتزمن و T دمای محیط می‌باشد. بنابراین می‌توان گفت که آنتروپی محیط به اندازهٔ $k_B T \ln 2$ افزایش پیدا کرده است.

بنابراین کاهش آنتروپی گاز با افزایش آنتروپی شیطانک همراه است و برای پاک کردن اطلاعاتی که شیطانک در فرایند اندازه‌گیری کسب کرده باید مقداری انرژی در محیط هدر رود. بنابراین بر طبق اصل لانداوو مصرف انرژی ناشی از اندازه‌گیری نیست بلکه ناشی از پاک کردن اطلاعات است.

1-1-3- رابطه پاک کردن اطلاعات و انجام کار

مثال زیر برای شرح اصل لانداوو مفید می‌باشد. فرض کنید اطلاعات در حالت یک سیستم فیزیکی ذخیره شود. مثلاً ممکن است یک بیت اطلاعات را با یک تک مولکول در جعبه تجسم کنیم. بنا بر قرارداد اگر مولکول در سمت چپ جعبه باشد، مقدار بیت برابر با **0** و اگر در سمت راست جعبه باشد، مقدار بیت برابر با 1 می‌باشد. قانون‌های ترمودینامیکی را می‌توان حتی برای یک مولکول نیز به‌کار برد. در ترمودینامیک داریم:

$$dE = dL + dQ \tag{1-1}$$

که dE تغییر انرژی درونی گاز، dL کار انجام شده روی گاز و dQ میزان گرمایی است که جذب گاز شده است. اگر یک تغییر حالت شبه پایا (تغییر حالتی چنان آهسته که بتوان حین آن سیستم را همواره در حال تعادل فرض کرد) در نظر بگیریم، می‌توان نوشت:

$$dS = \frac{dQ}{T} \tag{2-1}$$

که dS تغییر آنتروپی گاز می‌باشد. فرض می‌کنیم که جعبه در تماس گرمایی با حمام گرما در دمای T قرار دارد و گاز را با یک پیستون بدون اصطکاک فشرده می‌کنیم شکل (2-1). اگر پیستون به اندازه dx جابجا شود کار انجام شده روی گاز از رابطه زیر به‌دست می‌آید:

$$dL = -Fdx = -pAdx = -pdV \tag{3-1}$$

که F نیروی گاز روی پیستون، p فشار گاز، A سطح پیستون و V حجم گاز می‌باشد. البته چون ما تنها یک مولکول داریم مفاهیمی مانند فشار و نیرو باید در زمان متوسط در نظر گرفته شوند. متوسط-گیری روی تعداد زیاد برخورد مولکول با پیستون، انجام می‌گیرد. اینجا تغییر حالتی را در نظر می-گیریم که حجم گاز نصف حالت اولیه شود. معادله حالت برای گاز ایده‌آل را در نظر می‌گیریم :

$$pV = N\ \mathrm{k_B}T \tag{4-1}$$

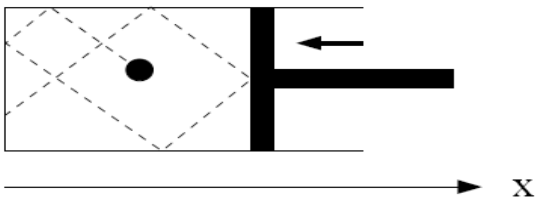
N تعداد مولکول‌های گاز می‌باشد (در اینجا $N=1$)، پس می‌توان کار انجام شده روی گاز را به صورت زیر محاسبه کرد:

$$L = -\int_V^{V/2} p\, dV' = -\int_V^{V/2} \frac{\mathrm{k_B}T}{V'} dV' = \mathrm{k_B}T\ \ln 2 \tag{5-1}$$

توجه کنید که در اینجا تبدیلات هم‌دما در نظر گرفته شده‌اند. چون سیستم در تماس گرمایی با حمام گرما در دمای T می‌باشد، انرژی درونی گاز تغییر نمی‌کند زیرا دما ثابت است. بنابراین بنا به قانون اول ترمودینامیک، معادله(1-1)، کار انجام شده روی گاز به گرمای تلف شده در محیط تبدیل می-شود: $\Delta Q = -L$. به این دلیل که گرما هدر رفته و جذب نشده است، $\Delta Q < 0$ می‌باشد. تغییر در آنتروپی گاز از معادلهٔ زیر به‌دست می‌آید.

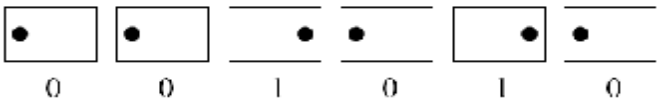
$$\Delta S = \frac{\Delta Q}{T} = \frac{-L}{T} = -\mathrm{k_B} \ln 2 \tag{6-1}$$

در اینجا $\Delta S < 0$، زیرا پس از متراکم کردن حجم قابل دسترس مولکول و بنابراین تعداد میکروحالت‌های قابل دسترس کاهش می‌یابد. آنتروپی سیستم کم می‌شود در حالی که آنتروپی محیط افزایش می‌یابد. چون آنتروپی کل جهان هرگز کاهش نمی‌یابد، داریم $\Delta S + \Delta S_{\mathrm{env}} \geq 0$. بنابراین



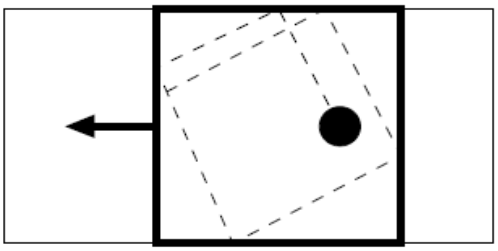
شکل(1-2) فشرده کردن گاز تک مولکولی به وسیله پیستون.

$\Delta S_{env} \geq k_B \ln 2$ که با اصل لانداوو در توافق است. حال فرض می‌کنیم یک پیام دودویی با دنباله‌ای از



شکل (1-3). دنباله‌ای از جعبه‌های تک مولکولی مشترک با رشتهٔ دودویی.

جعبه‌های تک مولکولی ذخیره شود. هر جعبه یک بیت از اطلاعات را ذخیره می‌کند. بسته به اینکه مولکول در چپ یا راست جعبه باشد به ترتیب حالت **0** یا 1 را به آن نسبت می‌دهیم شکل (1-3). حال باید اعتبار جملهٔ زیر نشان داده شود: اطلاعات نهفته در پیام باید با انرژی لازم برای پاک کردن پیام یعنی حرکت دادن همهٔ مولکول‌ها به سمت چپ یا راست جعبه برابر باشد. پیش از هر چیز باید محتوای اطلاعاتی پیام تعریف شود. در اینجا محتوای اطلاعاتی را به شکل مقدار اطلاعاتی تعریف می‌کنیم که با فهمیدن مقادیر بیت‌های تشکیل دهندهٔ پیام کسب می‌کنیم. بنابراین اطلاعات اندازه‌ای از ناآگاهی ما دربارهٔ پیام می‌باشد. اگر ما مقدار بیت‌ها را بدانیم هیچ اطلاعات بیشتری از پیام به‌دست نمی‌آوریم، در این مورد اطلاعاتی که پیام شامل می‌شود برابر صفر می‌باشد و بر طبق جملهٔ بالا هیچ انرژی‌ای صرف پاک کردن پیام نمی‌شود. در زیر نشان می‌دهیم که واقعاً هیچ کاری برای قرار دادن



شکل(1-4). روشی برای انتقال یک مولکول به سمت چپ جعبه بدون مصرف انرژی.

حالت هر بیت روی صفر لازم نیست و تنها کافی است مولکول را درون جعبه کوچکتری که در جعبه اولیه قرار دارد حبس کنیم و آن را همانند شکل (1-4) به چپ انتقال دهیم. هیچ کاری برای انجام این عمل لازم نیست زیرا مولکول به کرات به دیواره‌های داخلی چپ و راست جعبه برمی‌خورد و باز می‌گردد و جعبه به سمت راست رانده می‌شود. تنها در موردی که از پیش موقعیت مولکول را ندانیم، همانند مورد تک مولکولی، باید حجم گاز را نصف کنیم، این عمل به کاری برابر $L = \mathrm{k_B} T \ln 2$ نیاز دارد که باید روی گاز انجام شود. در این مورد محتوای اطلاعاتی گاز صفر نیست و برای پاک کردن آن باید انرژی مصرف کنیم[3].

1-2- استخراج کار از اطلاعات

مثال زیر برای درک بهتر رابطهٔ بین انرژی و اطلاعات آموزنده است. این مثال را بنت طراحی کرد. وی با این مثال نشان داد که از اطلاعات می‌توان در حکم سوخت برای حرکت ماشین استفاده کرد. فرض کنید گاری بارکشی در تماس با حمام گرمایی در دمای T قرار دارد و نواری ساخته شده از رشته‌ای از جعبه‌های تک مولکولی از آن عبور کرده است، شکل(1-5). اگر از ابتدا موقعیت چپ یا راست مولکول‌ها را بدانیم می‌توانیم برای حرکت گاری کار استخراج کنیم. برای این هدف تنها کافی است پیستونی را در وسط هر جعبه وارد کنیم. همان‌طور که در شکل (1-6) نشان داده شده است اگر مولکول در سمت چپ جعبه باشد پیستون می‌تواند به سمت راست حرکت کند ولی اگر مولکول در سمت راست جعبه باشد پیستون به چپ حرکت می‌کند. چون تمام سیستم در دمای T قرار دارد، از هر جعبه کار $L = \mathrm{k_B} T \ln 2$ قابل استخراج می‌باشد. اگر یک نوار N بیتی داشته باشیم کار کل $N\ \mathrm{k_B} T \ln 2$ است و برای جابجایی گاری می‌توان از آن استفاده کرد. تأکید می‌کنیم زمانی‌که نوار از گاری بیرون می‌آید مولکول‌ها می‌توانند هر جایی درون حجم V باشند، پس محتوای اطلاعاتی رشته-ای از جعبه‌ها کاملاً از بین رفته و در حکم سوخت برای حرکت گاری استفاده شده است. از طرف دیگر اگر در ابتدا موقعیت چپ یا راست مولکول‌ها مشخص نباشد نمی‌توان کار مفیدی استخراج کرد. در این مورد اگر پیستونی وارد جعبه‌ها کنیم نیمی از زمان‌ها گاز تولید کار می‌کند و نیم دیگر روی گاز کار انجام می‌شود. بنابراین کار انجام شده به طور متوسط برابر صفر می‌باشد[3].

1-3 - رابطه آنتروپی شانون[5] و تعریف آنتروپی در مکانیک آماری

آنتروپی شانون اگرچه تنها برای کاربرد در سیستم‌های ارتباطی ارائه گردید اما همانطور که خواهیم دید این آنتروپی در واقع همان آنتروپی کلاسیکی است که در مکانیک آماری گویای بی‌نظمی سیستم است. بنابراین مفهومی فیزیکی که رفتار سیستم‌ها را در ترمودینامیک و مکانیک آماری توجیه

[5] shannon

می‌کند در مبحث ارتباطات کاربرد بسیار مهمی پیدا کرده است. پس لازم می‌دانیم در این بخش اندکی راجع به رابطه بین آنتروپی آماری بولتزمن و آنتروپی شانون صحبت کنیم.

3-1- آنتروپی بولتزمن وشباهت آن با آنتروپی شانون

اطلاعات مستقیماً به کمیت‌های فیزیکی ربط داده نمی‌شود. در واقع اطلاعات نه ماده‌اند و نه انرژی، اگرچه ممکن است برای مکاتبه کردن آن‌ها به ماده و انرژی نیاز باشد. از این‌رو مقدار اطلاعات را نمی‌توان مستقیماً با دستگاهی اندازه گرفت و یا در واحدهایی که برای کمیت‌های فیزیکی اختصاص داده شده است بیان کرد[4]. در فیزیک کلاسیک معتبرترین رابطه برای نشان دادن میزان ناآگاهی ما دربارهٔ سیستم به صورت کمّی، با آنتروپی شانون داده می‌شود به همین دلیل شانون را پدر نظریهٔ اطلاعات می‌نامند. در فیزیک کلاسیک آنتروپی شانون همان آنتروپی بولتزمن، $S = \mathrm{k}\ln W$، است. در آنتروپی بولتزمن W تعداد میکروحالت‌های سیستم و $\mathrm{k} = 1{,}138\times 10^{-23}\ {}^{\mathrm{J}}\!/_{\mathrm{k}}$ ثابت بولتزمن می‌باشد.

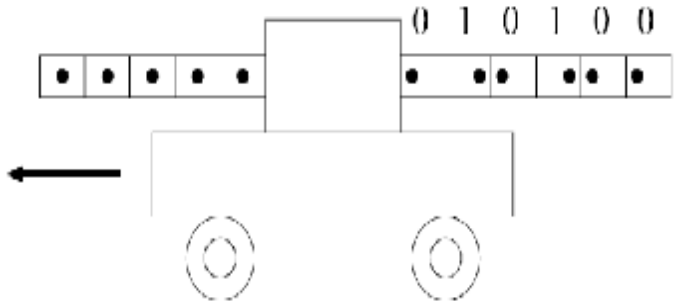
شانون (1916 - 2001) و لئو زیلارد[6](1898-1964) رابطه بین آنتروپی و فقدان یا کمبود اطلاعات را به صورت کمّی درآوردند. در سال 1946 نیاز AT&T برای کمّی کردن میزان اطلاعاتی که در یک سیم انتقال می‌یابد، باعث شد شانون نتیجهٔ مطالعات خود را دربارهٔ اطلاعاتی که در کامپیوترها و ارتباطات به‌کار می‌رود منتشر کند. او نخست تعداد ارقام دودویی لازم برای نشان دادن یک عدد صحیح را آزمود. با توجه به اینکه زبان معمول کامپیوتر یک سیستم دودویی است، بنابراین سؤالی که به ذهن شانون خطور کرد این بود: «برای ذخیره هر عدد صحیح چه تعداد رقم باینری لازم است؟». او سپس اعلام کرد که تعداد رقم باینری لازم برای ذخیرهٔ اعدادی که توان‌های صحیحی از 2 هستند یعنی، 2^n، برابر $n+1$ می‌باشد، ولی اگر عدد بزرگتر یا مساوی 2^{n-1} وکوچکتر از 2^n باشد، n بیت لازم است. مثلا عدد 105 در یک سیستم باینری به صورت زیر قابل نمایش است:

$$11\mathbf{0}1\mathbf{0}\mathbf{0}1 = 1\times 2^6 + 1\times 2^5 + \mathbf{0}\times 2^4 + 1\times 2^3 + \mathbf{0}\times 2^2 + \mathbf{0}\times 2^1 + 1\times 2^{\mathbf{0}}$$

[6] Leo Szilard

شانون برای راحتی کلمهٔ «بیت» را به جای «رقم باینری» معرفی کرد. در نظریهٔ شانون بیت واحد اطلاعات است. در واقع یک بیت مقدار اطلاعاتی است که یک سیستم دودویی کلاسیکی می‌تواند حمل کند. او تعداد بیت‌های لازم برای بیان یک عدد صحیح اختیاری Ω را به صورت زیر بیان کرد:

$$I = \log_2 \Omega \ \text{bits} = \frac{\ln \Omega}{\ln 2} \ \text{bits} = 1{,}442695 \ln \Omega \ \text{bits} \qquad (7\text{-}1)$$



شکل (1-5)، استفاده از اطلاعات برای انجام کار.

شکل (1-6)، استخراج کار از یک گاز تک‌مولکولی، که در آغاز در سمت چپ یا راست ظرف می‌باشد.

یا

$$I = k \ln \Omega \qquad (8\text{-}1)$$

فرض کنید سیستمی در اختیار داریم که دانش ما دربارهٔ آن کامل نیست، یا اطلاعات ما درمورد این سیستم ناکامل است.[7] شانون قصد داشت قبل از انجام دادن آزمایش مقدار اطلاعاتی را که ممکن است از سیستم به‌دست آید به صورت کمّی محاسبه کند. یعنی نتیجه‌ای که ما برای پیشگویی آن با قطعیت نامطمئن هستیم. (برای مثال ممکن است آزمایش پرتاب سکه باشد). شانون در آغاز اطلاعات مفقود شده یا اطلاعاتی را که از آزمایش به‌دست می‌آوریم برای N آزمایش مستقل و نه برای یک تک آزمایش محاسبه کرد. فرض کنید در یک تک آزمایش احتمال رخ دادن یک نتیجه ویژه i برابر با p_i

[7] در بعضی منابع گفته می‌شود که اطلاعات این سیستم از بین رفته است.

باشد. اگر آزمایش N بار تکرار شود، هنگامی که N خیلی بزرگ می‌شود، تعداد دفعاتی که نتیجهٔ i رخ می‌دهد، بیشتر و بیشتر به p_i نزدیک می‌شود. برای مثال اگر یک سکه N بار پرتاب شود، هنگامی که N فوق‌العاده بزرگ می‌شود تعداد دفعاتی که شیر می‌آید بیشتر و بیشتر به $1/2$ نزدیک می‌شود. اگرچه بعضی اطلاعات هنوز مفقود است، چون ما هنوز دنبالهٔ حاصل از نتایج را نمی‌دانیم. شانون نشان داد که این اطلاعات از دست رفته دربارهٔ دنبالهٔ نتایج از معادله زیر به‌دست می‌آید:

$$I_N = K \ln \Omega \tag{9-1}$$

که

$$\Omega = \frac{N!}{n_1! n_2! n_3! \ldots n_i! \ldots} \tag{10-1}$$

یا

$$I_N = K \ln \Omega = K \left[\ln (N!) - \sum_i \ln (n_i!) \right] \tag{11-1}$$

با استفاده از تقریب استرلینگ، $\ln(n_i!) \approx n_i (\ln n_i - 1)$، به نتیجهٔ زیر می‌رسیم:

$$I_N = -K N \sum_i P_i \ln P_i \tag{12-1}$$

سرانجام با تقسیم بر N به نتیجهٔ مطلوب می‌رسیم. که مقدار اطلاعات کمی نهفته در سیستم است که شانون قبل از انجام هر گونه آزمایشی به‌دست آورد.

هنگامی که شانون روی معادله‌اش کار می‌کرد به طور اتفاقی با فون‌نیومن ملاقات کرد که به او پیشنهاد کرد به جای «اطلاعات از دست رفته» نام «آنتروپی اطلاعات» را برای معادله خود برگزیند. شانون نیز پذیرفت و معادله‌اش «آنتروپی شانون» نام گرفت. ایده شانون همچنین می‌تواند برای ترمودینامیک نیز به‌کار گرفته شود.

1-3-2- بیان آنتروپی به شکل اطلاعات از دست رفته

از نقطه نظر نظریهٔ اطلاعات، برای آنسامبلی از N سیستم مشابه در یک ماکروحالت با برهمکنش خیلی ضعیف، می‌توان آنتروپی ترمودینامیکی S_N را به‌شکل اطلاعات مفقود شدهٔ مورد نیاز برای

مشخص کردن حالت هر سیستم، یعنی میکروحالت‌های آنسامبل، تفسیر کرد. با استفاده از فرمول شانون می‌توانیم این اطلاعات از دست رفته را به طور کمی اندازه بگیریم. با به‌کار بردن فرمول شانون، معادله (1-12)، برای اطلاعات از دست رفته درآمار بولتزمن، می‌توانیم هویت W را با Ω، و S_N را با I_N، و k را با K تعیین کنیم.

$$W \to \Omega \qquad S_N \to I_N \qquad k \to K = \frac{1}{\ln 2}\ \text{bit} \tag{1-13}$$

چنانکه

$$k \ln 2 = 1\ \text{bit} = 0{,}95697 \times 10^{-23}\ \frac{\text{joule}}{\text{kelvin}} \tag{1-14}$$

$$k = 1{,}442695\ \text{bits}$$

$$1\ \text{degree Kelvin} = 0{,}95697 \times 10^{-23}\ \frac{\text{joule}}{\text{bit}} \tag{1-15}$$

معادلهٔ بالا نشان می‌دهد که دما بُعد انرژی بر بیت دارد. در نتیجه:

$$1 \frac{\text{joule}}{\text{kelvin}} = 1{,}04496 \times 10^{23}\ \text{joule} \tag{1-16}$$

اگر معادله بالا را بر عدد آووگادرو تقسیم کنیم، داریم:

$$1\ \frac{\text{joule}}{\text{kelvin}} = \frac{1{,}04496 \times 10^{-23}\ \text{bits/mol}}{6{,}02217 \times 10^{-23}\ \text{molecule/mol}} = 0{,}17352\ \frac{\text{bits}}{\text{molecule}}. \tag{1-17}$$

آنتروپی که به طور آزمایشی از آمونیاک به‌دست آمده به عنوان تابعی از دما بر حسب کلوین در شکل (1-7) نمایش داده شده است. معمول است که آنتروپی بر حسب $\text{joule} / \text{Kelvin-mol}$ بیان شود؛ اما از معادله قبل برداشت می‌شود که می‌توان آنتروپی را بر حسب $\text{bit} / \text{molecule}$ نیز بیان کرد.

چون

$$1\ \text{electron volt} = 1{,}6023 \times 10^{-19}\ \text{joule} \tag{1-18}$$

همچنین از معادله (1-17) نتیجه می‌گیریم که:

$$1\frac{\text{electron volt}}{\text{Kelvin}} = 1{,}6743 \times 10^{4} \text{ bits} \qquad (19\text{-}1)$$

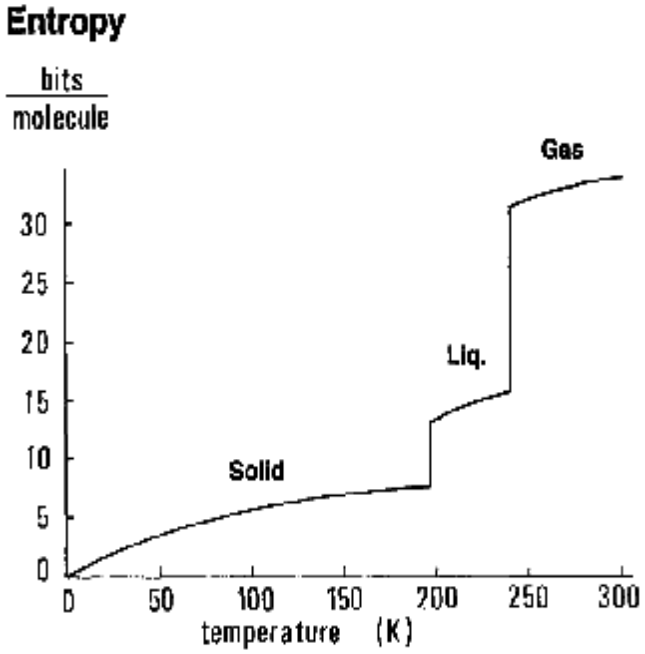
بنابراین یک الکترون- ولت انرژی در دمای اتاق $T = 298{,}15\ ^{\circ}K$، به گرما تبدیل شده و تغییر آنتروپی زیر ایجاد خواهد شد:

$$\frac{1 \text{ electron volt}}{298{,}15 \text{ kelvin}} = 56{,}157 \text{ bits} \qquad (20\text{-}1)$$

هر گاه سیستم در تعادل ترمودینامیکی باشد آنتروپی به مقدار بیشینه خود می‌رسد؛ اما اگر سیستم در تعادل ترمودینامیکی نباشد آنتروپی مقدار کمی دارد. برای مثال موردی را که گاوس هنگام معرفی مفهوم آنتروپی بررسی کرد در نظر می‌گیریم؛ گاوس یک سیستم ایزوله در نظرگرفت و آن را به دو بخش تقسیم کرد. یکی را در دمای T_1 و دیگری را در دمای پایین‌تر یعنی T_2 قرار داد. گرما از محیط گرم به محیط سرد منتقل می‌شود و آنتروپی افزایش می‌یابد. سرانجام هنگامی که سیستم در دمایی یکنواخت به تعادل برسد آنتروپی به مقدار بیشینه می‌رسد. تفاوت آنتروپی بین حالت اولیه سیستم گاوس و حالت نهایی اندازه‌ای است از اینکه سیستم چقدر از تعادل ترمودینامیکی آغازین، فاصله گرفته است. از بحث بالا می‌توان تفاوت آنتروپی را به عنوان محتوای اطلاعات ترمودینامیکی اولیه سیستم تفسیر کرد.

موردی مشابه را در نظر می‌گیریم که فوتونی از خورشید در قطرهٔ آبی بر روی زمین جذب می‌شود. انرژی فوتون بین مولکولهای آب تقسیم شده و در نتیجه دمای آب اندکی افزایش می‌یابد. بنابراین آنتروپی اولیهٔ سیستم آب و فوتون کمتر از حالتی است که فوتون در آب جذب شده است. این تفاوت آنتروپی را می‌توان به‌عنوان مقدار اطلاعات ترمودینامیکی تفسیر کرد که ابتدا سیستم قطره فوتون شامل می‌شده است. ولی زمانی که انرژی آزاد فوتون به گرما تغییر شکل داد این اطلاعات از دست رفته است. با استفاده از معادله (19-1) می‌توانیم تفاوت آنتروپی را برحسب بیت بیان کنیم. برای مثال اگر انرژی فوتون 2 الکترون ولت و دمای آب 15,289 درجه کلوین باشد، پس $\Delta S = 11{,}231\ bits$ می‌باشد. این مقدار اطلاعات ترمودینامیکی، در حالت اولیه سیستم موجود است.

اما اگر فوتون به برگ گیاهی برسد، به جای اینکه فوراً کاهش یابد به شکل قیدهای شیمیایی انرژی بالا تثبیت می‌شود. بنابراین دراین حالت قسمتی از انرژی فوتون تثبیت می‌شود و همه اطلاعات



شکل(1-7). آنتروپی که به طور آزمایشی از آمونیاک به دست آمده و به عنوان تابعی از دما بر حسب کلوین رسم شده. است

ترمودینامیکی که سیستم دربردارد، از بین نمی‌رود[1].

فصل دوم

اطلاعات کلاسیکی و آنتروپی شانون

مقدمه

آنتروپی شانون

تعاریفی مهم بر پایه آنتروپی شانون

متراکم سازی اطلاعات در سیستم‌های کلاسیکی (متراکم سازی رقمی)

نتیجه‌گیری

مقدمه

شانون با مقاله‌ای که در زمینه نظریهٔ اطلاعات ارائه داد خود را یکی از تأثیرگذارترین افراد در این زمینه معرفی کرد. نتایج مطالعات شانون در زمینه کامپیوترهای کلاسیکی و نظریهٔ اطلاعات هنوز هم اهمیت بسزایی در حل مشکلات این شاخه از علوم و مهندسی دارد. در این فصل در مورد آنتروپی شانون به عنوان رابطه‌ای که محتوای اطلاعاتی یک سیستم کلاسیکی را تعیین می‌کند صحبت کرده و ویژگی‌های مهم آن را بر می‌شماریم. همچنین نشان می‌دهیم که این آنتروپی اهمیت اساسی و بنیادین در تعیین ظرفیت کانال و آهنگ انتقال اطلاعات در کانال دارد.

2-1- آنتروپی شانون

2-1-1- وابستگی آنتروپی شانون به توزیع احتمال در تعیین محتوای اطلاعاتی

گفتیم که شانون را پدر علم اطلاعات می‌نامند و مفهوم کلیدی نظریهٔ اطلاعات کلاسیکی آنتروپی شانون می‌باشد. فرض کنید مقدار متغیر تصادفی X را می‌دانیم، آنتروپی شانون X میانگین مقدار اطلاعات به‌دست آمده از متغیر تصادفی X را به صورت کمی مشخص می‌کند. یا از دیدی دیگر آنتروپی شانون میزان عدم قطعیت دربارهٔ X پیش از دانستن مقدار آن را مشخص می‌کند. این دو دیدگاه مکملند[5].

محتوای اطلاعاتی یک متغیر تصادفی نباید به برچسب‌های مختلفی که آن متغیر می‌تواند بگیرد بستگی داشته باشد، و اگر یک متغیر تصادفی را به عنوان نمادی برای نشان دادن توزیع‌های احتمالی متفاوت بکار ببریم اطلاعات کسب شده نباید به نوع آن متغیر وابسته باشد. مثلاً متغیر تصادفی که می‌تواند شیر یا خط را با احتمال‌های 1/4,3/4 بگیرد، همان مقدار اطلاعاتی را در بر دارد که اگر مقادیر **0** یا 1 را با احتمال‌های 1/4,3/4 اختیار می‌کرد. به همین دلیل بنا به تعریف، آنتروپی متغیر تصادفی تابعی از احتمال مقادیر ممکن و مختلف است که متغیر تصادفی می‌تواند اختیار کند. اغلب آنتروپی به صورت تابعی از توزیع احتمال $p_1,\ldots\ldots,p_n$ نوشته می‌شود. شانون نشان داد که آنتروپی متناسب با این توزیع احتمال را می‌توان به صورت زیر تعریف کرد [5 و6]:

$$H(X) \equiv H(p_1,...,p_n) \equiv -\sum_x p_x \log p_x \qquad (1\text{-}2)$$

شانون برای سازگاری رابطه‌اش با سیستم‌های دودویی پیشنهاد کرد که مبنای لگاریتم 2 در نظر گرفته شود. ممکن است مورد $p_x = 0$ (چون $0\log 0$ تعریف نشده) تعجب برانگیز باشد. به طور شهودی حادثه‌ای که هیچ‌گاه اتفاق نمی‌افتد سهمی در آنتروپی ندارد، بنابراین با قرارداد توافق می‌کنیم که $0\log 0 \equiv 0$. یا به صورت فرمولی‌تر $\lim_{x\to 0} x\log x = 0$ [5] از رابطه شانون مشخص است که مقدار ماکزیمم $H(X)$ هنگامی است که همه خروجی‌های یک رویداد متساوی‌الاحتمال باشند. در این صورت ناآگاهی ما از سیستم بیشینه است. و در صورتی که همه احتمال‌ها به جز یکی صفر باشند عدم یقین ما از سیستم کمینه بوده و برابر صفر می‌شود که امری کاملاً طبیعی است. اکنون رویدادی مانند X را در نظر بگیرید که پیشامدها یا نتایج ممکن آن با مجموعهٔ $\{x_1, x_2, ..., x_n\}$ نشان داده می‌شود. اگر نتیجه حاصل از رخ دادن این رویداد به ما گزارش شود، مثلاً به طریقی بفهمیم که نتیجه x_i رخ داده است، در این صورت می‌توان پرسید که چقدر اطلاعات کسب کرده‌ایم یا چه اندازه از ناآگاهی ما کاسته شده است. کاملاً قابل درک است که هرچه پیش‌آمدی که رخ داده محتمل‌تر باشد اطلاعاتی که کسب کرده‌ایم کمتر و هرچه آن اتفاق دور از انتظار باشد، تعجب ما از وقوع آن بیشتر و در نتیجه اطلاعی که کسب کرده‌ایم بیشتر است. بنابراین میزان اطلاعات کسب شده نسبتی معکوس با احتمال رخ دادن آن پیش‌آمد دارد. این نتایج با رابطه شانون به کمال سازگار است.

2-1-2- اهمیت تعریف آنتروپی شانون به شکل کنونی

چرا آنتروپی اطلاعات به این روش تعریف می‌شود و یا به عبارتی چرا این رابطه را بسیاری پذیرفته‌اند؟ این فصل را با تأکید فراوان بر مناسب و مفید بودن این رابطه ادامه می‌دهیم. اما در اینجا برای پاسخ به این پرسش دو دلیل ساده به طور خلاصه ذکر می‌کنیم که می‌توانند تا حدی روشنگر باشند:

1) تعریف آنتروپی به این شکل در سیستم‌های دودویی کمترین فضای لازم برای ذخیرهٔ اطلاعات را محاسبه می‌کند[8] و اگر بخواهیم به هر روشی فرمولی ارائه کنیم که فضای کمتری برای ذخیره اطلاعات به دست دهد مطمئناً مقداری از اطلاعات را از دست خواهیم داد[5 و 6].

2) دلیل دیگر برای تعریف آنتروپی به صورت گفته شده در واقع توجیهی شهودی است که بیان می‌کند بهترین تعریف برای آنتروپی اطلاعات همان است که شانون ارائه داده است. فرض کنید هدف تعیین مقدار اطلاعاتی است که رویداد احتمالی E حمل می‌کند. اگر این حجم اطلاعات با $I(E)$ نشان داده شود، باید فرض‌های زیر دربارهٔ $I(E)$ صادق باشد:

الف) چون از هر رویداد E تنها احتمال رخ دادن آن را می‌دانیم پس در این مورد باید $I(E)$ها تنها تابعی از احتمال رویدادها باشد. بنابراین I ممکن است به صورت $I = I(p)$ نوشته شود.

ب) I باید تابعی هموار از احتمال‌ها باشد.

ج) هرگاه $p > 0$ و q، دو توزیع احتمال مفروض باشند باید $I(q\,p) = I(p) + I(q)$ برقرار باشد. (تفسیر: اطلاعات به دست آمده از دو رویداد با توزیع احتمال‌های منحصربفرد p و q، با مجموع اطلاعات حاصل از هر رویداد برابر است.)

از فرض‌های بالا کاملاً مشخص است که بهترین تابعی که شرط‌های بالا را برآورده می‌کند به صورت $I(p) = k \log p$ با مقدار ثابت k می‌باشد. دنباله‌ای از رویدادها با احتمال‌های $p_1, \ldots\ldots, p_n$ را درنظر می‌گیریم، پس مقدار اطلاعاتی که می‌توانیم از این رویدادها به‌دست آوریم به صورت $k \sum_i p_i \log p_i$

است. این رابطه درست همان آنتروپی شانون است که در ثابت k ضرب شده است.

2-2- تعاریفی مهم بر پایهٔ آنتروپی شانون

2-2-1- آنتروپی باینری

[8] این مطلب در بخش‌های آتی در نظریه کد کردن کانال بدون نوفه شانون اثبات می‌شود.

آنتروپی باینری مربوط به یک متغیر تصادفی است که می‌تواند دو خروجی اختیار کند. این آنتروپی چنان اهمیت بسزایی دارد که نام ویژهٔ آنتروپی باینری برای آن برگزیده شده است و به صورت زیر تعریف می‌شود:

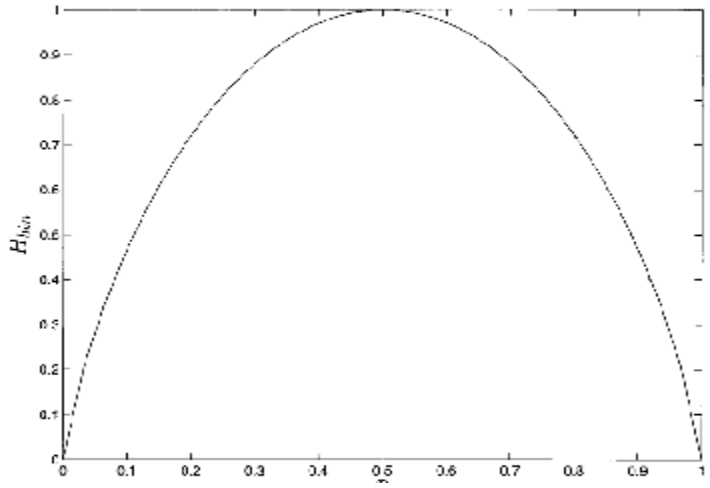
$$H_{bin} \equiv -p\log p-(1-p)\log(1-p) \qquad (2\text{-}2)$$

که p و $1-p$ مربوط به احتمال‌های ممکن دو خروجی می‌باشند. آنتروپی باینری در $p=1/2$ به مقدار بیشینهٔ خود می‌رسد شکل(2-1). بنابراین با متساوی‌الاحتمال شدن رویدادها عدم یقین ما از سیستم بیشینه شده است. این آنتروپی آزمایشی عالی برای فهمیدن بعضی ویژگیهای کلی آنتروپی می‌باشد. یک ویژگی جالب توجه آنتروپی باینری این است که نشان می‌دهد هنگامی که دو توزیع احتمال را مخلوط می‌کنیم رفتار کلی آنتروپی چگونه است. برای مثال تصور کنید که آلیس دو سکه در اختیار دارد، یکی سکه 25 سنتی آمریکا و دیگری 1 دلاری استرالیا. هر سکه طوری تغییر داده شده است که احتمال آمدن یک طرف آن بیشتر است. احتمال آمدن شیر برای سکه آمریکایی با p_U و برای سکه استرالیایی با p_A نشان داده می‌شود. سپس فرض کنید آلیس سکه آمریکایی را با احتمال q و سکه استرالیایی را با احتمال $1-q$ پرتاب می‌کند، و نتیجهٔ حاصل را که شیر آمده است یا خط به باب گزارش می‌دهد. باب به طور متوسط چقدر اطلاعات کسب می‌کند؟ روشن است که اطلاعات کسب شده برای باب باید حداقل برابر با میانگین اطلاعاتی باشد که از پرتاب سکهٔ آمریکایی یا استرالیایی به دست می‌آورد. این قضیه را می‌توان به‌صورت معادله زیر بیان کرد:

$$H\left(q\,p_U+(1-q)p_A\right)\geq q\,H\left(p_U\right)+(1-q)H\left(p_A\right) \qquad (3\text{-}2)$$

گاهی اوقات ممکن است نامساوی اکیداً برقرار باشد، یعنی مساوی آن برداشته شود. چون باب نه تنها اطلاعاتی دربارهٔ مقدار سکه به‌دست می‌آورد، بلکه اطلاعات اضافی‌ای نیز دربارهٔ هویت سکه کسب می‌کند. برای مثال اگر $p_U=1/6$ و $p_A=5/6$ باشد پس باب با دقت خوبی پی می‌برد که سکه

استرالیایی بوده است. ویژگی دیگر آنتروپی باینری آن است که یک تابع مقعر است[9] .که این ویژگی کاربردهای زیادی هم در کلاسیک و هم در کوانتوم دارد و تعداد زیادی از عمیق‌ترین نتایج در آنتروپی‌های کلاسیکی یا کوانتومی ریشه در کاربرد ماهرانهٔ این ویژگی دارد[5].



