Universidade de Federal de Goiás
Instituto de Física

Thiago Mureebe Carrijo

Correlações quânticas e generalização da entropia de von Neumann

Goiânia

2 de Abril de 2012

sistema de bibliotecas ufg

UFG

TERMO DE CIÊNCIA E DE AUTORIZAÇÃO PARA DISPONIBILIZAR AS TESES E DISSERTAÇÕES ELETRÔNICAS NA BIBLIOTECA DIGITAL DA UFG

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFG), regulamentada pela Resolução CEPEC nº 832/2007, sem ressarcimento dos direitos autorais, de acordo com a Lei nº 9610/98, o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou *download*, a título de divulgação da produção científica brasileira, a partir desta data.

1. Identificação do material bibliográfico: **[X] Dissertação** **[ ] Tese**

2. Identificação da Tese ou Dissertação

Nome completo do autor: Thiago Mureebe Carrijo

Título do trabalho: Correlações quânticas e generalização da entropia de von Neumann

3. Informações de acesso ao documento:

Concorda com a liberação total do documento [X] SIM [ ] NÃO[1]

Havendo concordância com a disponibilização eletrônica, torna-se imprescindível o envio do(s) arquivo(s) em formato digital PDF da tese ou dissertação.

Thiago M. Carrijo
Assinatura do (a) autor (a)

Data: 09/09/2016

[1] Neste caso o documento será embargado por até um ano a partir da data de defesa. A extensão deste prazo suscita justificativa junto à coordenação do curso. Os dados do documento não serão disponibilizados durante o período de embargo.

Thiago Mureebe Carrijo

Correlações Quânticas e uma Generalização da Entropia de von Neumann

Dissertação submetida ao Instituto de Física da Universidade de Federal de Goiás destinada à defesa mestrado.

Orientador: *Prof. Dr. Ardiley Torres Avelar*
Co-orientador: *Prof. Dr. Norton Gomes de Almeida*

Goiânia

2 de Abril de 2012

Ficha de identificação da obra elaborada pelo autor, através do
Programa de Geração Automática do Sistema de Bibliotecas da UFG.

Mureebe Carrijo, Thiago
Correlações quânticas e generalização da entropia de von Neumann [manuscrito] / Thiago Mureebe Carrijo. - 2012.
70 f.: il.

Orientador: Prof. Dr. Ardiley Torres Avelar; co-orientador Dr. Norton Gomes de Almeida.
Dissertação (Mestrado) - Universidade Federal de Goiás, Instituto de Física (IF), Programa de Pós-Graduação em Física, Goiânia, 2012.
Bibliografia.
Inclui lista de figuras.

1. Correlações quânticas. 2. Entropia de von Neumann. I. Torres Avelar, Ardiley , orient. II. Título.

CDU 530.1/145

Universidade Federal de Goiás
Instituto de Física
Programa de Pós-Graduação em Física

Ata **Nº 130** de defesa de dissertação de **Thiago Mureebe Carrijo** para obtenção do título de Mestre em Física.

Aos 10 dias do mês de abril de 2012, às 14h00min, no Mini-Auditório do Instituto de Física, reuniu-se a Banca Examinadora indicada pela Coordenadoria do Programa de Pós-Graduação, aprovada pelo Conselho Diretor e designada pela Diretoria do Instituto de Física da Universidade Federal de Goiás, composta pelo **Prof. Dr. Ardiley Torres Avelar** (orientador e presidente da Banca – IF/UFG), **Prof. Dr. Norton Gomes de Almeida** (co-orientador – IF/UFG), **Prof. Dr. Roberto Menezes Serra** (CCNH/UFABC) e **Prof. Dr. Wesley Bueno Cardoso** (IF/UFG), para julgar a dissertação de mestrado de **Thiago Mureebe Carrijo**, intitulada: "Correlações quânticas e generalização da entropia de Von Neumann". O Presidente abriu os trabalhos agradecendo a presença dos membros da Banca Examinadora e concedeu a palavra a **Thiago Mureebe Carrijo,** que expôs detalhadamente seu trabalho. Em seguida, os membros da Banca fizeram suas considerações e procederam à argüição do candidato. Concluída esta etapa, a Banca, em sessão fechada, deu prosseguimento ao julgamento do trabalho, atribuindo os seguintes conceitos:

Prof. Dr. Ardiley Torres Avelar (Orientador)	APROVADO
Prof. Dr. Norton Gomes de Almeida (Co-orientador)	APROVADO
Prof. Dr. Roberto Menezes Serra (CCNH/UFABC)	APROVADO
Prof. Dr. Wesley Bueno Cardoso (IF/UFG)	Wesley Bueno Cardoso (APROVADO)

Novamente em sessão aberta, o presidente da Banca anunciou o resultado final do julgamento, declarando o candidato Thiago Mureebe Carrijo APROVADO pela Banca Examinadora. Nada mais havendo a tratar, a sessão foi encerrada e lavrou-se a presente ata que segue assinada pelos membros da Banca Examinadora.

Goiânia, 10 de abril de 2012.

Ardiley Torres Avelar
Prof. Dr. Ardiley Torres Avelar (Orientador)

Prof. Dr. Norton Gomes de Almeida (Co-orientador)

Prof. Dr. Roberto Menezes Serra (CCNH/UFABC)

Wesley Bueno Cardoso
Prof. Dr. Wesley Bueno Cardoso (IF/UFG)

Agradecimentos

- Aos professores Dr. Ardiley T. Avelar e Dr. Norton G. de Almeida pela orientação.
- À Gisana Cristina, à minha mãe e ao meu pai.

Este trabalho foi financiado pela Capes.

Conteúdo

Lista de Figuras

Resumo

Este trabalho tem por objetivo tratar o conceito de correlação, tanto clássica quanto quântica, abordando o que já existe sobre o assunto e fornecendo uma nova perspectiva sobre o tema e, além disso, propor uma nova maneira de se calcular a quantidade de informação de um sistema quântico. Para isso, será discutido o paradoxo de EPR, o teorema de Bell, o conceito de emaranhamento, as entropias de Shannon e von Neumann, a teoria de códigos e a medida de correlação quântica conhecida como discórdia quântica. A partir desses temas, será proposta duas generalizações: da entropria de von Neumann e do conceito de correlação quântica. Alguns aspectos e propriedades são discutidos sobre essas novas definições, porém ainda há muito o que ser investigado sobre suas implicações.

Abstract

This work has the goal to treat the concept of correlation on both classical and quantum, addressing what already exists about the matter and giving a new perspective about it. Beyond this, we propose a new manner to calculate the amount of information of a quantum system. For this, the EPR paradox, Bell theorem, entanglement concept, shannon and von Neumann entropies, code theory and the quantum correlation measure so-called quantum discord will be discussed. From those subjects, two generalizations will be proposed: of the von Neumann entropy and the quantum correlation concept. Some features e properties will be discussed about those new definitions, however there is still much to be investigated about its implications.

Introdução

A Teoria de Informação Clássica, originalmente destinada à solução de problemas de engenharia ligados à telecomunicação via rádio e telégrafo [1, 2], obteve sua base formal matemática com o trabalho de Shannon [3], o qual definiu uma medida para a quantidade de informação de uma fonte de variáveis aleatórias, conhecida como entropia de Shannon, a qual tem uma aplicação direta em teoria de códigos. Essa medida tem a mesma forma funcional da entropia de Gibbs. Portanto, a generalização da entropia termodinâmica clássica para sistemas quânticos obtida por von Neumann [4] é também a generalização da entropia de Shannon. A entropia de von Neumann é a base da Teoria de Informação Quântica e tem aplicações em áreas como a compressão de dados [5–7] e a criptografia quântica [8, 9].

Observou-se que alguns tipos de estados quânticos, conhecidos como estados emaranhados [10], que descrevem sistemas com correlações, entre seus subsistemas, mais fortes do seria possível classicamente [11], têm aplicações interessantes em teoria de informação e computação quântica, como a correção quântica de erro [12–14], o teletransporte quântico [15] e a codificação superdensa [16]. Mas há estados que, embora não estejam emaranhados, tornam possível a realização de tarefas computacionais de forma mais eficiente que do seu análogo clássico por meio da computação quântica de estados mistos [17,18]. O que os caracteriza são as correlações puramente quânticas, as quais são corretamente medidas pela quantidade conhecida como discórdia quântica, definida por Zurek [19]. Em complementação, Henderson e Vedral [20] propuseram uma medida das correlações puramente clássicas. As correlações totais são a soma dessas duas. Shabani e colaboradores [21] mostraram que o estado de um sistema incluindo sua vizinhança deve ter discórdia nula para que a evolução do sistema seja um mapa completamente positivo, o que é um resultado de interesse em computação quântica. Isso significa que eles encontraram a classe de estados na qual qualquer evolução quântica é um mapa completamente positivo.

A discórdia também é capaz de captar assinaturas de transições quânticas de fase [22–24] em sistemas de muitos corpos. Com respeito à "quantidade"de estados com discórdia não-nula ou, mais precisamente, à cardinalidade do conjunto de estados com essa discórdia, ela é maior do que a cardinalidade do conjunto de estados que a anulam [25], ou seja, estados com discórdia nula são muito raros.

Capítulo 1

Conceitos Básicos

1.1 Teoria da Probabilidade

O objetivo dessa seção é expor brevemente as principais definições da teoria de probabilidades, as quais serão posteriormente no texto. Os principais conceitos são de σ-álgebra, medida, variável aleatória, medida de probabilidade, distribuição de probabilidade, densidade de probabilidade e valor esperado.

Definição 1.1.1 (σ-álgebra)**.** *Seja Ω um conjunto não vazio e 2^Ω o conjunto das partes de Ω. A coleção $\mathcal{F} \subseteq 2^\Omega$ é uma σ-álgebra se satisfaz as seguintes propriedades:*

1. $\Omega \in \mathcal{F}$;
2. $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$
3. $A, B \in \mathcal{F} \Rightarrow A \cup B \in \mathcal{F}$
4. $A_n \in \mathcal{F}, n \geq 1 \Rightarrow \bigcup_{n\geq 1} A_n \in \mathcal{F}$

Definição 1.1.2 (σ-álgebra de Borel)**.** *A σ-álgebra de Borel sobre um espaço topológico $\mathbb{T}$ é definida como a σ-álgebra gerada pelos abertos de $\mathbb{T}$.*

Definição 1.1.3 (Medida)**.** *Sejam Ω um conjunto não vazio e $\mathcal{F}$ uma σ-álgebra sobre Ω, a função $\mu : \mathcal{F} \to \overline{\mathbb{R}}_+$, em que $\overline{\mathbb{R}} \equiv \mathbb{R} \cup \{\infty\}$, com as propriedades:*

1. $\mu(\emptyset) \equiv 0$;

2. seja $A_n \in \mathcal{F}, n \geq 1$, uma coleção de conjuntos disjuntos tal que $\bigcup_{n\geq 1} A_n \in \mathcal{F}$,

$$\mu\left(\bigcup_{n\geq 1} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n).$$

é denominada medida.

Definição 1.1.4 (Espaço de Medida)**.** *Sejam Ω um conjunto não vazio e $\mathcal{F}$ uma σ-álgebra sobre Ω, o par $(\Omega, \mathcal{F})$ é denominado espaço mensurável. Se μ é uma medida em $(\Omega, \mathcal{F})$, então $(\Omega, \mathcal{F}, \mu)$ é um espaço de medida.*

Definição 1.1.5 (Função Mensurável)**.** *Sejam $(\Omega, \mathcal{F})$ e $(\Omega', \mathcal{F}')$ espaços mensuráveis, a função $f : \Omega \to \Omega'$ é denominada $\langle \mathcal{F}, \mathcal{F}' \rangle$-mensurável se, para cada $B \in \mathcal{F}'$, $f^{-1}(B) \in \mathcal{F}$*

Definição 1.1.6 (Medida de Probabilidade)**.** *A probabilidade P definida como uma medida sobre o espaço mensurável $(\Omega, \mathcal{F})$ em que $P(\Omega) = 1$. Um espaço de probabilidade é definido como um espaço de medida em que a medida é uma probabilidade.*

Definição 1.1.7 (Variável Aleatória)**.** *Seja $(\Omega, \mathcal{F}, P)$ um espaço de probabilidade $(\Omega', \mathcal{F}')$ um espaço mensurável. A função $X : \Omega \to \Omega'$ é uma variável aleatória se é $\langle \mathcal{F}, \mathcal{F}' \rangle$-mensurável.*

Definição 1.1.8 (Distribuição de Probabilidade)**.** *Seja X uma variável aleatória real em $(\Omega, \mathcal{F}, P)$. Define-se P_X, denominada distribuição de probabilidade de X, por*

$$P_X \equiv P(X^{-1}(A)), \quad \forall A \in \mathcal{B}(\mathbb{R})$$

Definição 1.1.9 (Valor Esperado)**.** *Seja X uma variável aleatória real em $(\Omega, \mathcal{F}, P)$, o valor esperado de X, $E(X)$ é definido como*

$$E(X) \equiv \int_{\Omega} X dP.$$

Definição 1.1.10 (Função Densidade de Probabilidade)**.** *Seja X uma variável aleatória real em $(\Omega, \mathcal{F})$. Sejam μ uma medida qualquer e P_X uma distribuição de probabilidade de X sobre $(\Omega, \mathcal{F})$. A função $\langle \mathcal{F}, \mathcal{B}(\mathbb{R}) \rangle$-mensurável $f_X : \Omega \to \mathbb{R}$, também escrita como $\frac{dP_X}{d\mu}$, é definida pela equação*

$$\int_{X^{-1}(A)} dP = \int_A f_X \, d\mu, \quad \forall A \in \mathcal{F}$$

Proposição 1.1.1. *Seja $\alpha : \mathbb{R} \to \mathbb{R}$ uma função contínua e $X : \Omega \to \mathbb{R}$ uma variável aleatória real em $(\Omega, \mathcal{F})$, a composição $\beta \equiv \alpha \otimes X$ também é uma variável aleatória real em $(\Omega, \mathcal{F})$.*

1.2 Medida

Na representação de Schrödinger, a evolução do estado de um sistema quântico A isolado se dá por meio da aplicação de um operador de evolução, que é unitário. Se o sistema A é aberto, a evolução pode não ser unitária. Porém, se o espaço de Hilbert do sistema é H_A e do ambiente S à sua volta é H_S, o sistema mais o ambiente com espaço de estados dado por $H_A \otimes H_S$ evolui unitariamente, pois está isolado. Se interpretarmos esse ambiente como sendo um aparato de medida, o estado reduzido do sistema A, após a interação, é o estado de A após a medida. Se estamos interessados apenas em A, é possível descrever essa operação quântica sem fazer referência a S. Há, entretanto, duas possíveis medidas: seletivas ou não. Medidas seletivas são aquelas em que há um resultado de saída concreto. Quando o aparato de medida interage, mas não fornece o resultado, não há seleção. Nesse caso, o estado de A é uma combinação convexa dos estados de saída, os quais podem ser calculados através do postulado da medida:

Postulado 1.2.1. *Seja $j \in J$, em que J é um conjunto de índices e, para cada j, $V_j \colon \mathcal{H} \to \mathcal{H}$ é um operador linear, sendo $\mathcal{H}$ um espaço de Hilbert, tal que $\sum_j V_j^\dagger V_j = I$. O conjunto $M \equiv \{V_j, j \in J\}$ é um conjunto de medidas com valores em J. Se a medida é realizada no estado ρ, a saída j ocorre com probabilidade $p_j = \mathrm{tr}(V_j \rho V_j^\dagger)$ e o estado do sistema após a medida seletiva é:*

$$\rho_j = \frac{V_j \rho V_j^\dagger}{\mathrm{tr}(V_j \rho V_j^\dagger)} \tag{1.1}$$

Quando a medida é não-seletiva, o estado final é $\langle \rho \rangle_M$, que é a média da variável aleatória $X \equiv \sum_{j \in J} \rho_j I_{\{\rho_j\}}$, em que $I_{\{\rho_j\}} : \Omega \to \{0, 1\}$, sendo $\Omega \equiv \{\rho_j, j \in J\}$ e $I_{\{\rho_j\}}(\rho_i) \equiv \delta_{i,j}$, no espaço de medida $(\Omega, 2^\Omega, P)$ com $P(\{\rho_j\}) = p_j$.

É comum trocar o uso dos operadores V_j pelos operadores $E_j \equiv V_j^\dagger V_j$. Quando a medida

é caracterizada dessa forma, ela recebe o nome de POVM, que significa medidas avaliadas por operadores positivos, uma vez que $E_j \geq 0$. Um caso particular são as medidas projetivas ortogonais $\{P_j | P_j^\dagger = P_j, P_j P_i = \delta_{i,j} P_j, \sum_j P_j = I\}$, as quais representam os observáveis físicos, uma vez que todo observável pode ser escrito como uma combinação convexa de projetores.

Um fato surpreendente é que toda medida POVM é equivalente a uma medida projetiva [30]. Isso é provado pelo teorema de Naimark, o qual estabelece que para todo conjunto $\{E_j | j \in J \subseteq \mathbb{N}\}$ de medidas POVM, as quais atuam em $\mathcal{H}_a$, existe um conjunto de medidas projetivas $\{P_j\}$, as quais atuam em $\mathcal{H}_a \otimes \mathcal{H}_b$, tais que $E_j = \mathrm{tr}_b\left[(I_a \otimes \rho_b) P_j\right]$ para algum $\rho_b = \mathrm{tr}_a \rho_{a,b}$ que atua em $\mathcal{H}_b$ e $\mathrm{tr}_a(E_j \rho_a) = \mathrm{tr}_{a,b}(P_j \rho_{a,b})$ qualquer que seja ρ_a. Esse resultado significa que sempre é possível obter um sistema auxiliar, que não precisa interagir com o sistema principal, de modo os resultados obtidos por medidas projetivas no sistema global sejam os mesmos que seriam obtidos medindo-se o sistema principal, qualquer que seja o POVM.

1.3 Emaranhamento

Nos primórdios da construção da mecânica quântica, questinou-se se seu formalismo fornecia uma descrição completa e precisa dos fenômenos de seu domínio. Em 1935, Einstein, Podolsky e Rosen (EPR) publicaram um artigo [26] em que mostravam, admitindo certas condições plausíveis, que a mecânica quântica não fornecia toda a informação física que se poderia prever de um sistema. Para provar esse ponto de vista, admitiram as seguintes hipóteses físicas:

1. Se, de modo algum houver distúrbio em um sistema, pode-se predizer com certeza (probabilidade igual a um) o valor de uma quantidade física, então existe um elemento da realidade física correspondendo a essa quantidade física (hipótese do realismo);

2. Não existe ação a distância na natureza (hipótese da localidade).

Em mecânica quântica, a primeira condição significa que se um sistema está no autoestado $|a^k\rangle$ de um observável A, com autovalor a_k, então pode-se predizer com certeza que o valor

da grandeza A é a_k, o que significa que A corresponde a um elemento da realidade física. Um outro observável B tal que $[A, B] \neq 0$ não possui $|a^k\rangle$ como seu autoestado, qualquer que seja k, logo não se pode predizer com certeza o resultado de uma medida B, o que significa que B não corresponde a um elemento da realidade física enquanto A corresponder e vice versa. Porém, considerando dois sistemas que estão em um estado conjunto do tipo

$$|\psi\rangle = a_k|a'^k\rangle|a^k\rangle = b_k|b'^k\rangle|b^k\rangle \tag{1.2}$$

em que A e B atuam apenas no segundo sistema. Sejam A' e B' observáveis que atuam no primeiro sistema com autovetores $|a'^k\rangle$ e $|b'^k\rangle$ respectivamente, caso se faça uma separação que impeça a interação entre os sistemas de modo que se realize uma medida de A' e se obtenha $|a'^r\rangle$ como saída, então o estado do segundo sistema será $|a^r\rangle$. Porém, medindo-se B', tendo-se $|b'^s\rangle$ de saída, o estado do outro sistema será $|b^s\rangle$. Como não há interação, os dois estados devem caracterizar simultaneamente o segundo sistema. No artigo de EPR é provado que existe um sistema de duas partículas em que se pode fazer as medidas consideradas, em que os estados obtidos para o segundo sistema são autoestados de observáveis que não comutam. Dessa forma, segundo a hipótese do realismo, A e B possuem realidade simultânea, pois seus valores podem ser preditos, porém a mecânica quântica não fornece meios determinar esses valores com certeza devido à não comutatividade. Logo, conclui-se que a mecânica quântica não é completa. Por completa, segundo a definição de EPR, entende-se uma teoria na qual cada elemento da realidade deve possuir seu correspondente na teoria. Assim, se o valor de uma grandeza pode ser predito com certeza, tal predição precisa estar inclusa na teoria. Isso sugere que podem haver parâmetros ocultos que determinam os valores das grandezas. Porém, em 1964, Bell [11] mostrou que se tais variáveis existem, e se vale a hipótese da localidade, não se pode reproduzir os resultados da mecânica quântica, ou seja, uma teoria local de variáveis ocultas não descreve os fenômenos quânticos.