شکل(2-1) تابع آنتروپی باینری $H(p)$

2-2-2-آنتروپی نسبی[10]

آنتروپی نسبی اندازه‌ای است با رفتاری همانند آنتروپی و بسیار مفید می‌باشد. این کمیت میزان نزدیکی دو توزیع احتمال $p(x)$ و $q(x)$ را که روی اندیسی یکسان مانند x تعریف شده‌اند، نشان می‌دهد. فرض کنید $p(x)$ و $q(x)$ دو توزیع احتمال روی اندیسی یکسان، مانند x باشند. آنتروپی نسبی $p(x)$ و $q(x)$ به صورت زیر تعریف می‌شود:

$$H(p(x)\|q(x)) \equiv \sum_x p(x)\log\frac{p(x)}{q(x)} \equiv -H(x) - \sum_x p(x)\log q(x) \qquad (4\text{-}2)$$

که تعریف می‌کنیم (با شرط $p(x)\geq 0$) $p(x)\log 0 \equiv +\infty$ و $0\log 0 \equiv 0$.

قضیه1. نامنفی بودن آنتروپی نسبی: آنتروپی نسبی یک کمیت نامنفی است، $H(p(x)\|q(x)) \geq 0$. و به ازای هر x اگر و تنها اگر $p(x)=q(x)$ باشد نامساوی تبدیل به مساوی می‌شود.

---

[9] تابع f مقعر گفته می‌شود اگر برای هر p در بازهٔ $0,1$ داشته باشیم:
$f(px+(1-p)y) \geq pf(x)+(1-p)f(y)$.

[10] Relative entropy

اثبات:

یک نامساوی خیلی مفید در نظریهٔ اطلاعات[1] $\log x \ \ln 2 = \ln 2^{\log x} = \ln x \le x - 1$ می‌باشد، که x همواره مثبت است. و اگر $x = 1$ باشد نامساوی به مساوی تبدیل می‌شود. از نامساوی بالا داریم: $-\log x \ge (1-x)/\ln 2$. پس می‌توان نوشت:

$$H(p(x)\|q(x)) = -\sum_x p(x)\log\frac{p(x)}{q(x)} \ge \frac{1}{\ln 2}\sum_x p(x)\left(1-\frac{q(x)}{p(x)}\right)$$
$$= \frac{1}{\ln 2}\sum_x (p(x)-q(x)) = \frac{1}{\ln 2}(1-1) = \mathbf{0} \qquad (5\text{-}2)$$

که نتیجه مطلوب است و نامساوی کامل می‌شود.

از آنتروپی نسبی می‌توان زیرافزایشی[11] بودن آنتروپی شانون را نتیجه گرفت:

نشان می‌دهیم که :

$$H(p(x,y)\|p(x)p(y)) = H(p(x)) + H(p(y)) - H(p(x,y)) \qquad (6\text{-}2)$$

$$H(p(x,y)\|p(x)p(y)) \equiv$$
$$\sum_x p(x,y)\log\frac{p(x,y)}{p(x)p(y)} \equiv \sum_{x,y} p(x,y)\log p(x,y)$$
$$-\sum_{x,y} p(x,y)\log p(x) - \sum_{x,y} p(x,y)\log p(y) \qquad (7\text{-}2)$$
$$= H(p(x)) + H(p(y)) - H(p(x,y)) \ge \mathbf{0} \Rightarrow$$
$$H(X) + H(Y) \ge H(X,Y)$$

که نتیجه آخر زیر افزایشی بودن آنتروپی شانون نامیده می‌شود.

2-2-3- آنتروپی شرطی و اطلاعات دوجانبه[12]

[1] subadditivity

1 Conditional entropy and Mutual information

X و Y را به عنوان دو متغیر تصادفی در نظر بگیرید. محتوای اطلاعاتی X و Y چگونه به هم مربوط می‌شوند؟ دو مفهوم آنتروپی شرطی و اطلاعات دوجانبه ما را در پاسخ دادن به این پرسش کمک می‌کنند.

آنتروپی توأم[13] یک جفت از متغیرهای تصادفی به صورت $H(X,Y)$ تعریف می‌شود که در واقع میزان اطلاعاتی است که از این دو متغیر تصادفی به دست می‌آوریم، یا عدم قطعیت کل دو متغیر تصادفی X و Y را مشخص می‌کند. همان گونه که پیش از این گفته شد به‌صورت زیر تعریف می‌شود:

$$H(X,Y) = -\sum_{x,y} p(x,y)\log p(x,y) \qquad (8\text{-}2)$$

دو متغیر تصادفی X و Y را در نظر بگیرید، آنتروپی شرطی به طور متوسط میزان عدم یقین ما را دربارهٔ X، به‌شرط دانستن مقدار Y، مشخص می‌کند.

$$H(X \mid Y) \equiv H(X,Y) - H(Y) \qquad (9\text{-}2)$$

محتوای اطلاعات دوجانبه[14] میزان اطلاعات مشترک X و Y را تعیین می‌کند و به صورت زیر تعریف می‌شود:

$$H(X:Y) \equiv H(X) + H(Y) - H(X,Y) \qquad (10\text{-}2)$$

قضیهٔ2. خصوصیات اصلی آنتروپی شانون:

$$(1)\ H(X,Y) = H(Y,X),\ H(X:Y) = H(Y:X). \qquad (11\text{-}2)$$

$$(2)\ H(Y \mid X) \geq 0 \Rightarrow H(X:Y) \leq H(Y) \qquad (12\text{-}2)$$

اگر و تنها اگر Y تابعی از X باشد نامساوی در معادله (2-12) به مساوی تبدیل می‌شود.

$$(3)\ H(X) \leq H(X,Y) \qquad (13\text{-}2)$$

$$(4)\ H(X) + H(Y) \geq H(X,Y) \qquad (14\text{-}2)$$

[13] joint entropy

[14] mutual information content

رابطه بالا همان قضیهٔ زیرافزایشی ($H(X)+H(Y)\geq H(X,Y)$) است که قبلاً ذکر شد.

(5) $H(Y\mid X)\leq H(Y)\Rightarrow H(X:Y)\geq 0$ (15-2)

(6) $H(X,Y,Z)+H(Y)\leq H(X,Y)+H(Y,Z).$ (16-2)

رابطه بالا را زیرافزایشی قوی می‌نامند. اگر $Z\rightarrow Y\rightarrow X$، یعنی این سه متغیر یک زنجیره مارکوف[15] تشکیل دهند (حادثه X تنها به Y وابسته بوده و مستقل از Z باشد و حادثه Y تنها به Z وابسته باشد)، نامساوی تبدیل به مساوی می‌شود.

(7) $H(X|Y,Z)\leq H(X|Y).$ (17-2)

که آنتروپی شرطی کاسته شده[16] نامیده می‌شود.

قضیه3. قانون حلقه‌ای برای آنتروپی شرطی: فرض کنید $X_1,\mathbf{K},X_n$ و Y یک سری از متغیرهای تصادفی باشند. پس:

$$H(X_1,\mathbf{K},X_n\mid Y)=\sum_{i=1}^{n}H(x_i\mid Y,X_1,\mathbf{K},X_{i-1})$$ (18-2)

[15] Markov chain

[16] conditioning reduces entropy

2-2-4- نامساوی پردازش داده (اطلاعات)[17]

در بسیاری از کاربردهای مورد نظر، محاسباتی را بر روی اطلاعات در دسترس انجام می‌دهیم. اما آن اطلاعات کامل نیستند و پیش از آنکه در اختیار ما قرار گیرد در معرض نوفه[18] واقع شده‌اند. بنا به نامساوی اساسی نظریهٔ اطلاعات (نامساوی پردازش داده)، اطلاعات دربارهٔ خروجی منبع تنها ممکن است با زمان کاهش یابد. یک بار که اطلاعاتی گم شود برای همیشه از دست رفته است و هیچ‌گاه قابل بازگشت نیست.

مفهوم شهودی پردازش اطلاعات در دل زنجیرهٔ مارکوف متغیرهای تصادفی قرار دارد[5]. ویژگی مارکوف در مبحث فرآیندهای تصادفی، ویژگی فرآیندهایی است که احتمال شرطی رخداد آینده فقط به آخرین رخداد موجود یعنی رخداد کنونی وابسته باشد و نه به دیگر رویدادهای گذشته.

به زبان ریاضی اگر $X(t)$، $t>\mathbf{0}$یک فرآیند تصادفی با ویژگی مارکوف باشد داریم:

$$p(X(t+h)=y \,|X(s)=x(s), \forall s \le t)$$
$$= p\,(X(t+h)=y \,|X(t)=x(t)), \quad \forall h>\mathbf{0} \qquad (19\text{-}2)$$

یا به عبارتی دیگر در فرآیندهای گسسته زمانی، اگر فرآیندی مثل $\{X_n\{n\in N\}\}$دارای ویژگی مارکوف باشد، آنگاه:

$$p\{X_{n+1}=x_{n+1} \,|\, X_n=x_n,\ldots, X_1=x_1\}=p\{X_{n+1}=x_{n+1} \,|\, X_n=x_n\} \qquad (20\text{-}2)$$

پس فرآیندهای این چنینی را زنجیره مارکوف خطاب می‌کنند [5 و7].

بنابراین اگر دنبالهٔ.... $X_1 \to X_2 \to$ از متغیرهای تصادفی یک زنجیره مارکوف تشکیل دهد پس X_{n+1} مستقل از $X_1,\ldots\ldots\ldots\ldots,X_{n-1}$ و تنها به X_n وابسته است.

[17] Data processing inequality

[18] Noise

تحت چه شرایطی در زنجیرهٔ مارکوف با پیشروی زمان اطلاعات از دست می‌رود؟ نامساوی پردازش داده (اطلاعات) که در زیر آمده است، روشی است مبتنی بر نظریهٔ اطلاعات که به این پرسش پاسخ می‌دهد.

قضیه4. نامساوی پردازش اطلاعات: فرض کنید $X \to Y \to Z$ یک زنجیره مارکوف باشد، آنگاه

$$H(X) \geq H(X:Y) \geq H(X:Z) \qquad (21\text{-}2)$$

علاوه براین، اگر و تنها اگر با Y داده شده امکان بازسازی X وجود داشته باشد، نامساوی اول اشباع می‌شود. این نتیجه به طور شهودی پذیرفتنی است. بنا بر این نتیجه، اگر متغیر تصادفی X در معرض نوفه قرار گیرد و متغیر Y ایجاد شود، آنگاه ممکن نیست که از اعمال بعدی در فرآیند پردازش اطلاعات برای افزایش میزان اطاعات دوجانبه بین خروجی این فرآیند و اطلاعات ورودی X استفاده کرد. در زیر اثباتی هر چند خلاصه برای این موضوع داده می‌شود.

نامساوی اول کاملاً واضح است زیرا همواره محتوای اطلاعاتی یک متغیر تصادفی بیشتر از یا مساوی با اطلاعات مشترک بین این متغیر و هر متغیر اختیاری دیگری است. از تعریف‌های ارائه شده در قضیه 2 درمی‌یابیم که $H(X:Z) \leq H(X:Y)$ با $H(X\ |Y\ ) \leq H(X\ |\ Z)$ معادل است. با استفاده از این امر که $X \to Y \to Z$ یک زنجیرهٔ مارکوف تشکیل می‌دهند می‌توان نتیجه گرفت که $Z \to Y \to X$ نیز یک زنجیرهٔ مارکوف می‌باشد، و بنابراین $H(X|Y) = H(X|Y,Z)$. در نتیجه مسأله به اثبات

$$H(X,Y,Z) - H(Y,Z) = H(X|Y,Z) \leq H(X|Z) = H(X,Z) - H(Z) \qquad (22\text{-}2)$$

تبدیل می‌شود این همان نامساوی زیر افزایشی قوی است که پیش از این ثابت شد.

فرض کنید $H(X:Y) < H(X)$. پس بازسازی X از Y ممکن نیست، زیرا اگر Z تلاش برای بازسازی تنها بر اساس آگاهی از Y پایه شود، پس $X \to Y \to Z$ باید یک زنجیرهٔ مارکوف باشد، و بنابراین به وسیلهٔ نامساوی پردازش اطلاعات $H(X) > H(X:Y)$ است. از این رو $Z \neq X$ است. از

سوی دیگر اگر $H(X:Y)=H(X)$ پس $H(X|Y)=0$ و بنابراین هرگاه $p(X=x,Y=y)>0$ باشد، $p(X=x|Y=y)=1$ است. اگر $Y=y$ می‌توان با قطعیت استنتاج کرد که X با x برابر بوده و بازسازی X مجاز می‌باشد [5].

2-3- متراکم سازی اطلاعات در سیستم‌های کلاسیکی (متراکم کردن رقمی)

متراکم کردن اطلاعات یک فرایند دینامیکی ابتدایی است که هم در نظریه اطلاعات کلاسیکی و هم در نظریه اطلاعات کوانتومی کاربرد دارد. کلی‌ترین صورت مسأله متراکم سازی اطلاعات تعیین کمینه دربایست‌های فیزیکی لازم برای ذخیره یک منبع فیزیکی است. در واقع تعیین کمینه منبع فیزیکی در ذخیره اطلاعات یکی از مسائل اساسی نظریه اطلاعات می‌باشد. در هر دو نظریه اطلاعات کلاسیکی و کوانتمی تکنیک‌هایی در حل این مسأله بکار می‌روند که تنها برای متراکم کردن اطلاعات بکار نمی‌رود و دامنه وسیعی از کاربردها را در بر می‌گیرد. در حوزه فیزیک کلاسیک شانون با تلاش-هایی که در زمینه نظریهٔ اطلاعات انجام داد توانست پاسخ بسیار مناسبی برای این تقاضا ارائه دهد که هنوز هم پس از گذشت سال‌ها از ارزش علمی و عملی آن کاسته نشده است. در این بخش به طور اجمالی نتایج شانون را در این زمینه از نظر خواهیم گذراند.

2-3-1- نظریه کد کردن کانال بدون نوفه شانون

نظریه کد کردن کانال بدن نوفه شانون تعیین می‌کند که اطلاعات کلاسیکی به دست آمده از منبعی کلاسیکی را تا چه حد می‌توان فشرده کرد؟ گونه‌های زیادی برای چنین منبعی وجود دارد. یک مدل مناسب و خیلی ساده منبعی است که شامل دنباله‌ای از متغیرهای تصادفی $X_1, X_2, \mathbf{K}$ می‌باشد. این مقادیر خروجی منبع را نمایش می‌دهند. ما این منبع را با این فرض در نظر می‌گیریم که مقادیری از نمادهای حروف الفبای محدود را اتخاذ می‌کند. اگرچه هنگامی که حروف نامحدود می‌شوند هنوز هم می‌توان از این منبع فرضی استفاده کرد. علاوه بر این فرض می‌کنیم که استفاده-های متفاوت از منبع به طور مستقل و یکسان صورت می‌گیرد. چنین منبعی با نام منبع اطلاعات

i.i.d[19] شناخته می‌شود. منبع‌های واقعی اغلب به این شکل رفتار نمی‌کنند. این رفتار را می‌توان به خوبی از متون انگلیسی که به عنوان منبع استفاده می‌شوند فهمید. حرف‌ها در متون انگلیسی به طور مستقل ظاهر نمی‌شوند و بین چگونگی اتفاق افتادن آنها رابطه‌ای قوی وجود دارد. مثلاً تعداد دفعاتی که حرف h بعد از حرف t ظاهر می‌شود خیلی بیشتر از تعداد دفعاتی است که حرف h به تنهایی می‌آید. بنابراین می‌توان گفت که t و h مستقل از هم اتفاق نمی‌افتند. با وجود این فرضیه منبع اطلاعاتی i.i.d برای گستره وسیعی از منابع اطلاعاتی به خوبی کار می‌کند. و ایده‌های معرفی شده برای منابع اطلاعاتی i.i.d را می‌توان برای منابع پیچیده‌تر نیز بکار برد.

پیش از بررسی جزئیات تکنیکی نظریه کد کردن بدون نوفه شانون از مثال ساده‌ای برای درک نتیجه پیشین استفاده می‌کنیم. فرض کنید یک منبع اطلاعاتی i.i.d بیت‌های $X_1, X_2, \mathbf{K}$ را تولید می‌کند. این بیت‌ها با احتمال p برابر **0** و با احتمال $1-p$ برابر 1 می‌باشد. ایدهٔ کلیدی پشت نظریه شانون این است که دنباله‌های ممکن $x_1, \mathbf{K}, x_n$ برای متغیرهای تصادفی $X_1, \mathbf{K}, X_N$ را به دو دسته تقسیم کنیم. یکی دنباله‌هایی که با احتمال بیشتری اتفاق می‌افتد و آنها را دنباله‌های بهنجار[20] می‌نامند و دیگری دنباله‌هایی که به ندرت اتفاق می‌افتند و آنها را دنباله‌های نابهنجار[21] می‌نامند. اکنون ممکن است این پرسش پیش آید که چگونه این اتفاق می‌افتد؟ انتظار می‌رود که با بزرگ شدن n کسر p از نمادهای خروجی از منبع برابر **0** و کسر $1-p$ از نمادهای خروجی از منبع برابر 1 باشند. با این فرضیه دنبالهٔ $x_1, \ldots, x_n$ بهنجار است. این تعریف را با فرضیه مستقل بودن منبع ترکیب می‌کنیم، پس به رابطهٔ زیر می‌رسیم:

$$p(x_1, \mathbf{K}, x_n) = p(x_1)\, p(x_2)\, \mathbf{K}\, p(x_n) \approx p^{np}(1-p)^{(1-p)n} \qquad (23\text{-}2)$$

که احتمال به دست آوردن یک دنباله بهنجار می‌باشد. اگر از دو طرف رابطه بالا در مبنای 2 لگاریتم بگیریم، آنگاه:

---

[19] Independent and Identically Distributed

[20] Typical sequences

[21] Atypical sequences

$$-\log p(x_1,\mathbf{K},x_n) \approx -n\,p\log p - n(1-p)\log(1-p) = n\,H(X) \qquad (24\text{-}2)$$

$H(X)$ آنتروپی توزیع منبع می‌باشد و همچنین به عنوان آهنگ آنتروپی منبع نیز شناخته می‌شود. بنابراین $p(x_1,\mathbf{K},x_2) \approx 2^{-nH(X)}$. از این موضوع می‌توان نتیجه گرفت که حداکثر $2^{nH(X)}$ دنباله بهنجار می‌تواند وجود داشته باشد. زیرا احتمال کل همه دنباله‌های بهنجار نمی‌تواند بزرگتر از 1 باشد.

اکنون ابزاری برای درک نمای ساده‌ای از متراکم سازی اطلاعات در دست داریم. فرض کنید خروجی منبع $x_1,\mathbf{K},x_n$ باشد. برای متراکم‌سازی این خروجی نخست بررسی می‌کنیم که آیا بهنجار است یا نه، اگر بهنجار نبود ثبت می‌کنیم «نادرست» و متراکم سازی صورت نمی‌گیرد. خوشبختانه هنگامی که n بزرگ می‌شود این اتفاق بندرت می‌افتد. چون همه دنباله‌ها در حد nهای بزرگ بهنجارند. اگر دنباله خروجی بهنجار باشد، ثبت می‌کنیم «درست». چون تعداد $2^{nH(X)}$ دنباله بهنجار وجود دارد، به طور یکتا تنها $nH(X)$ بیت برای مشخص کردن این دنباله بهنجار ویژه لازم است.

تاکنون دنباله‌های بهنجار باینری را مورد بررسی قرار دادیم زیرا در ورودی تنها دو مقدار **0** و **1** داشته‌ایم. اکنون آن را به موردهای کلی تعمیم می‌دهیم. فرض کنید $X_1,X_2,\mathbf{K}$ یک منبع اطلاعاتی i.i.d می‌باشد. هرگاه دنباله خروجی از منبع بهنجار باشد فراوانی رخ دادن هر حرف داده شده x در این دنباله به $p(x)$ نزدیک می‌شود، که $p(x)$ احتمال رخ دادن x در هر بار استفاده از منبع است. با این فهم شهودی تعریف زیر را از موضوع دنباله‌های بهنجار ارائه می‌دهیم. اگر $e>0$ معلوم باشد می‌گوییم که رشته‌ای از نمادهای منبع $x_1x_2\mathbf{K}x_n$، eبهنجار است اگر:

$$2^{-nH((X)+e)} \le p(x_1,\mathbf{K},x_n) \le 2^{-nH((X)-e)} \qquad (25\text{-}2)$$

و سری چنین دنباله‌های eبهنجاری را با $T(n,e)$ نشان می‌دهیم. تعریف بالا را به شکل معادل و مفید زیر بازنویسی می‌کنیم:

$$\left|\frac{1}{n}\log\frac{1}{p(x_1,\mathbf{K},x_n)} - H(X)\right| \le e \qquad (26\text{-}2)$$

این موضوع که در حد nهای بزرگ دنباله‌های خروجی از منبع بهنجارند در قضیه زیر اثبات می‌شود.

2-3-1-1- قانون اعداد بزرگ

فرض کنید یک آزمایش را به دفعات زیاد تکرار کنیم و هر بار مقدار بعضی پارمترها مانند X را اندازه بگیریم. نتایج آزمایش‌ها را با $X_1, X_2, \mathbf{K}$ برچسب می‌زنیم. با فرض اینکه نتیجه آزمایش‌ها مستقل از هم‌اند به طور شهودی انتظار داریم که مقدار تخمینی[22] $S_n = \sum_{i=1}^{n} X_i / n$ [23] [36] از میانگین $E(X)$ است، در حد $n \to \infty$ به مقدار $E(X)$ میل کند. قانون اعداد بزرگ بیان قاطعی از این تصویر شهودی می‌باشد.

قضیه5؛ قانون اعداد بزرگ. فرض کنید $X_1, X_2, \mathbf{K}$ متغیرهای تصادفی مستقلی با توزیع مشابهی به شکل متغیر تصادفی X با گشتاور اولیه و ثانویه[24] مشخص هستند $|E(X)| < \infty, |E(X^2)| < \infty$ [37]. پس به‌ازای هر $e > 0$ هر گاه $n \to \infty$ میل کند، $p(|S_n - E(X)| > e) \to 0$ است.

اثبات

در آغاز فرض می‌کنیم که $E(X) = 0$ و برای تکمیل اثبات درمورد رخداد ناشی از $E(X) \neq 0$ بحث می‌کنیم. چون متغیرهای تصادفی با میانگین صفر مستقل از هم می‌باشند، در نتیجه هرگاه $i \neq j$ باشد، $E(X_i X_j) = E(X_i)E(X_j) = 0$ و بنابراین

---

[22] estimator

[23] در علم آمار استیمیتور (برآورد) آماری (تابعی از رقم نمادی قابل مشاهده) است که برای تخمین زدن پارامتری از جمعیت ناشناخته (که qاستفاده می‌شود. استیمیت نتیجه اعمال تابع بر نمونه‌ای خاص از داده‌هاست. فرض کنید پارامتر ثابتی مانند نامیده می‌شود)*estimand* داریم که می‌خواهیم آنرا برآورد کنیم. پس استیمیتور تابعی است که یک طرح نمونه را به برآوردهای نمونه نگاشت می‌کند.

[24] ، با متغیری حقیقی $f(x)$ مین گشتاور تابع حقیقیN در ریاضی از مفهوم گشتاور در فیزیک گرفته می‌شود. moment مفهوم گشتاور یا تعریف می‌شود. می‌توان گشتاور را برای متغیرهای تصادفی در مدلی کلی‌تر $m'_n = \int_{-\infty}^{\infty} (x-c)^2 f(x)\,dx$ ، به‌صورت cحول مقدار می‌باشد. معمولاً در مفهوم ویژه از مسأله گشتاور، تابع ما یک $c = 0$نسبت به مقادیر حقیقی بیان کرد.گشتاور یک تابع همان عبارت بالا با می‌باشد و گشتاور سطری نامیده می- x^n مقدار انتظاری $f(x)$امین گشتاور حول صفر تابع توزیع احتمال n تابع توزیع احتمال می‌باشد. تابع توزیع احتمال از fامین نامیده می‌شود.به طور کلی اگر n تابع توزیع احتمال باشد پس مقدار انتگرال گشتاور توزیع احتمال fشود. اگر Riemann-Stieltgesمین گشتاور توزیع احتمال از انتگرال nهر توزیعی باشد، ممکن است تابع چگالی نداشته باشد پس عملگر انتظاری E یک متغیر تصادفی است که این توزیع را دارا می‌باشد و Xبه‌دست می‌آید که $m'_n = E(X^n) = \int_{-\infty}^{\infty} x^n d f(x)$ است.

$$E(S_n^2)=\frac{\sum_{i,j=1}^{n}E(X_iX_j)}{n^2}=\frac{\sum_{i=1}^{n}E(X_i^2)}{n^2}=\frac{E(X^2)}{n} \qquad (27\text{-}2)$$

که تساوی آخر نتیجه این امر است که $X_1,\mathbf{K},X_n$ بطور یکسان روی X توزیع شده اند. به طور مشابه از تعریف مقدار انتظاری داریم:

$$E(S_n^2)=\int d\,P\,S_n^2 \qquad (28\text{-}2)$$

که $d\,P$ اندازهٔ احتمال می‌باشد. آشکار است که $|S_n|<e$ یا $|S_n|>e$. بنابراین انتگرال را به دو قسمت می‌شکنیم و سپس یکی از این قسمت‌ها را با این توجیه که نامنفی است رها می‌کنیم،

$$E(S_n^2)=\int_{|S_n|\le e} d\,P\,S_n^2+\int_{|S_n|>e} d\,P\,S_n^2\ge\int_{|S_n|>e} d\,P\,S_n^2 \qquad (29\text{-}2)$$

در بازه انتگرال‌گیری $S_n^2>e$ و بنابراین:

$$E(S_n^2)\,e^2\int_{|S_n|>e}\mathrm{d}p=e^2\mathrm{p}(|S_n|>e) \qquad (30\text{-}2)$$

از مقایسه این نامساوی با معادله (2-27) درمی‌یابیم که

$$p(|S_n|>e)\le\frac{E(X^2)}{ne^2} \qquad (31\text{-}2)$$

با فرض $n\to\infty$ نامساوی کامل می‌شود. در موردی که $E(X)\neq\mathbf{0}$ باشد، با استفاده از تعریف زیر نتیجه آسانتر به‌دست می‌آید.

$$Y_i\equiv X_i-E(X),\qquad Y\equiv X-E(X) \qquad (32\text{-}2)$$

Y و $Y_1,Y_2,\mathbf{K}$ دنباله‌ای از متغیرهای تصادفی هستند که به طور مستقل و یکسان توزیع شده‌اند $\left(E(Y^2)<\infty\,,\,E(Y)=\mathbf{0}\right)$. نتیجه از استدلال پیشین پیروی می‌کند.

اکنون با استفاده از قانون اعداد بزرگ می‌توانیم نظریه دنباله‌های بهنجار را اثبات کنیم که نشان می‌دهد در حد nهای بزرگ بیشتر دنباله‌های خروجی از منبع اطلاعاتی بهنجار می‌باشند.

2-1-3-2- نظریه دنباله‌های بهنجار

(1) $e>0$ و ثابت در نظر می‌گیریم. پس برای هر $d>0$ و n به اندازه کافی بزرگ، احتمال اینکه یک دنباله eبهنجار باشد، حداقل $1-d$ است.

(2) برای هر $e>0$ و $d>0$ و n به اندازه کافی بزرگ تعداد $|T(n,e)|$ دنباله-های eبهنجار در معادله زیر صدق می‌کند.

$$(1-d)2^{n(H(X)-e)} \le |T(n,e)| \le 2^{n(H(X)+e)} \qquad (33\text{-}2)$$

(3) فرض کنید $S(n)$ مجموعهٔ دنباله‌ای با طول n باشد که اندازهٔ آن حداکثر 2^{nR} است. $R<H(X)$ ثابت است. پس برای هر $d>0$ و n به اندازه کافی بزرگ داریم:

$$\sum_{x\in S(n)} p(x) \le d \qquad (34\text{-}2)$$

اثبات:

قسمت 1: این قسمت کاربرد مستقیم قانون اعداد بزرگ می‌باشد. توجه کنید که $-\log p(X_i)$ متغیرهای تصادفی هستند که به طور مستقل و یکسان توزیع شده‌اند.

با استفاده از قانون اعداد بزرگ برای $e>0$ و $d>0$ و برای n به اندازه کافی بزرگ داریم:

$$p\left(\left|\sum_{i=1}^{n}\frac{-\log p(X_i)}{n} - E(-\log p(X))\right| \le e\right) \ge 1-d \qquad (35\text{-}2)$$

ولی $E(\log p(X)) = -H(X)$ و $\sum_{i=1}^{n}\log p(X_i) = \log(p(X_1,...,X_n))$. بنابراین

$$p(\left|-\log(p(X_1,...,X_n))/n - H(X)\right| \le e) \ge 1-d \qquad (36\text{-}2)$$

پس احتمال اینکه یک دنباله eبهنجار باشد $1-d$ است.

قسمت 2: این قسمت از تعریف بهنجاری، و ملاحظه اینکه جمع احتمالات دنباله‌های بهنجار باید در گسترهٔ $1-d$ تا 1 باشد[25] (زیرا جمع احتمال‌ها نمی‌تواند بیشتر از 1 شود) نتیجه می‌شود. بنابراین:

$$p(\left|-\log(p(X_1,...,X_n))/n - H(X)\right| \le e) \ge 1-d \qquad (37\text{-}2)$$

[25] با توجه به قسمت 1

$$1 \geq \sum_{x \in T(n,e)} p(x) \geq \sum_{x \in T(n,e)} 2^{-n(H(X)+e)} = |T(n,e)| 2^{-n(H(X)+e)} \qquad (38\text{-}2)$$

پس استنباط می‌کنیم که $|T(n,e)| \leq 2^{n(H(X)+e)}$ و

$$1-d \leq \sum_{x \in T(n,e)} p(x) \leq \sum_{x \in T(n,e)} 2^{-n(H(X)-e)} = |T(n,e)| 2^{-n(H(X)-e)} \qquad (39\text{-}2)$$

پس $|T(n,e)| \geq (1-d) 2^{n(H(X)-e)}$

قسمت 3: مفهوم اصلی در این قسمت، تقسیم دنباله‌های موجود در $S(n)$ به دنباله‌های بهنجار و نابهنجار می‌باشد. در حد n بزرگ دنباله نابهنجار احتمال کمتری دارد. تعداد دنباله‌های بهنجار در $S(n)$ آشکارا کمتر از تعداد کل دنباله‌ها در $S(n)$ که حداکثر 2^{nR} است، می‌باشد. احتمال هر دنباله بهنجار تقریباً $2^{-nH(X)}$ است. بنابراین احتمال کل دنباله بهنجار در $S(n)$ مقیاسی مانند $2^{n(R-H(X))}$ دارد که هر گاه $R < H(X)$ باشد به سمت صفر میل می‌کند.

نظریه کد کردن کانال بدون نوفه شانون کاربردی ساده از نظریه دنباله بهنجار است. در اینجا نسخهٔ خیلی ساده‌ای از نظریه کدکردن کانال بدون نوفه ارائه می‌دهیم. طرح کلی این است که فرض کنیم $X_1, X_2, \mathbf{K}$ یک منبع اطلاعاتی i.i.d روی تعدادی حروف الفبای محدود باشد. یک طرح فشرده-سازی با آهنگ R دنباله‌های ممکن $x = (x_1, \mathbf{K}, x_n)$ را به رشته بیتی به طول nR نگاشت می-کند. این عمل را با $C^n(x) = C^n(x_1, \mathbf{K}, x_n)$ نشان می‌دهیم. (اگر nR عددی صحیح نباشد با قرارداد فرض می‌کنیم که $nR = [nR]$). طرح پادمتراکم سازی، nR بیت فشرده را می‌گیرد و آن را در جهت برعکس عمل پیشین، برای تولید یک رشته n حرفی مناسب، نگاشت می‌کند، این عمل را با $D^n(C^n(x))$ نشان می‌دهیم. طرح کلی متراکم‌سازی-پادمتراکم‌سازی (C^n, D^n) در صورتی قابل اطمینان است که هنگامی که n به سمت بینهایت میل کند احتمال $D^n(C^n(x)) = x$ به یک نزدیک شود. نظریه کد کردن کانال بدون نوفه شانون مشخص می‌کند که برای چه مقادیری با آهنگ R یک طرح متراکم‌سازی قابل اطمینان وجود دارد. این کار را با تفسیر عملی قابل تصوری

برای آهنگ آنتروپی $H(X)$ انجام می‌دهد. آهنگ آنتروپی منابع فیزیکی لازم و کافی برای ذخیره منطقی خروجی منبع می‌باشد.

قضیه6؛ نظریه کد کردن کانال بدون نوفه شانون. فرض کنید $\{X_i\}$ یک منبع اطلاعاتی i.i.d با آهنگ آنتروپی $H(X)$ باشد. فرض کنید $R>H(X)$. پس طرح متراکم‌سازی منطقی با آهنگ R وجود دارد. برعکس اگر $R<H(X)$ هیچ طرح متراکم‌سازی منطقی نخواهد بود.

اثبات:

فرض کنید $R>H(X)$. $e>0$ را چنان انتخاب می‌کنیم که $R>H(X)+e$باشد. سری $T(n,e)$ از دنباله‌های eبهنجار را درنظر بگیرید. به‌ازای هر $d>0$ و n به اندازه کافی بزرگ، حداکثر $2^{n(H(X)+e)}<2^{nR}$ از چنین دنباله‌هایی وجود دارد. احتمال اینکه منبع چنین دنباله‌ای را تولید کند حداقل $1-d$می‌باشد. بنابراین روش متراکم سازی روشی ساده است که خروجی منبع را می‌آزماید که eبهنجار است یا نه. اگر eبهنجار نباشد در متراکم سازی به رشته بیتی به طول nR اشتباه صورت می‌گیرد و عمل پادمتراکم‌سازی دنباله‌های تصادفی حدسی $x_1,\mathbf{K},x_n$ را درحکم اطلاعات تولید شده منبع خارج می‌کند. در این مورد متراکم‌سازی کاملاً رها می‌شود. اگر خروجی منبع بهنجار باشد می‌توان آن را به سادگی با استفاده از nR بیت به روشی آشکار متراکم کرد.

فرض کنید $R<H(X)$. ترکیب عمل متراکم‌سازی – پادمتراکم‌سازی حداکثر 2^{nR} خروجی ممکن دارد، بنابراین حداکثر 2^{nR} دنباله خروجی از منبع بدون هیچ اشتباهی متراکم و پادمتراکم می‌شود. با نظریه دنباله‌های بهنجار ، برای n به اندازه کافی بزرگ احتمال اینکه دنباله خروجی از منبع برای $R<H(X)$ در زیر سری 2^{nR} دنباله قرار گیرد به سمت صفر میل می‌کند. بنابراین هیچ یک از چنین طرح‌هایی منطقی نمی‌باشند.

2-3-2- اطلاعات کلاسیکی در کانال‌های نوفه‌ای

تاکنون فقط در مورد ذخیره و انتقال اطلاعات در کانال بدون نوفه صحبت کردیم. اکنون می‌خواهیم به انتقال اطلاعات در کانال نوفه‌ای بپردازیم. در این مورد اطلاعات با اثرات نوفه مواجه می‌

شوند. هنگام صحبت با تلفن گاهی در فهمیدن صدای شخص مقابل دچار اشکال می‌شویم. این گونه-ای نوفه محسوب می‌شود که در سیستم‌های پردازش اطلاعات وجود دارد. گاهی می‌توان برای مقابله با اثرات نوفه از کدهایی مانند کدهای صحیح-غلط استفاده کرد. که این بحث‌ها در کار ما نمی‌گنجد و ما تنها به بررسی اطلاعات در کانال‌های نوفه‌ای و مقدار اطلاعاتی که در کانال‌های نوفه‌ای می‌تواند به طور معتبر انتقال یابد، می‌پردازیم. مثلاً گاهی هزار بیت برای انتقال پانصد بیت اطلاعات به‌کار می‌رود. به عبارتی می‌خواهیم ظرفیت کانال را مشخص کنیم یعنی می‌خواهیم آهنگ بیشینه ارتباط معتبر میان کانال را بیابیم.