Suponha que um sistema composto por dois subsistemas no estado singleto estejam separados de modo a não haver interação entre eles. Sejam $\mathcal{A}(\vec{a})$ e $\mathcal{B}(\vec{b})$ grandezas físicas que,

uma vez medidas, retornam como valores 1 ou -1. No formalismo de operadores, sejam $A(\vec{a})$ e $B(\vec{b})$ os observáveis que representam essas grandezas. Sejam $A(\vec{a})|a_A\rangle = |a_A\rangle$, $A(\vec{a})|-a_A\rangle = -|-a_A\rangle$, $B(\vec{a})|a_B\rangle = |a_B\rangle$ e $B(\vec{a})|-a_B\rangle = -|-a_B\rangle$, considerando o estado do sistema dado por

$$|\psi\rangle = \frac{|a_A\rangle|-a_B\rangle - |-a_A\rangle|a_B\rangle}{\sqrt{2}}, \tag{1.3}$$

uma medida de $A(\vec{a})$ sempre retorna o valor oposto de $B(\vec{a})$ e vice-versa. Sejam $A(\vec{a})$ e $B(\vec{b})$ operadores tais que

$$\langle\psi|A(\vec{a})B(\vec{b})|\psi\rangle = -\vec{a}\cdot\vec{b}, \tag{1.4}$$

isso ocorre para operadores de spin $\vec{\sigma} \equiv (\sigma_x, \sigma_y, \sigma_z)$, tomando $A(\vec{a}) \equiv \vec{\sigma}_A\cdot\vec{a}$ e $B(\vec{b}) \equiv \vec{\sigma}_B\cdot\vec{b}$. Agora suponha que as grandezas $\mathcal{A}(\vec{a})$ e $\mathcal{B}(\vec{b})$ dependam de um parâmetro X. Essa hipótese significa que existe um parâmetro que determina, a priori, os valores das grandezas. Pela suposição da localidade, os valores de $\mathcal{A}(\vec{a})$ não podem depender da escolha de $\vec{b}$, nem $\mathcal{B}(\vec{b})$ de $\vec{a}$. Além disso, considere que X é uma variável aleatória de um espaço de probabilidade $(\Omega, \mathcal{F}, P)$, que P_X seja sua distribuição de probabilidade e que as grandezas sejam o resultado da composição de uma função contínua com X, ou seja, $\mathcal{A}(\vec{a}) = \mathcal{A}_X(\vec{a}) \circ X$ e $\mathcal{B}(\vec{b}) = \mathcal{B}_X(\vec{b}) \circ X$. Essa última restrição tem por objetivo garantir que as grandezas sejam variáveis aleatórias. Assim, o valor esperado de $\mathcal{A}_X(\vec{a})\mathcal{B}_X(\vec{b})$ é

$$\langle\mathcal{A}_X(\vec{a})\mathcal{B}_X(\vec{b})\rangle = \int_{\mathbb{R}} \mathrm{d}P_X \mathcal{A}_X(\vec{a})\mathcal{B}_X(\vec{b}) \tag{1.5}$$

Sejam $\vec{a}'$ e $\vec{b}'$ vetores tais que $|\vec{a}'\cdot\vec{a}| << 1$ e $|\vec{b}'\cdot\vec{b}| << 1$. Suponha que

$$|\langle\mathcal{A}_X(\vec{a}')\mathcal{B}_X(\vec{b}')\rangle + \vec{a}\cdot\vec{b}| \leq \epsilon \tag{1.6}$$

e suponha também que, qualquer que sejam $\vec{a}$ e $\vec{b}$,

$$|\vec{a}'\cdot\vec{b}' - \vec{a}\cdot\vec{b}| \leq \delta, \tag{1.7}$$

somando as duas equações anteriores,

$$|\langle\mathcal{A}_X(\vec{a}')\mathcal{B}_X(\vec{b}')\rangle + \vec{a}'\cdot\vec{b}'| \leq \epsilon + \delta \tag{1.8}$$

de (1.5) em (1.8), fazendo $\vec{a}' = \vec{b}'$,

$$\int_{\mathbb{R}} \mathrm{d}P_X \mathcal{A}_X(\vec{b}')\mathcal{B}_X(\vec{b}') + 1 \leq \epsilon + \delta \tag{1.9}$$

e

$$\begin{aligned} \langle \mathcal{A}_X(\vec{a}')\mathcal{B}_X(\vec{b}')\rangle - \langle \mathcal{A}_X(\vec{a}')\mathcal{B}_X(\vec{c}')\rangle &= \int_{\mathbb{R}} \mathrm{d}P_X \mathcal{A}_X(\vec{a}')\mathcal{B}_X(\vec{b}') \left[1 + \mathcal{A}_X(\vec{b}')\mathcal{B}_X(\vec{c}')\right] \\ &- \int_{\mathbb{R}} \mathrm{d}P_X \mathcal{A}_X(\vec{a}')\mathcal{B}_X(\vec{b}') \left[1 + \mathcal{A}_X(\vec{b}')\mathcal{B}_X(\vec{b}')\right] \end{aligned} \tag{1.10}$$

$$\begin{aligned} |\langle \mathcal{A}_X(\vec{a}')\mathcal{B}_X(\vec{b}')\rangle - \langle \mathcal{A}_X(\vec{a}')\mathcal{B}_X(\vec{c}')\rangle| &\leq \int_{\mathbb{R}} \mathrm{d}P_X \left[1 + \mathcal{A}_X(\vec{b}')\mathcal{B}_X(\vec{c}')\right] \\ &- \int_{\mathbb{R}} \mathrm{d}P_X \left[1 + \mathcal{A}_X(\vec{b}')\mathcal{B}_X(\vec{b}')\right]. \end{aligned} \tag{1.11}$$

De (1.9),

$$|\langle \mathcal{A}_X(\vec{a}')\mathcal{B}_X(\vec{b}')\rangle - \langle \mathcal{A}_X(\vec{a}')\mathcal{B}_X(\vec{c}')\rangle| \leq 1 + \langle \mathcal{A}_X(\vec{b}')\mathcal{B}_X(\vec{c}')\rangle + \epsilon + \delta. \tag{1.12}$$

De (1.6) na última equação,

$$\epsilon + \delta \geq \frac{|\vec{a}\cdot\vec{c} - \vec{a}\cdot\vec{b}| + \vec{b}\cdot\vec{c} - 1}{4}, \tag{1.13}$$

pode-se escolher $\vec{a}\cdot\vec{c} = \frac{\sqrt{3}}{2}$, $\vec{a}\cdot\vec{b} = \frac{\sqrt{3}}{2}$ e $\vec{b}\cdot\vec{c} = 0$, de tal modo que

$$\epsilon + \delta \geq \frac{\sqrt{3} - 1}{4}, \tag{1.14}$$

Para δ muito pequeno, ϵ não pode ser tomado arbitrariamente pequeno, o que significa que a média $\langle \mathcal{A}_X(\vec{a}')\mathcal{B}_X(\vec{b}')\rangle$ não pode ser arbitrariamente aproximada do valor quântico. Isso implica que não existe teoria local de variáveis ocultas que reproduz os resultados estatísticos da mecânica quântica.

O estado do sistema composto utilizado para se fazer a média quântica é um tipo especial denominado estado emaranhado, o qual carrega correlação entre os subsistemas que permanece ainda que eles estejam longe entre si o suficiente para não haver interação física. Esses estados são interessantes pois não há análogo clássico.

Um sistema quântico composto descrito por um operador densidade ρ e constituído de N subsistemas é separável se, e somente se, puder ser escrito na forma [27]:

$$\rho = \sum_k p_k \bigotimes_{l=1}^{N} \rho_k^{(l)}, \tag{1.15}$$

em que ρ é a matriz densidade do estado do sistema composto, $\rho_{(l)}^k$ é uma matriz densidade do subsistema l e p_k é a probabilidade de o sistema composto estar na configuração k de todas as configurações possíveis. Deve-se ter também que $0 \leq p_k \leq 1$ para todo k e $\sum_k p_k = 1$.

Por definição, o estado de um sistema é emaranhado se, e somente se, não puder ser escrito na forma da equação (1.15).

A definição acima é aplicável tanto a estados puros quanto a estados mistos. Porém, para estados puros, descritos por vetores de estado, é mais simples utilizar outra definição de separabilidade. Seja $|\psi\rangle$ um estado puro de um sistema composto por N subsistemas, ele é separável se, e somente se, puder ser escrito na forma

$$|\psi\rangle = \bigotimes_{i=1}^{N} |\psi_i\rangle, \tag{1.16}$$

em que $|\psi_i\rangle$ é um estado qualquer pertencente ao espaço de Hilbert associado ao i-ésimo subsistema. Qualquer estado puro que não puder ser escrito na forma da equação (1.16) está emaranhado.

Capítulo 2

Informação Clássica

2.1 Códigos

O que é a Teoria da Informação Clássica? De modo não muito específico, ela é uma teoria que explica como os processos de armazenamento, envio, recepção e troca de informação ocorrem. Abaixo há um esquema gráfico que ilustra como esses processos estão relacionados num contexto de comunicação entre duas partes.

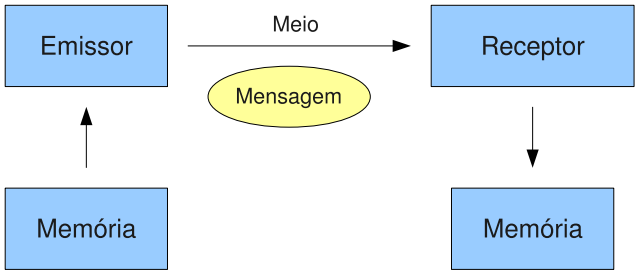


Figura 2.1: Esquema de comunicação

Na Figura (2.1), observa-se que o emissor envia uma mensagem que pode ser retirada de um banco de dados (memória) e enviada através de um meio físico ao receptor, o qual a recebe e pode ou não guarda-la em sua memória. Um exemplo é a transmissão de um texto armazenado no disco rígido de um computador pela internet, através de fibras óticas, o qual é recebido por outro computador e armazenado em sua memória. Em geral, porém, uma mensagem escrita com letras e números recebe uma codificação. No caso dos computadores atuais, usa-se o alfabeto

$\{0,1\}$. Cada caractere, então, é relacionado a uma palavra-código por meio de uma regra de codificação [1]. Por exemplo, $C_0 = \{(a,00),(b,01),(c,10),(d,11)\}$ é um código que relaciona cada letra a dois bits distintos. De modo simples, pode se definir um código $C\colon \Omega \to P_A$ como uma função com domínio no conjunto de caracteres Ω e contradomínio no conjunto de palavras-códigos P_A. Seja A um alfabeto d-ário, em que $d = |A|$, e $A \times \ldots \times A \equiv \times_{i=1}^{n} A \equiv A^n$. O conjunto P_A é definido então por $P_A \equiv \bigcup_{k=1}^{n} A^k$, em que $k \in \mathbb{N}$. Como regra, costuma-se escrever n-uplas $(a_1, a_2, \ldots, a_n)$ simplismente como $a_1 a_2 \ldots a_n$.

Uma questão que logo vem à mente é o de como criar um código. Em geral, há várias maneiras de se fazer isso. Um código simples, que se encaixa a qualquer conjunto Ω, é dado pelo algoritmo: Seja $\Omega = \{\omega_1, \ldots, \omega_k\}$, com uma indexação dos elementos realizada de forma arbitrária, e $A = \{0,1\}$. Seja $n = \log_2 k$ se $\log_2 k$ for inteiro ou $n = \text{int}\{\log_2 k\} + 1$ caso contrário. Seja $P_A \equiv A^n = \{a_1^{(n)}, \ldots, a_{2^n}^{(n)}\}$, também com indexação arbitrária. O código C_s, chamado código simples, será então definido pela relação $C_s(\omega_i) = a_i^{(n)}$. O código C_0 escrito anteriormente é um exemplo de código simples. Ocorre que nem sempre será conveniente utilizar esse algoritmo de codificação, uma vez que é interessante escolher um que minimize o tamanho das mensagens. Isso está relacionado com a minimização dos recursos físicos necessários para se enviar uma mensagem. Para abordar esse problema, é necessário calcular o tamanho de uma palavra-código. Seja $C(\Omega)$ a imagem do código $C\colon \Omega \to P_A$ e $C(\omega_i) = a_{i,1} \ldots a_{i,k}$, define-se o comprimento de palavra-código $l_C\colon C(\Omega) \to \mathbb{N}$ pela relação $l_C \circ C(\omega_i) = k$, o qual também pode ser escrito como $l_C \circ C = \sum_{i=1}^{n} l_C \circ C(\omega_i) I_{\{C(\omega_i)\}}$, em que $I_{\{C(\omega_i)\}}(C(\omega_i)) = 1$ e zero caso contrário. Seja P uma medida de probabilidade em $(\Omega, 2^\Omega)$ tal que $P(\{C(\omega_i)\}) = p(\omega_i)$. Como nem sempre os elementos de P_A têm o mesmo comprimento, define-se o comprimento médio $L(C)$ do código C pela expressão $L(C) \equiv \int_\Omega \mathrm{d}P l_C \circ C = \sum_{i=1}^{n} p(\omega_i) l_C \circ C(\omega_i)$.

Voltando ao problema da escolha de um código que minimiza os recursos, fica claro que deve-se escolher um que forneça o $L(C)$ mínimo. Se soubermos esse valor mínimo, basta, a

[1] Algumas definições apresentadas neste trabalho em relação à teoria dos códigos não são usuais. Uma boa referência para uma descrição usual é o livro de Cover e Thomas [29]

princípio, criar um código que o atinja. Esse problema está relacionado a outro bem próximo, que é o de determinar a quantidade de informação contida em um evento E_Ω. Um evento pode ser, por exemplo, um lançamento de dados ou de uma moeda ou um sorteio de loteria. Para especificá-lo, de modo geral, basta representá-lo por $E_\Omega \equiv \{(\omega, p(\{\omega\})), \omega \in \Omega\}$, em que $p(\omega)$ é a probabilidade de o elemento ω ser o resultado da realização do evento. Posteriormente, será mostrado que essa quantidade de informação é sempre menor ou igual a $L(C)$.

2.2 Entropia de Shannon

O que se quer dizer com a quantidade de informação contida em um evento? Suponha que o evento seja jogar uma moeda e observar se foi cara ou coroa que caiu para cima. Que informação que se tem sobre o resultado antes de a moeda cair? Se ela não for viciada, a única coisa que se pode saber é que as chances de sair cara ou coroa são as mesmas. Mas, se a moeda possuir cara nos dois lados, pode-se com certeza afirmar que sairá cara. Ou seja, na primeira experiência, a informação está oculta, sendo revelada apenas após a jogada, enquanto na segunda não há informação oculta. O quanto de informação oculta em um evento é o que se deseja saber. Pois, medindo-a, será possível dizer o quão incerto se está acerca do seu resultado.

Suponha que se queira saber quão surpreso se está de um elemento ω_i ser o resultado da realização de um evento E_Ω com espaço amostral Ω. Quanto maior a probabilidade $p(\{\omega_i\})$, menos surpreso se está. Entende-se por "surpresa"de ω_i a medida do quão não esperado é que o resultado de um evento seja ω_i. A essa quantidade, é dado também o nome autoinformação $h(p(\{\omega_i\}))$. Mas o que isso tem a ver com a incerteza do resultado de um evento? Quanto menos autoinformação seus elementos apresentarem, menor será sua incerteza, pois menos "surpresas"ocorrerão. Para quantificar, portanto, a quantidade de informação contida em um evento, que possui uma medida de probabilidade associada p, representada por $H(p)$, deve-se utilizar o conceito de autoinformação por meio dos postulados abaixo [2]:

[2]No trabalho de Shannon [3], os postulados são outros. O resultado, entretanto, é o mesmo obtido neste trabalho.

- 0. $h(p(\{\omega\}))$ é não-negativa $\forall\omega \in \Omega$;
- 1. $h(p(\{\omega\}))$ é contínua em $(0, 1]$ $\forall\omega \in \Omega$;
- 2. $h(p(\{\omega\}))$ é estritamente decrescente $\forall\omega \in \Omega$;
- 3. $h(p(\{\omega_1\})\cdot p(\{\omega_2\})) = h(p(\{\omega_1\})) + h(p(\{\omega_1\}))$ $\forall\omega_1, \omega_2 \in \Omega$;
- 4. $H(p) = \sum_\omega h(p(\{\omega\}))$.

O postulado zero é a exigência de que a quantidade de autoinformação não pode ser um número negativo. O primeiro é o que espera de uma boa medida de uma variável real. O segundo postulado é simplesmente a afirmação feita no início do parágrafo anterior. O terceiro surge do fato de que se um resultado ω puder ser descrito como a ocorrência dos resultados independentes ω_1 e ω_2, então a autoinformação de ω_1 mais a autoinformação de ω_2 é a autoinformação de ω. Por exemplo, seja ω o resultado: jogar um dado e sair o número 3 e jogar de novo e sair 4. A surpresa de sair x é a surpresa de sair 3 na primeira jogada mais a surpresa de sair 4 na segunda jogada. Como sair 4 independe de ter saído 3 na primeira jogada, $p(\{(\omega_1, \omega_2)\}) = p(\{\omega_1\})\cdot p(\{\omega_2\}) \Rightarrow h(p(\{(\omega_1, \omega_2)\})) = h(p(\{\omega_2\})) + h(p(\{\omega_2\}))$. O quarto postulado simplesmente diz que incerteza sobre o resultado de um evento é a média das incertezas sobre seus possíveis resultados, o que também é uma afirmação razoável. Decorre desses postulados o seguinte teorema:

Teorema 2.2.1 (Entropia de Shannon)**.** *A função que satisfaz os quatro postulados acima é* $H(p) = -K\sum_{i=1}^{n} p(\{\omega_i\}) \log_d p(\{\omega_i\})$*, em que* $|\Omega| = n$*, d é a base do logaritmo e K é uma constante positiva. Essa função é conhecida como Entropia de Shannon.*

Demonstração. Seja $p_1 \notin \{0, 1\}$ a probabilidade de um resultado ω_1 ocorrer e $p_2 \notin \{0, 1\}$ a de um resultado ω_2 ocorrer. Seja n um número natural arbitrário e m tal que:

$$p_2^m \geq p_1^n > p_2^{m+1} \tag{2.1}$$

De acordo com o postulado 3,

$$h(p_1^n) = nh(p_1) \quad \text{e} \quad h(p_2^m) = mh(p_2) \tag{2.2}$$

De acordo com a inequação (2.1), aplicando a função logaritmo,

$$m\log_d p_2 \geq n\log_d p_1 > (m+1)\log_d p_2 \Rightarrow \frac{m}{n} \leq \frac{\log_d p_1}{\log_d p_2} < \frac{m}{n} + \frac{1}{n} \Rightarrow \left|\frac{\log_d p_1}{\log_d p_2} - \frac{m}{n}\right| < \frac{1}{n} \equiv \epsilon, \tag{2.3}$$

como n é arbitrário, ϵ também é. Pelo postulado 2,

$$\begin{aligned} h(p_2^m) &\leq h(p_1^n) < h(p_2^{m+1}) \Rightarrow mh(p_2) \leq nh(p_1) < (m+1)h(p_2) \Rightarrow \frac{m}{n} \leq \frac{h(p_1)}{h(p_2)} < \frac{m}{n} + \frac{1}{n} \\ &\Rightarrow \left|\frac{h(p_1)}{h(p_2)} - \frac{m}{n}\right| < \frac{1}{n} \equiv \epsilon, \end{aligned} \tag{2.4}$$

O postulado zero permite que a divisão por $nh(p_2)$ não altere as desigualdades. Pela desigualdade triangular,

$$\left|\frac{h(p_1)}{h(p_2)} - \frac{\log_d p_1}{\log_d p_2}\right| \leq \left|\frac{h(p_1)}{h(p_2)} - \frac{m}{n}\right| + \left| - \frac{\log_d p_1}{\log_d p_2} + \frac{m}{n}\right| < 2\epsilon \tag{2.5}$$

Pelo postulado 1, $\dfrac{\log_d p_1}{\log_d p_2}$ é o limite de $-\dfrac{h(p_1)}{h(p_2)}$. Como p_1 e p_2 são arbitrários, $h(p_1) = -K\log_d p_1$ para $p_1 \in (0,1)$, em que K é uma constante positiva. O resultado não foi mostrado para $p_1 = 1$. Mas, se $p_1 = 1 \Rightarrow h(1) = 0$, o que está correto. Com isso, seja $E_\Omega = \{(\omega_i, p(\{\omega_i\}))|i \in \{1,\ldots,n\}\}$, pelo postulado 4,

$$H(p) = -K\sum_{i=1}^{n} p(\{\omega_i\})\log_d p(\{\omega_i\}), \tag{2.6}$$

em que $p(\{\omega_i\}) \neq 0$ para todo $i \in \{1,\ldots,n\}$. □

Como K é uma constante arbitrária, é costume fazer $K = 1$ e trabalhar com a base $d = 2$. Portanto, a entropia pode ser reescrita como:

$$H(E_\Omega) = -\sum_{i=1}^{n} p(\{\omega_i\})\log_2 p(\{\omega_i\}) \tag{2.7}$$

Anteriormente, afirmou-se que o problema de encontrar o melhor código para um evento está relacionado à quantidade de informação do evento, ou seja, à entropia de Shannon. Mas como se dá essa relação? Para mostra-la, serão apresentadas algumas definições e demonstrações.

Sejam E_Ω^p e E_Ω^q dois eventos no mesmo conjunto de caracteres, mas com medidas de probabilidade p e q, respectivamente, distintas. E_Ω^p poderia ser, por exemplo, o lançamento de um dado não viciado e E_Ω^q o lançamento de um que esteja viciado. Os resultados dos dois eventos são os mesmos, mas as medidas de probabilidade não. Espera-se que um dado "levemente"viciado apresente uma estatística de resultados semelhante à do dado não viciado. Como medir o "quão viciado"um dado está? Ou, de modo mais geral, como medir a distância entre duas medidas de probabilidade? Para isso, é necessário definir uma métrica no conjunto $\mathcal{W}(X)$, definido como a classe de todas as medidas de probabilidade com domínio em uma σ-álgebra de X. Entretanto, é comum usar uma função, denominada entropia relativa $H(p\|q)$, para calcular as distâncias em $\Gamma(X)$, ainda que ela não seja uma métrica.

A entropia relativa é uma função com domínio em $\Gamma(X) \times \Gamma(X)$ e contradomínio em $\mathbb{R}$ definida pela relação:

$$H(p\|q) \equiv -\sum_{i=1}^{n} p(\{\omega_i\}) \log_2 \left(\frac{q(\{\omega_i\})}{p(\{\omega_i\})} \right) \tag{2.8}$$

Teorema 2.2.2 (Não-negatividade da entropia relativa)**.** *A entropia relativa $H(p\|q)$ é não-negativa para todos elementos de seu domínio e é nula se, e somente se, $p = q$.*

Demonstração. Um fato crucial para a demonstração é provar que $\ln x \leq x - 1$. Seja $f(x) = \ln x + 1 - x \Rightarrow \dfrac{df(x)}{dx} = \dfrac{1}{x} - 1$ Se $\dfrac{df(x)}{dx} = 0 \Rightarrow x = 1$. $\dfrac{d^2 f(x)}{dx^2} = \dfrac{-1}{x^2} < 0 \Rightarrow x = 1$ é ponto de máximo. Como $f(x = 1) = 0$,

$$f(x) \leq 0 \Leftrightarrow \ln x \leq x - 1 \Leftrightarrow -\log_2 x \geq \frac{1 - x}{\ln 2}. \tag{2.9}$$

De (2.9),

$$\begin{aligned} H(p\|q) &= -\sum_{i=1}^{n} p(\{\omega_i\}) \log_2 \left(\frac{q(\{\omega_i\})}{p(\{\omega_i\})} \right) \geq \frac{1}{\ln 2} \sum_{i=1}^{n} p(\{\omega_i\}) \left(1 - \frac{q(\{\omega_i\})}{p_(\{\omega_i\})} \right) = \\ &= \frac{1}{\ln 2} \sum_{i=1}^{n} \left(p(\{\omega_i\}) - q(\{\omega_i\}) \right) = 0 \end{aligned} \tag{2.10}$$

Como $\ln x = x - 1 \Leftrightarrow x = 1 \Rightarrow -\log_2 \left(\dfrac{q(\{\omega_i\})}{p(\{\omega_i\})} \right) = \left(1 - \dfrac{q(\{\omega_i\})}{p(\{\omega_i\})} \right) \dfrac{1}{\ln 2} \Leftrightarrow p(\{\omega_i\}) = q(\{\omega_i\})$ para todo $\omega_i \in \Omega$ □

O teorema acima mostra que a entropia relativa tem duas propriedades de uma métrica, porém é claramente não simétrica, o que não a qualifica como métrica.

Voltando à questão dos códigos, um código, a princípio, não precisa ser uma bijeção. Entretanto, se ele for, é certo que, dado um conjunto de palavras código, sempre é possível obter os elementos de Ω associados e vice-versa. É conveniente, portanto, a seguinte definição: um código é unicamente decodificável se é uma bijeção. O código C_0 definido anteriormente é unicamente decodificável. O código $C_1 \equiv \{(a,0),(b,10),(c,110),(d,111)\}$ também é um exemplo de bijeção. O próximo teorema, cuja demonstração encontra-se em [29], é necessário para mostrar o resultado após ele:

Teorema 2.2.3 (Desigualdade de Kraft-MacMillan)**.** *Seja $C\colon \Omega \to P_A$ um código unicamente decodificável, tal que $\Omega = \{\omega_1, \ldots, \omega_n\}$ e $|A| = d$,*

$$\sum_{i=1}^{n} d^{-l_C \circ C(\omega_i)} \leq 1. \tag{2.11}$$

Teorema 2.2.4. *O comprimento médio $L(C)$ de qualquer código unicamente decodificável $C\colon \Omega \to P_A$ é maior ou igual à entropia $H(p)$*

Demonstração. Seja $F \equiv L(C) - H(p)$

$$\begin{aligned} F &= \sum_{i=1}^{n} p(\{\omega_i\}) l_C \circ C(\omega_i)) + \sum_{i=1}^{n} p(\{\omega_i\}) \log_2 p(\{\omega_i\}) \qquad (2.12)\\ &= -\sum_{i=1}^{n} p(\{\omega_i\}) \log_2 2^{-l_C \circ C(\omega_i))} + \sum_{i=1}^{n} p(\{\omega_i\}) \log_2 p(\{\omega_i\}) = \sum_{i=1}^{n} p(\{\omega_i\}) \log_2 \left(\frac{p(\{\omega_i\})}{r(\{\omega_i\})} \right) \\ &- \log_2 c, \end{aligned}$$

em que $r(\{\omega_i\}) = c^{-1} 2^{-l_C \circ C(\omega_i))}$ e $c = \sum_{i=1}^{n} 2^{-l_C \circ C(\omega_i)}$. A função $r(\{\omega_i\})$ é uma medida de probabilidade, pois $r(\{\omega_i\}) \geq 0$ e:

$$\sum_{i=1}^{n} r(\{\omega_i\}) = \sum_{i=1}^{n} \frac{2^{-l_C \circ C(\omega_i)}}{\sum_{j=1}^{n} 2^{-l_C \circ C(\omega_j)}} = 1 \tag{2.13}$$

Pela definição de entropia relativa,

$$F = H(p \| r) + \log_2 \frac{1}{c} \tag{2.14}$$

Pela desigualdade de Kraft-McMillan e pela não negatividade da entropia relativa, $c^{-1} \geq 1 \Rightarrow F \geq 0 \Leftrightarrow L(C) \geq H(p)$ □

Esse resultado revela a ligação entre entropia e códigos: nenhum código bijetor pode ter o comprimento médio de palavra-código menor do que a entropia de Shannon do evento que dá a medida de probabilidades das palavras-código. Ou seja, a entropia de Shannon limita a eficiência de um código. Por exemplo, seja $E_\Omega = \{(a, 1/2), (b, 1/4), (c, 1/8), (d, 1/8)\}$ e $C = C_0$, tem-se que $L(C_0) = 2$ bits e $H(p) = 1,75$ bits. O código C_0, aparentemente, não é o melhor código para o evento E_Ω. Suponha, então, que se utilize o código C_1, obtendo-se $L(C_1) = 1,75$ bits, que coincide com a entropia do evento. Logo, não há código bijetor mais eficiente que C_1 para o evento especificado, ainda que seja possível construir outro que tenha a mesma eficiência como, por exemplo, o código $C_1' = \{(a, 1), (b, 01), (c, 001), (d, 000)\}$.

2.2.1 Entropia condicional e informação mútua

Uma questão que surge naturalmente é como medir a quantidade de informação de um evento que é composto por dois ou mais eventos? Por exemplo: considere o evento de se jogar um dado e uma moeda. Se X são os possíveis resultados do dado e Y os da moeda, o evento é um conjunto do tipo: $E_{X,Y} = \{(x, y, p(x, y))|x \in X, y \in Y \text{ e } p \in \mathcal{W}(X, Y)\}$, fazendo $p(x, y) \equiv p(\{(x, y)\})$ e $p_1(x) \equiv p_1(\{x\}), p_2(y) \equiv p_2(\{y\})$, em que p_1 e p_2 são as medidas de probabilidade marginais de p e $\mathcal{W}(X, Y)$ é o conjunto das possíveis medidas de probabilidade com domínio em uma σ-álgebra de $X \times Y$. Essa definição vale para qualquer evento que dependa de dois conjuntos de caracteres. Como o fato de se jogar uma moeda independe do resultado do lançamento de um dado, diz-se que esses dois eventos são independentes, o que é matematicamente representado pela equação $p(x, y) = p_1(x)p_2(y)$. Caso os eventos não sejam independentes, a única coisa que se pode afirmar sobre as suas medidas de probabilidade é que: $p_1(x) = \sum_y p(x, y)$ e $p_2(y) = \sum_x p(x, y)$. Essa dependência decorre do fato de que os resultados de um evento influenciam nos resultados do outro. Com isso, surge o conceito de probabilidade condicional $p_1(x|y)$ ou $p_2(y|x)$, em que a primeira é a probabilidade do resultado de E_X ser x,

sendo conhecido que o resultado de E_Y é y.

Ao se querer especificar quantidade de informação de um evento dependente de uma medida de probabilidade conjunta p, usa-se a entropia $H(p)$, a qual também pode ser escrita como $H(X,Y)$:

$$H(X,Y) \equiv -\sum_{x,y} p(x,y)\log_2 p(x,y). \tag{2.15}$$

A incerteza a respeito do resultado de E_Y, dado que se conhece o resultado de E_X é:

$$H(Y|X) \equiv -\sum_{x} p_1(x)\sum_{y} p_2(y|x)\log_2 p_2(y|x), \tag{2.16}$$

A probabilidade de ocorrer (x,y) pode ser descrita como a probabilidade de ocorrer x e depois y dado que ocorreu x, ou seja, $p(x,y) = p_1(x)p_2(y|x)$ ou, equivalentemente, $p(x,y) = p_2(y)p_1(x|y)$. Sendo assim,

$$\begin{aligned} H(Y|X) &= -\sum_{x,y} p(x,y)\log_2 p(x,y) + \left(\sum_{y} p_2(y|x)\right)\sum_{x} p_1(x)\log_2 p_1(x) \\ &= H(X,Y) - H(X). \end{aligned} \tag{2.17}$$

Suponha agora que se deseja conhecer o quão dependente E_X é de E_Y. Se os eventos forem independentes:

$$\begin{aligned} H(X,Y) &= -\sum_{x,y} p_1(x)p_2(y)\log_2 p_1(x)p_2(y) \\ &= -\left(\sum_{x} p_1(x)\right)\sum_{y} p_2(y)\log_2 p_2(y) - \left(\sum_{y} p_2(y)\right)\sum_{x} p_1(x)\log_2 p_1(x) \\ &= H(X) + H(Y). \end{aligned} \tag{2.18}$$

Como a igualdade ocorre para eventos independentes, ela não deve se manter para eventos dependentes. A diferença dada por:

$$H(X\colon Y) \equiv H(X) + H(Y) - H(X,Y), \tag{2.19}$$

chamada de informação mútua de E_X e E_Y, ela mede quanta informação é compartilhada entre os dois eventos e só é nula se os eventos forem independentes, pois eles não compartilham informação. Mas a entropia conjunta poderá ser maior ou menor que a entropia das partes?

Teorema 2.2.5 (Subaditividade)**.** $H(X,Y) \leq H(X) + H(Y) \Leftrightarrow H(X\colon Y) \geq 0$.

Demonstração. Seja $q(x,y) = p_1(x)p_2(y)$ e pela equação (2.9),

$$\begin{aligned} -H(p\|q) &= \sum_{x,y} p(x,y) \log_2 \frac{p_1(x)p_2(y)}{p(x,y)} \leq \frac{1}{\ln 2} \sum_{x,y} p(x,y) \left(\frac{p_1(x)p_2(y)}{p(x,y)} - 1 \right) \\ &= \frac{1}{\ln 2} \sum_x p_1(x) \sum_y p_2(y) - \sum_{x,y} p(x,y) = 0. \end{aligned} \tag{2.20}$$

Mas,

$$\begin{aligned} H(p\|q) &= \sum_{x,y} p(x,y) \log_2 p(x,y) - \sum_x \left(\sum_y p(x,y) \right) \log_2 p_1(x) - \sum_y \left(\sum_x p(x,y) \right) \log_2 p_2(y) \\ &= H(X,Y). \end{aligned} \tag{2.21}$$

De (2.20) e (2.21), $-H(X,Y) \leq 0 \Leftrightarrow H(X,Y) \geq 0$. □

O que significa uma informação mútua positiva? Significa que a entropia conjunta é menor que a soma da entropia das partes, ou seja, a dependência entre E_X e E_Y reduz a incerteza sobre o resultado do evento $E_{X,Y}$.

Como exemplo, suponha que seja realizado um concurso para uma vaga de professor em uma universidade. Há dois candidatos, a_1 e a_2, que formam o conjunto X. No dia da entrega do resultado, os candidatos podem estar vestidos com uma camisa laranja c_1 ou amarela c_2, formando o conjunto Y. a_1 estará vestido com c_1 com probabilidade $p(c_1|a_1) = \frac{1}{4}$ e com c_2 com $p(c_2|a_1) = \frac{3}{4}$. Para a_2, tem-se $p(c_1|a_2) = \frac{1}{2}$ e $p(c_2|a_2) = \frac{1}{2}$. A probabilidade de a_1 ganhar o concurso é de $p(a_1) = \frac{1}{3}$ e a de a_2 é de $p(a_2) = \frac{2}{3}$. O evento E_X corresponde a quem irá ser aprovado no concurso. Suponha que E_Y corresponda a qual cor de camisa o ganhador estará usando no dia da apresentação do resultado. Deve-se ter que:

$$p(c_1) = \sum_{i=1}^{2} p(a_i)p(c_1|a_i) = \frac{5}{12} \quad \text{e} \quad p(c_2) = \sum_{i=1}^{2} p(a_i)p(c_2|a_i) = \frac{7}{12}. \tag{2.22}$$

Agora que as probabilidades do evento E_Y foram calculadas, resta saber se ele depende do evento E_Y. Sabendo que $p(a_i, c_j) = p(a_i)p(c_j|a_i)$, tem-se que:

$$p(a_1, c_1) = \frac{1}{3} \cdot \frac{1}{4} = \frac{1}{12} \neq \frac{5}{36} = \frac{1}{3} \cdot \frac{5}{12} = p(a_1)p(c_1) \tag{2.23}$$

Essa inequação mostra que os eventos dependem entre si. Se dependem, sua entropia conjunta deve ser menor que a soma das entropias das partes.

$$H(X,Y) \simeq 1,29 \quad \text{bits}, \quad H(X) + H(Y) \simeq 1,32 \quad \text{bits} \Rightarrow H(X:Y) \simeq 0,03 \quad \text{bits}. \tag{2.24}$$

A informação mútua, pela equação (2.17), pode ser escrita como:

$$H(X:Y) = H(X) - (H(X,Y) - H(Y)) = H(X) - H(X|Y). \tag{2.25}$$

A possibilidade de escrever a informação mútua de duas formas terá significativa importância nos desenvolvimentos futuros.

Capítulo 3

Informação Quântica

3.1 Entropia de von Neumann

Quando os estados são quânticos, alguma coisa é alterada na teoria da informação? Suponha que se queira saber o quão incerto se está sobre o estado de um sistema quântico. Para conhecê-lo, é necessário medi-lo. Ao se lançar um dado, não é possível saber, a princípio, qual será a face que cairá para cima. Esse conhecimento só é adquirido quando o dado é observado. Entretanto, a observação do dado não altera nenhuma propriedade do mesmo. Porém, é possível que a medida de um sistema quântica altere seu estado. Suponha que se queira medir a componente z do spin de um sistema de dois níveis. Se o estado do sistema for dado por $|\psi\rangle = \frac{|+\rangle + |-\rangle}{\sqrt{2}}$ e a medida dada pelo conjunto $M = \{P_+ = |+\rangle\langle +|, P_- = |-\rangle\langle -|\}$, caso a medida tenha como resultado $\frac{\hbar}{2}$, o estado $|\psi\rangle$ será colapsado no estado $|\psi_{out}\rangle = |+\rangle$. Sendo assim, o resultado da medida não corresponde diretamente ao estado do sistema, mas sim ao seu estado após ela. Isso significa que, mesmo após a medida, ainda não há certeza sobre qual estado estava o sistema, diferentemente do que acorre em sistemas clássicos. Porém, é possível escolher um conjunto de medidas que não altere estado do sistema. Seja $M' = \{P_1 = |\psi\rangle\langle\psi|, P_2 = I - |\psi\rangle\langle\psi|$, a probabilidade da medida ser P_1 é $p_1 = \mathrm{tr}(P_1|\psi\rangle\langle\psi|) = \mathrm{tr}(|\psi\rangle\langle\psi|) = 1$. O estado $|\psi\rangle$, portanto, não será alterado pela medida. Além disso, a medida deixa de ser um evento aleatório, o que significa que não há incerteza no resultado. Como sempre é possível, para todo estado $|\psi\rangle$, criar o conjunto de medidas M', não há incerteza associada a estados

puros. Mas qual é a incerteza associada a estados de mistura?