2-3-2-1- ارتباط روی کانال‌های کلاسیکی نوفه‌ای

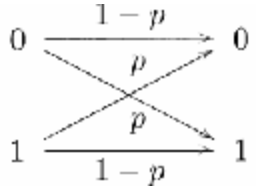
در آغاز به آهنگ بیشینهٔ انتقال معتبر اطلاعات میان کانال‌های متقارن باینری[26] می‌پردازیم و نشان می‌دهیم که این آهنگ برابر با $1-H(P)$ می‌باشد. فرض کنید همه ورودی‌های کانال به یکبارگی در بلوک‌های بزرگی به صورت کد درآیند، بزرگ بودن اندازه بلوک‌ها ایجاب می‌کند که احتمال خطا در انتقال اطلاعات به سمت صفر میل کند. این گفته می‌تواند تعبیری برای معتبر بودن کانال باشد.

آیا آهنگ بیشینه‌ای وجود دارد که هر نوع مخابره اطلاعات با بیش از آن آهنگ غیر ممکن باشد؟ (آیا ظرفیت بیشینه‌ای برای کانال وجود دارد؟) این پرسشی است که شانون آن را در قضیه دوم خود پاسخ می‌دهد. در اینجا بهتر است پیش از اشاره‌ای صریح به این قضیه به مثال‌های مفیدی بپردازیم که درک مطلب راحت‌تر باشد.

فرض کنید می‌خواهیم nR بیت از اطلاعات را با n بار استفاده از کانال متقارن باینری انتقال دهیم. یعنی آهنگ انتقال اطلاعات R می‌باشد. در روشی بسیار مفید برای انجام این کار از کدهای صحیح- غلط استفاده می‌شود. با استفاده از این کدها، در حد nهای بزرگ، احتمال رخ دادن اشتباه بسیار کم می‌شود. ساختن کدهای صحیح-غلط نیز می‌تواند به صورت‌های مختلفی انجام پذیرد. روشی که در اینجا استفاده می‌شود روش کدکردن تصادفی است. فرض می‌کنیم ورودی‌های ممکن کانال (1 و **0**) بوده و (q و $1-q$) توزیع احتمال ثابتی روی این ورودی‌ها باشند. برای هر کد موجود، کدواژه $x=(x_1,\mathbf{K},x_n)$ را انتخاب می‌کنیم که با احتمال q، مقدار $x_j=\mathbf{0}$ و با احتمال $1-q$، مقدار

---

[26] کانال متقارن باینری ، کانال ارتباطی نوفه‌ای برای یک بیت سیگنال از اطلاعات است که در آن یک بیت با احتمال $1-P$ انتقال می‌یابد و فلیپ می‌کند. مفاهیم زیادی از کد کردن کانال‌های نوفه‌ای، با استفاده از کانال متقارن باینری قابل درک است. $P>\mathbf{0}$ با احتمال



$x_j = 1$ می‌باشد. در حالت کلی اگر از n بیت برای انتقال 2^k پیام استفاده شود آهنگ انتقال پیام با رابطه $R := k/n$ داده می‌شود.

فرآیند انتخاب کدواژه‌ها را 2^{nR} بار تکرار می‌کنیم. بنابراین می‌توان گفت تعداد کدواژه‌ها برابر با 2^{nR} می‌باشد. شاید انتخاب کدواژه به طور تصادفی، مناسب به نظر نرسد ولی به طور متوسط فرآیند کد کردن تصادفی، کد صحیح-غلط خوبی ارائه می‌دهد.

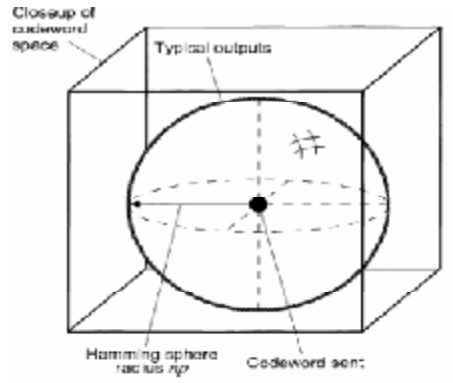
برای درک درستی این مطلب به عملی که کانال روی یک کدواژه در یک کد انجام می‌دهد توجه می‌کنیم. هر کدواژه را به صورت x^j نمایش می‌دهیم. مثلاً اثر کانال روی x^1 چه می‌باشد؟ در کدواژه-ای به طول n در کانال متقارن باینری انتظار داریم np از بیت‌ها فلیپ[27] کند بنابراین با احتمال بالایی خروجی کانال از کدواژه x^1 فاصله هامینگ[28] np دارد. ر.ک. شکل (2-2). می‌گوییم که چنین خروجی‌هایی کره هامینگی به شعاع np حول x^1 خواهند ساخت. گفته‌های بالا را می‌توان اینگونه بیان کرد که مجموعه نقاط تمام رشته‌های n بیتی را که 2^{nR} رشته است به عنوان نقاط یک شبکه ابرمکعبی n بعدی تصور کنیم. اگر همه این نقاط را به عنوان کدواژه‌های خود در نظر بگیریم در طول عبور از کانال یک کدواژه به یکی از کدواژه‌های همسایه‌اش تبدیل می‌شود و گیرنده واقعاً نمی‌فهمد چه کدواژه‌ای برایش ارسال شده است. بنابراین راه مقابله با خطا این است که کدواژه‌ها به فاصله مناسب از هم انتخاب شوند تا احتمال وقوع خطا پایین آید. در واقع هرگاه یک کدواژه را ارسال می-کنیم حول آن یک کره به اندازه کافی بزرگ رسم می‌کنیم و هنگامی که نقطه‌ای درون این کره دریافت می‌کنیم آن را به عنوان کدواژه‌ای که در مرکز آن قرار دارد تعبیر و خطاهای به وجود آمده را تصحیح می‌کنیم. این کار آهنگ را به طور طبیعی پایین می‌آورد، زیرا از همه نقاط شبکه استفاده نمی‌کنیم.

می‌توان تعداد عناصری را که در این فاصله هامینگ وجود دارد، از $\binom{n}{np} \approx 2^{nH(p)}$ محاسبه کرد.

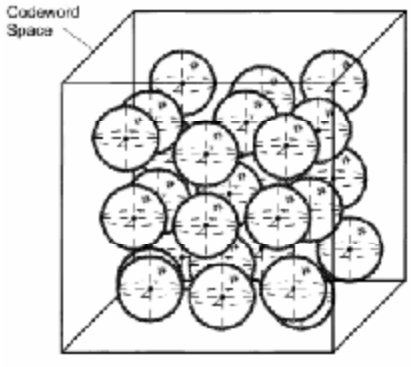
---

[27] Flip

[28] فاصلهٔ هامینگ با تعداد مکان‌هایی که دو رشتهٔ بیتی مساوی نباشند مشخص می‌شود. در این باره در فصل سوم بیشتر توضیح داد ه می‌شود.



بار استفاده از کانال متقارن باینری فرستاده شود.پس یک خروجی n با x^1شکل (2-2)، فرض کنید کد واژهٔ
است که فرستاده شده است . $n\,p$بهنجار کانال یک عنصر از کره همینگ به شعاع



شکل (2-3)، کد کلمه‌هایی که به طور تصادفی برای کانال متقارن باینری فرستاده شده‌اند.
آنها به وسیله کره‌های همینگ خروجی‌های بهنجار شان، احاطه شده‌اند.

چون فاصله همینگ همه خروجی‌های ممکن اتفاقی را شامل می‌شود، $y = x^1 \oplus e$ خروجی و e خطایی است که ممکن است در کانال رخ دهد. با نظریهٔ دنباله‌های بهنجار تعداد چنین خطاهای از نوع e حدوداً $2^{nH(p)}$ می‌باشد. این نوع انحراف‌ها برای همه کدواژه‌ها رخ می‌دهد. حال همه فضای کدواژه‌ها و فضای همینگ اطراف آنها را در نظر می‌گیریم شکل (2-3)، پس راهی آسان برای شخص دریافت کنندهٔ پیام وجود دارد که بتواند به کمک آن خروجی کانال را کدگشایی کند. در واقع او با بررسی اینکه آیا خروجی در یکی از کره‌های همینگ است خروجی را منطبق بر کدواژه می‌گیرد و اگر خروجی در کره همینگ نباشد اعلام خطا می‌کند. چون کره‌ها همپوشانی ندارند نتیجه کدگشایی

موفقیت آمیز است. در واقع اگر همپوشانی بسیار کوچکی وجود داشته باشد باز هم شخص دریافت-کننده پیام با شانس موفقیت خوبی می‌تواند کدگشایی صحیحی انجام دهد.

چگونه همپوشانی کوچکی بین کره‌ها رخ می‌دهد؟ برای پاسخ به این پرسش باید ساختار خروجی‌های ممکن کانال را بهتر بشناسیم. با 2^{nR} بار نمونه‌برداری از سری $(X_1,...,X_n)$ از متغیرهای تصادفی مستقل، کدواژه‌ها را به‌دست می‌آوریم ($X_j=\mathbf{0}$ با احتمال q و $X_j=1$ با احتمال $1-q$). فرض کنید Y_j نتیجه فرستادن X_j میان کانال متقارن باینری باشد، نظریهٔ دنباله‌های بهنجار بیان می‌کند که سری مقادیر بهنجار برای $Y_1,...,Y_n$ اندازه‌ای برابر $2^{nH(Y)}$ دارا می‌باشند که Y توزیع همه Y_i می‌باشد. همه خروجی‌های ممکن بهنجار متساوی‌الاحتمال‌اند.

اکنون موردی را در نظر بگیرید که یکصد بار از جمعیت یک میلیونی به طور یکنواخت نمونه برداری کنیم احتمال دست‌یابی به تکرار چندان محتمل نیست. حتی اگر هزار بار نمونه برداری کنیم باز هم تعداد تکرارها با دقت خوبی کوچک خواهد بود. هرگاه ما یک میلیون بار نمونه‌برداری کنیم تعداد تکرارها نسبت به اندازه نمونه بزرگ می‌شود. در مدلی مشابه تا هنگامی که تعداد عناصر ترکیب‌شده در همه کره‌ها به اندازه فضای $2^{nH(Y)}$ نزدیک نشود، مقدار همپوشانی بین کره‌های هامینگ با شعاع np، آغاز نمی‌گردد. چون هر کره $2^{nH(p)}$ عنصر دارد با احتمال بالایی کد صحیح-غلط خوبی داریم مشروط بر اینکه:

$$2^{nR}\times 2^{nH(p)}<2^{nH(Y)} \qquad (40\text{-}2)$$

که بر شرط زیر منطبق است

$$R<H(Y)-H(p) \qquad (41\text{-}2)$$

آنتروپی $H(Y)$ به توزیع پیشین (q و $1-q$) که برای X_j انتخاب شده، بستگی دارد. برای اینکه بتوانیم آهنگ انتقال اطلاعات را به اندازه کافی بزرگ کنیم باید سعی کنیم $H(Y)$ را بیشینه کنیم. اگر توزیع منطبق بر $q=1/2$ باشد، $H(Y)=1$ و بنابراین برای آهنگ انتقال اطلاعات به‌دست می-آید:

$$R < 1 - H(p) \qquad (42\text{-}2)$$

2-3-2-2- نظریهٔ کد کردن کانال نوفه‌ای شانون

نظریهٔ کد کردن کانال نوفه‌ای شانون. ظرفیت برای کانال نوفه‌ای N از رابطه زیر به‌دست می‌آید:

$$C(N) = \max_{p(x)} H(X : Y) \qquad (43\text{-}2)$$

که ماکزیمم روی همه توزیع‌های ورودی $p(x)$ برای X و هر استفاده از کانال گرفته می‌شود. و Y متغیر تصادفی منطبق در خروجی کانال می‌باشد. به عنوان یک مثال از نظریهٔ کد کردن کانال نوفه‌ای، مورد کانال متقارن باینری را در نظر بگیرید که در آن هر بیت با احتمال p، فلیپ می‌کند و با توزیع ورودی $p(\mathbf{0}) = q$ و $p(1) = 1 - q$ داریم:

$$\begin{aligned} H(X:Y) &= H(Y) - H(Y \mid X) \\ &= H(Y) - \sum_x p(x) H(Y \mid X = x). \end{aligned} \qquad (44\text{-}2)$$

به‌ازای هر x، $H(Y \mid X = x) = H(p)$ بنابراین $H(X:Y) = H(Y) - H(p)$، که با انتخاب $q = 1/2$ بیشینه می‌شود. بنابراین $H(Y) = 1$ و $C(N) = 1 - H(p)$. این نتیجه که از نظریهٔ کدکردن کانال نوفه‌ای شانون به دست آمد درست همان نتیجه‌ای است که با محاسبات پیشین برای کانال متقارن باینری به‌دست آوردیم[5 و6].

2-4- نتیجه‌گیری

در این فصل ضمن معرفی آنتروپی شانون به چند ویژگی مهم این آنتروپی اشاره کردیم و برخی از کاربردهای مهم آن را برشمردیم. بنابراین می‌توان نتیجه گرفت که در فیزیک کلاسیک بهترین و مفیدترین رابطه‌ای که برای کمی‌کردن اطلاعات بکار می‌رود آنتروپی شانون می‌باشد. لازم به ذکر است که چون آنتروپی شانون به احتمال رویدادهای مختلف وابسته است، محتوای اطلاعات سیستمی که بر اساس آن به‌دست می‌آید همواره تحت اندازه‌گیری‌های مختلف ناوردا باقی می‌ماند. موضوع دیگر اینکه اگر احتمال رویدادهای مختلف به زمان وابسته نباشد، محتوای اطلاعات به دست آمده از آنتروپی شانون در زمان بقا دارد. این دو ویژگی از مهمترین ویژگی‌هایی است که اندازه اطلاعات باید داشته باشد که همان‌گونه که گفتیم آنتروپی شانون به خوبی در این شرط‌ها صدق می‌کند.

فصل سوم

آنتروپی فون‌نیومن و اطلاعات کوانتمی

- ذخیرهٔ اطلاعات در فیزیک کلاسیک و کوانتمی
- غیر قابل کاربردبودن آنتروپی شانون در مکانیک کوانتمی
- اطلاعات کوانتمی
- ذخیره و انتقال اطلاعات در سیستم‌های کوانتمی
- اطلاعات کلاسیکی در کانال‌های کوانتمی
- نتیجه‌گیری

مقدمه

اطلاعات کوانتمی چیست؟ اطلاعات کوانتمی گونه‌ای از اطلاعات است که سیستم‌های کوانتمی حمل می‌کنند. در یک آزمایش کوانتمی این اطلاعات از وسیله‌ای که آن‌ها را آماده کرده به ابزار اندازه‌گیری منتقل می‌شود. بنابراین یک انتقال دهندهٔ اطلاعات کوانتمی چیزی نیست به جز وسیله‌ای که ذره‌های کوانتمی را آماده می‌کند، و دریافت کننده یک وسیلهٔ اندازه‌گیری است.

در این فصل به طور مفصل دربارهٔ انتقال اطلاعات کوانتمی بحث می‌کنیم. ولی پیش از پرداختن به این بحث باید بتوانیم اطلاعات را به صورت کد (رمز) درآوریم، یعنی به صورت کدهایی که حجم کمتری را نسبت به خود اطلاعات داشته باشند و ما قادر باشیم این کدها را به حالت اولیه یعنی حالت قبل از کد شدن (رمز شدن) برگردانیم. از طرفی برای دانستن میزان اطلاعات هر منبع باید آن را به صورت کمی درآوریم. بیان ریاضی منبع اطلاعاتی اصطلاحاً کمّی کردن اطلاعات نامیده می‌شود. گفتیم که در فیزیک کلاسیک بهترین روش برای کمی کردن اطلاعات استفاده از آنتروپی شانون می‌باشد، اما در مکانیک کوانتمی استفاده از این آنتروپی ما را با مشکلاتی مواجه می‌کند و به دلایل زیادی غیرقابل کاربرد اعلام می‌شود که ما به چند مورد از آنها اشاره کرده و سرانجام سعی بر معرفی آنتروپی‌هایی داریم که در مکانیک کوانتمی مفید و قابل استفاده باشند.

3-1 ذخیرهٔ اطلاعات در فیزیک کلاسیک و کوانتمی

در فیزیک کلاسیک اطلاعات به صورت دنباله‌ای دودویی یعنی دنباله‌ای از مقادیر **0** یا 1 نمایش داده می‌شوند. هرگاه اطلاعاتی را که یک سیستم کلاسیکی حمل می‌کند بازخوانی کنیم در واقع یک ویژگی از سیستم را که پیش از بازخوانی نیز وجود داشته است آشکار می‌کنیم. یعنی می‌توان گفت که در سیستم‌های کلاسیکی ویژگیهای سیستم است که دنباله‌ای از بیت‌ها را تشکیل می‌دهد. در این سیستم‌ها اطلاعات را به‌وسیلهٔ آنتروپی شانون اندازه می‌گیریم و یا به صورت کمی درمی‌آوریم. به تعبیری می‌توان اندازه شانون اطلاعات را تعداد پرسشهایی با پاسخ «بله» یا «خیر» دانست که برای تعیین دنبالهٔ خاصی از **0** و 1ها لازم است. در ادامه این مطلب بیشتر توضیح داده می‌شود.

در فیزیک کوانتمی اطلاعات به وسیلهٔ دنباله‌ای از کیوبیت‌ها نمایش داده می‌شود. هر کیوبیت در یک فضای هیلبرت دو بعدی تعریف می‌شود. یک کیوبیت مقدار اطلاعاتی است که یک سیستم کوانتمی دو حالته مانند ذره‌های اسپین 1/2 می‌توانند حمل کنند. بنابراین کیوبیت یکی از واحدهای اطلاعات کوانتمی است. اگر بخواهیم اطلاعاتی را که یک کیوبیت حمل می‌کند بازخوانی کنیم مجبوریم حالت کیوبیت را روی پایه‌های اندازه‌گیری، $\{|\mathbf{0}\rangle,|1\rangle\}$ تصویر کنیم که مقادیر بیتی **0** یا 1 را به ما می‌دهد. تنها در مورد استثنایی که کیوبیت در ویژه حالت ابزار اندازه‌گیری باشد مقدار بیتی که از اندازه‌گیری به‌دست می‌آید یک ویژگی سیستم را که پیش از اندازه‌گیری نیز وجود داشته آشکار می‌کند. ولی به طور کلی می‌توان گفت که نتیجه اندازه‌گیری به طور کاهش‌ناپذیری تصادفی است و نمی‌توان به هیچ روشی فرض کرد که آنچه در اندازه‌گیری حاصل می‌شود یک ویژگی سیستم را نشان می‌دهد. عدم وجود مقادیر بیتی خوش‌تعریف و مستقل از مشاهده بیانگر این است که اندازه شانون اطلاعات در مکانیک کوانتمی غیر قابل استفاده است و پیش از انجام دادن اندازه‌گیری نمی‌توان عدم یقین یک سیستم کوانتمی را با آنتروپی شانون محاسبه کرد [8 و11].

3-2- غیر قابل کاربرد بودن آنتروپی شانون در مکانیک کوانتمی

گفته شد که در اندازه‌گیری‌های کلاسیکی آنتروپی شانون میزان عدم یقین ما را دربارهٔ ویژگی‌های سیستم نشان می‌دهد. اما به دلیل ریشهٔ کاملاً متفاوت اندازه‌گیری‌های کوانتمی در مقایسه با اندازه-گیری‌های کلاسیکی واشکالات مفهومی که روی می‌دهد، استفاده از آنتروپی شانون در مکانیک کوانتمی مردود اعلام می‌شود. دلیلی که وجود دارد این است که در اندازه‌گیری‌های کوانتمی برخلاف اندازه‌گیری‌های کلاسیکی (به‌جز استثناهای اندکی که وجود دارد)، نمی‌توان ادعا کرد که با اندازه-گیری ویژگیهایی از سیستم که پیش از اندازه‌گیری نیز وجود داشته‌اند آشکار می‌شود.

3-2-1- تحلیل عملی آنتروپی شانون برای اندازه‌گیری‌های کوانتمی

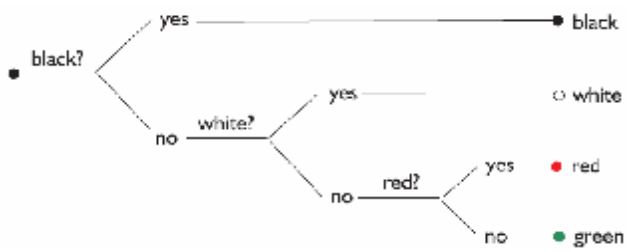
در اینجا به شکلی عملی کاربرد اندازه اطلاعات شانون بررسی می‌شود و نشان داده می‌شود که این اندازهٔ در تعیین محتوی اطلاعات سیستم‌های کوانتمی غیر قابل کاربرد است. در حقیقت خواهیم

دید که هرگاه بخواهیم اصل شانون را در اندازه‌گیری کوانتمی بکار بریم یا آنکه محتوی اطلاعاتی را با اطلاعات (آنتروپی) شانون تعیین کنیم، عنصری حاصل می‌شود که توصیف کامل آن از عهده مکانیک کوانتمی خارج است. این عنصر همیشه با تصادفی بودن عینی رخدادهای کوانتمی و با اصل مکملیت کوانتمی همراه است.

برای مشاهده‌های کلاسیکی، اندازه شانون اطلاعات از نظر مفهومی می‌تواند با روشی که بر سؤال پرسیدن استوار است، توضیح داده شود. برای این کار فرض کنید کوزه‌ای از توپ‌های رنگی پر شده باشد. پس $n_1, n_2, \mathbf{K}, n_m$ توپ با رنگ‌های مختلف: سیاه ، سفید،... قرمز وجود دارد. اگر کوزه تکان داده شده و یک توپ بیرون آوریم، تا چه حد می‌توان رنگ ویژه‌ای را که خارج شده پیش‌بینی کرد؟ اگر رنگ همهٔ توپ‌ها یکسان باشد آشکارا رنگ توپ بیرون آورده شده را دقیق پیش‌بینی خواهیم کرد. ازطرفی اگر رنگ‌ها متساوی‌الاحتمال باشند عدم یقین ما دربارهٔ رنگ بیرون آمده بیشینه می‌باشد.

به عنوان مثالی ویژه فرض کنید کوزه‌ای با توپ‌هایی از چهار رنگ مختلف سیاه، سفید، قرمز و سبز پر شده باشد، که احتمال آنها به ترتیب $p_1 = \frac{1}{2}$ ، $p_2 = \frac{1}{4}$ ، $p_3 = \frac{1}{8}$ و $p_4 = \frac{1}{8}$ می‌باشد. اکنون فرض کنید توپی از کوزه بیرون آورده می‌شود، می‌خواهیم با پرسیدن تعدادی سؤال که جواب آنها بله یا خیر است به رنگ بیرون آورده شده پی ببریم. اگرچه تعداد این پرسش‌ها به استراتژی پرسیدن بستگی دارد ولی بهینه‌ترین استراتژی که می‌توان به‌کار برد این است که سعی کنیم از هر پرسش اطلاعاتی بیشینه به‌دست آوریم. بنابراین مجبوریم سؤال‌هایی بپرسیم که جواب‌های آنها نیمی از امکانات را که احتمال رخدادن آنها وجود دارد حذف کند. درآغاز بهترین سؤالی که می‌توان پرسید این است: آیا رنگ سیاه بیرون آورده شده است؟ چه جواب این سؤال بله باشد چه خیر، نیمی از امکانات وقوع حذف خواهند شد. اگر جواب بله باشد ما به خواسته خود رسیده‌ایم. ولی اگر جواب خیر باشد می‌توان سری باقیمانده را به دو قسمت با احتمال‌های مساوی{سفید} و {قرمز و سبز} تقسیم کرد، و با پرسیدن اینکه « آیا رنگ بیرون آمده سفید است؟» جلو رفت. دوباره اگر جواب «بله» باشد، مطلوب ماست ولی اگر «خیر» باشد باز هم باید ادامه دهیم تا به رنگ بیرون آمده پی ببریم. یک خروجی ویژه

با به ترتیب نوشتن بله و خیرهایی که هنگام حرکت از ریشهٔ درخت باینری شکل(3-1) به برگ‌های آن به‌دست آمده مشخص می‌شود. به سادگی می‌توان دیدکه با این استراتژی بهینه، تعداد سؤال‌های باینری کمینه که برا ی تعیین رنگ بیرون آمده لازم است به صورت زیر می‌باشد:



شکل (3-1). درخت سؤالات باینری برای تعیین رنگ بیرون آورده شده.

$$p_1\times 1+p_2\times 2+(p_3+p_4)\times 3=\frac{1}{2}\times 1+\frac{1}{4}\times 2+(\frac{1}{8}+\frac{1}{8})\times 3=\frac{7}{4} \qquad (1\text{-}3)$$

معادلهٔ بالا اگر به صورت زیر نیز نوشته شود همان جواب پیشین را می‌دهد.

$$-\frac{1}{2}\log\frac{1}{2}-\frac{1}{4}\log\frac{1}{4}-\frac{1}{8}\log\frac{1}{8}-\frac{1}{8}\log\frac{1}{8}=-\sum_{i=1}^{4}p_i\log p_i \qquad (2\text{-}3)$$

یا موقعیتی کلی را تصورکنید که N توپ مختلف با احتمال‌های $p_1,p_2,...,p_m$ داریم که آنها را از دریای بینهایتی ازتوپ بیرون کشیده‌ایم. پس یک دنبالهٔ طولانی شامل p_1N توپ از رنگ اول p_2N توپ از رنگ دوم و... است (چنین دنباله‌ای دنبالهٔ بهنجار نامیده می‌شود). احتمال برای به‌دست آوردن یک دنبالهٔ بهنجار ویژه با رابطهٔ زیر داده می‌شود.

$$p\,(\text{sequence})=p_1^{\,p_1N}\,p_2^{\,p_2N}\;\mathbf{K}=\frac{1}{2^{N\,H}} \qquad (4\text{-}3)$$

که

$$H=-\sum_{i=1}^{m}p_i\log p_i \qquad (5\text{-}3)$$

H آنتروپی شانون می‌باشد و به دلیل تناسب با سیستم دوتایی یا بیت، log در پایهٔ 2 گرفته می‌شود. تعداد کل دنباله‌های بهنجار N توپ با رنگ‌های متفاوت با تعداد جایگشت‌های تمیزپذیر این توپ‌ها به‌دست می‌آید.

$$\frac{N\ !}{(p_1 N)!(p_2 N)!...(p_m N)!} \to 2^{N H} \qquad (6\text{-}3)$$

در به‌دست آوردن معادلهٔ بالا از تقریب استرلینگ $N\ ! \approx \sqrt{2p\ N}\ N^N e^{-N}$ استفاده کرده‌ایم. در واقع می‌توان گفت که چون $W \approx 2^{N H}$ دنبالهٔ ویژه وجود دارد. بنابراین تعداد سؤال‌های «بله، خیر» لازم برای به‌دست آوردن یک دنبالهٔ ویژه برابر NH می‌باشد. یا می‌توان گفت آنتروپی شانون به‌دست آمده در بیت‌ها از تقسیم تعداد سؤال‌های لازم برای به‌دست آوردن یک دنبالهٔ ویژه بر N حاصل می‌شود.

اگر به جای توپ‌ها با سیستم‌های کوانتمی که ویژگی‌های آنها پیش از انجام دادن اندازه‌گیری برای ما معلوم نیست سروکار داشته باشیم دیگر اندازهٔ شانون اطلاعات برای تعیین اطلاعات کسب شده از اندازه‌گیری مناسب نمی‌باشد و نمی‌توان $W \approx 2^{N H}$ را تعداد سؤال‌های لازم برای تعیین خروجی اندازه‌گیری دانست [8].

در فیزیک کلاسیک رفتار کل سیستم از رفتار تک‌تک اجزای تشکیل دهندهٔ سیستم پیروی می‌کند. اما این در مورد سیستم‌های کوانتمی درست نمی‌باشد. نامعینی اساسی در عدم وجود توصیفی جزئی و پیشگویی برای نتایج رخدادهای کوانتمی در اندازه‌گیری ویژه بیان می‌کند که در مکانیک کوانتمی پیش از انجام دادن اندازه‌گیری نمی‌توان دنبالهٔ ویژه‌ای از خروجی‌ها را با به ترتیب نوشتن بله وخیرها به‌دست آورد. و هیچ خروجی مشخصی قبل از انجام دادن اندازه‌گیری وجود ندارد. بنابراین تعداد خروجی‌های ممکن متفاوت نمی‌تواند عدم یقین ما را درباره سیستمی که اندازه‌گیری روی آن انجام نشده مشخص کند.

اما اگر دنباله‌ای از اندازه‌گیری‌ها روی سیستم کوانتمی انجام شود و نتیجهٔ اندازه‌گیری معلوم شود می‌توان اطلاعات به‌دست آمده را با آنتروپی شانون کمّی کرد.

3-2-2- غیرقابل استفاده بودن فرضیه‌های شانون در اندازه‌گیری‌های کوانتمی

جینز می‌نویسد: یک دلیل مهم برای ترجیح دادن اندازه شانون اطلاعات این است که به خوبی فرضیه‌های منطقی خودش را برآورده می‌کند[12].

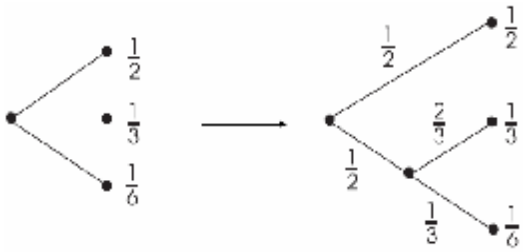
رابطهٔ مشهور شانون بر فرض‌های منطقی زیر استوار است:

اگر $H(p_1, p_2, \ldots\ldots p_n)$ عدم‌یقین موجود در خروجی اندازه‌گیریِ رویدادهایی با احتمال رخداد $p_n, \ldots\ldots, p_i$ باشد آنگاه،

الف) H باید پیوسته باشد.

ب) اگر همه p_i ها مساوی باشند و $p_i = 1/n$ پس H باید یک تابع اکیداً صعودی از n باشد.

ج) اگر یک رویداد احتمالی به دو رویداد شکسته شود، H اولیه باید با مجموع Hهای منحصر بفرد تولید شده برابر باشد. برای مثال:



شکل(3-2). شکستن یک امکان به سه

$$H\left(\frac{1}{2},\frac{1}{3},\frac{1}{6}\right)=H\left(\frac{1}{2},\frac{1}{2}\right)+\frac{1}{2}H\left(\frac{2}{3},\frac{1}{3}\right)$$

شانون نشان داد که تنها تابع $H=-\sum p_i \log p_i$ این فرضیه‌ها را برآورده می‌کند[6 و 8].

دو فرضیهٔ نخست برای هر اندازه‌ای از اطلاعات طبیعی برقرار می‌باشند. ولی فرضیهٔ سوم نیاز به توجیه بیشتری دارد. در واقع فرضیه سوم شانون شکل لگاریتمی H را تعیین می‌کند و همین فرضیه است که در مکانیک کوانتمی دچار اشکال می‌شود. این فرضیه را فادیو به صورت زیر فرمولبندی کرده است: برای هر $n \geq 2$

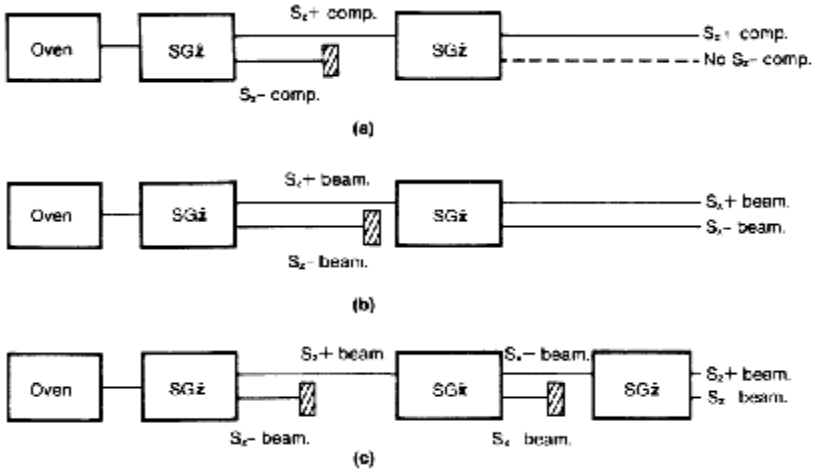
$$H(p_1,\ldots,p_n,q_1,q_2)=H(p_1,\ldots,p_{n-1},p_n)+p_nH(\frac{q_1}{p_n},\frac{q_2}{p_n}) \qquad (7\text{-}3)$$

که $p_n=q_1+q_2$ می‌باشد. تعبیر فیزیکی (7-3) به این قرار است: فرض کنید $H(p_1,p_2,\ldots\ldots\ldots p_n)$ مقدار اطلاعات حاصل از انجام دادن آزمایشی با خروجی‌های $a_1\ldots\ldots\ldots\ldots a_n$ باشد. اکنون اگر a_n به رخدادهای b_1,b_2 تبدیل شود یعنی برای رخ دادن b_1 و b_1 حتما نخست باید a_n رخ دهد، یعنی داریم $a_n\wedge b_1$ و $a_n\wedge b_2$ که ($\wedge$به‌معنای «و»). احتمال‌های خروجی $a_n\wedge b_1$ و $a_n\wedge b_2$ به‌ترتیب عبارتند از q_1 و q_2. پس سمت چپ معادله (7-3) مقدار اطلاعات حاصل از انجام آزمایشی با خروجی‌های $a_n\wedge b_1$ و $a_n\wedge b_2$ و $a_1\ldots\ldots\ldots a_n$ را مشخص می‌کند. معادله (7-3) و تفسیرهای بالا بیانگر آن است که در فیزیک کلاسیک همواره می‌توان همهٔ آزمایش‌ها را هم‌زمان روی سیستم انجام داد مثلا می‌توان مکان و اندازه‌حرکت جسم را هم‌زمان اندازه گرفت. ولی در مکانیک کوانتومی فقط زمانی می توان دو مشاهده‌پذیر را هم‌زمان روی سیستم اثر داد که آن دو مشاهده‌پذیر جابجاپذیر باشند. در غیر این صورت نتیجه یک مشاهده باعث از دست دادن اطلاعات مربوط به مشاهده‌پذیر دیگر می‌شود. مثلاً نمی‌توان مکان و اندازه حرکت یک سیستم کوانتمی را هم‌زمان دانست. در واقع اطلاعات راجع به مکان جسم اطلاعات مربوط به اندازه حرکت را از بین می‌برد و برعکس. پس فرضیه سوم شانون که با معادله (7-3) تفسیر شد در مکانیک کوانتومی قابل استفاده نیست. همچنین این فرضیه دربردارنده دو معنی است که در مکانیک کوانتومی رد می‌شوند: الف) کشف دانش جدید درباره سیستم سبب افزایش آگاهی ما می‌شود. در مکانیک کوانتومی اگر مشاهده‌پذیرها جابجاپذیر نباشند این مسأله درست نیست و کسب اطلاعات دربارهٔ یک مشاهده‌پذیر باعث از دست دادن اطلاعات مشاهده‌پذیر دیگر می‌شود. ب) اطلاعاتی که راجع به سیستم به‌دست می‌آید به ترتیب انجام دادن آزمایش بستگی ندارد. اگر چه این قضیه درمکانیک کلاسیک کاملاً برقرار است ولی در مکانیک کوانتومی اطلاعات حاصل از سیستم دقیقاً به ترتیب انجام دادن آزمایش‌ها بستگی دارد. آزمایش اشترن- گرلاخ مصداق بسیار خوبی برای نشان دادن هر دو مورد الف و ب است شکل (3-3) [16].

بنابراین دلیل‌های بالا همه فرضیه سوم شانون و در نتیجه آنتروپی شانون را در مکانیک کوانتومی مردود و غیر قابل استفاده می‌سازند.

3-2-3 مشکلات تعیین محتوای اطلاعات یک سیستم کوانتمی

به طور کلی محتوای اطلاعاتی سیستم باید تحت اندازه‌گیری‌های متفاوت ناوردا باشد. همچنین اگر جابجایی اطلاعات با محیط صورت نگیرد محتوای اطلاعاتی سیستم باید در بستر زمان پایسته باشد. آنتروپی شانون این دربایست‌ها را در تعیین محتوای اطلاعات کوانتمی برآورده نمی‌کند.