Seja ρ o operador densidade definido no espaço de Hilbert $\mathcal{H}$ de um sistema quântico. Como ρ é positivo, ele também é hermitiano e será um observável se $\mathcal{H}$ possuir dimensão finita. Suponha que esse seja o caso. Pode-se, portanto, medir ρ, cujos possíveis resultados são seus autovalores. Pela decomposição espectral, $\rho = \sum_k \lambda_k |\rho_k\rangle\langle\rho_k|$, no caso de os autovalores não serem degenerados. Os vetores $|\rho_k\rangle$ são normalizados e ortogonais entre si.

$$P_k \equiv |\rho_k\rangle\langle\rho_k| \Rightarrow P_k P_k = |\rho_k\rangle\langle\rho_k|\rho_k\rangle\langle\rho_k| = |\rho_k\rangle\langle\rho_k| \Rightarrow P_k P_k = P_k, \tag{3.1}$$

P_k é, portanto, um projetor e satisfaz a relação de completeza $\sum_k P_k = I$, em que I é o operador identidade. É possível, então, escrever $\rho = \sum_k \lambda_k P_k$. E, além disso, o conjunto $\{P_k\}$ é um conjunto completo de medida. Ao se medir ρ, portanto, a probabilidade p_k do resultado da medida ser λ_k é

$$p_k = \mathrm{tr}(P_k\rho) = \mathrm{tr}\left(\sum_{k'} \lambda_{k'} P_k P_{k'}\right) = \mathrm{tr}(P_k\lambda_k) = \lambda_k \,\mathrm{tr}(P_k) = \lambda_k. \tag{3.2}$$

Seja $E_X = \{(\lambda_k, p_k)\}$ para todo k, a incerteza do resultado da medida é dada por

$$H(X) = -\sum_k p_k \log_2 p_k = -\sum_k \lambda_k \log_2 \lambda_k. \tag{3.3}$$

O resultado está expresso em função dos autovalores do operador densidade. Há como expressar o resultado em termos diretamente de ρ? Bom, para responder a essa pergunta, a seguinte definição será útil: seja $d^B = A$, em que A e B são operadores pertencentes a $L(\mathcal{H})$, que é conjunto dos operadores lineares que atuam no espaço $\mathcal{H}$, o logaritmo de A é tal que $\log_d(A) \equiv B$. Se $A = \rho$,

$$\begin{aligned} d^{\sum_k \alpha_k P_k} &= e^{\sum_k \ln(d)\alpha_k P_k} = \sum_s \frac{\left(\sum_k \ln(d)\alpha_k P_k\right)^s}{s\,!} = \sum_k \left(\sum_s \frac{(\ln(d)\alpha_k)^s}{s\,!}\right) P_k = \sum_k e^{\ln(d)\alpha_k} \\ &= \sum_k d^{\alpha_k} P_k = \sum_k \lambda_k P_k \Rightarrow \alpha_k = \log_d \lambda_k \Rightarrow \log_d \rho = \sum_k \log_d(\lambda_k) P_k. \end{aligned} \tag{3.4}$$

Pela equação acima, pode-se demonstrar a seguinte proposição:

Proposição 3.1.1. $H(X) = -\sum_k \lambda_k \log_2 \lambda_k = -\,\mathrm{tr}(\rho \log_2 \rho)$

Demonstração.

$$\begin{aligned} -\operatorname{tr}(\rho \log_2 \rho) &= -\operatorname{tr}\left(\sum_k \lambda_k P_k \sum_s \log_2(\lambda_s) P_s\right) = -\operatorname{tr}\left(\sum_{k,s} \lambda_k \log_2(\lambda_s) P_k P_s\right) \\ &= -\operatorname{tr}\left(\sum_k \lambda_k \log_2(\lambda_k) P_k\right) = -\sum_k \lambda_k \log_2 \lambda_k \end{aligned} \tag{3.5}$$

□

Com a motivação dada pelo resultado acima, define-se a entropia quântica, ou entropia de von Neumann, por:

$$S(\rho) \equiv -\operatorname{tr}(\rho \log_2 \rho) \tag{3.6}$$

que é interpretada como a quantidade de informação de um sistema quântico no estado ρ. Essa generalização da entropia de Shannon pode ser obtida por meios termodinâmicos, como foi mostrado por von Neumann [4].

Para estados muitos sistemas, a entropia não sofre modificações em sua forma, uma vez que ρ descreve qualquer quantidade de sistemas. Entretanto, se há sistemas A e B é comum escrever $S(A, B)$ para a entropia conjunta.

3.1.1 Propriedades da entropia

Da mesma forma que é feito em teoria de informação clássica, é comum definir a entropia relativa quântica $S(\rho\|\sigma)$ por:

$$S(\rho\|\sigma) \equiv \operatorname{tr}(\rho \log_2 \rho) - \operatorname{tr}(\rho \log_2 \sigma) \tag{3.7}$$

A entropia relativa quântica tem as mesmas propriedades que fazem a clássica ser utilizada como uma medida de distância, embora não seja uma métrica. A prova disso está no teorema abaixo:

Teorema 3.1.2 (Desigualdade de Klein)**.** *Sejam ρ e σ operadores densidade, $S(\rho\|\sigma) \geq 0$, com a igualdade é satisfeita se, e somente se, $\rho = \sigma$*

Demonstração. Seja $\rho = \sum_k p_k|\rho_k\rangle\langle\rho_k|$ e $\sigma = \sum_l q_l|\sigma_l\rangle\langle\sigma_l|$,

$$\begin{aligned} S(\rho\|\sigma) &= \sum_k p_k \log_2 p_k - \mathrm{tr}\left(\sum_k \sum_l p_k \log_2(q_l)|\rho_k\rangle\langle\rho_k|\sigma_l\rangle\langle\sigma_l|\right) \\ &= \sum_k p_k \log_2 p_k - \sum_j \sum_k \sum_l p_k \log_2(q_l)\langle\rho_j|\rho_k\rangle\langle\rho_k|\sigma_l\rangle\langle\sigma_l|\rho_j\rangle \\ &= \sum_k p_k \log_2 p_k - \sum_k \sum_l p_k \log_2(q_l)|\langle\rho_k|\sigma_l\rangle|^2 \end{aligned} \tag{3.8}$$

Seja $Q_{k,l} = |\langle\rho_k|\sigma_l\rangle|^2$, $\sum_k Q_{k,l} = \langle\sigma_l|\left(\sum_k |\rho_k\rangle\langle\rho_k|\right)|\sigma_l\rangle = \langle\sigma_l|\sigma_l\rangle = 1 \Rightarrow \sum_l Q_{k,l}q_l \equiv r_k$ é uma combinação convexa de q_l e r é uma medida de probabilidade. Como $\log_2$ é uma função côncava,

$$\sum_l \log_2(q_l)Q_{k,l} \leq \log_2\left(\sum_l q_l Q_{k,l}\right) \Rightarrow S(\rho\|\sigma) \geq \sum_k p_k(\log_2 p_k - \log_2 r_k) = H(p\|r) \geq 0 \tag{3.9}$$

Se $\rho = \sigma$, $S(\rho\|\sigma) = 0$ de maneira óbvia. Mas, se $S(\rho\|\sigma) = 0$, $p = r \Rightarrow p_k = \sum_l Q_{k,l}q_l$. Também isso implica que

$$\begin{aligned} \sum_l \log_2(q_l)Q_{k,l} &= \log_2\left(\sum_l q_l Q_{k,l}\right) = \log_2 p_k \Rightarrow p_k = 2^{\sum_l \log_2(q_l)Q_{k,l}} = \prod_l 2^{\log_2(q_l)Q_{k,l}} = \prod_l q_l^{Q_{k,l}} \\ &\Rightarrow \sum_k p_k = \prod_l q_l^{\sum_k Q_{k,l}} \Rightarrow 1 = \prod_l q_l \Rightarrow q_l = 1, \quad \text{para todo } l, \end{aligned} \tag{3.10}$$

o que é um absurdo. Para que isso não ocorra, obrigatoriamente deve-se ter $Q_{k,l} = \delta_{k,l} \Rightarrow p_k = q_k \Rightarrow \rho = \sigma$ □

É comum classificar estados quânticos em estados puros e de mistura. Mas o que significa um estado estar maximamente misturado? O quantifica a mistura? Definindo que quanto maior a mistura estatística de estados, mais difícil é prever qual é o estado do sistema, então o estado de maior mistura é o que possui a maior entropia. Mas qual é o maior valor da entropia? Usando o fato de que o logaritmo é uma função côncava,

$$S(\rho) = -\sum_{i=1}^{d} \lambda_i \log_2 \lambda_i = \sum_{i=1}^{d} \lambda_i \log_2\left(\frac{1}{\lambda_i}\right) \leq \log_2\left(\sum_{i=1}^{d} \frac{1}{\lambda_i}\lambda_i\right) = \log_2 d. \tag{3.11}$$

Para qual estado a igualdade em (3.11) ocorre?

$$\log_2 d = \log_2\left(d^{\frac{1}{d}}\right)^d = \sum_{i=1}^{d} \log_2 d^{\frac{1}{d}} = -\sum_{i=1}^{d} \frac{1}{d}\log_2\frac{1}{d} = S(I/d), \tag{3.12}$$

ou seja, $\rho = I/d$ é o estado maximamente misturado.

Uma propriedade, que será utilizada mais adiante, é a concavidade da entropia, expressa no seguinte teorema:

Teorema 3.1.3 (Concavidade da entropia)**.** *Sejam ρ_i estados quânticos e p_i probabilidades, tem-se que $S\left(\sum_i p_i\rho_i\right) \geq \sum_i p_i S(\rho_i)$*

Outro resultado de interesse é dado por:

Teorema 3.1.4. *Sejam p_i probabilidades e ρ_i estados com suporte em subespaços ortogonais, tem-se que*

$$S\left(\sum_i p_i\rho_i\right) = H(p) + \sum_i p_i S(\rho_i) \tag{3.13}$$

Demonstração. Uma vez que ρ_i são estados com suporte em subespaços ortogonais, seja $\rho_i = \sum_j \lambda_j^{(i)} |\rho_j^{(j)}\rangle$, tem-se $\langle\rho_{i'}^{(j')}|\rho_i^{(j)}\rangle = \delta_{i,i'}\delta_{j,j'}$, o que significa que $p_i\lambda_j^{(i)}$ são os autovalores do estado $\rho \equiv \sum_i p_i\rho_i$. Assim,

$$\begin{aligned} S(\rho) &= -\sum_{i,j} p_i\lambda_j^{(i)} \log_2 p_i\lambda_j^{(i)} = -\sum_i \left(\sum_j \lambda_j^{(i)}\right) p_i \log_2 p_i - \sum_i p_i \sum_j \lambda_j^{(i)} \log_2 \lambda_j^{(i)} \\ &= H(p) + \sum_i p_i S(\rho_i). \end{aligned} \tag{3.14}$$

□

O próximo teorema é consequência direta do anterior.

Teorema 3.1.5 (Entropia conjunta)**.** *Sejam p_i probabilidades, $|b_i\rangle$ estados ortogonais de um sistema B e $\{\rho_A^{(i)}\}$ um conjunto de operadores densidade de um sistema A, vale que*

$$S\left(\sum_i \rho_A^{(i)} \otimes p_i |b_i\rangle\langle b_i|\right) = H(p) + \sum_i p_i S(\rho_A^{(i)}), \tag{3.15}$$

em que $H(p) = -\sum_i p_i \log_2 p_i$

Se for efetuada uma medida sobre um sistema, como sua entropia se altera? Ou melhor, se o resultado permanece desconhecido, ou seja, se o aparato de medida interagiu com o sistema, mas o observador não averiguou o resultado, qual será a entropia desse novo estado?

Teorema 3.1.6. *Seja* $\{P_k\}$ *um conjunto completo de medidas projetivas ortogonais e* $\rho^D \equiv \sum_k P_k \rho P_k$. $S(\rho^D) \geq S(\rho)$, *sendo a igualdade satisfeita se, e somente se,* $\rho^D = \rho$.

Demonstração. Seja $S(\rho\|\rho^D) = -S(\rho) - \mathrm{tr}(\rho \log_2 \rho^D)$, $\sum_k P_k = I$ e $P_k^2 = P_k$,

$$\begin{aligned} -\mathrm{tr}(\rho \log_2 \rho^D) &= -\mathrm{tr}\left(\sum_k P_k \rho \log_2 \rho^D\right) = -\mathrm{tr}\left(\sum_k P_k P_k \rho \log_2 \rho^D\right) \\ &= -\sum_k \mathrm{tr}\left(P_k P_k \rho \log_2 \rho^D\right) = -\sum_k \mathrm{tr}\left(P_k \rho \log_2 \rho^D P_k\right). \end{aligned} \tag{3.16}$$

Seja $[\rho^D, P_k] = \rho^D P_k - P_k \rho^D$,

$$[\rho^D, P_k] = \sum_l P_l \rho P_l P_k - \sum_s P_k P_s \rho P_s = P_k \rho P_k - P_k \rho P_k = 0 \Rightarrow [\log_2 \rho^D, P_k] = 0. \tag{3.17}$$

De (3.17) em (3.16),

$$\begin{aligned} &-\mathrm{tr}(\rho \log_2 \rho^D) = -\sum_k \mathrm{tr}\left(P_k \rho P_k \log_2 \rho^D\right) = -\mathrm{tr}\left(\sum_k P_k \rho P_k \log_2 \rho^D\right) = S(\rho^D) \\ &\Rightarrow S(\rho\|\rho^D) = S(\rho^D) - S(\rho) \geq 0, \end{aligned} \tag{3.18}$$

sendo que a igualdade na inequação só ocorre se, e somente se, $\rho^D = \rho$, pelo teorema (3.1.2). □

3.1.2 Entropia condicional

Uma questão importante é como definir uma entropia quântica condicional. Primeiramente, qual deve ser seu significado? Ela deve medir a incerteza sobre o estado de um sistema A, dado que se conhece o estado do sistema B. Mas como esse conhecimento é obtido? Por medidas, claro. Então, após medir o estado em B, qual a entropia em A? Com isso, surge uma questão: qual medida efetuar? A medida escolhida deverá, necessariamente, minimizar a incerteza em A. Então, a entropia condicional será definida como a média das incertezas de A, dado que se efetuou a melhor medida possível em B. Como a medida é um evento aleatório, a entropia condicional $\overline{S}(A|B)$ deve ser definida como uma média, dada por:

$$\overline{S}(A|B) \equiv \min_{\{B_k\}} \{\sum_k p_k S(\rho_A^{(k)})\}, \tag{3.19}$$

em que ρ é o estado do sistema AB, $\rho_A = \mathrm{tr}_B \rho$, $\{B_k\}$ é um conjunto de medidas locais em B, $\rho_A^{(k)} = \dfrac{1}{p_k} \mathrm{tr}_B[(I_A \otimes B_k)\rho(I_A \otimes B_k)]$ e $p_k = \mathrm{tr}(I_A \otimes B_k \rho I_A \otimes B_k)$.

Se o estado do sistema composto for $\rho = \rho_A \otimes \rho_B$, A e B são independentes um do outro. Qualquer medida que se efetuar em um, não irá alterar o estado do outro. Isso significa deve-se ter $S(A) = \overline{S}(A|B)$ ou $S(B) = \overline{S}(B|A)$. Seja $\{B_k\}$ um conjunto de medidas em B,

$$\begin{aligned}\rho^{(k)} &= \rho_A \otimes B_k \rho_B B_k^\dagger \Rightarrow \rho_A^{(k)} = \rho_A \,\mathrm{tr}\left(B_k \rho_B B_k^\dagger\right) = \rho_A \Rightarrow \overline{S}(A|B) = \min_{\{B_k\}}\{\sum_k p_k S(\rho_A^{(k)})\} \\ &= \min_{\{B_k\}}\{\left(\sum_k p_k\right) S(\rho_A)\} = S(A) \end{aligned} \tag{3.20}$$

Deve ser destacado o fato de que essa definição não é convencionalmente. Na literatura, define-se entropia condicional de modo análogo ao caso clássico: $S(A|B) \equiv S(A,B) - S(B)$. Uma diferença básica entre as duas entropias é que $\overline{S}(A|B) \geq 0, \forall A, B$, o que é evidente pelo modo como foi definida, enquanto $S(A|B)$ pode ter valores negativos. Por exemplo, se $|\psi\rangle \equiv \frac{|01\rangle + |10\rangle}{\sqrt{2}}$, $S(A|B) = -1$.

3.2 Discórdia Quântica

Para medir se existe dependência entre os estados, há duas maneiras, não equivalentes, de se definir a informação mútua quântica. Uma maneira é, de modo análogo ao tratamento clássico, defini-la como $T(A:B)$:

$$T(A:B) \equiv S(A) + S(B) - S(A,B), \tag{3.21}$$

ou $C(A:B)$ dada por [20]:

$$C(A|B) \equiv S(A) - S(A|B). \tag{3.22}$$

Qual informação mútua irá fornecer, portanto, dependência entre os subsistemas? Como não é necessário efetuar medidas para se obter T, as relações de dependência não podem sofrer nenhum tipo de perda. Logo, T deve medir a correlação total. Mas, que tipo de informação fornece C? Como a diferença entre T e C só pode ocorrer em sistemas quânticos, sendo iguais para sistemas clássicos, C deve ser uma medida de correlação puramente clássica, enquanto $T - C$ deve ser uma medida de correlação puramente quântica. Oliver e Zurek [19] definem

essa diferença:

$$D(A|B) = T(A:B) - C(A|B) \quad \text{ou} \quad D(B|A) = T(A:B) - C(B|A) \tag{3.23}$$

por discórdia quântica. A primeira coisa que se deve verificar é se realmente as correlações em T são maiores do que as obtidas em C. Mas, antes disso, é mostrar o seguinte resultado:

Teorema 3.2.1. *O conjunto de medidas POVM* $\{E_k\}$ *que minimiza a quantidade* $\sum_k p_k S(\rho_A^{(k)})$ *são elementos POVM de* $rank = 1$.

Demonstração. Seja o sistema de estado ρ constituído dos subsistemas A e B. O estado do subsistema A após a medida é $\rho_A^{(k)} = \dfrac{1}{p_k} \cdot \mathrm{tr}_B \left[(I_A \otimes E_k)\, \rho \right]$ e $p_k = \mathrm{tr}\left[(I_A \otimes E_k)\, \rho \right]$. Como E_k é um operador positivo, é possível lhe aplicar a decomposição espectral de modo que:

$$E_k = \sum_l \lambda_{k,l} |e_{k,l}\rangle\langle e_{k,l}|, \quad E_{k,l} \equiv \lambda_{k,l} |e_{k,l}\rangle\langle e_{k,l}| \Rightarrow E_k = \sum_l E_{k,l} \Rightarrow \sum_{k,l} E_{k,l} = I, \tag{3.24}$$

ou seja, $E_{k,l}$ são elementos POVM de $rank = 1$. Pode-se, portanto, utilizar esse conjunto para efetuar medidas no sistema,

$$\begin{aligned}
\rho_A^{(k,l)} &= \frac{1}{p_{k,l}} \cdot \mathrm{tr}_B \left[(I_A \otimes E_{k,l})\, \rho \right], \quad p_{k,l} = \mathrm{tr}\left[(I_A \otimes E_{k,l})\, \rho \right] \Rightarrow \sum_l p_{k,l} \\
&= \mathrm{tr}\left[\left(I_A \otimes \sum_l E_{k,l} \right) \rho \right] = \mathrm{tr}\left[(I_A \otimes E_k)\, \rho \right] = p_k \Rightarrow p_{l|k} = \frac{p_{l,k}}{p_k} \Rightarrow \rho_A^{(k)} \\
&= \frac{1}{p_k} \cdot \mathrm{tr}_B(E_k \rho) = \sum_k \frac{p_{l|k}}{p_{k,l}} \mathrm{tr}_B(\rho E_{k,l}) = \sum_l p_{l|k} \rho_A^{(k,l)} \Rightarrow \sum_k p_k S(\rho_A^{(k)}) \\
&= \sum_k p_k S\left(\sum_l p_{l|k} \rho_A^{(k,l)} \right) \geq \sum_{k,l} p_l p_{l|k} S(\rho_A^{(k,l)}) = \sum_{k,l} p_{k,l} S(\rho_A^{k,l})
\end{aligned} \tag{3.25}$$

□

Além disso, o artigo de D'Ariano e colaboradores [31] mostra que um subconjunto do conjunto de elementos POVM de $rank = 1$, que possui elementos linearmente independentes, minimiza a entropia condicional. Na tese de doutorado de Animesh Datta [32], ele afirma ser suficiente fazer a minimização utilizando projetores ortogonais como operadores de medida. A prova desses resultados, entretanto, não será discutida nesse trabalho.