از این آزمایش کاملاً واضح است که اطلاعات شکل (3-3). آزمایش اشترن گرلاخ.
سیستم‌های کوانتمی به ترتیب انجام آزمایش بستگی دارد

جهان کلاسیکی از ذره‌ها و میدان‌ها تشکیل شده است و ویژگیهای همه این ذرات تشکیل دهنده از روشی که آزمایشگر برای اندازه‌گیری برمی‌گزیند کاملاً مستقل است. در این صورت گفته می‌شود که جهان کلاسیکی یک جهان ناوابسته به قراین عینی[29] است. اما دنیای کوانتمی یک دنیای وابسته به قراین عینی[30] می‌باشد و محتوای اطلاعاتی سیستم دقیقاً به زمینه آزمایشی انجام شده روی آن بستگی دارد. در صورتی از سیستم کوانتمی بیشینهٔ اطلاعات را کسب می‌کنیم که پایه‌های اندازه-گیری $|i\rangle$ با ویژه‌پایه‌های ماتریس چگالی سیستم همزمان باشند یعنی $\hat{r}|i\rangle = w_i|i\rangle$. در این

[29] Noncontextual

[30] contextual

صورت $\hat{r}$ در پایه‌های آزمایش بهین قطری است و w_i عناصر قطر اصلی می‌باشند. درنتیجه در این حالت خاص آنتروپی شانون با آنتروپی فون‌نیومن[31]، $-Tr(\hat{r}\log\hat{r})$[32]، که برای سیستم‌های کوانتمی به‌کار می‌رود برابر است.

$$H=-\sum_i w_i \log w_i=-Tr(\hat{r}\log\hat{r}) \qquad (8\text{-}3)$$

در اینجا H تحت تبدیلات یکانی ناوردا‌ست. یعنی اگر پایه‌های آزمایش را تغییر دهیم $\hat{r}$ تحت تغییر پایه‌ها، که تبدیلاتی یکانی‌اند، ناوردا باقی می‌ماند و بنابراین H نیز ناوردا باقی می‌ماند. همچنین:

$$H(t)=-\sum_i w_i(t)\log w_i(t)=-\sum_i w_i \log w_i \qquad (9\text{-}3)$$

که نشان می‌دهد در این حالت H در بستر زمان پایسته می‌باشد. درصورتی می‌توان از (3-8) و (3-9) استفاده کرد که ویژه‌پایه‌های دستگاه اندازه‌گیری بر ویژه‌پایه‌های ماتریس چگالی منطبق باشند، در غیر این صورت H تحت تبدیلات یکانی ناوردا نیست. اما اگر ویژه‌پایه‌های ماتریس چگالی را ندانیم، نمی‌توانیم آزمایش بهین را پیدا کنیم و آنتروپی شانون برای سیستم‌های کوانتمی قابل استفاده نیست. بنا بر اصل تمامیت[33] کوانتمی محتوای اطلاعات سیستم در یک دنبالهٔ کامل مشاهده‌پذیرهای دو به دو جابجاناپذیر نهفته است. زیرا با اندازه‌گیری سری کامل مشاهده‌پذیرهای مکمل می‌توان ماتریس چگالی سیستم را بازسازی کرد [17]. ویژگی این مشاهده‌پذیرها آن است که دانش کامل دربارهٔ یکی، ما را از داشتن هر دانشی در مورد دیگری محروم می‌سازد. پیشنهاد می‌شود محتوای اطلاعاتی کل سیستم برابر مجموع مقدارهای منحصربفرد اطلاعات روی یک دنباله کامل از m مشاهده‌پذیر دوبه‌دو جابجاناپذیر باشد. پس اگر بخواهیم با استفاده از آنتروپی شانون محتوای اطلاعاتی یک سیستم اسپین $1/2$ با حالت $|y\rangle=\cos\frac{q}{2}|+\rangle+\sin\frac{q}{2}|-\rangle$ را حساب کنیم، عبارت زیر برقرار می‌شود:

$$H_{\text{total}}=H_1(p_x^+,p_x^-)+H_2(p_y^+,p_y^-)+H_3(p_z^+,p_z^-) \qquad (10\text{-}3)$$

---

[31] Von Neumann

[32] دربارهٔ آنتروپی فون‌نیومن و کاربرد آن در مکانیک کوانتومی در قسمت‌های بعد به تفصیل بحث خواهد شد.

[33] Complementarity

$$H_{\text{total}} = -\frac{1-\sin q}{2}\log\frac{1-\sin q}{2} - \frac{1+\sin q}{2}\log\frac{1+\sin q}{2}$$
$$-\cos^2\frac{q}{2}\log(\cos^2\frac{q}{2}) - \sin^2\frac{q}{2}\log\sin^2\frac{q}{2} \qquad (11\text{-}3)$$

از طرفی اگر $|y\rangle$ حول محوری مفروض به اندازه زاویه j دوران داده شود و محتوای اطلاعاتی سیستم با استفاده از آنتروپی شانون محاسبه گردد، H به دست آمده به j وابسته است. بنابراین محتوای اطلاعاتی سیستم اسپین 1/2 اگر با رابطه شانون محاسبه شود، ناوردا نمی باشد و با j تغییر می‌کند که تأکیدی دوباره بر رد آنتروپی شانون در مکانیک کوانتومی است.

3-3- اطلاعات کوانتمی

3-3-1- ماتریس چگالی

از آنجا که ماتریس چگالی در نظریهٔ اطلاعات کوانتمی اهمیت بسزایی دارد بنابراین بهتر است اندکی راجع به ویژگیهای آن بحث کنیم. به طور کلی حالت یک سیستم با ماتریس چگالی نمایش داده می‌شود.

اگر زیر سیستمی از یک سیستم بزرگ‌تر داشته باشیم، حتی اگر حالت[34] سیستم بزرگتر یک «رِی[35]» را نمایش دهد، حالت زیرسیستم لازم نیست با یک «ری» نمایش داده شود. ولی در موردی که حالت زیرسیستم یک ری را نشان دهد گفته می‌شود که حالت خالص می‌باشد. در غیر این صورت حالت آمیخته است. اگر حالت مورد نظر در حالت خالص $|y\rangle_A$ باشد، ماتریس چگالی $r_A = |y\rangle_{A\,A}\langle y|$ می‌باشد. یک ماتریس چگالی کلی را به صورت زیر می‌توان نشان داد:

$$r_A = \sum_a p_a |a\rangle\langle a| \qquad (12\text{-}3)$$

که $\sum_a p_a = 1$, $0 < p_a \le 1$، اگر [9].

[34] حالت توصیف کاملی از سیستم کوانتمی ارائه می‌دهد

[35] در فضای هیلبرت نمایش داده می‌شود. "ray" در مکانیک کوانتمی، یک حالت با یک ری

در بخش‌های پیش اندکی راجع به کیوبیت صحبت کردیم. هر کیوبیت برداری واحد در کره بلاخ[36] می‌باشد. کره بلاخ یک کره سه بعدی به شعاع واحد است. بنابراین هر نقطه $r=(x,y,z)$ با $|r|\le 1$ بر حالت کیوبیتی معتبر منطبق است. چنین نقطه‌هایی با

$$\begin{cases} x = r\sin q\cos j \\ y = r\sin q\sin j \\ z = r\sin q \end{cases} \qquad (13\text{-}3)$$

داده می‌شوند. که $0\le q\le p$ ، $0\le j\le 2p$ و $0\le r\le 1$ می‌باشد. نقاط روی سطح ($|r|=1$) بر حالت-های خالص و نقاطی با $|r|<1$ بر حالت‌های آمیخته منطبق‌اند. مبدأ با حالت کاملاً آمیخته تناظر دارد.

اگر $|r|=1$، حالت خالص $|y\rangle$ منطبق بر r به صورت زیر می‌باشد.

$$|y\rangle=\begin{pmatrix} \cos q/2 \\ e^{ij}\sin q/2 \end{pmatrix} \qquad (14\text{-}3)$$

ماتریس چگالی حالت خالص $|y\rangle$ به صورت زیر نمایش داده می‌شود:

$$r=|y\rangle\langle y|=\frac{1}{2}\begin{pmatrix} 1+z & x-iy \\ x+iy & 1-z \end{pmatrix}=\frac{1}{2}(I+x s_x+y s_y+z s_z) \qquad (15\text{-}3)$$

که I ماتریس یکانی است و

$$s_x=\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},\ s_y=\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix},\ s_z=\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad (16\text{-}3)$$

ماتریس‌های پائولی می‌باشند. با معرفی بردار $s=(s_x,s_y,s_z)$، معادله (16-3) به صورت زیر نوشته می‌شود.

$$r=\frac{1}{2}(I+r.s) \qquad (17\text{-}3)$$

ماتریس چگالی یک حالت آمیخته مخلوطی احتمالی از ماتریس‌های r_i حالت‌های خالص می‌باشند. $r=\sum_{i=1}^{k}c_i r_i$ که $0\le c_i\le 1$ و $\sum_{i=1}^{k}c_i=1$. توجه شود در صورتی که ویژه‌مقادیر r منفی باشند

[36] Bloch sphere

حالت‌های با $|r|\leq 1$ بر کیوبیت‌های معتبر منطبق نیستند. اضافه می‌کنیم که هر کیوبیت باید بهنجار باشد.

$$|y\rangle = a|\uparrow\rangle + b|\downarrow\rangle, \qquad a^2 + b^2 = 1 \tag{18-3}$$

3-3-2 - آنتروپی فون نیومن

آنتروپی شانون عدم قطعیت همراه با توزیع‌های احتمالی کلاسیکی را اندازه می‌گیرد. حالت‌های کوانتمی را می‌توان به روشی مشابه با عملگرهای چگالی که جایگزین توزیع‌های احتمالی می‌شوند، توصیف کرد. آنتروپی فون نیومنِ یک حالت کوانتمیِ r از رابطهٔ زیر به‌دست می‌آید:

$$S(r) = -\mathrm{Tr}(r \log r) \tag{20-3}$$

در رابطهٔ بالا لگاریتم در مبنای 2 گرفته می‌شود و اگر l_x ها ویژه‌مقادیر r باشند، تعریف فون نیومن می‌تواند به صورت زیر بیان شود:

$$S(r) = -\sum_x l_x \log l_x \tag{21-3}$$

که همانند آنتروپی شانون تعریف می‌کنیم: $\mathbf{0}\log\mathbf{0} \equiv \mathbf{0}$. مثلاً عملگر چگالی حالت کاملاً آمیخته در فضای d بعدی برابر $\frac{I}{d}$ بوده و آنتروپی برابر $\log d$ می‌باشد.

3-3-2-1- تعاریفی بر مبنای آنتروپی فون نیومن

1- آنتروپی نسبی کوانتمی. همان گونه که برای آنتروپی نسبی کلاسیکی داشتیم تعریف نسخهٔ کوانتمی آنتروپی نسبی نیز بسیار مفید می‌باشد. فرض کنید r و s عملگرهای چگالی باشند، پس آنتروپی نسبی r,s به صورت زیر تعریف می‌شود:

$$S(r\|s) \equiv \mathrm{Tr}(r\log r) - \mathrm{Tr}(r\log s) \qquad (22\text{-}3)$$

همانند آنتروپی نسبی کلاسیکی آنتروپی نسبی کوانتمی می‌تواند گاهی اوقات نامحدود باشد. به ویژه آنتروپی نسبی هنگامی $+\infty$ (نامحدود) می‌شود که کرنل[37] s (یعنی فضای برداری گسترده شده با ویژه‌بردارهایی از s که ویژه‌مقادیر صفر دارند) با فضای گسترده شده با ویژه‌بردارهایی از r که ویژه-مقادیر آنها مخالف صفر است)[38] اشتراک غیرجزئی داشته باشد. در غیر این صورت محدود است.

قضیهٔ1؛ نامساوی کلین[39]. آنتروپی نسبی کوانتمی نامنفی می‌باشد $S(r\|s) \geq \mathbf{0}$ [5 و 9].

2- نامساوی فانس[40] و پیوستگی آنتروپی. همان‌گونه که در مورد آنتروپی شانون داشتیم، آنتروپی فون نیومن نیز تابعی پیوسته می‌باشد.

اثبات:

فرض کنید r ماتریس چگالی باشد که به مقدار خیلی کوچکی تغییر داده می‌شود. به کمک نامساوی فانس نشان خواهیم داد که تغییر $S(r)$ خیلی کوچک بوده و پیوستگی $S(r)$ را نتیجه می‌گیریم.

قضیه2؛ نامساوی فانس. فرض کنید r و s ماتریس‌های چگالی باشند به‌گونه‌ای که «فاصلهٔ ردی[41]» بین آنها در رابطهٔ $T(r,s) \leq 1/e$[42] صدق کند. بنابراین نامساوی زیر برقرار است

$$\left|S(r)-S(s)\right| \leq T(r,s)\log d + h(T(r,s)) \qquad (23\text{-}3)$$

---

[37] Kernel

[38] r support

[39] Klein's inequality

[1] Fannes' inequality

[41] Trace distance

[42] در ادامه فاصلهٔ ردی (trace distanse) توضیح داده خواهد شد.

در معادلهٔ بالا d بُعد فضای هیلبرت بوده و $h(x) = -x\log x$ و $T(r,s)$ فاصلهٔ ردی می‌باشد.

پیش از پرداختن به بقیهٔ اثبات نامساوی فانس بهتر است اندکی راجع به فاصلهٔ ردی و اهمیتی که در نظریهٔ اطلاعات کوانتمی دارد صحبت کنیم.

معنی این گفته که دو دسته اطلاعات مشابه‌اند چیست؟ یا منظورمان از گفتن اینکه تحت بعضی فرایندها اطلاعات حفظ می‌شوند و یا پایسته‌اند چه می‌تواند باشد؟ این پرسش‌ها در نظریهٔ پردازش اطلاعات کوانتمی پرسش‌هایی اساسی‌اند و هدفِ بحث «اندازهٔ فاصله[43]» این است که به این پرسش‌ها جواب‌های کمّی بدهد.

3-3-2-2- اندازهٔ فاصله برای اطلاعات کلاسیکی

در نظریهٔ اطلاعات کلاسیکی چه اشیائی با هم مقایسه می‌شوند؟ ممکن است مقایسهٔ دو رشتهٔ بیتی مثل **0001**0 و 1**000**11 مد نظر باشد. یک روش برای تعیین تفاوت بین این رشته‌ها استفاده از « فاصلهٔ هامینگ» می‌باشد، که به وسیلهٔ تعداد مکان‌هایی که دو رشتهٔ بیتی مساوی نباشند مشخص می‌شود. برای مثال دو رشتهٔ بیتی بالا در دو مکان اول و آخر با هم فرق می‌کنند، بنابراین فاصلهٔ هامینگ بین آنها برابر دو می‌باشد. متأسفانه فاصلهٔ هامینگ بین دو شئ برچسب‌گذاری ساده‌ای می‌باشد، ولی در گسترهٔ فضای هیلبرت مکانیک کوانتمی چون از پیش هیچ برچسب ویژه‌ای وجود ندارد بنابراین نمی‌توان از فاصلهٔ هامینگ برای تعیین تفاوت بین دو مجموعه اطلاعات کوانتمی یا به عبارتی فاصلهٔ آن‌ها استفاده کرد.

مقایسهٔ اطلاعات کوانتمی با توزیع احتمال‌های کلاسیکی راه بهتری برای مطالعهٔ اندازه‌های فاصله فراهم می‌آورد. در واقع در نظریهٔ اطلاعات کلاسیکی، منبع اطلاعاتی معمولاً به شکل متغیرهای تصادفی مدل‌سازی می‌شود. مثلاً توزیع احتمال می‌تواند روی حروف الفبایی که منبع تولید می‌کند، تعریف شود. برای مثال منبعی ناشناخته از متنی انگلیسی به شکل دنباله‌ای از متغیرهای تصادفی روی حروف الفبا مدل‌سازی می‌شود. و پیش از آنکه متن خوانده شود می‌توان حدس‌های نسبتاً خوبی

[43] Distance measure

در مورد تکرار نسبی حروفی که در متن پدیدار می‌شوند و همچنین ارتباط ویژه‌ای که میان آنها وجود دارد، زد. مانند این واقعیت که جفت حرف‌های ''th'' خیلی بیشتر از جفت حرف‌های ''zx'' در متون انگلیسی ظاهر می‌شوند. این مشخصهٔ منبع اطلاعاتی که به صورت توزیع احتمال‌هایی روی بعضی حروف الفبا تعریف می‌شود، ما را رهنمون می‌کند که هنگام جستجو برای اندازهٔ فاصله، روی توزیع احتمال‌ها تمرکز کنیم.

منظور از بیان اینکه دو توزیع احتمال $\{p_x\}$ و $\{q_x\}$ که روی اندیس مشابه x تعریف می‌شوند با یکدیگر مشابه‌اند، چست؟ روشن است که پاسخ یکتا و درست به این پرسش‌ها مشکل است. بنابراین دو جواب مختلف پیشنهاد می‌شود که هر کدام به طور گسترده در محاسبات و اطلاعات کوانتمی مورد استفاده قرار می‌گیرند. نخستین اندازه « فاصلهٔ ردی» می‌باشد که با معادلهٔ زیر تعریف می‌شود.

$$D(p_x,q_x)\equiv\frac{1}{2}\sum_x|p_x-q_x| \qquad (24\text{-}3)$$

گاهی این معادله را با نام فاصلهٔ L_1 یا فاصلهٔ کولموگروف[44] می‌شناسند.

دومین اندازهٔ فاصله بین دو توزیع احتمال، فیدلیتیِ[45] توزیع‌های احتمالی $\{p_x\}$ و $\{q_x\}$ است و به صورت زیر تعریف می‌شود:

$$F(p_x,q_x)\equiv\sum_x\sqrt{p_x\,q_x} \qquad (25\text{-}3)$$

فیدلیتی نسبت به فاصلهٔ ردی برای اندازه‌گیری فاصلهٔ بین توزیع احتمال‌ها روشی بسیار متفاوت است [5]. یادآوری می‌کنیم که هرگاه $\{p_x\}$ و $\{q_x\}$ برابر باشند $F(p_x,q_x)=1$ [5].

3-3-2-3- اندازه فاصله برای حالت‌های کوانتمی

حالت‌های کوانتمی چقدر به همدیگر نزدیک می‌باشند؟ برای توضیح میزان نزدیکی حالت‌های کوانتمی اندیشهٔ کلاسیکی فاصلهٔ ردی و فیدلیتی برای سیستم‌های کوانتمی تعمیم داده می‌شود.

[44] kolmogorov

[45] Fidelity

الف) فاصلهٔ ردی: بین حالت‌های کوانتمی r و s به صورت زیر تعریف می‌شود:

$$T(r,s)\equiv\frac{1}{2}\mathrm{Tr}(r-s) \tag{26-3}$$

فاصلهٔ ردی کوانتمی تعمیمی است از فاصلهٔ ردی کلاسیکی و هرگاه r و s با همدیگر جابجا شوند فاصلهٔ ردی بین آنها با فاصلهٔ ردی کلاسیکی بین ویژه‌مقادیر آنها برابر است. به طور آشکارتر اگر r و s جابجاپذیر باشند پس در پایه‌های مشابه قطری می‌شوند،

$$T(r,s)\equiv\frac{1}{2}\mathrm{Tr}(r-s)\ \ r=\sum_i r_i|i\rangle\langle i|;\ \ s=\sum_i s_i|i\rangle\langle i| \tag{27-3}$$

که پایه‌های $|i\rangle$ متعامد می‌باشند. بنابراین

$$T(r,s)=\frac{1}{2}\mathrm{Tr}\left|\sum_i(r_i-S_i)|i\rangle\langle i\right|=T(r_i,s_i) \tag{28-3}$$

فیدلیتی کوانتمی: دومین اندازه فاصله بین حالت‌های کوانتمی فیدلیتی کوانتمی می‌باشد. فیدلیتی حالت r و s به صورت زیر تعریف می‌شود.

$$F(r,s)=\mathrm{Tr}\sqrt{r^{1/2}s\ r^{1/2}} \tag{29-3}$$

گرچه از رابطه بالا فوراً مشخص نمی‌شود که این عبارت مفید بوده و کاربردهای زیادی در توضیح مسائل مکانیک کوانتمی دارد ولی در ادامه با ذکر کاربردهایی از این مفهوم اهمیت آن به خوبی روشن خواهد شد. بنابراین به مثال‌هایی از این کمیت خواهیم پرداخت. مثال نخست زمانی است که r و s با هم جابجا شوند، یعنی در پایه‌های مشابه قطری باشند،

$$r=\sum_i r_i|i\rangle\langle i|,\qquad s=\sum_i s_i|i\rangle\langle i| \tag{30-3}$$

پایه‌های $|i\rangle$ متعامد می‌باشند.

$$F(r,s)=\mathrm{Tr}\sqrt{\sum_i r_i s_i|i\rangle\langle i|}\ \mathrm{Tr}(\sum_i\sqrt{r_i s_i}|i\rangle\langle i|)$$
$$=\sum_i\sqrt{r_i s_i}=F(r_i,s_i) \tag{31-3}$$

محاسبات بالا نشان می‌دهد که هرگاه r,s با همدیگر جابجا شوند فیدلیتی $F(r,s)$ به فیدلیتی کلاسیکی $F(r_i,s_i)$ بین توزیع ویژه‌مقادیر s_i و r_i از r و s کاهش پیدا می‌کند.

دومین مثال برای محاسبه فیدلیتی بین یک حالت خالص $|y\rangle$ و یک حالت اختیاری r می‌باشد. از معادله فیدلیتی داریم :

$$F(|y\rangle,r)=tr\sqrt{\langle y|r|y\rangle|y\rangle\langle y|}=\sqrt{\langle y|r|y\rangle} \qquad (32\text{-}3)$$

بنابراین فیدلیتی با ریشه مربعی همپوشانی بین $|y\rangle$ و r برابر است . این نتیجه‌ای مهم است که از آن استفاده‌های بسیاری می‌شود[5 و 9].

اکنون به اثبات قضیهٔ پیوستگی آنتروپی و نامساوی فانس می‌پردازیم.

اگر در معادلهٔ (23-3) محدودیت $T(r,s)\le 1/e$ را جایگزین کنیم به نامساوی ضعیف‌تر زیر می‌رسیم،

$$|S(r)-S(s)|\le T(r,s)\log d+\frac{1}{e} \qquad (33\text{-}3)$$

برای اثبات نامساوی فانس به نتیجهٔ ساده‌ای که فاصلهٔ ردی بین دو عملگر را به ویژه‌مقادیرشان ربط می‌دهد نیاز داریم. فرض کنید $r_1\ge r_2\ge\ldots\ge r_d$ ویژه‌مقادیر r به ترتیبِ نزولی باشد و $s_1\ge s_2\ge\ldots\ge s_d$ ویژه‌مقادیر s باز هم به ترتیبِ نزولی باشد. با توجه به تجزیهٔ طیفی[46] که در پیوست ب توضیح داده شده، $r-s=Q-R$ است. Q و R عملگرهای مثبتی هستند که ویژه فضای برداری آنها با شرط اینکه ویژه‌مقادیرشان صفر نباشد بر هم عمودند. بنابراین $T(r,s)=\mathrm{Tr}(R)+\mathrm{Tr}(Q)$. با تعریف $V\equiv R+r=Q+s$، رابطهٔ زیر برقرار است:

$$T(r,s)=\mathrm{Tr}(R)+\mathrm{Tr}(Q)=\mathrm{Tr}(2V)-\mathrm{Tr}(r)-\mathrm{Tr}(s) \qquad (34\text{-}3)$$

اکنون $t_1\ge t_2\ge\ldots\ge t_d$ را ویژه‌مقادیر T در نظر می‌گیریم. می‌دانیم که $t_i\ge\max(r_i,s_i)$ می‌باشد بنابراین $2t_i\ge r_i+s_i+|r_i-s_i|$.پس نتیجه می‌گیریم که:

[46] Spectral decomposition

$$T(r,s) \geq \sum_i |r_i - s_i| \qquad (35\text{-}3)$$

هرگاه $|r-s| \geq 1/2$ باشد، با محاسبه نتیجه می‌گیریم که $h(r)-h(s) \leq h|r-s|$. با کمی تأمل در می‌یابیم که به‌ازای همهٔ iها که $|r_i - s_i| \leq 1/2$، بنابراین

$$|S(r)-S(s)| = \left|\sum_i (h(r_i)-h(s_i))\right| \leq \sum_i h(|r_i - s_i|) \qquad (36\text{-}3)$$

با جایگزینی $\Delta \equiv \sum_i |r_i - s_i|$ و مشاهدهٔ $h(|r_i - s_i|) = \Delta h(|r_i - s_i|/\Delta) - |r_i - s_i| \log \Delta$ درمی‌یابیم که:

$$|S(r)-S(s)| \leq \Delta \sum h(|r_i - s_i|/\Delta) + h(\Delta) \leq \Delta \log d + h(\Delta) \qquad (37\text{-}3)$$

نامساوی دوم در معادلهٔ بالا از این موضوع به‌دست آمد که اگر X یک متغیر تصادفی با خروجی d باشد پس $H(X) \leq \log d$ است. اما با بکارگیری $T(r,s) \geq \sum_i |r_i - s_i|$، نامساوی $\Delta \leq T(r,s)$ برقرار می‌شود، بنابراین به کمک یکنوا بودن $h(\cdot)$ روی بازهٔ $[0, 1/e]$، به رابطهٔ زیر می‌رسیم:

$$|S(r)-S(s)| \leq T(r,s) \log d + h(T(r,s)) \qquad (38\text{-}3)$$

با شرط $T(r,s) \leq 1/e$، به نامساوی فانس دست می‌یابیم.

3-3-3- ویژگی‌های اصلی آنتروپی فون‌نیومن

1. آنتروپی نامنفی می‌باشد.
2. اگر و تنها اگر حالت خالص بوده و ماتریس چگالی بر اساس حالت پایه مورد نظر بسط داده شود برابر صفر می‌باشد.
3. در یک فضای هیلبرت d بعدی آنتروپی حداکثر می‌تواند برابر $\log d$ باشد. این امر در صورتی رخ می‌دهد که سیستم در حالت کاملاً آمیخته باشد.
4. اگرسیستم مرکب AB در حالت خالص باشد، بنابراین $S(A) = S(B)$
5. اگر فرض کنیم p_i ها احتمال باشند و روی زیر فضاهای متعامد تعریف شوند، پس:

$$S(\sum_i p_i r_i) = H(p_i) + \sum_i p_i S(r_i) \qquad (38\text{-}3)$$

5. نظریهٔ آنتروپی توأم[47]: با فرض آنکه p_i ها احتمال و $|i\rangle$ها حالت‌های متعامد برای سیستم A و r_i عملگرهای چگالی برای سیستم دیگری مثل B باشند، پس آنتروپی توأم به صورت زیر تعریف می‌شود:

$$S\left(\sum_i p_i |i\rangle\langle i| \otimes r_i\right) = H(p_i) + \sum_i p_i S(r_i) \qquad (40\text{-}3)$$

3-3-4- آنتروپی واندازه‌گیری

هرگاه اندازه‌گیری روی سیستم کوانتمی انجام گیرد، آنتروپی سیستم چگونه تغییر می‌کند؟ بسیار روشن است که پاسخ این پرسش به گونهٔ اندازه‌گیری که ممکن است روی سیستم اعمال شود بستگی دارد، ولی به طور کلی می‌توان بعضی ادعاهای شگفت‌انگیز دربارهٔ چگونگی رفتار آنتروپی تحت اندازه‌گیری‌های مختلف ارائه داد.

[47] Joint entropy

3-3-4-1- اندازه‌گیری تصویری و آنتروپی

اگر اندازه‌گیری تصویری که با عملگر p_i توصیف می‌شود، روی سیستم اعمال کنیم و نتیجهٔ اندازه‌گیری به هیچ وجه معلوم نباشد، آنگاه اگر حالت سیستم پیش از اندازه‌گیری با r نمایش داده شود، حالت سیستم پس از اندازه‌گیری به صورت زیر داده می‌شود:

$$r' = \sum_i P_i r P_i \qquad (41\text{-}3)$$

قضیهٔ زیر نشان می‌دهد که اگر اندازه‌گیری تصویری باشد آنتروپی هرگز کاهش نمی‌یابد.

قضیه3؛ اندازه‌گیری تصویری آنتروپی را افزایش می‌دهد. فرض کنید P_i یک سری کامل از تصویرگرهای متعامد و r عملگر چگالی باشد، پس آنتروپی سیستم پس از اندازه‌گیری حداقل می‌تواند مساوی آنتروپی پیش از اندازه‌گیری باشد،

$$S(r') \geq S(r) \qquad (42\text{-}3)$$

اثبات:

با بکارگیری نامساوی کلین، رابطهٔ زیر برقرار است:

$$\mathbf{0} \leq S(r' \| r) = -S(r) - \mathrm{Tr}(r \log r') \qquad (43\text{-}3)$$

اگر ثابت کنیم که $-\mathrm{Tr}(r \log r') = S(r')$، به نتیجهٔ مطلوب رسیده‌ایم. برای اینکه بتوانیم معادلهٔ پیشین را اثبات کنیم از رابطه کامل بودن $\sum_i P_i = I$ و $P_i{}^2 = P_i$ و ویژگی چرخه‌ای رد استفاده می‌کنیم، پس به‌دست می‌آوریم:

$$-\mathrm{Tr}(r \log r') = -\mathrm{Tr}(\sum_i P_i \; r \log r') = -\mathrm{Tr}(\mathrm{Tr} \sum P_i \; r \log r' P_i). \qquad (44\text{-}3)$$

با توجه به اینکه $r' P_i = P_i r P_i = P_i r'$. از آنجا که P_i با r' جابجا می‌شود، پس با $\log r'$ نیز جابجا می‌شود و بنابراین رابطهٔ زیر برقرار است:

$$-\mathrm{Tr}(r \log r') = -\mathrm{Tr}\left(\sum P_i \; r P_i \log r' \right) = -\mathrm{Tr}(r' \log r') = S(r') \qquad (45\text{-}3)$$

از (3-43) و (3-45) نتیجه می‌گیریم که اندازه‌گیری تصویری آنتروپی را افزایش می‌دهد در حالی که اندازه‌گیری تعمیم یافته[48] می‌تواند باعث کاهش آنتروپی شود.

3-4- ذخیره و انتقال اطلاعات در کانال‌های کوانتمی بدون نوفه

نظریهٔ اطلاعات کلاسیکی قویاً به مسألهٔ انتقال اطلاعات کلاسیکی وابسته است که این اطلاعات میان کانال‌های ارتباطی، که از قانون‌های فیزیک کلاسیک تبعیت می‌کنند، انتقال می‌یابند.
نظریهٔ اطلاعات کوانتمی با مطالعهٔ کانال‌های ارتباطی بیشتر برانگیخته می‌شود. به طور اساسی هنگام کار روی نظریهٔ اطلاعات کوانتمی سه هدف دنبال می‌شود: 1- تشخیص طبقه‌های ابتدایی منبع‌های استاتیک در مکانیک کوانتومی که نوع اطلاعات را مشخص می‌کند. 2- تشخیص طبقه‌های ابتدایی منبع‌های دینامیکی که نوع پردازش اطلاعات را مشخص می‌کند.3- تعیین منبع‌های سبک و سنگین کنندهٔ فرایندهای دینامیکی انجام‌شده. نظریهٔ اطلاعات کوانتمی غنی‌تر از نظریهٔ اطلاعات کلاسیکی است چون مکانیک کوانتمی طبقه‌های بیشتری از منبع‌های دینامیکی وکلاسیکی را دربر دارد.

پیش از پرداختن به مسألهٔ ذخیره و انتقال اطلاعات در سیستم‌های کوانتومی به دو مفهوم مهم در نظریهٔ اطلاعات کوانتمی می‌پردازیم که ما را در درک بقیه این بخش یاری می‌کنند.

3-4-1- اطلاعات قابل دستیابی وتمیز دادن حالت‌های کوانتمی[49]

مثال زیر تفاوت‌های بارز بین اطلاعات کلاسیکی وکوانتمی را بسیار خوب نشان می‌دهد. فرض کنید آلیس منبع کلاسیکی از اطلاعات در اختیار دارد که نمادهای $X = \mathbf{0},...,n$ را با احتمال‌های $P_0,...,P_n$ آماده می‌کند. سپس آلیس یک حالت کوانتمی r_x از سری ثابت $r_0,...,r_n$ انتخاب کرده و تحویل باب می‌دهد. هدف این است که باب مقدار X را به گونه‌ای تعیین کند که به واقعیت نزدیکتر باشد. بنابراین باب یک اندازه‌گیری کوانتمی روی حالتی که تحویل گرفته انجام می‌دهد و سپس تلاش می‌کند هویت X را با خروجی اندازه‌گیری خودش که Y می‌باشد تعیین کند. محتوای

[48] Generalized mesurement

[49] Distinguishing quantum states and the accessible information

«اطلاعات دوجانبه»، $H(X:Y)$، اندازهٔ خوبی از مقدار اطلاعاتی است که باب در اندازه‌گیری کسب کرده است. نامساوی پردازش اطلاعات بیان می‌کند که تنها هنگامی باب می‌تواند X را از Y به‌دست آورد که $H(X:Y)=H(X)$ باشد. ولی به‌طور کلی داریم $H(X:Y)\le H(X)$. در واقع نزدیکی $H(X:Y)$ به $H(X)$ اندازهٔ کمّی از این واقعیت است که باب تا چه اندازه می‌تواند X را تعیین کند. پس هدف باب انتخاب اندازه‌گیری است که $H(X:Y)$ را بیشینه کند. بنابراین اطلاعات قابل دستیابی به صورت بیشینه اطلاعات دوجانبه $H(X:Y)$ روی همهٔ طرح‌های ممکن اندازه‌گیری تعریف می‌شود. در واقع اطلاعات قابل دستیابی اندازه‌ای است از اینکه باب تا چه حد می‌تواند X را تعیین کند. در نظریهٔ اطلاعات کلاسیکی دستیابی به اطلاعات اگرچه مشکل باشد، اما در اصل همیشه ممکن است. ولی در نظریهٔ اطلاعات کوانتمی تمیز دادن حالت‌های مجزا همیشه ممکن نیست. مثلاً در مکانیک کوانتومی روشی برای تشخیص حالت‌های کوانتمی نامتعامد وجود ندارد[3]. اگر آلیس دو حالت $|y\rangle$ و $|f\rangle$ را به ترتیب با احتمال‌های p و $1-p$ آماده کند اطلاعات قابل دستیابی از این آماده‌سازی لزوماً کمتر از $H(p)$ است. اما اگر حالت‌های کلاسیکی بیت‌های 1 و **0** را با احتمال‌های p و $1-p$ آماده کند دلیل اساسی برای اینکه باب نتواند این حالت‌ها را تمیز دهد وجود ندارد. بنابراین اطلاعات قابل دستیابی مشابه آنتروپی آماده سازی یعنی $H(p)$ است.

3-4-1-1- نظریهٔ تولید مثل ناپذیری[50]

نظریهٔ تولید مثل ناپذیری بیان می‌کند که حالت‌های کوانتمی نامتعامد به هیچ روشی قابل کپی کردن نیستند. ولی حالت‌های متعامد را می‌توان کپی کرد و این با قضیه‌ای که فیزیک کلاسیک موردی خاص از مکانیک کوانتمی است منافاتی ندارد زیرا حالت‌های کلاسیکی را می‌توان حالت‌های متعامدی در نظر گرفت که قابل کپی کردن می‌باشند.

چه ارتباطی بین قابل دستیابی بودن اطلاعات و تولید نسل یا کپی‌پذیر بودن حالت‌ها وجود دارد؟ فرض کنید آلیس یکی از دوحالت کوانتمی $|y\rangle$ و $|f\rangle$ را به ترتیب با احتمال‌های p و $1-p$ آماده کرده و تحویل باب می‌دهد. اگر موردی را فرض کنیم که اطلاعات قابل دستیابی باب $H(p)$ باشد. یعنی قانون‌های مکانیک کوانتومی به باب اجازه داده باشند که با اندازه‌گیری به اندازهٔ کافی اطلاعات کسب کرده و تعیین کند که آلیس کدام یک از حالت‌های $|y\rangle$ و $|f\rangle$ را آماده کرده‌است، پس باب حالت‌ها را به روشی خیلی ساده تولید مثل کرده است. در واقع او با اندازه‌گیری که انجام داده است توانسته است حالت‌ها را تشخیص دهد و می‌تواند چندین کپی از حالت‌های $|y\rangle$ و $|f\rangle$ که آلیس به او داده آماده کند. بنابراین نظریهٔ تولید مثل ناپذیری نتیجه‌ای از این واقعیت است که اطلاعات قابل دسترس برای حالت‌های نامتعامد اکیداً کمتر از $H(p)$ می‌باشد. یا اینکه اطلاعات قابل‌دستیابی همیشه کمتر از $H(p)$ است نتیجه‌ای از نظریهٔ تولید مثل ناپذیری می‌باشد. با استفاده از نظریهٔ تولید مثل ناپذیری و مبحث اطلاعات قابل دستیابی می‌توان نتیجه گرفت که اگر سیگنال‌های فرستاده شده از طرف باب متعامد نباشند همیشه $S(r)<H(p)$ می‌باشد[21]. در این حالت که سیگنال‌های فرستاده شده متعامد نیستند هیچ روشی وجود ندارد که بتواند اطلاعات فرستاده شده را به طور کامل کدگشایی کند.

متأسفانه هیچ روشی برای محاسبهٔ اطلاعات قابل دستیابی وجود ندارد ولی قید هالوو حد بالایی بر میزان اطلاعات قابل دستیابی ممکن اعمال می‌کند

[50] No-cloning

3-4-1-2- قید هالوو

قید هالوو همان‌طور که گفتیم حد بالایی برای میزان اطلاعات قابل دستیابی قرار می‌دهد و این قید اهمیت بسزایی در نظریهٔ اطلاعات کوانتمی دارد.

فرض کنید آلیس حالت r_x را که $X = \mathbf{0},...,n$ می‌باشد با احتمال $p_0,...,p_n$ آماده کند. باب اندازه‌گیری‌ای را که با عناصر POVM، $\{E_y\}=\{E_0,...,E_m\}$ توصیف می‌شود روی آن حالت‌ها اعمال می‌کند و خروجی اندازه‌گیری Y به‌دست می‌آید. قید هالوو بیان می‌کند که برای اندازه‌گیریی که باب انجام می‌دهد

$$H(X:Y) \le S(r) - \sum_x p_x S(r_x) \qquad (51\text{-}3)$$

در رابطه بالا $r = \sum_x p_x r_x$ می‌باشد. این قید گاهی با c نشان داده می‌شود[3 و9].

3-4-2- نظریهٔ کد کردن کانال بدون نوفه شوماخر[51]

3-4-2-1- تبدیلات اطلاعات از منبع تا دریافت‌کننده

فرض کنید کانالی مانند X داشته باشیم که برای انتقال اطلاعات کوانتمی از آن استفاده می‌شود. سیگنال ورودی کانال، از منبعی مانند M گرفته می‌شود و از طریق کانال به دریافت‌کننده‌ای که برای مثال M' نامیده می‌شود، می‌رود. بنابر نظریهٔ تولید مثل ناپذیری، عمل گذار باید تبدیلی یکانی باشد. بنابراین ضرب داخلی حالت‌هایی که در M تولید می‌شوند باید با ضرب داخلی این حالت‌ها هنگامی که وارد کانال X شده‌آند برابر باشد. مثلاً اگر حالت‌های اختیاری $|b_M\rangle$ و $|a_M\rangle$ در ابتدا تولید شده باشند باید $\langle a_M | b_M \rangle = \langle a_X | b_X \rangle$ برقرار باشد. این امر در صورتی امکان‌پذیر است که بعد فضای هیلبرت H_X حداقل با بعد فضای هیلبرت H_M برابر باشد. برای مشخص کردن تحول یکانی U که عمل گذار را انجام می‌دهد، تنها لازم است که چگونگی نگاشت پایه‌های متعامد H_M به پایه‌های متعامد H_X را تعیین کنیم. عمل گذار معکوس پذیر است. چون حالت‌های سیگنال می‌توانند

[51] Schumacher

با یک تبدیل یکانی U^{-1} از X به M برگردانده شوند. بنابراین سیستم منبع به وسیله تبدیل یکانی U به سیستم کد شده X می‌رود و سپس تبدیل معکوسی برای برگرداندن سیگنال حالت از X به M' بکار گرفته می‌شود. می‌توان گفت که X کانالی ارتباطی بین منبع و دریافت‌کننده می‌باشد. این گفته به طور شماتیک به صورت $M \xrightarrow[U]{} X \xrightarrow[U^{-1}]{} M'$ نمایش می‌دهیم. چون روی گذار حالت‌های کوانتمی تمرکز کرده‌ایم بنابراین لازم است که کانال کوانتمی X به اندازه کافی بزرگ باشد تا برای نمایش سیگنال M مناسب باشد. یعنی برای یک گذار کامل $\dim H_X \geq \dim H_M$ است. اگرچه ممکن است گذار کامل غیر ممکن باشد و با کانال X تنها گذاری تقریبی از M به M' صورت گیرد. بسته به مشخصه‌های منبع سیگنال، ممکن است بتوانیم از کانال کوچکتری استفاده کنیم و بتوانیم سیگنال‌ها را به طور مناسب انتقال دهیم.

3-4-2-2- گذار تقریبی و فیدلیتی

کانال ارتباطی کوانتمی مانند آنچه در بالا خلاصه شد، در نظر بگیرید. حالت سیگنال M به طور یکانی به X تبدیل شده و می‌تواند در سیستم M' بازیابی شود. اگرچه در عمل بهتر است این‌گونه فرض کنیم که همه سیستم X از انتقال دهنده به دریافت کننده منتقل نمی‌شود. در واقع بهتر است سیستم X را مرکب از دو زیرسیستم در نظر بگیریم که آن‌ها را C و E می‌نامیم. تنها زیرسیستم C به دریافت کننده M' منتقل شده و برای کدگشایی مورد استفاده قرار می‌گیرد. زیر سیستم E درحکم عضوی اضافی نادیده گرفته شده و دور انداخته می‌شود. آشکارا سیگنال اولیه ممکن است به طور دقیق از C تنها به دست نیاید. زیرا ممکن است $\dim H_C < \dim H_M$ باشد. از طرفی ممکن است بعضی سیگنال‌های تقریبی از بازیافت C حاصل شود. بنابراین کانالی از این نوع را کانال انتقال تقریبی از M به M' می‌نامیم. برای دریافت سیگنال از C به M' به زیرسیستم C یک سیستم کمکی E' اضافه می‌کنیم که یک کپی از زیرسیستم نادیده گرفته شده E می‌باشد. و سپس گذاری از $C+E'$ به M با عملگر تبدیل یکانی U' انجام می‌دهیم.

$$M \xrightarrow[U]{} C + E \underset{E}{\downarrow} C \underset{E'}{\uparrow} C + E' \xrightarrow[U']{} M' \qquad (52\text{-}3)$$

ممکن است گزینش $U' = U^{-1}$ مناسب به نظر آید. برای تعیین کارایی طرح گذار معرفی شده، لازم است که فیدلیتی اندازه‌گیری شود. فرض کنید $|a_M\rangle$ سیگنال اولیه M باشد که با عملگر چگالی $r_a = |a_M\rangle\langle a_M|$ قابل نمایش است. سیگنال نهایی در M' با عملگر چگالی w_a نمایش داده می‌شود. چون حالت سیگنال خروجی لزوماً خالص نمی‌باشد، بنابراین w_a در حالت کلی یک عملگر تصویر نمی‌باشد.

برای نشان دادن میزان نزدیکی حالت اولیه و حالت نهایی، با مشاهده‌پذیر p_a یک اندازه‌گیری روی حالت نهایی انجام می‌دهیم. این اندازه‌گیری را گاهی اندازه‌گیری اعتبار سنجی می‌نامند. اندازه‌گیری دو نتیجه ممکن می‌تواند داشته باشد: مقدار1 نشان می‌دهد که سیگنال حالت نهایی بر سیگنال اولیه منطبق است؛ یا **0**، که نشان می‌دهد سیگنال نهایی با سیگنال اولیه تفاوت دارد. احتمال اینکه سیگنال نهایی از این آزمایش اعتبار سنجی با موفقیت عبور کند، از رابطه $\mathrm{Tr}p_a w_a$ به دست می‌آید. پس می‌توان تصور کرد اگر سیگنالی از آنسامبل M داشته باشیم که به M' انتقال داده می‌شود، آنگاه فیدلیتی F را می‌توان احتمال کل مربوط به عبور سیگنال M' از آزمون اعتبار سنجی در مقایسه با سیگنال اولیه تعریف کرد.

$$F = \sum_a p(a)\,\mathrm{Tr}p_a w_a \qquad (53\text{-}3)$$

فیدلیتی آشکارا بین صفر و یک می‌باشد. و در صورتی که گذار کاملی داشته باشیم برابر یک می‌باشد. اگر سیگنال‌هایی با احتمال بزرگ $p(a)$ هنگام گذار به اندازه کمی مختل شوند باز هم فیدلیتی نزدیک به یک می‌باشد، و w_a بسیار نزدیک به p_a است. همچنین اگر سری سیگنال‌هایی که زیاد مختل می‌شوند، احتمال کوچکی داشته باشند بار دیگر فیدلیتی به 1 نزدیک بوده و w_a تقریباً برابر p_a است.

ردیابی چگونگی تغیر سیگنال در این کانال شایان توجه است. در مرحله نخست گذار کد کردن یکانی از M به $C+E$ با عملگر U انجام می‌شود. اگر p_a حالت سیگنال اولیه M باشد پس می‌توان حالت سیگنال $C+E$ را با $\Pi_a = U p_a U^{-1}$ نشان داد. زمانی که زیرسیستم E نادیده گرفته شود، بقیه سیستم به حالت $\mathrm{Tr}_E \Pi_a$ واگذار می‌شود. تریس جزئی Π_a روی E گرفته می‌شود. پس از اینکه E' به سیستم C بپیوندد سیستم مرکب در حالت $\mathrm{W}_a = \mathrm{Tr}_E \Pi_a \otimes \left| o_E \right\rangle \left\langle o_E \right|$ قرار می‌گیرد. سرانجام گذار کدگشایی یکانی اتفاق می‌افتد و $w_a = U' W_a (U')^{-1}$ می‌شود. فرض کنید به طور منطقی $U' = U^{-1}$ برگزیده شود، یعنی گذار کدگشایی عملگری معکوس گذار کدکننده در نظر گرفته شود. پس فیدلیتی به صورت زیر می‌باشد:

$$F = \sum_a p(a) \mathrm{Tr} p_a w_a \qquad (54\text{-}3)$$
$$= \sum_a p(a) \mathrm{Tr}(U^{-1} p_a U)\left(U^{-1} W_a U\right) = \sum_a p(a) \mathrm{Tr} p_a W_a$$

بنابراین فیدلیتی گذار از M به M' تنها با امتحان کردن حالت‌های سیگنال $C+E$ و $C+E'$ محاسبه می‌شود.

3-4-2-3- اصول موضوعه فیدلیتی

به طور شهودی مشخص است که اگر زیرسیستم C کانال خیلی کوچک باشد، فیدلیتی باید به صفر نزدیک باشد و در نتیجه اگر C به اندازه کافی بزرگ باشد، می‌توان فیدلیتی را به یک نزدیک کرد. در اینجا به یک جفت اصول موضوعه فیدلیتی می‌پردازیم که ما را در اثبات نظریهٔ کد کردن کوانتمی یاری می‌کنند. ابتدا با این فرض که کانال C خیلی کوچک باشد آغاز کرده و قضیه زیر را ثابت می‌کنیم.

اصل موضوعه 1. فرض کنید $\dim H_C = d$ بوده و آنسامبل سیگنال‌های موجود M که با $r = \sum_a p(a) p_a$ نمایش داده می‌شوند، دارای خصوصیتی باشد که برای هر عملگر تصویر Γ روی یک زیر فضای d بعدی H_M رابطهٔ زیر برقرار باشد :

$$\mathrm{Tr}\, r\, \Gamma < h \tag{55-3}$$

h ثابت می‌باشد. پس فیدلیتی $F < h$ می‌باشد.

اگر آنسامبل سیگنال در زیرفضاهای زیادی با اندازه‌ای مشابه H_C، وزن کوچکی داشته باشد بنابراین متناظراً فیدلیتی گذار کوچک خواهد بود. سیگنال حالت $\left| a_M \right\rangle$ را در نظر بگیرید که بر طبق طرح کلی که در بخش پیشین اشاره شد انتقال داده می‌شود. با این فرض که E' ابتدا در حالت خالص $\left| o_{E'} \right\rangle$ باشد، حالت سیگنال W_a مربوط به $C+E'$ تنها تحت زیرفضای dبعدی $H_{C+E'}=H_C \otimes H_{E'}$ حمایت می‌شود. این زیر فضا مربوط به حالت‌هایی به شکل $\left| y_C, o_{E'} \right\rangle$ می‌باشد.

بنابراین حالت سیگنال کدگشایی شده w_a از M' تنها روی زیرفضای d بعدی $H_M{}'$ تعریف می‌شود. یادآوری می‌شود که عملگر تصویر روی این زیرفضا Γ می‌باشد.

فرض کنید $\left| f_k \right\rangle$، $k=1,...,d$، پایه‌های متعامد برای زیرفضایی که از ویژه‌مقادیر w_a تشکیل شده است باشد. پس w_a می‌تواند به شکل زیر نوشته شود:

$$w_a = \sum_k q_k \left| f_k \right\rangle \left\langle f_k \right| \tag{56-3}$$

که q_k ویژه‌مقادیر w_a بوده و دربردارندهٔ همه موردهای مخالف صفر می‌باشد و آشکارا $q_k \le 1$.

عملگر تصویر Γ به صورت زیر نمایش داده می‌شود.

$$\Gamma = \sum_k \left| f_k \right\rangle \left\langle f_k \right| \tag{57-3}$$

حال جملهٔ $\mathrm{Tr} p_a w_a$ را که در عبارت فیدلیتی ظاهر می‌شود در نظر بگیرید.

$$\mathrm{Tr} p_a w_a = \mathrm{Tr} p_a \left[w_a = \sum_k q_k \left| f_k \right\rangle \left\langle f_k \right| \right] = \sum_k q_k \mathrm{Tr} p_a \left| f_k \right\rangle \left\langle f_k \right| \tag{57-3}$$

$$\le \sum_k \mathrm{Tr} p_a \left| f_k \right\rangle \left\langle f_k \right| = \mathrm{Tr} p_a \left[\sum_k \left| f_k \right\rangle \left\langle f_k \right| \right] = \mathrm{Tr} p_a \Gamma$$

فیدلیتی F به دست می‌آید:

$$F=\sum_a p(a)\mathrm{Tr}\,p_a w_a \le \sum_a p(a)\mathrm{Tr}\,p_a\Gamma=\mathrm{Tr}\left[\sum_a p(a)p_a\right]\Gamma=\mathrm{Tr}\,r\,\Gamma \qquad (58\text{-}3)$$

و بنابراین $F<h$.

اگر سیستم E' در حالت خالص نباشد، حالت سیگنال نهایی w_a مخلوطی از حالت‌ها می‌باشد که هر کدام از آن‌ها روی زیرفضای d بعدی حمایت می‌شوند. فیدلیتی کل میانگین وزنی همه جمله-هایی است که در شرایط بالا مقید شده‌اند.

شایان ذکر است که برای همه تصویرگرهای Γ می‌توان شرط لازم $\mathrm{Tr}\,r\,\Gamma<h$ را در جمله‌هایی از ویژه‌مقادیر r بازنویسی کرد. اگر P_n ویژه‌مقادیر r و $|n\rangle$ ویژه‌حالت‌های متناظر باشد، کمیت‌های $Q_n=\langle n|\Gamma|n\rangle$، شرط‌های $0\le Q_n\le 1$ و $\sum_n Q_n=\mathrm{Tr}\Gamma=d$ را برآورده می‌کند. پس

$$\mathrm{Tr}\,r\,\Gamma=\sum_n P_n Q_n \qquad (59\text{-}3)$$

به سادگی دیده می‌شود هنگامی که Q_n به‌ازای همه مقادیر n منطبق بر بزرگ‌ترین ویژه‌مقادیر P_n، برابر 1 برگزیده شود، جمع بالا بیشینه می‌شود و به‌ازای دیگر مقادیر n صفر می‌شود. اگر Γ به صورتی برگزیده شود که عملگر تصویر روی زیر فضای گسترده شده با ویژه‌حالت‌هایی باشدکه بر بزرگ‌ترین ویژه‌مقادیر r منطبق‌اند، به مقدار بیشینه اشاره شده دست می‌یابیم. بنابراین اگر و تنها اگر جمع هر ویژه‌مقدار d مربوط به r کمتر از h باشد، عبارت $\mathrm{Tr}\,r\,\Gamma<h$ برقرار است.

اکنون به موردی برمی‌گردیم که کانال C به اندازه کافی بزرگ می‌باشد و گذاری با فیدلیتی بالا مجاز می‌شود.

اصل موضوعه 2. فرض کنید که $\dim H_C=d$ باشد و عملگر تصویر Γ روی زیرفضای d بعدی H_M چنان باشد که $\mathrm{Tr}\,r\,\Gamma>1-h$. پس طرح گذاری با فیدلیتی $F>1-2h$ وجود دارد. اگر آنسامبل سیگنال به طور کافی روی زیرفضایی با اندازه مشابه H_C حمایت شود، پس می‌توان گذاری با فیدلیتی بسیار نزدیک به 1 ساخت.

اگر عملگر Γ تصویری باشد روی زیرفضای Λ از H_M که با ویژه‌حالت‌های d مربوط به r گسترده شده است به کلیت مسأله لطمه‌ای وارد نمی‌شود. ویژه‌حالت‌های r که Λ را می‌گسترانند با $|1\rangle,\mathbf{K},|d\rangle$ نشان داده شوند در این صورت $|d+1\rangle,\mathbf{K},|D\rangle$ بر Λ عمود بوده و با $\Lambda^{\perp}$ نمایش داده می‌شود. پس:

$$\sum_{n=1}^{d}|n\rangle\langle n|=\Gamma \tag{60-3}$$

$$\sum_{n=1}^{d}P_n>1-h \tag{61-3}$$

$$\sum_{n=d+1}^{D}P_n<h \tag{62-3}$$

رهیافت مورد نظر در زیر آورده شده است. ویژه‌حالت‌های r را که با Λ مشترک‌اند، انتقال می‌دهیم، به‌گونه‌ای که به خوبی با حالت‌هایی از کانال C نمایش داده و به درستی در M' بازسازی شوند. چون $\dim\Lambda=\dim H_C$ است، انجام دادن این کار ممکن است. اگرچه، ویژه‌حالت‌های r لزوماً حالت-های سیگنال نمی‌باشند، و هیچ تضمینی نداریم که واقعاً همه سیگنال درون Λ قرار گیرند و بنابراین گذار بدون اختلال صورت گیرد. با این وجود چون Λ بیشترین وزن آنسامبل سیگنال را شامل می-شود (به‌جز برای تکه‌های کوچکی از اندازه‌گیری کمتر از h) می‌توان نشان داد که برای رسیدن به فیدلیتی لازم، سیگنال‌ها نزدیک به زیرفضای Λ می‌باشد.

برای مشخص کردن تبدیلات یکانی U که گذار کدکردن را انجام می‌دهد، در واقع تعیین می‌کنیم که چگونه پایه‌های متعامد r به حالت‌های متعامد $C+E$ نگاشت می‌شود. نگاشت زیر را در نظر بگیرید:

$$|n\rangle\begin{cases}|n_C,o_E\rangle, & n=1,\ldots,d\\ |o_C,n_E\rangle, & n=d+1,\ldots,D,\end{cases} \tag{63-3}$$

که $|n_C\rangle$ و $|n_E\rangle$ به ترتیب سری‌های متعامد حالت‌های سیستم C و E و $|o_C\rangle$ و $|o_E\rangle$ حالت-های خنثی ثابت می‌باشند. لازم است که حالت‌های خنثی $|o_E\rangle$ بر هر یک از حالت‌های $|n_E\rangle$

برای $n = d+1,...,D$ عمود باشد. می‌توان گفت که حالت‌های Λ به حالت‌هایی از C و حالت‌های در $\Lambda^{\perp}$ به حالت‌هایی از E نگاشت می‌شوند. به طور کلی‌تر تمایز بین حالت‌هایی در Λ اکنون بین حالت‌هایی از C ساخته می‌شود، و تمایز بین حالت‌هایی در $\Lambda^{\perp}$ اکنون بین حالت‌هایی در E ساخته می‌شود.

اکنون سیستم اضافی E دور انداخته شده و کپی جدید E' به سیستم پیوند داده می‌شود. مشخص می‌کنیم که E' در آغاز در حالت $\left| o_{E'} \right\rangle$ بوده، چنانکه ویژه‌حالت‌های r اکنون به حالت-های زیر نگاشت می‌شوند:

$$\left| n \right\rangle \begin{cases} \left| n_C, o_{E'} \right\rangle, n = 1, \mathbf{K}, d \\ \left| o_C, n_{E'} \right\rangle, n = d+1, \mathbf{K}, D \end{cases} \qquad (64\text{-}3)$$

سرانجام سیگنال را در M' با استفاده از عملگر معکوس U^{-1} کدگشایی می‌کنیم. حالت سیگنال ویژه $\left| a_M \right\rangle$ در این طرح گذار تقریبی چگونه عبور می‌کند؟ هر حالت M را می‌توان به صورت برهم‌نهی از حالت‌های:

$$\left| a_M \right\rangle = l_a \left| l(a)_M \right\rangle + m_a \left| m(a)_M \right\rangle \qquad (65\text{-}3)$$

نوشت که $\left| l(a)_M \right\rangle$ در Λ و $\left| m(a)_M \right\rangle$ در $\Lambda^{\perp}$ می‌باشد و $\left| l_a \right|^2 + \left| m_a \right|^2 = 1$. حالت‌های $\left| l(a)_M \right\rangle$ و $\left| m(a)_M \right\rangle$ را می‌توان در جمله‌های از ویژه‌پایه‌های r بسط داد:

$$\left| a_M \right\rangle = l_a \left[\sum_{n=1}^{d} \left\langle n \mid l(a)_M \right\rangle \left| n \right\rangle \right] + m_a \left[\sum_{n=d+1}^{D} \left\langle n \mid m(a)_M \right\rangle \left| n \right\rangle \right] \qquad (66\text{-}3)$$

این حالت با گذار کد کردن U به حالت $\left| a_{C+E} \right\rangle$ نگاشت می‌شود، که :

$$\left| a_M \right\rangle \to \left| a_{C+E} \right\rangle = l_a \left[\sum_{n=1}^{d} \left\langle n \mid l(a)_M \right\rangle \left| n_C, o_E \right\rangle \right] \qquad (67\text{-}3)$$

$$+ m_a \left[\sum_{n=d+1}^{D} \left\langle n \mid m(a)_M \right\rangle \left| o_C, n_E \right\rangle \right] = l_a \left| l(a)_C, o_E \right\rangle + m_a \left| o_C, m(a)_E \right\rangle$$

که $\left| l(a)_C \right\rangle$ و $\left| m(a)_E \right\rangle$ تعریفی آشکار دارند. آشکارا می‌دانیم که $\left\langle m(a)_E \left| o_E \right\rangle = \mathbf{0} \right.$ است.

حالت سیگنال $C+E$ درحکم عملگر تصویر $\Pi_a = \left| C+E \right\rangle\left\langle C+E \right|$ نوشته می‌شود. که

$$\Pi_a = \left| l_a \right|^2 \left| l(a)_C, o_E \right\rangle\left\langle l(a)_C, o_E \right| + l_a m_a^* \left| l(a)_C, o_E \right\rangle\left\langle o_C, m(a)_E \right| \qquad (68\text{-}3)$$
$$+ l_a^* m_a \left| o_C, m(a)_E \right\rangle\left\langle l(a)_C, o_E \right| + \left| m_a \right|^2 \left| o_C, m(a)_E \right\rangle\left\langle o_C, m(a)_E \right|$$

هرگاه E نادیده گرفته شود، حالت کانال C با اعمال تریس جزئی روی Π_a به دست می‌آید.

$$\mathrm{Tr}_E \Pi_a = \left| l_a \right|^2 \left| l(a)_C \right\rangle\left\langle l(a)_C \right| + \left| m_a \right|^2 \left| o(a)_C \right\rangle\left\langle o(a)_c \right| \qquad (69\text{-}3)$$

با پیوستن سیستم E' به حالت $\left| o_{E'} \right\rangle$، W_a حاصل می‌شود:

$$W_a = \left| l_a \right|^2 \left| l(a)_C, o_{E'} \right\rangle\left\langle l(a)_C, o_{E'} \right| + \left| m_a \right|^2 \left| o_C, o_{E'} \right\rangle\left\langle o_C, o_{E'} \right| \qquad (70\text{-}3)$$

چون این سیگنال در M با استفاده از عملگر معکوس گذار کدکردن، کدگشایی می‌شود فیدلیتی کل درست برابر با $F = \sum_a p(a)\,\mathrm{Tr}\,\Pi_a W_a$ است. برای سیگنال داده شده:

$$\mathrm{Tr}\,\Pi_a W_a = \left| l_a \right|^4 + \left| l_a \right|^2 \left| m_a \right|^2 \left| \left\langle l(a)_C \left| o_C \right\rangle \right| \geq \left| l_a \right|^4 = \left(1 - \left| m_a \right|^2\right)^2 \qquad (71\text{-}3)$$
$$\geq 1 - 2\left| m_a \right|^2$$

بنابراین فیدلیتی به صورت $F \geq 1 - 2\sum_a p(a)\left| m_a \right|^2$ می‌باشد. لازم است که $\mathrm{Tr}\,r\,\Gamma > 1-h$ باشد.

$$\mathrm{Tr}\,r\,\Gamma = \sum_a p(a)\,\mathrm{Tr}\,p_a \Gamma = \sum_a p(a)\left| l_a \right|^2 = \sum_a p(a)\left(1 - \left| m_a \right|^2\right) \qquad (72\text{-}3)$$
$$= 1 - \sum_a p(a)\left| l_a \right|^2$$

بنابراین شرط روی مقدار $\mathrm{Tr}\,r\,\Gamma$ مستلزم آن است که $\sum_a p(a)\left| l_a \right|^2 < h$ باشد. یعنی فیدلیتی طرح کد کردن در رابطه $F > 1-h$ صدق می‌کند و نتیجه مطلوب حاصل می‌شود.

خلاصه می‌کنیم که فیدلیتی طرح گذار را به بعد d فضای حالت هیلبرت کانال C ربط داده‌ایم. اگر d به اندازه کافی کوچک باشد به گونه‌ای که آنسامبل سیگنال r وزنی کمتر از h روی هر زیرفضای d-بعدی H_M داشته باشد، پس فیدلیتی باید در رابطه $F < h$ را صدق کند. از طرف دیگر

اگر زیرفضای d-بعدی از H_M پیدا شود که r وزنی بیشتر از $1-h$ داشته باشد، می‌توان گذاری با فیدلیتی $F > 1-2h$ داشت.

3-4-2-4- کد کردن کانال کوانتمی بدون نوفه شوماخر

در نظریهٔ کد کردن بدون نوفه کلاسیکی، سه ویژگی اصلی وجود دارد. نخست یک سیستم ابتدایی کد کردن یکتا مانند بیت انتخاب می‌کنیم. همه پیام‌ها با استفاده از بیت رمزی می‌شوند. دوم مجاز هستیم که پیام‌ها را به‌شکل دنباله‌هایی از N پیام که از منبع‌های مستقل و یکسان گرفته شده‌اند به شکل کد درآوریم (هرچند این شیوه کد کردن منحصر بفرد نمی‌باشد). سوم نیازی نیست که کد کردن کاملاً خالی از اشتباه باشد. تنها کافی است که فیدلیتی کلاسیکی اختیاری به 1 نزدیک باشد. هر یک از این سه مرحله باید برای مفهوم کوانتمی نیز حفظ شوند. تا کنون این کار را برای سومین مورد، معیار فیدلیتی، انجام داده‌ایم. در کدکردن کوانتمی فیدلیتی احتمالی است که یک حالت سیگنال انتقال داده شده از آزمون اعتبارسنجی که آن را با سیگنال اولیه مقایسه می‌کند، به خوبی عبور کند. اصول موضوعه فیدلیتی بخش پیشین، در موقعیت‌های مختلف قیدهایی روی F قرار می‌دهد. بنابراین تنها تعمیم کلی ویژگی‌های اول و دوم باقی می‌ماند. برای سیستم کد کردن ابتدایی کوانتمی همان‌گونه که پیش از این گفته شد از سیستم‌های دو حالتی مانند ذره اسپین $1/2$ استفاده می‌شود. که کیوبیت نامیده می‌شود. اگر ما سیستم تک کیوبیتی را با Q نشان دهیم، کانال کوانتمی ما از تعداد K کیوبیت تشکیل شده است که K معمولاً عددی بزرگ است. بنابراین رابطهٔ زیر برقرار است:

$$H_C = H_Q \otimes ... \otimes H_Q \qquad (73\text{-}3)$$

تعداد فاکتورها برابر K و بُعد H_C، 2^k می‌باشد.

برای بحث کد کردن بلوکی کوانتمی، باید یک منبع تک کوانتمی بسط داده شده را در نظر بگیریم. که سیستمی است که از N کپی مستقل از سیستم M تشکیل شده است. بنابراین M^N سیستم مستقل داریم. هر زیر سیستم M_K در یک حالت سیگنالی $|a_K\rangle$ می‌باشد که هر کدام

سازگار با توزیع احتمال $p(a_K)$ تولید شده‌اند. بنابراین سیستم M^N در حالت $|a\rangle = |a_1,...,a_N\rangle$ می‌باشد. احتمال کل با $p(a) = p(a_1)p(a_2)...p(a_N)$ به دست می‌آید. باید حالت سیگنال را از سیستم توأم M^N به سیستم مشابه M'^N و با استفاده از یک کانال تشکیل شده از کیوبیت‌ها انتقال دهیم.

عملگر چگالی r^N که آنسامبل M^N را توصیف می‌کند به سادگی حاصل‌ضرب مستقیم عملگرهای چگالی برای آنسامبل‌های تکی از زیرسیستم‌های منحصربفرد می‌باشد: $r^N = r_1 \otimes ... \otimes r_N$. یعنی ویژه‌حالت‌های r^N، برابر با حاصل‌ضرب ویژه‌حالت‌های r_i ها، $|n_1,...n_N\rangle$، و ویژه‌مقادیر r^N برابر با حاصل‌ضرب ویژه‌مقادیر r_i ها، $P_{n_1,...,n_N} = P_{n_1}...P_{n_N}$، هستند.

اکنون برای سیستم M ویژه‌مقادیر P_n $(n = 1,...,D)$ عملگر چگالی آنسامبل r، همه ویژگیهای توزیع احتمالی روی اعداد صحیح $D-1$ را دارد. (در حقیقت آن‌ها توزیع احتمال برای خروجی‌های یک اندازه‌گیری کامل با ویژه‌حالت‌های $|n\rangle$ هستند). علاوه بر این آنتروپی فون نیومن برای این توزیع احتمال درست همان آنتروپی شانون می‌باشد.

$$S(r) = -\sum_n P_n \log_2 P_n \qquad (74\text{-}3)$$

یک ویژه‌حالت r^N بر دنباله $n_1, \mathbf{K}, n_N$ منطبق است که Nها اعداد صحیح می‌باشد و ویژه‌مقادیر این ویژه‌حالت‌ها درست برابر با احتمال این است که آیا این دنباله با N استفاده مستقل با توزیع احتمال P_N به دست می‌آید.

تنها به ویژه‌مقادیر و ویژه‌حالت‌های عملگرهای چگالی توجه کرده‌ایم. می‌توان ادعا کرد که هر منبع سیگنال کوانتمی یک منبع پیام کلاسیکی است که از تعداد $D-1$ حروف الفبا استفاده کرده است و توزیع احتمال پیام P_n دارد. $D-1$ عددی صحیح است. آنتروپی شانون این توزیع برابر با آنتروپی فون نیومن می‌باشد. منبع پیام، از این دنباله‌ها گسترده شده به‌وجود آمده است. از بحث بالا

می‌توان فهمید که به‌ازای Nهای به اندازه کافی بزرگ دنباله‌ها به دو دنباله تقسیم می‌شوند: الف)یک سری که حدوداً $2^{N S(r)}$ دنباله بهنجار دارد. ب) یک سری از دنباله‌ها به احتمال کلی کوچکتر. این دو سری دو زیرفضای متعامد از فضای هیلبرت H_{M^N} را مشخص می‌کنند. اینها حدوداً بعدی برابر $2^{N S(r)}$ دارند و می‌توانند با اطمینان خوبی به حالتی از مجموعه $N S(r)$ کیوبیت انتقال داده شوند. زیرفضای دیگر $\Lambda^{\perp}$ می‌باشد، که وزن کوچکی با توجه r^N دارد و بنابراین فیدلیتی آن اثر زیادی ندارد.

اکنون می‌توانیم نظریهٔ بدون نوفه کوانتمی را ثابت کنیم. M را منبع سیگنال کوانتمی در نظر بگیرید که با عملگر چگالی آنسامبل سیگنال r توصیف می‌شود. همچنین $d, e > 0$ را برگزینید. الف) با فرض آنکه برای هر سیگنال M، $S(r)+d$ کیوبیت موجود باشد. پس برای Nهای به اندازه کافی بزرگ، گروه‌های N سیگنالی از منبع سیگنال M می‌توانند با کیوبیت‌های موجود با فیدلیتی $F > 1-e$ انتقال داده شود. ب) با فرض آنکه $S(r)-d$ کیوبیت برای هر M سیگنال موجود باشد. برای Nهای به اندازه کافی بزرگ، اگر گروه N سیگنالی منبع سیگنال M با کیوبیت‌های موجود انتقال داده شوند، پس فیدلیتی $F > e$ می‌باشد.

قسمت الف به روش زیر ثابت می‌شود. ابتدا توجه کنید که اگر کانال کوانتمی C، $N(S(r)+d)$ کیوبیت، Q، باشد پس $\dim H_C = 2^{N(S(r)+d)}$. از تشابه با نظریهٔ کلاسیکی و ویژه-مقادیر احتمالی می‌دانیم که به‌ازای Nهای به اندازه کافی بزرگ، تعداد دنباله‌های ویژه‌حالت‌های بهنجار برابر با $n \le 2^{N(S(r)+d)}$ می‌باشد و جمع ویژه‌مقادیر باقی‌مانده r^N کمتر از $e/2$ است. اگر Γ تصویر روی زیرفضاهای nبعدی Λ باشد که به وسیله ویژه‌حالت‌های بهنجار گسترده شده است، پس $\mathrm{Tr} r^N \Gamma > 1-e/2$ می‌باشد. به کمک اصل موضوعه 2 طرح گذاری با فیدلیتی $F > 1-e$ وجود دارد.

برای قسمت ب توجه کنید که بحث کلاسیکی به ما می‌گوید که به‌ازای Nهای به اندازه کافی بزرگ، که هیچ یک از $2^{N(S(r)-d)}$ دنباله ویژه‌حالت‌ها برای M^N ویژه‌مقادیری با حاصل جمعی به بزرگی اندازه e ندارند. بنابراین برای هر تصویر Γ روی زیرفضای $2^{N(S(r)-d)} = \dim H_C$

رابطهٔ $\mathrm{Tr} r^{N} \Gamma < e$ برقرار است. بنابراین با اصل موضوعه 1 هر طرح گذاری فیدلیتی $F < e$ دارد[22].

3-4-2-5- انتقال حالت‌های در هم تنیده

در بخش پیشین موقعیتی را تصور کردیم که حالت‌های کوانتمی مختلف به طور احتمالی تولید می‌شدند. بنابراین عملگر چگالی r لزوماً یک حالت آمیخته را توصیف می‌کند. حتی هنگامی که سیستم M بخشی از سیستم بزرگ‌تر $M+Z$ باشد باز هم عملگر چگالی به وجود می‌آید. حالت کوانتمی $\left| y_{M+Z} \right\rangle$ یک حالت خالص است ولی درهم‌تنیده می‌باشد. چنین حالتی را همیشه می‌توان با تجزیه قطبی[52] به صورت زیر نوشت:

$$\left| y_{M+Z} \right\rangle = \sum_{n} \sqrt{P_n} \left| n_M , n_Z \right\rangle \qquad (75\text{-}3)$$

که $\left| n_M \right\rangle$ و $\left| n_Z \right\rangle$ به ترتیب سری‌ها حالت‌های متعامد M و Z هستند. برای نوشتن حالت M باید روی Z تریس جزئی بگیریم که در این صورت عملگر زیر حاصل می‌شود:

$$r = \sum_{n} P_n \left| n_M \right\rangle \left\langle n_M \right| \qquad (76\text{-}3)$$

که ویژه‌حالت‌های این عملگر $\left| n_M \right\rangle$ و ویژه‌مقادیرهای آن P_n نشان داده می‌شود. آنتروپی فون نیومن $S(r)$ این عملگر چگالی گاهی در نظریهٔ اطلاعات با عنوان درجه درهم‌تنیدگی بین سیستم‌های کوانتمی M و Z شناخته می‌شود.