Para verificar se $T \geq C$, é necessário enunciar os seguintes resultados:

Teorema 3.2.2 (Operações Quânticas nunca aumentam a informação mútua T)**.** *Seja AB um sistema composto e E uma operação quântica no sistema B que preserva o traço. Seja $T(A:B)$ a informação mútua antes da realização da operação quântica e $T(A':B')$ a informação mútua após a operação, resulta que $T(A':B') \leq T(A:B)$.*

Sejam $\mathcal{H}^a$ e $\mathcal{H}^b$ os espaços de Hilbert dos subsistemas A e B. Sejam $L(\mathcal{H}^a)$ e $L(\mathcal{H}^b)$ os conjuntos de todos os operadores lineares que atuam nos espaços $\mathcal{H}^a$ e $\mathcal{H}^b$, respectivamente. Sejam $X, Y \in L(\mathcal{H}^a)\text{ou}L(\mathcal{H}^b)\text{ou}L(\mathcal{H}^a \otimes \mathcal{H}^b)$, o produto interno é definido por:

$$\langle X, Y \rangle \equiv \text{tr}\left(X^\dagger Y\right) \tag{3.26}$$

Sejam $\{X_i\}$ e $\{Y_j\}$ bases de $L(\mathcal{H}^a)$ e $L(\mathcal{H}^b)$, respectivamente, cujos elementos são operadores hermitianos ortonormais. O conjunto $\{X_i \otimes Y_j\}$ constitui uma base ortonormal de $L(\mathcal{H}^a \otimes \mathcal{H}^b)$.

O seguinte teorema é encontrado no artigo de Hassan e colaboradores [33]:

Teorema 3.2.3. *Seja ρ o estado de um sistema AB, pode-se escrever:*

$$\rho = \sum_{i,j} c_{i,j} X_i \otimes Y_j, \tag{3.27}$$

sendo $c_{i,j} = \text{tr}(\rho X_i \otimes Y_j)$

Teorema 3.2.4. $D(A|B) \geq 0$

Demonstração. Seja B_k um conjunto de medidas projetivas ortogonais em B, um subsistema do sistema AB. Seja $\rho^{(k)} = \dfrac{1}{p_k} \cdot (I \otimes B_k)\rho(I \otimes B_k)$, em que $p_k = \text{tr}\,(I \otimes B_k \rho)$, tem-se:

$$\rho^D \equiv \sum_k p_k \rho^{(k)} = \sum_k (I \otimes B_k)\rho(I \otimes B_k) = \sum_k \sum_{i,j} c_{i,j} X_i \otimes B_k Y_j B_k \tag{3.28}$$

Como Y_j é hermitiano, pode ser decomposto em termos de seus autovetores $|y_l^{(j)}\rangle$. Seja $B_k = |b_k\rangle\langle b_k|$,

$$\begin{aligned} \sum_k \sum_{i,j} c_{i,j} X_i \otimes B_k Y_j B_k &= \sum_k \sum_{i,j} c_{i,j} X_i \otimes \sum_l \lambda_l^{(j)} |b_k\rangle\langle b_k| y_l^{(j)}\rangle\langle y_l^{(j)}| b_k\rangle\langle b_k| \\ = \sum_k \sum_{i,j} c_{i,j}\alpha_{k,j} X_i \otimes B_k &= \sum_k \sum_{i,j} \frac{c_{i,j}\alpha_{k,j}}{p_k} X_i \otimes p_k B_k = \sum_k \rho_A^{(k)} \otimes p_k B_k, \end{aligned} \tag{3.29}$$

em que $\rho_A^{(k)} = \mathrm{tr}_B\, \rho^{(k)}$. De acordo com o Teorema (3.1.5),

$$S(\rho^D) = \sum_k p_k S(\rho_A^{(k)}) + H(p_k). \tag{3.30}$$

Seja $\rho_B^{(k)}$=$\mathrm{tr}_A\, \rho^{(k)} = \mathrm{tr}_A(\rho_A^{(k)} \otimes B_k) = B_k$,

$$\rho_B^D \equiv \sum_k p_k B_k = \mathrm{tr}_A\, \rho^D \Rightarrow S(\rho_B^D) = H(p_k) \tag{3.31}$$

De (3.30) e (3.31), resulta que:

$$\sum_k p_k S(\rho_A^{(k)}) = S(\rho^D) - S(\rho_B^D) \tag{3.32}$$

A entropia condicional surge do primeiro termo da equação acima após o processo de minimização:

$$S(A|B) = \min_{\{B_k\}} \sum_k p_k S(\rho_A^{(k)}) = \min_{\{B_k\}} S(\rho^D) - S(\rho_B^D). \tag{3.33}$$

Seja $\rho_A^D \equiv \mathrm{tr}_B\, \rho^D$,

$$\begin{aligned}
\rho_A^D &= \sum_k p_k \rho_A^{(k)} = \sum_k \sum_{i,j} c_{i,j}\alpha_{k,j} X_i = \sum_k \sum_{i,j} c_{i,j} X_i \sum_l \lambda_l^{(j)} \langle b_k | y_l^{(j)} \rangle \langle y_l^{(j)} | b_k \rangle \\
&= \sum_{i,j} c_{i,j} X_i \sum_l \lambda_l^{(j)} \langle y_l^{(j)}| \left(\sum_k |b_k\rangle\langle b_k| \right) |y_l^{(j)}\rangle = \sum_{i,j} c_{i,j} X_i \sum_l \lambda_l^{(j)} = \sum_{i,j} c_{i,j} X_i\, \mathrm{tr}(Y_j) \\
&= \mathrm{tr}_B\, \rho = \rho_A.
\end{aligned} \tag{3.34}$$

De (3.33) e (3.34),

$$C(A|B) = S(A) - S(A|B) = S(\rho_A^D) + S(\rho_B^D) - S(\rho^D) = T(A' : B'), \tag{3.35}$$

pois a operação quântica, representada pelo mapa $\varepsilon(\rho) \equiv \rho^D$, preserva o traço, o que faz valer o teorema (3.2.2):

$$D(A|B) = T(A : B) - C(A|B) = T(A : B) - T(A' : B') \geq 0. \tag{3.36}$$

□

De fato, deve haver perda de correlação quando medidas locais são efetuadas no sistema, o que é embasado pelo teorema acima. As medidas devem destruir a correlação quântica que

existe entre os subsistemas. O que deve sobrar são apenas as correlações clássicas, se existirem. Mas, e se um estado ρ não for afetado pela medida? Ou seja, e se o estado permanecer o mesmo após a medida? Espera-se que esses estados não devam apresentar correlações quânticas, pois essa invariância é uma propriedade de sistemas clássicos. Portanto, se $\rho = \rho^D$, de (3.36) $T(A:B) = T(A',B') \Rightarrow D(A|B) = 0$, o que era esperado.

Mas, e se a discórdia quântica for nula? Será que há mais critérios sobre o conjunto dos estados que a anulam? Se houver mais critérios, esses são satisfeitos por estados os quais espera-se terem um comportamento clássico? De acordo com o artigo de Zurek [19], a nulidade da discórdia implica que estado global não é alterado pela medida local. Uma demonstração formal desse resultado encontra-se na tese de Animesh Datta.

$$D(A|B) = 0 \Leftrightarrow \rho = \rho^D. \tag{3.37}$$

É claro então que as correlações quânticas se anulam somente para estados invariantes por medidas locais. Se for definido que correlações clássicas são aquelas invariantes por medidas locais, então, de fato, a discórdia quântica é uma boa medida de correlação quântica.

Com o resultado (3.37) é possível mostrar para que tipo de estado a discórdia se anula.

Teorema 3.2.5. *Seja ρ a matriz densidade do sistema AB, $D(A|B) = 0 \Leftrightarrow \rho = \sum_k p_k \rho^{(k)} \otimes B_k$, em que $B_k = |b_k\rangle\langle b_k|$ e $\{|b_k\rangle\}$ é um conjunto de vetores ortogonais.*

Demonstração. Seja $\rho = \sum_{i,j} c_{i,j} X_i \otimes Y_j$ e $\{B_k\}$ um conjunto completo de medidas projetivas em B. Pelo resultado (3.37),

$$D(A|B) = 0 \Leftrightarrow \rho = \sum_k (I \otimes B_k)\,\rho\,(I \otimes B_k) \Rightarrow \sum_{i,j} c_{i,j} X_i \otimes Y_j = \sum_{i,j} c_{i,j} X_i \otimes \left(\sum_k B_k Y_j B_k \right)$$
$$\Rightarrow Y_j = \sum_k B_k Y_j B_k \tag{3.38}$$

Seja $Y_j = \sum_l \lambda_l^{(j)} |y_l^{(i)}\rangle\langle y_l^{(i)}|$ e $B_k = |b_k\rangle\langle b_k|$,

$$Y_j = \sum_l \lambda_l^{(j)} \sum_k |b_k\rangle\langle b_k|y_l^{(i)}\rangle\langle y_l^{(i)}|b_k\rangle\langle b_k| = \sum_k \left(\sum_l \lambda_l^{(j)} |\langle b_k|y_l^{(i)}\rangle|^2 \right) B_k \Rightarrow |y_l^{(i)}\rangle\langle y_l^{(i)}| = B_k \tag{3.39}$$

pois a decomposição espectral é única. Pode-se escrever $\rho = \sum_k p_k \dfrac{(I \otimes B_k)\,\rho\,(I \otimes B_k)}{p_k}$, em que $p_k = \mathrm{tr}\,(I \otimes B_k \rho)$,

$$\rho = \sum_k p_k \sum_{i,j} \frac{c_{i,j}}{p_k} X_i \otimes B_k \left(\sum_l \lambda_l^{(j)} B_l \right) B_k = \sum_k \left(\sum_{i,j} \frac{c_{i,j}\lambda_k^{(j)}}{p_k} X_i \right) \otimes p_k B_k = \sum_k p_k \rho^{(k)} \otimes B_k, \tag{3.40}$$

sendo $\rho^{(k)} = \sum_{i,j} \dfrac{c_{i,j}\lambda_k^{(j)}}{p_k} X_i$. Agora, partindo da hipótese de que o estado a ser medido seja dado por $\sum_k p_k \rho^{(k)} \otimes B_k$, basta escolher o conjunto de medidas locais $\{B_k\}$ em B para manter o estado inalterado após uma medida cujo resultado permanece desconhecido, ou seja, a discórdia anula-se para esse estado. □

Bom, então basta calcular a discórdia para se saber se a correlação quântica existe e qual a sua intensidade, correto? Não é tão simples assim. O cálculo da discórdia não é simples, devido ao processo de minimização inserido na definição de entropia relativa. A dificuldade dessa tarefa será mostrada no exemplo a seguir.

O exemplo encontra-se no artigo de Luo [34]. Alguns cálculos, realizados computacionalmente, e algumas passagens demonstrativas serão omitidas, pois os detalhes se encontram no artigo fonte. Considere um sistema AB de dois qubits no estado τ escrito na base computacional $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$. De acordo com o artigo de Fano [35], esse estado pode ser escrito da forma:

$$\tau = \frac{1}{4}\left(I + \vec{u}\vec{\sigma} \otimes I + I \otimes \vec{v}\vec{\sigma} + \sum_{j,k=1}^{3} w_{jk}\sigma_j \otimes \sigma_k \right), \tag{3.41}$$

em que σ_k, $k \in \{1,2,3\}$ são as matrizes de Pauli, em que os índices x,y e z foram trocados por, respectivamente, 1,2 e 3, $\vec{\sigma} = (\sigma_1, \sigma_1, \sigma_1) \in L(\mathcal{H}) \times L(\mathcal{H}) \times L(\mathcal{H})$, $\vec{u}$ e $\vec{v}$ pertencem a $\mathbb{R}^3$, o produto $\vec{u}\vec{\sigma}$ é definido por $\vec{u}\vec{\sigma} = \sum_{i=1}^{3} u_i\sigma_i$ e $w_{j,k}$ são coeficientes reais.

Por meio de transformações unitárias locais nos dois subsistemas, o estado τ é reduzido em:

$$\gamma = \frac{1}{4}\left(I + \vec{a}\vec{\sigma} \otimes I + I \otimes \vec{b}\vec{\sigma} + \sum_{j=1}^{3} c_j\sigma_j \otimes \sigma_j \right), \tag{3.42}$$

em que c_j são coeficientes reais. O cálculo da discórdia quântica para esse estado ainda é muito trabalhoso. Assim, será imposta a condição de que os subsistemas tenham estados

maximamente misturados. Pois, desse modo, a correlação quântica obtida é certamente a que existe apenas entre os subsistemas, pois esses estados não possuem termos de coerência, o que representa uma correlação quântica interna entre os estados da base em cada subsistema. Com essa imposição, tem-se o estado global dado por:

$$\rho = \frac{1}{4}\left(I + \sum_{j=1}^{3} c_j \sigma_j \otimes \sigma_j\right) \tag{3.43}$$

cujos autovalores, calculados com o auxílio de computação algébrica, são dados por:

$$\lambda_0 = \frac{1}{4}(1 - \sum_{k=1}^{3} c_k), \quad \lambda_j = \frac{1}{4}(1 + \sum_{k=1}^{3}(-1)^{\delta_{j,k}} c_k), \quad j \in 1,2,3. \tag{3.44}$$

O que implica em

$$T(A:B) = 2 + \sum_{k=0}^{3} \lambda_k \log \lambda_k \tag{3.45}$$

Considere um conjunto de medidas dado por $\{B_k\}$, em que $B_k = V\Pi_k V^\dagger$, em que V é unitário e Π_k são projetores na base computacional. Mas como escrever uma matriz unitária genérica? Utilizando a propriedade $V^{-1} = V^\dagger$, sendo $V = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$,

$$\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} a^* & c^* \\ b^* & d^* \end{bmatrix} \Rightarrow V = \begin{bmatrix} a & b \\ -b^* & a^* \end{bmatrix}, \tag{3.46}$$

usando também a condição $\det V = 1$, tem-se ainda que $|a|^2 + |b|^2 = 1$. Seja $a \equiv t + \mathrm{i}\, y_3$ e $b \equiv y_2 + \mathrm{i}\, y_1$, tem-se

$$V = (tI + \mathrm{i}\, \vec{y}\vec{\sigma}), \quad t^2 + \sum_{k=1}^{3} y_k^2 = 1, \tag{3.47}$$

em que $\vec{y} \equiv (y_1, y_2, y_3)$. Após a medida local, tem-se

$$\begin{aligned}
p_k\rho_k &= (I \otimes B_k)\,\rho\,(I \otimes B_k) = (I \otimes B_k)\,\frac{1}{4}\left(I + \sum_{j=1}^{3} c_j\sigma_j \otimes \sigma_j\right)(I \otimes B_k) \\
&= \left(I \otimes V\Pi_k V^\dagger\right)\frac{1}{4}\left(I + \sum_{j=1}^{3} c_j\sigma_j \otimes \sigma_j\right)\left(I \otimes V\Pi_k V^\dagger\right) \\
&= (I \otimes V\Pi_k)\left(I \otimes V^\dagger\right)\frac{1}{4}\left(I + \sum_{j=1}^{3} c_j\sigma_j \otimes \sigma_j\right)(I \otimes V)\left(I \otimes \Pi_k V^\dagger\right) \\
&= (I \otimes V\Pi_k)\,\frac{1}{4}\left(I + \sum_{j=1}^{3} c_j\sigma_j \otimes V^\dagger\sigma_j V\right)\left(I \otimes \Pi_k V^\dagger\right),
\end{aligned} \tag{3.48}$$

sendo

$$\begin{aligned} V^\dagger\sigma_j V &= (tI - \mathrm{i}\,\vec{y}\vec{\sigma})\sigma_j(tI + \mathrm{i}\,\vec{y}\vec{\sigma}) = \left(t\sigma_j - \mathrm{i}\sum_{k=1}^{3} y_k\sigma_k\sigma_j\right)(tI - \mathrm{i}\,\vec{y}\vec{\sigma}) \\ &= t^2\sigma_j + \sum_{k,l} y_k y_l \sigma_k\sigma_j\sigma_l + \mathrm{i}\sum_{l=1}^{3} t y_l\sigma_j\sigma_l - \mathrm{i}\sum_{k=1}^{3} t y_k\sigma_k\sigma_j \\ &= t^2\sigma_j + \sum_{k,l} y_k y_l \sigma_k\sigma_j\sigma_l + \mathrm{i}\,t\sum_{l=1}^{3} y_l[\sigma_j,\sigma_l]. \end{aligned} \tag{3.49}$$

Seja $\varepsilon_{ijk} = \dfrac{(j-i)(k-i)(k-j)}{2}$ o símbolo de Levi-Civita, as seguintes relações são válidas:

$$[\sigma_i,\sigma_j] = 2\,\mathrm{i}\,\varepsilon_{ijk}\sigma_k, \tag{3.50}$$

$$\sigma_i\sigma_i\sigma_i = \sigma_i; \quad \sigma_i\sigma_i\sigma_j = \sigma_j\sigma_i\sigma_i; \quad \sigma_i\sigma_j\sigma_i = -\sigma_j; \quad \sigma_i\sigma_j\sigma_k = \varepsilon_{ijk}\,\mathrm{i}\,I, \quad \text{se } i\neq j\neq k. \tag{3.51}$$

Utilizando esses resultados na equação (3.49), obtém-se

$$V^\dagger\sigma_j V = t^2\sigma_j - \sum_{\substack{k=1\\k\neq j}}^{3} y_k^2\sigma_j + 2\sum_{\substack{k=1\\k\neq j}}^{3} y_j y_k\sigma_k + \sum_{\substack{k,l=1\\k\neq j\neq l}}^{3} y_k y_l\varepsilon_{kjl}\,\mathrm{i}\,I - 2t\sum_{l=1} y_l\varepsilon_{jls}\sigma_s + y_j^2\sigma_j. \tag{3.52}$$

Sejam

$$\Pi_j\sigma_k\Pi_j = \delta_{3,k}(-1)^j\Pi_j, \quad p_k = \frac{1}{2}, \tag{3.53}$$

de (3.52) e (3.53) em (3.48),

$$\rho_0 = \frac{1}{2}(I + \vec{x}\vec{\sigma})\otimes B_0, \quad \rho_1 = \frac{1}{2}(I + \vec{x}\vec{\sigma})\otimes B_1 \tag{3.54}$$

$$x_1 \equiv 2c_1(-ty_2 + y_1y_3), x_2 \equiv 2c_2(ty_1 + y_2y_3), x_2 \equiv c_3(t^2 + y_3^2 - y_1^2 - y_2^2). \tag{3.55}$$

Usando (3.54) e (3.55) para calcular as entropias, obtém-se

$$J(A:B)_{\{B_k\}} = \frac{1-|\vec{x}|}{2}\log(1-|\vec{x}|) + \frac{1+|\vec{x}|}{2}\log(1+|\vec{x}|) \tag{3.56}$$

$$c \equiv \max\{|c_1|,|c_2|,|c_3|\} \Rightarrow |\vec{x}| \leq c \tag{3.57}$$

$$C(A:B) = \frac{1-c}{2}\log(1-c) + \frac{1+c}{2}\log(1+c) \tag{3.58}$$

A discórdia quântica é, portanto:

$$D(A:B) = 2 + \sum_{l=0}^{3} \lambda_l \log \lambda_l - \sum_{k=0}^{1} \frac{(1+(-1)^{k+1}c)}{2} \log(1+(-1)^{k+1}c) \tag{3.59}$$

Seja o estado de Werner dado por:

$$\rho = (1-c)\frac{I}{4} + c|\psi^-\rangle\langle\psi^-|, \quad c \in [0,1] \tag{3.60}$$

Tem-se então que $c_1 = c_2 = c_3 = -c$, em que $|\psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$. A discórdia associada a esse estado é dada por:

$$D(A:B) = \frac{1-c}{4}\log_2(1-c) - \frac{1+c}{2}\log_2(1+c) + \frac{1+3c}{4}\log_2(1+3c) \tag{3.61}$$

Algo interessante a se observar é como a discórdia quântica se comporta em relação ao emaranhamento, pois esse é um tipo de correlação quântica, presente em estados que não são produto. Uma medida utilizada para se calcular o grau de emaranhamento é o emaranhamento de formação [36,37]. Seja $\breve{\rho} \equiv (\sigma_y \otimes \sigma_y)\rho^*(\sigma_y \otimes \sigma_y)$ e $\omega \equiv \sqrt{\rho}\sqrt{\breve{\rho}}$. Sejam g_1, g_2, g_3, g_4 os valores singulares de ω em ordem decrescente, a concorrência é dada por:

$$Cn(\rho) = \max\{0, g_1 - g_2 - g_3 - g_4\} = \frac{3c-1}{2} \tag{3.62}$$

Seja $H(x) \equiv -x\log_2 x - (1-x)\log_2(1-x)$ a entropia de Shannon binária, emaranhamento de formação é dado por:

$$E_F(\rho) = H\left(\frac{1+\sqrt{1-(Cn(\rho))^2}}{2}\right) \tag{3.63}$$

O gráfico seguinte mostra a relação entre discórdia e emaranhamento para o estado de Werner. Como é esperado, a discórdia quântica permanece não nula mesmo para estados não emaranhados.