اکنون گذار تقریبی از M به M' را در نظر می‌گیریم. آیا این گذار حالت کوانتمی $M+Z$ را به طور مناسب به $M'+Z$ انتقال می‌دهد؟ و یا به عبارتی دیگر آیا طرح گذار معرفی شده درهم‌تنیگی سیستم با بقیه محیط را از M به M' انتقال می‌دهد؟ پاسخ این پرسش مثبت می‌باشد. حالت $\left| y_{M+Z} \right\rangle$ که با عملگر تصویر $p = \left| y_{M+Z} \right\rangle \left\langle y_{M+Z} \right|$ نمایش داده می‌شود به حالت نهایی w، از $M'+Z$ می‌رود. فیدلیتی این گذار با $F = \mathrm{Tr} p\, w$ بیان می‌گردد. در طرح گذار خود، ویژه‌حالت-

[52] Polar decomposition

های r را که به طور محتمل در زیر فضای Λ از H_M با $\dim\Lambda = \dim H_C$ بودند، انتقال دادیم. همانند قبل می‌توان تعداد ویژه‌حالت‌های n_M را شمرد چنانچه نخست d تا از آن‌ها در Λ قرار می‌گیرد. اکنون حالت کلی را می‌توان به صورت زیر نوشت:

$$\left| y_{M+Z} \right\rangle = l \left| l_{M+Z} \right\rangle + m \left| m_{M+Z} \right\rangle \qquad (77\text{-}3)$$

که $\left| l_{M+Z} \right\rangle$ و $\left| m_{M+Z} \right\rangle$ حالت‌های متعامد بهنجار می‌باشند.

$$\left| l_{M+Z} \right\rangle = \sum_{n=1}^{d} c_n \left| n_M, n_Z \right\rangle$$
$$\left| m_{M+Z} \right\rangle = \sum_{n=d+1}^{D} c'_n \left| n_M, n_Z \right\rangle \qquad (78\text{-}3)$$

که $|l|^2 + |m|^2 = 1$ است. می‌توان این حالت را با تصویرگر $p = \left| y_{M+Z} \right\rangle\left\langle y_{M+Z} \right|$ نمایش داد. در طرح گذار حالت $\left| l_{M+Z} \right\rangle$ کاملاً انتقال داده می‌شود و حالت $\left| m_{M+Z} \right\rangle$ به بعضی حالت‌های آمیخته منتقل شده تا حالت آمیخته w از $M'+Z$ حاصل شود. این حالت آمیخته :

$$y = |l|^2 \left| l_{M'+Z} \right\rangle\left\langle l_{M'+Z} \right| + |m|^2 \left[\sum_{n=d+1}^{D} \left| c'_n \right|^2 \left| o_{M'}, n_Z \right\rangle\left\langle o_{M'}, n_Z \right| \right] \qquad (79\text{-}3)$$

از عبارت بالا می‌توان نتیجه گرفت که $F = \mathrm{Tr} p\, w \geq |l|^4$ است. اگر $|l|^2 > 1-h$ باشد پس $F > 1-2h$ که همانند نتیجه‌ای است که در اصل موضوعه 2 حاصل شد.

یک بار که این نتیجه حاصل شد می‌توان بحث نظریهٔ کد کردن را برای N کپی از سیستم‌های در هم تنیده $M+Z$ تکرار کرد و نتیجه گرفت درهم‌تنیدگی بین M و Z را می‌توان با استفاده از کانال C با $S(r)$ کیوبیت به طور مطمئن از M به M' انتقال یابد. به دیگر سخن می‌توان درهم‌تنیدگی تعداد زیادی سیستم را با اطمینان منتقل کرد، البته در صورتی که برای هر سیستم $S(r)$ داشته باشیم. بنابراین می‌توان $S(r)$ را به عنوان اندازه منبع‌های فیزیکی لازم برای حرکت درهم‌تنیدگی کوانتمی بین M و بقیه جهان (سیستم Z) از یک سیستم به دیگری تفسیر کرد.

3-7- اطلاعات کلاسیکی در کانال‌های کوانتمی

3-7-1- اطلاعات کلاسیکی در کانال‌های بدون نوفه کوانتمی

در این بخش در مورد حداکثر ظرفیت یک کانال کوانتمی برای اطلاعات کلاسیکی صحبت می-کنیم. فرض کنید آلیس می‌خواهد اطلاعات کلاسیکی را با استفاده از کانال ارتباطی Q منتقل کند. بنابراین آلیس کانال را در یکی از حالت‌های کوانتمی W_x و با احتمال p_x آماده کرده و باب سعی می‌کند با اندازه‌گیری روی Q، حالتی را که آلیس فرستاده تشخیص دهد (برای مثال کانال می‌تواند با حالت‌های کوانتمی قطبش فوتون‌های ویژه مشخص شود که گاهی این حالت‌های کوانتمی را حروف الفبا می‌نامند). نظریه‌ای را گوردن[53] و لویتاین[54] مطرح کردند و هالوو[55] آن را اثبات کرد، که در آن حد بالایی روی حداکثر اطلاعاتی که باب می‌تواند کسب کند قرار می‌دهند[28]. اگر $W = \sum_x p_x W_x$ ماتریس چگالی توصیف کننده سیگنال آلیس می‌باشد، حد بالای گفته شده روی حداکثر اطلاعاتی که باب می‌تواند به آن دست یابد، صورت زیر می‌باشد:

$$H(X:Y) \le H(W) - \sum_x p_x H(W_x) \qquad (80\text{-}3)$$

در رابطه بالا $H(X:Y)$ اطلاعات دو جانبه یا مشترک بین ورودی به کانال یعنی X و خروجی کانال یعنی Y و $H(W)$ آنتروپی فون نیومن برای عملگر چگالی W می‌باشد[29]. در حاصل شدن این نتیجه از کارهای مفیدی که در[30، 31، 32، 33، 34 و 35] آمده استفاده مستقیم شده است. اگر حالت‌های سیگنال آلیس حالت‌های کاملاً خالص باشند جمله دوم در رابطه (80-3) برابر صفر شده و اطلاعات قابل دستیابی برای باب ممکن است به $H(W)$ بسیار نزدیک شود. در حالتی که W_x حالت‌های آمیخته باشند باز هم راهی برای باب وجود دارد که اطلاعات دریافتی‌اش به حد هالوو نزدیک شود. برای دستیابی به این هدف آلیس باید اطلاعات خود را به گونه‌های مناسب کد کند و باب مشاهده‌پذیر کدگشاینده مناسبی برای اعمال روی کانال برگزیند. مثلاً آلیس می‌تواند با گزینش

[53] Gordon

[54] Levitin

[55] Kolevo

کدواژه‌هایی به طول L (که به اندازه کافی بزرگ می‌باشد) کدهایی به دست آورد که اگر مشاهده-پذیرهای کد گشاینده به صورتی مناسب برگزیده شوند بتوان با احتمال خطای $P_E \le 9e + N\ 2^{-L(c-4d)}$ اطلاعات قابل دستیابی برای هر حرف را به $c-d$ رساند که $e,d>0$ و $c = H(W) - \sum_x p_x H(W_x)$ فرض شده که همان قید هالوو بوده و N تعداد کدواژه‌ها می‌باشد. بنابراین می‌توان گفت که ظرفیت کانال کوانتمی برای اطلاعات کلاسیکی حداکثر می‌تواند برابر با c باشد.

3-7-2- اطلاعات کلاسیکی در کانال‌های نوفه‌ای کوانتمی

در بخش پیشین تنها دربارهٔ انتقال اطلاعات کلاسیکی در کانال‌های بدون نوفه کوانتمی صحبت کردیم اما همان‌گونه که می‌دانیم اطلاعات ممکن است در کانال دچار نوفه شده و مقداری از آن از بین برود. مثلاً اگر ورودی کانال حالتی خالص باشد ممکن است در اثر عبور از کانال به حالتی آمیخته تبدیل شود. با این توضیح ممکن است حالت‌های آمیخته‌ای که در بخش پیشین ذکر شد گاهی خروجی کانال نوفه‌ای محسوب شوند. بنابراین با فرض آنکه اطلاعات رسیده به باب در حالت‌های آمیخته می‌باشند آغاز می‌کنیم. در حین انتقال اطلاعات از آلیس به باب، کانال Q ممکن است تحول یکانی درونی را تحمل کند (باب می‌تواند با گزینش مشاهده‌پذیرهای چرخیده مناسب کدگشایی صحیحی انجام دهد). در حالتی دیگر ممکن است کانال با محیط برهمکنش داشته باشد (باب نمی-تواند روی خروجی تصحیح انجام دهد).

کلی‌ترین توصیف تحول سیستم کوانتمی Q که با یک محیط برهمکنش می‌کند به کمک نگاشت خطی مثبت روی سری عملگرهای چگالی Q حاصل می‌شود، مشروط بر اینکه این نگاشت رد[56] عملگرها را حفظ کند. چنین نگاشتی با ابرعملگر[57] e توصیف می‌شود:

[56] Trace

[57] Superoperator

$$r \to r' = e(r) \qquad (81\text{-}3)$$

که r حالت اولیه سیستم و r' حالت نهایی می‌باشد. ابرعملگر e به طور خطی روی سیتم عمل کرده و باعث تحول یکانی سیستم می‌شود. بنابر آنچه گفته شد کانال کوانتمی نوفه‌ای با یک ابرعملگر e مشخص می‌شود که تحول هر حرف الفبایی را که از آلیس به باب انتقال داده می‌شود، توصیف می‌کند. همچنین فرض می‌کنیم که تحول هر حرف مستقل از دیگری است. یعنی اگر ورودی کانال به شکل حاصلضرب حالت‌های حروف در نظر گرفته شود، خروجی کانال نیز حاصلضرب حالت‌ها می‌باشد. مسأله اساسی آلیس این است که از حالت‌های ورودی w_x چنان استفاده کند که حالت‌های خروجی $W_x = e(w_x)$ برای باب تشخیص‌پذیر باشند. اگر حالت‌های ورودی آلیس حروف الفبای ثابت $\{w_x\}$ باشد پس به ازای هر حرف بیشینه اطلاعاتی که می‌تواند برای باب قابل دسترس باشد با C_Γ داده می‌شود، که Γ حروف الفبای حالت‌های خروجی می‌باشد. در [29] نشان داده شده است که بهینه‌ترین ورودی برای کانال‌های نوفه‌ای استفاده از حالت‌های خالص می‌باشد و در این صورت باب می‌تواند بیشینه اطلاعات را در خروجی کسب کند.

به خلاصه گفته می‌شود که اگر آلیس از حالت‌های حاصلضرب برای ورودی کانال استفاده کند، آنگاه ظرفیت کانال کوانتمی نوفه‌ای به صورت زیر می‌باشد:

$$C^{(1)} = \max c \qquad (82\text{-}3)$$

که $c = H(W) - \sum_x p_x H\left(e\left(w_x\right)\right)$ می‌باشد. در رابطهٔ (82-3) ماکزیمم روی حالت‌های ورودی خالص در نظر گرفته می‌شود. پس آلیس به طور منطقی می‌تواند اطلاعات را با آهنگی که از $C^{(1)}$ کمتر است به باب منتقل کند. گاهی $C^{(1)}$ را ظرفیت حالت‌های حاصل‌ضرب در نظر می‌گیرند.

3-6- نتیجه‌گیری

پس از بررسی‌های بسیار به این نتیجه می‌رسیم که آنتروپی فون نیومن با وجود اینکه پیش از آنتروپی شانون معرفی شده و هدف فون نیومن تنها معرفی آنتروپی کوانتمی بود که نیاز مکانیک کوانتمی

آماری برآورده شود، پس از سال‌ها ثابت شد که این رابطه در نظریهٔ اطلاعات کوانتمی بسیار مفید بوده و نیازهای زیادی را برآورده می‌کند. با توجه به کاربرد فوق‌العاده‌ای که آنتروپی فون نیومن در تعیین کانال‌های کوانتمی پیدا کرده است گروهی بر این باوردند که علی‌رغم ایرادهایی که به آنتروپی فون نیومن گرفته می‌شود، تاکنون هیچ آنتروپی دیگری نتوانسته است جای آنتروپی فون نیومن را در مکانیک کوانتمی بگیرد.

فصل چهارم

معرفی دو آنتروپی برای کمّی کردن اطلاعات کوانتمی و آنتروپی اطلاعات سیستم‌های تریتی و کیوتریتی

- اندازه اطلاعات پایه شده روی پیشگویی احتمالی آزمایشگر
- آنتروپی مرکب درحکم تصحیحی برای آنتروپی فون نیومن
- مقایسه اهمیت دو آنتروپی معرفی شده با آنتروپی فون نیومن و کارایی آن‌ها
- سیستم‌های سه‌تایی کلاسیکی
- سیستم‌های کوانتمی سه حالتی
- نتیجه گیری

مقدمه

تاکنون تنها در مورد آنتروپی فون نیومن درحکم آنتروپی که محتوای اطلاعاتی یک سیستم کوانتمی را مشخص می‌کند، صحبت کردیم. در این فصل قصد داریم به اندازه‌های دیگری از اطلاعات بپردازیم که برای تعیین محتوای اطلاعاتی سیستم‌های کوانتمی مفید می‌باشند و در بعضی کاربردها مفیدتر از آنتروپی فون نیومن به نظر می‌رسند. همچنین در انتهای این فصل در مورد سیستم‌های سه حالتی کلاسیکی و کوانتمی و رفتاری که آنتروپی اطلاعات هنگام استفاده از این سیستم‌ها از خود نشان می‌دهد، صحبت خواهیم کرد.

4-1 اندازه اطلاعات پایه شده روی پیشگویی احتمالی آزمایشگر

مکانیک کوانتمی ذاتاً توصیف احتمالی طبیعت است و پیش از انجام دادن هر آزمایش کوانتمی احتمال رخ دادن همه خروجی‌های ممکن برای هر آزمایشگر مشخص است. ولی به طور کلی رخ دادن هر خروجی کاملاً تصادفی می‌باشد. در اینجا اندازه‌ای از اطلاعات برای یک اندازه‌گیری منحصربفرد آورده می‌شود که براساس این حقیقت بنا شده که پیشگویی احتمالی که آزمایشگر می‌تواند بسازد برای هیچ آزمایش منحصربفردی اهمیت عملی ندارد و تنها پیشگویی دربارهٔ تعداد رخ دادن یک خروجی مشخص در تکرارهای آتی آزمایش اهمیت دارد.

4-1-1- عدم یقین در تعداد تکرارهای یک رویداد

آزمایشی با دو خروجی ممکن را در نظر بگیرید که احتمال رخ دادن یک خروجی را با p و احتمال رخداد خروجی دیگر را با $1-p$ نشان می‌دهیم. با دانستن احتمال‌ها برای دو خروجی یک آزمایشگر می‌تواند پیش‌بینی کند که چند بار یک خروجی اتفاق می‌افتد. تنها تعداد محدودی سیستم وجود دارند که با پیش‌بینی هماهنگی دارند. به سبب افت و خیزهای آماری سهیم با هر تعداد محدودی آزمایش تجربی، تعداد رخ دادن یک خروجی مشخص در N تکرار بعدی آزمایش به طور کلی قابل پیش‌بینی نیست. در N آزمایش تجربی مستقل، دنباله منظم ویژه از نتایج «بله» و «خیر»، که دقیقاً شامل رخ دادن n بار «بله» و $N-n$ بار «خیر» باشد احتمالی به صورت زیر دارد:

$$p.(1-p)(1-p)...p = p^n(1-p)^{N-n} \quad (1\text{-}4)$$

جایگشت‌های متفاوت دنباله‌های مختلف حادثه‌هایی مستقل‌اند بنابراین می‌توان احتمال‌های آن‌ها را با هم جمع کرد و نتیجهٔ زیر را به‌دست آورد:

$$P_N(n) = \binom{N}{n} p^n(1-p)^{N-n} \quad (2\text{-}4)$$

کمیت بالا احتمال این است که از N آزمایش تجربی مستقل n بار «بله» و $N-n$ بار «خیر»مشاهده کنیم. رابطهٔ بالا با نام توزیع دوجمله‌ای شناخته می‌شود[13]. اگر روی نتیجه‌ای مشخص شرط‌بندی کنیم، برای مثال روی تعداد خروجی‌ها که با بالاترین احتمال «بله» باشد، $n^{\max} = pN$ است، هنوز هم احتمال موفقیت به p بستگی دارد. با $p = 0{,}5$ احتمال اینکه در 10 بار آزمایش 5 بار خروجی «بله» داشته باشیم تنها برابر 0,25است.

عدم یقین آزمایش یا کمبود اطلاعات در مقدار n به وسیله انحراف مجذور میانگین که به شکل مقدار انتظاری مجذور انحراف n از مقدار میانگین pN تعریف شده ، به‌دست می‌آید.

$$s^2 = \sum_{n=1}^{N} P_N(n)(n-pN)^2 = p(1-p)N \quad (3\text{-}4)$$

در حقیقت اگر s کوچک باشد پس هر جمله‌ای در معادله (3-4) کوچک است. برای اینکه $(n-pN)$ بزرگ باشد مقدار n دارای احتمال کوچک $P_N(n)$ است. به عبارتی دیگر در موردی که s کوچک است انحراف تعداد رخداد خروجی «بله» از مقدار میانگین pN غیرممکن است. در این مورد آزمایشگر ترکیب تعداد رخ دادن خروجی‌ها را با یقین بالایی می‌تواند حدس بزند. برعکس انحراف بزرگ نشان می‌دهد که با احتمال بالایی مقادیر n نزدیک مقدار متوسط pN قرار نمی‌گیرد. اکنون آزمایشی با سه خروجی ممکن a_1، a_2 و a_3 را در نظر بگیرید که احتمال رخ دادن آن‌ها بترتیب p_1، p_2 و p_3 باشد. روشی وجود دارد که به کمک آن مشاهده‌گر می‌تواند این گروه سه‌تایی را به یک گروه باینری تبدیل کند و سپس از اندازه اطلاعات (3-4) استفاده کند. برای مثال او می‌تواند

دو خروجی a_1 و a_2 را به صورت یک خروجی تک $a_1 \vee a_2$ با احتمال $p_1 + p_2$ و خروجی a_3 را به صورتی یک خروجی منحصربفرد که با احتمال $p_3 = 1 - p_1 - p_2$ رخ می‌دهد، فرض کند. اکنون ممکن است مشاهده‌گر ابتدا در مورد تعداد رخ دادن‌های خروجی $a_1 \vee a_2$ در N آزمایش تجربی بپرسد، و بعد از اینکه $a_1 \vee a_2$ رخ داد در مورد تعداد دفعه‌هایی که a_1 یا a_2 رخ می‌دهد تحقیق کند.

در بخش نخست این روش اندازه کمبود اطلاعات آزمایشگر درباره تعداد رخدادن خروجی $a_1 \vee a_2$ در N آزمایش تجربی به صورت زیر داده می‌شود.

$$s^2(a_1 \vee a_2, a_3) = (p_1 + p_2) p_3 N \qquad (4\text{-}4)$$

در بخش دوم روش با دانستن اینکه $a_1 \vee a_2$ رخ داده است، اندازه کمبود اطلاعات درباره رخدادن a_1 یا a_2 را جستجو کنیم. بنابراین اندازه کمبود اطلاعات در این مورد از رابطه زیر به‌دست می‌آید

$$s^2(a_1, a_2 \mid a_1 \vee a_2) = \frac{p_1}{p_1 + p_2} \frac{p_2}{p_1 + p_2} N (p_1 + p_2) \qquad (5\text{-}4)$$

توجه کنید هرگاه $a_1 \vee a_2$ اتفاق افتاده باشد احتمال‌های شرطی برای a_1 یا a_2 به‌ترتیب به صورت $\frac{p_1}{p_1 + p_2}$ و $\frac{p_2}{p_1 + p_2}$ می‌باشد. همچنین تعداد آزمایش‌های تجربی که آزمایشگر در بخش دوم مجبور است انجام دهد برابر $N(p_1 + p_2)$ می‌باشد. بخش دوم روش شرطی بوده و تنها برای کسر $p_1 + p_2$ از زمان انتظار می‌رود که رخ دهد. بنابراین به طور کلی اندازه کمبود اطلاعات انتظاری $U_N(a_1, a_2, a_3)$ با توجه به تعداد دفعه‌های رخدادن سه خروجی در N بار تکرار آزمایش به صورت زیر محاسبه می‌شود:

$$\begin{aligned} U_N(a_1, a_2, a_3) &= s^2(a_1 \vee a_2, a_3) + (p_1 + p_2) s^2(a_1, a_2 \mid a_1 \vee a_2) \qquad (6\text{-}4) \\ &= (p_1 p_2 + p_1 p_3 + p_2 p_3) N \end{aligned}$$

گفته‌های بالا را می‌توان به‌راحتی برای n خروجی ممکن $a_1, \mathbf{K}, a_n$ که احتمال‌های رخدادن آن‌ها $\vec{p} = (p_1, \mathbf{K}, p_n)$ می‌باشد تعمیم داد.

$$U_N(a_1,\mathbf{K},a_n)=\sum_{i<j}^{n} p_i p_j\, N \qquad (7\text{-}4)$$

می‌بینیم که کمبود اطلاعات آزمایشگر متناسب با تعداد آزمایش‌های تجربی است که انجام می‌دهد. این خصوصیت بیان می‌کند که با انجام دادن هر آزمایش منحصربفرد، مقدار اطلاعات مشابهی کسب می‌شود و تعداد دفعه‌هایی که تاکنون آزمایش انجام شده اهمیتی ندارد و بنابراین پس از هر آزمایش تجربی کمبود اطلاعات به مقدار مشابهی کاهش می‌یابد.

$$U(\mathbf{r}\vec{p})=\frac{U_N(a_1,\mathbf{K},a_n)}{N}=\sum_{i<j}^{n} p_i p_j=\frac{1}{2}\left(1-\sum_{i=1}^{n} p_i^2\right) \qquad (8\text{-}4)$$

این معادله کمبود اطلاعات مربوط به یک تک آزمایش تجربی آتی را نشان می‌دهد.

بنا به تعریف $U(\vec{p})$ دانش یا اطلاعاتی که یک آزمایشگر پیش از انجام دادن آزمایش ارائه می‌دهد تا حدودی مکمل $U(\vec{p})$ است و علاوه بر این دانش یا اطلاعات گفته شده تابعی از مجذور احتمال‌ها می‌باشند.

$$I(\mathbf{r}\vec{p})=\sum_{i=1}^{n} p_i^2 \qquad (9\text{-}4)$$

عبارت‌های کلی از این نوع را هاردی[58] و همکارانش با جزئیات کامل مورد مطالعه قرار دادند و یوفینک[59] برای اینکه نشان دهد که $I(\vec{p})$ یک اندازه اطلاعات منطقی است چندین ویژگی آن را اثبات کرد. برای مثال،

1. $I(\vec{p})$ تحت برچسب‌گذاری جدید سری خروجی‌های ممکن ناوردا باقی می‌ماند.
2. $I(\vec{p})$ تابع پیوسته‌ای از p_j هاست و تنها موقعی که همه p_jها به‌جز یکی صفر هستند، دارای بیشینه‌ای برابر 1 می‌باشد. هنگامی که احتمالها به طور یکنواخت بین همه خروجی-های ممکن توزیع شده باشد، یعنی زمانی که به‌ازای همه jها $p_j=1/n$ باشد، $I(\vec{p})$ به مقدار کمینه خود که برابر $1/n$ می‌رسد.

[58] Hardy

[59] Uffink

3. اگر دو آنسامبل که با توزیع احتمال‌های $\vec{p}=(p_1,\mathbf{K},p_n)$ و $\vec{q}=(q_1,\mathbf{K},q_n)$ مشخص شده‌اند، به یکدیگر ملحق شوند (به زبان ریاضی این مطلب با ترکیب مقعر $r_j=a\,p_j+(1-a)q_j$ و $0\le a\le 1$توصیف می‌شود.) اطلاعات مربوط به آنسامبلی که نمونه بخصوصی از آن می‌آید، از دست می‌رود، و در نتیجه اندازهٔ اطلاعات کاهش می‌یابد.

4. هرگاه $\vec{p}=(p_1,\mathbf{K},p_n)$ و $\vec{q}=(q_1,\mathbf{K},q_n)$ دو توزیع احتمال بوده و $\vec{r}=(r_{11},\mathbf{K},r_{nm})$ حاصلضرب مستقل‌شان باشد، یعنی $r_{ij}=p_iq_j$، آنگاه:

$$I(\vec{r})=I(\vec{p})I(\vec{q}) \qquad (10\text{-}4)$$

در حقیقت هاردی و همکارانش طبقه‌ای کلی از عبارت‌های ریاضی معرفی کردند که از نقطه نظر نظریهٔ اطلاعات می‌توانند برای تعیین اطلاعات بکار روند.

$$M_a(\vec{p})=\left(\sum_{i=1}^{n}p_i^a\right)^{a-1} \qquad \text{و} \quad 0\le a\le 1 \qquad (11\text{-}4)$$

بنابراین می‌توان $I(\vec{p})$ را که به‌تازگی معرفی کردیم، موردی خاص از رابطه (11-4) به‌ازای $a=2$ دانست.

این عبارت‌ها را می‌توان به آنتروپی رینی[60] نیز ربط داد.

$$H_a(\vec{p})=-\log M_a(\vec{p})=-\frac{1}{1-a}\log\sum_{i=1}^{n}p_i^a \qquad (12\text{-}4)$$

همچنین توجه شود که می‌توان آنتروپی اطلاعات شانون:

$$H(\vec{p})=\lim_{a\to 0}H_a(\vec{p}) \qquad (13\text{-}4)$$

و همچنین منفی لگاریتم اندازه اطلاعاتی را که به تازگی معرفی شد:

$$-\log I(\vec{p})=H_2(\vec{p}) \qquad (14\text{-}4)$$

[60] Rányi

موردی ویژه از آنتروپی رینی دانست. همچنین اگر به جای $I(\vec{p})$ عبارت $-\log I(\vec{p})$ را به عنوان اندازه عدم یقین مناسب در یک اندازه‌گیری منحصربفرد در نظر بگیریم، پس عدم یقین درباره خروجی مشترکی که از حاصلضرب اعضای دو توزیع احتمال به دست می‌آید، $r_{ij}=p_i q_j$، با جمع عدم یقین‌های دو دو توزیع احتمال مستقل برابر است.

$$-\log I(\vec{r})=-\log I(\vec{p})-\log I(\vec{q}) \qquad (15\text{-}4)$$

توجه شود که می‌توان عبارت (9-4) را توصیف کننده طول بردار احتمال $\vec{p}$ نیز محسوب کرد. آشکارا به خاطر $\sum_i \mathbf{r}\, p_i =1$ همه بردارها در فضای برداری مجاز نمی‌باشند. در حقیقت طول کمینه $\vec{p}$ زمانی است که همه p_i مساوی $(p_i=1/n)$ باشند. این حالت بر موقعیت فقدان اطلاعات کامل در یک آزمایش سازگار است.

4-1-2- اندازه اطلاعات بروکنر[61] و ژلینگر[62]

بروکنر و ژلینگر توانستند با استفاده از نتایج کارهای هاردی، یوفینک، رینی و حتی استفاده از آنتروپی شانون، اندازه جدیدی از اطلاعات ارائه دهند که بسیار مفید می‌باشد. اندازه اطلاعات جدید معرفی شده، در شکل بهنجار[63] شده برای یک اندازه‌گیری کوانتمی منحصربفرد به صورت زیر می‌باشد:

$$I(\vec{p}) = N \sum_{i=1}^{n} \left(p_i - \frac{1}{n} \right)^2 \qquad (16\text{-}4)$$

با فرض آن‌که به طور بیشینه k بیت اطلاعات می‌توانند به صورت کد درآیند، یعنی $n = 2^k$، ثابت بهنجارش $N = 2^k k / (2^k - 1)$ می‌باشد[15].

$$I(\vec{p}) = \frac{2^k k}{2^k - 1} \sum_{i=1}^{2^k} \left(p_i - \frac{1}{2^k} \right)^2 \qquad (17\text{-}4)$$

که این رابطه به ازای $k = 1$ و $k = 2$ به شکل $I(\vec{p}) = 2\sum_{i=1}^{2} p_i^2 - 1 = (p_1 - p_2)^2$ و $I(\vec{p}) = 8/3 \sum_{i=1}^{4} p_i^2 - 2/3$ درمی‌آید البته این مقادیر برای سیستم‌های کلاسیکی است.

4-1-3- ویژگی ناوردایی اطلاعات تحت آنتروپی بروکنر و ژلینگر

بروکنر و ژلینگر بعضی از ویژگیهای جالب آنتروپی جدید خود را در کاری مشترک ارائه دادند که در زیر به برخی از آن‌ها اشاره می‌شود.

در اندازه‌گیری کوانتمی منحصربفرد تنها ویژگی تعریف شده پیش از انجام دادن آزمایش، احتمال-های مشخص برای هر خروجی ممکن منحصر بفرد می‌باشد. همان‌گونه که در فصل سوم اشاره کردیم طبیعی به نظر می‌رسد که محتوای اطلاعات کل یک سیستم کوانتمی از جمع مقادیر منحصر بفرد اطلاعات روی یک سری کامل از m مشاهده‌پذیر مکمل[64](مشاهده‌پذیرهایی که با یکدیگر جابجا

[61] Brukner

[62] Zeilinger

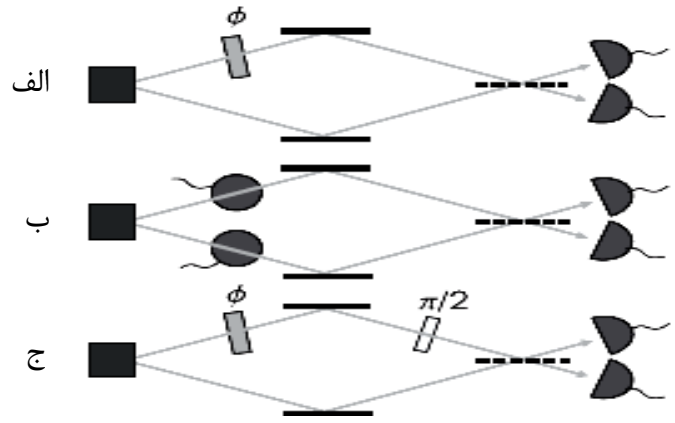
[63] Normalize

[64] Complementary

نشوند) به‌دست آید[17]. برای مثال ممکن است محتوای اطلاعاتی یک سیستم اسپین $1/2$ از جمع اطلاعات به دست آمده از سه مختصه متعامد اسپینی (s_x, s_y, s_z) به‌دست آید. زیرا این سه مشاهده‌پذیر جابجاناپذیر و یا به عبارتی مکملند [12 و 15]. در فصل سوم نشان دادیم که اگر محتوای اطلاعاتی یک سیستم اسپین $1/2$ با استفاده از آنتروپی شانون به دست آید مقدار به دست آمده تحت تبدیلات یکانی ناوردا نمی‌باشد. اما به طور شهودی انتظار داریم که محتوای اطلاعات حمل شده به وسیله یک سیستم خواه کلاسیکی باشد و خواه کوانتمی ناوردا باشد. یک ویژگی مهم اندازه اطلاعات بروکنر و ژلینگر این است که اگر محتوای اطلاعات کل سیستم کوانتمی با استفاده از آن محاسبه شود تحت تغییر سری مشاهده‌پذیرهای جابجاناپذیر ناوردا باقی می‌ماند. ما در اینجا این ناوردایی را نشان می‌دهیم. همچنین اگر جابجایی اطلاعات با محیط صورت نگیرد اطلاعات به دست آمده با گذشت زمان پایسته است.

هنگام انجام دادن آزمایش‌های کوانتمی با قضیه‌ها یا موقعیت‌های مختلفی مواجه می‌شویم که یک سری از این موقعیت‌ها، مکملیت دوجانبه بین گروهی از مشاهده‌پذیرهاست که روی سیستم اثر می‌کنند. ویژگی بارز این گروه این است که دانش کامل در مورد یکی باعث عدم یقین بیشینه در مورد دیگری می‌شود. برای مثال موقعیت مکملی برای یک ذره اسپین $1/2$ می‌تواند به صورت زیر باشد. الف) اسپین در جهت مثبت محور x باشد. ب) اسپین در جهت مثبت محور y باشد. ج) اسپین در جهت مثبت محور z باشد.

آزمایش دیگری که موقعیت‌های مکمل یک سیستم دو حالته کوانتمی را نشان می‌دهد، تداخل کوانتمی می‌باشد. یک تداخل‌سنج ماخ-زندر[65] ایده‌آل را در نظر بگیرید شکل (4-1). سه موقعیت مکمل که ممکن است در هنگام آزمایش رخ دهد در این شکل نشان داده شده است. فرض کنید که برای انتقال فاز f بین دو باریکه درونی تداخل سنج شکل (4-1الف) ذره با قطعیت توسط آشکارساز بالایی (پایینی) پشت باریکه شکافنده آشکار می‌شود. در این مورد با وجود اینکه ما هیچ اطلاعاتی از



شکل (4-1)، آرایه‌ای برای مشخص کردن موقعیت‌های مکمل دوجانبه در یک تداخل سنج ماخ-زندر.
در اینجا فرض می‌شود که آشکار سازها ذره‌ها را بدون جذب آشکار می‌کنند

مسیر ذره در داخل تداخل‌سنج نداریم ولی دانش کاملی در مورد ذرات باریکه که در پشت باریکه شکافنده دیده می‌شوند در دست داریم. بنابراین حالت ذره به وسیله مقدار صحیحی (صحیح یا غلط) از موقعیت: (1) «ذره در حضور انتقال فاز f مسیر به سمت آشکار ساز بالایی را دنبال می‌کند.» نمایش داده می‌شود.

در مقابل اگر مسیری را که باریکه در تداخل‌سنج طی می‌کند، بدانیم (شکل 4-1ب)، هیچ تداخلی اتفاق نمی‌افتد و از این رو در مورد مسیری که ذره در پشت شکافنده باریکه طی می‌کند کاملاً نامطمئن هستیم. بنابراین حالت ذره می‌تواند به وسیله مقدار صحیحی از موقعیت: (2)«ذره مسیر بالایی را درون تداخل‌سنج در پیش می‌گیرد.» نمایش داده می‌شود.

[65] mach-zehnder

با دانستن اینکه اسپین $1/2$ مدلی از مکانیک کوانتمی برای همه سیستم‌های دو حالتی یعنی کیوبیت‌ها آماده می‌کند، پس همیشه انتظار داریم که هر گاه تناوبی باینری داشته باشیم سه موقعیت مکمل دوجانبه برای آن وجود داشته باشد. به راحتی می‌توان نشان داد که حتی بدون داشتن اطلاعات مسیر، در صورتی که انتقال فاز $p/2$ بین دو باریکه درون تداخل‌سنج داشته باشیم تمام اطلاعات شکل (4-1الف) کاملاً از بین می‌رود. پس در آرایه‌ای جدید در شکل (4-1) هر دو باریکه خروجی کاملاً متساوی‌الاحتمال‌اند. اکنون فرض کنید که در حضور انتقال فاز $f+p/2$ ذره با قطعیت به سمت آشکارساز بالایی (پایینی)خارج خواهد شد. بنابراین باز هم حالت سیستم با حالت صحیحی از موقعیت: (3) «ذره در حضور انتقال فازی $f+p/2$ با قطعیت مسیر خروجی را به سمت آشکارساز بالایی می‌پیماید.» نمایش داده می‌شود. برای ذره‌ای در این حالت آگاهی کامل در مورد مسیر خروجی ذره داریم، بدون اینکه دانشی از مسیر ذره در داخل تداخل‌سنج داشته باشیم و یا بدون اینکه آرایه شکل (4-1الف) را داشته باشیم. می‌توانیم سری‌های مختلف سه موقعیت مکمل دوجانبه را با مقادیر f انتقال فاز برچسب بزنیم. این سه موقعیت مکمل برای تداخل‌سنج در واقع معادل موقعیت‌های مکمل اسپین $1/2$ می‌باشند. (1) اسپین در امتداد f در صفحه $x-y$ به سمت بالاست. (2) اسپین در طول محور z به سمت بالاست. (3) اسپین $f+p/2$ در صفحه $x-y$ به سمت بالاست. در اینجا جهت f در صفحه $x-y$ در امتداد زاویه‌ای که با محور x ساخته می‌شود قرار می‌گیرد. بنابراین از حالا می‌توانیم در مورد اندازه‌گیری‌های اسپین بحث کنیم و تنها کاربرد این ایده را برای تداخل‌سنج در ذهن داشته باشیم. اکنون اگر مسیر بالایی در داخل تداخل سنج با $|UP\rangle$ و مسیر پایینی با $|LP\rangle$ نمایش داده شوند، تحول کت‌ها به صورت زیر داده می‌شود:

$$|UP\rangle \to e^{if}|UP\rangle \to \frac{1}{\sqrt{2}}e^{if}\left(|D_2\rangle + i|D_1\rangle\right)$$
$$|LP\rangle \to \frac{1}{\sqrt{2}}\left(|D_1\rangle + i|D_2\rangle\right) \qquad (18\text{-}4)$$

که $|D_1\rangle$ ذره‌ای را که به سمت آشکارساز بالایی می‌رود و $|D_2\rangle$ ذره‌ای که به سمت آشکارساز پایینی نشان می‌دهد. اکنون اگر سیستمی را که در حالت $\hat{r}$ قرار دارد در پایه‌های $|UP\rangle$ و $|LP\rangle$ نمایش دهیم:

$$r_{11}=\langle UP|r|UP\rangle,\ r_{12}=\langle UP|r|LP\rangle$$
$$r_{21}=\langle LP|r|UP\rangle,\ r_{22}=\langle LP|r|LP\rangle \qquad (19\text{-}4)$$

اگر آزمایش را در شکل (4-1ب) برای تعیین مسیری که ذره می‌پیماید انجام دهیم، ذره با احتمال $p_z^+=r_{11}$ در مسیر بالایی و با احتمال $p_z^-=r_{22}$ در مسیر پایینی یافت خواهد شد. بنابراین اندازه اطلاعات دربارهٔ مسیر ذره با توجه به تک کیوبیتی بودن سیستم به صورت:

$$I_3=(p_z^+-p_z^-)^2=(r_{11}-r_{22})^2 \qquad (20\text{-}4)$$

به دست می‌آید.

با استفاده از تبدیلات تحولی (4-18) احتمال پیدا کردن ذره در باریکه‌ای که به آشکارساز بالایی وارد می‌شود، همانند آرایه آزمایشی شکل (4-1الف) نشان داده شده است و به آسانی از $p_x^+=1/2(1+2|r_{21}|^2\sin(f+a))$ محاسبه می‌گردد. احتمال پیدا کردن ذره در باریکه‌ای که به سمت آشکارساز پایینی می‌رود به صورت $p_x^-=1/2(1-2|r_{21}|^2\sin(f+a))$ به دست می‌آید. در اینجا $r_{21}=e^{ia}|r_{21}|$ معرفی شده است. در این شرایط اندازه اطلاعات درباره ذره‌ای که در پشت شکافنده باریکه‌ها پیدا می‌شود، برابر

$$I_3=(p_x^+-p_x^-)^2=4|r_{21}|^2\sin^2(f+a) \qquad (21\text{-}4)$$

می‌باشد. با وارد کردن $f+p/2\to f$ در عبارت پیشین، اندازه اطلاعات درباره ذرات در پشت شکافنده باریکه‌ها (شکل (4-1ج)) به صورت:

$$I_2=(p_y^+-p_y^-)^2=4|r_{21}|^2\cos^2(f+a) \qquad (22\text{-}4)$$

به دست می‌آید. سرانجام با استفاده از مطالب گفته شده، محتوای اطلاعات کل این سیستم برابر مقدار زیر است:

$$I_{\text{total}} = I_1(p_1^+, p_1^-) + I_2(p_2^+, p_2^-) + I_3(p_3^+, p_3^-) = 2\text{Tr}\,\hat{r}^2 - 1 \qquad (23\text{-}4)$$

این رابطه برای همه سیستم‌های دوحالته از جمله سیستم‌های اسپین $1/2$ صادق است. در رابطهٔ بالا p_1^+ احتمال پیدا کردن ذره در حالت $\hat{r}$ با اسپین بالا در طول f می‌باشد. آشکارا رابطه (23-4) تحت تبدیلات یکانی ناوردا است زیرا $\hat{r}$ یک عملگر هرمیتی است. همچنین هنگامی که موقعیت‌هایی با مقادیر صحیح به سیستم نسبت داده شود این رابطه برای حالت خالص یک بیت اطلاعات را نتیجه می‌دهد. برای حالت کاملاً آمیخته که هیچ موقعیتی با مقادیر صحیح به سیستم نسبت داده نمی‌شود، صفر بیت از اطلاعات را نتیجه می‌دهد. بنابراین محتوای اطلاعات کل سیستم کاملاً با حالت سیستم مشخص می‌شود و از پارامتر فیزیکی f که سری‌های مختلف مشاهده‌پذیرهای دوبدو جابجاناپذیر را برچسب می‌زند مستقل می‌باشد و می‌توان از هر سری از موقعیت‌های مکمل دوجانبه برای نمایش میزان عدم یقین سیستم استفاده کرد. بنابراین محتوای اطلاعات تحت انتخاب سری مشاهده‌پذیرهای دوبدو جابجا ناپذیر مستقل است. بنابراین منطقی است که از اصطلاح «محتوای اطلاعات کل» استفاده کنیم بدون اینکه آشکارا سری موقعیت‌های مکمل دوجانبه مرجع را مشخص کرده باشیم.

اکنون حالتی را در نظر بگیرید که سیستمی مرکب از دو کیوبیت داشته باشیم. در این حالت 5 جفت موقعیت مکمل دو جانبه داریم.[66] یک مثال مشخص با دو ذره اسپین $1/2$ به دست می‌آید. سری کاملی از جفت موقعیت‌های مکمل دوجانبه را برای این سیستم در نظر می‌گیریم. در مورد دو ذره‌ای نیز دانش کامل درباره یک جفت از سری مشاهده‌پذیرهای جابجاناپذیر ما را از داشتن هر گونه دانشی درباره بقیه محروم می‌کند. یک گزینش ممکن برای سری کامل جفت موقعیت‌های مکمل برای دو ذره به صورت زیر است: (1) «اسپین ذرهٔ 1در طول محور z به سمت بالا باشد» و « اسپین ذرهٔ2 در طول محور z بالا باشد؛» (2) «اسپین ذرهٔ 1در طول f_1 به سمت بالا باشد» و « اسپین ذره 2 در

[66] برای $n = 2^k$ ذره $2^k + 1$ مشاهده‌پذیر مکمل دوجانبه وجود دارد.

طول f_2 بالا باشد؛»، (3) « اسپین ذره 1در طول $f_1+p/2$ به سمت بالا باشد» و « اسپین ذرهٔ 2 در طول $f_2+p/2$ بالا باشد؛» (4) «اسپین ذرهٔ 1در طول z و اسپین ذره 2 در طول f_2 مشابه باشند؛» و «اسپین ذرهٔ 1در طول f_1 به سمت بالا باشد» و «اسپین ذرهٔ 2 در طول $f_2+p/2$ مشابه باشند؛» (5) « اسپین ذرهٔ 1در طول z و اسپین ذرهٔ 2 در طول $f_2+p/2$ مشابه باشند؛» و « اسپین ذرهٔ 1در طول f_1 و اسپین ذرهٔ 2 در طول f_2 مشابه باشند.» باز هم فرض می‌شود جهت‌های f_1 و f_2 در صفحه $x-y$ به ترتیب در طول زاویه f_1 و f_2 که نسبت به محور xها ساخته می‌شوند قرار گیرد.

محتوای اطلاعات کل حمل شده از رابطهٔ زیر به‌دست می‌آید:

$$I_{\text{total}}=\sum_{j=1}^{5} I_j(\mathbf{p}^j)=\frac{2}{3}(4\,\mathrm{Tr}\,\hat{r}^2-1) \qquad (24\text{-}4)$$

بنابراین در مورد سیستم دو کیوبیتی نیز محتوای اطلاعات کل سیستم تحت تبدیلات یکانی ناوردا باقی می‌ماند و از جهت‌های f_1 و f_2 مستقل است.

اگر به افزایش کیوبیت‌ها ادامه دهیم برای سیستمی مرکب از kکیوبیت، محتوای اطلاعات کل از رابطه:

$$I_{\text{total}}=\sum_{i=1}^{2^k-1} I_j(\mathbf{p}^j)=\frac{k}{2^k-1}(2^k\,\mathrm{Tr}\,\hat{r}^2-1) \qquad (25\text{-}4)$$

به‌دست می‌آید که آشکارا تحت تبدیلات یکانی ناوردا است.

گفته‌های بالا همه مؤید آن است که اندازه اطلاعات معرفی شده، رابطه‌ای مناسب برای کمی کردن اطلاعات کوانتمی می‌باشد و مشکلات آنتروپی شانون را ندارد. همچنین اگر سیستم برهمکنشی با محیط اطراف نداشته باشد و جابجایی اطلاعات با محیط صورت نگیرد اندازه اطلاعات معرفی شده در زمان پایسته است. بدون در نظر گرفتن ضریب بهنجارش می‌توان محتوای اطلاعات کل را به صورت $I_{\text{total}}:=\sum_{j=1}^{m} I(p_1^j,\mathbf{K},p_n^j)=\sum_{j=1}^{m}\sum_{i=1}^{n}(p_i^j-1/n)=\mathrm{Tr}\hat{r}^2-1/n$ [67] خلاصه کرد. که p_i^j احتمال مشاهده iمین خروجی از jمین مشاهده‌پذیر می‌باشد. لازم به ذکر است که محتوای اطلاعات نمی-

[67] $I(p_1,\mathbf{K},p_n)=\sum_{i=1}^{n}(p_i-1/n)^2$ اطلاعات به دست آمده از اندازه‌گیری یک مشاهده‌پذیر می‌باشد.

تواند از مقدار بیشینه‌ای که همه اطلاعات سیستم در یک تک مشاهده‌پذیر ذخیره شده‌اند تجاوز کنند. در این حالت $p_i = 1$ وسیستم در حالت خالص و $I_{\text{total}} = (n-1)/n$ است. اما اگر سیستم در حالت کاملاً آمیخته $p_i = 1/n$ باشد، اطلاعات کل کمینه و مقدار آن برابر با $\mathbf{0}$ است. بنا به ناوردایی تحت تبدیلات یکانی، محتوای اطلاعات کل سیستم به سری ویژهٔ مشاهده‌پذیرهای دوبه‌دو جابجاناپذیر بستگی ندارد و تنها مشخصه‌ای از حالت سیستم است. همچنین چون تحول زمانی سیستم با عملگری یکانی توصیف می‌شود، اگر جابجایی اطلاعات با محیط صورت نگیرد، اطلاعات کل سیستم در زمان پایسته است[8 و 12].

4-2- آنتروپی مرکب درحکم تصحیحی برای آنتروپی فون نیومن

گفتیم که گاهی آنتروپی فون‌نیومن درحکم معادل آنتروپی شانون در مکانیک کوانتومی به‌کار می‌رود، اما گروهی بر این باورند که این کاربرد از نقطه نظر مکانیک کوانتومی با اشکال مواجه می‌شود، زیرا آنتروپی فون نیومن برای حالت خالصی که در حالت پایه باشد، مقدار $S = \mathbf{0}$ را نتیجه می‌دهد. مثلاً اگر حالت مورد نظر ما مربوط به ذره‌ای اسپین $1/2$ باشد که در امتداد محور z قطبیده شده است و فرض کنیم که حالت پایه ویژه‌حالت‌های s_z باشد دراین صورت ماترس چگالی به شکل:

$$\begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \qquad (26\text{-}4)$$

درمی‌آید. که آنتروپی فون نیومن آن صفر است[68]. درست است که حالت خالصی با این شرایط، حالتی با کمترین عدم قطعیت است ولی به این معنی نیست که باید $S = \mathbf{0}$ باشد. زیرا ما هنوز هم مقداری در خروجی اندازه‌گیری نامطمئن هستیم و بنابراین آشکارا آنتروپی فون‌نیومن عدم قطعیت ذاتی را که با این گروه از حالت‌های کوانتومی وجود دارد، منعکس نمی‌کند[24 و 25]. بنابراین گاهی پیشنهاد می‌شود که به دنبال آنتروپی باشیم که مشکلات این چنینی نداشته باشد. چون آنتروپی فون

[68] هنگامی که سیستم خالص در حالت پایه نباشد، آنتروپی فون نیمن آن مخاف صفر و بیشینه است.

نیومن برای حالت‌های خالص اشاره شده، مقدار صفر را نتیجه می‌دهد بنابراین می‌توان آن را درحکم اندازه‌ای برای فقدان خالص بودن ملاحظه کرد.

بنابر مشکل اشاره شده برای آنتروپی فون نیومن ممکن است احساس نیاز به آنتروپی دیگری شود که اینگونه مشکلات را نداشته باشد و برای حالت‌های خالص به $S=0$ منتهی نشود. آنتروپی که معرفی می‌شود، همانند مورد فون نیومن، می‌تواند درحکم اندازه‌ای از کمبود خالص بودن یک حالت کلی تلقی شود ولی برخلاف آنتروپی فون نیومن برای همه حالت‌های خالص، مقدار $S=0$ نتیجه نمی‌شود. روشی که برای دستیابی به این اندازه اطلاعات اتخاذ می‌شود بر پایه آنتروپی شانون بنا می-شود.

می‌دانیم که حالت آماری یک سیستم کلاسیکی که می‌تواند در یکی از N حالت ممکن r باشد با احتمال‌های متناظر p_r و شرط بهنجارش $\sum p_r =1$ مشخص می‌شود. مقدار اطلاعاتی که در خروجی اندازه‌گیری به دست می‌آید از آنتروپی شانون، $H=-\sum_r p_r \ln p_r$، محاسبه می‌شود. توجه شود که اگر سیستم در یک حالت مشخص باشد، $H=0$ است و در بدترین حالت که توزیعی روی همه احتمال‌ها یکنواخت باشد $H=\ln(N)$ است. اگر r سلول‌های فضای فاز در نظر گرفته شود، تعریف شانون بر تعریف بولتزمن از آنتروپی منطبق است.

در مکانیک کوانتمی که حالت آماری سیستم به وسیله یک ماتریس r مشخص می‌شود، هر اندازه‌گیری نیازمند این است که حالت‌های پایه خالص $|a\rangle$ را مشخص کنیم. بدون از دست دادن کلیت مسأله مناسب است که پایه‌های داده شده را با مشخص کردن عملگر هرمیتی A تعریف کنیم. گاهی در تفسیر نیمه کلاسیکی پایه‌های A می‌توانند تفکیکی از فضای فاز به سلول‌ها در نظر گرفته شوند. احتمال اینکه خروجی اندازه‌گیری a باشد با $\langle a|r|a\rangle$ داده می‌شود. بنابراین آنتروپی اطلاعات برای چنین اندازه‌گیری‌ای با:

$$S[\,r\,|A\,]=-\sum_a \langle a|r|a\rangle \ln(\langle a|r|a\rangle) \qquad (27\text{-}4)$$

مشخص می‌شود. نمادگذاری[69] بالا تأکیدی بر این امر است که در حقیقت این آنتروپی یک آنتروپی شرطی می‌باشد زیرا لازم است از ابتدا ابزار اندازه‌گیری معلوم باشد. به طور خاص پایه‌های H وجود دارد که r در آن‌ها قطری است، $r=\mathrm{diag}\{p_r\}$. در این پایه‌ها $S[\,r\,|A\,]$ کمترین مقدار خود را می‌گیرد.

$$S_H[\,r\,]=S[\,r\,|A\,]=-\sum_a p_r \ln p_r =-\mathrm{Tr}(r\ln r) \qquad (28\text{-}4)$$

بنابراین همان‌گونه که در فصل سوم اشاره شد اگر ماتریس چگالی سیستم در پایه‌های ابزار اندازه-گیری قطری باشد آنتروپی شانون با آنتروپی فون نیومن برابر می‌شود. در [23] تأکید شده است که از نقطه نظر نظریهٔ اطلاعات، تنها در صورتی می‌توان کمیت $S_H[\,r\,]$ را آنتروپی اطلاعات تفسیر کرد که از پایه‌های مرجحی که r را قطری می‌سازند آگاهی داشته باشیم. در مکانیک آماری این حالت به حالت سکون موسوم است. بنابراین r در پایه‌هایی که با هامیلتونی H مشخص می‌شود، قطری می-شود. بنابراین اگر انرژی سیستم را اندازه بگیریم آنتروپی اطلاعات واقعاً $S_H[\,r\,]$ می‌شود. زیرا r در این پایه‌ها قطری است.

گفتیم که برای یک حالت خالص که بر پایه‌های مفروض منطبق باشند، بنا به تعریف فون نیومن مقدار $S_H[\,r\,]=\mathbf{0}$ به‌دست می‌آید. ممکن است به نظر آید که یک حالت کوانتمی خالص که در پایه-های H قطری است طبیعت آماری ندارد، که البته این گفته درست نیست و برای یک اندازه‌گیری کلی هنوز هم مقداری عدم یقین با این حالت وجود دارد. یک تعریف مطلق برای آنتروپی اطلاعات حالات کوانتمی نباید هیچ پایه مشخصی را ترجیح دهد. این گفته بیان می‌کند که باید تعریفی یکتا از آنتروپی مطلق داشته باشیم.

اگر ابزار اندازه‌گیری را بخشی از سیستم در نظر بگیریم نظریه استاندارد اطلاعات بیان می‌کند که آنتروپی $S_{\text{total}}=S[\,A\,]+\sum_A P(\,A\,)S[\,r\,|A\,]$. احتمال $P(\,A\,)$ کمبود دانش ما را با در نظر

[69] Notation

گرفتن حالت ابزار اندازه‌گیری توصیف می‌کند و $S[A]$ آنتروپی متناظر است. با استفاده از این بحث نتیجه می‌گیریم که:

$$S[r]=\overline{S[r|A]}=S_0(N)+F(p_1,p_2,\mathbf{K})\equiv S_0(N)+S_F[r] \qquad (29\text{-}4)$$

که خط بالای جمله بعد از تساوی اول متوسط‌گیری روی همه سری‌های پایه ممکن با اندازه یکنواخت را و نه پایه خاصی را نشان می‌دهد. تأکید می‌کنیم که روش متوسط‌گیری یکتا می‌باشد. گزینش یک پایه همانند گزینش یک جهت در فضای $2N-1$ بعدی است. تساوی دوم در معادله (27-4) عبارتی آشکار برای آنتروپی مطلق می‌دهد که در زیر استنتاج می‌شود. نتیجه به شکل جمع دو جمله نوشته می‌شود: جمله اول آنتروپی عدم یقین کمینه یک حالت کوانتمی است، در حالی که جمله دوم انحراف از خالص بودن است. جمله دوم را آنتروپی آماری اضافه می‌نامیم و نماد $S_F[r]$ برای آن برمی‌گزینیم. در نتیجه منطقی که بپرسیم چه مقدار آنتروپی $S_F[r]$ به آنتروپی $S_H[r]$ مربوط است.

فرض کنید که $r=\mathrm{diag}\{p_r\}$ در بعضی پایه‌های H قطری است. ما می‌توانیم همه سری پایه‌های ممکن A را چرخش یکانی H در نظر بگیریم. یعنی هر $a\in A$ در پایه چرخیده از تأثیر عملگر یکانی U روی پایه‌های مرجع $r\in H$ به‌دست می‌آید. در نتیجه

$$\begin{aligned}
S &= \overline{\sum_a f\left(\sum_r p_r\left|\langle r\,|\,a\rangle\right|^2\right)}^{A} = \overline{\sum_s f\left(\sum_r p_r\left|\langle r\,|U|\,s\rangle\right|^2\right)}^{U} \\
&= N\,\overline{f\left(\sum_r p_r\left|\langle r\,|\,y\rangle\right|^2\right)}^{y} = N\,\overline{f\left(\sum_r p_r\left(x_r^2+y_r^2\right)\right)}^{\text{sphere}} \\
&= N\int_0^\infty f(s)P(s)\,\mathrm{d}s
\end{aligned} \qquad (30\text{-}4)$$

که از نمادگذاری $f(s)=-s\ln s$ استفاده شده است. هر $\left|\langle r\,|U|\,y\rangle\right|^2$ میانگین‌گیری شده در خط اول معادله (24-4) با $\left|\langle r\,|\,y\rangle\right|^2$ برابر است که روی همه yهای ممکن میانگین‌گیری شده است. مکانیک کوانتمی اندازه‌ای را برای متوسط y پیش‌بینی می‌کند. اگر x_r و y_r بخش‌های

حقیقی و موهومی $y_r = \langle r \mid y \rangle$ فرض شوند مطلب گفته شده در جمله پیش واضح‌تر به نظر می‌آید. شرط بهنجارش $\sum_r x_r^2 + y_r^2 = 1$ می‌باشد. از این رو در عبارت آخر متوسط‌گیری روی همه ابعاد ممکن در فضای $2N - 1$ بعدی است. در عبارت نهایی از نمادگذاری:

$$s = \sum_r p_r \left| y_r \right|^2 = \sum_r p_r \left(x_r + y_r \right)^2 \qquad (31\text{-}4)$$

استفاده شده است. $P(s)$ توزیع احتمال می‌باشد. در ادامه در مورد محاسبه $P(s)$ و $f(s)$ بحث می‌کنیم. در مورد یک حالت کاملاً آمیخته $P(s)$ توزیعی دلتایی حول $s = 1/N$ دارد، و از این رو $f(s) = \ln N / N$. و همانند انتظار آنتروپی اطلاعات متناظر $S(r) = \ln N$ می‌باشد. اگر حالت کاملاً آمیخته نباشد $P(s)$ ناچیز نمی‌باشد. در مورد یک حالت کاملاً خالص $s = \left| y_1 \right|^2$ و توزیع احتمال شناخته شده است[18]:

$$P(s) = (N-1)(1-s)^{N-2} \qquad (32\text{-}4)$$

بنابراین عبارتی به شکل زیر برای آنتروپی عدم یقین کمینه به‌دست می‌آوریم که به N وابسته است:

$$S_0(N) = \sum_{k=2}^{N} \frac{1}{k} \approx \ln(N) - (1-g) + \frac{1}{2N} \qquad (33\text{-}4)$$

با توجه به تقریب مجانبی در تساوی آخر می‌بینیم که تفاوت S بین حالت کاملاً آمیخته و یک حالت خالص، به مقدار جهانی $(1-g)$، (g ثابت اولر می‌باشد) نزدیک می‌شود. با استفاده از اصطلاحی متفاوت، می‌بینیم که آنتروپی آماری اضافه به صورت زیر مقید می‌شود:

$$S_F(r) < (1-g) \qquad (34\text{-}4)$$

اما اگر حالت ما نه کاملاً خالص و نه کاملاً آمیخته باشد کار مشکل می‌شود و باید $P(s)$ را محاسبه کرده تا بتوانیم انتگرال (4-30) را محاسبه کنیم. در حالت کلی داریم:

$$p(s) = (N-1) \sum_{p_r > s} \left[\prod_{r' \neq r} \frac{1}{p_r - p_{r'}} \right] (p_r - s)^{N-2} \qquad (35\text{-}4)$$

و با استفاده از این رابطه می‌توانیم انتگرال(4-28) را محاسبه کنیم:

$$\int_0^p (p-s)^{N-2} s \ln(s)\,ds = \frac{p^N}{N(N-1)}\left[\ln(p) - \sum_{k=2}^{n} \frac{1}{2}\right] \quad (36\text{-}4)$$

و سپس با استفاده از :

$$\sum_r p_r^N \prod_{r'(\neq r)} \frac{1}{p_r - p_{r'}} = \sum_r p_r = 1 \quad (37\text{-}4)$$

به دست می‌آوریم:

$$F(p_1, p_2, \mathbf{K}) = -\sum_r \left[\prod_{r' \neq r} \frac{p_r}{p_r - p_{r'}}\right] p_r \ln(p_r) \quad (38\text{-}4)$$

برای ترکیبی از n حالت کاملاً آمیخته داریم:

$$S_f[r] = \ln(n) - \sum_{k=2}^{n} \frac{1}{n} \quad (39\text{-}4)$$

$$S[r] = \ln(n) + \sum_{n<k\leq N}^{n} \frac{1}{k} \quad (40\text{-}4)$$

هم $S_H[r]$ و هم $S_F[r]$ می‌توانند اندازه‌ای از کمبود خلوص محسوب می‌شوند. در شکل (2-4) نتایج محاسبه $S_F[r]$ در مقابل $S_H[r]$ برای یک سری از حالت‌ها که به عنوان نماینده انتخاب می‌شوند رسم شده است (هم حالت‌های آمیخته یکنواخت و هم حالت‌های آمیخته نایکنواخت). کاملاً واضح است که رابطه‌ای خیلی قوی بین این دو اندازه که هر دو نشانه کمبود خالص بودن‌اند وجود دارد.

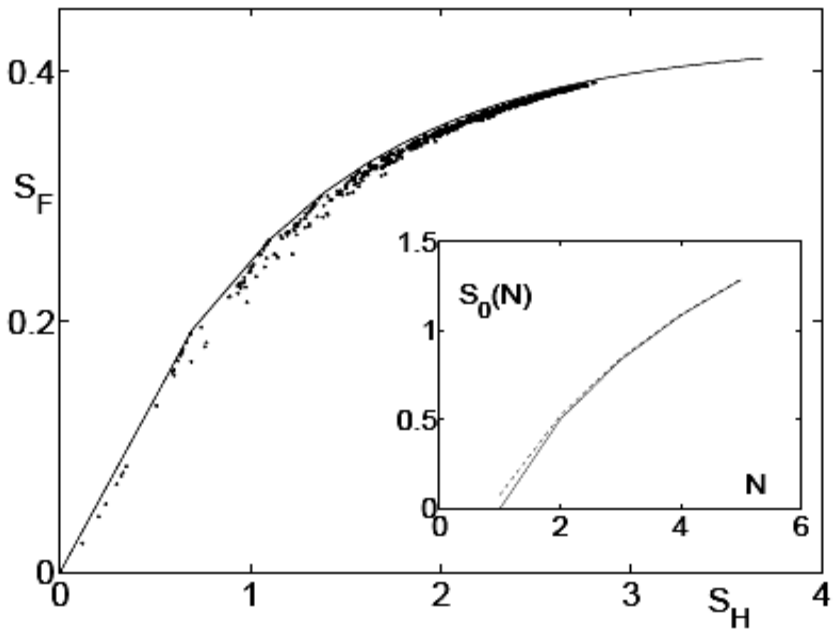
تعریفی که از آنتروپی ارائه شد بعضی ویژگیهای جالب دارد. یک ویژگی جالب آن مقعر بودن آن می‌باشد. مثلاً برای $0 < l < 1$ و دو سری از احتمال‌ها داریم:

$$F(l p_r + (1-l) q_r) \geq l F(p_r) + (1-l) F(q_r) \quad (41\text{-}4)$$

زیرا $f(s)$ مقعر می‌باشد. تقعر و تقارن با توجه به متغیرهای p_i بیان می‌کنند که $S[r]$ بیشینه‌ای برای حالت‌های کاملاً آمیخته و کمینه‌ای برای حالت‌های خالص دارد. این ویژگی ما را در توجیه کردن بحث‌هایی که روی بدترین حالت پایه شدند کمک می‌کند[23].

3-4- مقایسه اهمیت دو آنتروپی معرفی شده با آنتروپی فون نیومن و کارایی آن‌ها

در دو بخش پیش به تفصیل در مورد آنتروپی‌های جدیدی که اطلاعات کوانتمی را کمّی می‌کنند صحبت کردیم. با وجود اینکه این اندازه‌ها در برخی موارد مفید واقع می‌شوند، نمی‌توان آنها را جایگزین آنتروپی فون نیومن کرد. زیرا آنتروپی فون نیومن علی‌رغم بعضی اشکالات وارده با گسترشی که در حوزه تعیین کانال‌های کوانتمی کمینه یافته، اندازه‌ای بی‌نظیر در نظریهٔ اطلاعات کوانتمی می‌باشد. از طرفی همان‌گونه که در بخش پیشین اشاره شد گروهی عقیده دارند که آنتروپی فون نیومن عدم یقین ذاتی حالت‌های کاملاً خالص را برآورده نمی‌کند. اگر به استفاده‌ای که در نظریهٔ اطلاعات از آنتروپی فون نیومن می‌شود توجه کنیم، کاملاً واضح به نظر می‌رسد که هدف تعیین محتوای اطلاعاتی یک سیستم کوانتمی به صورت کمّی می‌باشد. چون محتوای اطلاعاتی هر سیستم با توجه به سیستمی که درحکم مرجع برگزیده می‌شود، سنجیده شده و کاملاً نسبی می‌باشد. بنابراین به‌جاست که آنتروپی فون نیومن یک حالت کاملاً خالص را حالتی مرجع برگزیده و عدم یقین یا محتوای اطلاعاتی بقیه سیستم‌ها با توجه به آن محاسبه کنیم. می‌توان ادعا کرد که با وجود ایراد وارد بر آنتروپی فون نیومن چیزی از اهمیت این آنتروپی در حوزه نظریهٔ اطلاعات کاسته نمی‌شود. در واقع



شکل(4-2).آنتروپی اطلاعات اضافه، S_F ، یک حالت کوانتمی آمیخته در مقابل آنتروپی فون نیومن، S_H . خط پر برای مخلوط‌های یکنواخت می‌باشد، در حالی که نقطه‌ها حالت‌های آمیخته‌ایست که تصادفاً انتخاب شده‌اند. منحنی درونی آنتروپی حالت کوانتمی خاص به‌شکل بعدی می‌باشد. N تابعی از فضای هیلبرت

می‌توان گفت که اگرچه آنتروپی فون نیومن عدم یقین ذاتی بعضی سیستم‌ها را نشان نمی‌دهد اما اشکالات این چنینی بهتر است در محدوده مکانیک آماری کوانتمی بررسی شود و نه در حوزه نظریهٔ اطلاعات.

4-4- سیستم‌های سه‌تایی کلاسیکی

تاکنون همه بحث‌های ارائه شده بر پایهٔ سیستم‌های دوتایی و یا به بیانی دیگر سیستم‌های دو حالتی بوده است. ولی در عمل سیستم‌های مورد استفاده در کامپیوترها می‌توانند به جای دوحالتی بودن سه حالتی باشند یا به اقتضای استفاده‌ای که دارند مبنای بیشتری را اختیار کنند. مسلماً استفاده از سیستم‌های با مبنای سه و بالاتر هم از نظر هزینه‌های اقتصادی و هم از نظر هزینه‌های زمانی در عمل بسیار مفیدتر از سیستم‌های با مبنای دو می‌باشند. این مطلب به طوری شهودی کاملاً

طبیعی و قابل درک می‌باشد. مثلاً اگر سیستمی 8 بیتی در اختیار داشته باشیم حداکثر 2^8 بیت اطلاعات را می‌تواند ذخیره کند اما اگر سیستمی 8 تریتی[70] در اختیار داشته باشیم 3^8 تریت اطلاعات را می‌توان در این سیستم ذخیره کرد که این مقدار از سیستم بیتی بسیار بیشتر می‌باشد.

پیش از پرداختن به آنتروپی سیستم‌های سه تایی مثالی را که در فصل 1 برای نشان دادن استخراج کار از اطلاعات ذکر کردیم این بار برای سیستم‌های سه حالتی تکرار می‌کنیم و به طور ساده نشان می‌دهیم که سیستم‌های سه حالتی مفیدتر از سیستم‌های دو حالتی می‌باشند. در این مثال نشان داده شد که از اطلاعات می‌توان در حکم سوخت برای حرکت ماشین استفاده کرد. در آنجا گفته شد که اگر یک نوار N بیتی داشته باشیم کاری برابر $N\ k_B T \ln 2$ می‌توان از آن استخراج کرد. به همان روش می‌توان ثابت کرد که اگر یک نوار N تریتی داشته باشیم، کاری معادل $N K_B T \ln 3$ قابل استخراج است. بنابراین واضح است که از سیستم تریتی کار بیشتری می‌توان استخراج کرد.

در این بخش به سیستم‌های سه حالتی کلاسیکی در نظریهٔ اطلاعات پرداخته و ویژگیهای مطلوب آن‌ها را با جزئیات بیشتر از نظر می‌گذرانیم.

4-4-1- آنتروپی اطلاعات سیستم‌های سه‌حالتی کلاسیکی یا سیستم‌های تریناری[71]

گفتیم که بهترین روش برای تعیین محتوای اطلاعاتی یک سیستم کلاسیکی استفاده از آنتروپی شانون می‌باشد و اظهار کردیم که برای سازگاری این آنتروپی با سیستم‌های دو حالتی، مبنای لگاریتم را در رابطه شانون 2 اختیار می‌کنیم. گزینش مبنای 2 در رابطهٔ شانون را خود او پیشنهاد کرد. اکنون پرسشی که در ذهن مطرح می‌شود این است که آیا آنتروپی شانون برای تعیین محتوای اطلاعاتی یک سیستم کلاسیکی سه‌حالتی نیز مفید و قابل استفاده است یا خیر؟

برای بررسی مناسب بودن آنتروپی شانون برای سیستم‌های تریتی ابتدا بهتر است مفهوم واحد اطلاعات کلاسیکی را یادآوری کنیم. در کامپیوترهای کلاسیکی واحد اطلاعات بیت می‌باشد. در واقع

[70] سیستم‌های کلاسیکی که مبنای آن‌ها سه می‌باشد، سیستم‌های تریتی نامیده می‌شوند.

[71] Trinary

یک بیت مقدار ماکزیمم اطلاعاتی است که یک وسیله یا سیستم فیزیکی که دو حالت گسسته دارد می‌تواند ذخیره شود. در نظریهٔ اطلاعات یک بیت عدم یقین یک متغیر تصادفی باینری است که با احتمال‌های مساوی صفر یا یک می‌باشد. اگر از آنتروپی شانون برای تعیین محتوای اطلاعات یک سیستم کلاسیکی مبنا بیتی استفاده کنیم مقدار این کمیت هرچه که به دست آید برحسب بیت بیان می‌شود.

$$H = -\sum_i p_i \log_2 p_i \qquad (42\text{-}4)$$

حال اگر سیستم ما به جای بیتی بودن تریتی باشد، برای انطباق رابطه بالا با سیستم‌های تریتی مبنای لگاریتم را 3 انتخاب می‌کنیم. با انتخاب این مبنا مشخصاً مقدار عددی که برای یک توزیع احتمالی (مثلاً توزیع احتمالی مربوط به یک سری رویدادها که به صورت $(p_1, p_2, \mathbf{K}, p_n)$داده شده است.) بدست می‌آید کمتر از مقداری است که برای همین توزیع احتمالی در سیستم‌های بیتی بدست می‌آید. $(-\sum_i p_i \log_2 p_i < -\sum_i p_i \log_3 p_i)$. ولی این بار واحد اطلاعات تریت می‌باشد که از بیت بزرگ‌تر و حدوداً 1,58 برابر آن است. پس یک تریت مقدار اطلاعاتی است که یک سیستم سه حالتی می‌تواند در خود ذخیره کند. می‌توان نتیجه گرفت که کمتر بودن عدد سیستم تریتی(بدون در نظر گرفتن واحد) نسبت به سیستم بیتی، برای یک توزیع احتمال کاملاً منطقی بوده و مناسب‌تر بودن سیستم تریتی از سیستم بیتی را نشان می‌دهد و یا نشان می‌دهد که محتوای اطلاعاتی یک تریت از یک بیت بیشتر می‌باشد. زیرا با استفاده از تعداد تریت کمتر می‌توان اطلاعات بیشتری را ذخیره کرد.

4-4-2- برخی از مزیت‌های تریت بر بیت

4-4-2-1- مقایسه میزان متراکم سازی اطلاعات در سیستم‌های کلاسیکی بیتی و تریتی

همانگونه که می‌دانیم اطلاعات باید هنگام ذخیره‌سازی تا جایی که ممکن است متراکم شوند تا فضای کمتری برای ذخیره لازم باشد. چه دربایست‌های فیزیکی کمینه‌ای برای تعیین منبع اطلاعاتی لازم است؟ در سیستم‌های مبنا بیتی، نظریهٔ کدکردن کانال بدون نوفه شانون گستره‌ای را که می‌توانیم اطلاعات را فشرده کنیم مشخص می‌کند. می‌توان همین نظریهٔ را برای سیستم‌های مبنا تریتی و بالاتر بکار برد و تنها لازم است اصلاحات اندکی انجام داد.

در فصل دوم نشان دادیم که تعداد دنباله‌های بهنجار یک منبع اطلاعاتی i.i.d در یک سیستم باینری برابر $2^{nH(X)}$ می‌باشد، اکنون می‌خواهیم تعداد دنباله‌های بهنجار یک منبع i.i.d را در یک سیستم تریتی به دست آورده و از این طریق رهیافتی به سمت میزان متراک سازی اطلاعات در همین سیستم‌ها پیدا کنیم.

برای تعریف دنباله‌های بهنجار در سیستم‌های تریتی ابتدا یک منبع اطلاعاتی i.i.d را در نظر می‌گیریم که تریت‌های $X_1, X_2, X_3, \mathbf{K}$ را تولید می‌کند. هر یک از این تریت‌ها با احتمال p صفر، با احتمال q یک و با احتمال $1-p-q$ دو می‌باشند. اگر دنباله‌های ممکن برای متغیرهای تصادفی $X_1, X_2, X_3, \mathbf{K}$ را با $x_1, \mathbf{K}, x_n$ نشان دهیم، هنگامی که n بسیار بزرگ شود یعنی آزمایش به دفعات زیادی انجام گیرد، کسر p از نمادهای خروجی از منبع برابر صفر، کسر q از نمادهای خروجی از منبع برابر یک، و کسر $1-p-q$ از نمادهای خروجی از منبع برابر دو خواهد بود. چنین دنباله‌هایی را دنباله‌های بهنجار تریتی می‌نامیم. اگر شرط مستقل بودن خروجی‌ها را نیز به شرط بهنجار بودن دنباله‌ها اضافه کنیم، رابطهٔ زیر برقرار است:

$$p(x_1, \mathbf{K}, x_n) = p(x_1)p(x_2)\mathbf{K}\,p(x_n) \approx p^{np}q^{nq}(1-p-q)^{(1-p-q)n} \qquad (43\text{-}4)$$

حال از طرفین رابطه بالا در مبنای 3 لگاریتم می‌گیریم، درنتیجه:

[72]

(44-4)

$$-\log p(x_1,...,x_n) \approx -n\,p\log p - n\,q\log q - (1-p-q)\log(1-p-q) = n\,H'(X)$$

که X متغیری تصادفی است که مطابق توزیع منبع توزیع شده است و $H'(X)$ آنتروپی توزیع منبع می‌باشد. بنابراین $p(x_1,\mathbf{K},x_n) = 3^{-nH'(x)}$ است. پس می‌توان گفت حداکثر $3^{nH(X)}$ دنباله بهنجار برای سیستم‌های تریتی وجود دارد. چون احتمال کل همه دنباله‌ها نمی‌تواند بزرگتر از یک باشد.