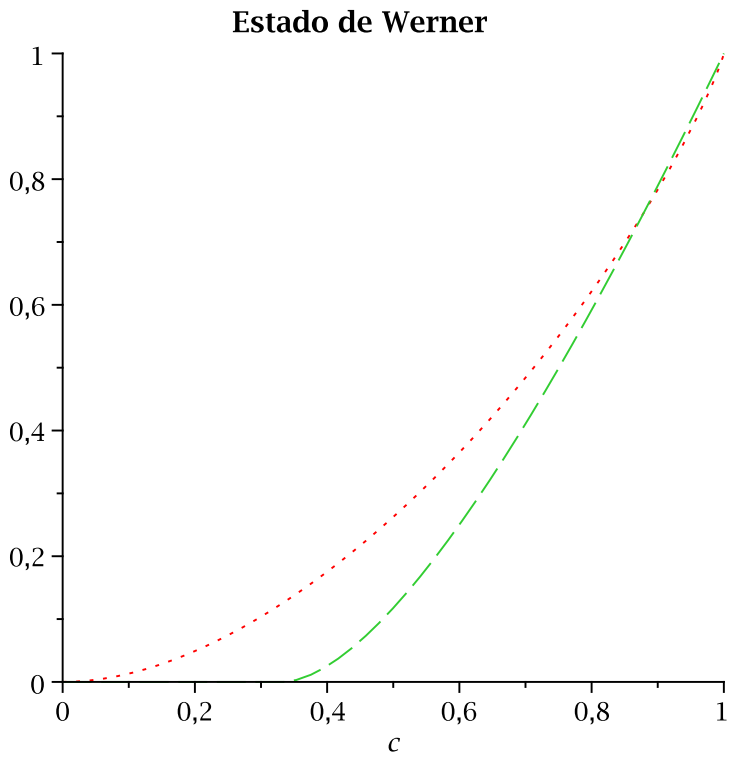


Figura 3.1: A linha pontilhada é a discórdia quântica, enquanto a tracejada é o emaranhamento de formação.

Capítulo 4

Resultados

4.1 Geração de POVMs

Uma questão de cunho prático a se perguntar é: como gerar um conjunto de medidas POVM? A proposição abaixo fornece um ponto de partida para essa construção.

Proposição 4.1.1. *Sejam $\mathcal{H}_\mathcal{A}$ e $\mathcal{H}_\mathcal{B}$ espaços de Hilbert dos sistemas A e B, respectivamente. Sejam $K_A \equiv \{|\alpha_i\rangle \colon i \in \{1,\ldots,m\}\}$ e $K_B \equiv \{|\beta_j\rangle \colon j \in \{1,\ldots,n\}\}$ bases de $\mathcal{H}_\mathcal{A}$ e $\mathcal{H}_\mathcal{B}$ respectivamente e $K_{AB} \equiv \{|\alpha_i\rangle \otimes |\beta_j\rangle \colon i \in \{1,\ldots,m\}, j \in \{1,\ldots,n\}\}$ base de $\mathcal{H}_\mathcal{A} \otimes \mathcal{H}_\mathcal{B}$. O sequência $\psi \equiv (|\psi_1\rangle,\ldots,|\psi_n\rangle)$ tal que $\langle\psi_j|\psi_i\rangle = \delta_{i,j} \quad \forall i,j \in \{1,\ldots,n\}$ determina um único conjunto POVM: $M \equiv \{V_k \colon i \in \{1,\ldots,m\}\}$ na base $K_{BB} \equiv \{|\beta_i\rangle\langle\beta_j| \colon i,j \in \{1,\ldots,n\}\}$.*

Demonstração. Sejam

$$|\psi_j\rangle \equiv \sum_{i,k} v_{i,j}^{(k)} |\alpha_k\rangle|\beta_i\rangle \quad \text{e} \quad V_k \equiv \sum_{i,j} v_{i,j}^{(k)} |\beta_i\rangle\langle\beta_j| \tag{4.1}$$

$$\begin{aligned} \langle\psi_{j'}|\psi_j\rangle &= \sum_{i,k,i',k'} v_{i,j}^{(k)} v_{i',j'}^{(k')*} \left(\langle\alpha_{k'}|\langle\beta_{i'}|\right)\left(|\alpha_k\rangle|\beta_i\rangle\right) = \sum_{i,k,i',k'} v_{i,j}^{(k)} v_{i',j'}^{(k)*} \delta_{i,i'}\delta_{k,k'} = \sum_{k,i} v_{i,j}^{(k)} v_{i,j'}^{(k)*} = \delta_{j,j'} \\ \Rightarrow \sum_k V_k^\dagger V_k &= \sum_{k,i,j,i',j'} v_{i,j}^{(k)*} v_{i',j'}^{(k)} |\beta_j\rangle\langle\beta_i|\beta_{i'}\rangle\langle\beta_{j'}| = \sum_{j,j'} \delta_{j,j'} |\beta_j\rangle\langle\beta_{j'}| = I \end{aligned} \tag{4.2}$$

A sequência ψ determina um conjunto de coeficientes $C_\psi \equiv \{v_{i,j}^{(k)} \colon i,j \in \{1,\ldots,n\}, k \in \{1,\ldots,m\}\}$ que é único, pois um conjunto diferente só pode ser gerado por uma sequência diferente de ψ. Seja $M' \equiv \{V_k' \colon i \in \{1,\ldots,m\}\} \neq M$, existe $V_l, M' \ni V_l' \notin M$ para algum l.

Isso significa que, mesmo que os coeficientes de V_l' pertençam a C_ψ, eles estarão associados a índices diferentes. Mas trocar os índices dos coeficientes altera a sequência ψ. Logo, M é único.

□

4.2 Correlações

Qual o significado de correlação? Pode-se dizer que dois sistemas estão correlacionados quando um interfere no outro. Essa interferência não precisa ser determinística. No caso de sistemas quânticos, cujos estados são descritos por operadores densidade, se um sistema provoca uma alteração no estado do outro, o resultado é mudança na estatística no que foi afetado. Um sistema AB, composto de duas partes, é representado por seu estado global ρ e seus componentes por seus operadores densidade reduzidos ρ_A e ρ_B. Pode-se questionar, então, como uma operação em uma parte afeta o estado global ou da outra parte ou como uma operação no todo afeta as partes. É possível considerar três tipos de operações: evolução unitária, medida não-seletiva e medida seletiva. É imediato questionar quais as classes de estados invariantes sob essas operações. Além disso, tomando como exemplo uma evolução unitária em A, a classe que se quer determinar deve ser invariante para toda evolução ou apenas deve existir uma? Quando se trata de uma medida não-seletiva, a exigência de que a invariância deve permanecer para toda medida significa que, de modo algum, essa operação em um sistema influencia os outros. No entanto, se a exigência é apenas existir uma medida, interpreta-se que há um modo de se medir em que os sistemas ficam descorrelacionados.

Após essa discussão, percebe-se que a correlação depende de quatro parâmetros, a saber, X, Y, α e β, em que:

- Y é sistema no qual se deseja observar uma possível alteração no estado;
- X é o sistema em que se realiza a operação;
- α é o tipo de operação;

- β define se a operação é do tipo: para toda ou existir ao menos uma.

As classes de estados invariantes são denotadas por $\Gamma_X\left(Y,\alpha,\beta\right)$. A notação para os 'valores' dos parâmetros α e β é:

- evolução unitária = EU;
- medida não-seletiva = MnS;
- medida seletiva = MS;
- para toda = $\forall$;
- existe ao menos uma = $\exists$.

Como exemplo, a classe de estados do sistema AB em que A não é afetada pela operação $(MnS,\exists)$ em B é $\Gamma_A\left(B,MnS,\exists\right)$. Com isso, define-se que, seja $\mathcal{D}$ o conjunto de todos os operadores densidade que atuam em $\mathcal{H}_{AB}$, se $\mathcal{D}\ni\rho\notin\Gamma_X\left(Y,\alpha,\beta\right)$, o sistema X está correlacionado com o sistema Y pela operação (α,β).

Uma importante constação é que um estado de um sistema AB possui discórdia quântica nula se, e somente se, pertencer à classe $\Gamma_{AB}\left(B,MnS,\exists\right)$ [19].

Um primeiro problema é saber quantas classes existem. Como um sistema AB possui espaço de Hilbert dado por $\mathcal{H}_A\otimes\mathcal{H}_B$ ou, equivalentemente, por $\mathcal{H}_B\otimes\mathcal{H}_A$, as classes que possuem como parâmetros A ou B na mesma posição são equivalentes. Com base nessa consideração, pode-se facilmente calcular que há 18 classes de carrelação. As mais simples de caracterizar são, aparentemente, as que envolvem o operador de evolução unitária. A primeira classe a ser caracterizada será $\Gamma_{AB}\left(B,EU,\exists\right)$.

Sejam $\{|a^k\rangle\}$ e $\{|b^l\rangle\}$ bases de $\mathcal{H}_A$ e $\mathcal{H}_B$ respectivamente,

$$\rho\equiv c_{k,l}|a^k\rangle|b^l\rangle \tag{4.3}$$

Seja $U \equiv \mathrm{e}^{\mathrm{i}\,\theta}\, I$ uma matriz unitária que atua em $\mathcal{H}_B$, é óbvio que $(I \otimes U)\,\rho\left(I \otimes U^{\dagger}\right) = \rho$, logo $\Gamma_{AB}\,(B, EU, \exists) = \mathcal{D}$. Pelo mesmo argumento, $\Gamma_B\,(AB, EU, \exists) = \mathcal{D}$. Como $\mathrm{tr}_B\,(\rho) = \mathrm{tr}_B\left[(I \otimes U)\,\rho\left(I \otimes U^{\dagger}\right)\right]$, $\Gamma_A\,(B, EU, \exists) = \Gamma_A\,(B, EU, \forall) = \mathcal{D}$.

Uma classe interessante é $\Gamma_{AB}\,(B, MS, \exists)$, pois seus elementos são os estados que, após uma medida seletiva específica em um subsistema, permanecem os mesmos, ou seja, são estados em que é possível se fazer uma medida local seletiva sem que haja qualquer colapso. Os cálculos abaixo tem o objetivo de caracterizar essa classe.

Uma operação que será utilizada posteriormente está definida abaixo:

Definição 4.2.1. *Sejam $\mathcal{M}_{n\times k}$ e $\mathcal{M}_{n\times l}$ os conjuntos das matrizes $n\times k$ e $n\times l$ respectivamente. A operação de concatenação horizontal $\oplus\colon \mathcal{M}_{n\times k} \times \mathcal{M}_{n\times l} \longrightarrow \mathcal{M}_{n\times(k+l)}$ é dada por $A \oplus B \equiv [AB]$, em que $A \in \mathcal{M}_{n\times k}$ e $B \in \mathcal{M}_{n\times l}$. A operação de concatenação vertical $\ominus\colon \mathcal{M}_{k\times n} \times \mathcal{M}_{l\times n} \longrightarrow \mathcal{M}_{(k+l)\times n}$ é dada por $A \ominus B \equiv \begin{bmatrix} A \\ B \end{bmatrix}$, em que $A \in \mathcal{M}_{k\times n}$ e $B \in \mathcal{M}_{l\times n}$.*

A operação de concatenação não é comutativa, não é distributiva, mas é associativa:

- $(A \oplus B) \oplus C = A \oplus (B \oplus C)$ e $(A \ominus B) \ominus C = A \ominus (B \ominus C)$
- $\forall A \in \mathcal{M}_{m\times n}, A \equiv [a_{i,j}]_{m\times n} = \oplus_{j=1}^{n}\left(\ominus_{i=1}^{m}\,[a_{i,j}]\right)$

4.2.1 Equação $AX = A$

Agora vou encontrar a solução para a seguinte equação: sejam $A \in \mathcal{M}_{m\times n}$ e $X \in \mathcal{M}_{n\times n}$,

$$AX = A, \tag{4.4}$$

em que X é a variável. Suponha que $\mathrm{rank}(A) = n$, sendo $m \geq n$. Isso implica que existe uma inversa à esquerda $A_{esq.}^{-1} = \left(A^T A\right)^{-1} A^T\colon A_{esq.}^{-1} A = I \Rightarrow X = I$, que é a única solução. Se $\mathrm{rank}(A) = k < n$, sendo $m \leq n$, não existe inversa à esquerda. Com isso, apesar de $X = I$ ser uma solução, não se sabe se ela é a única ou não. Seja $\vec{a}_j = \begin{bmatrix} a_{1,j} \\ \vdots \\ a_{m,j} \end{bmatrix}, i \in \{1, \ldots, m\}$, em que $A = [a_{i,j}]_{m\times n}$ tem-se $A = \oplus_{j=1}^{n} \vec{a}_j$. Pelo rank de A, existem k vetores coluna que formam

um conjunto LI e o restante pode ser escrito como combinações lineares dos primeiros. Sejam $\gamma' \equiv \{\vec{a}_j : j \in J' \subset J\}$ e $\gamma'' \equiv \{\vec{a}_l : l \in J'' = J - J'\}$ em que $J = \{1, \ldots, n\}$, $|\gamma'| = k$ e γ' é L.I.. Sejam $J' = \{y_1, \ldots, y_k\}$, $J'' = \{y_{k+1}, \ldots, y_n\}$, E_{i,y_i} uma matriz que permuta a coluna i com a y_i se multiplicada pela esquerda ou que permuta a linha i com a y_i se multiplicada pela direita. Com isso, pode-se observar que resolver a equação (4.4) é equivalente a resolver a equação:

$$A'X' = A' \prod_{i=1}^{n} E_{i,y_i}, \tag{4.5}$$

em que $A' \equiv A \prod_{i=1}^{n} E_{i,y_i}$ e $X' \equiv \left(\prod_{i=1}^{n} E_{i,y_i}\right) X$. Pode-se escrever os $n-k$ vetores coluna de A' da seguinte forma: $\vec{a}'_{k+j} = \sum_{i=1}^{k} \alpha_{i,j} \vec{a}'_i$. Sejam $A'' \equiv \oplus_{i=1}^{k} \vec{a}'_i$, $\alpha \equiv [\alpha_{i,j}]_{k \times (n-k)}$ e $\alpha' \equiv [I\alpha]$, em que $I \in \mathcal{M}_{k \times k}$,

$$A' = A''\alpha' \tag{4.6}$$

Como $\text{rank}(A'') = k$, existe uma inversa à esquerda $(A'')^{-1}_{esq.}$ de A''. Aplicando essa inversa nos dois lados da Equação (4.5), tem-se:

$$\alpha' X' = \alpha' \prod_{i=1}^{n} E_{i,y_i}. \tag{4.7}$$

Antes de buscar uma solução diferente da identidade, é preciso saber se ela existe. É fácil observar que (4.7) é equivalente a:

$$(I \otimes \alpha') \,\text{vec}\,(X) = \text{vec}\left(\alpha' \prod_{i=1}^{n} E_{i,y_i}\right), \tag{4.8}$$

em que $I \in \mathcal{M}_{n \times n}$. Como $\text{vec}\,(X') = \text{vec}\left(\prod_{i=1}^{n} E^L_{i,y_i}\right)$ é solução de (4.8), a equação é consistente. Como o número de colunas de $I \otimes \alpha'$ é maior que o de linhas, há infinitas soluções. Voltando à (4.7), sejam $\vec{x'}_i, i \in \{1, \ldots, n\}$ vetores linha de X, $X' = \ominus_{i=1}^{n} \vec{x'}_i \Rightarrow \alpha' X' = \sum_{i=1}^{n} \left[\vec{\alpha'_i} \vec{x'_i}\right]$. Pode-se tomar as $n-k$ últimas linhas de X' e definir que seus $(n-k)k$ coeficientes serão variáveis livres. A equação

$$\sum_{i=1}^{k} \left[\vec{\alpha'_i} \vec{x'_i}\right] = \alpha' \prod_{i=1}^{n} E_{i,y_i} - \sum_{i=1}^{n} \left[\vec{\alpha'_i} \vec{x'_i}\right] \tag{4.9}$$

é a Equação (4.7), mas com as variáveis dependentes no lado esquerdo e as livres no direito. Sejam $\beta \equiv \oplus_{i=1}^{k} \vec{\alpha}'_i$, $Y \equiv \ominus_{i=1}^{k} \vec{x}'_i$, $\beta' \equiv \oplus_{i=k+1}^{n} \vec{\alpha}'_i$ e $Y' \equiv \ominus_{i=k+1}^{n} \vec{x}'_i$, a Equação (4.9) pode ser

reescrita, portanto, como

$$\beta Y = \alpha' \prod_{i=1}^{n} E_{i,y_i} - \beta' Y'. \tag{4.10}$$

Uma vez que $\beta = I$,

$$Y = \alpha' \prod_{i=1}^{n} E_{i,y_i} - \beta' Y'. \tag{4.11}$$

A solução de (4.7) é, portanto:

$$X' = \begin{bmatrix} \alpha' \prod_{i=1}^{n} E_{i,y_i} - \beta' Y' \\ Y' \end{bmatrix} \Rightarrow X = \left(\prod_{i=1}^{n} E_{i,y_i} \right) \begin{bmatrix} \alpha' \prod_{i=1}^{n} E_{i,y_i} - \beta' Y' \\ Y' \end{bmatrix}. \tag{4.12}$$