از بحث بالا نتیجه می‌گیریم که در سیستم‌های تریتی تعداد دنباله‌های بهنجار بیشتر از سیستم‌های بیتی می‌باشد، زیرا هنگامی که n بسیار بزرگ شود، $3^{nH'(X)} > 2^{nH(X)}$ است.

نظریهٔ کدکردن کانال بدون نوفه شانون بیان می‌کند که فرض کنید $\{X_i\}$ یک منبع اطلاعاتی i.i.d با آهنگ آنتروپی $H(X)$ باشد. $R>H(X)$ را در نظر بگیرید. پس طرح متراکم سازی معتبری با آهنگ R برای منبع وجود دارد. برعکس اگر $R<H(X)$، چنین طرح متراکم‌سازی وجود نخواهد داشت[5].

اگر سیستم در مبنای تریتی باشد مسلماً این نظریهٔ برای آن نیز کاربرد خواهد داشت با این تفاوت که در اینجا اطلاعات را می توان فشرده‌تر کرد. اگر در اینجا نرخ فشرده‌سازی اطلاعات را با R' نمایش دهیم با توجه به $H(X) > H'(X)$، می‌توان نتیجه گرفت که لزوماً $R>R'$. بنابراین نرخ فشرده سازی اطلاعات در سیستم‌های تریتی بیشتر از سیستم‌های بیتی می‌باشد.

4-4-2-2- بررسی انتقال اطلاعات کلاسیکی در کانال‌های نوفه‌ای کلاسیکی

در فصل دوم نرخ انتقال اطلاعات برای سیستم‌های بیتی را، در کانال‌های نوفه‌ای بررسی کردیم. گفتیم که نظریهٔ کدکردن کانال نوفه‌ای شانون نشان می‌دهد که برای کانال نوفه‌ای N ظرفیت از رابطهٔ $C(\mathrm{N}) = \underset{p(x)}{\mathrm{M}\mathrm{ax}} H(X:Y)$ به‌دست می‌آید. در این رابطه بیشینه روی همه توزیع‌های ورودی $p(x)$ برای X و هر استفاده از کانال گرفته می‌شود. Y متغیر تصادفی منطبق در خروجی کانال می‌باشد. مثالی از نظریهٔ کدکردن کانال نوفه‌ای، مورد کانال متقارن باینری را در نظر بگیرید که در آن هر بیت با احتمال p، فلیپ می‌کند. اگر در این کانال توزیع ورودی $p(\mathbf{0}) = q$, $p(1) = 1-q$ داریم:

[72] $H'(X)$ را آنتروپی سیستم‌های تریتی در نظر می‌گیریم.

$$H(X:Y)=H(Y)-H(Y\mid X)=H(Y)-\sum_{x}p(x)H(Y\mid X=x). \qquad (45\text{-}4)$$

به‌ازای هر x، $H(Y\mid X=x)=H(p)$ بنابراین $H(X:Y)=H(Y)-H(p)$است، که با گـزینش $q=1/2$ بیشینه می‌شود. بنابراین $H(Y)=1$ و $C(N)=1-H(p)$.

در صحبت از سیستم‌های تریتی ادعا کردیم که این سیستم‌ها مناسب‌تر از سیستم‌های بیتی مـی-باشند. اگر بخواهیم بر همه‌جانبه و همیشگی بودن این ادعا تأکیـدی داشـته باشـیم لازم اسـت نـشان دهیم که استفاده از سیستم‌های تریتی نرخ انتقال اطلاعات را افزایش دهد.

فرض کنید سه نماد **0**، 1 و 2 موجود باشند و این نمادهـا را بـا آهنـگ 1000 نمـاد بـر ثانیـه و بـا احتمال‌های $p_0=p_1=p_2=1/3$ انتقال می‌دهیم. بنابراین منبعی داریم کـه اطلاعـات را بـا آهنـگ 1000 تریت بر ثانیه تولید می‌کند. اگر در هنگام انتقال اطلاعات، نوفه درون کانال باعث بروز خطـایی شود، برای مثال 0٫1 نمادهای دریافت شده صحیح نباشند، بنابراین آهنگ انتقال اطلاعات 1000 بیـت بر ثانیه نخواهد بود و برای اینکه بتوانیم اطلاعات فرسـتاده شـده را بـه طورکامـل در اختیـار داشـته باشیم مجبوریم بر روی اطلاعات دریافت شده اصلاحاتی انجام دهیم.

شانون اظهار داشت که آشکارا تصحیح مناسب بر روی مقـدار اطلاعـات انتقـال داده شـده، مقـدار اطلاعاتی است که در سیگنال دریافت شده از بین رفته و یا اصطلاحاً گم شده اسـت. و یـا بـه عبـارتی تصحیح مناسب عدم یقینی است که سیگنال دریافتی نسبت به از آنچه واقعـاً فرسـتاده شـده دارد. بـا نگاهی به تعریف آنتروپی شرطی می‌توان از این آنتروپی درحکم اندازه اطلاعات گم شده استفاده کـرد مشروط براینکه سیگنال دریافت شده را بدانیم. بنابراین اندیشهٔ شانون برای آهنگ انتقال اطلاعـات در کانال نوفه‌ای به صورت

$$R=H(X)-H(X\mid Y) \qquad (46\text{-}4)$$

فرمولبندی شد. مثلاً اگر تأثیر کانال به گونه‌ای باشد مانند مثال گفته شده 0٫1 اطلاعات دریافـت شـده صحیح نباشند هنگامی که خروجی ما به فرض **0** باشد بـا احتمـال 0٫99ورودی منطبـق **0** بـوده و بـا احتمال 0٫01 برابربا 1 یا 2 می‌باشد. بنابراین می‌توان آنتروپی تصحیح را به صورت زیر محاسبه کرد:

$$H(X\ |Y\ ) = -(0{,}99\log_3^{0{,}99} + {,}01\log_3^{0{,}01}) = {,}051\ \ \text{trit/symbol} \qquad (47\text{-}4)$$

یا 51 تریت بر ثانیه دچار اثرات مخرب نوفه شده‌اند بنابراین می‌توان گفت که آهنگ انتقال اطلاعات $1000 - 51 = 949$ تریت بر ثانیه می‌باشد. اگر همین مثال برای یک سیستم بیتی بکار برده شود و دو نماد **0** و 1 درحکم ورودی کانال منظور شوند با دریافت صفر در خروجی با احتمال $0{,}99$ ورودی **0** بوده و با احتمال $0{,}01$ ورودی 1 بوده است. آنتروپی تصحیح مقداری برابر

$$H(X\ |Y\ ) = -(0{,}99\log_2^{0{,}99} + 0{,}01\log_2^{0{,}01}) = 0{,}081\ \ \text{trit/symbol} \qquad (48\text{-}4)$$

می‌گیرد. و چون $1000 - 81 = 919$، بنابراین آهنگ انتقال اطلاعات در این کانال 919 بیت بر ثانیه است. این محاسبات ساده تاحدودی بیشتر بودن ظرفیت کانال تریتی را نسبت به کانال بیتی نشان می‌دهد، اگرچه درک این مطلب از نظر شهودی نیز کار سختی نیست.

در مثال دیگر فرض کنید اثرات نوفه آنقدر قوی باشد که نمادهای رسیده کاملاً مستقل از نمادهای ورودی باشند در یک کانال نوفه‌ای، بیتی اگر **0** دریافت شود، با احتمال $1/2$ ورودی کانال 1 و با احتمال $1/2$ ورودی **0** بوده است. بنابراین مقدار آنتروپی تصحیح 1 بیت بر نماد و یا 1000 بیت بر ثانیه می‌باشد. بنابراین آهنگ انتقال اطلاعات در این کانال برابر **0** می‌باشد. اگر همین مثال را برای کانال تریتی بکار بریم با فرض اینکه خروجی **0** باشد، ورودی کانال با احتمال $1/3$، برابر با **0** و با احتمال $1/3$ برابر با 1 و با احتمال $1/3$برابر با 2 می‌باشد. در اینجا نیز آهنگ انتقال اطلاعات برابر **0** به دست می‌آید که با منطق نیز سازگار است.

مثال‌های بالا همه شواهدی هستند بر این گفته که ظرفیت یک کانال نوفه‌ای تریتی همانند ظرفیت یک کانال نوفه‌ای بیتی از رابطه پیشنهادی شانون

$$C = \text{Max}\left(H(X\ ) - H(X\ |Y\ )\right) = \text{Max}\left(H(Y\ ) - H(Y\ |X\ )\right) \qquad (49\text{-}4)$$

به دست می‌آید و از نظر عددی از ظرفیت کانال بیتی بیشتر است. در فصل دوم برای یک کانال متقارن باینری مقدار این عبارت $1 - H(p)$ به دست آمد. برای یک کانال متقارن ترینری نیز همین رابطه برقرار است اما به این دلیل که $H(p)$ در یک سیستم تریتی از نظر عددی کمتر از یک

سیستم بیتی است مقدار $1-H(p)$ در سیستم بیتی بزرگ‌تر است و در این مورد نیز بیشتر بودن ظرفیت کانال تریتی ثابت می‌شود.

4-5 سیستم‌های کوانتمی سه حالتی

متناظر با سیستم‌های تریتی در کلاسیک، سیستم‌های سه حالتی‌ای نیز در فیزیک کوانتمی می‌توان در نظر گرفت که استفاده از آن‌ها در کامپیوترهای کوانتمی شرایط بسیار بهینه‌تری نسبت به سیستم‌های کیوبیتی به وجود می‌آورد. بنابراین به یک سیستم سه‌حالتی در مکانیک کوانتمی نام ویژه کیوتریت[73] نسبت داده می‌شود و به صورت زیر تعریف می‌شود:

$$|y\rangle = a|0\rangle + b|1\rangle + g|2\rangle \qquad (50\text{-}4)$$

با این شرط که ضرایب a,b,g می‌توانند اعداد حقیقی یا مختلط باشند. و همچنین شرط $|a|^2+|b|^2+|g|^2=1$ برقرار باشد.

در فصل سوم گفتیم که هر نقطه در کره بلاخ نشان‌دهندهٔ یک کیوبیت می‌باشد و توضیحاتی در مورد این تناظر و ماتریس چگالی کیوبیت‌ها ارائه دادیم. اکنون زمان آن است که در مورد این نکته صحبت کنیم که آیا برای کیوتریت‌ها نیز نقطه متناظری در فضا وجود دارد یا خیر؟ اگر چنین تناظری یافت شود ماتریس چگالی یک کیوتریت با توجه به این فضا چگونه نمایش داده می‌شود؟

در مورد کیوتریت مجبوریم از ماتریس‌های گلمن[74] به جای ماتریس‌های پائولی برای مشخص کردن ماتریس چگالی کلی یک کیوتریت استفاده کنیم [19 و 20]:

$$r = \frac{1}{3}\left(I + \sqrt{3} r_i . l_i\right) \qquad (51\text{-}4)$$

که $r_i \in \mathbf{i}^8$ بردار بلاخ حالت کیوتریت r بوده و l یک بردار صوری شامل ماتریس‌های گلمن است:

[73] Qutrit

[74] Gell-Mann

$$l_1=\begin{pmatrix}0&1&0\\1&0&0\\0&0&0\end{pmatrix},\quad l_2=\begin{pmatrix}0&-i&0\\i&0&0\\0&0&0\end{pmatrix},\quad l_3=\begin{pmatrix}1&0&0\\0&-1&0\\0&0&0\end{pmatrix},\quad l_4=\begin{pmatrix}0&0&1\\0&0&0\\1&0&0\end{pmatrix}$$
$$l_5=\begin{pmatrix}0&0&-i\\i&0&0\\0&0&0\end{pmatrix},\quad l_6=\begin{pmatrix}0&0&0\\0&0&1\\0&1&0\end{pmatrix},\quad l_7=\begin{pmatrix}0&0&0\\0&0&-i\\0&i&0\end{pmatrix},\quad l_8=\frac{1}{\sqrt{3}}\begin{pmatrix}1&0&0\\0&1&0\\0&0&-2\end{pmatrix} \qquad (52\text{-}4)$$

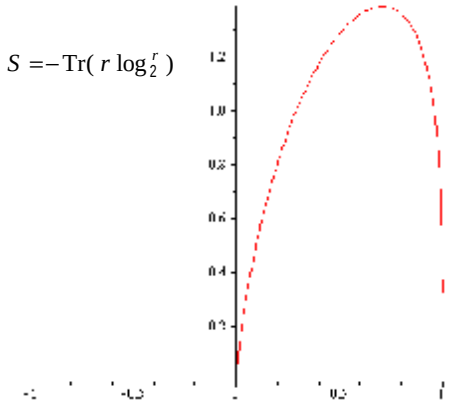
برای حالت‌های کیوتریت خالص $|r|=1$، و برای حالت‌های کیوتریت آمیخته $|r|<1$ می‌باشد. اگرچه حالت‌های با $|r|\le 1$ همیشه بر حالت‌های کیوتریت معتبر منطبق نیستند زیرا حالت‌های کیوتریت معتبر حالت‌هایی هستند که برای آن‌ها ویژه‌مقادیر ماتریس چگالی نامنفی باشد. مثلاً بردار بلاخ $r=(0,0,0,0,0,0,0,1)$ یک کیوتریت معتبر نمی‌سازد زیرا ویژه مقادیر ماتریس چگالی حاصل از آن برابر $2/3$, $2/3$ و $1/3$- می‌باشد.

4-5-1 آنتروپی اطلاعات یک سیستم کیوتریتی

پس از معرفی کیوتریت اکنون زمان آن رسیده که محتوای اطلاعاتی یک سیستم کیوتریتی را بیابیم و نشان دهیم که از نظر نظریهٔ اطلاعات کوانتمی، برای ذخیره اطلاعات یک کیوتریت مناسب‌تر از یک کیوبیت می‌باشد. در شکل(4-3) نموداری می‌بینیم که به طور تقریبی می‌توان با استفاده از آن بازه تغییرات محتوای اطلاعاتی سیستم‌های کیوبیتی را که بر اساس آنتروپی فون نیومن به‌دست آمده‌اند، استخراج کرد. لازم به ذکر است که این منحنی کاملاً تقریبی بوده و نمودار تنها برای r های حقیقی رسم شده است. با وجود این تقریب‌ها می‌توان به نتایجی مطلوب دست یافت.

$$|y\rangle=a|0\rangle+b|1\rangle,\qquad |a|^2+|b|^2=1,\qquad r=\begin{pmatrix}a^2&ab\\ba&b^2\end{pmatrix} \qquad (53\text{-}4)$$

مشاهده می‌کنیم که بسته به کیوبیتی که به سیستم مورد نظرمان نسبت می‌دهیم، محتوای اطلاعاتی سیستم ما می‌تواند از صفر تا حدود 2 تغییر کند. بنابراین محتوای اطلاعاتی یک کیوبیت حداکثر می‌تواند 2 باشد. که مقدار بیشینه مربوط به یک کیوبیت خالص می‌باشد. این بحث‌ها بدون توجه به نمودار نیز به طور شهودی قابل درک‌اند.

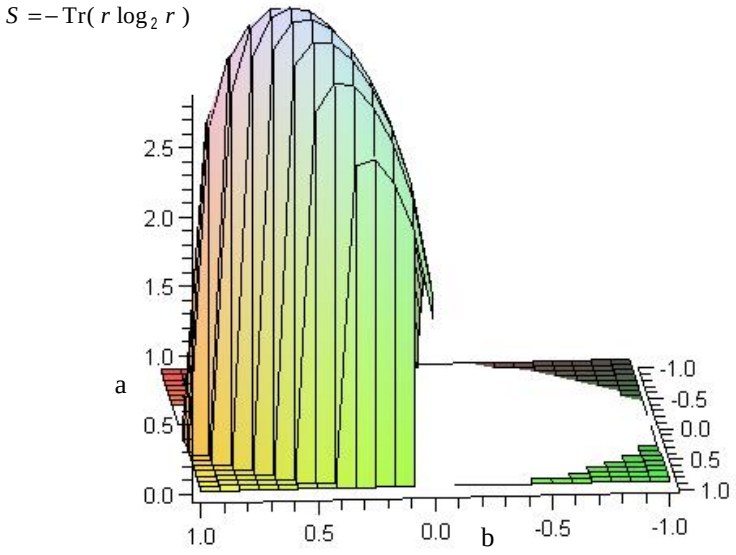


)، نمودار آنتروپی فون نیومن برحسب تغییرات یکی از دامنه‌های احتمال با این شرط که تنها دامنه‌های3شکل(4-
احتمال حقیقی در نظر گرفته شده‌اند.

اکنون به سیستم‌های کیوتریتی می‌پردازیم. برای مشاهده رفتار کلی محتوای اطلاعاتی سیستم‌های کیوتریتی و یافتن نموداری متناظر نمودار بالا برای این سیستم‌ها توجه خود را به شکل(4-4) جلب می‌کنیم. در این شکل بازه‌ای که محتوای اطلاعاتی یک سیستم کیوتریتی می‌تواند اختیار کند نشان داده شده است. همانند پیش باز هم نمودار برای r های حقیقی رسم شده است. بنابراین مقادیر به-دست آمده از این نمودار اگرچه تقریبی می‌باشند ولی می‌توانند ما را به نتایج صحیح رهنمون سازند.

$$|y\rangle = a|\mathbf{0}\rangle + b|1\rangle + c|2\rangle\,, \quad |a|^2 + |b|^2 + |c|^2 = 1\,, \quad r = \begin{pmatrix} a^2 & ab & ac \\ ba & b^2 & bc \\ ca & cb & c^2 \end{pmatrix} \qquad (54\text{-}4)$$

بنابراین همان‌گونه که انتظار می‌رفت سیستم‌های کیوتریتی می‌توانند دربردارندهٔ اطلاعات بیشتری نسبت به سیستم‌های کیوبیتی باشند. این مطلب از نظر نظریهٔ اطلاعات بسیار خوشایند است زیرا به کمک سیستم‌های کیوتریتی حجم ذخیره اطلاعات را تا حد بسیار زیادی گسترش داد. نکته دیگر اینکه بزرگ بودن بازه تغییرات آنتروپی فون نیومن نشان می‌دهد که که ظرفیت کانال‌های کیوتریتی بدون نوفه و نوفه‌ای بیشتر از کانال‌های کیوتریتی است. این مطلب از نظریهٔ شوماخر کاملاً آشکار است.



شکل(4-4)، نمودار آنتروپی فون نیومن بر حسب دوتا از دامنه‌های احتمال با این شرط که هر سه دامنه احتمال حقیقی در نظر گرفته شوند

4-6- نتیجه گیری

در این فصل با معرفی اندازه‌های جدید اطلاعات نشان دادیم که اگرچه آنتروپی شانون در فیزیک کلاسیک همتا ندارد، اما در فیزیک کوانتمی به جر آنتروپی فون نیومن می‌توان آنتروپی‌هایی را معرفی کرد که اگرچه در بعضی زمینه‌ها به اندازه آنتروپی فون نیومن مناسب نیستند ولی گاهی کارکردی مفیدتر از آنتروپی فون نیومن دارند. نتیجه دیگری که از این فصل می‌توان گرفت این است که سیستم‌های سه حالتی در عمل مناسب‌تر از سیستم‌های دو حالتی می‌باشند.

نتیجه‌گیری کلی

در این رساله آنتروپی شانون و آنتروپی فون نیومن را شرح داده و در مورد کاربرد آن‌ها در تئوری اطلاعات صحبت کردیم. نشان دادیم که اگرچه آنتروپی شانون گویای رابطه‌ای بین اطلاعات و مکانیک آماری است ولی بدون در نظر گرفتن آنتروپی شانون هم می‌توان به ارتباط بین اطلاعات و انرژی (آنتروپی) پی برد. آنتروپی شانون در واقع به این رابطه رنگ و بویی ریاضی بخشید و به ساخت کامپیوترها کمک شایانی کرد. همچنین در طی این رساله پی بردیم که در حوزه اطلاعات کلاسیکی تاکنون هیچ رابطه‌ای مفیدتر از آنتروپی شانون ظاهر نشده و آنتروپی شانون هنوز هم سنگ بنای سیستم‌های محاسباتی کلاسیکی می‌باشد.

در حوزه مکانیک کوانتمی بهترین رابطه‌ای که برای کمّی کردن اطلاعات به کار می‌رود آنتروپی فون نیومن می‌باشد. این رابطه همتای آنتروپی شانون در مکانیک کوانتمی است و گسترش آن در تعیین کانال‌های کوانتمی، آن را به رابطه‌ای بسیار مفید تبدیل کرده به‌گونه‌ای که با وجود ایرادهایی که گاهی بر آن گرفته می‌شود همچنان اهمیت خود را از دست نداده است.

در فصل آخر در مورد سیستم‌های سه حالتی در نظریهٔ اطلاعات صحبت کرده و نشان دادیم که این سیستم‌ها در عمل بسیار مفیدتر از سیستم‌های دو حالتی می‌باشند. همچنین نشان دادیم که هر دو آنتروپی شانون و فون نیومن برای سیستم‌های سه حالتی قابل استفاده و مفید هستند.

در آخر به‌عنوان نتیجه‌ای مهم به این نکته اشاره می‌کنیم که تمام اندازه‌هایی که برای کمّی کردن اطلاعات به کار می‌روند باید به‌گونه‌ای باشند که استفاده از آن‌ها در سیستم‌های سه حالتی بلامانع بوده و بهینه بودن این سیستم‌ها را آشکارا نشان دهند.

پیوست

پیوست الف

اندازه‌گیری‌های کوانتمی

هر گاه بخواهیم در مکانیک کوانتمی مشاهده‌پذیری را اندازه‌بگیریم، مکانیک کوانتمی می‌تواند پیشگویی کند که چه نتایجی ممکن است به دست آید و احتمال رخ دادن هر نتیجه را مشخص کند. در اندازه‌گیری‌های کوانتمی حالت سیستم بعد از اندازه‌گیری با حالت اولیه متفاوت است مگر اینکه حالت اولیه ویژه‌حالت ابزار اندازه‌گیری باشد. در حالی که در فیزیک کلاسیک اندازه‌گیری اثری روی سیستم نداشته و باعث تغییر آن نمی‌شود. بنابراین می‌توان گفت که اندازه‌گیری سیستم کوانتمی را به روشی معکوس ناپذیر تغییر می‌دهد.

تشخیص حالت‌های کوانتمی و اندازه‌گیری

اندازه‌گیری به شکلی اساسی سیستم را مختل می‌کند. برای مثال کیوبیتی کلی را در نظر بگیرید:

$$|y\rangle = a|0\rangle + b|1\rangle \qquad \text{(الف-1)}$$

هر گاه اندازه‌گیری روی این سیستم انجام شود، کیوبیت به حالت $|y\rangle \to |0\rangle$ یا $|y\rangle \to |1\rangle$ می‌رود. بعد از اندازه‌گیری حالت اولیه از بین می‌رود یا اصطلاحاً گم می‌شود. برای تعیین a و b لازم است روی کپی‌های زیادی از $|y\rangle$ تعداد زیادی اندازه‌گیری انجام دهیم. اندازه‌گیری سیستم کوانتمی باعث برهمکنش یا جفت شدگی سیستم با وسیله اندازگیری می‌شود.

اندازه‌گیری‌های تصویری[75]

فهمیدن اندازه‌گیری تصویری ساده‌تر از هر نوع اندازه‌گیری است. اندازه‌گیری تصویری گاهی با عنوان اندازه‌گیری فون نیومن شناخته می‌شود زیرا ریاضی‌دانی بود که برای اولین بار این اندازه‌گیری را توصیف کرد.

اگر برای سیستمی یک سری حالت‌ها ممکن باشد با اندازه‌گیری تصویری تعیین می‌کنیم که سیستم واقعاً در چه حالتی است. برای مثال یک اتم می‌تواند در دو حالت منحصربفرد باشد، یکی حالتی با انرژی پایین‌تر یا حات پایه و دیگری با اندژی بالاتر یا حالت برانگیخته. ما از اندازه‌گیری تصویری استفاده می‌کنیم که بفهمیم سیستم واقعاً در کدام یک از این حالت‌ها است. یک عملگر تصویری هرمیتی لست و با مجذور خودش مساوی می‌باشد.

اگر حاصل‌ضرب عملگرهای تصویری P_1 و P_2 صفر باشد این عملگرها متعمد نامیده می‌شوند. حاصل‌جمع یک سری کامل از عملگرهای تصویری متعامد برابر 1 می‌باشد.

$$\sum_i P_i = 1 \qquad \text{(الف-2)}$$

هر سری کامل از عملگرهای متعامد اندازه‌گیری را مشخص می‌کنند که می‌تواند درک شود. اگر یک سری عملگرهای متعامد کامل باشند حداقل یکی از نتایج اندازه‌گیری ممکن باید درست[76] باشد. این گفته بیان دیگری از حقیقتی است که جمع احتمال‌ها باید صفر باشد. تعداد عملگرهای تصویری با بعد فضای هیلبرتی که سیستم را توصیف می‌کند، مشخص می‌شود. اگر بعد فضای هیلبرت d و بعد عملگرهای تصویر m باشدف داریم $m \le d$. اگر یک سری از عملگرهای تصویر $\{P_1, P_2, ...\}$ داشته باشیم، شرط تعامد با $P_i P_j = d_{ij} P_i$ داده می‌شود.

[75] Projective Measurements
[76] True

اگر سیستمی با بعد n و حالت $|y\rangle$ داشته باشیم و یک سری از عملگرهای تصویر متعامد $\{P_1,P_2,...,P_n\}$در نظر بگیریم، احتمال پیدا کردن خروجی موقعی که اندازه‌گیری انجام شود به صورت زیر داده می‌شود:

$$p_i=\left|P_i|y\rangle\right|^2=\left(P_i|y\rangle\right)^\dagger\left(P_i|y\rangle\right)=\langle y|P_i^2|y\rangle \qquad \text{(الف-3)}$$

سیستمی که ابتدا در حالت $|y\rangle$ باشد پس از اعمال اندازه‌گیری تصویری در حالت:

$$|y'\rangle=\frac{P_i|y\rangle}{\sqrt{\langle y|P_i|y\rangle}} \qquad \text{(الف-4)}$$

مخرج کسر نسان می‌دهد که $|y\rangle$ بهنجار است.

اگر ویژه‌پایه‌های عملگر A با که با ویژه‌مقادیر a_i تناظر دارد با $|u_i\rangle$ نمایش داده شوند، تجزیه طیفی A اجازه می‌دهد که آن را به شکل $A=\sum_{i=1}^n a_i|u_i\rangle\langle u_i|=\sum_{i=1}^n a_iP_i$ بازنویسی کنیم. که عملگر تصویر منطبق بر خروجی a_i با $P_i=|u_i\rangle\langle u_i|$ داده می‌شود.

اندازه‌گیری تعمیم یافته[77]

اندازه‌گیری می‌تواند به روشی کلی‌تر توصیف شود. عموماً یک عملگر اندازه‌گیری با M_m نمایش داده می‌شود. که m اندیسی است که نتیجه یک اندازه‌گیری ممکن را نشان می‌دهد. اگر حالت $|y\rangle$ داده شده باشد احتمال پیدا کردن نتیجه m به صورت زیر داده می‌شود:

$$p_m=\langle y|M_m^\dagger M_m|y\rangle \qquad \text{(الف-5)}$$

اندازه‌گیری تصویری زیر مجموعه اندازه‌گیری تعمیم یافتته است. بعد از یک اندازه‌گیری تصویری حالت سیستم با

$$|y'\rangle=\frac{M_m|y\rangle}{\sqrt{\langle y|M_m^\dagger M_m|y\rangle}} \qquad \text{(الف-6)}$$

[77] Generalized Measurements

عملگرهای اندازه‌گیری رابطه کامل بودن را برآورده می‌کند. این رابطه برای اندازه‌گیری تعمیم یافته به صورت زیر نوشته می‌شود:

$$\sum_m M_m^\dagger M_m = I \qquad \text{(الف-7)}$$

اگر سیستم به وسیله عملگر چگالی r توصیف شود احتمال به دست آوردن نتیجه m بعد از اندازه-گیری به صورت زیر است:

$$p_m = \mathrm{Tr}\left(M_m^\dagger M_m r \right) \qquad \text{(الف-8)}$$

حالت سیتم بعداز به دست آورن نتیجه m به شکل

$$r' = \frac{M_m r M_m^\dagger}{Tr\left(M_m^\dagger M_m r \right)} \qquad \text{(الف-9)}$$

اندازه‌گیری‌های POVM [78]

اندازه‌گیری‌های POVM کلی‌تر از اندازه‌گیری‌های تصویری هستند. یک POVM یک سری از عملگرهای مثبت را شامل می‌شود که عموماً با E_m نمایش داده می‌شوند. احتمال به دست آوردن نتیجه اندازه‌گیری m در این مورد با

$$p_m = \langle y | E_m | y \rangle \qquad \text{(الف-10)}$$

به دست می‌آید. هر گاه سیستم در حالت آمیخته r باشد، احتمال به دست اوردن نتیجه نتیجه m با $\mathrm{Tr}\left(E_m r \right)$ به دست می‌آید. به علاوه سرس عملگرهای مثبت E_m شرط $\sum_m E_m = I$ را برآورده می‌کنند.

[78] Positive Operator- valued Measures

پیوست ب

تجزیه طیفی

عملگر نرمال A به بعضی فضاهای برداری تعلق دارد و با توجه به بعضی پایه‌های فضای برداری قطری می‌شود. این نتیجه به عنوان تئوری تجزیه طیفی شناخته می‌شود. فرض کنید عملگر A تئوری تجزیه طیفی را برای پایه‌های $\left| u_i \right\rangle$ برآورده کند، پس می‌توان عملگر A را به شکل زیر نوشت:

$$A = \sum_{i=1}^{n} a_i \left| u_i \right\rangle \left\langle u_i \right| \qquad \text{(ب-1)}$$

که a_i ویژه‌مقادیر عملگر می‌باشند.

فهرست منابع

[1] John Avery, (1933), *Information theory and evolution.*

[2] R. K. Pathria, (1996), *Statistical Mechanics*, Second Edition.

[3] Harvey S Leff, Andrew F Rex (2003), *Maxwell's Demon 2 Entropy, Classical and Quantum Information, Computing*, IOP, Bristol and Philadelphia

[4] C. S. Wallace, (2004), *Statistical and Inductive Inference by Minimum Message Length*, springer.

[5] Michael A. Nielsen And Issac L. chuang, (2000), *Quantum Computation and Quantum information.*

[6] C.E. Shannon, Bell Syst. Tech. J. 27, 379 (1948)

[7] Sean P. Meyn, Richard L. Tweedie, (1993), *Markov Chains and Stochastic Stability*, Springer-Verlag, London.

[8] C. Brukner and A. Zeilinger, *arXiv: quant-ph/0006087.*

[9] John Preskill (1998), *Lecture Notes for Physics: Quantum Information and Computation*, California Institute of Technology. http://www.theory.caltech.edu/people/preskill/ph229/

[10] Māris Ozols And Laura Mančinska, *generalized Bloch Vector and the Eigenvalues of a Density Matrix.*

[11] C. Brukner, *Information Individual Quantum systems*, PHD Thesis (Vienna 1999).

[12] E. T. Jaynes, *Probability Theory: The Logic of Science*, in quant-ph/0006087 By C. Burner and A. Zeilinger.

[13] B. V. Gnedenko, *The Theory of Probability* (Mir Publishers, Moscow, 1976).

[14] J. Uffink, *Measures of Uncertainty and the Uncertainty Principle*, PhD thesis (Utrecht, 1990).

[15] Cáslav brukner And Anton Zeilinger, *Operationally Ivariant Quantum Information*, Phys. Rev. Lett (1999)

[16] J. J. Sakurai, *Modern Quantum Mechanic*

[17] I. Ivanovic, J. Phys. A 14, 3241 (1981).

[18] F. Haake (springer, 2000), *Quantum signature of Chaos*

[19] Arvnid, Mallesh, K.S., Mukunda, N.: *A generalized Pancharatnam geometric phase formula for three–level quantum systems*. J. Phys. A: Math. Gen. 30 (1997) 2417–2431; quant-ph/9605042

[20] Klimov, A.B., S´anchez-Soto L.L., de Guise, H., Bj¨ork, G.: *Quantum phases of a qutrit*. J. Phys. A: Math. Gen. 37 (2004) 4097–4106

[21] A. Wehrl, Rev. Mod.Phys, 50, 221,(1978)

[22] Benjamin Schumacher, *Quantum Coding*,Physical Review A, volume 51,(1993)

[23] Alexander Stotland, Andrei A. Pomeransky, Eitan Bachmat and Doron Cohen, *Information Entropy of Quantum mechanical states*, arXive: quant-ph/0401021

[24] E. T. Jaynes, Phys Rev 108, 171 (1957) In *Information Entropy of Quantum mechanical states*,By Alexander Stotland, Andrei A. Pomeransky, Eitan Bachmat and Doron Cohen

[25] R. G. Newton, Am. J. Phys. 72, 348 (2004).

[26] E. T. Jaynes, Information Theory in *Statistical Physics*, Brandeis Summer Institute (W.A. Benjamin inc, NewYork, 1962).

[27] For $n = 2k$ there are $2k + 1$ mutually complementary observables. See W. K. Wootters and B. D. Fields, Ann. of Phys. **191**, 363 (1989).

[28] Benjamin Schumacher1 and Michael D. Westmoreland, Phys Rev A, **56**, **1**, (1997).

[29] L. B. Levitin, *Information, Complexity, and Control in Quantum Physics*, edited by A. Blaquie`re, S. Diner, and G. Lochak ~Springer, Vienna, 1987, pp. 111–115.

[30] A. S. Kholevo, Probl. Peredachi Inf. **9**, 177 ~1973.

[31] C. A. Fuchs and C. M. Caves, Phys. Rev. Lett. **73**, 3047, 1994.

[32] P. Hausladen, R. Josza, B. Schumacher, M. Westmoreland, and W. K. Wootters, Phys. Rev. A **54**, 1869 ~1996.

[33] A. S. Kholevo, IEEE Trans. Inf. Theory ~to be published.

[34] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, Wiley, New York, 1991.

[35] B. Schumacher, Phys. Rev. A **51**, 2738 , 1995.

[36] Marcos Martín Fernández, *Fundamentals on Estimation Theory*, May 4, 2004, http://lmi.bwh.harvard.edu/papers/pdfs/2004/martin-fernandezCOURSE04b.pdf

Bol'shev, L.N. (2001), "Statistical Estimator", in Hazewinkel, Michiel, *Encyclopaedia of Mathematics*, Kluwer Academic Publishers, ISBN 978-1556080104

[37] Casella, George; Berger, Roger L. (2002). *Statistical Inference* (2ed.). Pacific Grove: Duxbury. Ballanda, Kevin P.; MacGillivray, H. L. (1988). "Kurtosis: A Critical Review". *The American Statistician* **42** (2): 111-119. W. Feller, *An introduction to probability theory and its applications* , **2** , Wiley (1971) pp. Chapt. 1

[38] David McMahon (2008), *Quantum Computing Explained*, John Wiley & Sons, Hoboken, New Jersey

Abstract:

The progressive development of science and human needs makes an obligation to expedite technology. Computers and calculating machines are everlasting human needs involved in his mind, so that engineers and scientists often try to offer more advanced ones. The applications of quantum mechanics in industry make scientists to make advantage of particular quantum properties in development of computers. In this manner computers can be optimized in a way which is incomparable with our classical computers.

The above mentioned needs cause the increasing research in application of quantum mechanics for computers. Here we try to clear some aspects of quantum computers from physics point of view.

The information theory plays the fundamental rule in both classical and quantum computers. In chapter one of this thesis, the connection of information theory with physical concepts especially work and energy is discussed. Chapter two is devoted to classical information theory, especially Shannon information and its important features in classical computers. In chapter three quantum information theory and von Neumann entropy as the most effective relation are discussed. It also includes some applications of von Neumann entropy. In chapter four some suggestions are offered to develop the quantification of information in quantum computers. Furthermore some supposed relations have been compared with von Neumann entropy. At the last sections of chapter four, the development and obvious better performance three –states systems in compared with two –states ones and also the profits of three states usages are discussed.

Key Words: “statistical entropy, Shannon entropy, von Neumann entropy, classical information, quantum information, bit, trit, qubit, qutrit.”

Shahrood University of Technology

Faculty of Physics

Shannon Information and Quantum Information

Halime Vahid

Supervisor(s):

Dr. Hossein Movahedyan

Date: February 2010