4.2.2 Classe $\Gamma_{AB}(B, MS, \exists)$

A pergunta que se pretende responder é: que estados pertencem a $\Gamma^{p}_{AB}(B, MS, \exists)$? O índice p significa que a classe está restrita a estados puros. Essa pergunta pode ser expressa na forma de uma outra: para que tipo de estado $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ existe V tal que

$$(I \otimes V) P_\psi \left(I \otimes V^\dagger \right) = |\lambda|^2 P_\psi, 0 < |\lambda| \leq 1, \tag{4.13}$$

e $V \in M$ para algum conjunto M de medidas POVM que atua em $\mathcal{H}_B$, dados $\dim \mathcal{H}_A = m$, $\dim \mathcal{H}_B = n$, I atua em $\mathcal{H}_A$? A restrição de λ vem do fato de que $|\lambda|^2 = \mathrm{tr}\left(V P_\psi V^\dagger\right)$, que é a probabilidade de a medida V ser selecionada. Se existe uma solução para a equação

$$(I \otimes V) |\psi\rangle = \lambda' |\psi\rangle, \tag{4.14}$$

em que $|\lambda'|^2 = |\lambda|^2$, também existe para a equação (4.13), pois suponha que $\forall |\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ e $\forall V, (I \otimes V)|\psi\rangle = |\psi'_V\rangle \neq \lambda'|\psi\rangle \Rightarrow (I \otimes V) P_\psi \left(I \otimes V^\dagger \right) = |\psi'_V\rangle\langle\psi'_V|$. Deve-se ter $|\psi'_V\rangle\langle\psi'_V| = |\lambda|^2 P_\psi$ somente se $|\psi'_V\rangle = \gamma|\psi\rangle$, onde $|\gamma|^2 = |\lambda|^2$. Mas se isso ocorre, então (4.14) possui solução, o que contraria a hipótese. Logo, não existe γ tal que $|\psi'_V\rangle = \gamma|\psi\rangle$ e, portanto, $|\psi'_V\rangle\langle\psi'_V| \neq |\lambda|^2 P_\psi$. Sejam $\{|\alpha_i\rangle : i \in \{1, \ldots, m\}\}$ e $\{|\beta_j\rangle : j \in \{1, \ldots, n\}\}$ bases para $\mathcal{H}_A$ e $\mathcal{H}_B$ respectivamente. Tem-se então que $|\psi\rangle = \sum_{i,j} c_{i,j} |\alpha_i\rangle|\beta_j\rangle$ e $V = \sum_{r,s} v_{r,s} |\beta_r\rangle\langle\beta_s|$. Trocando-se λ' por λ em (4.14), tem-se:

$$c_{i,j} v_{r,s} |\alpha^i\rangle|\beta^r\rangle\langle\beta^s|\beta^j\rangle = \lambda c_{i',j'} |\alpha^{i'}\rangle|\beta^{j'}\rangle. \tag{4.15}$$

Na equação acima foi usada a notação de Einstein, a qual se continuará a usar. De (4.15), tem-se

$$\sum_r c_{i,r} v_{j,r} |\alpha^i\rangle|\beta^j\rangle = \lambda c_{i',j'} |\alpha_{i'}\rangle|\beta_{j'}\rangle \Leftrightarrow \left[CV^T\right]_{i,j} |\alpha^i\rangle|\beta^j\rangle = \lambda \left[C\right]_{i,j} |\alpha^i\rangle|\beta^j\rangle \Leftrightarrow CV^T = \lambda C, \tag{4.16}$$

em que C é a matriz $\left[c_{i,j}\right]_{m\times n}$. Dado C, a solução é, de acordo com a equação (4.12), fazendo $E \equiv \prod_{i=1}^{n} E_{i,y_i}$,

$$V = \left(E \begin{bmatrix} \lambda\alpha' E - \beta' Y' \\ Y' \end{bmatrix}\right)^T = \begin{bmatrix} \lambda E\alpha'^T - Y'^T\beta'^T & Y'^T \end{bmatrix} E. \tag{4.17}$$

Essa solução, porém, não leva em consideração a condição de que $V \in M$, a qual implica que deve-se ter

$$V^\dagger V \leq I, V^\dagger V \neq I. \tag{4.18}$$

Agora vou mostrar que sempre é possível, para $V \neq \lambda I$, construir Y' e escolher λ de tal modo que (4.18) seja satisfeita.

$$V^\dagger V = E \begin{bmatrix} |\lambda|^2\alpha'^*\alpha'^T - \lambda^*\alpha'^* E Y'^T\beta'^T - \lambda\beta'^* Y'^* E\alpha'^T - \beta'^* Y'^* Y'^T\beta'^T & \lambda^*\alpha'^* E Y'^T - \beta'^* Y'^* Y'^T \\ \lambda^* Y'^* E\alpha'^* E\alpha'^T - Y'^* Y'^T\beta'^T & Y'^* Y'^T \end{bmatrix} E. \tag{4.19}$$

Sejam $\Theta_1 \equiv \alpha'^*\alpha'^T$, $\Theta_2 \equiv \alpha'^* E Y'^T\beta'^T$, $\Theta_3 \equiv \beta'^* Y'^* E\alpha'^T$, $\Theta_4 \equiv Y'^* E\alpha'^* E\alpha'^T$, $\Theta_5 \equiv \alpha'^* E Y'^T$, $\Omega_1 \equiv \beta'^* Y'^* Y'^T\beta'^T$, $\Omega_2 \equiv Y'^* Y'^T\beta'^T$, $\Omega_3 \equiv \beta'^* Y'^* Y'^T$ e $\Omega_4 \equiv Y'^* Y'^T$. Restringindo o valor de λ a $\lambda \in \mathbb{R}$, tem-se, de (4.19),

$$V^\dagger V = E \begin{bmatrix} \lambda^2\Theta_1 - \lambda\Theta_2 - \lambda\Theta_3 - \Omega_1 & \lambda\Theta_5 - \Omega_3 \\ \lambda\Theta_4 - \Omega_2 & \Omega_4 \end{bmatrix} E \tag{4.20}$$

Com a base fixada, sejam θ_i e ω_j as componentes de maior módulo de Θ e Ω, respectivamente, ou seja, $\theta_i = \left[\Theta\right]_{r',s'}, |\theta_i| \geq \left[\Theta\right]_{r,s}, \forall(r,s), \forall i$, para algum par (r', s') e $\omega_j = \left[\Omega\right]_{p',q'}, |\omega_j| \geq \left[\Omega\right]_{p,q}, \forall(p,q), \forall j$, para algum par (p', q'). Sejam $F_1 \equiv \lambda^2\Theta_1 - \lambda\Theta_2 - \lambda\Theta_3 - \Omega_1$, $F_2 \equiv \lambda\Theta_4 - \Omega_2$,

$F_3 \equiv \lambda\Theta_5 - \Omega_3$ e $F_4 \equiv \Omega_4$. Seja f_i a componente de maior módulo de F_i, $\forall i$. Tem-se que

$$\begin{aligned} |f_1| &\leq \lambda^2|\theta_1| + \lambda|\theta_2| + \lambda|\theta_3| + |\omega_1| \\ |f_2| &\leq \lambda|\theta_5| + |\omega_3| \\ |f_3| &\leq \lambda|\theta_4| + |\omega_2| \\ |f_4| &= |\omega_4| \end{aligned} \tag{4.21}$$

Observa-se que

$$[\Omega_1]_{p,q} = \sum_{u,l,k} b'^*_{p,k} y'^*_{k,u} y'^*_{l,u} b'_{q,l} \Rightarrow |\,[\Omega_1]_{p,q}\,| \leq \sum_{u,l,k} |b'^*_{p,k} b'_{q,l}||y'^*_{k,u} y'^*_{l,u}|. \tag{4.22}$$

Seja $\frac{y'}{\sqrt{n}} \equiv |y'_{i',j'}|$ para algum par (i', j') tal que $\frac{y'}{\sqrt{n}} \geq |y'_{i,j}|$, $\forall i, j$. Da equação acima,

$$|\,[\Omega_1]_{p,q}\,| \leq (y')^2 \sum_{l,k} |b'^*_{p,k} b'_{q.l}|. \tag{4.23}$$

Seja $b'_1 \equiv \sum_{l,k} |b'^*_{p',k} b'_{q',l}| \geq \sum_{l,k} |b'^*_{p,k} b'_{q.l}|$, $\forall p, q$ e para algum par (p', q'), isso implica que

$$|\,[\Omega_1]_{p,q}\,| \leq h_1 \equiv (y')^2 b'_1 \tag{4.24}$$

Também se observa que

$$[\Omega_2]_{p,q} = \sum_{k,l} y'^*_{p,k} y'_{l,k} b'_{q,l} \Rightarrow |\,[\Omega_2]_{p,q}\,| \leq (y')^2 \sum_{l} |b'_{q,l}|. \tag{4.25}$$

Seja $b'_2 \equiv \sum_l |b'_{q',l}| \geq \sum_l |b'_{q,l}|$, $\forall q$ para algum q',

$$|\,[\Omega_2]_{p,q}\,| \leq h_2 \equiv (y')^2 b'_2 \tag{4.26}$$

$$[\Omega_3]_{p,q} = \sum_{l,k} b'^* y'^*_{k,l} y'_{q,l} \Rightarrow |\,[\Omega_3]_{p,q}\,| \leq (y')^2 \sum_{l} |b'^*_{p,k}|. \tag{4.27}$$

Seja $b'_3 \equiv \sum_k |b'^*_{p',k}| \geq \sum_l |b'^*_{p,k}|$, $\forall p$ para algum p',

$$|\,[\Omega_3]_{p,q}\,| \leq h_3 \equiv (y')^2 b'_3 \tag{4.28}$$

$$[\Omega_4]_{p,q} = \sum_{l} y'^*_{p,l} y'_{q,l} \Rightarrow |\,[\Omega_4]_{p,q}\,| \leq (y')^2 \tag{4.29}$$

Das Equações (4.24), (4.26), (4.28) e (4.29), verifica-se que sempre é possível escolher um valor de y' que satisfaça, simultaneamente, as inequações:

$$h_1 < \frac{1}{4n}, \quad h_1 < \frac{1}{2n}, \quad h_1 < \frac{1}{2n}, \quad h_1 < \frac{1}{n} \tag{4.30}$$

Após escolher y', escolhe-se λ tal que

$$\lambda^2|\theta_1| < \frac{1}{4n}, \quad \lambda|\theta_2| < \frac{1}{4n}, \quad \lambda|\theta_3| < \frac{1}{4n}, \quad \lambda|\theta_4| < \frac{1}{2n}, \quad \lambda|\theta_5| < \frac{1}{2n} \tag{4.31}$$

De (4.30) e (4.31), sendo γ a matrix em (4.19) que está multiplicada à direita e à esquerda por E e $\sigma_{max}\,(A)$ a função que dá o maior autovalor de uma matrix A qualquer,

$$\begin{aligned} |f_i| < \frac{1}{n}, \forall i \quad &\Rightarrow \quad \mathrm{tr}\left(V^\dagger V\right) = \mathrm{tr}\,(\gamma) < 1 \Rightarrow \sigma_{max}\left(V^\dagger V\right) < 1 \Rightarrow \langle x|V^\dagger V|x\rangle < \langle x|x\rangle, \forall |x\rangle \\ &\Rightarrow \quad V^\dagger V < I \end{aligned} \tag{4.32}$$

A equação acima mostra que todo estado puro associado a uma matriz $C \in \mathcal{M}_{n\times m}, m < n$, de rank menor que m, existe uma medida seletiva não trivial sob a qual o estado é invariante. Isso significa que sistemas puros $N \times 2$ são invariantes por medida seletiva não trivial se, e somente se, forem separáveis. Agora, considere uma matriz 3×3 $C \equiv [c_{i,j}]_{3\times 3}$ tal que $\mathrm{rank}\, C = 2$. Suponha, por simplicidade, que os vetores coluna $\vec{c}_1$ e $\vec{c}_2$ formem um conjunto L.I.. Tem-se então, usando (4.16):

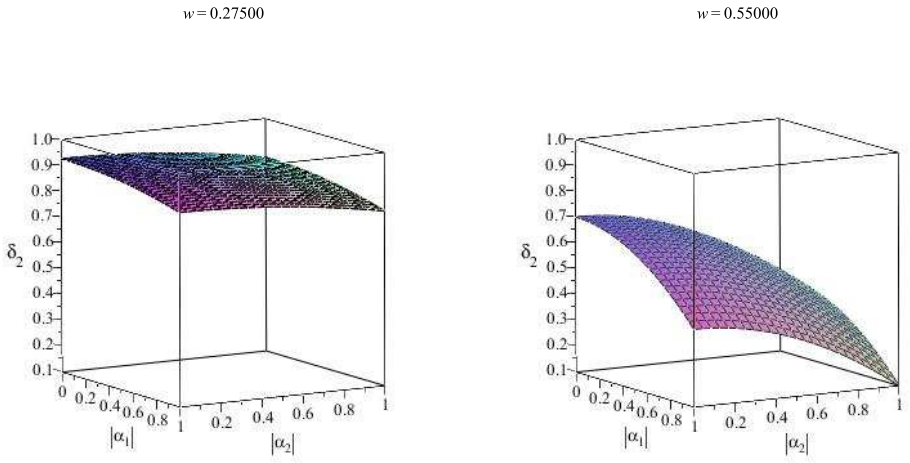
$$\begin{bmatrix} c_{11} & c_{12} & \alpha_1 c_{11} + \alpha_2 c_{12} \\ c_{21} & c_{22} & \alpha_1 c_{21} + \alpha_2 c_{22} \\ c_{31} & c_{32} & \alpha_1 c_{31} + \alpha_2 c_{32} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \\ c_{31} & c_{32} \end{bmatrix} \begin{bmatrix} 1 & 0 & \alpha_1 \\ 0 & 1 & \alpha_2 \end{bmatrix} \Leftrightarrow C'\alpha V^T = \lambda C'\alpha, \tag{4.33}$$

em que $C' \equiv [\vec{c}_1 \vec{c}_2]$ e $\alpha \equiv \begin{bmatrix} 1 & 0 & \alpha_1 \\ 0 & 1 & \alpha_2 \end{bmatrix}$. Como C' possui inversa à esquerda,

$$\alpha V^T = \lambda \alpha \tag{4.34}$$

A equação acima implica no seguinte sistema de equações:

$$\begin{aligned} &v_{11} + \alpha_1 v_{13} = \lambda \quad v_{21} + \alpha_1 v_{23} = 0 \quad v_{31} + \alpha_1 v_{33} = \lambda\alpha_1 \\ &v_{12} + \alpha_2 v_{13} = 0 \quad v_{22} + \alpha_2 v_{23} = \lambda \quad v_{32} + \alpha_2 v_{33} = \lambda\alpha_2, \end{aligned} \tag{4.35}$$



(a) (b)

Figura 4.1: Autovalor de $I - V^\dagger V$

as quais definem V como sendo

$$V = \begin{bmatrix} \lambda - \alpha_1 v_1 & -\alpha_2 v_2 & v_1 \\ -\alpha_1 v_2 & \lambda - \alpha_2 v_2 & v_2 \\ \alpha_1 (\lambda - v_3) & \alpha_2 (\lambda - v_3) & v_3 \end{bmatrix}, \quad v_i \equiv v_{i3}, i \in \{1, 2, 3\}. \tag{4.36}$$

Escolhendo $v_i = 0, \forall i$,

$$V = \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ \alpha_1 \lambda & \alpha_2 \lambda & 0 \end{bmatrix} \Rightarrow I - V^\dagger V = \begin{bmatrix} 1 - |\lambda|^2 (1 + |\alpha_1|^2) & \alpha_1^* \alpha_2 |\lambda|^2 & 0 \\ \alpha_2^* \alpha_1 |\lambda|^2 & 1 - |\lambda|^2 (1 + |\alpha_1|^2) & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{4.37}$$

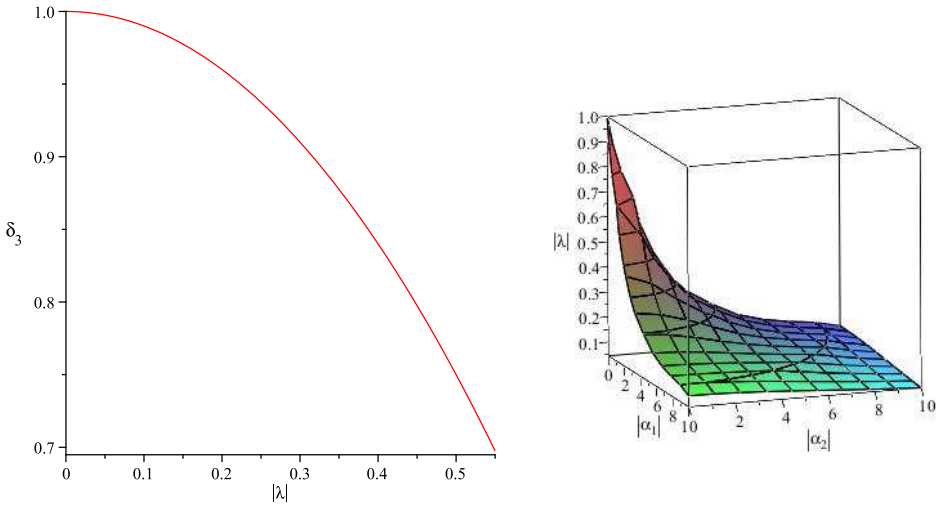
Os autovalores da matrix acima são:

$$\delta_1 = 1, \quad \delta_2 = 1 - |\lambda|^2 \left(1 + |\alpha_1|^2 + |\alpha_2|^2\right), \quad \delta_3 = 1 - |\lambda|^2, \tag{4.38}$$

Nos Gráficos (4.1a) e (4.1b), $w \equiv |\lambda|$. A Figura (4.2b) é o gráfico de $\delta_2 = 0$. Qualquer ponto acima daquela superfície representa um autovalor negativo. Logo, ela limita superiormente os valores do terno ordenado $(|\lambda|, |\alpha_1|, |\alpha_2|)$.

4.3 Entropia Generalizada

O resultado abaixo mostra como a entropia de von Neumann pode ser obtida diretamente da entropia de Shannon:



(a) (b)

Figura 4.2: (a) Autovalor de $I - V^{\dagger}V$ e (b) $\delta_2 = 0$

Teorema 4.3.1. *Seja* $\{P_k\}$ *um conjunto completo de medidas projetivas unidimensionais ortogonais,* $p_k = \mathrm{tr}(P_k\rho)$ *a probabilidade de se medir* P_k *em um sistema no estado* ρ *e* p *a medida de probabilidade tal que* $p(\{P_k\}) = p_k$ *tem-se:*

$$S(\rho) = \min_{\{P_k\}} H(p), \tag{4.39}$$

em que o mínimo é tomado em todos os possíveis conjuntos $\{P_k\}$.

Demonstração. Sejam $\rho^D = \sum_k P_k\rho P_k$, $P_k = |p_k\rangle\langle p_k|$ e $\rho = \sum_l \lambda_l|\rho_l\rangle\langle\rho_l|$,

$$p_k = \mathrm{tr}(P_k\rho) = \mathrm{tr}\left(\sum_l \lambda_l\langle\rho_l|p_k\rangle|\rho_l\rangle\langle p_k|\right) = \sum_l \lambda_l|\langle\rho_l|p_k\rangle|^2 \tag{4.40}$$

$$S(\rho^D) = S\left(\sum_k\left(\sum_l \lambda_l|\langle\rho_l|p_k\rangle|^2\right)P_k\right) = S\left(\sum_k p_k P_k\right). \tag{4.41}$$

De acordo com (3.1.4), $S(\rho^D) = H(p_k)$. Logo, pelo teorema (3.1.6),

$$S(\rho) \leq H(p), \tag{4.42}$$

sendo que, pela Proposição 3.1.1, a igualdade ocorre para $P_k = |\rho_k\rangle\langle\rho_k|$. □

Esse teorema fornece uma interpretação da entropia quântica em função da clássica. Suponha que se deseje conhecer qual é o estado de um sistema quântico. Para isso, é necessário

medi-lo. Como a medida é um evento aleatório, existe uma incerteza sobre qual será a saída obtida. Existe, entretanto, uma arbitrariedade na escolha de qual medida será efetuada. É preferível que se escolha uma que minimize a incerteza sobre o resultado. É conveniente, portanto, escolher essa incerteza mínima como sendo a entropia do sistema quântico. O teorema, então, mostra que a entropia de von Neumann é a escolha certa dentro do domínio das medidas projetivas.

A partir da Proposição 4.3.1, questiona-se, então, o que acontece se o mínimo for tomado entre todas as medidas POVM possíveis, ao invés de usar a restrição dos projetores unidimensionais. Primeiramente, qual a interpretação física de uma generalização como essa, dada pela equação abaixo?

$$S_E(\rho) \equiv \inf_{M_n^E} H(p), \tag{4.43}$$

em que p é a medida de probabilidade tal que $p(\{E_k\}) \equiv p_k \equiv \mathrm{tr}\left(E_k \rho E_k^\dagger\right)$, $\{E_k^\dagger E_k\}$ é um conjunto de medidas POVM, M_n^E é o conjunto de todos os conjuntos de medidas POVM que atuam em um estado ρ de um sistema de dimensão n. Pode-se interpretar $S_E(\rho)$ como a menor entropia dos dados de saída de uma medida efetuada em ρ considerando todas as possíveis teoricamente. O problema dessa definição é que pode ocorrer que, após uma medida, $E_k^\dagger E_k$ por exemplo, o estado ser colapsado em um que não seja puro, ou seja, um estado que ainda carrega informação consigo. Dessa forma, seria possível obter $S(\rho) > S_E(\rho)$ com $S(E_k \rho E_k^\dagger) > 0$ para algum k, mas através de medidas que não extraem toda a informação possível. Para evitar isso, será utilizada, ao invés da classe M_E a seguinte classe:

$$M_n^{(1)}(\rho) \equiv \{\{E_k^\dagger E_k\} \in M_n^E : \mathrm{rank}\left(E_k \rho E_k^\dagger\right) = 1, \forall k \in K\}, \tag{4.44}$$

em que K é o conjunto de índices associado a $\{E_k^\dagger E_k\}$. A redefinição da entropia é dada por:

$$S_E^{(1)}(\rho) \equiv \inf_{M_n^{(1)}(\rho)} H(p). \tag{4.45}$$

Seja $P_n^{(1)}$ o conjunto de todos os conjuntos de medidas projetivas de rank 1 que atuam em um sistema de dimensão n, será que $P_n^{(1)} \neq M_n^{(1)}(\rho)$? Obviamente, $P_n^{(1)} \subseteq M_n^{(1)}(\rho)$. De acordo

com [38], como ρ é representado por uma matrix não-negativa hermitiana,

$$\operatorname{rank}\left(E_k \rho E_k^\dagger\right) \geq \operatorname{rank}(E_k) + \operatorname{rank}(\rho) - n \quad \forall k \in K, \tag{4.46}$$

em que ρ e E_k são matrizes quadradas de dimensão n. Se $\operatorname{rank}(\rho) = 1$, o valor máximo de $\operatorname{rank}(E_k)$ é $n-1$. Porém, nesse caso, não é necessário verificar se $P_n^{(1)} = M_n^{(1)}(\rho)$, pois como $S_E^{(1)}(\rho) \geq 0$, $S(\rho) = 0 \Rightarrow S_E^{(1)}(\rho) = 0$. Se $\operatorname{rank}\rho = n$, obrigatoriamente $\operatorname{rank}(E_k) = 1$. Nesse caso, pode-se escrever $E_k = |\psi\rangle\langle\phi| \Rightarrow E_k^\dagger E_k = |\phi\rangle\langle\phi|$, ou seja, a medida é necessariamente projetiva. Mas e se $1 < \operatorname{rank}(\rho) < n$?

Seja $\{E_k^\dagger E_k\}$ um conjunto POVM, $E_k \in M_{n\times n}$. Qualquer que seja $U \in SU(n)$, em que $SU(n)$ é o grupo especial unitário,

$$\sum_k (E_k U)^\dagger E_k U = U^\dagger \left(\sum_k E_k^\dagger E_k\right) U = I, \tag{4.47}$$

ou seja, uma trasformação de $SU(n)$ em um conjunto POVM leva a outro conjunto POVM. Levando isso em consideração, seja U a matrix mudança de base que leva ρ em sua forma diagonal,

$$E_k \rho E_k^\dagger = E_k U \rho' U^\dagger E_k^\dagger = E_k' \rho' E_k'^\dagger, \tag{4.48}$$

em que $\rho' \equiv \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Como o número de autovalores não nulos de ρ é igual a seu rank, deve haver pelo menos um autovalor nulo. Considere $\rho \in M_{3\times 3}$. Todo estado ρ de rank 2 pode ser escrito da forma:

$$\rho = \lambda_1 |\lambda_1\rangle\langle\lambda_1| + \lambda_3 |\lambda_3\rangle\langle\lambda_3|. \tag{4.49}$$

Seja $M \equiv \{E^\dagger E, F^\dagger F\}$ o conjunto de medida que será utilizado. A restrição $M \in M_3^{(1)}(\rho)$ implica que deve-se ter

$$E\rho E^\dagger = \nu |\psi\rangle\langle\psi| \quad \text{e} \quad F\rho F^\dagger = (1-\nu)|\phi\rangle\langle\phi| \tag{4.50}$$

para algum estado $|\psi\rangle$ e $|\phi\rangle$ e para $0 \leq \nu \leq 1$, sendo que essa última condição é necessária para que o postulado da medida seja satisfeito. Com isso, seja $|\psi\rangle \equiv \psi_1|\lambda_1\rangle + \psi_2\lambda_3\rangle$ e $E \equiv \sum_{i,j}^2 e_{i,j}|\lambda_i\rangle\langle\lambda_j|$, tem-se as seguintes equações:

$$\nu|\lambda_1|^2 = |e_{1,1}|^2\lambda_1, \quad \nu\psi_1\psi_2^* = e_{1,1}e_{2,1}^*\lambda_1, \quad \nu|\lambda_1|^2 = |e_{2,1}|^2\lambda_1 \tag{4.51}$$

Com as soluções

$$\psi_1 = e_{1,1}\sqrt{\frac{\lambda_1}{\nu}}, \quad \psi_2 = e_{2,1}\sqrt{\frac{\lambda_1}{\nu}}. \tag{4.52}$$

A normalização de $|\psi\rangle$ exige que

$$|e_{1,1}|^2 + |e_{2,1}|^2 = \frac{\nu}{\lambda_1}. \tag{4.53}$$

Pode-se tomar $F = |\lambda_3\rangle\langle\lambda_3|$. A condição $E^\dagger E + F^\dagger F = I$ implica que

$$|e_{1,1}|^2 + |e_{2,1}|^2 = 1 \Rightarrow \nu = \lambda_1, \quad |e_{1,2}|^2 + |e_{2,2}|^2 = 1, \quad e^*_{1,1}e_{1,2} + e^*_{2,1}e_{2,2} = 0. \tag{4.54}$$

Como as restrições impostas por meio das equações acima permitem que exista uma matriz E de rank=2 que as satisfaça. A matriz A dada por

$$A \equiv \begin{bmatrix} \frac{1}{3} & -\frac{2\sqrt{2}}{3} & 0 \\ \frac{2\sqrt{2}}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{4.55}$$

claramente satisfaz as equações em (4.54) e tem rank 2, logo $P_3^{(1)} \neq M_3^{(1)}(\rho)$. Esse resultado pode ser facilmente generalizado. Sejam $\rho_3 = \mathrm{diag}(\lambda_1, 0, \lambda_3)$, $O_k \in M_{k\times 3}$ uma matriz preenchida totalmente com zeros, $O'_k \in M_{k\times k}$ tal que $A'_k \equiv \mathrm{diag}(\lambda_4, \cdots, \lambda_{k+3})$ e $E_3 \equiv \begin{bmatrix} e_{1,1} & e_{1,2} & 0 \\ e_{2,1} & e_{2,2} & 0 \\ 0 & 0 & 0 \end{bmatrix}$, define-se:

$$E \equiv \begin{bmatrix} E_3 & O_k^T \\ O_k & O'_k \end{bmatrix}, \quad \rho \equiv \begin{bmatrix} \rho_3 & O_k^T \\ O_k & A'_k \end{bmatrix} \tag{4.56}$$

Essas definições implicam

$$E\rho E^\dagger = \begin{bmatrix} E_3 A_3 E_3^\dagger & O_k^T \\ O_k & O'_k \end{bmatrix} = \nu|\psi\rangle\langle\psi|. \tag{4.57}$$

Fazendo os outros elementos $F_j, j \in \{1, \cdots, k+1\}$ do conjunto POVM serem definidos por $F_j \equiv |\lambda_{j+2}\rangle\langle\lambda_{j+2}|$, tem-se que encontrar E que satisfaça 4.57 é equivalente a encontrar uma solução de (4.54). Logo,

$$P_n^{(1)} \neq M_n^{(1)}(\rho), \quad n \geq 3. \tag{4.58}$$

A restrição $n \geq 3$ vem do fato de que, em dimensão 2, a matriz E com $\mathrm{rank}\, E = 2$ não pode existir para $\mathrm{rank}\, \rho = 2$, uma vez que a inequação (4.46) deixa de ser satisfeita. Um

questionamento que pode ser feito é qual a forma geral E_k que seja solução de $E_k\rho E_k^\dagger = \nu_k|\psi_k\rangle\langle\psi_k|$. No caso de um sistema de dimensão 3 com o conjunto de medida $\{E^\dagger E, F^\dagger F\}$, sendo $E \equiv e_{i,j}|\lambda^i\rangle\langle\lambda^j|$, deve-se ter

$$\sum_{l\in\{1,3\}} e_{l,i}^* e_{l,j}\lambda_l = \nu\psi_i\psi_j^* \tag{4.59}$$

Uma possível solução para esse sistema de equações é dada por

$$\psi_i = \frac{e_{1,i}\sqrt{\lambda_1} + e_{3,i}^*\sqrt{\lambda_3}}{\sqrt{\nu}}, \quad \text{para} \quad e_{1,i}^* e_{2,j} + e_{3,i}^* e_{1,j} = 0, \tag{4.60}$$

Uma outra maneira de expressar a restrição acima é, definindo $H \equiv \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$,

$$E^\dagger H E = O, \tag{4.61}$$

em que O é a matriz nula. Seja $e_{r,s} = a_{r,s} + \mathrm{i}\, b_{r,s}$, fazendo $e_{1,s} = a_{1,s}(1+\mathrm{i})$ e $e_{3,s} = a_{3,s}(1-\mathrm{i})$ e $a_{1,r}a_{3,s} = a_{1,s}a_{3r}$, a equação (4.61) é satisfeita. A condição de normalização de $|\psi\rangle$ exige que

$$\lambda_1 \sum_{j=1}^{3} |e_{1,j}|^2 + \lambda_3 \sum_{j=1}^{3} |e_{3,j}|^2 = \nu. \tag{4.62}$$

Os resultados obtidos para E valem, de modo análogo, para F. A última condição a ser satisfeita é

$$E^\dagger E + F^\dagger F = I. \tag{4.63}$$

Há também uma imposição que não surge da busca da solução de (4.50), mas sim de um fato físico, que a exigência de que

$$E\rho E^\dagger \neq \alpha F\rho F^\dagger, \tag{4.64}$$

em que α é uma costante complexa de proporcionalidade, pois, caso haja a igualdade, as duas medidas retornarão o mesmo estado, sendo, portanto, totalmente indistinguíveis.

Seja $E = \ominus_{i=1}^{3}\vec{e}_i$, em que $\vec{e}_i$ são vetores correspondentes às linhas de E. Podemos escolher um produto de matrizes de permutação C de tal modo que $E' \equiv CE = \vec{e}_{y_1} \ominus \vec{e}_{y_1} \ominus (\alpha_1\vec{e}_{y_2} + \alpha_2\vec{e}_{y_2})$, em que $y_i \in \{1,2,3\}$. Seja $E'' \equiv \vec{e}_{y_1} \ominus \vec{e}_{y_1}$, $\alpha \equiv \begin{bmatrix}\alpha_1 & \alpha_2\end{bmatrix}$ e $\alpha' \equiv I \ominus \alpha$, então:

$$E = C\alpha' E''. \tag{4.65}$$

Como rank $E'' = 2$, ela possui inversa à esquerda e $E''^{\dagger}$ inversa à direta, logo

$$\alpha'^{\dagger} CHC\alpha' = 0 \tag{4.66}$$

Suponha o caso em que $C = I$:

$$\alpha'^{\dagger} H\alpha' = 0. \tag{4.67}$$

Da equação acima implica que

$$\Re(\alpha_1) = 0, \quad \alpha_2 = 0. \tag{4.68}$$

As condições acima são necessárias, porém não suficientes. Uma condição necessária e suficiente é fazer

$$\alpha_1 = \mathrm{e}^{\mathrm{i}(\frac{\pi}{2}+n\pi)}, \quad n \in \mathbb{Z} \tag{4.69}$$

Agora é preciso satisfazer a equação (4.63). Se ela é satisfeita, não é preciso impor condições sobre λ, pois elas são automaticamente satisfeitas. Seja

$$E = \begin{bmatrix} e_{1,1} & e_{1,2} & e_{1,3} \\ e_{2,1} & e_{2,2} & e_{2,3} \\ \mathrm{e}^{\mathrm{i}(\frac{\pi}{2}+n\pi)} e_{1,1} & \mathrm{e}^{\mathrm{i}(\frac{\pi}{2}+n\pi)} e_{1,2} & \mathrm{e}^{\mathrm{i}(\frac{\pi}{2}+n\pi)} e_{1,3} \end{bmatrix} \tag{4.70}$$

e

$$\mathcal{E} \equiv \begin{bmatrix} \sqrt{2}e_{1,1} & \sqrt{2}e_{1,2} & \sqrt{2}e_{1,3} \\ e_{2,1} & e_{2,2} & e_{2,3} \end{bmatrix}, \tag{4.71}$$

tem-se

$$E^{\dagger}E = \mathcal{E}^{\dagger}\mathcal{E} \tag{4.72}$$

Seja $\mathcal{F}$ definido de forma análoga à definição de $\mathcal{E}$, então a solução da equação

$$\mathcal{E}^{\dagger}\mathcal{E} + \mathcal{F}^{\dagger}\mathcal{F} = I \tag{4.73}$$

determina a solução de (4.63). Outra implicação é que

$$\lambda = \mathrm{tr}(\mathcal{E}\rho\mathcal{E}^{\dagger}). \tag{4.74}$$

Agora, o problema se reduz a encontrar λ sob as condições (4.73) e (4.64). A constante $\sqrt{2}$ pode ser omitida das matrizes para que se possa usar matrizes genéricas 2×3 X e Y e trabalhar na solução de

$$X^{\dagger}X + Y^{\dagger}Y = I. \tag{4.75}$$

Capítulo 5

Conclusão

Neste trabalho, procurou-se apresentar o tema das correlações quânticas e da medida de informação quântica. Foi discutido como as correlações quânticas foram identificadas por meio de seu caráter sem análogo clássico e que isso deu início ao estudo do emaranhamento, um tipo de correlação mais forte do que seria possível pela Física Clássica. Depois, foi mostrado que havia um tipo de correlação quântica que ocorre em estado separáveis, porém ainda assim mais forte do que a clássica e que ela pode ser medida através de uma quantidade conhecida como discórdia quântica, a qual possui interessantes propriedades. Foi por meio de uma dessas propriedades, a de que estados de discórdia nula são invariantes por medidas locais não seletivas, que desenvolvi uma generalização do conceito de correlação quântica. Por meio desse conceito, as correlações foram dividas em classes e algumas dessas foram caracterizadas, sendo que uma em especial revelou ter ligações com o emaranhamento em estados puros. Em um outro caminho, foi apresentado o conceito de entropia como uma medida da quantidade de informação, mostrando seu significado em teoria da informação. Também foi discutida a versão quântica dessa entropia, a entropia de von Neumann, comparando-a com a anterior. Foi a partir dessa comparação que mostrou-se como a clássica dá origem à quântica e, indo além, fez uma generalização desta a partir da classe mais geral de medidas quânticas. Observou-se que essa classe deveria restringir-se a medidas que colapsassem o estado do sistema em um que fosse puro. Por fim, obteve-se que essa classe é mais abrangente do que a classe das medidas projetivas, as quais são usadas para se obter a entropia de von Neumann.

Muito ainda há o que se trabalhar a partir dessas duas generalizações, como caracterizar todas as classes de correlação e obter as propriedades da entropia generalizada, comparando-a com a de von Neumann, buscar aplicações para ambas e verificar suas implicações em mecânica quântica e teoria de informação quântica.

Bibliografia

[1] H. Nyquist, Bell System Technical Journal **324** (1924).

[2] R. V. L. Hartley, Bell System Technical Journal **535** (1928).

[3] C. E. Shannon, Bell System Technical Journal **27**, 379 (19480).

[4] J. von Neumann, *Mathematische Grundlagen der Quantenmechanic* (Springer, Berlin 1932).

[5] B. Schumacher, Phys. Rev. A **51**, 2738 (1995).

[6] B. Schumacher e M. B. Westmoreland, Phys. Rev. A **56(1)**, 131 (1997).

[7] A. S. Holevo, IEEE. Trans. Inf. Theory **44(1)**, 269 (1998).

[8] S. Wiesner, SIGACT News **15**, 77 (1983).

[9] N. Gisin, G. Ribordy, W. Tittel e H. Zbinden, Rev. Mod. Phys. **74**, 145 (2002).

[10] R. Horodecki, P. Horodecki, M. Horodecki, K. Horodecki, Rev. Mod. Phys. **81**, (2009).

[11] J. S. Bell, Physics **1**, 195 (1964).

[12] P. W. Shor, Phys. Rev. A **52**, 2493 (1995).

[13] A. M. Steane, Phys. Rev. Lett. **77**, 793 (1996).

[14] D. Gottesman, arXiv:quant-ph/9705052 (1997).

[15] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres e W. K. Wootters, Phys. Rev. Lett. **70**, 1895 (1993).

[16] C. H. Bennett e S. J. Wiesner, Phys. Rev. Lett. **69**, 2881 (1992).

[17] A. Datta, A. Shaji e C. M. Caves, Phys. Rev. Lett. **100**, 050502 (2008).

[18] A. Datta, A. Shaji, Int. Quant. Inf. **9**, 1787 (2011).

[19] H. Ollivier e W. H. Zurek, Phys. Rev. Lett. **88**, 017901 (2001).

[20] L. Henderson e V. Vedral, J. Phys. A: Math. Gen. **34**, 6899 (2001).

[21] A. Shabani e D. A. Lidar, Phys. Rev. Lett. **102**, 100402 (2009).

[22] R. Dillenschneider, Phys. Rev. B **78**, 224413 (2008)

[23] M. S. Sarandy, Phys. Rev. A **80**, 022108 (2009).

[24] M. Allegra, P. Giorda, A. Montorsi, Phys. Rev. B **84**, 245133 (2011)

[25] A. Ferraro, L. Aolita, D. Cavalcanti, F. M. Cucchietti, A. Acin, Phys. Rev. A **81**, 052318 (2010).

[26] A. Einstein, B. Podolsky, N. Rosen, Phys. Rev. **47**, 777 (1935).

[27] R. F. Werner, Phys. Rev. A **40**, 4277 (1989b).

[28] M. Mosca, arXiv:quant-ph/0808.0369v1 (2008).

[29] T. M. Cover, J. A. Thomas, *Elements of Information Theory*, Wiley Series in Telecommunications (Wiley-Interscience, United States of American 1991).

[30] C. W. Helstrom, *Quantum Detection and Estimation Theory*, Academic Press.

[31] G. M. DAriano, P. L. Presti e P. Perinotti, J. Phys. A **38**, 5979 (2005).

[32] A. Datta, *Studies on the Role of Entanglement in Mixed-state Quantum Computation*, arXiv:0807.4490v1 [quant-ph] (2008).

[33] A. S. M. Hassan, B. Lari, P. S. Joag, arXiv:quant-ph/1010.1920v2 (2010).

[34] S. Luo, Phys. Rev. A **77**, 042303 (2008).

[35] U. Fano, Rev. Mod. Phys **55**, 855 (1983).

[36] W. K. Wooters, Phys. Rev. Lett. **80**, 2245 (1998).

[37] W. K. Wootters, Quantum Inf. Comput. I **59** (2001).

[38] G. A. F. Seber, *A Matrix Handbook for Statisticians*, Wiley Series in Probability and Statistics (Wiley, United States of American 2007).